INTERSPEECH 2014

Study of Changes in Glottal Vibration Characteristics During Laughter

Vinay Kumar Mittal[1] *and B. Yegnanarayana*[2]

International Institute of Information Technology, Hyderabad, India
[1]`vinay.mittal@iiit.ac.in`, [2]`yegna@iiit.ac.in`

10.21437/Interspeech.2014-405

Abstract

Laughter in speech has been studied mostly relying upon the spectral representation like formants and harmonics derived from the short-time spectrum. Significant changes appear to take place in the characteristics of glottal source of excitation during the production of laughter, but these changes have not been explored much. In this study, we examine the changes in the glottal vibration characteristics in laughter production, using the electroglottograph (EGG) signals. The excitation source characteristics are also examined from the corresponding acoustic signal, using a modified zero-frequency filtering method. Changes are examined in three categories, namely, normal speech, laughed-speech and nonspeech-laugh. Results using both EGG and acoustic signals are similar. The closed phase quotient in each glottal cycle is observed to decrease more for nonspeech-laugh calls than laughed-speech, with reference to normal speech. Correspondingly, the instantaneous fundamental frequency, hence pitch, increases more for nonspeech-laugh.

Index Terms: laughed-speech, nonspeech-laugh, glottal vibration characteristics, modified zero-frequency filtering, EGG

1. Introduction

Humans use nonlinguistic communication to convey representational messages like emotions or intentions [1]. Laughter, a paralinguistic event, is one of the most variable acoustic expression of a human being [2, 3]. Laughter is a special nonlinguistic vocalization because it induces positive affective state in the listeners, thereby affecting their behaviour [1, 4]. Detection of paralinguistic events like laughter has potentially diverse applications such as indexing/search of audio-visual databases, healthcare and biometrics etc. Detection of laughter regions in continuous speech can also help in classification of emotional states of a speaker [5]. Hence, researchers have been attracted in last few years towards finding the distinguishing features of laughter, to develop systems for detecting laughter regions in continuous speech [6, 7, 8, 9, 10, 11, 12].

Laughter signals are usually studied at episode, bout, call or segment levels [13, 14]. An *episode* consists of laughter bouts produced during one exhalation. Each laughter *bout* is a behaviour-acoustic event consisting of several laugh cycles or calls. A *call* consists of periods of laughter vocalization [13]. *Segments* are regions of audible changes during the production of a laughter call [14]. Laughter bouts are subdivided into onset (short/steep explosive laughter), apex-period (vocalization) and offset (post-vocalization) [13]. The number of laugh pulses in a laughter call depends upon the lung volume during it [13]. In this study, we examine the laughter signals at bout/call levels.

Laughter has been categorised in different studies, in different ways. Laughter was categorised in three classes: spontaneous laughter (with no voluntary restrain), voluntary laughter (a kind of faked laughter) and singing/speaking laughter [13]. In another study [14], three types of laugh bouts were discussed as song-like laugh (like giggle or chuckle), snort-like laugh (the unvoiced call) and unvoiced grunt-like laugh (like breathy pants and harsher cackles) [14]. Laughter was also categorised as per three classes of vowel quality: ‘ha’, ‘he’ and ‘ho’ [15, 16]. Four phonetic types of laughter were studied as: voiced, chuckle, breathy (ingression) and nasal-grunt [9, 17]. Two broad categories of laughter, based upon voicing, are: *voiced laughter* (song-like bouts, chuckles and giggles) and *unvoiced laughter* (open-mouth pant-like sounds, closed-mouth grunts and nasal-snorts) [1]. The speech-laugh continuum was divided as speech, laugh-speech and laugh [4]. In this study, we analyse laughter in speech in three categories, namely, *normal speech*, *laughed-speech* and *nonspeech-laugh*. Laughed-speech may have both linguistic vocalization and nonlinguistic laugh content interspersed in some degree. Normal speech is used for reference.

Acoustic analysis of laughter was carried out using features such as fundamental frequency (F_0), rms amplitude and duration, and breathiness [18]. Source-energy and temporal/spectral features of laughter were analyzed using F_0 and formant clusters [14]. The role of F_0 and rhythm (duration) in laughter was investigated [19]. The formants space, pitch range and voice quality features for speech-laughter continuum were also studied [4]. Combinations of pitch, energy, voicing, perceptual linear prediction (PLP) and modulation spectrum features were used, to model the laughter and speech [5]. The acoustic spectral and perceptual features have been used in applications such as detection of laughter events in meetings [8], distinguishing the four (phonetic) types of laughter [9], and speech/laughter classification in meetings audio [11] etc. Other diverse applications include ‘hot-spotter’ [7], automatic laughter detection [5, 10] and ‘AVLaughterCycle’ project [12].

Production of laughter was studied from respiratory dynamics point of view [20]. “Laughter generally takes place when the lung volume is low” (p442) [21]. Since, laughter is produced by the human speech production mechanism, the laughter signal also can be analyzed in terms of the excitation source and vocal tract system characteristics, like normal speech. In the production of laughter, apparently significant changes occur in the characteristics of the glottal source of excitation [4]. But, most studies on the acoustic analysis of laughter have focused mainly on the spectral and perceptual features. Changes in the excitation source characteristics in laughter have not been studied much. A preliminary study of laugh signals used source features like instantaneous pitch periods (T_0) and strength of excitation [22]. But, changes in the glottal vibration characteristics in laughter have not been studied using glottal features. Hence, there is a need to investigate the changes in the vibration characteristics of the vocal folds during laughter, in order to understand the laughter production characteristics in detail.

In this paper, we examine the changes in the glottal vibration characteristics of laughter in speech, using the EGG sig-

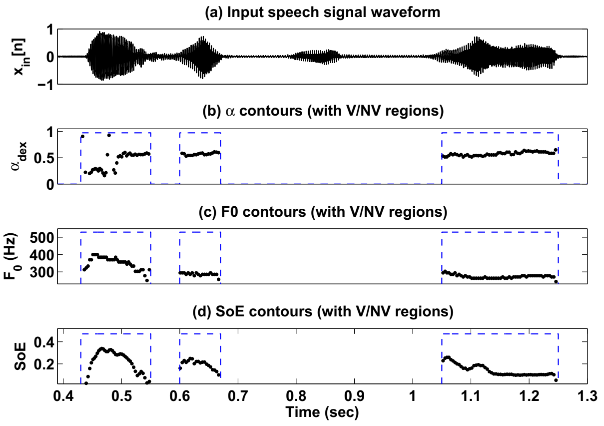

Figure 1: Illustration of (a) speech signal, and (b) α, (c) F_0 and (d) SoE contours for a *laughed-speech* segment of the utterance of text "it is a good joke", by a female speaker.

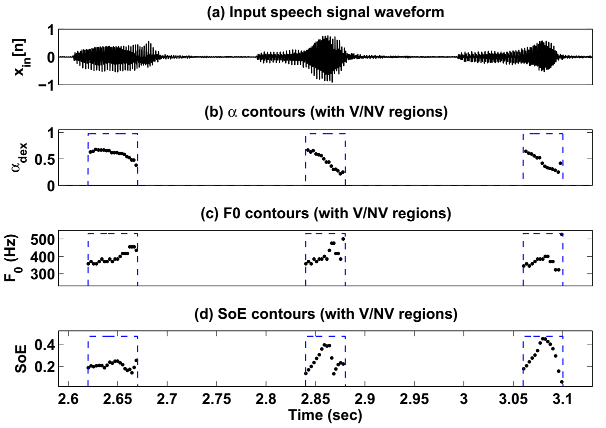

Figure 2: Illustration of (a) speech signal, and (b) α, (c) F_0 and (d) SoE contours for three calls in a *nonspeech-laugh* bout *after* utterance of text "it is a good joke", by a female speaker.

nal [23] first. Results are then compared with the excitation source features derived from the corresponding acoustic signal. Data for the study was collected by eliciting the natural laughter responses, by playing hilarious videos and joke audios to the subjects. Three pre-defined texts were used for the responses. Both EGG and acoustic signals were recorded for normal speech, laughed-speech and nonspeech-laugh. Differences in the features derived from laughed-speech and nonspeech-laugh are examined with reference to normal speech. The closed phase quotient (α) and instantaneous fundamental frequency (${F_0}_{EGG}$) are derived from the open/closed phase durations of glottal cycles in voiced regions, using the differenced EGG (dEGG) signal. Changes in the source characteristics are also analyzed by computing the ${F_0}_{ZFF}$ and strength of excitation (SoE) around glottal closure instants, from the corresponding acoustic signals, using a few modifications proposed in the zero-frequency filtering method [24]. Features from both EGG and acoustic signals (e.g., ${F_0}_{EGG}$ and ${F_0}_{ZFF}$) are compared.

The paper is organized as follows. Section 2 discusses the details of data collected for the study. The glottal characteristics of laughter are analyzed from EGG signal in Section 3. In Section 4, the excitation source characteristics of laughter are obtained from acoustic signal, using the modified zero-frequency filtering method. The results are discussed in Section 5. A summary is given in Section 6, along with scope of further work.

2. Data for the study

In order to examine changes in the glottal vibration characteristics during production of laughter, the data was collected in three categories: (i) normal speech (NS), (ii) *laughed-speech (LS)*, and (iii) *non-speech laugh (NSL)*. The term *non-speech laugh (NSL)* corresponds to 'pure laugh' in [4]. It is difficult to quantify the degree of laughter content in *laughed-speech*. Hence, wide range of variations can be expected in the feature values derived for laughed-speech. Since the aim is to examine changes in the glottal vibration, the unvoiced laughter calls (like nasal-snorts or breathy-grunts) are not focused, in this study.

Natural spontaneous laughter signals data was collected by eliciting natural laughter response in the subjects. A series of audio-visual comedy clips and/or joke audios (in English/Telugu), collected from online media sources, were played to each subject. In order to minimize the text related variability in features, the subjects were suggested to use 3 pre-defined English texts in their responses or to express themselves, in case they really liked the joke or comedy. The texts are: (i) "It is a good joke", (ii) "It is really funny.", (iii) "I have enjoyed." Both acoustic and EGG signals were recorded for the laughter responses of each subject. Data was recorded for total 11 speakers (7 males, 4 females), all research students at IIIT, Hyderabad. The natural responses data consists of both nonspeech-laugh calls and laughed-speech. The data has total 191 laughter calls in 32 utterances produced by 11 speakers. The nonspeech-laugh calls occur mostly prior to (or sometimes after) the laughed-speech for a text. Each utterance consists of 3 to 8 laugh calls.

The EGG signal was recorded using a EGG recording device [25], and the corresponding acoustic signal using a close speaking (stand mounted) microphone (at a distance around 6" from speaker). The data was recorded in normal lab-room conditions, at a sampling rate of 48 KHz. Data was downsampled to 10 KHz for the analysis, for computational convenience. The ground truth for the data was established manually by listening to each file and segregating the nonspeech-laugh and laughed-speech calls in each utterance.

3. Glottal characteristics from EGG signal

It is known from earlier studies that F_0 is higher for laughter than for normal speech [18, 14, 4, 19]. Like normal speech, the laughter signals are also produced by the vibration of the vocal folds at the glottis. Hence, it is possible that in the case of laughter the vocal folds vibrate faster, with a changed rate of their opening/closing. In order to examine the glottal vibration characteristics of laughter in more detail, the relative open/closed phase durations in each glottal cycle are studied using the EGG signal [23, 25]. These durations are better computed using the dEGG signal [26]. The proportion (α) of the closed phase region within each glottal cycle period is computed for laugh calls in the voiced regions. The closed phase quotient (α) values are compared for laughed-speech and nonspeech-laugh bouts.

An illustration of α contours for a laughed-speech segment and 3 calls in a nonspeech-laugh bout of a female speaker is shown in Fig. 1(b) and Fig. 2(b), respectively. The voiced/nonvoiced (V/NV) regions (Fig. 1 and Fig. 2) are de-

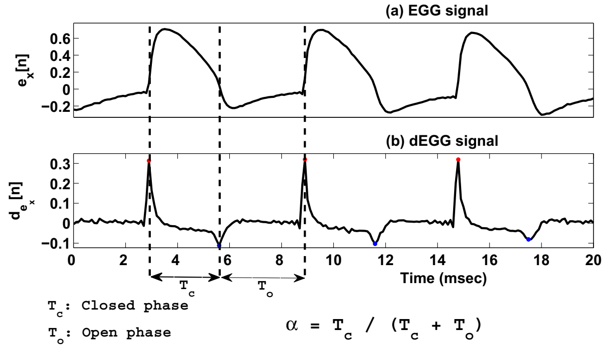

Figure 3: Illustration of α-*computation* from durations of closed phase (T_C) and open phase (T_O), derived using dEGG signal.

Table 1: Average α values for each speaker, using *EGG signal*, for the case of: (a) normal speech (NS), (b) laughed-speech (LS), (c) nonspeech-laugh (NSL), (d) changes for laughed-speech ($\Delta\alpha_{LS}$), and (e) changes for nonspeech-laugh ($\Delta\alpha_{NSL}$). Note: M/F indicates male/female speaker.

Speaker (M/F)	(a) α_{NS}	(b) α_{LS}	(c) α_{NSL}	(d)(%) $\Delta\alpha_{LS}$	(e)(%) $\Delta\alpha_{NSL}$
S1 (M)	.504	.480	.395	-4.73	-21.53
S2 (M)	.465	.408	.360	-12.26	-22.58
S3 (F)	.440	.489	.379	11.09	-13.89
S4 (M)	.511	.486	.469	-4.76	-8.19
S5 (M)	.436	.373	.338	-14.39	-22.36
S6 (F)	.479	.478	.464	-0.01	-3.03
S7 (M)	.524	.467	.419	-10.93	-20.03
S8 (M)	.153	.526	.460	243.92	200.46
S9 (F)	.379	.493	.422	30.05	11.23
S10(F)	.485	.443	.416	-8.67	-14.38
S11(M)	.412	.425	.350	3.30	-14.89

cided using frame-wise signal-energy and strength of excitation based criteria. Voiced regions above a threshold (20%) are chosen for the analysis. The proportion α is computed as $\alpha = \frac{T_C}{(T_C+T_O)}$, where T_C and T_O are closed and open phase durations, respectively (Fig. 3) [26]. The durations T_C and T_O are obtained using the dEGG signal ($d_{ex}[n]$), as in Fig. 3 [26].

It is interesting to observe that the glottal closed phase quotient (α) decreases for nonspeech-laugh (Fig. 2(b)), in comparison to normal speech. Also, the gradient of α-contour increases over successive calls in a laugh bout. Both observations can be contrasted with higher and mostly uniform α-contour for laughed-speech (Fig. 1(b)). It indicates relatively longer open phase and shorter closing phase of the vocal folds, during glottal cycles in nonspeech-laugh. It is possibly related to large subglottal pressure, which causes increased lateral compression of the glottis, and the resultant greater airflow during nonspeech-laugh. Consequently, the rate of vibration of vocal folds, i.e., F_0 increases in nonspeech-laugh. The F_0 from EGG signal (i.e., ground truth F_0) is computed as $F0_{EGG} = 1/T0_{EGG}$, where $T0_{EGG}$ is glottal cycle period ($T0_{EGG} = T_C + T_O$).

In Table 1, the average α values for each speaker are given in columns (a), (b) and (c), for the cases of normal speech, laughed-speech and nonspeech-laugh, respectively. Changes in average α for laughed-speech with reference to normal speech ($\Delta\alpha_{LS}\% = \frac{(\alpha_{LS} - \alpha_{NS})}{\alpha_{NS}}$x100) are given in column (d). Similarly, changes in average α for nonspeech-laugh ($\Delta\alpha_{NSL}\% = \frac{(\alpha_{NSL} - \alpha_{NS})}{\alpha_{NS}}$x100) are given in column (e). It may be observed from columns (b) and (c) that average α values are lower for nonspeech-laugh than laughed-speech. Also, the average α reduces more for nonspeech-laugh (column (e)) than for laughed-speech (column (d)), across speakers. The reduction in α for laughed-speech is due to laughter interspersed with speech. Variation in α seems due to different phonation styles.

Table 2: Average $F0_{EGG}$ values (Hz) for each speaker, using *EGG signal*: (a) normal speech (NS), (b) laughed-speech (LS), (c) nonspeech-laugh (NSL), (d) changes for laughed-speech ($\Delta F0_{LS}$), and (e) changes for nonspeech-laugh ($\Delta F0_{NSL}$).

Speaker (M/F)	(a)(Hz) $F0_{NS}$	(b)(Hz) $F0_{LS}$	(c)(Hz) $F0_{NSL}$	(d)(EGG) $\Delta F0_{LS}$ (%)	(e)(EGG) $\Delta F0_{NSL}$ (%)
S1 (M)	168.1	168.3	255.5	0.11	51.95
S2 (M)	148.2	178.5	183.2	20.44	23.64
S3 (F)	249.1	279.9	281.4	12.33	12.94
S4 (M)	146.6	148.9	182.9	1.60	24.80
S5 (M)	152.2	218.1	234.5	43.27	54.07
S6 (F)	234.6	281.9	378.4	20.20	61.31
S7 (M)	137.7	125.8	187.0	-8.56	35.87
S8 (M)	133.2	194.1	220.7	45.68	65.61
S9 (F)	211.9	255.5	290.9	20.55	37.29
S10(F)	207.8	354.7	383.8	70.72	84.72
S11(M)	142.6	175.8	181.8	23.29	27.44
Average change in F_0 (EGG)				*22.70*	*43.60*

Changes in the closed phase quotient (α) in each glottal cycle period, also explain the rise in pitch (F_0) in laughter that can be perceived easily by human listeners. In Table 2, average F_0 values computed from EGG signal ($F0_{EGG}$) for each speaker are given (in Hz) in columns (a), (b) and (c), for normal speech, laughed-speech and non-speech laugh, respectively. Changes in average $F0_{EGG}$ for laughed-speech with reference to normal speech ($\Delta F0_{LS_{EGG}}\% = \frac{(F0_{LS} - F0_{NS})}{F0_{NS}}$x100), are given in column (d). Similarly, changes in the average $F0_{EGG}$ for nonspeech-laugh ($\Delta F0_{NSL_{EGG}}\% = \frac{(F0_{NSL} - F0_{NS})}{F0_{NS}}$x100)) are given in column (e). In columns (d) and (e), the changes in $F0_{EGG}$ with reference to normal speech are more for nonspeech-laugh than for laughed-speech, across speakers. The wide variation in $\Delta F0_{EGG}$ for laughed-speech is due to varying degree of laughter content interspersed with speech.

4. Source characteristics from speech signal

Since it is difficult to collect the EGG signal practically in real-life, the excitation source characteristics of laughter are also derived from the acoustic signal using the zero-frequency filtering (ZFF) method [24]. In the ZFF method [24] for normal speech, a cascade of two ideal digital filters, called *zero-frequency resonators (ZFRs)*, is used. Each ZFR has a pair of poles at the Unit-circle in the $z-$plane. The pre-emphasized signal $s[n]$ (obtained from the input speech signal $x_{in}[n]$) is passed through this cascade of ZFRs, each giving output as $y[n] = -\sum_{k=1}^{2} a_k y[n-k] + s[n]$, where $a_1 = -2$, $a_2 = 1$. The trend in the output of the cascade of ZFRs is removed by subtracting the local mean computed over a window. The window size is taken as about 1.3 times the average pitch period estimated a priori for each voiced region using the autocorrelation method [27]. The mean-subtracted output signal is called the *zero-frequency filtered (ZFF) signal* [24]. The lo-

Table 3: Average $F0_{ZFF}$ (Hz) for each speaker, using *acoustic signal*: (a) normal speech (NS), (b) laughed-speech (LS), (c) nonspeech-laugh (NSL), (d) changes for laughed-speech ($\Delta F0_{LS}$), and (e) changes for nonspeech-laugh ($\Delta F0_{NSL}$).

Speaker (M/F)	(a)(Hz) $F0_{NS}$	(b)(Hz) $F0_{LS}$	(c)(Hz) $F0_{NSL}$	(d)(ZFF) $\Delta F0_{LS}$ (%)	(e)(ZFF) $\Delta F0_{NSL}$(%)
S1 (M)	184.3	192.6	269.2	4.46	46.02
S2 (M)	186.6	217.0	223.3	16.34	19.68
S3 (F)	273.3	278.6	307.6	1.91	12.54
S4 (M)	145.7	150.4	236.0	3.19	61.94
S5 (M)	157.8	218.5	258.9	38.48	64.16
S6 (F)	261.0	275.7	378.4	5.62	44.99
S7 (M)	157.2	151.8	293.8	-3.40	86.96
S8 (M)	141.9	211.2	260.2	48.79	83.28
S9 (F)	239.7	259.3	298.2	8.15	24.40
S10(F)	262.3	339.3	344.1	29.33	31.19
S11(M)	168.5	185.9	208.5	10.36	23.74
Average change in F_0 (ZFF)				*14.84*	*45.35*

cations of zero crossings of the ZFF signal correspond to the glottal closure instants (GCIs), called *epochs* [24]. The *instantaneous fundamental frequency* (F_0, i.e., $F0_{ZFF}$) is computed from the inverse of period (T_0) between successive epochs [28].

The ZFF method [24, 28], that was also used in [22] (applying ZFF separately to each voiced region), was proposed mainly for modal voicing. A few modifications are carried out in it, for the analysis of nonverbal/expressive voices such as laughter. The *modified ZFF method* uses gradually reducing window lengths for trend removal (details in [29]). The slope of the ZFF signal at GCIs (using the modified ZFF method), gives the relative *strength of* the impulse-like *excitation (SoE)* around epochs [24]. Changes in $F0_{ZFF}$ and SoE contours for the calls in a nonspeech-laugh bout (Fig. 2(c),(d)) can be contrasted from those for a laughed-speech segment (Fig. 1(c),(d)), by a female speaker. The increasing SoE and $F0_{ZFF}$ contours for successive calls in a nonspeech-laugh bout (Fig. 2), are perhaps related to the correspondingly increasing α-contours. This possibly explains the changes in pitch and the strength of excitation over the successive calls in a laughter bout. The variations in $F0_{ZFF}$ and SoE for laughed-speech (Fig. 1) are due to laughter content present (interspersed) with speech (see [29] for SoE study).

In Table 3, average F_0 derived from the acoustic signal using the modified ZFF method ($F0_{ZFF}$) for each speaker are given for normal speech, laughed-speech and nonspeech-laugh (in Hz) in columns (a), (b) and (c), respectively. Changes in average $F0_{ZFF}$ for laughed-speech with reference to normal speech ($\Delta F0_{LS_{ZFF}}\% = \frac{(F0_{LS} - F0_{NS})}{F0_{NS}}$x100), are given in column (d). Similarly, changes in $F0_{ZFF}$ for laughed-speech ($\Delta F0_{NSL_{ZFF}}\% = \frac{(F0_{NSL} - F0_{NS})}{F0_{NS}}$x100) are given in column (e). It may be observed from columns (d) and (e) that changes in $F0_{ZFF}$ (with reference to normal speech) are more for nonspeech-laugh than laughed-speech, across speakers. The differences between $F0_{EGG}$ and $F0_{ZFF}$ (for the three cases) in Table 2 and Table 3, respectively, are within a range of around 10%, which is acceptable because $\Delta F0_{NSL}\%$ is around 44% . Thus, changes in $F0_{ZFF}$ are in-line with the changes in $F0_{EGG}$.

5. Discussion on Results

The average α values are lower for nonspeech-laugh than those for laughed-speech bouts, across speakers (in columns (d) and (e) in Table 1). The results verify an earlier proposition [4] that the open quotient (derived approximately in [4] using spectral harmonics) is larger for laughter than speech in vowel regions. The detailed analyses of EGG along with acoustic signals in this study, establishes clearly that in the production of laughter the closed phase of glottal cycles is reduced. The varying degree of reduction in the closed phase quotient for laughed-speech (α_{LS}), is due to different phonation styles and varying extents of laughter present in speech. Larger $\Delta\alpha_{LS}$ indicates the presence of laughter interspersed with speech to a larger extent.

The lower closed phase quotient (α) for nonspeech-laugh, possibly causes the wide open vocal folds to abruptly slap together during the closing phase of the glottal cycle. Resultant higher subglottal pressure causes higher lateral compression of the vocal folds. Consequently, the rate of airflow through the glottis is high, causing the vocal folds to close rapidly. The related increased rate of vibration of the vocal folds at the glottis is manifested as the rise in pitch, i.e., increased F_0 for nonspeech-laugh. The higher average $F0_{EGG}$ (in Table 2) for nonspeech-laugh as compared to laughed-speech bouts, strengthens the possibility of this phenomenon occurring in the production of laughter. Results in Tables 1 and 2 also infer that, reduction in α is indeed related to increased F_0 and hence higher pitch in laughter as compared to normal speech.

Changes in average $F0_{ZFF}$ derived from the acoustic signal using the modified ZFF method [29] (in Table 3), are in line with changes in average $F0_{EGG}$ derived from the EGG signal (in Table 2). It validates the approaches and signal processing methods adopted for the analyses. The lower closed phase quotient (α) of glottal cycles is also perhaps related to shorter glottal cycle periods in nonspeech-laugh, which is manifested as increased rate of glottal vibration (i.e., F_0), in the case of laughter. The range of F_0 values observed using both EGG and acoustic signals also confirm earlier results [18, 14, 4, 19].

6. Summary and conclusion

In this paper, we have examined the changes in the glottal vibration characteristics in the production of *nonspeech laugh* and *laughed-speech*, with reference to normal speech. Analyses of the EGG and acoustic signals reveal that significant changes indeed take place in the vibration characteristics of the vocal folds at the glottis, during the production of laughter. The proportion (α) of the closed phase region to the glottal cycle period is lower for nonspeech-laugh than for laughed-speech. Also, the closed phase quotient for laughed-speech (α_{LS}) varies between that for normal speech and nonspeech-laugh, because it has speech mixed with laughter, to some extent.

The comparative study of results from the EGG signal and the acoustic signal confirms that in the case of nonspeech-laugh the vocal folds at the glottis have relatively shorter closing phase and smaller period (T_0), thereby leading to the perception of higher pitch (F_0). The consistently higher average F_0 is observed for nonspeech-laugh bouts in comparison to those for laughed-speech. A modified ZFF method is proposed for extracting source features from the acoustic signal for laughter.

The changes in the closed phase quotient (α) within each glottal cycle period in the production of laughter may have associated changes in the vocal tract system characteristics also, which we aim to study in our further work. This study may be helpful in understanding in detail the role of glottal source of excitation in the production of laughter, and *emotions* like ‘happy’ or ‘joy’, which may in turn help further in developing better systems for detecting these emotions in continuous speech.

7. References

[1] M. Owren and J. A. Bachorowski, “Reconsidering the evolution of nonlinguistic communication: The case of laughter,” *Journal of Nonverbal Behavior*, vol. 27, no. 3, pp. 183–200, 2003.

[2] H. Rothganger, G. Hauser, A. C. Cappellini, and A. Guidotti, “Analysis of laughter and speech sounds in Italian and German students,” *Naturwissenschaften*, vol. 85, no. 8, pp. 394–402, 1998.

[3] E. E. Nwokah, H. C. Hsu, P. Davies, and A. Fogel, “The integration of laughter and speech in vocal communication: A dynamic systems perspective,” *J. Speech Lang. Hear. Res.*, vol. 42, no. 4, pp. 880–894, 1999.

[4] C. Menezes and Y. Igarashi, “The speech laugh spectrum,” in *Proc. 6th International Seminar on Speech Production, 2006 (ISSP’06)*, Dec. 13-15, 2006, pp. 157–524.

[5] K. P. Truong and D. A. V. Leeuwen, “Automatic detection of laughter,” in *Proc. of 9th European Conference on Speech Communication and Technology, 2005 (INTERSPEECH’05)*, Sep. 4-8, 2005, pp. 485–488.

[6] A. Lockerd and F. Mueller, “LAFCam: Leveraging affective feedback camcorder,” in *CHI ’02 Extended Abstracts on Human Factors in Computing Systems*, 2002, pp. 574–575.

[7] R. Cai, L. Lu, H. J. Zhang, and L. H. Cai, “Highlight sound effects detection in audio stream,” in *Proc. IEEE International Conference on Multimedia and Expo, 2003 (ICME’03)*, vol. 3, July 2003, pp. 37–40.

[8] L. S. Kennedy and D. P. W. Ellis, “Laughter detection in meetings,” in *Proc. NIST ICASSP 2004 Meeting Recognition Workshop*, 2004.

[9] N. Campbell, H. Kashioka, and R. Ohara, “No laughing matter,” in *Proc. 9th European Conference on Speech Communication and Technology, 2005 (INTERSPEECH’05)*, Sep. 4-8, 2005, pp. 465–468.

[10] M. T. Knox and N. Mirghafori, “Automatic laughter detection using neural networks,” in *Proc. 8th Annual Conference of the International Speech Communication Association, 2007 (INTERSPEECH’07)*, Aug. 27-31, 2007, pp. 2973–2976.

[11] S. Z. K. Khine, T. L. Nwe, and H. Li, “Speech/laughter classification in meeting audio,” in *Proc. 9th Annual Conference of the International Speech Communication Association, 2008 (INTERSPEECH’08)*, Sep. 22-26, 2008, pp. 793–796.

[12] J. Urbain, R. Niewiadomski, E. Bevacqua, T. Dutoit, A. Moinet, C. Pelachaud, B. Picart, J. Tilmanne, and J. Wagner, “AVLaughterCycle,” *Journal on Multimodal User Interfaces*, vol. 4, no. 1, pp. 47–58, 2010.

[13] W. Ruch and P. Ekman, “The Expressive Pattern of Laughter,” *Emotion, Qualia, and Consciousness*, pp. 426–443, 2001, edited by A. W. Kaszniak (Word Scientific, Tokyo).

[14] J. A. Bachorowski, M. J. Smoski1, and M. J. Owren, “The acoustic features of human laughter,” *J. Acoust. Soc. Am.*, vol. 110, no. 3, pp. 1581–1597, 2001.

[15] R. R. Provine and Y. L. Yong, “Laughter: A stereotyped human vocalization,” *Ethology*, vol. 89, no. 2, pp. 115–124, 1991, (published by Blackwell Publishing Ltd).

[16] R. Provine, *Laughter: A Scientific Investigation*. Viking, 2000, (published by Viking).

[17] H. Tanaka and N. Campbell, “Acoustic features of four types of laughter in natural conversational speech,” in *Proc. 17th International Congress of Phonetic Sciences, 2011 (ICPhS XVII)*, Aug. 17-21, 2011, pp. 1958–1961.

[18] C. A. Bickley and S. Hunnicutt, “Acoustic analysis of laughter,” in *Proc. Second International Conference on Spoken Language Processing, 1992 (ICSLP’92)*. ISCA, Oct 13-16, 1992, pp. 927–930.

[19] S. Kipper and D. Todt, “The role of rhythm and pitch in the evaluation of human laughter,” *Journal of Nonverbal Behavior*, vol. 27, no. 4, pp. 255–272, 2003.

[20] M. Filippelli1, R. Pellegrino, I. Iandelli1, G. Misuri1, J. R. Rodarte, R. Duranti, V. Brusasco, and G. Scano1, “Respiratory dynamics during laughter,” *Journal of Applied Physiology*, vol. 90, no. 4, pp. 1441–1446, Apr. 1, 2001.

[21] E. S. Luschei1, L. O. Ramig, E. M. Finnegan, K. K. Baker, and M. E. Smith, “Patterns of laryngeal electromyography and the activity of the respiratory system during spontaneous laughter,” *Journal of Neurophysiology*, vol. 96, no. 1, pp. 442–450, Jul. 2006.

[22] K. S. Kumar, S. H. R. Mallidi, K. S. R. Murty, and B. Yegnanarayana, “Analysis of laugh signals for detecting in continuous speech,” in *Proc. 10th Annual Conference of the International Speech Communication Association,2009 (INTERSPEECH’09)*. ISCA, Sep. 6-10, 2009, pp. 1591–1594.

[23] G. Fant, Q. Lin, and C. Gobl, “Notes on glottal flow interaction,” *STL-QPSR, KTH, Sweden*, vol. 26, no. 2-3, pp. 21–45, 1985.

[24] K. S. R. Murty and B. Yegnanarayana, “Epoch extraction from speech signals,” *IEEE Transactions on Audio, Speech, and Language Processing*, vol. 16, no. 8, pp. 1602 –1613, Nov. 2008.

[25] D. G. Miller, “EGGs for Singers,” 2012, (last viewed May 29, 2013). [Online]. Available: http://www.eggsforsingers.eu/

[26] V. K. Mittal and B. Yegnanarayana, “Effect of glottal dynamics in the production of shouted speech,” *The Journal of the Acoustical Society of America*, vol. 133, no. 5, pp. 3050–3061, May 2013.

[27] J. Makhoul, “Linear prediction: A tutorial review,” *Proc. IEEE*, vol. 63, pp. 561–580, Apr. 1975.

[28] B. Yegnanarayana and K. S. R. Murty, “Event-based instantaneous fundamental frequency estimation from speech signals,” *IEEE Trans. Audio, Speech, and Language Processing*, vol. 17, no. 4, pp. 614–624, May 2009.

[29] V. K. Mittal and B. Yegnanarayana, “Analysis of production characteristics of laughter,” *Computer Speech and Language*, 2014, (paper is under second review).