Descriptive Statistical Evaluation of Zero Crossing Rate (ZCR) Features in Indonesian and Buginese Speech Signals for Speech Processing Applications

Sriwijanaka Yudi Hartono[1,a], Saidah Suyuti[2,*,b], Muhammad Iqbal[3,c]
[1,2,3] Program Studi Teknik Elektro Fakultas Teknik Universitas Muslim Indonesia
Jalan Urip Sumoharjo Km 04 Makassar Sulawesi Selatan
[a]Sriwijanaka.hartono@umi.ac.id, [*,b]saidahsuyuti06@gmail.com, [3]iqbalsmanlan@gmail.com

***Abstract*— This study analyzes the statistical characteristics of the Zero Crossing Rate (ZCR) in Indonesian and Buginese speech signals produced by both male and female speakers. The objectives of this research are to identify the statistical patterns of ZCR, compare the distribution of its key descriptive measures between the two languages, and provide a scientific foundation for developing speech processing systems that are adaptive to regional languages. Three types of utterances were examined, and ZCR features were extracted using descriptive statistical parameters including mean, variance, standard deviation, skewness, and kurtosis. The findings show that Buginese speech exhibits higher and more variable ZCR values compared to Indonesian speech. A statistical summary from all tables indicates that the mean ZCR of Buginese speech ranges from 0.0727 to 0.2464, whereas Indonesian speech ranges from 0.0338 to 0.1577. The variance in Buginese speech reaches 0.009023–0.022207, significantly higher than Indonesian speech, which ranges from 0.000738 to 0.012896. The standard deviation of Buginese speech reaches 0.149019, while Indonesian speech ranges from 0.027173 to 0.112889. These wider distributions in Buginese speech reflect more intense polarity fluctuations, influenced by its distinctive phonetic features such as geminate consonants, sharp fricatives, and glottal stops. In contrast, Indonesian speech shows more stable ZCR characteristics, with narrower distributions of skewness and kurtosis, consistent with its dominance of vowels and softer voiced consonants. These statistical differences are consistent across both male and female speakers, although female speakers tend to exhibit slightly more stable patterns. The results confirm that ZCR is an effective acoustic feature for distinguishing the speech characteristics of Indonesian and Buginese. Overall, this study provides a strong scientific basis for the development of speech processing and automatic speech recognition (ASR) systems that are more adaptive to the phonetic characteristics of regional languages.**

Keywords*— *algoritma, zero crossing rate, staistika, descriptive, buginese

I. Introduction

Speech signal processing is one of the strategic areas in the development of artificial intelligence technology, particularly in speech recognition systems, language identification, and voice-bot applications. One of the most widely used temporal features in signal analysis is the Zero Crossing Rate (ZCR), which functions as an indicator of signal sign changes and can describe the temporal dynamics of speech signals [1], [2].

Bahasa Indonesia has been extensively studied from acoustic and phonetic perspectives, with emphasis on fricative features and syllabic structures that tend to produce more complex signal fluctuations [3], [4]. However, studies on regional languages, including the Bugis language, remain relatively limited. In fact, Bugis is one of the major languages in South Sulawesi, characterized by unique phonetic properties such as long vowels, soft-voiced consonants, and prominent glottal stops [5], [6].

These phonetic characteristics are presumed to influence the temporal patterns of speech signals, which are reflected in ZCR values. However, no research to date has specifically compared the statistical characteristics of ZCR between Bahasa Indonesia and Bugis. Furthermore, dialectal variations in Bugis may introduce important articulatory differences that should be considered in speech recognition systems [7].

This study is designed to address this gap by conducting a descriptive statistical analysis of ZCR across four groups of speakers. Thus, the research not only provides deeper phonetic insights but also establishes a scientific foundation for the development of automatic speech recognition (ASR) technologies and AI systems for regional languages.

The objectives of this research include: (1) identifying the statistical characteristics of ZCR in Bahasa Indonesia and Bugis; (2) comparing mean, variance, standard deviation, and range between the two languages; and (3) providing a scientific basis for developing speech recognition systems that adapt to regional language variations.

II. Research Methodology

This section describes the research methodology, including the tools and materials used, research equipment, objects and location of the study, as well as the stages of the research procedure.

A. Tools, Materials, and Research Location

1. The tools used in this study include a high-sensitivity condenser microphone, a laptop/PC with a minimum specification of 8 GB RAM, and software such as Wavesurfer or Audacity for audio recording and preprocessing. MATLAB programming language, Microsoft Word, and Microsoft Excel were also used.
2. Materials:The materials consisted of recorded speech from 5–10 speakers (Bahasa Indonesia speakers and Bugis language speakers), each producing recordings with a minimum duration of 3 seconds consisting of general narrative and descriptive sentences in .wav format. The recording process was conducted in a quiet environment with controlled acoustics to ensure high-quality speech signals.
3. Object: The object of this research is speech signals from native speakers of Bahasa Indonesia and the Bugis language.
4. Research Location: The research was conducted within the Faculty of Engineering, Department of Electrical Engineering, specifically at the Computational Engineering Laboratory, Electrical Engineering Study Program, Faculty of Engineering, Universitas Muslim Indonesia.

B. Research Procedure

The stages of the research procedure are outlined as follows:

1. Stage 1: Preparation and Data Collection

   A list of test sentences in Bahasa Indonesia and the Bugis language with similar length and structure was prepared. Native speakers aged 20–40 years were recruited, consisting of 10 respondents (5 males and 5 females) who had no speech impairments and were fluent in one or both languages. Speech recordings were carried out in a closed room with minimal noise and stored in WAV format following standard technical parameters. MATLAB and Wavesurfer were used to organize therecordings into a database. The sentences recorded included:

   - Bahasa Indonesia: “membaca”, “berbicara bahasa Indonesia”, and “satu kata satu perbuatan”
   - Bugis language: “na’baca”, “ma’bicara ogi”, and “taro ada taro gau”.
2. Stage 2: Signal Preprocessing

   Signal preprocessing included removing background noise using Audacity, normalizing audio volume, and segmenting recordings into sentence or phrase units. Framing and windowing were applied to each speech signal using MATLAB and Wavesurfer as part of ZCR feature extraction. The waveform of each signal was also visualized.
3. Stage 3: ZCR Feature Extraction

   Feature extraction was performed using the Zero Crossing Rate (ZCR) algorithm on speech signals from both Bahasa Indonesia and Bugis speakers. MATLAB was used to compute ZCR values and their sequences. The results were saved as numerical arrays and visualized for descriptive statistical analysis and signal representation.
4. Stage 4: Descriptive Statistical Analysis

   The extracted ZCR values were analyzed using descriptive statistical parameters including mean, variance, standard deviation, skewness, and kurtosis (if required). Boxplots were generated for each language to observe data distribution. A two-sample difference test was also conducted to examine whether the ZCR differences between the two languages were statistically significant.
5. Stage 5: Interpretation and Reporting

   The final stage involved interpreting the ZCR differences across languages, drawing conclusions, and providing recommendations for applying the findings to locally adapted speech recognition systems.

   A block diagram illustrating the research workflow is presented in Figure 1.

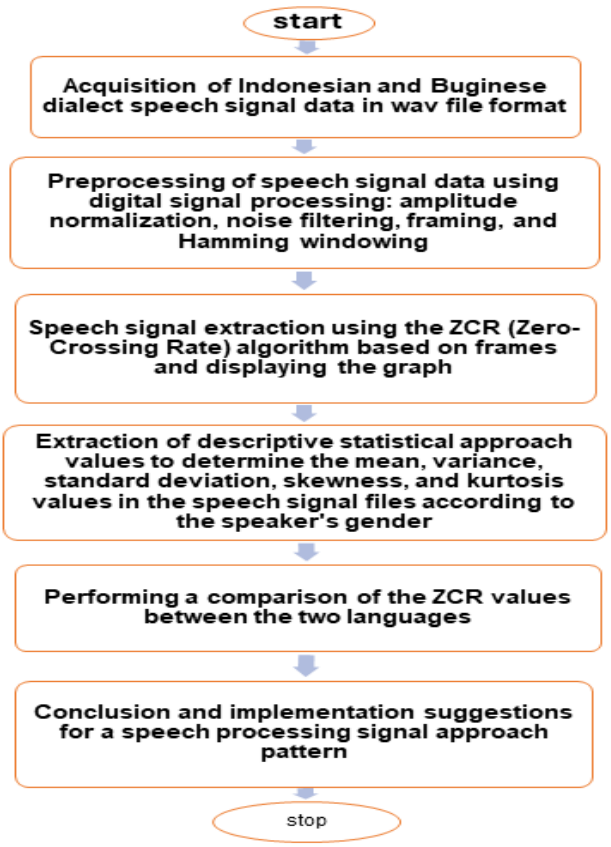

Figure 1. Block diagram illustrating the research workflow

III. Result and Discussion

A. Results of Speech Signal Acquisition

The speech signal data used for analysis were obtained through recordings conducted using MATLAB and Wavesurfer. The recorded words and sentences in Bahasa Indonesia and the Bugis language were sampled at 16 kHz, 16-bit, mono channel to produce high-quality *.wav* files for the database.

Based on the collected samples from both language groups, Figure 2 presents the waveform of one speech sample from Bahasa Indonesia, specifically the utterance *"membaca"*.

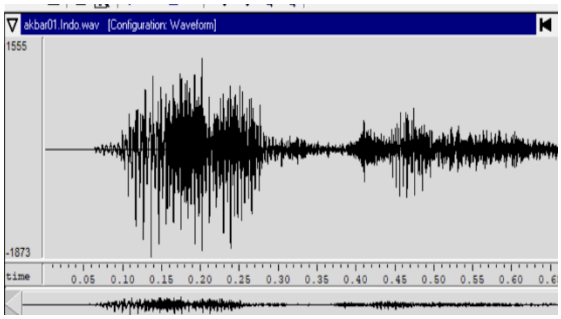

Figure 2. Waveform of the Indonesian speech utterance "membaca

Meanwhile, Figure 3 shows the waveform of the Bugis-language utterance *"na'baca"*.

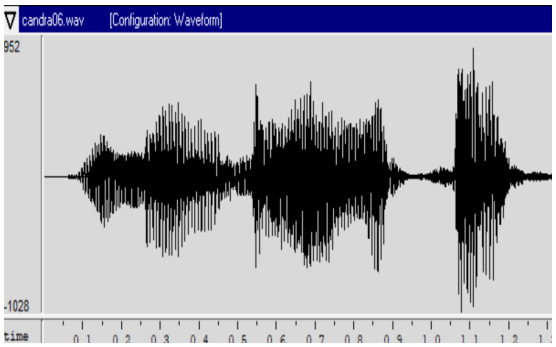

Figure 3. Waveform of the Bugis speech utterance "na'baca"

These waveforms demonstrate that speech signals exhibit distinctive characteristics within specific time regions. Speech signals are inherently time-varying waveforms that carry information generated by human vocal tract movements, represented through continuous variations in air pressure.

When observed over very short durations (5–100 ms), speech signals can be considered quasi-stationary. At shorter durations, the characteristics remain stable; however, over longer durations, the characteristics change and reflect the dynamic content of the spoken utterance. Thus, speech signals are generally depicted as complex, continuous acoustic waveforms.

Additionally, speech signals do not consist solely of isolated syllables; overlapping transitions between consecutive sounds often occur, contributing to the complexity of the waveform.

B. Results of Digital Signal Processing

Preprocessing is a crucial stage in speech-signal extraction, ensuring that noise is minimized and the signals are prepared for feature computation. This stage

improves the quality of the speech data for subsequent processing.

For both Bahasa Indonesia and Bugis signals recorded at a sampling frequency of 16 kHz, framing was performed using a frame length of 400 samples with an overlap of 160 samples. Frame duration, typically 10–30 ms, plays a critical role in capturing the quasi-stationary behavior of speech.

In this study, the speech signals were divided into frames of size 366 × 480. Figure 4 illustrates an example of the Indonesian speech signal before windowing.

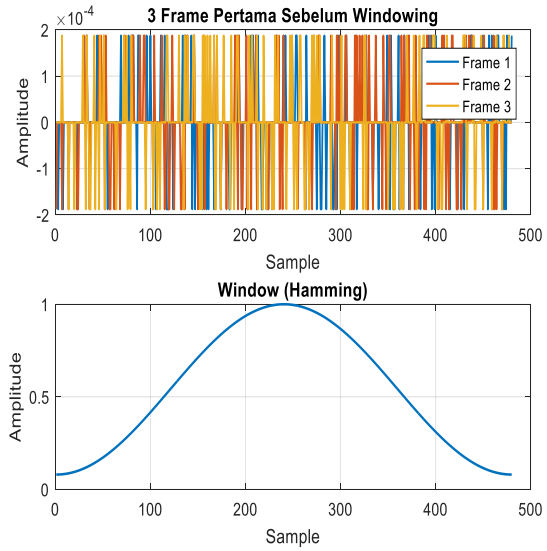

Figure 4. Frame graph before windowing and Hamming window applied.

Windowing was performed using the **Hamming window**, which reduces discontinuities between frames by assigning lower weights at the edges and higher weights at the center. This smooth transition minimizes artifacts and prepares the signal for reliable further analysis.

Figure 5 shows the waveform after applying the Hamming window.

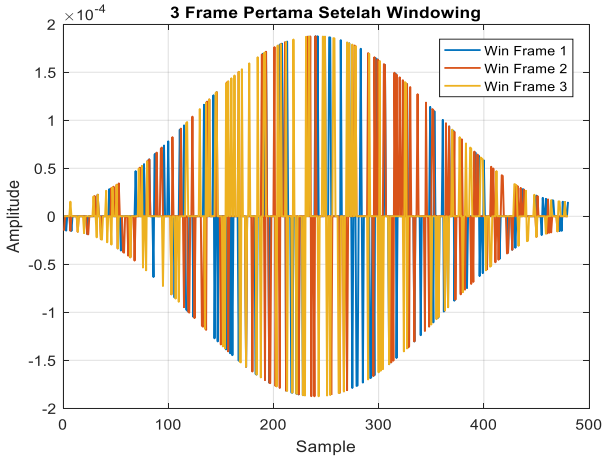

Figure 5. Frame graph after applying the Hamming window

After segmentation and windowing, the signals were ready for feature extraction.

C. Results of ZCR Feature Extraction

Zero Crossing Rate (ZCR) extraction measures the rate at which the speech signal crosses the zero-amplitude axis—changing from positive to negative or vice versa. ZCR is effective for distinguishing voiced from unvoiced segments. Voiced segments such as vowels typically have low ZCR due to periodic low-frequency content, whereas unvoiced segments or noise-like components exhibit high ZCR.

Figure 6 provides an example of ZCR extraction for the Indonesian utterance *"membaca"* (male speaker).

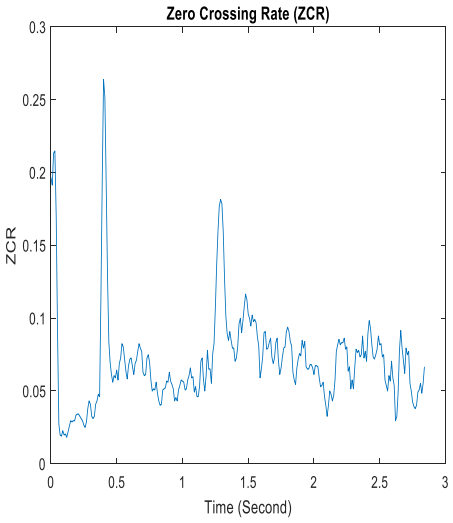

Figure 6. ZCR extraction graph for Indonesian speech signal

Figure 7 presents the ZCR extraction for the Bugis-language utterance *"na'baca"*.

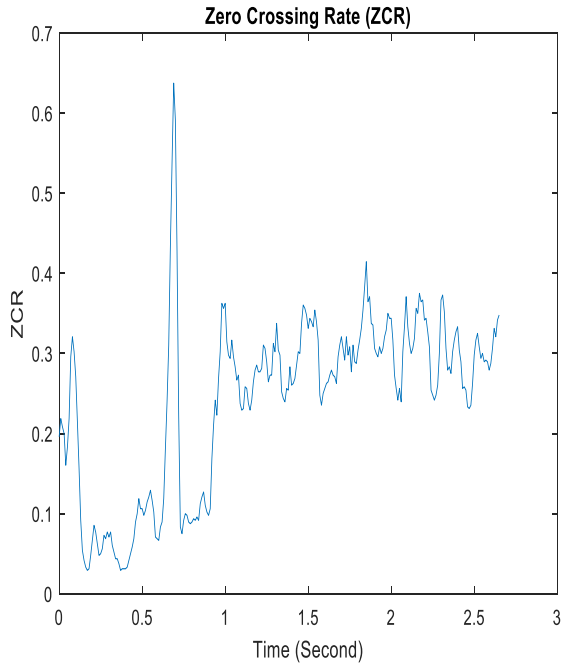

Figure 7. ZCR extraction graph for Bugis speech signal

Based on the test results presented, it can be observed that the ZCR values in voiced signal frames tend to be lower, whereas in unvoiced segments that resemble noise, the ZCR values tend to be higher. This indicates that the ZCR varies depending on the nature and context of the uttered speech.

Speech-signal analysis is one of the important approaches in digital signal processing, particularly for understanding the phonetic, acoustic, and prosodic characteristics of a language. A speech signal contains complex information stemming from both linguistic aspects and the physiological characteristics of the speaker. Therefore, various acoustic features have been developed to quantitatively describe the structure of speech signals. One of the fundamental yet highly significant features is the Zero Crossing Rate (ZCR), which measures the number of times a signal crosses the zero axis within a given time unit.

D. Results of the ZCR Algorithm for Descriptive Statistical Analysis of Indonesian and Bugis Speech Signals

This subsection presents the results obtained from the application of the Zero Crossing Rate (ZCR) algorithm to speech-signal data produced by male speakers in both Bahasa Indonesia and the Bugis language. All computational processes were carried out using MATLAB, which was used to calculate the descriptive statistical parameters—including mean, variance, standard deviation, skewness, and kurtosis—for each recorded speech signal stored in the database. These statistical indicators allow for a quantitative characterization of the temporal behavior of the speech signals and provide a basis for comparing the acoustic properties of the two languages.

1. Results of the ZCR Algorithm Test on Indonesian Speech Signals Using Descriptive Statistical Extraction. The ZCR algorithm testing on Indonesian speech-signal data for descriptive statistical extraction is presented in three main tables, each containing descriptive statistical values for the word *"membaca"*, the sentence *"berbicara bahasa Indonesia"*, and the phrase *"satu kata satu perbuatan"*. These values illustrate variations in the acoustic characteristics produced by each speaker, as reflected in the changes in mean values, the spread of variance, and the distribution patterns indicated by skewness and kurtosis. The descriptive statistical results obtained from the testing process are presented in the corresponding tables in this study.

   a) Results for Male Speakers

      The results obtained from the descriptive statistical analysis using the ZCR algorithm for Indonesian speech signals produced by male speakers are presented in the following tables:

Table 1. Descriptive statistical results of the ZCR algorithm for the Indonesian utterance *"membaca"* male speaker

SIGNAL DATA LABEL	MEAN	VARIANCE	STANDARD DEVIATION	SKEWNESS	KURTOSIS
DATA AK01	0.071400	0.001280	0.035779	2.322900	10.700800
DATA AS01	0.045300	0.000738	0.027173	3.083400	17.597200
DATA Ch01	0.304500	0.012896	0.113561	-0.587500	2.266500
DATA Dd01	0.178600	.0189030	0.137487	0.801700	2.935600
DATA Fd01	0.085000	0.001496	0.038680	1.053100	5.945500

Table 2. Descriptive statistical results of the ZCR algorithm for the Indonesian utterance *"berbicara bahasa Indonesia"* male speaker

SIGNAL DATA LABEL	MEAN	VARIANCE	STANDARD DEVIATION	SKEWNESS	KURTOSIS
DATA AK02	0.071300	0.003245	0.056964	1.710000	5.278400
DATA AS02	0.033800	0.001125	0.033535	2.755900	11.541500
DATA Ch02	0.232500	0.024879	0.157731	0.533400	2.464200
DATA Dd02	0.171900	0.020652	0.143707	0.959200	3.005900
DATA Fd02	0.102300	0.012515	0.111872	3.759700	18.672000

Table 3. Descriptive statistical results of the ZCR algorithm for the Indonesian utterance *"satu kata satu perbuatan "* male speaker

SIGNAL DATA LABEL	MEAN	VARIANCE	STANDARD DEVIATION	SKEWNESS	KURTOSIS
DATA AK03	0.060300	0.001682	0.041017	1.664500	5.552100
DATA AS03	0.038900	0.001165	0.034126	3.193300	13.922300
DATA Ch03	0.156500	0.011955	0.109340	1.986900	7.612400
DATA Dd03	0.243900	0.036144	0.190115	0.884900	3.248500
DATA Fd03	0.129700	0.020378	0.142751	2.825600	11.382100

Based on the results presented in these tables, the striking differences observed across all statistical parameters—particularly the mean, standard deviation (SD), and kurtosis—confirm that the ZCR feature has strong discriminative power for distinguishing speaker identities in Bahasa Indonesia.

Speaker AS is characterized by a very low mean ZCR and extremely high kurtosis, indicating a speech signal dominated by highly pure vowel components but containing very sharp outliers. In contrast, speaker Ch is characterized by a higher mean and a distribution that is closer to normal due to lower kurtosis, suggesting a signal more evenly dominated by higher-frequency components.

b) Results for Female Speakers

The results of the descriptive statistical analysis using the ZCR algorithm for Indonesian speech signals produced by female speakers are presented in the tables below:

Table 4. Descriptive statistical results of the ZCR algorithm for the Indonesian utterance *"membaca"* female speaker

SIGNAL DATA LABEL	MEAN	VARIANCE	STANDARD DEVIATION	SKEWNESS	KURTOSIS
DATA AF01	0.066300	0.001883	0.043390	3.801700	24.149600
DATA AL01	0.091500	0.006627	0.081408	4.992000	33.090500
DATA DE01	0.099005	0.008241	0.090780	3.298796	16.904398
DATA MU01	0.125600	0.006049	0.077778	4.482900	30.666600
DATA NA01	0.091331	0.003103	0.055709	2.319472	12.681080

Table 5. Descriptive statistical results of the ZCR algorithm for the Indonesian utterance *"brbicara bahasa Indonesia"* female speaker

SIGNAL DATA LABEL	MEAN	VARIANCE	STANDARD DEVIATION	SKEWNESS	KURTOSIS
DATA AF01	0.095200	0.010914	0.104468	3.159700	13.402500
DATA AL01	0.119200	0.012744	0.112889	3.228300	14.166300
DATA DE01	0.081100	0.003385	0.058180	2.456700	10.278200
DATA MU01	0.117300	0. 018947	0. 137649	3. 058100	11.949700
DATA NA01	0.103617	0.011983	0.109466	3.579934	18.094299

Table 6 Descriptive statistical results of the ZCR algorithm for the Indonesian utterance *"satu kata satu perbuatan"* female speaker

SIGNAL DATA LABEL	MEAN	VARIANCE	STANDARD DEVIATION	SKEWNESS	KURTOSIS
DATA AF01	0.096000	0.005461	0.073899	4.469800	28.720500
DATA AL01	0.102400	0.010330	0.101638	4.566900	26.042000
DATA DE01	0.064900	0.001384	0.037205	3.877000	23.933900
DATA MU01	0.106500	0.006669	0.081664	3.014900	15.954900
DATA NA01	0.093952	0.011987	0.109486	5.202344	32.259556

The ZCR data analysis for female speakers shows highly consistent characteristics, particularly in the distribution-shape parameters, which differ significantly from the ZCR characteristics of male speakers. The generally low and concentrated ZCR values indicate that female speech signals are dominated by voiced segments (vowels). This observation is consistent with the fact that the fundamental frequency (pitch) of female voices is typically higher, whereas ZCR,which measures zero crossingsis influenced primarily by harmonic frequencies that are usually lower in vowel sounds.

Comparatively (implicitly), the substantially higher kurtosis values observed in female speakers compared to the average male speakers indicate that the ZCR feature is highly effective as a gender discriminator, as the distribution profile is much sharper and more concentrated for female voices.

2. Results of the ZCR Algorithm Test on Bugis Speech Signals for Descriptive Statistical Extraction

The following tables present the results of the ZCR algorithm using a descriptive statistical approach,

generated through MATLAB processing of several words and sentences in the Bugis language. The subjects, or speakers, consist of both male and female respondents.

a) Statistical Results for Male Speakers

The tables below present the results of the ZCR algorithm applied to Bugis speech-signal data produced by male speakers:

Table 7 Descriptive statistical results of the ZCR algorithm for the Indonesian utterance *"na'baca"* male speaker

SIGNAL DATA LABEL	MEAN	VARIANCE	STANDARD DEVIATION	SKEWNESS	KURTOSIS
DATA A01	0.239600	0.012302	**0.110913**	-0.262900	2.949100
DATA B01	0.113600	0.006663	0.081627	1.907100	6.267200
DATA C01	0.207700	0.009228	0.096063	-0.612700	1.636500
DATA D01	0.137500	0.004953	0.070375	1.400900	5.028300
DATA U01	0.087600	0.003024	0.054994	5.210800	3.834280

Table 8 Descriptive statistical results of the ZCR algorithm for the Indonesian utterance *"mabi'cara ogi"* male speaker

SIGNAL DATA LABEL	MEAN	VARIANCE	STANDARD DEVIATION	SKEWNESS	KURTOSIS
DATA A02	0.072700	0.002224	0.047161	0.887300	3.985300
DATA B02	0.076700	0.000730	0.027012	2.222100	11.190700
DATA C02	0.153600	0.010050	0.100249	0.862300	2.933730
DATA D02	0.104100	0.009556	0.097757	1.432900	3.846100
DATA U02	0.090000	0.003621	0.060177	3.968600	26.930400

Table 9 Descriptive statistical results of the ZCR algorithm for the Indonesian utterance *"taro ada taro gau"* male speaker

SIGNAL DATA LABEL	MEAN	VARIANCE	STANDARD DEVIATION	SKEWNESS	KURTOSIS
DATA A03	0.220300	0.022207	0.149019	0.209800	1.326400
DATA B03	0.119200	0.005956	0.077176	1.697200	5.132500
DATA C03	0.206900	0.017166	0.131019	0.543100	1.915000
DATA D03	0.129500	0.005013	0.070805	0.803500	3.281800
DATA U03	0.080300	0.001180	0.034357	1.223400	5.355100

The following tables present the results of the ZCR algorithm applied to Bugis speech-signal data produced by male speakers:

The highest mean ZCR values were obtained by DATA A01 when pronouncing the word *"na'baca"*, and by DATA A03 when pronouncing the phrase *"taro ada taro gau"*. Both recordings produced mean values above 0.2000. This indicates that the speech signals of speaker A contain a larger proportion of consonantal elements or exhibit clearer consonant segmentation within the signal duration compared to other speakers.

Speaker **U** shows the most extreme skewness and kurtosis values (U01, U02), although the standard deviation remains low. This suggests that, although the signal is predominantly vowel-based (indicated by high skewness), the ZCR transitions between vowels and consonants occur infrequently but are very sharp (extreme kurtosis).

b) Statistical Results for Female Speakers

The results obtained from the descriptive statistical analysis using the ZCR algorithm for Bugis speech signals produced by female speakers are presented in the tables below:

Table 10 Descriptive statistical results of the ZCR algorithm for the Indonesian utterance *"na'baca"* female speaker

SIGNAL DATA LABEL	MEAN	VARIANCE	STANDARD DEVIATION	SKEWNESS	KURTOSIS
DATA A01	0.093000	0.009897	3.526900	3.526900	15.757400
DATA M01	0.242500	0.011055	0.105145	-0.396300	2.774000
DATA N01	0.246400	0.014640	0.120996	1.008200	5.829900
DATA R01	0.052100	0.001595	0.039941	3.528400	16.996500
DATA T01	0.086200	0.002024	0.044989	3.946500	28.512100

Table 11 Descriptive statistical results of the ZCR algorithm for the Indonesian utterance *"mabi'cara ogi"* female speaker

SIGNAL DATA LABEL	MEAN	VARIANCE	STANDARD DEVIATION	SKEWNESS	KURTOSIS
DATA A02	0.078100	0.002633	0.051314	4.550500	33.567200
DATA M02	0.191600	0.010615	0.103030	0.552700	5.451100
DATA N02	0.157600	0.009023	0.094988	1.624200	8.879900
DATA R02	0.112500	0.010307	0.101521	3.221800	14.488300
DATA T02	0.088200	0.004305	0.065616	5.192000	38.861100

Table 12 Descriptive statistical results of the ZCR algorithm for the Indonesian utterance *"taro ada taro gau"* female speaker

SIGNAL DATA LABEL	MEAN	VARIANCE	STANDARD DEVIATION	SKEWNESS	KURTOSIS
DATA A03	0.089100	0.001743	0.041754	1.918600	8.657000
DATA M03	0.207300	0.009987	0.099937	-0.206600	1.485300
DATA N03	0.206900	0.017166	0.131019	0.543100	1.915000
DATA R03	0.102200	0.001322	0.036362	1.543000	9.370500
DATA T03	0.080300	0.001180	0.034357	1.223400	5.355100

Speaker M (M01, M02, M03) consistently produced signals with the highest mean values among all speakers for the three phrases. This suggests that speaker M may have generated the strongest signal energy or spoken with higher amplitude compared to the other speakers during the recording process.

The descriptive statistical analysis of the ZCR algorithm applied to speech signals in Bahasa Indonesia and the Bugis language demonstrates that although both languages exhibit

ZCR mean values within relatively similar ranges, there are significant and consistent differences in signal variability and distributional shape. These differences make ZCR an effective feature for distinguishing between the two languages. Therefore, the ZCR mean, standard deviation, and skewness can be utilized as key feature vectors for effective recognition or classification systems involving Bahasa Indonesia and the Bugis language

IV. Conclusion

Based on the simulation results and data analysis conducted in this study, several conclusions can be drawn as follows:

1. The statistical characteristics of the ZCR feature in speech signals of Bahasa Indonesia and the Bugis language are proven to be highly informative and varied, both between the two languages and between genders. The ZCR of male speakers in Bahasa Indonesia is marked by a highly dynamic mean or the widest mean range, as well as the highest standard deviation (SD), with a maximum value of 0.190115. This also implies the presence of highly contrasting acoustic transitions occurring between vowel and consonant segments. Meanwhile, in female speakers of Bahasa Indonesia, ZCR reflects distinct acoustic transitions influenced by vocal articulation patterns.
2. The ZCR feature demonstrates significant discriminative power between Bahasa Indonesia and the Bugis language. In the Bugis language, ZCR values more frequently exhibit negative skewness observed in two cases which indicates that Bugis speech tends to show ZCR distributions concentrated at higher frequencies. This may be attributed to specific consonantal phonetic patterns.

Acknowledgement

We would like to express our sincere gratitude to the Rector of Universitas Muslim Indonesia, the LP2S Research Institute of Universitas Muslim Indonesia, the Dean of the Faculty of Engineering, and the Head of the Electrical Engineering Study Program for their support and assistance throughout the completion of this research.

References

[1] Yudi Hartono S, Suyuti S. Method of extracting speech characteristics of Bugis regional language. J Teknol Elektron. 2025;22(1):35–40.
[2] Yudi Hartono S, Suyuti S. Ekstraksi pitch sinyal wicara (speech) bahasa daerah Bugis dengan algoritma autokorelasi. J INSTEK. 2024;9(2):125–54.
[3] Hartono YS, Basalamah A. Ekstraksi pitch pada suara laki-laki dan perempuan dengan metode algoritma autokorelasi dan AMDF untuk penentuan tipe jenis jangkauan suara berdasarkan tipe suara. Penelitian Dosen Intern PUF, Fakultas Teknik, Universitas Muslim Indonesia; 2016.
[4] Sun Y, Zhang Z, Schuller BW. On the analysis of speech emotion recognition using deep spectrum features with LSTM networks. IEEE Access. 2020;8:165811–23.
[5] Mohd Hanifa, R., Isa, K., Mohamad, S., Shah, S. M., Nathan, S. S., Ramle, R., & Berahim, M. (2024). Voiced and unvoiced separation in Malay speech using zero crossing rate and energy**.** *Indonesian Journal of Electrical Engineering and Computer Science:16(2):775-780.*
[6] Endah, S. N. et al. **(2022)**. Continuous speech segmentation menggunakan dynamic thresholding pada fitur short-term (ZCR, energy, spectral flux); akurasi segmentasi kata mencapai 97,5 %, Jurnal of Engineering Science and Technology. 2022.:4(7):2719-2935
[7] Gunawan A, Riza LS, Arifianto A. Speech recognition system using MFCC and Dynamic Time Warping for Bahasa Indonesia. J Ilmu Komput dan Inform. 2021;14(1):37–44.
[8] Chen J, Wang Q, Xiao X, et al. End-to-end speech emotion recognition using 1D convolutional long short-term memory networks. Int J Speech Technol. 2021;24:695–704.
[9] Pardede HF, Adhi P, Zilvan V, Ramdan A, Krisnandi D. Deep convolutional neural networks-based features for Indonesian large vocabulary speech recognition. *International Journal of Artificial Intelligence (IJ-AI)*. 2023;12(2):610–oi 10.11591/ijai.v12.i2.pp610-61710.
[10] Dura M, Kadiri MS, Diouri O. An improved voice command recognition system using MFCC and deep learning. Procedia Comput Sci. 2020;170:718–23.
[11] Putra WR, Amalia F. Analisis ciri akustik vokal Bahasa Bugis dan Bahasa Indonesia menggunakan metode Zero Crossing Rate dan Spectral Centroid. J Ilm Teknol Inform. 2023;9(2):150–9.
[12] Rahman A, Nugroho A. Deteksi gender suara manusia menggunakan metode Zero Crossing Rate dan algoritma k-NN. J Fotonik. 2022;10(1):11–8.
[13] Nugraha R, Setyawan I. Penerapan ekstraksi ciri ZCR dan Energy untuk klasifikasi suara manusia. J Teknol dan Sistem Komput. 2022;10(3):321–7.
[14] Huang Z, Epps J, He L. An investigation of cross-language speech emotion recognition. Speech Commun. 2020;120:20–36.
[15] Putri DM, Siregar A. Klasifikasi suara berdasarkan jenis kelamin menggunakan algoritma Zero Crossing Rate dan Naive Bayes. J Teknol Inf dan Komun. 2023;7(1):75–81.
[16] Arsyad M, Nurfadillah S. Implementasi Zero Crossing Rate dan Energy untuk identifikasi ujaran frustasi. J Rekayasa Elektron dan Komput. 2024;5(2):88–93.
[17] Iskandar Y, Pratama S. Analisis karakteristik sinyal wicara berbasis bahasa daerah menggunakan metode MFCC dan ZCR. J Teknol Informasi. 2023;12(2):141–8.

[18] Latief R, Salam U. Pengenalan ucapan bahasa Bugis dengan metode deep learning berbasis fitur ZCR dan MFCC. J Teknol Informasi dan Ilmu Komputer. 2025;8(1):101–8.
[19] Safitri R, Wirawan B. Klasifikasi emosi suara Bahasa Indonesia menggunakan CNN dan fitur ZCR. Jurnal Sains Komputer. 2023;4(1):23–31.
[20] Lee J., Kim S.,Statistical Analysis of Acoustic Features Including Zero Crossing Rate for Speaker Identification IEEE Acess,2021;9;14567-14577.