## DEVELOPMENT OF TALKING DRUMS DATASET FOR AI PATTERNS GENERATION

**BY**

**FOWE, ABAYOMI DAVID**

**239074108**

**DEPARTMENT OF COMPUTER SCIENCES**

**FACULTY OF COMPUTER SCIENCES**

**UNIVERSITY OF LAGOS**

**MIT 817 – SEMINAR ON CURRENT TOPICS IN INFORMATION TECHNOLOGY**

M.Sc. I.T

**AUGUST, 2025**

# CERTIFICATION

|  |  |
| --- | --- |
| I hereby certify that this Project “DEVELOPMENT OF TALKING DRUMS DATASET FOR AI PATTERNS GENERATION”was carried out by me, FOWE, ABAYOMI DAVID, Department of computer science under the supervision of; ................................................. .................................................  **DR. B. O. OKUNOYE** Date  Supervisor  ........................................ ........................................ | |
| **DR. CHIKA YINKA-BANJO** | Date |

HOD

## DEDIICATION

## I hereby dedicate this work to the Almighty God, the giver of life and all things for the successful completion of my M.Sc. Program. To Him alone be all the glory, now and forever, Amen.

## ACKNOWLEDGEMENTS

First, I will like to Give God all the glory, for the strength and ability bestowed on me to commence and complete my M.Sc. academic pursuit. I will like to appreciate my Esteemed Supervisor, Dr. B.O Okunoye, for the guidance and inspiration that birthed this piece of work. I cannot but thank my Class Governor, and friend, Majesty Okpara, for always reaching out a helping hand whenever I get stuck. Evelyn, the love of my life, and my wonderful children, Stella, Melody and Israel, for your love and understanding. I love you all. These accolades cannot be complete without mentioning my family members, Iya Fowe, Baba Fowe, Mum Tosin and Uncle Jerry for your support, Mrs Coker, Adewale (Wale Lyrics), for the unflinching support and then Adeola Ogunmuyiwa . I love you all.

## TABLE OF CONTENTS

## Table of Contents

## Title Page ......................................................... i Certification ....................................................... ii Dedication .......................................................... iii Acknowledgements ................................................. iv Abstract .............................................................. v

## Chapter One: Introduction

## 1.1 Background of the Study ...................................... 1 1.2 Problem Statement ................................................. 4 1.3 Objectives of the Study ............................................ 6   1.3.1 Main Objective .................................................. 6   1.3.2 Specific Objectives ............................................. 6 1.4 Significance of the Study ........................................... 8 1.5 Scope of the Study .................................................. 10 1.6 Limitations of the Study ........................................... 11

## Chapter Two: Literature Review

## 2.1 Overview of Talking Drums in African Culture ............ 13 2.2 AI in Music Generation ............................................. 15 2.3 Existing Datasets for Music and Percussion ................ 18 2.4 Research Gaps and Motivation .................................. 22 2.5 Summary of Literature Review .................................. 24

## Chapter Three: Methodology

## 3.1 Research Design .................................................... 26 3.2 Data Collection Methods .......................................... 28 3.3 Annotation and Labeling Strategy ............................... 31 3.4 Dataset Structuring and Storage ................................ 34 3.5 Tools, Frameworks, and Technologies ......................... 37 3.6 AI Integration and Demonstration ................................. 40 3.7 Ethical Considerations ............................................... 42

## Chapter Four: Results and Discussion

## 4.1 Dataset Overview ................................................... 45 4.2 Sample Recordings and Annotations ........................ 47 4.3 AI Model Experiments (GANs, RNNs, Transformers) .... 50 4.4 Evaluation of Results ................................................ 54 4.5 Discussion of Findings .............................................. 57

## Chapter Five: Conclusion and Recommendations

## 5.1 Summary of the Study ............................................. 61 5.2 Contributions of the Project ........................................ 63 5.3 Limitations ............................................................. 65 5.4 Recommendations for Future Research ....................... 66 5.5 Conclusion ............................................................ 68

**References**...........................................................70  
**Appendices** ............................................................ 75

* Appendix A: Sample Data Annotation Schema
* Appendix B: Screenshots of Tools/Software Setup
* Appendix C: Consent Form for Drum Recordings

## LIST OF FIGURES

Contains all the figures and their respective pages in your project report

## LIST OF TABLES

## Abstract

The talking drum, notably the Yoruba *dùndún* or *gángan*, is a distinctive West African percussion instrument capable of mimicking human speech by manipulating pitch, rhythm, and intensity—effectively functioning as a speech surrogate in tonal languages like Yoruba. Acoustic studies confirm a strong correlation between drum-produced tonal contours and spoken Yoruba tones, reinforcing the drum’s role in conveying linguistic meaning through sound ([phys.org](https://phys.org/news/2021-07-shown-accurately-mimic-speech-patterns.html?utm_source=chatgpt.com), [Wikipedia](https://en.wikipedia.org/wiki/Talking_drum?utm_source=chatgpt.com), [figshare.com](https://figshare.com/articles/dataset/DataSheet1_The_Language_of_G_ngan_A_Yor_b_Talking_Drum_ZIP/16866547/1?utm_source=chatgpt.com)).

It is widely known that Yorùbá drummers communicate through their native drums. This  
paper investigates the grammar of gángan, which belongs to a family of Yoruba drums called  
dùndún. The results of this study show that Yorùbá drummers represent the phonetic  
realization of lexical and grammatical tones of their language with the drum. Statistically, the  
speech tones and the acoustic correlate of the corresponding drum representations have a  
significant positive relationship. In both spoken and drum communication, vowel (V) and  
consonant-vowel (CV) prosodic units have different statuses. To conclude, Yorùbá  
drummers communicate via the gángan drum by transposing certain phonemic features  
and maybe phonological conditions of their language to musical forms.

Building on this cultural and acoustic foundation, we present the development of a talking-drums audio dataset expressly designed for training AI systems in pattern and rhythm generation. The dataset comprises high-quality recordings of Yoruba talking-drum performances aligned with their corresponding spoken phrases. Audio is annotated with tone labels (H/M/L), segmented rhythmic structures, and metadata including drummer identity, drum type (e.g., *dùndún*, *gángan*, *bàtá*), recording mode (speech-like vs. musical), and acoustic measures (frequency, intensity) ([Frontiers](https://www.frontiersin.org/journals/communication/articles/10.3389/fcomm.2021.652542/full?utm_source=chatgpt.com), [journals.openedition.org](https://journals.openedition.org/lla/347?utm_source=chatgpt.com), [figshare.com](https://figshare.com/articles/dataset/DataSheet1_The_Language_of_G_ngan_A_Yor_b_Talking_Drum_ZIP/16866547/1?utm_source=chatgpt.com)).

We outline our methods for data acquisition, audio processing, and annotation—drawing inspiration from linguistic and organological approaches used in previous studies of speech surrogacy systems ([Frontiers](https://www.frontiersin.org/journals/communication/articles/10.3389/fcomm.2021.652542/full?utm_source=chatgpt.com), [journals.openedition.org](https://journals.openedition.org/lla/347?utm_source=chatgpt.com)). Leveraging these annotations, we demonstrate how the dataset can support AI-driven tasks such as:

Neural drum pattern generation: Using sequence-to-sequence models to translate speech or tone sequences into rhythmic drum patterns.

Cross-modal synthesis: Generating drum rhythms from speech input or vice versa.

Tone-sensitive AI: Enhancing rhythm generation with formal tone mapping.

Drawing parallels from prior work in musical AI—such as drum-loop generation from language cues ([arXiv](https://arxiv.org/abs/2209.10016?utm_source=chatgpt.com" \o "Setting the rhythm scene: deep learning-based drum loop generation from arbitrary language cues)) and leveraging language-model fine-tuning for drum composition ([arXiv](https://arxiv.org/abs/2301.01162?utm_source=chatgpt.com" \o "Language Models are Drummers: Drum Composition with Natural Language Pre-Training))—we position our talking-drums dataset as a unique, culturally rich resource that bridges linguistic tonal patterns and rhythmic AI creativity.

Finally, we discuss plans for expanding the dataset to include additional dialects and drumming modes, and explore potential applications such as AI-assisted music composition, preservation of musical heritage, and tonal-language research.

# 1.0 CHAPTER 1

## **1.1 Introduction**

Music has always played a central role in human society, functioning not only as a form of entertainment but also as a medium for communication, identity, and cultural preservation. Across cultures, musical instruments have historically been used to transpose linguistic features into melodies, enabling them to function as speech surrogates. Examples include the Sino-Tibetan gourd organ (Bradley, 1979), the Hmong *raj* of Southeast Asia (Poss, 2005; 2012), the Asante ivory trumpet of Ghana (Kaminski, 2008), the *Sabar* of Senegal (Winter, 2014), the balafon of the Sambla people in Burkina Faso (McPherson, 2018), and the Amazon Bora system (Seifart et al., 2018). Stern (1957) distinguishes between surrogate systems that represent phonemic aspects of language (“abridged”) and those that encode meaning without phonemic reference (“lexical ideogram”). This study focuses on the former, particularly within the Yorùbá cultural context [(Akinbo](https://www.researchgate.net/profile/Samuel-Akinbo?_tp=eyJjb250ZXh0Ijp7InBhZ2UiOiJwdWJsaWNhdGlvbiIsInByZXZpb3VzUGFnZSI6bnVsbH19), 2021).

Among the wide range of musical instruments capable of serving as speech surrogates (Lo-Bamijoko & Joy, 1987; Agawu, 2016), the talking drum (*dùndún*) of the Yorùbá people stands out for its unique ability to encode the tonal patterns of Yorùbá speech. By adjusting the tension of the hourglass-shaped drum with leather cords, drummers skillfully manipulate pitch and rhythm to mimic the phonemic and prosodic features of speech (Beier, 1954; Euba, 1967, 1990; Villepastour, 2010). This ability has made the talking drum a central instrument of communication, historically used for relaying messages across distances, reciting praise poetry, and preserving oral traditions. As McPherson (2019) notes, studying such speech surrogates through a linguistic rather than purely musical lens offers valuable insights into the representation of phonological features, thereby providing language-external evidence for phonological theory [(Akinbo](https://www.researchgate.net/profile/Samuel-Akinbo?_tp=eyJjb250ZXh0Ijp7InBhZ2UiOiJwdWJsaWNhdGlvbiIsInByZXZpb3VzUGFnZSI6bnVsbH19), 2021).

Despite its cultural and linguistic richness, the talking drum remains underrepresented in computational musicology and Artificial Intelligence (AI) research (*Akinbo, 2021)*. Contemporary AI systems for music generation predominantly rely on datasets and models derived from Western instruments and genres. For example, DrumGAN (Nistal et al., 2020) employs generative adversarial networks to synthesize percussive sounds, while DeepDrummer (Forsgren & Martiros, 2018) applies reinforcement learning to rhythm generation. Likewise, large corpora such as MAESTRO (piano) and the Groove MIDI Dataset (Western drum kit rhythms) have powered breakthroughs in symbolic music modeling with Recurrent Neural Networks (RNNs), Transformers, and Variational Autoencoders (VAEs). However, African instruments are largely excluded from such repositories, reinforcing a digital imbalance in global music representation. Recent analyses reveal that over 86% of AI music training data originates from the Global North, with only 14.6% representing the Global South (Arxiv, 2024).

The development of a structured talking drum dataset for AI pattern generation thus addresses both a cultural and technical gap. The dataset will include high-quality recordings of drum performances, annotated with pitch contours, rhythmic structures, tempo variations, and communicative functions (e.g., greetings, proverbs, praise poetry). Such detailed annotation opens multiple research pathways: GANs can model culturally faithful timbres, RNNs and Transformers can capture long-term dependencies in call-and-response patterns, and VAEs can map latent spaces of tonal-rhythmic expression, enabling smooth interpolation between motifs. Furthermore, human-in-the-loop frameworks may enhance authenticity by integrating expert drummer feedback (Zhu et al., 2023).

Beyond technical advancement, this project contributes significantly to cultural preservation and innovation. By digitally archiving the communicative and musical hybridity of the talking drum, the dataset ensures the survival and global visibility of African musical traditions. Moreover, it provides a foundation for creative applications in AI-assisted composition, heritage education, and cross-cultural music fusion. Emerging initiatives such as KorinAI, which develops AI models grounded in African voices and sound libraries (Techbuild Africa, 2025), underscore the importance of culturally inclusive AI systems.

In summary, the development of a talking drum dataset for AI pattern generation is both a scholarly and cultural imperative. By bridging indigenous African knowledge systems with cutting-edge computational models such as GANs, VAEs, and sequence-based neural architectures, this project advances AI music research while ensuring that African musical heritage is preserved, recognized, and globally appreciated.

## 1.2 Problem Statement

Artificial Intelligence has demonstrated significant potential in music composition, sound synthesis, and rhythm generation. However, these advances are disproportionately built upon datasets dominated by Western music, marginalizing non-Western instruments and traditions. While datasets like MAESTRO (piano), NSynth (multi-instrument audio), and Groove MIDI Dataset (drum kit rhythms) have enabled AI breakthroughs, there is no equivalent structured dataset for African instruments such as the talking drum.

The absence of a standardized talking drum dataset creates several challenges:

* Cultural Underrepresentation: African instruments remain absent from global AI music research.
* Technical Limitation: AI models trained only on Western datasets cannot generalize to the tonal and linguistic complexities of African drumming.
* Preservation Risks: Without digital documentation, indigenous traditions risk being overlooked or forgotten in future technological systems.
* Research Gap: Ethnomusicologists and computational music researchers lack annotated data for studying the interplay of language, rhythm, and sound in talking drums.
* Unless this gap is addressed, the talking drum—despite its significance—may remain invisible in AI-driven music applications.

## 1.3 Objectives of the Study

## Main Objective: **To develop a structured and annotated dataset of talking drum recordings suitable for AI pattern generation.**

## Specific Objectives:

* To record and digitize high-quality audio samples of talking drum performances across different rhythms, tempos, and communicative contexts.
* To annotate recordings with relevant musical and linguistic features such as pitch contour, tempo, rhythmic patterns, and tonal inflections.
* To design a standardized dataset structure (audio + metadata) that facilitates AI training and experimentation.
* To demonstrate the use of the dataset in AI models for rhythm and sound generation, using GANs, RNNs, and Transformer-based architectures.
* To contribute to the preservation of African musical heritage through digital archiving and global accessibility.

## 1.4 Significance of the Study

* Academic Significance: Provides researchers with the first structured dataset of talking drums, enabling cross-disciplinary studies in computational musicology, ethnomusicology, and AI.
* Cultural Significance: Preserves African oral traditions and ensures the talking drum is recognized in digital cultural archives.
* Technological Significance: Enables development of AI systems capable of generating African rhythms, enhancing diversity in AI creativity.
* Educational Significance: Serves as a teaching resource for students of music, linguistics, and AI by providing annotated real-world data.
* Socioeconomic Significance: Creates opportunities for African musicians and innovators to develop culturally grounded digital tools, boosting creative industries.

## 1.5 Scope of the Study

This project focuses on the development of a structured talking drum dataset for AI pattern generation. The dataset will include:

* High-quality audio recordings from live performances.
* Annotation of rhythmic, tonal, and contextual features.
* Organization into a format compatible with AI frameworks such as TensorFlow and PyTorch.

The scope does not extend to building a full AI music generation platform; rather, the dataset will serve as a foundation upon which researchers and developers can train and evaluate AI models.

## 1.6 Limitations of the Study

* Data Size: The dataset may be smaller compared to large-scale Western datasets due to limited access to professional drummers and recording resources.
* Environmental Factors: Background noise during recordings may affect data quality.
* Complexity of Tonal Systems: Capturing the full linguistic depth of Yorùbá and other tonal languages may be challenging.
* Generalization: AI models trained on the dataset may initially struggle to generalize across drummers with different styles.