

MAKERERE



UNIVERSITY

**ANALYSIS OF TONE ERRORS IN THE ACQUISITION OF CHINESE AMONG
UGANDAN L2 SPEAKERS**

BY

WANG WEI

2022/HD03/1541X

**A DISSERTATION SUBMITTED TO THE DIRECTORATE OF GRADUATE
TRAINING IN PARTIAL FULFILMENT OF THE REQUIREMENTS FOR THE
AWARD OF THE DEGREE OF MASTER OF ARTS IN LINGUISTICS OF
MAKERERE UNIVERSITY**

SEPTEMBER, 2026

DECLARATION

By submitting this dissertation, I WANG WEI, declare that the entirety of this dissertation titled *Analysis of Tone Errors in the Acquisition of Chinese among Ugandan L2 Speakers* is my original work, and has not been submitted for any other degree award to any other university before.

Signature: Wang Wei

Wang Wei

Date: 28/9/2026

APPROVAL

This is to certify that this dissertation titled *Analysis of Tone Errors in the Acquisition of Chinese among Ugandan L2 Speakers* has been submitted for examination with our approval as supervisors.

Name:

Signature:

Date:

Celestino Oriikiriza (PhD)

Celestino Oriikiriza

28.9.2026

Deo Kawalya (PhD)

Deo Kawalya

28/09/2026

DEDICATION

This research project is dedicated to my parents for their support. It is also dedicated to my dear husband for his unconditional love and support.

ACKNOWLEDGEMENTS

I so much appreciate my supervisors Dr. Celestino Oriikiriza and Dr. Deo Kawalya for their patience, useful guidance, and enormous support during the process of my research project.

Special thanks go to all my lecturers, Dr. Levis Mugumya, Dr. Michael Wangotta, Dr. Bayiga Florence and Dr. Lukwago Moses, best wishes to you all.

My sincere gratitude goes to the Head of Department of Linguistics, English Language Studies and Communication Skills, Dr. Levis Mugumya, and all members of the Department for your support. I will be grateful forever.

I also thank Prof. Oswald Ndoleriire for his support and guidance.

I also would like to thank my fellow classmates -- thank you for your understanding, sharing of materials and ideas, accompanying me to new places and helping me to overcome the language barrier. May God bless you!

TABLE OF CONTENTS

DECLARATION	I
APPROVAL	I
DEDICATION	II
ACKNOWLEDGEMENTS	III
LIST OF TABLES	VII
LIST OF FIGURES	VII
LIST OF CHARTS	IIIX
LIST OF APPENDICES	X
LIST OF ABBREVIATIONS AND ACRONYMS	XI
LIST OF SYMBOLS	XII
ABSTRACT	1
CHAPTER ONE: INTRODUCTION	2
1.1 Background to the Study	2
1.1.1 Chinese Language Education in Uganda	2
1.1.2 Mandarin Chinese Tone	2
1.1.3 Tone in the Bantu Languages	3
1.1.4 Tone in the Central Sudanic Language	4
1.1.5 L2 Chinese Tone Production and the Research Gap	4
1.2 Statement of the Problem	5
1.3 Objectives of the Study	6
1.3.1 General Objective	6
1.3.2 Specific Objectives	6
1.4 Significance of the Study	6
1.5 Justification of the Study	7
1.6 Theoretical Framework	7
1.7 Scope of the Study	14
1.7.1 Geographical Scope	14
1.7.2 Content Scope	14
1.7.3 Theoretical Scope	15

1.7.4 Temporal Scope	15
1.7.5 Linguistic Scope	15
1.8 Summary	15
CHAPTER TWO: LITERATURE REVIEW	16
2.1 Introduction	16
2.2 Tone in Chinese	17
2.3 Non-native Speakers' Production of Chinese Tones	18
2.4 Production of Chinese Tones by Its Learners in Uganda	22
2.5 Critical Evaluation	26
CHAPTER THREE: METHODOLOGY	28
3.1 Research Design	28
3.2 Research Approach	28
3.3 Population and Sampling	30
3.3.1 Target Population	30
3.3.2 Sample Size	30
3.3.3 Sampling Technique	31
3.4 Data Collection Procedures	32
3.4.1 Data Collection Techniques	32
3.4.2 Instruments	32
3.4.3 Materials	33
3.4.4 Research Procedure	33
3.5 Quality Control	34
3.5.1 Validity	34
3.5.2 Reliability	35
3.6 Data Analysis	35
3.7 Ethical Considerations	36
3.8 Summary	37
CHAPTER FOUR: DATA PRESENTATION AND ANALYSIS	38
4.1 Introduction	38
4.2 Tone Production Errors Made by Ugandan L2 Speakers of Chinese	41

4.2.1 Pitch-range Expansion and Overshoot	42
4.2.2 Dipping-contour Truncation and Reduction of Complex Movement	45
4.2.3 Tone 2 and Tone 3 Confusion	48
4.2.4 Duration Reduction	51
4.2.5 Delayed Rise Onset and Alignment in Tone 2	54
4.2.6 Citation-form Persistence in Sandhi Contexts	55
4.2.7 Neutral-tone Reduction Failure	56
4.3 Distribution of Tone Production Errors	57
4.3.1 Acoustic Basis of Tone Production Errors	57
4.3.2 Distributional Patterns of Tone Production Errors	59
4.4 Comparison with Previous Studies	68
4.5 Summary	70
CHAPTER FIVE: SUMMARY OF FINDINGS, CONCLUSIONS, AND RECOMMENDATIONS	71
5.1 Summary of Findings	71
5.1.1 Tone Production Errors	72
5.1.2 Distribution of Tone Production Errors	73
5.2 Limitations of the Study	76
5.2.1 Modest and Institutionally Bounded Participant Pool	76
5.2.2 Production-Centered and Controlled Evidence	76
5.3 Original Contribution to Knowledge	76
5.4 Conclusions	77
5.5 Recommendations	77
5.6 Pointers for Future Research	78
REFERENCES	79
APPENDICES	86

LIST OF TABLES

Table No.	Title
Table 1	Participant Profiles and Monosyllabic Tonal Performance Targets
Table 2	Mean Acoustic Deviation in Pitch Range (ΔF_0) for Monosyllabic Tones
Table 3	Production Confusion Matrix (Row-Normalized Probabilities)
Table 4	T2 and T3 Confusion Share Calculation
Table 5	Duration Comparison between Speaker and Native Tone Productions
Table 6	Accuracy Rates for Sandhi and Neutral Tone Contexts
Table 7	Monosyllabic Range-Similarity by L1 Background (Derived from ΔF_0)
Table 8	Context Comparison (Per-syllable Normalized) between Monosyllabic and Disyllabic Items
Table 9	Distribution of Errors by Nature and Mechanism
Table 10	Cross-Speaker Tone 1 Register Anchoring and Migration

LIST OF FIGURES

Figure No.	Title
Figure 1	Five-Degree Standard Tone System
Figure 2	Student D's Tone 2 Structures
Figure 3	Student L's Tone 2 Structures
Figure 4	Student M's Tone 3 Structures
Figure 5	Delayed Rise Onset in F_0 Trajectories (Tone 2)

LIST OF CHARTS

Chart No.	Title
Chart 1	Chart 1 Error Types of Sandhi <i>yi</i> 'one'
Chart 2	Chart 2 Error Types of Sandhi <i>bu</i> 'not'
Chart 3	Chart 3 Error Types of Neutral Tones

LIST OF APPENDICES

Appendix No.	Title
Appendix I	Field Acceptance Letter from the Department
Appendix II	Experimental Object (Learners of Mandarin in Uganda)
Appendix III	Experimental Object (Native Speaker of Mandarin)
Appendix IV	Recording Materials
Appendix V	Data Structure and Variable Organization for the Production Dataset

LIST OF ABBREVIATIONS AND ACRONYMS

A-LEVEL	— Advanced Level
CHUSS	— College of Humanities and Social Sciences
CI	— Confidence Interval
CI at MAK	— Confucius Institute at Makerere University
DGT	— Directorate of Graduate Training
F ₀	— Fundamental Frequency
HSK	— Hanyu Shuiping Kaoshi (Chinese Proficiency Test)
L1	— First Language
L2	— Second Language
LIT	— Luyanzi Institute of Technology
LME	— Linear Mixed-Effects Model
Mak	— Makerere University
MP3	— MPEG-1 Audio Layer III (Audio Format)
NCDC	— National Curriculum Development Center
PPM	— Pitch Perception Model
PSC	— Putonghua Shuiping Ceshi (Mandarin Proficiency Test)
R	— R (Statistical Computing Language)
SPSS	— Statistical Package for the Social Sciences
ST	— Semitone
TCSL	— Teaching Chinese as a Second Language
NCHE	— National Council for Higher Education
WAV	— Waveform Audio File Format

LIST OF SYMBOLS

Acoustic Measurement Symbols

- F_0 — Fundamental Frequency (the physical correlate of pitch, measured in Hertz)
- ΔF_0 — Tonal Range Difference (the difference between maximum and minimum F_0 within a syllable, measured in Hertz)
- Hz — Hertz (the unit of frequency, one cycle per second)
- ms — Millisecond (one-thousandth of a second, used for measuring phonetic duration)
- ST — Semitone (a unit of measure for intervals in pitch, used for normalized acoustic analysis)

Tonal System Analysis Symbols

- H — Entropy (a measure of uncertainty or disorder in a tonal system, measured in bits)
- MS — Migration Strength (the proportion of tones that are misclassified in a system)
- Stab — Stability (a measure of the internal consistency of a tonal system)

Statistical Analysis Symbols

- CI — Confidence Interval (statistical precision range)
- n — Sample Size (number of observations or participants)
- p — Probability value for statistical significance

ABSTRACT

The acquisition of Chinese tones by Ugandan L2 speakers remains insufficiently studied. In teaching practice, Ugandan L2 speakers of Chinese persistently struggled to produce tones in isolated syllables, disyllabic words and tone-conditioned environments. Therefore, the study aimed at analyzing tone production errors made by Ugandan L2 speakers of Chinese by identifying these errors and describing their distribution. The study was guided by the Pitch Perception Model. The scope of analysis covered the production of monosyllabic tones, disyllabic tone combinations, and tone-conditioning environments such as *yī* 'one', *bù* 'not', and neutral tone. Speech was recorded from 19 intermediate participants from six first-language backgrounds (Luganda, Lusoga, Lugisu, Runyankore, Samia, and Lugbarati), and the resulting recordings provided the data. The data were analyzed acoustically against a native Mandarin Chinese baseline using F_0 contour, pitch-range excursion, and duration as the main descriptive measures. The findings show that tonal errors were systematic rather than random. The overall difficulty hierarchy in production of Chinese tone by Ugandan speakers of Chinese was observed in the production of Tone 2, Tone 3, Tone 4, and Tone 1 of Chinese, with the most persistent instability concentrated in the Tone 2 and Tone 3 contrast. Recurrent error patterns included contour overshoot, truncation of Tone 3, delayed rise onset in Tone 2, duration compression, citation-form persistence in sandhi environments (i.e. context-conditioned tonal alternation), and failure to reduce neutral tone appropriately. Error distribution also varied across L1 backgrounds and became more pronounced in disyllabic and morphophonemic contexts. The study argues that these patterns are explained through Pitch Perception Model (PPM), since tone-production difficulty was closely related to perceptual salience, contour complexity, pitch-register control, temporal support, and contextual tonal pressure. The study contributes empirical evidence from an underrepresented Ugandan context and suggests that Chinese tone instruction should prioritize the Tone 2 and Tone 3 contrast, contour timing, and context-sensitive tonal realization.

CHAPTER ONE: INTRODUCTION

1.1 Background to the Study

1.1.1 Chinese Language Education in Uganda

Over the past decade, economic, educational, and cultural relations between China and Uganda have deepened significantly. In the field of education, the establishment of the Confucius Institute at Makerere University (CI at Mak) in 2014 marked an important stage in the institutional development of Chinese language education in Uganda (Makerere University, 2014, 2024). Chinese language education has since expanded through university-based programmes, cultural exchange activities, and formal language training. Makerere University (Mak) later introduced the Bachelor of Chinese and Asian Studies (BCAS), a three-year programme approved by Mak and the National Council for Higher Education (NCHE), which started in the 2019/2020 academic year (Makerere University College of Humanities and Social Sciences (CHUSS)). These developments are broadly consistent with Uganda Vision 2040, which emphasizes the transformation of Uganda into a modern and prosperous country through human-capital development, international engagement and competitiveness (National Planning Authority, 2013). In this wider educational context, Chinese language learning has become increasingly relevant in Uganda.

1.1.2 Mandarin Chinese Tone

With the growth of Chinese language education, pronunciation has become an important area of pedagogical and linguistic concern. In Mandarin Chinese, lexical tone is especially important because pitch variation functions at the lexical level. Standard Mandarin Tones¹ are different from sentence-level intonation: they are syllable-level contrasts that help distinguish lexical meaning. The four citation tones of Standard Mandarin are usually described as Tone 1 (T1), a high-level tone; Tone 2 (T2), a rising tone; Tone 3 (T3), a falling-rising or low-dipping tone; and Tone 4 (T4), a falling tone.

¹ In this thesis, the term *Standard Mandarin Tones* refers to the four lexical tones of Standard Mandarin, whose tonal differences are meaning-distinguishing rather than merely sentence-level intonational variation.

In the traditional five-degree notation, these tones are commonly represented as 55, 35, 214 and 51 respectively (Chao, 1948; Yip, 2002). For example, the syllable *ma* may correspond to different lexical meanings when produced as *mā* ‘mother’, *má* ‘hemp’, *mǎ* ‘horse’, or *mà* ‘scold’. This shows that tone is not an optional prosodic feature in Chinese, but an essential component of syllable meaning.

Learning Mandarin Chinese tone becomes more complex when tones are examined across different phonological environments. In monosyllabic words, tones are usually produced in citation form, so speakers mainly need to control the pitch height and pitch contour of a single syllable. In disyllabic words, however, one tone may be affected by the neighbouring tone, syllable duration, speech rhythm and word position. Mandarin also contains context-conditioned tonal phenomena, including tone sandhi² involving *yī* ‘one’ and *bù* ‘not’, as well as neutral-tone syllables (Liu, 2000; Shao, 2016; Yip, 2002). For example, *yī* may surface differently in forms such as *yíkuài* ‘one piece’ and *yìzhí* ‘continuously’, while *bù* may change in forms such as *búhuì* ‘not know how to’. Neutral-tone forms such as *māma* ‘mother’ and *dìdi* ‘younger brother’ also require reduced pitch prominence and shorter duration. These examples show that Mandarin tone acquisition involves not only the production of individual tones, but also the ability to adjust tones according to phonological context.

1.1.3 Tone in the Bantu Languages

Ugandan L2 speakers of Chinese³ approach Mandarin tone learning from multilingual linguistic backgrounds. Uganda is a multilingual country, the participants from Luganda, Lusoga, Lugisu, Runyankore, and Samia backgrounds speak Bantu languages. Existing descriptions show that tone in these languages may interact with register, grammatical structure, tone spreading, and morphotonological alternation rather than patterning as a Mandarin-like four-way lexical contour system (Hyman & Katamba,

² Tone Sandhi refers to context-conditioned tonal alternation across adjacent syllables or morphemes. In Mandarin, examples include *yī* changing in forms such as *yíkuài* ‘one piece’ and *yìzhí* ‘continuously’. This is a common feature of many tonal languages such as Mandarin.

³ Ugandan L2 Speakers (of Chinese) refers specifically to Ugandan learners who have acquired Mandarin as a second or additional language.

1993, 2010; Marlo, 2013; Hyman, 2018, 2022). This background may support sensitivity to pitch, but it does not by itself ensure target-like control of Mandarin contour direction, pitch range, or turning points.

1.1.4 Tone in the Central Sudanic Language

Lugbarati, often described as Lugbara in the linguistic literature, is the only Central Sudanic language background represented in the sample. Andersen (1986) reports dialectal variation in Lugbara tone, including three tone levels in the Terego variety and four levels in some western varieties. This level-tone experience may support lexical pitch awareness, but it differs from the dynamic contour contrasts required for Chinese Tone 2 and Tone 3.

1.1.5 L2 Chinese Tone Production and the Research Gap

Previous studies on second-language Chinese tone acquisition have repeatedly shown that tone learning is shaped by both phonetic properties and first-language backgrounds. Research on speakers from non-tonal and tonal language backgrounds has reported difficulty with contour tones, pitch direction, pitch height, tone-category formation and tone realization in connected speech (Wang, 2006; Chen, 2011; Zhang, 2020). Other studies have also shown that intonation and stress patterns from the first language may influence Chinese tone production, especially when speakers move from isolated syllables to connected speech (Lafontaine, 2017). These findings suggest that Chinese tone learning should not be treated as a purely articulatory problem. It also involves how speakers perceive, organize and apply pitch contrasts as linguistically meaningful categories.

However, the African and Ugandan contexts remain less fully documented in Chinese tone acquisition research. Existing studies on African learners of Chinese have identified recurring problems with T2, T3, disyllabic tone combinations and contextual tonal realization, but the available literature is still limited compared with research on Asian or European speaker populations (Gai, 2014; Zhao, 2020; Zhu, 2018; Zhang, 2020). For Uganda specifically, Namulondo (2014) provides an important early discussion of Chinese pronunciation errors among Ugandan speakers, including tone-

related difficulties, but further acoustic and context-sensitive analysis is still needed. There is therefore limited understanding of how Ugandan L2 speakers of Chinese, particularly those from Bantu and Central Sudanic language backgrounds, acquire and produce Chinese tones. This study seeks to help address that gap by examining Chinese tone production among Ugandan L2 speakers of Chinese within a perception-oriented framework.

1.2 Statement of the Problem

Although Chinese education in Uganda has expanded in recent years, the acquisition of Chinese tones by Ugandan L2 speakers remains insufficiently studied. In teaching practice, learners of Chinese in Uganda appear to experience recurring difficulty in producing Chinese tones, not only in isolated syllables but also in disyllabic words and tone-conditioned environments, e.g. *yī* changing in forms such as *yīkuài* ‘one piece’ and *yìzhí* ‘continuously’). Such difficulty can be observed in forms involving individual lexical tones, tone combinations, and contextual tone changes such as *yī* ‘one’, *bù* ‘not’, and neutral-tone syllables. However, classroom observation alone cannot fully explain whether these difficulties are occasional pronunciation mistakes or whether they represent systematic patterns in Ugandan learners’ Chinese tone production.

Existing research provides a useful starting point but remains limited for the present purpose. Namulondo (2014) examined errors in Chinese pronunciation among learners of Chinese in Uganda and discussed both segmental and tonal problems. However, the study did not provide a focused acoustic account of how Ugandan learners realize Chinese tonal contrasts across monosyllables, disyllabic sequences, and tone-conditioned environments. Nor did it systematically examine whether learners’ non-target tone productions show stable distributions across tone categories, phonological contexts, and first-language backgrounds.

This gap is important because Chinese tone production cannot be adequately understood only by listing general pronunciation errors. Accordingly, this study investigates how Ugandan L2 speakers of Chinese produce lexical tones and how their non-target outputs pattern across contexts. It identifies recurrent tone production errors

and describes their distribution across monosyllabic, disyllabic, and tone-conditioned environments.

1.3 Objectives of the Study

1.3.1 General Objective

The general objective of this research was to analyze tone production errors made by Ugandan L2 speakers of Chinese.

1.3.2 Specific Objectives

This study specifically seeks to:

- (i) Identify tone production errors made by Ugandan L2 speakers of Chinese.
- (ii) Describe the distribution of tone production errors among Ugandan L2 speakers of Chinese.

1.4 Significance of the Study

This study has both theoretical and applied significance for research on Chinese in under-documented learning contexts.

Theoretically, it extends empirical coverage of L2 Chinese tone acquisition to a Ugandan setting and places the analysis within a PPM-based perspective. Rather than treating tonal deviations as isolated mistakes, the study considers them as patterns that may reflect uneven perceptual access to tonal cues such as pitch height, directional movement, contour complexity, and syllable-internal turning points. It also brings the five-degree Standard tone system (see Figure 1 *Five-Degree Standard Tone System*) and the contrast between Chinese contour tones and more register-oriented tonal tendencies into a single interpretive frame that is relevant to Ugandan L2 speakers' (of Chinese) linguistic backgrounds.

Practically, the findings can support Chinese pronunciation work in Uganda by identifying which tones and contexts appear least stable and by clarifying why some tonal contrasts, especially those involving dynamic contour movement, remain difficult. Because the discussion is anchored in perceptual processing rather than in impressionistic classroom description, the study offers a more principled basis for prioritizing practice targets such as contour discrimination, pitch-range control, and

context-sensitive tonal realization.

1.5 Justification of the Study

The rapid growth of Chinese programs in Uganda creates a clear need for empirical, context-sensitive descriptions of tone production that go beyond general claims about tonal difficulty. Without such evidence, it is difficult to determine whether observed pronunciation challenges reflect momentary performance variation, stable group-level tendencies, or specific pressures associated with connected speech, tone sandhi, and neutral-tone environments. A focused investigation in the Ugandan context is therefore warranted in order to establish a reliable descriptive baseline and to evaluate whether a perception-oriented account can explain recurrent production problems.

A further motivation concerns cross-linguistic phonological diversity within Uganda. Many Ugandan languages make use of tone or pitch-based prominence, but they seem not pattern with Chinese in exactly the same way. Chinese relies heavily on lexical contour contrasts, whereas many Ugandan languages are more often described in terms of level or register-like tonal organization. This makes Uganda a particularly informative context for examining how prior tonal experience may support some aspects of learning while also creating perceptual bias in the processing of Chinese contour tones, especially in contrasts that require fine control of pitch movement.

1.6 Theoretical Framework

The Pitch Perception Model (PPM) is defined as a perception-centered framework that explains how the auditory system encodes, organizes, and interprets acoustic periodicity as pitch before such input is mapped onto linguistically meaningful categories (Licklider, 1951; de Cheveigné, 1986; Meddis & Hewitt, 1991; Yost, 2009). In this view, pitch perception is not treated as a passive reflection of surface acoustics, but as an active process in which temporal regularities in the speech signal are detected, compared, and stabilized in the listener's perceptual system. On this basis, the present study adopts PPM as the sole theoretical framework for explaining Chinese tone production errors among Ugandan L2 speakers of Chinese, on the assumption that accurate tone production depends fundamentally on how learners perceive, categorize,

and retrieve pitch contrasts rather than on articulation alone.

More specifically, PPM is adopted here on the assumption that tone production is shaped by prior perceptual organization of pitch rather than by articulation alone. The present dissertation adopts PPM as its sole theoretical orientation for explaining tone acquisition, recurrent tone production errors, and the perceptual factors that constrain accurate tonal realization. In this sense, PPM highlights the cognitive and perceptual foundations of tone learning by proposing that target-like production depends on the learner's ability to reorganize pitch information into linguistically meaningful tonal categories. In the case of Chinese, this involves remapping pre-existing first-language pitch routines to the localized lexical tone contrasts required by Chinese (Yip, 2002; Lin, 1996). Previous work further suggests that tone perception depends on sensitivity to rapidly unfolding pitch movement within a limited temporal window (Gandour, 1983). When such temporal and category-level perception remains unstable, acoustically similar contrasts may be insufficiently differentiated, which in turn increases the likelihood of confusion between tones such as Chinese T2 and T3. On this view, perceptual reorganization precedes articulatory accuracy: learners must establish new perceptual boundaries before they can reproduce tonal contours faithfully in production.

In Yip's (2002) discussion of tone perception and acquisition, adult non-native speakers are shown to perform substantially less accurately than native speakers in tone identification, while listeners from non-tonal backgrounds also display greater variability in locating tonal category crossover points. At the same time, the fact that training can markedly improve performance suggests that the central difficulty lies not in hearing pitch in a purely auditory sense, but in establishing stable linguistic categories for tonal contrast (Yip, 2002). This point is directly relevant to the present study: if learners do not reliably perceive and organize Chinese tonal cues in terms of relative height, contour direction, and dynamic pitch movement, later production is likely to remain vulnerable to confusion, compression, and inaccurate retrieval.

For the purposes of this dissertation, the PPM is adopted as the sole theoretical framework for explaining how Ugandan learners acquire, confuse, and produce Chinese

lexical tones. In the present study, PPM is used as a perception-centered account of tone learning in which accurate production depends on how pitch contrasts are heard, categorized, organized, and retrieved in speech rather than on articulation alone. Chinese tone acquisition is therefore treated as a process of perceptual remapping: learners must move from first-language pitch routines to the lexical pitch contrasts required by Chinese.

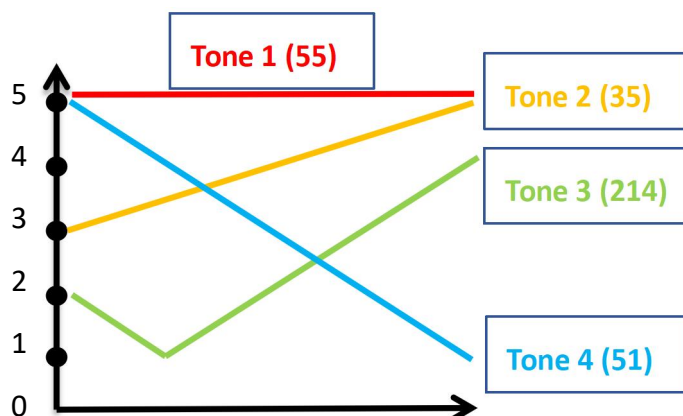
Within this study, PPM is not treated as a loose background idea about hearing, but as the main explanatory frame through which recurrent tonal deviations are interpreted. This position is important because the study is production-centered, yet production is not assumed to be self-explanatory. A non-target tone is read here as evidence that the underlying tonal category may still be weakly represented, insufficiently separated from a neighboring category, or inconsistently accessed under real-time speech conditions.

The first core claim of PPM is that stable production presupposes stable perceptual representation. Learners do not succeed merely by imitating surface melody; they must identify which aspects of pitch are linguistically contrastive, assign those cues to phonological categories, and retrieve them rapidly during speech. The second claim is that cue weighting is selective rather than uniform. Learners do not rely on every acoustic property equally, and persistent errors often arise when they overweight a global impression of pitch movement while underweighting finer cues such as contour direction, low turning region, rise onset, pitch range, or alignment. The third claim is that tonal categories are especially vulnerable when processing pressure increases. As Ye and Connine (1999), Tong et al. (2008), and Gao et al. (2019) show in different ways, lexical tone is not always accessed as easily or as early as segmental information, particularly in low-context conditions. From a PPM perspective, tonal difficulty is therefore shaped jointly by cue salience and processing conditions. A learner may know that a tone should rise or dip in principle, yet still fail to realize it accurately when the syllable is short, when another syllable follows, or when sandhi and neutral-tone adjustments must also be computed.

A perception-oriented framework is particularly suitable for Chinese because Chinese lexical tone is a syllable-level contrast whose acoustic realization depends on multiple

interacting cues. In the traditional five-degree standard tone system, Tone 1 is high level (55), Tone 2 is rising (35), Tone 3 is dipping or fall-rise (214), and Tone 4 is falling (51) (Chao, 1968). In the present study, these values are not treated merely as pedagogical labels, but as an operational framework for describing the relative height, starting point, endpoint, and directional movement of Chinese tonal categories. As Yip (2002) makes clear in her broader discussion of tonal systems, lexical tone is not reducible to static pitch labels alone; rather, tonal contrast depends on how pitch is organized and interpreted as a linguistically relevant pattern. For this reason, accurate tone production requires learners not only to hear that pitch changes, but to perceive and reconstruct the specific contour relations that distinguish one lexical category from another.

Figure 1: Five-Degree Standard Tone System



These labels in Figure 1, however, are only a shorthand description. Tone identity depends first on F_0 height and register, because tones must occupy distinguishable regions of a speaker's pitch space. It also depends on contour shape, including whether the pitch remains level, rises, dips, or falls. For the dynamic tones, pitch range and duration are crucial, since a rise or dip cannot be fully realized without sufficient melodic space and temporal support. In addition, the timing of movement matters. A rising contour that begins too late, or a dipping contour whose low turning region is too shallow or poorly aligned, may drift toward another category even when the intended lexical target is correct. For this reason, the present study treats F_0 , pitch range, duration, turning-point realization, and alignment as interrelated indicators of tonal organization rather than as isolated phonetic variables.

Several strands of previous research support the present PPM-based interpretation. Gandour (1983) showed that speakers from different language backgrounds do not perceive pitch patterns in the same way, especially when contour and level distinctions are involved. Ortega-Llebaria et al. (2015) further demonstrated that long-term experience with a tonal language reshapes later F_0 processing even in a non-tonal language, suggesting that earlier pitch experience leaves durable perceptual routines rather than disappearing once learners enter a new phonological system. Strange (2011), from the perspective of automatic selective perception, likewise argues that second-language listeners often continue to rely on first-language perceptual routines and therefore attend to the wrong or less reliable acoustic information in the new language. Although these studies are not identical in scope, they converge on one point central to PPM: tone learning depends on how pitch is perceptually organized, not merely on how it is physically produced.

Additional support comes from research on Chinese tone perception and L2 tone acquisition. Chen et al. (2017) report that Chinese listeners show stronger categorical perception of pitch direction than English listeners and that stimulus duration substantially affects contour perception, especially for tonal listeners. This finding is highly relevant because it links contour perception directly to temporal support. Yip's (2002) review of second-language tone acquisition points in a similar direction: non-native speakers are disadvantaged in tone identification, but that disadvantage is tied to unstable category perception rather than to an absolute inability to process pitch. Evidence from first-language acquisition in Chinese further supports this perception-centered account. Reviewing Li and Thompson (1977) and Clumeck (1980), Yip (2002) notes that Chinese-speaking children acquire the high level and high falling tones earlier, whereas rising and dipping tones emerge later and are frequently avoided, leveled, or confused, especially in utterance-final position where both retain a rising component. Yip also cites Yue-Hashimoto (1980), whose Chinese-speaking child manipulated tonal melodies independently of segmental material in early tone play, suggesting that tonal categories can already be represented as separable cognitive units. These observations are important here because they show that successful tone production presupposes an

underlying organization of tonal categories, and that recurrent confusion between rising-type tones is not arbitrary but reflects a deeper representational challenge. Empirical studies of Chinese tone learning also show that recurrent confusions are structured rather than random. Huang and Liu (2023), for example, show that Vietnamese learners display systematic error patterns in Chinese T1 and T4 across perception, pronunciation, and memory, with different error rates in monosyllabic and disyllabic positions. Chen (2011), working with learners from non-tonal first-language backgrounds, finds that tone-category development changes with proficiency and that *Yangping* and *Shangsheng* are repeatedly drawn together because learners group tones according to partially shared phonetic properties. Zhang (2020), in an experimental study of Congolese French speakers, identifies T2 and T3 as the greatest area of difficulty and reports that tone pattern is often better preserved than tone value, while connected speech further destabilizes performance. Lafontaine (2017) similarly argues that French stress and sentence intonation can negatively transfer to Chinese tone accuracy, especially once learners move beyond isolated syllables into connected speech and tonal alternation. Taken together, these studies support the use of a single perception-centered framework for interpreting recurrent tonal errors in production data. PPM is especially well suited to the present dissertation because the study is centered on production data but seeks to explain production through perceptual organization. Under this framework, recurrent acoustic deviations are treated as traces of how tonal categories are perceived, weighted, and accessed. This is particularly useful for explaining why the T2 and T3 contrast emerges as a major difficulty in the Ugandan data. Both categories involve dynamic pitch movement, but they require different cue constellations: T2 requires a sufficiently early and adequately scaled rise, whereas T3 requires attainment of a low turning region together with a more complex internal contour. If speakers detect only the broad fact of movement while failing to stabilize cue timing, low-point realization, or range, the two categories can converge into the same perceptual and acoustic corridor. The same logic also helps explain why error distributions differ across monosyllables, disyllables, sandhi environments, and neutral-tone contexts. Monosyllables provide clearer tonal space, whereas disyllables

impose greater timing pressure; sandhi requires contextual recomputation of tonal targets; and neutral tone requires reduction of prominence rather than citation-form preservation. In all these cases, the relevant question is not simply whether the learner knows the dictionary tone, but whether the learner can retrieve and implement the appropriate pitch pattern under changing phonetic and prosodic conditions.

The acoustic measures used in later chapters are interpreted within one unified PPM frame. F_0 contour is used to show the overall melodic pattern of a tone; pitch range indicates whether the learner establishes sufficient contrastive space; duration reveals whether the syllable provides adequate temporal support for contour realization; turning point and alignment indicate whether internal tonal movement is launched and coordinated at the appropriate moment; contour complexity helps explain why some categories are more vulnerable to reduction or simplification than others; and contextual pressure accounts for the extra instability found in disyllabic sequences, sandhi, and neutral-tone environments. In practical terms, this framework makes it possible to distinguish category confusion from phonetic distortion, that is, to separate instability in tonal categorization from difficulty in phonetic implementation. Recent work further suggests that deviations in L2 tonal performance may arise either from difficulty in establishing tone categories or from later difficulty in phonetic implementation, while actual error patterns also vary with tonal sequence, prosodic position, and speech context (Zhou and Veríssimo, 2025; Yang, 2016). In this way, the framework helps explain why tonal errors are unevenly distributed across tonal categories and phonological environments, and how cue salience, contour complexity, pitch-register control, temporal support, and possible first-language-related perceptual bias interact in the Ugandan data. PPM therefore does more than provide background theory: it organizes error identification and distribution while supporting cautious interpretation within a single, coherent explanatory frame.

It is at this point that theoretical comparison becomes necessary. A number of frameworks have been used in previous studies of tone acquisition and second-language speech. Interlanguage Theory helps explain learner speech as a developing system rather than as a simple collection of random mistakes. Speech Learning Model (SLM)

offers useful insight into how learners relate unfamiliar or partially similar L2 sounds to existing L1 categories. Working Memory and Automatization perspectives are relevant when repeated production practice, processing load, or limited control of complex forms are discussed. Optimality Theory (OT) has also been used in some studies to model error patterns in terms of competing constraints (Markedness constraint and faithfulness constraint) and output deviation. However, these frameworks do not contribute equally to the present study, and OT is not retained here as the primary explanatory framework.

For the purposes of the present study, a framework is needed that can connect the observable distribution of production errors with the perceptual and prosodic properties of the tones themselves. This is where Pitch Perception Model (PPM) becomes especially useful. The literature reviewed above repeatedly points to issues of tone salience, contour complexity, pitch direction, temporal support, perceptual discrimination difficulty, and first-language-related perceptual bias. These dimensions make it possible to move from a simple description of error to a more coherent explanation of why certain Chinese tones, or certain tonal sequences, are more difficult for learners/speakers of Chinese in Uganda to produce accurately. PPM therefore emerges from the review as the most suitable primary framework for linking this chapter to the methodological procedures in Chapter Three and the analysis developed in Chapters Four and Five.

1.7 Scope of the Study

1.7.1 Geographical Scope

Geographically, the study is limited to Luyanzi Institute of Technology in Uganda, where the 19 participants were drawn from one intact intermediate Chinese class.

1.7.2 Content Scope

The content scope corresponds to the two objectives: identifying tone production errors and describing their distribution among Ugandan L2 speakers of Chinese. It covers the four lexical tones across monosyllabic items, disyllabic combinations, and selected tone-conditioning environments involving T3 sandhi, *yī* ‘one’, *bù* ‘not’, and neutral

tone. The participants' Luganda, Lusoga, Lugisu, Runyankore, Samia, and Lugbarati backgrounds are included as descriptive sociolinguistic characteristics of the sample, not as balanced comparison groups.

1.7.3 Theoretical Scope

The theoretical scope is limited to the Pitch Perception Model (PPM), which guides cautious interpretation of the production patterns. The study does not include a separate perception experiment or an instructional intervention.

1.7.4 Temporal Scope

The temporal scope is limited to the single cross-sectional recording period represented by the dataset. It describes participant performance at the time of recording and does not track longitudinal development.

1.7.5 Linguistic Scope

The linguistic scope is limited to Ugandan L2 speakers of Chinese whose first language (L1) is one of the following Ugandan languages: Luganda, Lugisu, Lusoga, Runyankore, Samia, and Lugbarati, rather than other Ugandan language backgrounds, such as Acholi and Langi. This scope reflects the natural composition of the participants: the 19 participants were recruited from a single intact Chinese class, and their L1 backgrounds are limited within the 6 languages listed above.

1.8 Summary

This chapter has introduced the study by outlining its context, research problem, objectives, significance, justification, theoretical orientation, and scope. It has also situated the study within a PPM-based account in which recurrent production errors are interpreted in relation to tonal perception, pitch-register processing, contour complexity, and contextual pressure. The next chapters develop the necessary background (Chapter Two), methods and data (Chapter Three), and the presentation and discussion of the production results in Chapters Four and Five.

CHAPTER TWO: LITERATURE REVIEW

2.1 Introduction

The literature reviewed in this study is organized across three major dimensions. First, since the primary aim of this research is to analyze Chinese tone production errors in Chinese as a second language, studies on Chinese language provides an essential starting point for the review. These studies offer systematic descriptions of Chinese tone categories, tone perception and production, and pedagogical issues within the field of Teaching Chinese as a Second Language (TCSL). They are therefore central to understanding what has already been established about the linguistic properties of tones and the recurrent problems reported in learner speech.

Second, because the present data are drawn from Ugandan speakers and because Chinese is taught in Uganda largely through English-medium educational settings, scholarship on English language also forms an important part of the review. This body of literature helps situate Chinese tone learning within broader discussions of second-language phonetic acquisition, cross-linguistic transfer, and the relationship between perception and production. It is particularly useful for evaluating how speakers from different phonological backgrounds process pitch contrasts and why some tonal categories are consistently more difficult than others.

Third, Uganda must be understood within a wider African context. African-focused studies on Chinese language learning reveal both common tonal challenges and important regional differences, while research on Luganda and related languages provides necessary background for evaluating possible first-language influence.

Taken together, these three strands of literature allow the chapter to move from a general account of Chinese tones, to findings on foreign speakers' tone production, to the more limited evidence available for African and specifically Ugandan speakers. The purpose of this chapter is not merely to list previous studies, but to clarify what they collectively show, where they remain incomplete, and why the present study focuses specifically on error types and their distribution in Ugandan speakers' Chinese tone production. In this way, the chapter establishes the literature-based rationale for Chapter Three, the

organization of findings in Chapter Four, and the PPM-oriented interpretation developed later in the thesis.

2.2 Tone in Chinese

Celce-Murcia et al. (2010) explained that pronunciation is a critical component of second-language instruction because of its connections to phonology, morphology, syntax, semantics, and communicative competence. They further emphasized that inaccurate pronunciation may confuse lexical meaning, a point that is especially significant in tonal languages. In Chinese, pitch pattern is not a peripheral feature but part of the lexical identity of the syllable. For this reason, any study of speaker pronunciation that aims to identify tonal errors must begin with a clear understanding of how the Chinese tone system is structured.

Chomsky and Halle (1968) treated pitch as one of the relevant distinguishing features in phonological representation, while Wang (1967) proposed a tone-feature system derived from tonal phenomena in Chinese. Taken together, these studies suggest that tone should be treated as a systematic phonological category rather than as a purely phonetic ornament.

Pike (1948) distinguished Chinese from many African languages by classifying Chinese as a contour-tone language and many African languages as register-tone languages. This distinction is important for the present study because the existence of tone in a speaker's first language does not automatically guarantee ease in acquiring Chinese tones. What matters is not simply whether the first language is tonal, but also how pitch is organized, how contrasts are realized, and how tonal categories are distributed across the syllable. Liu (2000) observed that Chinese has approximately 400 basic syllables, each typically consisting of an initial and a final, and that these syllables expand substantially when combined with the four lexical tones. Thus, a basic syllable such as *fāng* 'square' can occur with Tone 1 (T1), Tone 2 (T2), Tone 3 (T3), or Tone 4 (T4), producing distinct lexical items. Shao (2016) further emphasized that Chinese is a tonal language in which every syllable carries a specific tonal pattern, including neutral-tone realizations whose surface form may vary with grammatical or phonological context. Examples involving

yi ‘one’ illustrate that tonal realization in Chinese is not static but conditioned by structural environments, as in *yìxiē* ‘some’, *yíhuìr* ‘a while’, and *xiǎng yì xiǎng* ‘to think, but last shortly’.

This foundational knowledge is essential for any analysis of speaker tone production. It shows why tonal errors cannot be reduced to vague problems of accent and why they must instead be examined in terms of contour shape, pitch direction, tonal category, and contextual realization. It also helps explain why later chapters of the present study distinguish among individual tones, tone combinations, neutral-tone sequences, and sandhi-related forms rather than treating all tonal deviation as a single undifferentiated phenomenon.

2.3 Non-native Speakers’ Production of Chinese Tones

Research on Non-native speakers production of Chinese tones has consistently shown that tonal difficulty is not distributed evenly across tone categories. Rather, some tones and some tonal combinations are repeatedly reported as more vulnerable to error than others, and such patterns are often linked to speakers’ first-language phonological background, to the phonetic properties of Chinese tones, and to the relationship between tonal perception and tonal production. This line of work provides the broader comparative context for the present study.

The challenges faced by Asian language speakers refer mainly to difficulties in Chinese tone acquisition reported among Korean, Thai and other Asian-background L2 speakers of Chinese. These challenges include difficulty in producing individual tone categories, especially T2 and T3; instability in pitch height and pitch range; confusion between rising and dipping contours; reduced accuracy in disyllabic tone combinations; and difficulty with tone-conditioned forms such as *yi* ‘one’, *bu* ‘not’ and neutral-tone syllables. The purpose of reviewing these studies is not to treat Asian speakers as identical to Ugandan L2 speakers of Chinese, but to identify common patterns in L2 Chinese tone acquisition that can serve as a comparative background for the present study.

Ni and Wang (1992) concluded that difficulties encountered by speakers from English-

speaking countries were strongly affected by first-language interference. Using a questionnaire that included both sound combinations and individual tones, they found that participants often struggled with segmental combinations as well as with tone production. Tone 3 was identified as especially difficult, and many speakers showed problems in allocating pitch movement and syllable duration appropriately. Although the study was not conducted in Uganda, its findings remain relevant because English-medium educational experience is one part of the linguistic environment in which many Ugandan L2 speakers of Chinese study Chinese.

Jin (2011) conducted a questionnaire-based investigation into pronunciation difficulties encountered by Korean speakers of Chinese, with a focus on tones. The study found that tone 3 posed the greatest challenge, followed by T2, and Li (2020) reported a similar pattern. Song (1998) also pointed out that tones were significantly more difficult for Korean speakers than initials and finals because Korean lacks lexical tone as a contrastive phonological category. In particular, the rising movement of T2 and the low dipping contour of tone 3 were difficult for speakers to produce in a stable way.

Iuliia (2020) found that tone problems existed across proficiency levels among Russian speakers of Chinese. One prominent problem was insufficient pitch range, for example when T2 was realized as a relatively low rise. Zhang (2011) similarly observed that T3 and T4 were especially problematic for Russian speakers and further noted difficulties with neutral tone, retroflex sounds, and continuous sandhi. Wang and Sun (2003) explained some of these difficulties by contrasting the role of stress in Russian with the role of lexical tone in Chinese.

Darya (2020) provided a more detailed account by identifying persistent problems in monosyllabic T2, in disyllabic combinations such as T2 + T1 in *shíjiān* ‘time’ and T1 + T2 in *gōngyuán* ‘park’, in T3 when paired with other tones, and in neutral-tone and sandhi contexts involving forms such as *shénme* ‘what’, *yìxiē* ‘some’, and *búshì* ‘not’. These studies collectively suggest that speakers often struggle not only with isolated tonal categories but also with contextual tonal realization.

The Thai case has attracted substantial attention because Thai, like Chinese, is a tonal language, yet Thai speakers still produce Chinese with noticeable tonal deviation if

sufficient practice and perceptual adjustment are lacking. Li (1995) analyzed speakers' abilities through a comparison of the Chinese and Thai tone systems and found serious problems particularly with T1 and T4, mainly in terms of pitch deviation, a point later echoed by Feng (2016). Zhang and Peng (2021) compared the two systems in detail and argued that confusion arose not because tone was entirely unfamiliar, but because the specific phonetic implementation of tonal categories differed across the two languages. Further studies by Zhang (2008), Xu (2013), and Feng (2016) on disyllabic words showed that Thai speakers were especially likely to confuse T2 and the rising portion of T3. Diao (2020) also found that tone sandhi, especially with *yi* 'one' and *bu* 'not', remained challenging even in instructional settings in Bangkok. Examples discussed in this line of research include disyllabic combinations such as T2 + T1 and T3 + T2, as well as sandhi-sensitive forms such as *yìxiē* 'some' and *búshì* 'not'.

Taken together, research on Asian speakers demonstrates two important points. First, tone-related difficulty often clusters around T2 and T3, rising movement, and context-sensitive tonal realization. Second, tonal experience in the first language may help in some respects but does not eliminate difficulty, because speakers must still adapt to the perceptual salience, contour shape, and temporal realization of Chinese tone categories. It is also important to note that most research on Chinese tones has been concentrated in Asian contexts, where Chinese language teaching has a longer history and where more detailed phonetic work has been conducted.

The Asian-language literature shows that Chinese tone errors are often related to pitch movement, contour timing, tone-category separation and contextual tonal adjustment. These issues are also relevant when examining Ugandan L2 speakers of Chinese, although the first-language backgrounds of Ugandan L2 speakers of Chinese are different from those of Korean, Thai or other Asian-background speakers. African language backgrounds involve different tone systems, different uses of pitch, and different relationships between lexical and grammatical tone. It is therefore necessary to consider how Chinese tone acquisition has been studied in African contexts.

African language speakers refer to L2 speakers of Chinese from African linguistic backgrounds that have been discussed in previous Chinese tone acquisition studies. The

groups reviewed here include Hausa-speaking, Malagasy-speaking, Egyptian, Botswana and Nigerian speakers of Chinese, as well as broader groups of African international students. These groups are relevant because many African languages use tone, but their tone systems differ from Chinese in tonal inventory, contour structure, pitch range and functional load. Reviewing these groups helps show why tone experience in an African language does not automatically lead to accurate Chinese tone production. Most African languages make use of tone, but the functions of tone are not always the same as those found in Chinese. In Chinese, tones are primarily lexical, whereas in many African languages, tones may mark both lexical distinctions and grammatical functions. This difference means that African speakers' prior experience with tone cannot be assumed to transfer directly into Chinese production. It may instead create a more complex relationship in which some tonal sensitivity is available, but the target contrasts are reorganized under a different phonological system.

Li (1988) compared Hausa and Chinese tones and noted that Hausa tones served not only lexical but also grammatical purposes. Unlike Chinese, Hausa was described as having three main tonal categories: high, low, and dipping. Gai (2014) investigated African international students' production of monosyllabic words and identified particular difficulty with T2 and T3. Sylviano (2016) also found that Malagasy students experienced significant problems with T3 and T2, especially in tonal combinations where the first portion of T3 proved unstable.

Zhao (2020) argued that Egyptian learners' local accent in Chinese pronunciation was largely associated with tonal errors. Through acoustic testing of monosyllabic and disyllabic words, Zhao found that T2 was the most difficult monosyllabic tone, followed by T1, while disyllabic problems were especially evident in combinations such as T3 + T2. Zhu (2018), working on Botswana speakers, similarly identified difficulties in monosyllabic T2 and T4 and in disyllabic sequences involving T3 at the beginning and T2 at the end, as well as T2 followed by T4.

Zhan (2017) found that Nigerian speakers of Chinese primarily struggled with disyllabic tone combinations, while Zhou (2013), Zhu (2018), and Zhao (2020) all employed experimental-phonetic methods to model learners' tonal deviation in African

settings. Even though the countries and language backgrounds differed, the existing African literature converges on several points: research remains relatively limited, studies are often country-specific, and tonal problems repeatedly centre on T2 and T3, on rising or dipping movement, and on disyllabic combinations rather than only isolated tones.

Compared with Asian scholarship, African research on Chinese tones remains narrower in scope. Asian studies have examined a wider range of related issues, including intonation, retroflexion⁴, and tone sandhi, whereas African studies have more often concentrated on monosyllabic and disyllabic tone production. This imbalance does not reduce the value of African research; rather, it highlights the need for more detailed, context-sensitive work that can connect observable tonal errors with broader explanatory factors. That need is particularly clear in the Ugandan context, where the multilingual environment differs from those reported in the better-documented Asian settings. The review of African-background speakers therefore narrows the literature gap toward the present study. Existing African studies show that tone-related problems often involve T2, T3, disyllabic combinations and pitch-contour realization, but Ugandan L2 speakers of Chinese remain insufficiently represented. This provides the basis for examining whether similar or different tone-error patterns appear among Ugandan L2 speakers of Chinese, especially across monosyllabic, disyllabic and tone-conditioned environments.

2.4 Production of Chinese Tones by Its Learners in Uganda

The Ugandan context is especially important for the present study because it brings together at least three relevant linguistic dimensions: local African tonal experience, English-mediated formal education, and the acquisition of Chinese as an additional language. Existing literature on production of Chinese tones by learners of Chinese in Uganda is limited, but the studies that do exist already indicate that the Ugandan case cannot be fully inferred from findings on either Asian learners or other African learner

⁴ In this work, forms such as *huār* illustrate erhua or rhotacization, that is, the addition of a rhotic element /-r/ to the end of a syllable, which may affect pronunciation and sometimes grammatical or lexical interpretation.

groups. This section therefore reviews the available Ugandan evidence as the most direct point of connection to the present research.

Before discussing previous studies on learners of Chinese in Uganda, it is necessary to consider the tonal background of the Ugandan languages represented in this study. The participants come from different first-language backgrounds, including Luganda, Lusoga, Lugisu, Runyankore, Samia and Lugbarati. The first five languages belong to the Bantu language family, while Lugbarati is commonly classified as Central Sudanic. Since Chinese is also a tonal language, the learners' first-language tonal experience may influence how they perceive and produce Chinese tones. However, tonal experience in a Ugandan language cannot be assumed to transfer directly into Chinese, because Ugandan languages and Chinese organize tone in different ways.

Luganda is one of the most extensively studied Ugandan languages in relation to tone. Hyman and Katamba (1993) proposed an influential analysis of Luganda tone, showing that Luganda tone should not be understood simply as surface high and low pitch, but as a system involving underlying tonal contrasts, tone-bearing units and tonal processes. Hyman and Katamba (2010) further examined the relationship between tone, syntax and prosodic domains in Luganda, showing that Luganda tonal patterns are shaped by prosodic grouping and syntactic structure. Their studies indicate that Luganda-speaking learners may already have experience with linguistic uses of pitch, but this experience is organized differently from the four lexical contour tones of Chinese.

Most of these languages belong to the Bantu language family, while Lugbarati is commonly classified as Central Sudanic. Since Chinese is also a tonal language, the learners' first-language tonal experience may influence how they perceive and produce Chinese tones. However, tonal experience in a Ugandan language cannot be assumed to transfer directly into Chinese, because Ugandan languages and Chinese organize tone in different ways.

Lugbarati, often referred to as Lugbara in the linguistic literature, deserves separate mention because it is the only non-Bantu first-language background represented in the present sample. The tonal organization of Lugbarati differs from the Bantu languages discussed above. Andersen (1986) shows that Lugbarati has dialectal variation in tone:

the Terego dialect of Ugandan Lugbarati has three tone levels, while some western varieties have been described as having four tone levels. This means that Lugbarati-speaking Ugandan L2 speakers of Chinese may have prior experience with lexical pitch distinctions, but that experience is organized primarily through level-tone contrasts rather than through Chinese-like four-way lexical contour tones.

This distinction is important for the present study. Chinese T2 and T3 require control of dynamic contour movement, including rise timing, low turning-point realization and recovery movement. By contrast, a level-tone system such as that described for Ugandan Lugbarati places greater emphasis on pitch level or register distinctions. Therefore, Lugbarati-speaking learners of Chinese in Uganda may show general sensitivity to pitch contrast, but may still experience difficulty with Chinese tones that require dynamic contour implementation. This is especially relevant when interpreting errors such as T2 and T3 confusion, delayed rise onset, dipping-contour truncation and possible tone-identification instability.

Other Bantu languages represented in this study, including Lusoga, Lugisu, Samia and Runyankore, have also been discussed in studies of Bantu tone. Hyman (2018), for example, used Lusoga to illustrate the autosegmental approach to tone and showed that Bantu tonal systems often involve abstract tonal representations rather than simple surface pitch patterns. Marlo (2013) examined verb tone across Bantu languages and emphasized that Bantu tone is frequently connected with morphology, verb structure, tense-aspect categories and other grammatical environments. Hyman (2022) also discussed tone in Runyankore verb-stem reduplication, showing that Runyankore tone interacts with morphological structure and stem shape. These studies suggest that speakers of Ugandan Bantu languages may have general sensitivity to pitch, but their first-language tonal experience may involve register tone, grammatical tone, tone spreading, downstep or morphotonological alternation rather than Chinese-like four-way lexical contour-tone contrasts. Therefore, such learners may still experience difficulty mapping Chinese T1, T2, T3 and T4 onto their existing tonal experience.

The inclusion of Lugbarati therefore widens the linguistic background of the present study beyond Bantu tone systems. While the Bantu languages represented in the sample

are relevant for understanding register tone, grammatical tone, tone spreading and morphotonological alternation, Lugbarati adds a Central Sudanic comparison point in which tonal contrast is also present but organized through a different historical and typological system. For this reason, the study does not treat Ugandan tonal experience as a single homogeneous background. Instead, it treats L1 background as an exploratory factor that may shape how Ugandan L2 speakers of Chinese perceive and implement Chinese pitch contrasts.

Namulondo (2014) examined similarities and differences between Luganda and Chinese and identified a range of pronunciation problems among learners of Chinese in Uganda, including both segmental and tonal difficulties. The study reported that learners' of Chinese pronunciation was shaped by their phonetic system, language environment, and individual factors. Segmental difficulties were reported for the alveolo-palatal initials *j*, *q*, and *x*, the dental affricate *c*, the retroflex initials⁵ *zh*, *ch*, *sh*, and *r*, the fricative *h*, and several finals⁶, such as *e*, *o*, *ü*, *er*, *uo*, *ou*, *ia*, *ie*, and *iu*. More importantly for the present study, Namulondo also highlighted tone production errors, especially those involving tones 2 and 3, as well as sandhi problems with *yi* and *bu*. Earlier classroom observation reported similar pressure points⁷ in forms such as *shíjiān* 'time', *xǐnglái* 'wake up', *chúle* 'except', *yìxiē* 'some', and *búshì* 'not'; however, these observations are better treated as motivation for systematic analysis than as evidence in themselves.

⁵ Retroflex initials here refer to the Mandarin consonants *zh*, *ch*, *sh*, and *r*, which are produced with the tongue tip curled or raised backward toward the area behind the alveolar ridge / toward the hard palate. See Fu, Chengbo, editor. 普通话教程 [Putonghua Jiaocheng; A Course in Putonghua]. Beijing University of Posts and Telecommunications Press, Nov. 2011.

⁶ Finals refer to the second part of a Mandarin syllable (i.e., the *yunmu*), usually consisting of one or more vowel elements, sometimes followed by *-n*, *-ng*, or *-r*; a final combines with an initial to form a complete syllable. See Fu, Chengbo, editor. 普通话教程 [Putonghua Jiaocheng; A Course in Putonghua]. Beijing University of Posts and Telecommunications Press, Nov. 2011.

⁷ Pressure points is used here in a figurative sense to mean recurring problem areas or especially vulnerable items in learners' pronunciation, rather than anatomical pressure points on the body. See Fu, Chengbo, editor. 普通话教程 [Putonghua Jiaocheng; A Course in Putonghua]. Beijing University of Posts and Telecommunications Press, Nov. 2011.

This tonal background is relevant to the present study because it provides a basis for interpreting possible first-language influence in Chapter Four. If learners of Chinese in Uganda show recurring problems such as T2 and T3 confusion, pitch-range instability, delayed rising movement, contour simplification, tone-identification errors, or difficulty with tone sandhi, these patterns should not be treated only as isolated pronunciation mistakes. They may also be discussed in relation to the participants' prior experience with Ugandan tonal systems, where pitch may be organized through register tone, grammatical tone, tone spreading, morphotonological alternation, or level-tone contrasts rather than through Chinese-like four-way lexical contour-tone categories. For this reason, reviewing the tone systems of Luganda, Lusoga, Lugisu, Samia, Runyankore and Lugbarati helps contextualize the later description of Chinese tone-error patterns among learners of Chinese in Uganda. In Chapter Four, this section provides the linguistic basis for explaining why prior tonal experience may support general pitch awareness but may not be sufficient for accurate Chinese contour-tone production or tone identification.

2.5 Critical Evaluation

The preceding review shows that research on Chinese tone learning has generated several recurring findings. Across learner populations, tonal difficulty is rarely random. Certain categories, especially T2 and T3 and combinations involving rising or dipping movement, are repeatedly identified as problematic. These findings are directly relevant to the two objectives of the present study: identifying tone production errors among Ugandan L2 speakers of Chinese and describing how those errors are distributed across their spoken output.

At the same time, the literature also reveals a gap. Much of the existing scholarship either concentrates on Asian learners or provides only relatively narrow country-based descriptions in African settings. Even where African learners are studied, the discussion often remains at the level of reporting which tones are difficult, without fully integrating error types and their distribution into one coherent descriptive account. For Uganda in particular, the literature remains too limited to show clearly how Chinese tone

production errors are patterned across speakers' spoken output, how those patterns vary across first-language backgrounds within Uganda, or how they relate to the interaction between local tonal experience, English-medium learning, and the perceptual demands of Chinese tone categories.

CHAPTER THREE: METHODOLOGY

3.1 Research Design

This study adopted an integrated descriptive-analytical design centered on controlled Mandarin Chinese tone production data from Ugandan participants. The descriptive component documents recurrent tonal patterns across monosyllabic items, disyllabic combinations, and tone-conditioning environments. The analytical component examines the acoustic evidence and organizes the documented patterns in relation to the two objectives. The two components therefore operate together: description establishes what errors occur and how they are distributed, while analysis provides the acoustic basis for that description without introducing a separate causal objective.

This combined design is appropriate because the study does not test an intervention or establish causal relationships through experimental manipulation. It first identifies recurrent non-target productions and then analyzes their distribution across tone categories, first-language backgrounds, and phonological environments. Acoustic measurements support both components by making the observed contour, range, timing, and duration patterns explicit.

3.2 Research Approach

This study adopts a primarily qualitative and descriptive research approach, supported by acoustic measurement and limited quantitative verification. Its central purpose is to identify recurrent tone-production errors and describe their distribution among Ugandan L2 speakers of Chinese. Quantitative acoustic evidence is used to support these observations rather than to build a predictive model or establish causal relationships.

A qualitative research approach was adopted because the study is primarily concerned with the nature, patterning, and interpretation of tone production errors rather than with making broad statistical generalizations. Although acoustic measurements such as F_0 range, pitch contour and duration are used to support the analysis, these measurements serve mainly as quantitative evidence for identifying and describing error patterns. The

focus of the study is therefore on interpreting how Ugandan L2 speakers of Chinese realize Mandarin Chinese tones in specific phonological contexts, including monosyllabic words, disyllabic sequences, tone sandhi and neutral-tone environments. At the same time, the study incorporates instrumental acoustic analysis and limited corroborative statistical procedures. These quantitative elements are not treated as the main methodological orientation of the study; rather, they serve a supportive role in strengthening the descriptive interpretation of the production data. Acoustic measurements make it possible to observe how tonal categories are realized in terms of pitch movement, pitch range, contour shape, and duration, while limited statistical checks help determine whether the main descriptive patterns remain sufficiently stable across speakers and items. The overall methodological emphasis, however, remains on explaining how these observed production patterns relate to the study objectives.

This approach is appropriate for the present research for three reasons. First, the study seeks to document and interpret actual learner output rather than to test a narrowly defined experimental hypothesis. Second, the research questions require close attention to how tonal errors are distributed across tone categories, prosodic contexts, and first-language backgrounds. Third, the approach is consistent with the Pitch Perception Model (PPM), which provides an interpretive framework for relating production errors to perceptual salience, contour complexity, pitch-register control, temporal support, and possible first-language-based perceptual bias. For these reasons, a primarily qualitative and descriptive approach, supported by acoustic evidence, is the most suitable methodological orientation for the study.

Although this study is qualitatively oriented, it also uses descriptive quantitative acoustic evidence to support the analysis. This combination is necessary because Mandarin Chinese tone errors cannot be adequately described through auditory impression alone. Measurements such as F_0 range, pitch contour, duration, accuracy rates and confusion patterns make it possible to observe how tone production differs from the expected Mandarin Chinese targets. At the same time, these measurements are not used for broad statistical generalization, given the modest sample size. Instead, they are used to support qualitative interpretation of the error patterns produced by Ugandan

L2 speakers of Chinese.

3.3 Population and Sampling

3.3.1 Target Population

The accessible population consisted of intermediate learners enrolled in the Diploma in Chinese Language and Culture programme at Luyanzi Institute of Technology in Uganda. The 19 participants came from one intact Chinese class whose members had completed foundational instruction in Mandarin Chinese phonetics and had broadly comparable exposure to lexical tone contrasts. The programme and class therefore define the classroom population from which the recorded participants were drawn.

3.3.2 Sample Size

Nineteen (19) participants (10 males, 9 females; aged 20–27) were recorded. They reported six L1 backgrounds: Luganda, Lusoga, Lugisu, Runyankore, Samia, and Lugbarati. The first five are Bantu languages, while Lugbarati is commonly classified within Central Sudanic; together, these backgrounds provide a useful range for describing cross-speaker variability in Chinese tone production. From the perspective of PPM, this diversity is relevant because participants approach Chinese tones through different prior experiences of pitch organization, including differences in register-oriented and lexical uses of pitch. For acoustic reference, one native baseline Chinese speaker (First Class, Level 2, National Chinese Proficiency Test/PSC) was recorded under identical conditions to provide target-norm measurements for comparison. The baseline speaker was not treated as a participant group in the study, but as an acoustic reference point against which the Ugandan production data could be interpreted.

The sociolinguistic information available for the participants comprises first-language background, gender, age, Chinese proficiency level, and programme of study, as reported in Table 1. No additional individual characteristics are inferred beyond the information recorded in the manuscript.

The sample consisted of 19 Ugandan L2 speakers of Chinese who were selected from one intact Chinese class available at the time of data collection. The class was chosen because it provided a practically accessible group of students who had received

comparable classroom instruction, had similar exposure to Chinese learning materials, and were available for controlled recording under the same institutional conditions. Within this class, participants were recruited on the basis of availability and willingness to participate. Therefore, the sample should be understood as a feasible classroom-based sample rather than a statistically representative sample of all Ugandan L2 speakers of Chinese.

Although the participants came from six first-language backgrounds, namely Luganda, Lusoga, Lugisu, Runyankore, Samia and Lugbarati, the purpose of including these backgrounds was not to make broad generalizations about each language group. Rather, the first-language information was recorded in order to describe the linguistic diversity of the sample and to provide a possible reference point for interpreting individual and group patterns in Mandarin Chinese tone production. Given the small number of speakers in each first-language subgroup, the study treats L1 background as a descriptive and exploratory factor, not as a basis for strong statistical comparison among Ugandan languages.

The presence of only one non-Bantu first-language background, Lugbarati, is a result of the actual composition of the class from which the data were collected. Since the study relied on an available intact class, it was not possible to balance the sample across Bantu and non-Bantu language groups. For this reason, Lugbarati is included as part of the linguistic profile of the participants, but it is not used as the sole basis for making general claims about non-Bantu speakers' acquisition of Chinese tones. Any observations related to Lugbarati-speaking participants are therefore treated as preliminary and descriptive.

3.3.3 Sampling Technique

Stratified random sampling was used to ensure proportional representation across the reported L1 groups. Participants were first stratified by L1 backgrounds, after which random selection within each stratum minimized sampling bias while maintaining coverage of the cohort's linguistic diversity. The production materials comprised 74 elicitation items: 16 monosyllabic syllables, 32 disyllabic syllables representing the 16

tone combinations, and 26 items targeting tone-conditioning environments involving *yī* ‘one’, *bù* ‘not’, and neutral tone. Each item was produced three times by each participant, resulting in approximately 4,218 recorded syllables. This corpus size supports stable acoustic summarization and provides sufficient repeated measures for the limited statistical verification procedures reported in Chapter Four.

3.4 Data Collection Procedures

3.4.1 Data Collection Techniques

Audio recording through a controlled production task was the data-collection technique for both objectives. For Objective 1, repeated productions of monosyllabic tones, disyllabic combinations, and tone-conditioning items supplied the acoustic evidence needed to identify recurrent non-target patterns. For Objective 2, the same recordings were coded and grouped by tone category, phonological context, and participant background so that the distribution of the identified errors could be described. This method was selected because F0 contour, pitch range, alignment, and duration require an inspectable speech signal rather than impressionistic classroom observation alone.

3.4.2 Instruments

Data collection relied on controlled recording and a standardized processing pipeline. Speech was recorded using a Xiaomi Mi 10 Pro condenser microphone (44.1 kHz, mono) in a quiet conference room at Luyanzi Institute of Technology (ambient noise < 35 dB). Recordings were saved in WAV format and inspected immediately after each session; unclear tokens were re-recorded. The WAV files were then imported into Praat (v6.3), which was used for manual segmentation, TextGrid annotation, and F0 extraction at the syllable level. Microsoft Excel (2021) was used to organize raw measurements, check data-entry consistency, and prepare the working dataset. R (v4.3.1) was used for descriptive summaries and limited mixed-effects verification of repeated-measures patterns, while SPSS (v28) was used as a supplementary environment for data checking and confirmation of summary outputs. In methodological terms, this instrument chain made it possible to compare pitch height, pitch movement, contour shape, and duration in ways that remained compatible with PPM-based interpretation

in later chapters.

3.4.3 Materials

The elicitation materials were designed primarily for the production task and were selected to sample tonal realization across structural contexts that are central to the later results chapters. The materials included monosyllabic and disyllabic words, as well as items targeting tone sandhi and neutral-tone environments. Monosyllabic materials comprised 16 items representing four segmental templates across the four lexical tones (Tone 1 (T1) –Tone 4 (T4)). Disyllabic materials comprised 32 items covering all 16 tone combinations (Tone 1 (T1)–Tone 4 (T4) × Tone 1 (T1)–Tone 4 (T4)). In addition, 26 items targeted tone-conditioning environments involving *yī* ‘one’, *bù* ‘not’, and neutral-tone tokens (see Table 1). Together, these materials provided the basis for identifying error types and describing their distribution across tone categories and phonological contexts.

3.4.4 Research Procedure

Data collection and processing were carried out in a structured sequence designed to ensure consistency from elicitation to analysis. First, participants were briefed on the task and recorded individually in a controlled environment using the prepared elicitation materials. Each participant produced the target items under the same recording conditions, and all responses were saved immediately in WAV format. After each session, the recordings were checked for clarity, background interference, and token completeness. Any unclear or incomplete productions were re-recorded at once so that the final dataset would remain comparable across speakers.

Second, the completed recordings were transferred into the processing stage. The WAV files were imported into Praat, where the speech signal was examined and segmented manually at the syllable level using TextGrid annotation. This stage made it possible to align each token with its intended tonal target and relevant phonological context. Acoustic measurements were then extracted in a systematic manner, with attention to the tonal properties most relevant to the study, including pitch movement, pitch range, contour realization, and duration.

Third, the extracted measurements were organized into a working dataset for further interpretation. Microsoft Excel was used for data arrangement, consistency checks, and preliminary coding, after which the data were prepared for descriptive analysis and limited statistical verification in R and SPSS. At this stage, tokens were grouped according to tone category, prosodic environment, and speaker background so that recurrent patterns could be compared across conditions.

Finally, the organized results were interpreted in light of the study objectives and the explanatory dimensions emphasized in PPM. In practical terms, the procedure moved from controlled elicitation, to acoustic processing, to dataset organization, and then to interpretive analysis. This sequence ensured that the later discussion of tone errors, their distribution, and the cautious interpretation of those patterns remained closely grounded in systematically collected production evidence.

3.5 Quality Control

3.5.1 Validity

Validity concerns whether the procedures capture the intended properties of Chinese tone production and whether the resulting interpretations are warranted. In this study, content validity was supported by drawing stimuli from official standards to ensure coverage of major tonal patterns and tone-conditioning environments. Construct validity was addressed by focusing on established acoustic correlates of lexical tone, including F_0 level and trajectory, pitch range, turning-point behavior, and duration. Criterion validity was operationalized through comparison with a certified native baseline speaker recorded under identical conditions. Internal validity was strengthened by standardized recording procedures and a controlled environment designed to reduce confounding variation, while external validity was supported by including participants from multiple L1 backgrounds within the Ugandan context. Methodologically, the combination of controlled production data, acoustic comparison, and careful contextual coding provided a sound basis for examining how perceptual salience, contour complexity, register alignment, and temporal support may be associated with recurrent tone-production errors.

3.5.2 Reliability

Reliability concerns the consistency of measurement and annotation across speakers and repetitions. Instrumental consistency was supported by using the same recording device, sampling rate, and Praat settings for all sessions. Each item was produced three times so that the analysis could observe whether a non-target production recurred, while reducing the influence of nervousness, slips of the tongue, and incidental pronunciation fluctuation. Stability across repetitions or comparable items provided a basis for distinguishing a systematic error from an occasional mistake; a single deviant token was not treated as sufficient evidence of an error. Pilot recordings conducted before the main data collection also helped confirm material clarity and stable pitch tracking.

3.6 Data Analysis

Data analysis proceeded in staged form. Praat-based acoustic measurements were used to extract F_0 trajectories and related metrics from the recorded production tokens, after which descriptive summaries were produced to characterize recurrent output patterns across tones and contexts. Where appropriate for the repeated-measures structure of the dataset, limited statistical verification was used to corroborate whether observed differences remained stable when controlling for speaker- and item-level variability; any follow-up comparisons were treated as confirmatory rather than as the central analytic goal. The main interpretive emphasis remained on how the production data could be organized to identify error types, describe their distribution across first-language backgrounds (Luganda, Lugisu, Lusoga, Runyankore, Samia, and Lugbarati), and interpret the observed patterns cautiously within PPM.

All recordings underwent initial acoustics quality inspection, with background noise and amplitude irregularities addressed to maintain signal comparability across sessions. Cleaned files were imported into Praat (v6.3) and manually segmented using TextGrids to align syllable intervals with tonal targets and relevant segmental landmarks. F_0 contours were extracted using standardized settings consistent with prior tone-acquisition research (Shi, 2018; Mitkov, 2021). For each token, F_0 was sampled at three analytically motivated points—onset, turning point, and offset—to capture directional

movement within the syllable, and tone-specific summaries were computed (mean F_0 , pitch range, and duration). Pitch values were normalized in semitones relative to speaker-specific medians to enable cross-speaker comparison. Descriptive statistics were then used to summarize central tendencies and dispersion by tone, context, and L1 group.

The use of both SPSS and R was necessary because the two tools complemented each other. SPSS was used mainly for organizing coded data, producing descriptive statistics, and summarizing frequency-based patterns, while R was used for more flexible acoustic-data processing, visualization, and reproducible calculation. This combination allowed the study to describe the distribution of tone production errors among Ugandan L2 speakers of Chinese and to support the acoustic interpretation of F_0 , pitch range and duration patterns in Chapter Four.

The combination of qualitative and descriptive quantitative analysis directly supports the two objectives. Quantitative acoustic measures identify observable patterns such as pitch-range deviation, tone-category confusion, duration reduction, and accuracy differences across tonal contexts. Qualitative analysis then describes how these patterns appear in particular tone categories, lexical items, sandhi environments, and neutral-tone productions. The numerical evidence supports and checks the observed patterns; it is not treated as causal proof.

3.7 Ethical Considerations

The study followed Makerere University's Directorate of Graduate Training (DGT) ethical standards. Participants received a clear explanation of the study procedures and provided informed consent prior to recording. Confidentiality was maintained by removing personal identifiers and assigning participant codes (e.g., Student A–Student S) in all files and analyses. The study posed no physical or psychological risk, and participation was voluntary. All audio files and derived datasets were stored securely and will be retained for five years before deletion in accordance with institutional requirements. Ethical clearance was obtained through the Department of Linguistics, English Language Studies and Communication Skills.

3.8 Summary

This chapter has specified the methodological foundation for examining Chinese tone production by Ugandan participants, including participant selection, elicitation materials, recording conditions, and the acoustic procedures used to obtain comparable measurements. It links the controlled production recordings to the two objectives: identifying recurrent tone-production errors and describing their distribution across tones, contexts, and participant backgrounds. Perception remains a limited theoretical reference within PPM rather than a separately tested empirical component.

CHAPTER FOUR: DATA PRESENTATION AND ANALYSIS

4.1 Introduction

The analysis is based on the productions of nineteen Ugandan participants from six L1 backgrounds: Luganda, Lusoga, Lugisu, Runyankore, Samia, and Lugbarati. Three acoustic dimensions are used throughout the analysis: F_0 , syllable duration, and pitch-range excursion (ΔF_0). Together, these measures make it possible to identify recurrent non-target patterns, compare them with the native baseline, and describe how they vary across tones, L1 groups, and phonological contexts within a single PPM-oriented account. Because absolute Hertz values differ across individuals, the chapter relies on normalized comparison with the native baseline so that the reported patterns reflect production tendencies rather than physiological pitch differences. This normalization allows the discussion to remain descriptive and empirically transparent while avoiding over-interpretation of raw pitch height alone.

For the identification of salient deviations in Objective 1, the analysis adopts a practical threshold: deviations in pitch height or trajectory exceeding two semitones relative to the native baseline are treated as sufficiently large to warrant classification as tonal errors. This threshold helps separate recurrent patterns from minor token-level fluctuation and provides a consistent criterion for identifying the main error types.

Table 1 is retained here only as a compact reference for participant profiles and monosyllabic production targets used in the subsequent analysis. The discussion therefore moves from the identification of tone production errors to an account of their distribution, then to comparison with previous studies, and finally to a brief summary. Given the small and uneven number of participants across first-language backgrounds, the patterns discussed here were interpreted cautiously. The L1 background comparison is intended to identify possible tendencies among the Ugandan L2 speakers of Chinese in this sample, rather than to make broad claims about the Chinese tone acquisition of each Ugandan language group.

Table 1: Participant Profiles and Monosyllabic Tonal Performance Targets

ID	L1 Backgro und	T1 (ma) F₀ (Hz)	T1 (ma) Dur (ms)	T2 (hai) F₀ (Hz)	T2 (hai) Dur (ms)	T3 (jiu) F₀ (Hz)	T3 (jiu) Dur (ms)	T4 (tui) F₀ (Hz)	T4 (tui) Dur (ms)
Native	Mandarin	126.6	540	110.4	752	104.4	879	136.5	582
A	Luganda	219.6	534	181.1	610	182.1	587	221.5	308
B	Luganda	97.0	540	84.9	407	85.4	455	95.8	427
C	Luganda	238.3	463	174.2	568	174.1	705	239.1	339
D	Luganda	157.2	376	108.5	341	113.0	470	155.2	236
E	Lugbarati	117.4	522	96.8	536	98.7	619	142.0	277
F	Lugisu	205.0	548	161.5	360	139.0	443	205.2	275
G	Lugisu	244.7	436	201.0	411	202.1	567	251.6	317
H	Lugisu	131.6	405	116.8	525	114.6	502	165.8	274
I	Lugisu	252.4	688	209.7	731	208.1	716	227.2	435
J	Lusoga	128.8	279	128.7	314	102.5	522	125.2	339
K	Lusoga	192.7	528	129.4	554	137.3	594	175.0	353

ID	L1 Backgro und	T1 (ma) F₀ (Hz)	T1 (ma) Dur (ms)	T2 (hai) F₀ (Hz)	T2 (hai) Dur (ms)	T3 (jiu) F₀ (Hz)	T3 (jiu) Dur (ms)	T4 (tui) F₀ (Hz)	T4 (tui) Dur (ms)
L	Lusoga	168.6	551	163.4	498	167.8	757	227.5	437
M	Lusoga	118.6	511	193.0	702	175.1	675	216.4	350
N	Lusoga	122.9	502	90.5	833	90.1	831	116.5	442
O	Runyan kore	169.2	505	146.6	550	123.0	787	153.0	384
P	Runyan kore	125.9	581	100.4	606	97.7	744	153.3	373
Q	Runyan kore	248.5	617	188.4	681	193.3	719	243.9	441
R	Runyan kore	138.5	407	100.9	485	99.9	571	132.0	283
S	Samia	195.0	514	113.4	567	103.4	747	160.7	343

With this structure in place, Section 4.2 begins by identifying the tone production errors observed in the production data.

4.2 Tone Production Errors Made by Ugandan L2 Speakers of Chinese

Uganda L2 speakers of Chinese are able to articulate the basic pitch movements of Chinese tones to a certain extent but often fail to produce them in a stable, target-like manner that conforms to native Chinese tonal norms, which constitutes tone production errors. This section addresses Objective 1 by identifying the tone production errors observed in the participants' Mandarin output. In the present chapter, the seven errors are derived inductively from recurrent acoustic patterns in the dataset itself. Their labels are therefore used as descriptive terms created for this study rather than as the direct adoption of any single prior error taxonomy, although the identification relies on acoustic dimensions that are also central in earlier research. An error is treated here as a recurrent non-target pattern that occurs often enough, and with clear enough acoustic evidence, to be described as a stable tendency rather than an isolated deviation. Operationally, the identification relies on pitch height, contour shape, turning-point realization, duration, and context-sensitive tonal realization. Within a PPM perspective, these recurrent patterns are important because they point to areas where tonal categories are not yet stably perceived, differentiated, or mapped onto production. The focus of this section, however, remains descriptive: the purpose of this section is to identify the tone production errors clearly before their distribution is examined in Section 4.3.

The first deviation pattern involves monosyllabic tones, where the tonal target is realized without the additional prosodic load of coordinating multiple syllables. Monosyllabic productions therefore provide the clearest window into whether speakers can approximate Chinese's canonical contour categories and whether deviations cluster into recurrent errors. Two broad monosyllabic tendencies are prominent in the dataset: (i) pitch-range expansion that produces overshoot and exaggerated excursions, and (ii) contour reshaping that reduces or truncates complex movements, especially for the dipping contour.

The identification begins with monosyllabic tones because they provide the clearest baseline for recognizing recurrent non-target patterns before additional connected-speech pressures are introduced. These isolated-syllable patterns lay the foundation for

understanding more complex prosodic errors. With this baseline established, the analysis now focuses on specific occurring misarticulations. The more detailed identification of tone production errors is reserved for the coming sections.

4.2.1 Pitch-range Expansion and Overshoot

A recurrent deviation in monosyllabic contour tones is pitch-range expansion and overshoot. This pattern is summarized in Table 2, which reports mean acoustic deviation in pitch range (ΔF_0) for monosyllabic contours.

Table 2: Mean Acoustic Deviation in Pitch Range (ΔF_0) for Monosyllabic Tones

Tone Category	Native Mean ΔF_0 (Hz)	Speaker Mean ΔF_0 (Hz)	Deviation (Hz)	p-value
Tone 1 (Level)	16.95	24.81	+7.86	P<0.051
Tone 2 (Rising)	53.6	67.9	+14.3	P<0.069
Tone 3 (Dipping)	37.5	69.8	+32.3	P<0.001
Tone 4 (Falling)	76.7	97.3	+20.6	P<0.027

Note. Table 2 reports tone-level mean ΔF_0 values aggregated across monosyllabic items and speakers. The specific lexical examples discussed below, such as *cōng* ‘clever’, *bēi* ‘cup’, *lěng* ‘cold’, *shuǎng* ‘cool’ and *fó* ‘buddha’, are drawn from the item-level acoustic measurements underlying the values in the table, not from the aggregated values displayed in Table 2.

Table 2 provides the aggregate acoustic basis for identifying pitch-range expansion and overshoot. It shows that the participant mean ΔF_0 values are higher than the native

baseline for all four tone categories, especially Tone 2 (T2), Tone 3 (T3) and Tone 4 (T4). However, because Table 2 reports tone-level averages, it does not display individual word-level values for particular speakers. The following lexical examples are therefore used to interpret the aggregate pattern qualitatively and to show how pitch-range expansion appeared in specific productions.

At the item level, pitch-range expansion was especially visible among some Luganda, Lusoga and Runyankore L1 speakers. In the Luganda group, this pattern appeared in Tone 1 (T1) items such as *cōng* ‘clever’ and *bēi* ‘cup’. In standard Mandarin production, T1 should be realized as a relatively stable high-level tone with limited pitch movement. In the item-level acoustic measurements, the native speaker’s production of *cōng* ‘clever’ had a narrow tonal range of 6.7 Hz, whereas Student A produced the same item with a tonal range of 30.2 Hz. Similarly, the native speaker’s production of *bēi* ‘cup’ had a tonal range of 23.4 Hz, while Student C produced it with a tonal range of 55.8 Hz. These examples suggest that the speakers did not simply raise the pitch level; rather, they introduced excessive pitch movement into a tone that should have remained comparatively level.

For the Lusoga group, overshoot was more evident in T3 items such as *lěng* ‘cold’ and *shuǎng* ‘cool’. In standard Mandarin, T3 requires a controlled low or dipping contour, rather than an exaggerated fall-rise movement. In the item-level acoustic measurements, the native speaker’s tonal range for *lěng* ‘cold’ was 32.6 Hz, while Student M produced it with a much wider range of 159.0 Hz. The same speaker produced *shuǎng* ‘cool’ with a tonal range of 182.8 Hz, compared with the native speaker’s 38.1 Hz. This indicates that the speaker expanded the pitch excursion far beyond the target range, possibly in an attempt to make the dipping contour more perceptually salient.

A similar pattern can be seen in the Runyankore group. For the T2 item *fó* ‘buddha’, the native speaker produced a moderate rising contour with a tonal range of 55.4 Hz, while Student O produced it with a tonal range of 152.6 Hz. For the T3 item *lěng* ‘cold’, the same speaker produced a tonal range of 86.1 Hz, compared with the native speaker’s 32.6 Hz. These examples show that pitch-range expansion affected both rising and dipping tones. The error therefore lies not only in producing an incorrect tone category,

but also in over-scaling the pitch movement required for Chinese tonal contrast. Acoustically, the most consistent sign of this error is the expansion of pitch excursion beyond the native baseline. Table 2 supports this interpretation at the aggregate level, while the item-level examples above show how the pattern appears in actual productions. Ugandan L2 speakers of Chinese often implemented contour movement with enlarged excursions, producing trajectories that were more extreme than the native target. The pattern is therefore identified as a distinct error because the productions repeatedly preserved tonal movement but overshoot the target-like range.

Figure 2: Student D's Tone 2 Structures

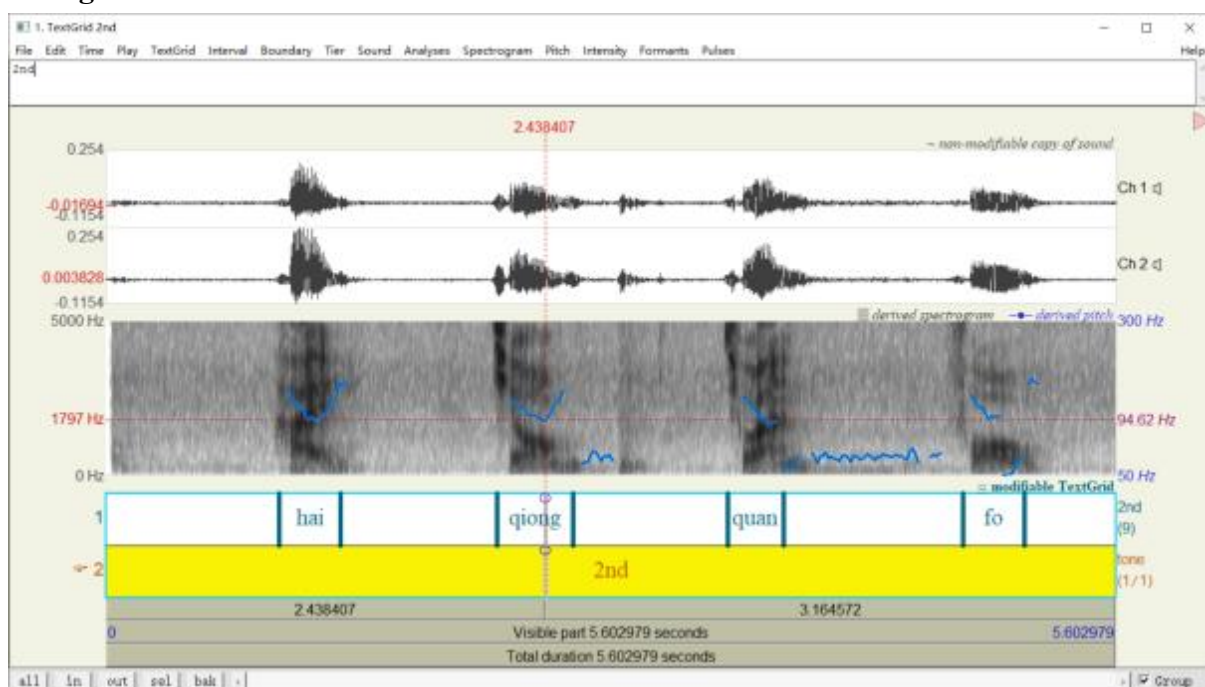


Figure 3: Student L's Tone 2 Structures

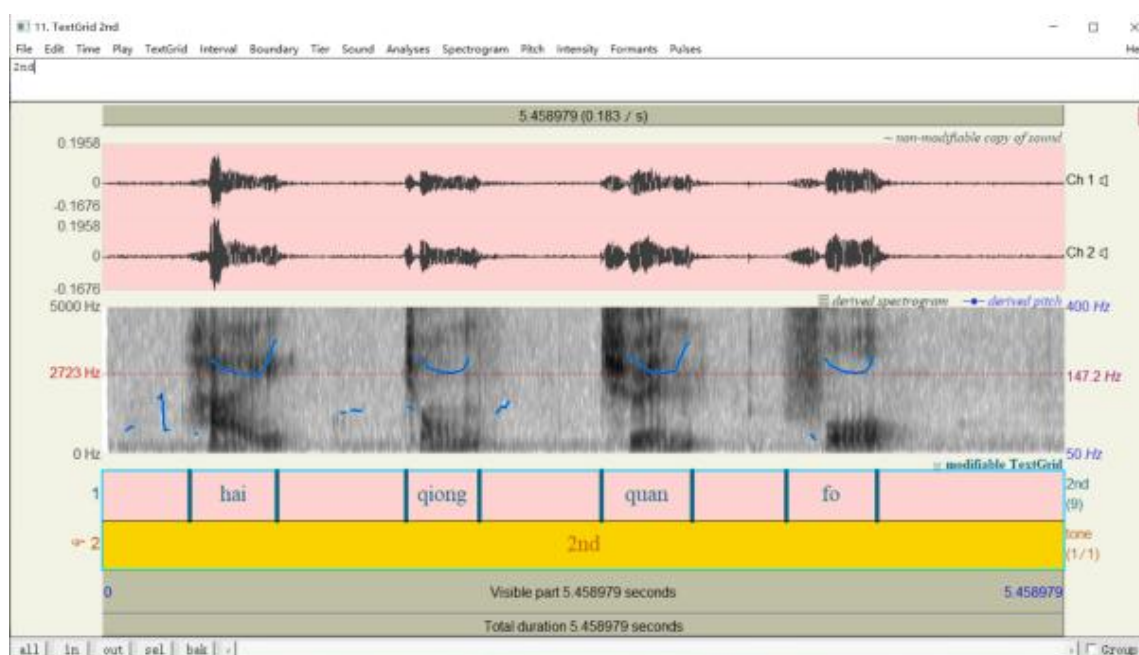


Figure 2 (*Student D's Tone 2 Structures*) and Figure 3 (*Student L's Tone 2 Structures*) provide representative illustrations of enlarged rising movement. They are included as visual examples of the same overscaling pattern discussed above, not as the direct source of the individual item-level numerical values.

At the identification stage, the key point is that the productions remain recognizably dynamic but depart from the native target in scaling. This justifies treating pitch-range expansion and overshoot as a separate recurrent error type.

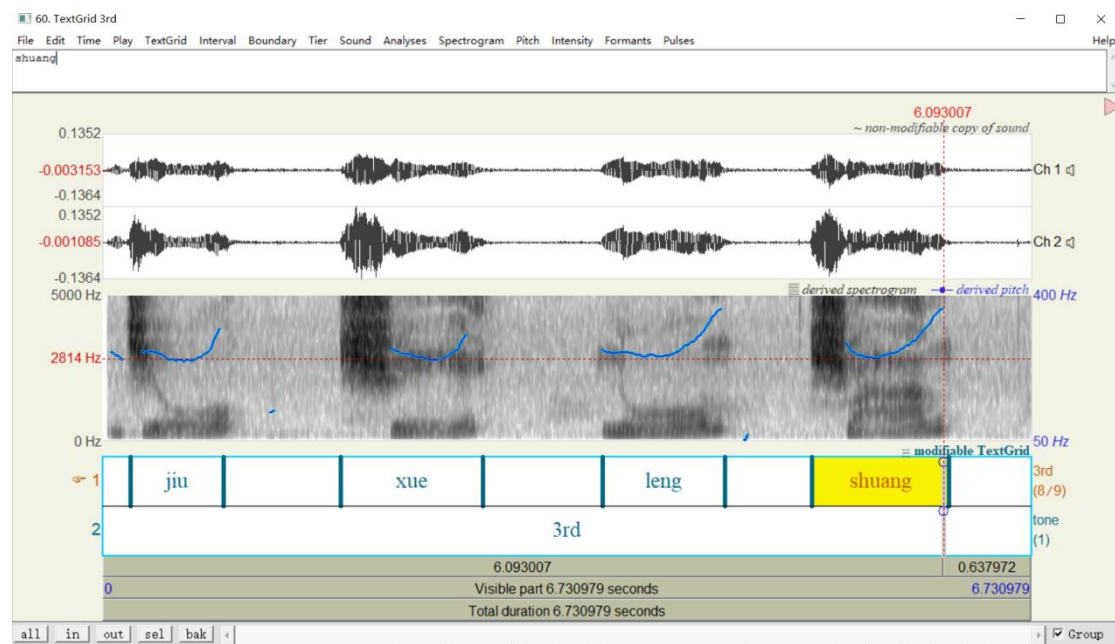
4.2.2 Dipping-contour Truncation and Reduction of Complex Movement

A second recurrent monosyllabic syllable deviation concerns the dipping contour, which is widely recognized as the most complex canonical tone production error in Chinese. The dataset shows that speakers often reduce or truncate the canonical fall–rise structure. Instead of implementing the full dip and recovery, productions frequently show partial movement, such as a fall without a target-like re-rise, or a shallow dip followed by a delayed rise.

A representative illustration is provided in Figure 4 (*Student M's Tone 3 Structures*). The figure shows a contour pattern that does not fully realize the canonical fall–rise structure, either because the low turning point is insufficiently reached or because the

recovery rise is incomplete.

Figure 4: Student M's Tone 3 Structures



For identification purposes, the defining evidence is reduced realization of the canonical fall–rise contour. The error is therefore distinguished from a simple register difference because it affects the internal shape of the tone rather than pitch height alone. From a PPM perspective, this error reflects incomplete encoding of the internal two-stage structure of T3. If the speaker does not stably perceive both the low target and the recovery rise as obligatory parts of the same category, production tends to preserve only the more perceptually available portion of the contour often the fall or shallow dip, while the re-rise is weakened, delayed, or omitted. At this stage, it is sufficient to note that the pattern is systematic enough to constitute a distinct contour-structure error type. The error arises because T3 requires the speaker to coordinate a relatively complex fall–rise movement within one syllable. In the controlled monosyllabic reading task, the target Tone 3 items included *jiǔ* ‘nine’, *xuě* ‘snow’, *lěng* ‘cold’ and *shuǎng* ‘cool’. In standard Mandarin citation-form production, these items should show a low dipping contour: the pitch first moves downward toward a low turning point and then rises again. Therefore, the difficulty is not only to produce a low pitch, but to maintain the internal two-stage structure of the contour.

In the data, different Ugandan L2 speakers of Chinese simplified this complex movement in different ways. Some speakers shortened the syllable so that there was not enough time for the full fall–rise contour to unfold. For example, the native speaker produced *jiǔ* ‘nine’, *xuě* ‘snow’, *lěng* ‘cold’ and *shuǎng* ‘cool’ with relatively long durations, but some speakers produced the same items with much shorter durations. Student B, whose L1 is Luganda, produced *jiǔ* ‘nine’ and *xuě* ‘snow’ with shortened durations and relatively limited tonal movement. In such cases, the contour tended to sound like a short low or falling tone rather than a full dipping tone with recovery. The learner therefore preserved only part of the T3 movement.

Other speakers simplified T3 by reducing the size of the dip or by weakening the final rise. For example, Student J, whose L1 is Lusoga, produced T3 items such as *jiǔ* ‘nine’, *xuě* ‘snow’, *lěng* ‘cold’ and *shuǎng* ‘cool’ with compressed pitch movement, especially in *shuǎng* ‘cool’, where the tonal range was much smaller than the native baseline. Instead of a clear fall followed by recovery, the production was closer to a flattened or weakly falling contour. This suggests that the speaker did not fully realize the low turning point and the subsequent rising movement as two necessary components of T3. A third pattern can be seen in cases where the speaker produced a delayed or abrupt recovery rise. Figure 4, which illustrates Student M’s production of the T3 sequence *jiǔ* ‘nine’–*xuě* ‘snow’–*lěng* ‘cold’–*shuǎng* ‘cool’, shows that the fall–rise movement was not evenly distributed across the syllable. The speaker appeared to place much of the rising movement late in the syllable. This kind of production does not simply omit the rise; rather, it compresses the complex movement into a shorter portion of the syllable. As a result, the contour becomes less target-like because the low turning point and recovery rise are not smoothly coordinated.

A similar simplification also appeared among some Runyankore L1 speakers. For example, Student R produced *jiǔ* ‘nine’, *xuě* ‘snow’ and *shuǎng* ‘cool’ with relatively compressed tonal ranges and shortened durations compared with the native production. This suggests that the learner reduced the full dipping contour into a simpler low or slightly falling pattern. Student O, by contrast, showed a wider pitch movement in items such as *lěng* ‘cool’, but the movement was still difficult to control as a balanced fall–

rise contour. These examples show that dipping-contour truncation does not take only one form. It may appear as shortened duration, reduced dip, missing or weakened recovery rise, delayed re-rise, or an unevenly scaled fall–rise movement.

These examples indicate that T3 errors are produced when Ugandan L2 speakers of Chinese simplify the internal structure of the Chinese dipping tone. Instead of producing T3 as a coordinated fall–rise contour, some speakers reduce it to a fall, a low-level tone, a shallow dip, or a late-rising contour. The error therefore reflects difficulty in controlling both the temporal and pitch dimensions of T3.

4.2.3 Tone 2 and Tone 3 Confusion

A high-frequency confusion exists between T2 (rising) and T3 (dipping), such that productions intended for one category frequently resemble the other. Table 3 reports the production confusion matrix (row-normalized probabilities). The matrix reveals that T2 shows the lowest accuracy (0.51), with a substantial portion (0.24) produced in a way that patterns with T3. T3 is likewise produced as T2 at a comparable rate (0.23). This bidirectional confusion suggests that the distinction between these two categories is unstable in production, producing a prominent error class centered on overlap in the rising-dipping region.

Table 3: Production Confusion Matrix (Row-Normalized Probabilities)

Target / Produced	Tone 1	Tone 2	Tone 3	Tone 4
Tone 1	0.89	0.04	0.02	0.06
Tone 2	0.06	0.51	0.24	0.19
Tone 3	0.02	0.23	0.59	0.15
Tone 4	0.03	0.10	0.11	0.77

The scale of this confusion is quantified in Table 4, which calculates the share of total errors accounted for by T2 and T3 cross-confusions. The result indicates that T2 and

T3 cross-confusions constitute 36.4% of total production errors (91 out of 250; 95% CI [30.4%, 42.4%]). This proportion underscores that T2 and T3 confusion is not a marginal phenomenon but one of the dominant instability loci in the dataset.

Table 4: T2 and T3 Confusion Share Calculation

Metric	Value
Total Tone Production Errors (E_{total})	250
T2 and T3 Cross-Confusions ($E_{2\leftrightarrow 3}$)	91
Proportion of Total Errors	36.4%
95% Confidence Interval	[30.4%, 42.4%]

From an acoustic perspective, the confusion can be described as a reduction in contrast along two dimensions: (i) reduced differentiation in turning-point structure (e.g., weakened initial fall for T3), and (ii) overlap in rising trajectories (e.g., T2 realized with atypical onset or delayed movement that makes it resemble a dipping contour).

The confusion between T2 and T3 occurred mainly in the internal pitch movement of the syllable, rather than in the segmental form of the word. In standard Mandarin, T2 should be produced as a rising tone. It usually begins from a mid or relatively low pitch level and then rises clearly toward the end of the syllable. In the recording materials, this pattern was expected in T2 items such as *hái* ‘still/also’, *qióng* ‘poor’, *quán* ‘whole’ and *fó* ‘buddha’. However, some Ugandan L2 speakers of Chinese did not produce these syllables with a smooth and continuous rising contour. Instead, the pitch either started too low, showed a slight initial dip, or delayed the rise until the later part of the syllable. As a result, the intended T2 production became acoustically similar to T3, because it contained a low point or a delayed rise that resembled the dipping structure of T3.

The reverse pattern was also observed in T3 items such as *jiǔ* ‘nine’, *xuě* ‘snow’, *lěng*

'cold' and *shuǎng* 'cool'. In standard Mandarin citation-form production, T3 should show a low dipping contour, with an initial fall toward a low turning point and a subsequent recovery rise. However, some speakers simplified this contour by weakening the initial fall or by producing only a rising movement from a low starting point. When the low turning point was not clearly realized, items such as *jiǔ* 'nine', *xuě* 'snow', *lěng* 'cold' and *shuǎng* 'cool' sounded closer to T2 than to a full dipping T3. In these cases, the problem was not that the speakers failed to use pitch movement altogether, but that they failed to maintain the contrast between a rising contour and a dipping contour.

This confusion was also visible in disyllabic contexts. In items such as *shíjiān* 'time', *yínháng* 'bank', *cídiǎn* 'dictionary' and *lán sè* 'blue', the first syllable carries T2. The target production should preserve a clear rising movement on the T2 syllable before moving into the following tone. However, under the temporal pressure of disyllabic production, some speakers delayed the rise or compressed the contour, making the T2 syllable sound closer to a shortened or shallow T3. Conversely, in items such as *shǒujī* 'mobile phone', *xǐnglái* 'wake up', *suǒyǐ* 'therefore' and *shǒutào* 'glove', the first syllable involves T3 or a T3-related context. Some speakers reduced the dipping movement and produced the first syllable with a low-to-rising shape, which made it perceptually close to T2.

Therefore, the T2 and T3 confusion should be understood as a confusion at the level of contour structure. T2 and T3 both involve dynamic pitch movement, and both may contain a rising portion. The main difference is that T2 is primarily a rise, whereas T3 requires a low turning point before recovery. The Ugandan L2 speakers of Chinese who produced this error appeared to rely mainly on the presence of rising movement, while giving insufficient control to the timing and depth of the low turning point. This explains why intended T2 items could be heard as T3 when the rise was delayed or preceded by a dip, and why intended T3 items could be heard as T2 when the initial fall and low turning point were weakened.

In other words, the confusion did not occur randomly across all tones. It was concentrated in the acoustic corridor where rising and dipping movements overlap. The

data in Table 5 show the overall direction of this problem, but the concrete lexical evidence shows where it happened: in T2 syllables such as *hái* ‘still/also’, *qióng* ‘poor’, *quán* ‘whole’ and *fó* ‘buddha’ when the rise was delayed or made dip-like, and in T3 syllables such as *jiǔ* ‘nine’, *xuě* ‘snow’, *lěng* ‘cold’ and *shuǎng* ‘cool’ when the dipping contour was reduced to a simple rise or shallow low-rising movement. This pattern supports the interpretation that the speakers had not fully stabilized the phonological boundary between T2 and T3.

Within a PPM interpretation, this error points to compression of the contrastive space between two neighboring dynamic categories. T2 and T3 depend on partially overlapping cues -rise direction, turning-point placement, and temporal deployment, but they require different cue-weighting profiles. For identification purposes, the crucial point is that the speakers repeatedly produce the two tones in an overlapping acoustic corridor. The pattern is therefore classified as a distinct confusion-based error type supported jointly by the confusion matrix and the contour observations.

The distribution of this overlap across speakers and contexts is examined in Section 4.3.

4.2.4 Duration Reduction

Table 5 summarizes the duration comparison between participant and native productions. The results show that duration for all four tones are shorter in the participant data than in the native baseline. The most pronounced deficits are observed for T3 and T2, with reductions of 344.03 ms and 272.17 ms respectively. The magnitude of these reductions suggests a recurrent temporal limitation rather than occasional fast speech.

The p-values in Table 5 are included only as supplementary checks and are not treated as proof of causation. In this dataset, the smaller p-values for T2, T3, and T4 indicate that the observed duration differences are sufficiently consistent to support closer qualitative interpretation, while the T1 result should be read cautiously. The main analytical issue is how shortened duration constrains the time available for a rising or dipping contour to unfold in particular tone items.

Since this study is primarily qualitative and descriptive, the p-values are not used to

make broad inferential claims about all Ugandan L2 speakers of Chinese. Rather, they show that the duration differences observed in this dataset are consistent enough to support qualitative interpretation. In other words, the statistical values help confirm that duration reduction is not limited to a few isolated tokens, but the main analytical question remains how shortened duration affects the realization of specific tone items. For the purposes of Objective 1, these reductions define a distinct temporal error type: the tonal contour is produced within a shortened syllabic window, which limits full contour realization and increases the likelihood of undershoot.

Table 5: Duration Comparison between Speaker and Native Tone Productions

Tone	Native Duration (Ms)	Speaker Duration (Ms)	Diff (ms)	Sig. (p)
Tone 1	582.25	530.24	-52.01	p < 0.05
Tone 2	848.50	576.33	-272.17	p < 0.001
Tone 3	1023.75	679.72	-344.03	p < 0.001
Tone 4	622.00	386.12	-235.88	p < 0.001

Note: For each speaker, durations were averaged across four monosyllabic items per tone (T1: *ma* ‘mother’/*bei* ‘cup’/*cong* ‘clever’/*guo* ‘pot’; T2: *hai* ‘still/also’/*qiong* ‘poor’/*quan* ‘whole’/*fo* ‘buddha’; T3: *jiu* ‘nine’/*xue* ‘snow’/*leng* ‘cold’/*shuang* ‘cool’; T4: *tui* ‘retreat’/*run* ‘moist’/*yuan* ‘yard’/*wang* ‘forget’). Group means were computed across 19 speakers.

This temporal compression is treated as a distinct error type because it repeatedly constrains contour realization within a shortened syllabic window. Its broader interpretive significance is taken up later in the comparison of error patterns across tones and contexts.

The qualitative significance of duration reduction becomes clearer when the figures in Table 5 are connected with the specific recording items. T1 items such as *mā* ‘mother’, *bēi* ‘cup’, *cōng* ‘clever’ and *guō* ‘pot’ were less affected because T1 is a high-level tone

and does not require complex internal pitch movement. Although these syllables were also slightly shorter than the native baseline, the reduction was relatively small. Therefore, the main effect of duration reduction in T1 was not category confusion, but a possible reduction in steadiness and naturalness.

The effect was more serious in T2 items such as *hái* ‘still/also’, *qióng* ‘poor’, *quán* ‘whole’ and *fó* ‘buddha’. In standard Mandarin, these syllables should be produced with a clear rising contour. However, when the syllable duration was shortened, the speaker had less time to begin and complete the rise. As a result, the rise could become delayed, compressed or incomplete. For example, *hái* ‘still/also’ or *qióng* ‘poor’ may still contain some upward pitch movement, but the rise may start too late or fail to reach a sufficiently clear endpoint. In such cases, the error is not simply that the syllable is short; the shortened duration weakens the acoustic shape of T2 itself.

The strongest effect appeared in T3 items such as *jiǔ* ‘nine’, *xuě* ‘snow’, *lěng* ‘cold’ and *shuǎng* ‘cool’. In citation-form Chinese, T3 requires a fall toward a low turning point followed by some recovery movement. This contour needs more temporal space than a level or simple falling tone. When Ugandan L2 speakers of Chinese produced these syllables with reduced duration, the full dipping contour often could not unfold. Some productions preserved only the falling or low portion, while the recovery rise was weakened, delayed or omitted. As a result, *jiǔ* ‘nine’, *xuě* ‘snow’, *lěng* ‘cold’ and *shuǎng* ‘cool’ could sound like a low-level tone, a short falling tone, or a shallow dip rather than a full T3 contour.

T4 items such as *tuì* ‘retreat’, *rùn* ‘moist’, *yuàn* ‘yard’ and *wàng* ‘to forget’ were also shortened, but the effect was different. Since T4 is a falling tone, speakers could often preserve the general falling direction even within a shorter syllable. However, the fall could become abrupt and overly compressed, making the contour less natural and less fully scaled than the native baseline. Therefore, duration reduction affected all four tones, but it had the strongest qualitative impact on T2 and T3 because these tones depend more heavily on the timing of internal pitch movement.

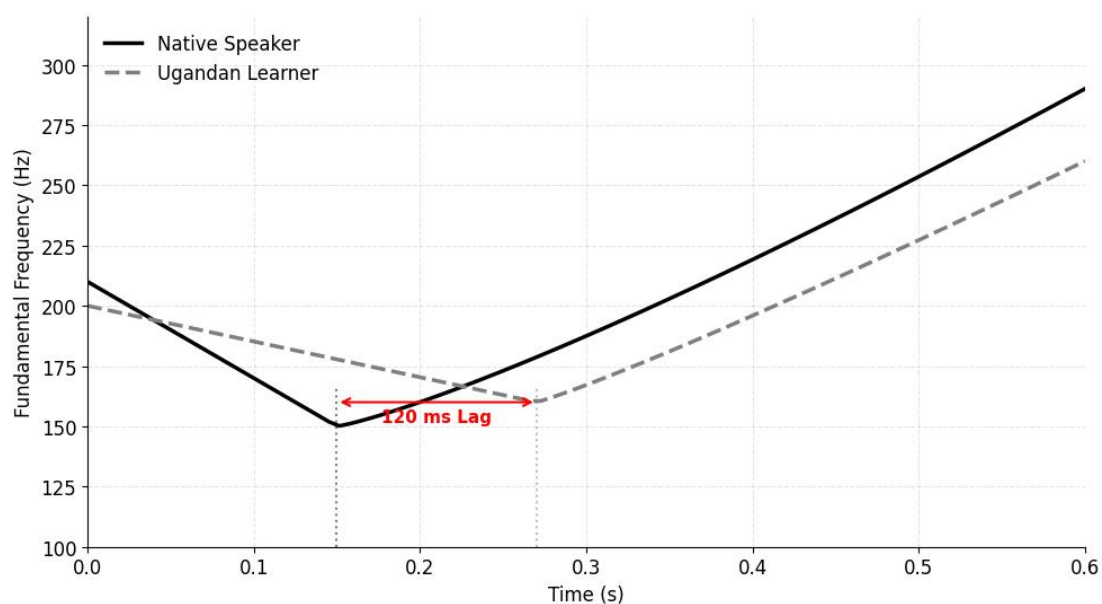
For this reason, duration reduction should not be interpreted merely as faster speech. It is a temporal-implementation error that reduces the acoustic space available for tonal

realization. In the present data, shortened duration helps explain why T2 may show delayed or incomplete rising movement, why T3 may lose its dipping structure, and why the T2 and T3 contrast becomes less stable.

4.2.5 Delayed Rise Onset and Alignment in Tone 2

Beyond overall duration reduction, the dataset provides evidence for delayed alignment of the rising gesture in T2. Figure 5 illustrates a 120 ms lag in rise onset relative to the native baseline.

Figure 5: Delayed Rise Onset in F₀ Trajectories (Tone 2)



Note : Figure 5 illustrates the timing difference between the native and participant contours. The lag is analytically useful because it provides a concrete acoustic signature of delayed contour deployment rather than a vague impression of inaccuracy.

For T2 words such as *hái* ‘still/also’, *qióng* ‘poor’, *quán* ‘whole’, and *fó* ‘Buddha’, the target rise should begin early enough to be heard as a sustained rising contour. In delayed-rise tokens, the syllable may begin too low or too level and only rise near the end. The resulting contour is not simply short; it is mistimed. This timing mismatch helps explain why some T2 tokens drift toward T3-like realizations.

Within a PPM perspective, this pattern suggests difficulty in coordinating the onset of rising movement with the available temporal window of the syllable. When the rise begins too late, the contour may still be identifiable as dynamic, but it is less likely to

match the target timing and shape.

At the identification stage, the relevant point is that the rise is present but mistimed. This recurrent delay justifies separating rise-alignment error from general duration reduction.

4.2.6 Citation-form Persistence in Sandhi Contexts

Table 6 reports the results for sandhi and neutral-tone contexts. T3 sandhi shows the lowest rate of target-like realization (46.8%), while sandhi *yi* ‘one’ and *bu* ‘not’ show 61.3% and 58.7% respectively. These results indicate that context-conditioned tone changes are less stable than citation-form productions⁸.

Examples from the elicitation list show how this error occurs. In *shíyī* ‘eleven’, *yìxiē* ‘some’, *yìzhí* ‘continuously’, *yìqǐ* ‘together’, and *yígòng* ‘altogether’, *yī* ‘one’ should change according to the following tone rather than remain in its citation form. Similarly, in *bùgāo* ‘not high’, *bùlái* ‘not come’, *bùmǎi* ‘not buy’, and *búhuì* ‘will not/cannot’, *bù* ‘not’ should be adjusted according to the following tone. When participants retained the citation tone instead, the error reflected weak context-sensitive tonal remapping rather than failure to pronounce an isolated tone.

⁸ Citation-form productions here refer to tones pronounced in their isolation form, that is, the default lexical tone shape produced when a syllable is said alone rather than in connected speech or in a context that triggers tonal alternation. In the present study, the term contrasts with context-conditioned surface realizations in environments such as Tone 3 sandhi, sandhi *yi*, sandhi *bu*, and neutral-tone reduction, where the target output is not the unchanged citation form. See Moisik S R, Lin H, Esling J H. A study of laryngeal gestures in Mandarin citation tones using simultaneous laryngoscopy and laryngeal ultrasound (SLLUS)[J]. Journal of the International Phonetic Association, 2014, 44(1): 21-58.

Table 6: Accuracy Rates for Sandhi and Neutral Tone Contexts

Phonological Category	Primary Error Pattern	Accuracy (%)
Tone 3 Sandhi	Failure to change first T3 to T2	46.8%
Sandhi "yi 'one'"	Retention of Tone 1 citation form	61.3%
Sandhi "bu 'not'"	Retention of Tone 4 citation form	58.7%
Neutral Tone	Insufficient duration/pitch reduction	52.4%

This error reflects sensitivity to phonological environment rather than the intrinsic difficulty of a single citation tone. In other words, the problem is not only how a tone is produced in isolation, but also whether the speaker adjusts that tone when the surrounding context requires modification. For identification purposes, the defining feature here is repeated retention of citation-form behavior in environments where contextual tonal adjustment is required.

4.2.7 Neutral-tone Reduction Failure

A common non-target pattern is reduction failure: speakers realize the neutral-tone syllable with insufficient reduction in duration and/or pitch prominence, so that it remains closer to a full lexical tone than to the expected weak form. This error is identified here as a contextual reduction problem.

The neutral-tone items *bāozi* 'steamed bun', *chūle* 'except', *diǎnxin* 'snack', and *dàifu* 'doctor' illustrate this pattern. In target-like Chinese, the second syllable should be reduced in prominence and duration relative to a full lexical tone. In several participant productions, however, the neutral syllable retained too much pitch prominence or syllable length, making it sound closer to a full tone. This is why the error is treated as neutral-tone reduction failure rather than as a simple wrong-tone choice.

From a PPM perspective, this error suggests that contextual environments place extra pressure on tonal perception and production because speakers must track both the

underlying category and the context-driven change in pitch pattern or prominence. Its fuller distribution across related morphophonemic environments is discussed in Section 4.3.2. This suggests that the reduced status of neutral tone is not yet being represented as a reliable prominence contrast in connected speech.

4.3 Distribution of Tone Production Errors

Having identified the major errors, this section addresses Objective 2 by describing their distribution. It first clarifies the acoustic basis on which the seven errors are distinguished, and then shows how those errors are distributed across L1 backgrounds, prosodic contexts, and tone categories.

4.3.1 Acoustic Basis of Tone Production Errors

The findings show that the tonal errors produced by learners of Chinese in Uganda are mainly of seven kinds: pitch-range expansion and overshoot, truncation of the dipping contour, T2 and T3 confusion, duration reduction, delayed rise onset in T2, persistence of citation forms in sandhi contexts, and failure to weaken the neutral tone appropriately. Put simply, the speakers' main problem is not that they produce no pitch movement at all, but that they do not yet produce Chinese tones in a stable, fully differentiated, and context-appropriate way. Their errors show that the tonal system in interlanguage remains only partially established, especially when speakers are required to control contour shape, timing, and contextual adjustment at the same time.

At the most basic level, the errors are phonological in nature because they reveal instability in tone-category formation. This is most visible in the recurring confusion between T2 and T3. These two tones are both dynamic contours and therefore place greater demands on speakers than the relatively more static T1 or the more clearly falling T4. In the present data, speakers often failed to keep the two categories distinct in production: T2 was not always realized with a sufficiently clear and timely rise, while T3 frequently lost its full dipping contour and was produced in a reduced or simplified form. This means that the problem is not merely one of inaccurate surface pronunciation, but one of incomplete category separation. In other words, the speakers seem to have access to the general existence of tonal contrast, but the internal boundaries between

some tone categories remain unstable.

The errors are also phonetic in nature because they involve difficulties in the acoustic implementation of tonal categories. A number of the observed errors reflect problems in executing tonal targets in real time. Pitch-range expansion and overshoot suggest that some speakers attempt to compensate for uncertainty by producing exaggerated excursions, but these excursions do not necessarily improve tonal accuracy. By contrast, dipping-contour truncation and duration reduction show the opposite tendency: speakers compress the contour so that the full tonal movement cannot unfold. Delayed rise onset in T2 is especially revealing in this respect, because it shows that the rising movement is not entirely absent, but begins too late to be perceived as fully target-like. These patterns indicate that tonal errors are not limited to choosing the wrong tone category; they also arise when speakers fail to align pitch movement, temporal support, and contour realization in a manner consistent with native production.

A further aspect of the nature of tonal error is that it is strongly context-sensitive. The data show that many speakers perform less accurately when tones appear in disyllabic strings, sandhi environments, or neutral-tone contexts than when they are produced in isolation. Citation-form persistence in sandhi contexts demonstrates that speakers often rely on the stored citation shape of a tone even when the phonological environment requires modification. Similarly, failure to reduce the neutral tone appropriately shows that speakers have difficulty moving from isolated tonal targets to context-governed prosodic realization. This suggests that the tonal problem is not confined to the level of individual citation tones. Rather, it extends to the ability to recalculate and adjust tonal realization under morphophonemic and prosodic pressure. The tonal categories may be partially available in isolation, but they become less stable when contextual processing is required.

Taken together, these findings indicate that the nature of tonal errors in this study is systematic rather than accidental. The errors cluster around three closely related dimensions: instability in tonal category contrast, weakness in phonetic implementation, and limited control of context-dependent tonal adjustment. For this reason, the tonal errors observed among learners of Chinese in Uganda should not be interpreted as

isolated slips of pronunciation. They are better understood as evidence of an interlanguage tonal system that is still developing and that remains particularly vulnerable in the dynamic T2 and T3 region and in prosodically conditioned environments. From the perspective of the PPM, these results suggest that production difficulty is closely tied to the speakers' incomplete control of perceptually salient pitch cues, contour complexity, temporal organization, and contextual tonal restructuring.

4.3.2 Distributional Patterns of Tone Production Errors

(1) Distribution across L1 backgrounds

Table 7 presents the first distributional view of tonal errors across L1 backgrounds by using a ΔF_0 -based acoustic similarity measure. Here, "L1 Group" identifies the participant groups by first-language background, while "Range-Similarity (%)" indicates how closely each group's monosyllabic pitch-range behavior approximates the native baseline. This column should not be read as categorical correctness; rather, it reflects the degree of acoustic proximity in pitch-range control. A higher value therefore suggests that the group is closer to the native target in overall scaling of tonal movement, whereas a lower value indicates a wider mismatch in the control of contrastive range.

The column "Predominant Deviation (from ΔF_0 profile)" shows where the largest mismatch is concentrated for each group. Its function is interpretive rather than merely descriptive: it identifies the tonal area in which the group departs most clearly from the native pattern. In this way, the table shows not only how much each group differs from the native baseline, but also what kind of range problem is most typical for that group. For example, T3 range overshoot in Lugisu, Lusoga, Runyankore, and Samia suggests that the main instability for these groups lies in the scaling of complex contour tones rather than in a uniform reduction across all tones. By contrast, T1 range mismatch in Luganda and T4 range compression in Lugbarati indicate different dominant loci of deviation, which means that L1-related error distribution is not homogeneous across the cohort.

Table 7: Monosyllabic Range-Similarity by L1 Background (Derived from ΔF_0)

L1 Group	Range-Similarity (%)	Predominant Deviation (from ΔF_0 profile)	Stability Index
Luganda	68.89	T1 range mismatch (largest mismatch)	0.689
Lugbarati	69.24	T4 range compression (largest mismatch)	0.692
Lugisu	31.54	T3 range overshoot (largest mismatch)	0.315
Lusoga	40.25	T3 range overshoot (largest mismatch)	0.402
Runyankore	71.83	T3 range overshoot (largest mismatch)	0.718
Samia	39.11	T3 range overshoot (largest mismatch)	0.391

Note. Range-Similarity (%) is a ΔF_0 -based measure of acoustic proximity to the native baseline. Stability Index expresses the same value on a 0–1 scale (Range-Similarity/100). This table is based on ΔF_0 only and does not represent categorical correctness from perceptual labeling.

The final column, “Stability Index,” restates the same acoustic similarity relationship on a 0–1 scale. Its value is not analytically separate from Range-Similarity (%), but it helps present the same pattern in a compact proportional form. Interpreted together, the three columns show that L1 background affects tonal production in two linked ways: first, by influencing how closely speakers approximate the native pitch-range profile overall, and second, by shaping which tonal category becomes the main site of deviation. Table 7 therefore does more than rank groups numerically; it shows that the distribution of tonal errors across L1 backgrounds is patterned both in degree and in dominant error locus.

These between-group differences provide the first distributional layer for Objective 2. They suggest that some L1 backgrounds are associated with closer approximation to the native pitch-range profile, whereas others show more persistent divergence in the realization of contour tones and in the stabilization of tonal contrast.

(2) Distribution across prosodic contexts

Prosodic context also conditions the distribution of tonal errors. For Ugandan L2 speakers of Chinese, tonal errors become more frequent and more structurally complex when tones are produced in longer or context-conditioned environments rather than in isolated citation forms. In other words, L2 speakers perform relatively better when they produce a single tone in isolation, but their tonal control becomes less stable when they must distribute pitch movement across two syllables or modify a citation tone according to sandhi and neutral-tone rules.

Table 8 compares isolated monosyllables with disyllabic sequences on a per-syllable basis. Relative to monosyllables, disyllables show reduced mean ΔF_0 range (64.99 Hz vs. 51.24 Hz) and shorter mean duration per syllable (543.16 ms vs. 455.10 ms), alongside a small reduction in the range-similarity index (48.92% vs. 46.45%). These values indicate that, for L2 speakers, prosodic expansion does not simply add another syllable; it also reduces the amount of pitch space and temporal support available for each syllable. As a result, tonal contrasts become harder to maintain clearly.

Table 8: Context Comparison (Per-syllable Normalized) between Monosyllabic and Disyllabic Items

Metric	Monosyllabic (Speakers)	Disyllabic (Speakers, per syllable)	Difference
Mean ΔF_0 (Hz)	64.99	51.24	-13.75
Native Mean ΔF_0 (Hz)	46.18	36.08	-10.1
Range-Similarity Index (%)	48.92	46.45	-2.47
Mean Duration (Ms)	543.16	455.10	-88.06
Native Mean Duration (Ms)	769.13	619.59	-149.54

Note. To make monosyllabic and disyllabic measures comparable, disyllabic ΔF_0 and duration are normalized per syllable by dividing by 2. Range-Similarity Index (%) is computed against the corresponding native baseline using the ΔF_0 -based comparison applied in this section.

This pattern can be explained from the L2 speakers' production burden. In a monosyllabic item, the speaker only needs to realize one tonal target within one syllabic frame. By contrast, in a disyllabic item, the speaker must preserve tonal identity while also coordinating two successive syllables within one lexical unit. This requires the L2 speaker to manage pitch movement, duration, and inter-syllabic timing simultaneously. When this burden increases, the contour often becomes compressed, and the tonal target is realized with less detail than in isolated production.

This contextual effect can be illustrated more concretely by the elicitation materials. In the monosyllabic list, items such as *hái* 'still/also', *qióng* 'poor', *quán* 'whole', *fó* 'buddha', *jiǔ* 'nine', *xuě* 'snow', *lěng* 'cold', and *shuǎng* 'cool' allowed L2 speakers to realize each tone within a single syllable, without having to coordinate tonal movement across adjacent syllables. In the disyllabic list, however, items such as *shíjiān* 'time',

yínháng ‘bank’, *cídiǎn* ‘dictionary’, *lánsè* ‘blue’, *shǒujī* ‘mobile phone’, *xǐnglái* ‘wake up’, *suǒyǐ* ‘therefore’, and *shǒutào* ‘glove’ required L2 speakers to maintain tonal contrasts across two consecutive syllables within one syllable. For L2 speakers whose tonal categories are not yet fully stabilized, this shift from one syllable to two syllables increases the likelihood of contour compression, shortened duration, and loss of fine tonal detail. Thus, the lower ΔF_0 range and shorter per-syllable duration in disyllabic items are not merely numerical differences; they reflect the practical difficulty L2 speakers face in preserving tonal accuracy when the prosodic frame becomes more complex.

A similar distributional effect appears in morphophonemic environments. In these contexts, L2 speakers are required not only to produce a tone, but also to replace a stored citation form with a context-conditioned surface form. Table 6 shows that T3 sandhi has the lowest rate of target-like realization (46.8%), while sandhi *yi* ‘one’ and *bu* ‘not’ show 61.3% and 58.7% respectively, and neutral tone shows 52.4%. These figures suggest that many Ugandan L2 speakers still rely more heavily on citation-form retrieval than on contextual tonal adjustment. Put differently, the difficulty is not simply producing T2, T3, or T4 as isolated categories; it is producing them in the form required by the surrounding phonological context.

This is especially clear in cases such as T3 sandhi, where the first T3 should be changed before another T3, and in the tonal behavior of *yi* ‘one’ and *bu* ‘not’, where the citation tone must be altered according to the following tone. For an L2 speakers, this means that successful production depends on two operations at once: recognizing the lexical tone and recalculating its context-appropriate realization. Neutral tone introduces a similar challenge, since the speakers must suppress full tonal prominence and produce a reduced syllable instead. When L2 speakers fail in these environments, the error is not just a tone-production error in the narrow sense; it is evidence that context-sensitive tonal control remains underdeveloped.

Chart 1: Error Types of Sandhi yi “one”

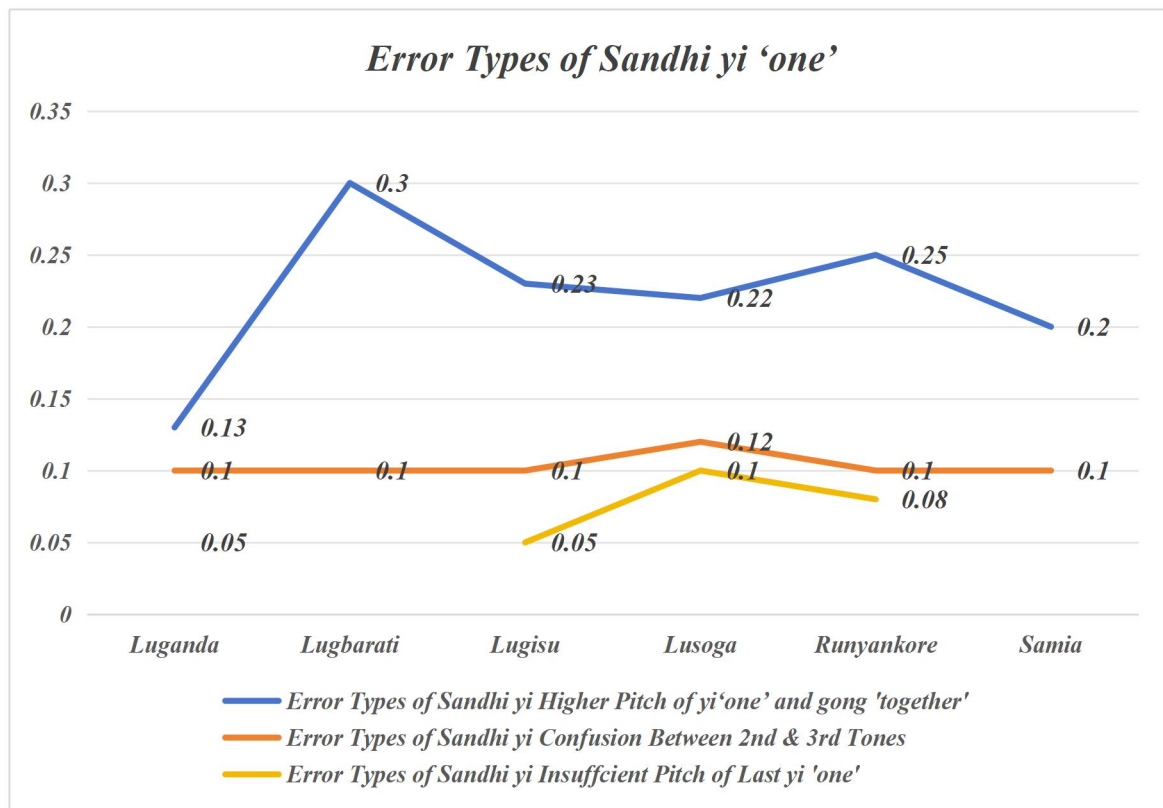


Chart 2: Error Types of Sandhi bu “not”

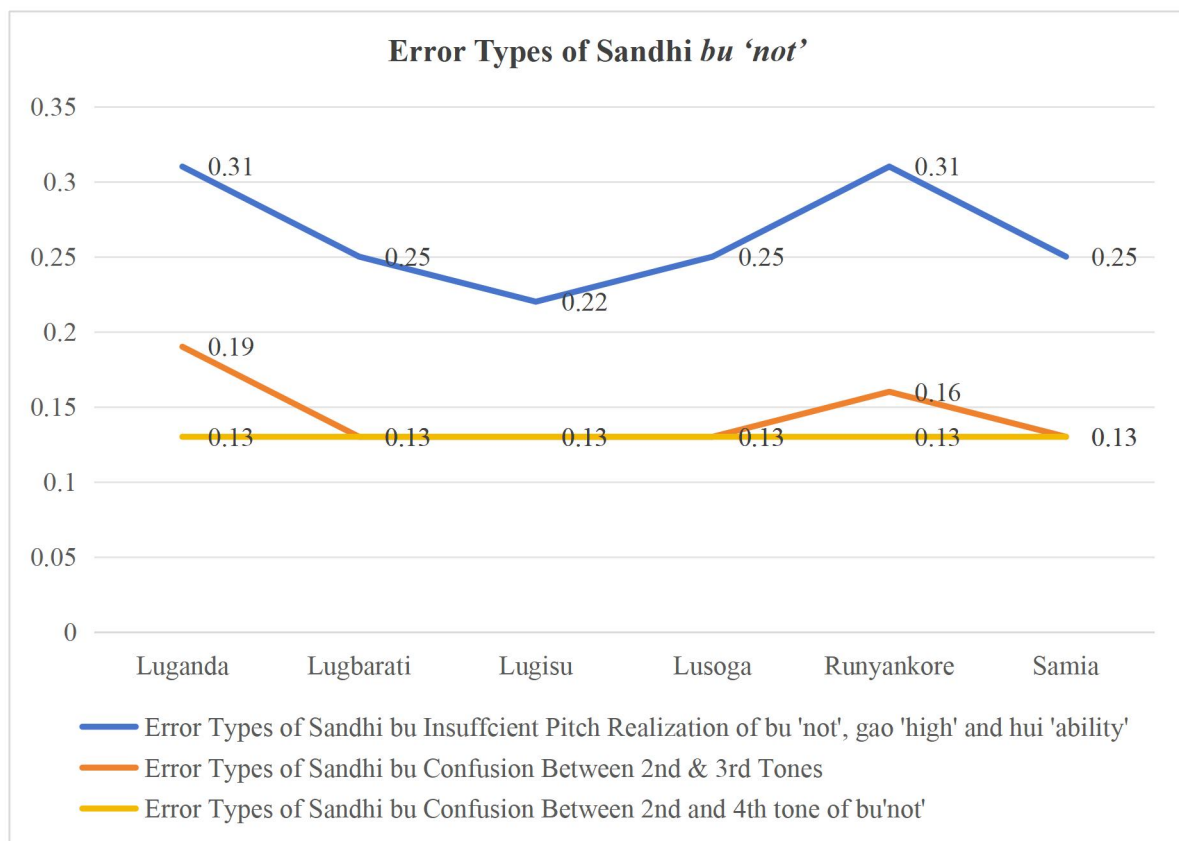


Chart 3: Error Types of Neutral Tone

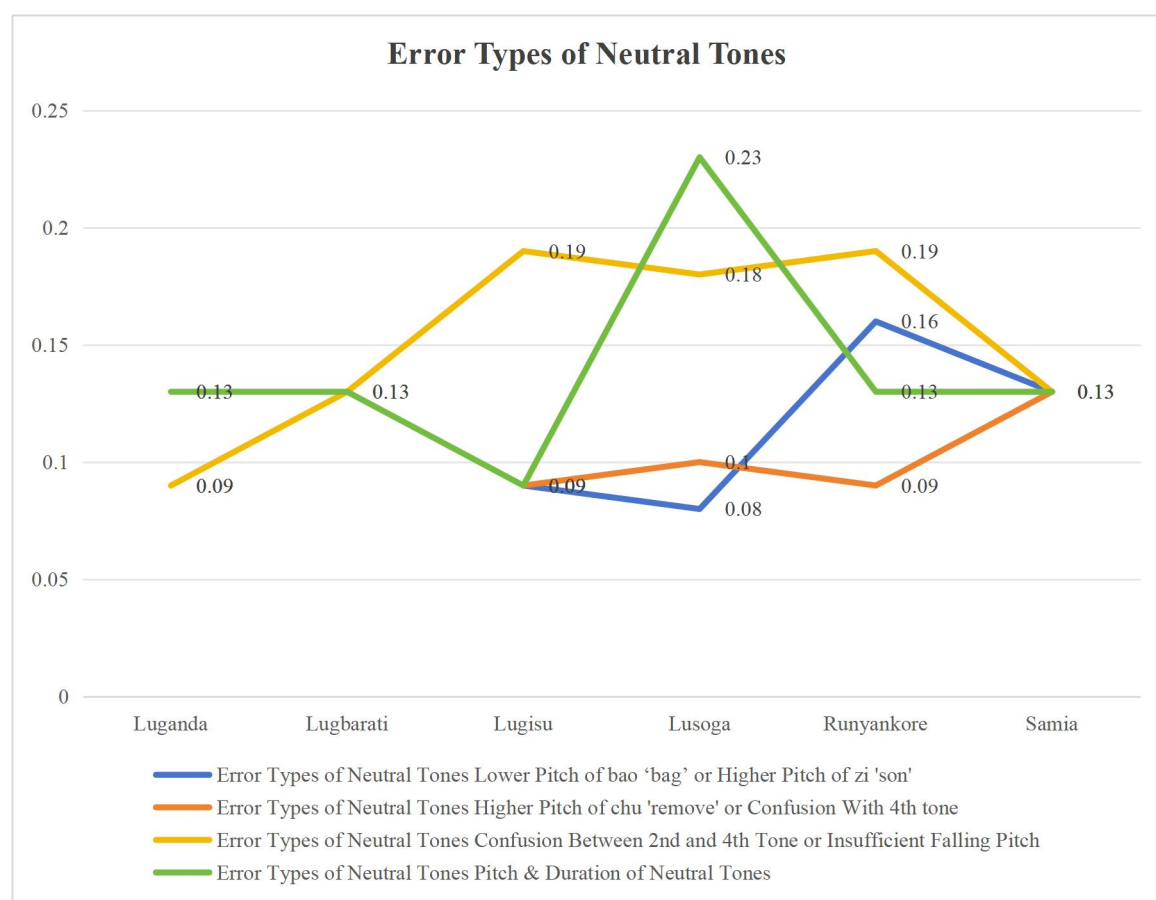


Chart 1 *Error Types of Sandhi yi 'one'*, Chart 2 *Error Types of Sandhi bu 'not'*, and Chart 3 *Error Types of Neutral Tone* illustrate this more specifically by showing the internal error-type composition in the three morphophonemic environments. Taken together, the charts show that once learners move beyond citation-form production, error distribution becomes more strongly shaped by context-sensitive alternation and reduction requirements.

Taken together, these results show that the distribution of tonal errors among Ugandan L2 speakers is strongly conditioned by prosodic environment. Errors are less concentrated in isolated citation forms and become more frequent when L2 speakers must coordinate tones across syllables or modify citation tones in accordance with contextual rules. The distribution of tonal error is therefore not uniform across the dataset: it shifts systematically toward disyllabic compression, sandhi failure, and neutral-tone reduction failure as prosodic demands increase.

(3) Distribution across tone categories and error mechanisms

The distribution of tonal errors is also uneven across the tonal inventory: for Ugandan L2 speakers, the heaviest error burden falls on T2 and T3, whereas T1 and T4 are more likely to retain their broad category identity but deviate in phonetic implementation. This means that L2 speakers do not mispronounce all tones in the same way. Instead, different tones attract different kinds of errors, and this reveals where the L2 tonal system is most unstable.

Table 9 makes this pattern clear. T2 is associated primarily with phonological error, especially categorical confusion (49.0%), while T3 is also dominated by a phonological problem, namely dipping truncation (41.0%). By contrast, T1 is linked mainly to a phonetic problem of register migration (11.0%), and Tone 4 more often shows phonetic range compression (23.0%). From the perspective of L2 use, this means that Ugandan L2 speakers experience their greatest difficulty not with every tone equally, but with those tones that require more delicate control of dynamic contour shape and category contrast.

Table 9 provides a compact summary of the distribution of error types by tone and mechanism.

Table 9: Distribution of Errors by Nature and Mechanism

Tone	Error Nature	Primary Mechanism	Distributional Frequency
Tone 1	Phonetic	Register Migration	11.0%
Tone 2	Phonological	Categorical Confusion	49.0%
Tone 3	Phonological	Dipping Truncation	41.0%
Tone 4	Phonetic	Range Compression	23.0%

For L2 speakers, T2 and T3 form the most unstable region of the system because both

are dynamic tones and both depend on successful control of contour movement over time. T2 requires a sufficiently clear and timely rise, whereas T3 requires a more complex dipping contour. When L2 speakers do not control timing and contour structure reliably, the two tones become especially vulnerable to overlap. This is why the T2 and T3 confusion emerges as the main site of phonological instability in the dataset. The problem here is category-level: L2 speakers are not only producing non-native-like contours, but are also failing to keep the two tonal categories fully separate. By contrast, T1 and T4 show a different distributional pattern. These tones are less often replaced by another tonal category, but they are still produced in non-target ways. For L2 speakers, T1 is more likely to drift in overall register, while T4 may preserve its falling identity but lose target-like scaling through range compression. In these cases, the speaker appears to retain the intended tone category, but the acoustic realization is still imperfect. The problem is therefore primarily phonetic rather than phonological. This distinction is important because it shows that some L2 tonal errors arise from unstable category selection, whereas others arise from incomplete control over the phonetic implementation of an already selected category.

From the standpoint of error mechanism, the distribution is therefore not only tone-specific but also mechanism-specific. For Ugandan L2 speakers, cross-category confusion clusters most strongly in T2 and T3, while implementation problems such as register displacement, range mismatch, or contextual compression are more visible in T1 and T4. In addition, context-sensitive mechanisms such as citation-form persistence and neutral-tone reduction failure become especially salient in morphophonemic environments. This means that the L2 tonal system is not characterized by one single type of weakness. Rather, different parts of the tonal inventory break down in different ways depending on the tone involved and the type of control required.

The descriptive implication is that tonal errors among Ugandan L2 speakers are distributed selectively rather than evenly. T2 and T3 are the main locus of phonological instability, while T1 and T4 more often show phonetic distortion within the intended category. This uneven distribution helps explain why some tones sound categorically wrong and others sound recognizably correct but still non-native-like. It also reinforces

the broader claim of this section: the tonal errors produced by Ugandan L2 speakers are systematically patterned across tones and mechanisms, rather than randomly scattered across the corpus.

Taken together, the distribution analysis shows that the identified error types are not randomly scattered across the dataset. They recur in structured ways across L1 backgrounds, prosodic conditions, and tone-specific mechanisms.

This interim summary completes Objective 2: the chapter has clarified what kinds of errors are involved and how those errors are distributed, providing the basis for comparison with previous studies.

4.4 Comparison with Previous Studies

This section compares the present findings with the relevant studies reviewed in Chapter Two. The aim is not to re-open the literature review, but to show more explicitly where the Ugandan data converge with earlier work on Chinese tone acquisition and where they contribute a different perspective.

First, the present chapter is broadly consistent with previous studies in showing that contour tones are less stable than the high-level tone. The current results likewise confirm that the main bottleneck lies in the T2 and T3 region, where rising and dipping categories are most likely to converge. In this respect, the Ugandan data fit the wider pattern reported in earlier L2 Chinese tone research, especially studies that identify contour timing and turning-point control as persistent sources of difficulty.

At the same time, the present dataset also shows a more specific pattern that deserves emphasis: T2 emerges as the most vulnerable category in the observed hierarchy ($T2 > T3 > T4 > T1$). This does not fully erase the difficulty of T3, but it suggests that, for the participants in this study, the sustained and timely realization of a rising contour may be at least as fragile as the execution of a dipping contour. This point adds nuance to earlier work that often foregrounds T3 alone as the main problem category.

Table 10: Cross-Speaker Tone 1 Register Anchoring and Migration

Participant Group	Target	Mean F₀ (Hz)	Baseline (127.8 Hz)	Deviation (Hz)	L1 Influence
Native Speaker	Tone 1	127.8	127.8	0	Native
Luganda Group	Tone 1	176.1	127.8	48.3	Elevated
Lugbarati	Tone 1	121.3	127.8	-6.5	Near-native
Lugisu Group	Tone 1	211.2	127.8	83.4	Highly elevated
Lusoga Group	Tone 1	166.4	127.8	38.6	Elevated
Runyankore Group	Tone 1	170.5	127.8	42.7	Elevated
Samia	Tone 1	196.9	127.8	69.1	Highly elevated

Second, the present findings are in line with previous studies that emphasize the importance of L1 background in shaping tonal production. However, the Ugandan data extend that discussion by bringing together participants from several East African L1 backgrounds within a single acoustic framework. The cross-group pattern summarized in Table 10 shows that pitch-register anchoring does not operate uniformly across groups, which helps explain why some participants remain closer to the native baseline while others display stronger deviation in contour scaling.

Third, the contextual findings also converge with earlier research. Previous studies have often reported that connected speech, sandhi environments, and neutral-tone contexts are more challenging than isolated citation forms. The present chapter confirms that pattern: disyllabic items show reduced range and greater timing pressure, while T3 sandhi, *yi* ‘one’, *bu* ‘not’, and neutral-tone contexts all remain comparatively unstable. A further contribution of the present study lies in the way these recurrent patterns are interpreted. Rather than treating the chapter as a catalogue of isolated errors or as an OT-based ranking exercise, the current version maintains a single PPM framework and shows how pitch-range control, cue instability, temporal pressure, and contextual adjustment jointly shape tonal production. This makes the comparison with previous studies more coherent with the overall theoretical design of the thesis.

More specifically, the Ugandan dataset contributes evidence from an underrepresented

linguistic setting in the Chinese tone-acquisition literature. The inclusion of both Bantu and Central Sudanic backgrounds, together with the combination of monosyllabic, disyllabic, sandhi, and neutral-tone environments, provides a broader view of how tonal difficulty is distributed across categories and contexts than is often available in single-group learner studies.

Overall, the present findings both align with and extend previous research. They align with earlier work in showing that contour tones, contextual alternations, and connected speech remain major sources of difficulty; they extend that work by demonstrating how these difficulties are distributed across Ugandan L1 backgrounds and by interpreting them within a unified PPM account that remains consistent with the wider argument of the study.

4.5 Summary

This chapter has shown that tonal errors in the Ugandan data are systematic rather than incidental. Objective 1 identified seven recurrent errors in the participants' Chinese production, while Objective 2 clarified how those errors are distributed across L1 backgrounds, prosodic contexts, tone categories, and error mechanisms. These findings establish the empirical basis for the conclusions and implications developed in Chapter Five.

CHAPTER FIVE: SUMMARY OF FINDINGS, CONCLUSIONS, AND RECOMMENDATIONS

5.1 Summary of Findings

This dissertation investigates tone production errors in Chinese Chinese among Ugandan L2 speakers and interprets these errors within the framework of the Pitch Perception Model (PPM). The study focuses on how Ugandan speakers produce Chinese tones across three major domains: monosyllabic citation tones, disyllabic tone combinations, and context-conditioned tonal environments, especially Tone 3 (T3) sandhi, the tonal alternations of *yi* ‘one’ and *bu* ‘not’, and neutral tone. By examining tonal performance across these different structures, the study seeks to provide a systematic account of what kinds of tonal errors Ugandan speakers produce, how these errors are distributed.

The study was guided by two specific objectives. The first objective was to identify the tone production errors made by Ugandan L2 speakers of Chinese. The second objective was to describe the distribution of these tonal errors across tone categories, participant backgrounds, and prosodic contexts. The study therefore moves from identifying recurrent non-target patterns to describing where and how those patterns occur in the production data.

Methodologically, the study adopted an experimental phonetic approach combining controlled speech elicitation with acoustic analysis. The participants were intermediate learners of Chinese in Uganda drawn from different first-language backgrounds, and their productions were compared with a native Chinese baseline. The dataset was collected through structured production tasks covering isolated monosyllables, disyllabic items, and selected morphophonemic environments. The recorded materials were then analyzed acoustically, with particular attention to F_0 , contour shape, pitch-range excursion, duration, and alignment. These measures made it possible to evaluate tonal accuracy not only impressionistically, but also in terms of objective phonetic evidence. Guided by PPM, the analysis treated tone production as dependent on the

speakers' ability to perceive, organize, and implement linguistically relevant pitch cues in a stable and context-sensitive manner.

On this basis, the following sections summarize the major findings of the study in relation to tonal errors and their distribution.

5.1.1 Tone Production Errors

Although Ugandan L2 speakers of Chinese are able to pronounce approximate trajectories of tonal contours, they have challenges in achieving stable, target-like precision expected of native production, which are hereby outlined below as tone production errors.

The major findings may be stated more explicitly as four linked results: a tone-specific hierarchy of stability, a concentrated Tone 2 (T2)–Tone 3 (T3) confusion, systematic variation across L1 backgrounds and prosodic contexts, and a speaker tonal system that is emerging but not yet fully consolidated.

(1) Tone-specific hierarchy is clear

Tone 1 (T1) is the most stable category, Tone 4 (T4) is relatively stable but not fully target-like, and the greatest vulnerability lies in T2 and T3. This matters because it shows that difficulty is concentrated in dynamic contour categories rather than evenly distributed across the tonal inventory.

T2 (Rising) > T3 (Dipping) > T4 (Falling) > T1 (High-Level)

The most robust overall pattern remains the hierarchy $T2 > T3 > T4 > T1$. This hierarchy is important because it reflects the present dataset rather than a generalized assumption imported from earlier CSL literature. It suggests that the demands of Chinese contour tones are not uniform for this cohort: the greatest instability was concentrated in the rising and low-rising region, whereas the high-level tone remained comparatively stable.

(2) T2 and T3 Confusion

The recurrent T2 and T3 confusion is especially significant because it points to difficulty not merely in executing a contour, but in maintaining distinct perceptual–productive mappings for two neighboring tonal categories. The most persistent errors arise where learners must distinguish an early, sustained rise from a contour involving

a low target plus recovery movement. In PPM terms, the contrast becomes unstable when rise onset, low turning-point attainment, and recovery timing are not weighted in a sufficiently differentiated way.

(3) Tonal errors are distributional rather than random

Performance varied across L1 backgrounds. Using the acoustic similarity values reported in Chapter Four (Table 7), Runyankore (71.83), Lugbarati (69.24), and Luganda (68.89) pattern closer to the native baseline, whereas Lusoga (40.25), Samia (39.11), and Lugisu (31.54) show lower similarity. These differences suggest that L1 influence was not uniform across the cohort and that speakers entered Chinese tone learning with different cue-weighting tendencies for pitch range, contour direction, and register control.

They vary across L1 backgrounds and become more pronounced in disyllabic, sandhi, and neutral-tone contexts. The Ugandan data therefore support a layered interpretation in which tone-specific vulnerability interacts with L1-related cue bias and with contextual pressure from connected speech.

(4) The speaker system is partially stabilized but not yet consolidated

The confusion matrix points to an emerging tonal system in which categories exist and are often retrievable, but remain permeable under pressure. The production evidence therefore supports an interpretation of structured interlanguage instability rather than chaotic inaccuracy.

The broader implication is that greater instability accompanies greater contour complexity and greater dependence on syllable-internal timing. Tones whose identity can be preserved through a relatively stable level target or a clear falling direction remain more accessible, whereas tones that require a rise, a turn, or tightly coordinated cue integration are more likely to drift toward non-target realizations.

5.1.2 Distribution of Tone Production Errors

The central finding of this section is that tonal errors among Ugandan L2 speakers of Chinese are not randomly dispersed across speakers, tones, or contexts. Rather, they are distributed in a systematic and interpretable way: they vary across L1 backgrounds,

increase in prosodically and morphophonemically demanding environments, cluster most heavily in the T2 and T3 region, and manifest through different error mechanisms depending on the tone involved. This distributional pattern shows that the tonal difficulties observed in the present study are best understood as structured properties of an evolving L2 tonal system rather than as isolated pronunciation failures.

(1) The distribution of tonal errors was uneven across L1 backgrounds among Ugandan L2 speakers

The tonal performance of Ugandan L2 speakers did not pattern uniformly across the cohort. Instead, speakers from different L1 backgrounds showed different degrees of proximity to the native Chinese baseline, particularly in pitch-range control and overall contour realization. This indicates that the distribution of tonal errors was partly conditioned by speakers' prior phonological experience. In other words, Ugandan L2 speakers did not enter Chinese tone acquisition with identical perceptual and production resources, and the resulting error patterns reflected this variation. The significance of this finding is that tonal difficulty in the present cohort cannot be treated as a single undifferentiated L2 problem; rather, it is mediated by the phonological background from which each group of L2 speakers approaches Chinese tone.

(2) For Ugandan L2 speakers, tonal errors increased noticeably when tones were produced in prosodically and morphophonemically demanding environments rather than in isolated citation forms

A major distributional finding of this study is that Ugandan L2 speakers were relatively more stable in isolated monosyllabic production, but became less accurate when tones had to be realized in disyllabic sequences, sandhi environments, and neutral-tone contexts. This means that tonal difficulty for L2 speakers was not constant across speaking conditions. When the prosodic frame became more complex, L2 speakers had to manage pitch movement, duration, inter-syllabic coordination, and contextual tonal adjustment simultaneously, and tonal accuracy decreased accordingly. The same tendency appeared in morphophonemic environments, where many L2 speakers continued to rely on citation-form retrieval instead of applying the contextually required tonal modification. The distribution of errors therefore shifted systematically toward

contexts that imposed greater prosodic and processing pressure on L2 speakers.

(3) The heaviest error burden for Ugandan L2 speakers clustered in the T2 and T3 region, while Tone 1 and Tone 4 more often showed phonetic deviation within a broadly retained category

The distribution of tonal errors across the tonal inventory was clearly uneven. For Ugandan L2 speakers, T2 and T3 constituted the most unstable part of the system, and the greatest concentration of error occurred in this dynamic contrastive region. These two tones require fine control of contour direction, temporal unfolding, and pitch alignment, and many L2 speakers did not yet maintain them as consistently distinct categories in production. By contrast, T1 and T4 were less often replaced by another tonal category, but were still affected by non-target-like phonetic realization, such as register displacement or range compression. This distinction is important because it shows that different tones were vulnerable in different ways: some were prone to category instability, while others were more likely to remain recognizably intact but phonetically non-native-like.

(4) The distribution of tonal errors was also mechanism-specific, showing that different parts of the L2 tonal system broke down in different ways

The errors produced by Ugandan L2 speakers did not arise from one single source. Instead, they were distributed across several mechanisms, including category confusion, contour truncation, delayed alignment, range distortion, citation-form persistence, and failure of neutral-tone reduction. These mechanisms did not contribute equally to all tones or all contexts. Rather, they patterned selectively according to the tonal category and the production demands placed on the L2 speakers. This means that the distribution of tonal errors in the present study was not only tone-specific and context-specific, but also mechanism-specific.

Taken together, these findings confirm that the L2 tonal system of Ugandan speakers remains only partially stabilized: some errors reflect weak category separation, some reflect incomplete phonetic implementation, and others reflect limited control of context-sensitive tonal restructuring.

5.2 Limitations of the Study

5.2.1 Modest and Institutionally Bounded Participant Pool

The study involved 19 Ugandan L2 speakers of Chinese drawn from one intact Chinese class at a single institution. This provided comparable learning and recording conditions, but the participant pool was modest, the six L1 groups were unevenly represented, and the institutional setting was narrow. The findings should therefore be interpreted as tendencies in this specific cohort rather than as a representative account of all Ugandan learners or of each language group.

5.2.2 Production-Centered and Controlled Evidence

The evidence is production-centered and was collected through a controlled word-list task, with one native Chinese speaker serving as the acoustic baseline. This design supports consistent comparison of the recorded tokens, but it does not independently test perception, represent the full range of native variation, reproduce spontaneous interaction, or trace change over time. PPM-based explanations should therefore be read as cautious interpretations of cross-sectional production evidence rather than as direct proof of perceptual processes or longitudinal development.

5.3 Original Contribution to Knowledge

The contribution of the study lies not in proposing an entirely new theory of Chinese tone acquisition, but in bringing together a specific dataset, a coherent production-centered analytic route, and a single PPM-based interpretive frame for a context that remains underrepresented in the literature.

Empirically, the study provides a production-centered account of Chinese tone errors among Ugandan participants and shows that the major instability cluster lies in the T2 and T3 region rather than in T3 alone. Theoretically, it demonstrates that the findings can be synthesized under PPM without fragmenting the discussion across competing explanatory models. Methodologically, it shows the value of keeping acoustic evidence, distributional description, and chapter-level synthesis distinct but mutually supportive.

5.4 Conclusions

The conclusions of the chapter follow directly from the major findings rather than from any additional analytical framework. They can be grouped into overall, theoretical, and pedagogical conclusions.

Overall, the study shows that Ugandan participants had developed an emerging but only partially stabilized Chinese tonal system. The error patterns were not random: they clustered around contour tones, especially in the T2 and T3 region, and became more fragile under greater contextual and temporal pressure.

At the theoretical level, the findings are most coherent when interpreted within a single PPM framework. The strongest explanatory dimensions are perceptual salience, contour complexity, pitch-register alignment, temporal support, and L1-related cue bias. This integrated reading is more consistent with the present work.

5.5 Recommendations

The findings suggest that instructional support should concentrate on the most fragile contrast region, on contour realization under reduced temporal support, and on tonal control in connected or morphophonemically conditioned contexts. The key implication is specificity: not all tones are equally problematic, and not all difficulties arise from the same source.

The patterns of difficulty observed above carry actionable implications for second language tonal teaching. The following two recommendations target the most persistent challenges observed in the speakers tone production data.

First, future teaching practice for Ugandan or comparable L2 speaker groups should prioritize contrastive work on the T2 and T3 region. This includes guided practice on rise onset, turning-point location, low target attainment, and the maintenance of sufficient temporal space for contour realization.

Second, tonal instruction should move progressively from isolated syllables to disyllabic combinations and short connected sequences so that speakers can maintain tonal identity under greater contextual pressure, including sandhi and neutral-tone environments.

5.6 Pointers for Future Research

Future research could strengthen the empirical base by using a larger and more evenly distributed participant pool drawn from multiple Chinese-learning institutions and L1 backgrounds in Uganda. This would allow more confident comparison of group-level tendencies and reduce the interpretive asymmetry caused by very small subgroup sizes. Further work could also examine tone production and perception in a more explicitly integrated design, provided that both components are implemented and reported with equal methodological clarity. Such research would help clarify how perceptual cue instability develops into recurrent production error patterns in African Chinese-learning contexts.

Finally, future studies may broaden the evidence by including several native Chinese reference speakers, more connected and spontaneous speech materials, systematically balanced sandhi environments, and longitudinal observation of tonal stabilization. These directions would extend the present controlled, cross-sectional production account without obscuring its current limitations.

REFERENCES

- Andersen, T. (1986). Tone splitting and vowel quality: Evidence from Lugbara. *Studies in African Linguistics*, 17(1), 55 – 68.
- Celce-Murcia, M. Brinton, D.M, & Goodwin, J.M. (2010). *Teaching pronunciation: A course book and reference guide*. Cambridge University Press.
- Chao, Y. R. (1933). Tone and intonation in Chinese. *Bulletin of the Institute of History and Philology, Academia Sinica*, 4(2), 121–134.
- Chao, Y. R. (1948). *Mandarin primer: An intensive course in spoken Chinese*. Harvard University Press.
- Chen, L., Liu, H., & Kager, R. (2017). Categorical perception of pitch direction in Mandarin and English listeners. *Journal of Phonetics*, 62, 1-12.
- Chen, Y. (2011). Tone category development in L2 Mandarin learners with non-tonal first languages. *Applied Psycholinguistics*, 32(4), 765-792.
- Chomsky, N., & Halle, M. (1968). *The sound pattern of English*. Harper & Row.
- Clumeck, H. (1980). The acquisition of tone. In J. H. Greenberg (Ed.), *Universals of Human Language: Volume. 2, Phonology* (pp. 257-275). Stanford University Press.
- Darya, L. (2020). *Analysis of Chinese intonation errors of advanced Russian students*. Shannxi Normal University.
- Clumeck, H. (1980). The acquisition of tone. In J. Greenberg (Ed.), *Universals of Human Language: Volume 4, Phonology* (pp.251-276). Standard University Press.
- de Cheveigné, A. (1986). A pitch perception model. In 1986 *IEEE International Conference on Acoustics, Speech, and Signal Processing (ICASSP)* (pp. 897 – 900). IEEE.
- Diao, X. R. (2020). *Analysis and teaching strategies on errors in acquiring Chinese intonation among Thai speakers*. Heilongjiang University.
- Dimmendaal, G. J., Ahland, C., Jakobi, A., & Kutsch Lojenga, C. (2019). Linguistic features and typologies in languages commonly referred to as ‘Nilo-Saharan’. In H. E. Wolff (Ed.), *The Cambridge Handbook of African Linguistics* (pp. 326 – 381). Cambridge University Press.

- Feng, X. (2016). *Experimental study on the intonation patterns of Thai international students' Chinese acquisition of two- and three-character groups*. Southwest University.
- Flege, J. E. (1995). Second language speech learning: Theory, findings, and problems. In W. Strange (Ed.), *Speech perception and linguistic experience: Issues in cross-language research* (pp. 233–277). York Press.
- Flege, J. E. (2003). Assessing constraints on second-language segmental production and perception. In N. O. Schiller & A. S. Meyer (Eds.), *Phonetics and phonology in language comprehension and production: Differences and similarities* (pp.319-358). De Gruyter Mouton.
- Fu, C. B. (2011). *Putonghua jiaocheng* [A course in putonghua]. Beijing University of Posts and Telecommunications Press.
- Gai, Q. Y. (2014). A preliminary study on the tone acquisition of Chinese monosyllabic words by African students. *Journal of Chengdu Institute of Information Engineering*. China Academic Journal Electronic Publishing House.
- Gandour, J. T. (1983). Tone perception in far eastern languages. *Journal of Phonetics*, 11, 149–175.
- Gao, X., Yan, T. T, Tang, D. L, Huang, T., Shu, H., Nan, Y., & Zhang, Y. X. (2019). What makes lexical tone special: A reverse accessing model for tonal speech perception. *Frontiers in Psychology*, 10, 2830.
- Goldsmith, J.O.H.N. (1976). *Autosegmental phonology*. Garland Press.
- Huang, Y., & Liu, Y. (2023). Mandarin tone acquisition in Vietnamese learners: Error patterns across perception, production, and memory. *Journal of Second Language Acquisition*, 45(3),789-812.
- Hyman, L. M., & Katamba, F. X. (1993). A new approach to tone in Luganda. *Language*, 69(1), 34 – 67.
- Hyman, L. M., & Katamba, F. X. (2010). Tone, syntax, and prosodic domains in Luganda. *ZAS Papers in Linguistics*, 53, 69 – 98.
- Hyman, L. M. (2018). The autosegmental approach to tone in Lusoga. In D. Brentari &

- J. L. Lee (Eds.), *Shaping phonology* (pp. 47 – 69). University of Chicago Press.
- Hyman, L. M. (2022). Tone in runyankore verb stem reduplication. *Stellenbosch Papers in Linguistics Plus*, 62, 1 – 24.
- Iuliia, T. (2020). *Research on the problems and countermeasures existing in Chinese pronunciation learning for Russian students*. Shaanxi Normal University.
- Jin, Y. Z. (2011). *The study on learning Chinese pronunciation of Korean students*. Northeast Normal University.
- Lafontaine, A. (2017). Negative transfer of French stress and intonation on L2 Mandarin tone production. *Canadian Modern Language Review*, 73(3).345-368.
- Levis, J. M., & Moyer, A. (Eds.). (2014). *Social dynamics in second language accent*. De Gruyter Mouton.
- Li, B. H.(1988). Comparison of tones between Chinese and Hausa and teaching of tones. *Journal of World Chinese Teaching*, 3.
- Li, C. N., & Thompson, S. A. (1977). The acquisition of tone in Mandarin-speaking children. *Journal of Child Language*, 4 (2), 185-199.
- Li, H. Y. (1995). Phonetic errors in Thai students learning Chinese. *Journal of World Chinese Teaching*, 2, 66-71.
- Li, Y. F. (2020). *Problems and teaching strategies in Chinese phonetics teaching in Gongzhou Girls' High School of South Korea*. Zhengzhou University.
- Licklider, J. C. R. (1951). A duplex theory of pitch perception. *Experientia*, 7 (4), 128-134.
- Lin, H. (1996). *Mandarin tonology*. Pyramid Press.
- Liu, X. (2000). *Introduction to teaching Chinese as a foreign language*. Beijing Language and Culture University Press.
- Lu, Y., Liang, C., & Kong, J.(2025). Contribution of F₀ and phonation to tone perception in the Zaiwa language. *Journal of Phonetics*, 110, 101413. <https://doi.org/10.1016/j.wocn.2025.101413>.
- Lugbarati Language Board. (2013). *Lugbarati orthography guide*. National Curriculum Development Centre.
- Makerere University. (2014, December 21). Mak Confucius Institute to bridge Uganda-

- China cultural gap. *Makerere University News*.
<https://news.mak.ac.ug/2014/12/mak-confucius-institute-to-bridge-uganda-china-cultural-gap/>
- Makerere University. (2024, November 28). Mak celebrates 10 years of the Confucius Institute with launch of three associations. *Makerere University News*.
<https://news.mak.ac.ug/2024/11/mak-celebrates-10-years-of-the-confucius-institute-with-launch-of-three-associations/>
- Makerere University College of Humanities and Social Sciences. (n.d.). *Bachelor of Chinese and Asian Studies*. Makerere University. Retrieved July 1, 2026, from <https://chuss.mak.ac.ug/en/course/bachelor-of-chinese-and-asian-studies/>
- Marlo, M. R. (2013). Verb tone in Bantu languages: Micro-typological patterns and research methods. *Africana Linguistica*, 19, 137 – 234.
- McCarthy, J. J., & Prince, A. (2001). Prosodic morphology. In A. Spencer & A. M. Zwicky (Eds.), *The handbook of morphology* (pp. 283–305). Blackwell.
- Meddis, R., & Hewitt, M. J. (1991). Virtual pitch and phase sensitivity of a computer model of the auditory periphery. I: Pitch identification. *The Journal of the Acoustical Society of America*, 89(6), 2866-2882.
- Ministry of Education & State Language Commission of China. (2021). *Chinese proficiency grading standards for international Chinese language education* (GF 00252021). The Commercial Press.
- Mitkov, A. (2021). *Using machine learning in an experimental study on the Chinese tones acquisition by Russian students*. Nankai University.
- Moisik, S. R., Lin, H., & Esling, J. H. (2014). A study of laryngeal gestures in Mandarin citation tones using simultaneous laryngoscopy and laryngeal ultrasound (SLLUS). *Journal of the International Phonetics Association*, 44(1), 21-58.
- Namulondo, S. (2014). *Analysis of errors in Chinese pronunciation of language speaking students and suggested teaching strategies*. Shandong University.
- Ni, Y. & Wang, X. K. (1992). Analysis of difficulties in learning Chinese pronunciation for students from English-speaking countries. *Journal of Nankai University*, 3, 72-76.

- National Planning Authority. (2013). *Uganda vision 2040*. National Planning Authority.
<https://npa.go.ug/uganda-vision-2040/>
- Ortega-Llebaria, M., Nemogá, M., & Presson, N. (2015). Long-term experience with a tonal language shapes the perception of intonation in English words: How Chinese-English bilinguals perceive “Rose?” vs “Rose”. *Bilingualism: Language and Cognition*, 18 (1), 56-72.
- Pike, K. L. (1948). *Tone languages*. University of Michigan Press.
- Prince, A., & Smolensky, P. (1993). *Optimality theory: Constraint interaction in generative grammar*. Blackwell.
- Shao, J. M. (2016). *General introduction to modern Chinese* (3rd ed.). Shanghai Education Press.
- Shi, H. (2018). *A method of teaching English-speaking speakers to produce Mandarin Chinese tones*. West Virginia University.
- Song, C. Y. (1998). On pronunciation teaching for Korean students. *Journal of Nankai University*. 3, 72-76.
- Strange, W. (2011). Automatic selective perception (ASP) of first and second language speech: A working model. *Journal of Phonetics*, 39(4), 456-466.
- Sylviano, R. A. A. (2016). *Malagasy students studying Chinese rising tone modulation error analysis*. Jiangxi Normal University.
- Tong, Y., Francis, A. L., & Gandour, J. T. (2008). Processing dependencies between segmental and suprasegmental features in Mandarin Chinese. *Language and Cognitive Processes*, 23 (5), 689-708.
- Wang, M. & Sun, F. B. (2003). An exploration of Russian-Chinese pronunciation and intonation teaching. *Journal of Heilongjiang Institute of Education*, 5, 59-60, 62.
- Wang, S.Y. (1967). Phonological features of tone. *International Journal of America Linguistics*, 33(2).
- Wang, Y. J. (2006). Second language learners’ perceptual acquisition of Mandarin lexical tones. *In Proceedings of Speech Prosody 2006* (pp.687-690). ISCA.
- Wang, Y. J., & Li, M. J. (2010). The effects of tone contour and register on the perception of Mandarin Tone 2 and Tone 3. *Acta Psychologica Sinica*, 42(9), 899-

908.

- Xu, H. (2013). *Experimental analysis of Thai students' acquisition of Chinese double-character tones*. Tianjin Normal University.
- Yang, J. (2016). Contextual effects on L2 Mandarin tone error patterns. *Studies in Second Language Acquisition*, 38 (2), 321-348.
- Ye, Y., & Connine, C. M. (1999). Processing spoken Chinese: The role of tone information. *Language and Cognitive Processes*, 14(5-6), 609-630.
- Yip, M. J. W. (2002). *Tone*. Cambridge University Press.
- Yost, W. A. (2009). Pitch perception. *Attention, Perception & Psychophysics*, 71(8), 1701-1715. <https://doi.org/10.3758/APP.71.8.1701>.
- Yue-Hashimoto, A. O. (1980). *Chinese tone: Its linguistic function and its acquisition*. University of California Press.
- Zhan, D. D. (2017). *A study of nigerian Chinese language speakers' disyllabic tone acquisition*. Beijing Institute of Technology.
- Zhang, D. (2011). *Chinese pronunciation errors and teaching strategies of Russian students studying abroad*. Hei Longjiang University.
- Zhang, L. (2020). Mandarin tone learning difficulties in Congolese French speakers: Tone pattern vs. Tone value. *Language Teaching Research*, 24(5), 689-708.
- Zhang, T. & Peng, Q. (2021). A comparative study of tones in Chinese and Thai and suggestions for Chinese tone teaching in Thailand. *Journal of Lanzhou Jiaotong University*, 40 (4).
- Zhang, Z. Y. (2008). *An experimental study on the disyllabic Chinese tonal errors of the junior Thai students*. Jinan University.
- Zhao, J. P. (2020). *A study on Chinese tonal errors of primary level Egyptian students and suggested countermeasures*. Jilin University
- Zhou, H. (2013). *Analysis and countermeasures on Chinese intonation errors of Ghanaian international students*. Liaoning Normal University.
- Zhou, H., & Verissimo, M. (2025). L2 Mandarin tone errors: Category establishment vs. phonetic implementation. *Journal of Multilingual and Multicultural Development*, 46(1), 45-62.

Zhu, H. (2018). *A method of teaching English-speaking speakers to produce Mandarin Chinese tones*. Beijing Language and Culture University.

APPENDICES

APPENDIX I:

FIELD ACCEPTANCE LETTER FROM THE DEPARTMENT

MAKERERE

P.O. Box 7062, Kampala, Uganda
Upper Building, Room 11



UNIVERSITY

Telephone: +256 414 530 106
E-mail: dlec.chuss@mak.ac.ug

**DEPARTMENT OF LINGUISTICS, ENGLISH LANGUAGE STUDIES
AND COMMUNICATION SKILLS**

Office of the Chair

Your Ref:
Our Ref:

20 February 2025

TO WHOM IT MAY CONCERN

Dear Sir/Madam,

RE: FIELDWORK RESEARCH FOR MS WANG WEI, REG. 2022/HD03/1541X

This is to introduce to you Ms Wang Wei, Reg. No. 2022/HD03/1541X, an MA student registered in the Department of Linguistics, English Language Studies and Communication Skills, Makerere University.

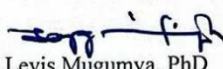
Ms Wang is preparing to conduct her fieldwork research for her MA Linguistics studies. Her research project title is *Analysis of tone errors in the acquisition of Chinese as a second language: a case of Ugandan learners of Chinese*.

As part of this project, she will be carrying out fieldwork at the Luyanzi Institute of Technology in Bweyogerere, Wakiso district. The information to be gathered from this fieldwork is exclusively for academic purposes and it is not intended, in any way, for economic or other purposes.

This is therefore to request you to accord Ms Wang the necessary cooperation she may require to carry out the fieldwork activities.

Any assistance given to her will be highly appreciated.

Yours faithfully,


Levis Mugumya, PhD
Head of Department



APPENDIX II: EXPERIMENTAL OBJECT

(Ugandan L2 speakers of Mandarin)

Language Branch	Language	Gender	Age Years	Chinese Proficiency Level	Major	NO.
Bantu	Luganda	F	23	HSK4	Diploma in Chinese Language and Culture	Student A
		M	24	HSK4		Student B
		F	20	HSK4		Student C
		M	25	HSK4		Student D
	Lugisu	F	25	HSK4		Student E
		F	26	HSK4		Student F
		M	25	HSK4		Student G
		F	21	HSK4		Student H
	Lusoga	M	22	HSK4		Student I
		F	21	HSK4		Student J
		F	26	HSK4		Student K
		F	23	HSK3		Student L
		M	22	HSK4		Student M
	Runyankore	M	25	HSK4		Student N
		M	23	HSK4		Student O
		F	23	HSK4		Student P
		M	24	HSK4		Student Q
	Samia	M	27	HSK4		Student R
Central Sudanic	Lugbarati	M	24	HSK4		Student S

**APPENDIX III: EXPERIMENTAL OBJECT (Native Speakers of
Mandarin)**

Language Branch	Language	Gender	Age Years	Chinese Proficiency Level	Major	NO.
Sinitic Branch	Mandarin Group	M	34	First Class of Level 2 in the Mandarin Proficiency Test	African Law	Native Speaker

APPENDIX IV: RECORDING MATERIALS

Recording Materials						
Category	Tones					
Monosyllabic Syllables	Tone 1	mā	bēi	cōng	guō	
	Tone 2	hái	qióng	quán	fó	
	Tone 3	jiǔ	xuě	lěng	shuǎng	
	Tone 4	tuì	rùn	yuàn	wàng	
Disyllabic Syllables	Tone 1+4 Individual Tones	kāfēi	gōngyuán	jīchǎng	chēzhàn	
	Tone 2+4 Individual Tones	shíjiān	yínháng	cídiǎn	lánsè	
	Tone 3+4 Individual Tones	shǒujī	xǐnglái	suǒyǐ	shǒutào	
	Tone 4+4 Individual Tones	miànbāo	miàntiáo	diànnǎo	fàngqì	
Tone Sandhi	Sandhi <i>yī</i>	shíyī	yìxiē	yìzhí	yìqǐ	yígòng
	Sandhi <i>bu</i>	bùgāo	bùlái	bùmǎi	búhuì	
Neutral	4 Individual Tones + Neutral	bāozi	chúle	diǎnxin	dàifu	

APPENDIX V: DATA STRUCTURE AND VARIABLE ORGANIZATION FOR THE PRODUCTION DATASET

Variable	Unit of analysis	Type	Coding / range	Definition / computation
participant_id	participant	string	S01–S19	Unique participant code.
L1_group	participant	categorical	Luganda/Lusoga/Lugisu/Runyankore/Samia/Lugbarati	Self-reported first language; used to define groups.
stimulus_id	trial	string	S001–S074	Unique token for each recording item.
syllable	stimulus	string	e.g., ma, ba, ta ...	Monosyllable segment corresponding to the Mandarin syllable inventory used.
target_tone	trial	categorical	T1/T2/T3/T4	Intended Mandarin lexical tone for the item.
repetition	trial	integer	1–3	Within-participant repetition index (each item recorded three times).
audio_file	trial	string	*.wav	Recorded waveform filename (linked to participant_id, stimulus_id, repetition).
F ₀ _track	trial	numeric vector	time-normalized F ₀ points	Praat-extracted F ₀ contour after preprocessing (voicing checks, smoothing, outlier handling).
F ₀ _mean_hz	trial	float	Hz	Mean F ₀ across the voiced portion of the syllable.
F ₀ _min_h	trial	float	Hz	Minimum F ₀ across the

Variable	Unit of analysis	Type	Coding / range	Definition / computation
z				syllable.
F ₀ _max_ hz	trial	float	Hz	Maximum F ₀ across the syllable.
F ₀ _range_ _hz	trial	float	Hz	F ₀ _max_hz – F ₀ _min_hz.
duration_ ms	trial	float	ms	Syllable duration from segmented onset to offset.
deltaF ₀ _r ange_sim ilarity	participant ×tone	float	0–100%	ΔF ₀ -based range-similarity index relative to the native baseline; averaged across tokens within tone and participant, then summarized by group.
native_ba seline_id	speaker	string	NB01	Certified native speaker baseline recording used for similarity computations.
rater_id	annotation	string	R1/R2	Annotator code for segmentation and tone labeling.
kappa	study-level	float	0–1	Cohen's κ computed on a stratified subset (n=820 tokens) to assess inter-rater agreement.
qc_subset _n	study-level	integer	820	Number of tokens sampled for inter-rater reliability; stratified across tones and L1 groups.