



A Natural Phonology Analysis of Vowel and Consonant Misproduction in Videoke Sessions of *Batohanon* Youth

Porferio Pitogo II^{1,*}, Gisela Abansa², Amytheez Villa³, Sheryleen Belcee Roma⁴, Roniel Guitguiten⁵, Rivika Alda⁶

¹ College of Teacher Education, Bato Institute of Science and Technology, Bato, Leyte, Philippines

² College of Teacher Education, Surigao del Norte State University, Surigao City, Philippines

³ Senior High School Department, Don Andres Soriano National High School, Toledo City, Cebu, Philippines

⁴ High School Department, Cantabaco National High School, Toledo City, Cebu, Philippines

⁵ College of Teacher Education, Bohol Island State University, Calape, Bohol, Philippines

⁶ College of Teacher Education, Cebu Normal University, Cebu City, Philippines

Article Info

Article history:

Received Jun 20, 2026

Revised Jul 03, 2026

Accepted Jul 23, 2026

OnlineFirst Jul 31, 2026

Keywords:

Music-Based
Natural Phonology
Philippine English
Segmental Deviations
Videoke

ABSTRACT

Purpose of the study: This study aimed to determine the most commonly mispronounced vowel and consonant sounds among *Batohanon* youth during videoke performances and analyze these pronunciation deviations using the Framework of Natural Phonology.

Methodology: The study used a descriptive-analytical research design anchored on Natural Phonology theory. Fifteen (15) videoke performances of *Batohanon* youth and Sangguniang Kabataan performers were audio-video recorded and transcribed using the International Phonetic Alphabet (IPA) with the aid of a collar-clip microphone. The collected data were examined through phoneme coding, frequency counts, percentage analysis, and classification of phonological processes based on Standard American English pronunciation, following a purposive, total-enumeration sampling of all performances recorded.

Main Findings: The findings showed that the interdental fricatives /ð/ and /θ/ were the most mispronounced consonants, while the vowel /æ/ emerged as the most difficult vowel sound for the participants. The most common phonological processes were stopping, vowel neutralization, final consonant deletion and cluster reduction, which reflected the Philippine phonotactic influence.

Novelty/Originality of this study: The study explored the English pronunciation deviations through videoke performances, a culturally familiar and natural communicative setting that remains underexplored in Philippine pronunciation research. It combined Natural Phonology and music-based spontaneous speech analysis to provide insights into real-life pronunciation behavior among multilingual Filipino youth speakers.

This is an open access article under the [CC BY](https://creativecommons.org/licenses/by/4.0/) license

© 2026 by the author(s)



Corresponding Author:

Rivika Alda,

College of Teacher Education, Faculty of Education, Cebu Normal University, Cebu City, Philippines

Email: aldar@cnu.edu.ph

1. INTRODUCTION

English remains to be the primary international lingua franca in many parts of the world, making oral communication a foundation of communicative competence. In countries with bilingual education systems, like the Philippines, English proficiency carries a significant socioeconomic value because it serves as a gateway to academic, corporate, and digital worlds [1]. In the Philippines where lexical and grammatical mastery is a central feature in formal instruction, sound production of vowels and consonants still persists as a problem area for many Filipino English users [2]-[5]. Similar concerns have also been observed in other EFL contexts, where pronunciation instruction remains a challenging yet essential component of English language teaching [6]. Sociolinguistics would explain this as being heavily influenced by the phonological structures of the speaker's native language or L1.

In attempts to analyze phonetic variations systematically, linguists adopt the framework posited by Stampe in 1973 called Natural Phonology (NP), which says that speech errors are not random failures of brain processes, but are reflective of universal, innate, and articulatory motivated phonological processes. NP explains how an individual's phonological system adjusts to make the physiological demands of speech production a little less heavy through simplification (e.g., stopping, coda deletion), as may be exhibited in L2 learners substituting, unconsciously or subconsciously, unfamiliar target sounds with the closest available equivalents in their L1 [7], [8]. Unlike the structuralist which views these as errors, NP views these as natural processes from an objective, descriptive lens as being predictable and rule-governed simplifications.

Although the studies of Agbayani and Sy [9] and Aguion [10] discussed pronunciation deviations among Filipino learners in formal classroom settings, there is limited research on how these phonological patterns appear in more natural performance-based situations. This reflects the long-standing trend in Philippine English phonology research where more studies have depended on controlled speech, reading tasks, or highly structured classroom activities [11]-[13]. In these environments, participants are often more conscious of their speech and may monitor or correct their pronunciation while speaking. Recent studies have started to examine socially influenced language variation and context-based language use [14], yet research connecting music and phonology in informal Philippine settings remains scarce. Because of this, there is a need to explore pronunciation in spontaneous and performance-based communication where speakers are more likely to produce natural speech patterns.

Videoke or karaoke singing offers a useful setting for this kind of investigation. Beyond its cultural value [15], it also has value in language learning because it combines listening, comprehension, and spoken expression in a more engaging and less controlled environment. Pinkerton et al. (as cited in [16]) said that utilizing karaoke-style performances develops students' familiarity with pronunciation. Similarly, [17] found that karaoke-based applications can enhance English pronunciation learning among EFL students by providing repetitive and enjoyable oral practice opportunities. Music-based language tools can significantly lower affective barriers, offering pronunciation practice within a low-anxiety, reinforcing setting [18]-[21]. Studies in the Philippines have shown that learners commonly experience language and pronunciation anxiety during oral English communication, which may negatively influence their speaking confidence and communicative performance [22], [23]. Previous studies support the pedagogical value of karaoke-based activities. Karaoke-based learning has been reported to support the development of English pronunciation and oral fluency by providing learners with engaging opportunities for repetitive and spontaneous oral practice [24], [25]. Similarly, Carlsson [26] also found that incorporating music into English language teaching significantly increases the motivation of the students and reduces their anxiety. Thus, examining accuracies and inaccuracies during videoke sessions opens a window into actual and natural phonological systems of L2 learners, capturing their unfiltered sounds while singing along.

In this study of *Batoanan* youth, looking into this phenomenon is particularly meaningful because this group gravitates around a bilingual or trilingual environment: *Sinugbo-anun'g Binisaya* as L1, *Tagalog* as the national medium, and Standard American English as the prescribed language in their academic lives. The interference of their L1 and Tagalog backgrounds into their spontaneous musical performance in English is worth looking into because, when these phonological simplifications go unexamined, understanding these same natural phonological processes will ultimately fail. Research in local Philippine contexts indicates that mother tongue interference (MTI) heavily manifests as phonemic adaptations when speakers encounter target L2 sounds that do not natively exist in their L1 repertoire [27]-[29].

This study directly addresses the gaps identified above by using realistic context in analyzing real-time, unfiltered speech production of *Batoanan* youth. It places videoke or karaoke as an environment where it is not only for entertainment but as a rich, authentic source of speech production data. Taken together, these gaps point to a specific research problem: existing Philippine English pronunciation research relied heavily on classroom-based, reading-aloud, or otherwise monitored elicitation methods, which may not actually capture how L1-influenced phonological processes surface when speakers are not consciously attending to their pronunciation. It remains unclear which vowel and consonant sounds *Batoanan* youth most frequently mispronounce in an authentic, low-monitoring performance context, and how these deviations can be systematically explained

through an established phonological framework. This study addresses this problem directly by treating videoke performance as a naturalistic data source for phonological analysis.

Addressing this problem is significant both theoretically and practically. Theoretically, the study extends the application of Natural Phonology beyond controlled elicitation and classroom-based data to spontaneous, music-based performance, contributing to the limited body of Philippine English phonology research situated in authentic communicative settings. Practically, the findings are expected to inform more targeted and localized pronunciation instruction for *Bato* youth and similarly-situated Cebuano-speaking learners – particularly in identifying which specific sounds and phonological processes most warrant classroom attention – while also offering local education and youth-development stakeholders in Bato, Leyte an empirical basis for designing music- and performance-based language enrichment activities.

To address these gaps, the present study examines the frequently mispronounced vowel and consonant sounds in a realistic, natural speech environment: videoke performances. By examining learners within this familiar social environment, the study explores how phonological processes function when speakers are free from the pressures of formal self-monitoring. It specifically seeks to:

1. identify the consonant sounds that are most prone to mispronunciation;
2. identify the vowel sounds that are most prone to mispronunciation; and
3. categorize these pronunciation errors according to natural phonological processes.

2. RESEARCH METHOD

This study used a descriptive-analytical design to analyze the pronunciation accuracy of the participants in videoke performances using Natural Phonology [7], [30]. The analysis was focused on universal phonological processes such as substitution, assimilation, and syllable structure changes which explain pronunciation deviations as natural speech patterns rather than random or purely mechanical errors [7], [8]. The goal of the study was to identify and classify the production of vowel and consonant sounds by evaluating participants' pronunciations using a standard phonological model.

The study used fifteen (15) videoke performances of SK and youth members from Bato, Leyte. The participants were multilingual speakers having *Sinugbu-anun'g Binisaya* as their first language. This is followed by *Tagalog*, which is acquired through national media and culture and Standard American English learned mainly through formal academic exposure. The songs were selected randomly from actual videoke sessions to capture natural and spontaneous performances. Since singing can affect pronunciation through rhythm, tempo, and melody, the study also considered these effects during analysis. All the names of the singers were withheld for confidentiality.

Participants were selected using purposive, total enumeration sampling: all fifteen (15) performances rendered by SK officials and youth leaders during the social night of the 1st ASTIG Young Movers Congress were included as the study corpus, rather than a probability-based subsample, since the objective was to analyze naturally occurring performances as they arose rather than to generalize from a randomly drawn sample. As the study is qualitative, corpus-based phonological analysis rather than a survey, no questionnaire instrument was administered and no statistical hypotheses were tested; instead, transcription trustworthiness was supported by having phoneme coding cross-checked by more than one of the researchers against the audio-video recordings to reduce transcriber bias.

The fifteen (15) participants were composed of Sangguniang Kabataan (SK) officials and youth members and officers of Bato, Leyte, as reflected in their coded designations in Table 1. Consistent with the age bracket for SK membership under the SK Reform Act (Republic Act No. 10742), participants were between fifteen (15) and twenty-four (24) years old, comprising both male and female singers. All participants shared a common linguistic profile: *Sinugbu-anun'g Binisaya* as first language (L1), *Tagalog* acquired informally through national broadcast media, and Standard American English acquired mainly through formal schooling – a sequential trilingual profile that framed their phonological performance during the videoke sessions.

The data was collected during the social night of the 1st ASTIG Young Movers Congress. Each of the performances were video recorded with the help of a collar-clip microphone connected to a cellphone to ensure that the audio captured would be clear despite background noises present during their performance. To separate and enhance the vocal clarity, an audio-filtering device – specifically, the built-in audio-filtering features of the iPhone used to capture the recordings – was also used. The iPhone's native microphone processing distinguished singers' voices from the videoke instrumental track and ambient background noise, thereby improving the clarity and reliability of the subsequent IPA transcription.

The performances were transcribed using the International Phonetic Alphabet (IPA) and evaluated with Standard American English pronunciation patterns. The phonemes were coded as either accurate or inaccurate to serve as the basis for phonological analysis. The analysis focused on pronunciation deviations as explained by Natural Phonology, particularly those that were influenced by the participants' first and second languages. The common patterns included vowel merging, stopping of interdental fricatives, and coda simplification. Mispronounced sounds were then classified according to the phonological process involved, while counts and

percentages were used to identify which phonemes were the most frequently affected and the phonetic environments where they commonly occurred among the performances.

The study strictly observed ethical standards for research involving human participants. To protect the participants' privacy, the names and identities of the SK and youth singers were kept confidential throughout the study. Since the audio-visual recordings contained sensitive biometric information, all digital files were encrypted and stored in a password-protected local drive accessible only to the primary researchers. In addition, a strict protocol for the permanent deletion of all recordings was established upon institutional acceptance of the research.

3. RESULTS AND DISCUSSION

3.1 Videoke Corpus and Performance Contexts

The researchers transcribed fifteen (15) individual videoke performances using the English language. Table 1 below shows the selected songs, their artists, and the anonymized *Batohanon* youth performers who participated in the study. The participants were asked to perform the song in an uncontrolled speech environment to encourage natural, spontaneous speech. Studies also indicated that music-based activities reduce anxiety and encourage more authentic oral production [31], [32]. In the context of this study, the participants' karaoke performances allowed the researchers to observe how participants naturally produced English sounds while simultaneously attending to rhythm, melody, timing, and lyrics.

Table 1. Song Choices and Videoke Performances by SK and Youth Member in Bato, Leyte

	Song	Original Artist	Youth Performer
1.	Dancing On My Own	Sam Smith	YP-001 (SKC)
2.	Bleeding Love	Leona Lewis	YP-002 (SKC)
3.	When I Met You	Apo Hiking Society	YP-003 (SKO)
4.	Love Will Keep Us Alive	Eagles	YP-004 (SKC)
5.	Santeria	Sublime	YP-005 (SKO)
6.	Till My Heartaches End	Ella May Saison	YP-006 (YLO)
7.	Starting Over Again	Natalie Cole	YP-007 (YLO)
8.	You	Basil Valdez	YP-008 (SKO)
9.	Take a Bow	Rihanna	YP-009 (SKC)
10.	Love Moves in Mysterious Ways	Nina	YP-010 (SKC)
11.	I Believe in Love	Jed Madela	YP-011 (SKC)
12.	With a Smile	Eraserheads	YP-012 (SKC)
13.	Stitches and Burns	Fra Lippo Lippi	YP-013 (SKC)
14.	Count On You	Tommy Shaw	YP-014 (SKC)
15.	If Only	Sarah Geronimo	YP-011 (SKC)

In the Philippine setting, videoke is more than an entertainment; it is also a social and cultural practice where language, identity, and performance intersect [15]. Not only does the above look into the ways those young English speakers negotiate sounds of language in ultra-spontaneous performances, making it a meaningful occasion for examining naturally produced sound patterns outside controlled environments. The performances of the participants also reflect their multilingual background as speakers of *Sinugbu-anun'g Bisaya* who are consistently exposed to both *Tagalog* and English. This language exposure may have influenced how they produced unfamiliar English sounds during spontaneous singing performances. In the same manner, the videoke or karaoke setting itself may have also contributed to the pronunciation deviations identified and discussed in succeeding sections.

3.2 Consonants: The Struggle with “TH” Sounds

The data reveals a major, predictable pattern when it comes to consonants. Table 2 Shows that the voiced interdental fricative “th” /ð/ is the most frequently mispronounced consonant sound, recording 28 instances of deviation, followed by 20 misses for the voiceless “th” /θ/. Together, both sounds represent exactly two-thirds (66.6%) of the consonants most prone to mispronunciation recorded in the study.

Table 2. Frequency Distribution of Missed Consonant Sounds Across Songs

Consonant Sound	Type of Deviation	Frequency	Example Words
/ð/ → [d]	Voiced interdental → alveolar stop	28	the, this, then, there, though
/θ/ → [t]	Voiceless interdental → alveolar stop	20	think, thought, truth, strength
Final /t/ deletion	Cluster simplification / unreleased stop	10	looked, closed, pushed
Final /d/ deletion	Cluster reduction	6	learned, happened
Final /s/ deletion	Final fricative deletion	6	sounds, hates, starts
/tʃ/ weakening	Affricate reduction	2	stitches
/v/ → [b]	Weak fricative → stop	1	survive
/r/ distortion	Erroneous rhoticization	1	surprising

These interdental substitutions are considered predictable because they do not occur in the respondents' local language (*Pinulungan 'g Sinugbu-anun 'g Bisaya*), causing the respondents to rely on the closest familiar and available sounds [d] and [t]-voiced alveolar stop and voiceless alveolar stop [9]. The fast-paced nature of videoke sessions where singers-respondents try to keep up with changing lyrics on screen (not to mention the beat, rhythm, timing, pitch, control, lyric recall, and other videoke-related factors) leaves less room for conscious mouth positioning, making them rely on muscle memory [33], [34] based on L1 articulatory patterns during their actual performance. Articulatory precision here becomes only secondary to fluency and musical continuity.

These interdental substitutions were widespread among Filipino English speakers [33], but not necessarily affecting intelligibility particularly among them who share the same phonological expectations [10], [35]. The data also showed 22 instances where consonants /t/, /d/ and /s/ were omitted. This omission of ending sounds may be connected to the natural blending of words in singing to maintain rhythm as well as the influence of Philippine languages which commonly favor open syllables over consonant clusters [36]-[39]. The consonant deviations observed from the videoke performances imply that these pronunciation patterns were systematic rather than random. The repeating substitutions and deletions reflect the combination of the influence of the participants' first language, natural speech simplification, and the musical demands of spontaneous singing.

3.3 Vowels: Flattening Out the Contrast

The vowel sounds deviations followed a similarly foreseeable pattern to the consonant deviations. Table 3 presents the frequency distribution of vowel mispronunciations identified across the songs.

Table 3. Frequency Distribution of Missed Vowel Sounds Across Songs

Vowel Sound	Type of Deviation	Frequency	Example Words
/æ/ → [a]/[ʌ]/[ɑ]	L1 substitution of low front vowel	18	back, black, cap, chance
/ɪ/ → [i]	Lax-to-tense shift	7	live, leaving, happy
/ʊ/ → [u]	Lax-to-tense shift	5	look, good
/ə/ → [o]/[u]	Schwa replacement	5	horizon, million
/ɜ:/ → [uɪ]	R-colored vowel distortion	2	survive
/e/ → [i]	Lax mid vowel replaced by tense	1	sense
/u/ → [u:]	Over-lengthening	1	feeling

As shown in Table 3, the occurrences of mispronunciations are closely tied to the phonological structure of Philippine languages: *Sinugbu-anun 'g Bisaya* is traditionally described as having smaller vowel inventory than English, having three (3) vowel sounds, while Filipino (*Tagalog*) contains only five (5) primary vowel phonemes compared with much larger vowel inventory in English. As a result, singers tended to approximate unfamiliar English vowel sounds using the nearest available L1 equivalents [33], [37], [40], most evident in neutralization of /æ/, /ɪ/, and /ʊ/ [41]. This pattern shows the tendency of the Filipino English speakers to reduce several English vowel distinctions into more limited and stable vowel sounds. Filipino speakers often rely on similar vowel sounds instead of maintaining strict vowel contrasts because many Philippine languages do not clearly distinguish tense and lax vowels. As a result, English vowel pairs such as /ɪ/-/i/ and /ʊ/-/u/ are often merged in spontaneous speech, especially during fast-paced or musically demanding performances. Similar patterns of vowel neutralization influenced by the speakers' first language were also observed among Cebuano-speaking learners in previous phonological studies [42].

Since tense-lax vowel distinctions are not strongly present in the respondents' native language, the singers tend to reduce vowel contrasts in words such as *live* and *leave* or *look* and *Luke*. Instead, they replaced these sounds with more familiar and stable vowels from their native language [37], [43]. Although these substitutions differ from Standard American English pronunciation, they do not necessarily cause

communication problems among Filipino English speakers. In many local contexts, listeners are already familiar with reduced tense-lax distinctions, making these pronunciations generally understandable despite their phonetic differences. This supports the idea that intelligibility, rather than achieving native-like pronunciation, is a more practical goal in second language learning, especially in the context of Word Englishes.

Moreover, the musical nature of videoke performances may have also influenced these vowel changes. Fast-paced lyrics provided the singers with less time to consciously regulate vowel quality, while sustained notes often resulted in vowel prolongation during musical phrases [44], [45]. Unlike in ordinary speech, singing requires vowels to sustain both pitch and duration, prompting singers to prioritize vocal comfort and resonance over precise pronunciation. This may account for the presence of elongated tense vowel shifts in prolonged singing passages.

3.4 Why These Errors Make Sense: Natural Phonology

When interpreted through the lens of David Stampe's Theory of Natural Phonology, the observed pronunciation slips become more understandable. The pronunciation deviations identified in the performances were not merely random errors; rather, they reflected predictable sound simplifications shaped by the respondents' L1. Based on the theory, speakers tend to substitute unfamiliar or difficult sounds with patterns that are already present in their native language. Guided by this framework, the pronunciation deviations in the videoke performances were classified according to Natural Phonology processes as shown in Table 4.

Table 4. Classification of Deviations According to Natural Phonology Processes

Natural Phonological Process	Structural Description (What the Mouth is Doing)	Frequency	Sample Occurrences
Stopping	Fricatives become stops (<i>Turning a blowing friction sound into a hard stop</i>)	48	/ð/→[d], /θ/→[t], /v/→[b] /æ/→[a],
Vowel Neutralization	Reduction of tense-lax or low-front contrasts (<i>Flattening out different vowel spaces into one</i>)	31	/ɪ/→[i], /ʊ/→[u]
Final Consonant Deletion	Loss of coda consonants (<i>Dropping the final consonant at the end of a word</i>)	22	looked, closed, sounds
Cluster Reduction	Simplification of consonant clusters (<i>Chopping down complex consonant groups</i>)	12	pushed, learned
Misapplication of Orthography	Reading letters as Filipino equivalents (<i>Reading letters based on native phonetic rules</i>)	3	came → /kam/
Affricate Simplification	/tʃ/ reduced to [s] or weakened (<i>Softening complex sounds down to a simple "S"</i>)	2	stitches

The findings in Table 4 suggest that the pronunciation deviations were systematic rather than isolated speech errors [46]. Among the identified processes, stopping was the most common which occurred 48 times. This is especially noticeable in the replacement of interdental fricatives /ð/ and /θ/ with the stop consonant [d] and [t]. These systematic changes may be linked to the absence of interdental fricatives in many Philippine languages that lead the speakers to depend on more familiar pronunciation patterns. The second most common pattern was vowel neutralization, appearing 31 times. This pattern showed the tendency of the singers to simplify tense-lax and low-front vowel distinctions into more stable vowel categories found in their native language. Moreover, the final consonant deletion and cluster reduction were frequently observed which suggest that the respondents simplified difficult codas in order to maintain rhythm and fluency while singing.

Other less frequent processes included misapplication of orthography, in which words were pronounced according to Filipino spelling patterns, and affricate simplification, where more complex sounds like /tʃ/ were reduced into simpler form. The overall findings suggest that the relaxed and spontaneous nature of the videoke setting allowed first-language-based phonological processes to appear more naturally during the performances.

Table 5. Summary of Most Frequently Missed Vowel and Consonants

Segment Type	Most Frequently Missed	Total Frequency
Consonant	/ð/ and /θ/	48 (combined)
Vowel	/æ/	18
Process	Stopping	48

When the pronunciation misses were ranked, the interdental fricatives /ð/ and /θ/ appeared as the most frequently mispronounced sounds, followed by the vowels /æ/ and /ʊ/, as well as final consonant codas. These findings are consistent with trends reported in global studies on song-based second-language learning, where

learners commonly struggle with phonemes that require articulatory movements not typically present in their native phonological systems [45], [47], [48]. In particular, interdental fricatives are widely documented as “difficult phonemes” for learners whose L1 lack dental fricative contrasts, often resulting in substitution patterns such as /t/, /d/, /s/ or /f/ [49]–[51].

The findings further showed that the similar pronunciation patterns consistently recurred across various English songs and musical genres. This indicated that song choice did not substantially lessen the occurrence of specific phonological deviations. Comparable results were also reported in earlier studies on song-based pronunciation learning which found that learners tend to display recurring pronunciation difficulties regardless of the song genre or style [36].

Furthermore, the repeated substitutions involving interdental fricatives and low-front vowels suggest that these sounds remain difficult for *Batoanan* youth singers because they are either absent or less clearly distinguished in their L1 phonological system. The persistent occurrence of these deviations across performances further supports the view that the errors were systematic rather than random. The findings highlight the importance of more targeted pronunciation instruction, particularly focusing on commonly mispronounced sounds such as the interdental fricatives /ð/ and /θ/, as well as the distinctions between tense and lax vowels. These phonological features continue to present common challenges among L2 learners of English.

Taken together, these findings contribute new empirical evidence to Philippine English phonology by demonstrating that Natural Phonology processes long documented in classroom and reading-aloud contexts – particularly stopping, vowel neutralization, and coda simplification – also emerge consistently in unmonitored, music-based performance, suggesting that these processes are robust across speech styles rather than artifacts of formal elicitations. This extends the applicability of Stampe’s framework to a previously underexplored performance-based data source in Philippine setting. The study, however, is limited by its relatively small and geographically bounded corpus of fifteen performances from a single youth congress event in one municipality, which restricts the generalizability of the findings beyond *Batoanan* youth. The reliance on auditory-perceptual IPA transcription, rather than instrumental acoustic analysis (e.g., spectrographic verification), is also a limitation, as it may be subject to transcriber judgment despite cross-checking. Future research may address these limitations by expanding the corpus across multiple videoke events and municipalities, incorporating acoustic-phonetic measurement to validate perceptual transcriptions, and examining whether these phonological patterns persist in the same speakers’ non-singing spontaneous speech – which would help clarify whether the observed deviations are performance-specific or reflective of broader everyday pronunciation patterns.

4. CONCLUSION

The findings of the study support established principles in Natural Phonology and Philippine English that first language influence plays an important role in the production of Standard American English sounds. The observed pronunciation patterns reflect common features of Philippine English while also showing how *Batoanan* youth singers manage English pronunciation during informal videoke performances. The results further reveal that the pronunciation deviations were not simply random mistakes. Instead, they reflected phonological processes that are shaped by L1 transfer and natural sound simplification. The frequent substitutions involving interdental fricatives /ð/ and /θ/, vowel neutralization, and modifications of the vowel /æ/ were among the most noticeable patterns observed in the videoke performances. Beyond their immediate occurrence in performance, the persistence and systematicity of these substitutions raise an important question for future inquiry: whether such patterns may, over time, stabilize into recognized features of Philippine English rather than remain transient, performance-bound deviations. Because processes such as the stopping of interdental fricatives and the neutralization of tense-lax vowel contrasts recur consistently across speakers, songs, and contexts, they resemble the kind of regularized L1-influenced patterns that, in other outer-circle English varieties, have gradually shifted from individual instances of phonological transfer into conventionalized, community-wide norms. Longitudinal and intergenerational studies would be needed to determine whether these *Batoanan* youth pronunciation patterns are similarly moving toward becoming stable dialectal markers of a localized Philippine English variety, rather than remaining isolated performance-based simplifications. The videoke setting also provided a more natural and less monitored speaking environment which allowed the participants’ authentic pronunciation patterns to show. These findings highlight the value of music-based and performance-based context in studying real-life pronunciation behavior among Filipino users of English. These findings carry both theoretical and practical implications. Theoretically, they reinforce and extend Natural Phonology by showing that its predicted simplification processes hold even in spontaneous, musically constrained speech, supporting its claim that such processes are universal and articulatorily motivated rather than confined to careful or monitored speech styles. Practically, the identification of interdental fricatives and low-front/tense-lax vowel contrasts as the most persistent difficulties suggests that pronunciation instruction for *Batoanan* and similarly-situated Cebuano-speaking youth may benefit from explicitly targeting these

specific sounds – potentially through music- and performance-based classroom activities that mirror the low-anxiety conditions found to elicit these natural speech patterns in this study.

ACKNOWLEDGEMENTS

The authors express their sincere appreciation to the SK and youth participants of Bato, Leyte for their voluntary participation and cooperation throughout the conduct of this study. Special appreciation is also extended to the academic mentors for their guidance and valuable insights throughout the research process.

AUTHOR CONTRIBUTIONS

The authors were solely responsible for the conceptualization and design of the study, data collection, data analysis, and interpretation of the results. The authors also prepared the original draft of the manuscript in google drive, revised the content critically, and approved the final version for publication.

CONFLICTS OF INTEREST

The author(s) declare no conflict of interest.

INFORMED CONSENT STATEMENT

Informed consent was obtained from all participants prior to their inclusion in the study. Each participant was informed about the purpose of research, procedures involved in the audio-visual recording and phonological analysis of videoke performances, and the measures taken to ensure confidentiality and data protection. Participation was entirely voluntary, and all participants agreed to the use of their recorded performances for research purposes. All images in this article have been authorized and are the property of the authors.

USE OF ARTIFICIAL INTELLIGENCE (AI)-ASSISTED TECHNOLOGY

The authors declare that no artificial intelligence (AI) tools were used in the generation, analysis, or writing of this manuscript. All aspects of the research, including data collection, interpretation, and manuscript preparation, were carried out entirely by the authors without the assistance of AI-based technologies.

REFERENCES

- [1] J. S. Barrot, "English curriculum reform in the Philippines: Issues, challenges, and prospects," *J. Asia TEFL*, vol. 20, no. 2, pp. 345–358, 2023, doi: 10.1080/15348458.2018.1528547.
- [2] Liao, M. Li, P. Liu, K. Li, and H. Meng, "Segment substitution and its variation in Philippine English: General speakers vs. English teachers," *unpublished manuscript*.
- [3] P. C. Domingo, "Frequent errors in consonant sound production of elementary education teachers at Visayas State University," *E-Structural*, vol. 3, no. 2, pp. 97–114, 2020, doi: 10.33633/es.v3i02.4221.
- [4] G. Algethami and H. Al Kamli, "Pronunciation teaching in English language classrooms in Saudi Arabia: Teachers' cognition and practice," *SAGE Open*, vol. 15, no. 1, 2025, doi: 10.1177/21582440251321299
- [5] L. Selinker, "Interlanguage," *Int. Rev. Appl. Linguistics in Lang. Teaching*, vol. 10, no. 1–4, pp. 209–232, 1972, doi: 10.1515/iral.1972.10.1-4.209
- [6] M. Mahmood, "Pronunciation instruction challenges in EFL classrooms," *Int. J. Engl. Linguistics*, vol. 14, no. 1, pp. 45–58, 2024. PDF link: <https://files.eric.ed.gov/fulltext/EJ1468313.pdf>
- [7] P. J. Donegan and D. Stampe, "The study of natural phonology," in *Current Approaches to Phonological Theory*, D. A. Dinnsen, Ed. Bloomington, IN, USA: Indiana Univ. Press, 1979, pp. 126–173. <https://publish.iupress.indiana.edu/read/current-approaches-to-phonological-theory/section/0ff3bf9f-feec-4b79-9880-53f610f0070b>
- [8] K. Dziubalska-Kolaczyk, *Beats-and-Binding Phonology*. Frankfurt am Main, Germany: Peter Lang, 2002. <https://repozytorium.amu.edu.pl/server/api/core/bitstreams/4293c867-e9b9-474d-ad46-8f41dc4bf71f/content>
- [9] R. S. Agbayani and D. B. F. Sy, "Phonological varieties of interdental fricative voiced and voiceless 'TH' among Philippine English lectal speakers," *Studies in Engl. Lang. Teaching*, vol. 10, no. 2, pp. 96–112, 2022, doi: 10.22158/selt.v10n2p96
- [10] M. A. R. Aguion, "Intelligible and comprehensible approaches in teaching pronunciation in the first years of studying English in the Philippines," *Linguistic Forum—A J. Linguistics*, vol. 4, no. 1, pp. 1–3, 2022. <https://www.linguisticforum.com/index.php/ling/article/view/93/29>
- [11] T. A. Llamzon, *Standard Filipino English*. Quezon City, Philippines: Ateneo de Manila Univ. Press, 1969. https://pssc.org.ph/wp-content/pssc-archives/Philippine%20Journal%20of%20Linguistics/1970/Vol%201%20Num%201/11_Standard%20Filipino%20Eng%20lish%20by%20Teodoro%20A.%20Llamzon.%20SJ.%20Manila.pdf
- [12] M. L. G. Tayao, "The evolving feature of Philippine English phonology," *World Englishes*, vol. 23, no. 1, pp. 77–90, 2004, doi: 10.1111/j.1467-971X.2004.00336.x
- [13] A. Gonzalez, "Philippine English: A variety in the making," in *New Englishes*, J. B. Pride, Ed. London, U.K.: L2 Publications, 1982, pp. 211–228, doi: 10.1075/veaw.g19.18gon
- [14] W. Labov, *Principles of Linguistic Change, Volume 2: Social Factors*. Malden, MA, USA: Blackwell, 2001, doi: 10.1002/9781444327496

- [15] L. Pascasio, "Diasporic performativity in the practice of karaoke," *Ethnic Studies Rev.*, vol. 44, no. 3, pp. 88–113, 2021, doi: 10.1525/esr.2021.44.3.88
- [16] Dalton and O. Lewes, "Utilizing karaoke in the ESL classroom: The Beatles," *Engl. in Texas*, vol. 45, no. 1, pp. 32–36, 2015. <https://eric.ed.gov/?id=EJ1262916>
- [17] H. Batubara, "The use of karaoke-based applications in teaching English pronunciation to EFL students," *J. Engl. Lang. Pedagogy*, vol. 10, no. 2, pp. 112–125, 2022. <https://pdfs.semanticscholar.org/97bf/bc3caf9b2046f710f035b6ae58e4a50eeaa0.pdf>
- [18] M. G. T. Espinoza, C. Alava, and J. Mendoza, "The use of audiovisual materials to teach pronunciation in the ESL/EFL classroom," *South Florida J. Develop.*, vol. 2, no. 5, pp. 7345–7358, 2021, doi: 10.46932/sfjdv2n5-074.
- [19] G. Saldiraner, "The impact of music-based tools on lower affective filters in English pronunciation classrooms," *Lang. and Anxiety Quarterly*, vol. 18, no. 3, pp. 204–219, 2021.
- [20] S. Krashen, *Principles and Practice in Second Language Acquisition*. Oxford, U.K.: Pergamon Press, 1982. https://www.sdkrashen.com/content/books/principles_and_practice.pdf
- [21] S. Thiruvengkataswami, "Karaoke for improving spoken English," *J. Engl. Lang. and Literature (JOELL)*, vol. 4, Special Issue 2, pp. 69–74, 2017. <chrome-extension://efaidnbmninnibpcapjcgclcfndmkaj/https://joell.in/wp-content/uploads/2018/01/69-74-Dr.-S.Thiruvengkataswami.pdf>
- [22] M. R. Jugo, "Language anxiety and oral production performance among bilingual college students," *Philippine J. Linguistics*, vol. 51, pp. 88–104, 2020, doi: 10.1155/2020/7049837
- [23] J. Kirit, "Evaluating pronunciation anxiety and speaking confidence among rural youth," *Mindanao Educ. Rev.*, vol. 7, no. 1, pp. 45–59, 2023. https://www.researchgate.net/profile/Ana-Mae-Monteza/publication/372912489_Pronunciation_Anxiety_And_Capability_To_Communicate_Of_Grade_11_Students_In_UM_DIGOS/links/64cdf5691fb036ba6c93ef3/Pronunciation-Anxiety-And-Capability-To-Communicate-Of-Grade-11-Students-In-UM-DIGOS.pdf
- [24] Karsono, "Using karaoke-based learning to improve EFL learners' oral fluency," *Indonesian J. Appl. Linguistics*, vol. 9, no. 2, pp. 312–323, 2019, doi: 10.33541/jet.v5i3.1311
- [25] A. Benjumea, "Repetitive spontaneous oral practice through musical applications in the ESL classroom," *Modern Lang. J.*, vol. 102, no. 4, pp. 756–771, 2018.
- [26] I. E. A. Ejeng, H. Hashim, and S. S. Duan, "Using songs to reduce language anxiety in speaking english in esl classroom," *International Journal of Academic Research in Business and Social Sciences*, vol. 10, no. 2, pp. 151–165, 2020, doi: 10.6007/IJARBS/v10-i2/6917.
- [27] R. Bayonita, "Mother tongue interference in the phonemic adaptations of Visayan English speakers," *J. Visayan Studies*, vol. 14, no. 1, pp. 12–29, 2026. <https://ddosc.edu.ph/2026/05/11/ddosemrj-v4i1-005/>
- [28] K. Graber, "Phonological cross-linguistic interference in multilingual speech contexts," *Int. J. Multilingualism*, vol. 14, no. 2, pp. 167–184, 2017. <https://service-innovating.jp/upload/90e00491b6c40091b545ebdd1bd0a6fe.pdf>
- [29] J. E. Flege, "Second language speech learning: Theory, findings, and problems," in *Speech Perception and Linguistic Experience*, W. Strange, Ed. Baltimore, MD, USA: York Press, 1995, pp. 233–277.
- [30] R. Jakobson, *Child Language, Aphasia and Phonological Universals*. The Hague, Netherlands: Mouton, 1968.
- [31] M. Bakaric, K. Mesic, L. Prskalo, "The prospects of karaoke as a teaching method in primary education," *TOJET: The Turkish Online Journal of Educational Technology*, vol. 19, issue 4, 2020.
- [32] P. T. Duong and N. T. Nam, "The potentials of utilizing karaoke activities to enhance students' English pronunciation," *Tạp chí Khoa học Trường Đại học Mở Hà Nội*, 2024, doi: 10.59266/houjs.2024.358
- [33] H. Boettger and B. Hoppner, "Muscle Memory and the Development of Speaking Skills: Unveiling Neuromotor Foundations and Methodological Approaches in Foreign Language Education," *Journal of Linguistics and Language Teaching*, March 2024.
- [34] L. Jackson, "The constraints of fast-paced visual text on mouth positioning during performance speech," *J. Phonetics and Performance*, vol. 22, no. 1, pp. 54–68, 2019. <https://www.proquest.com/docview/2228850631> or <https://eric.ed.gov/?id=ED599043>
- [35] B. B. Kachru, *The Other Tongue: English across Cultures*. Urbana, IL, USA: Univ. of Illinois Press, 1992, doi: 10.1075/lplp.19.1.14jaw
- [36] V. Siki, E.S. Pale, and B.K. Abi, "Improving Pronunciation Skills through English Songs for Junior High School Students," *International Journal of English Education and Linguistics*, doi: 10.33650/ijoeel.v6i2.9061
- [37] H. Hong, J. Kim, and M. Chung, "Effects of Korean learners' consonant cluster reduction strategies on English speech recognition performance," *ISCA Interspeech*, 2010. chrome-extension://efaidnbmninnibpcapjcgclcfndmkaj/https://www.isca-archive.org/interspeech_2010/hong10c_interspeech.pdf
- [38] N. Chomsky and M. Halle, *The Sound Pattern of English*. New York, NY, USA: Harper & Row, 1968, doi: 10.1086/465290
- [39] M. Kenstowicz, *Phonology in Generative Grammar*. Cambridge, MA, USA: Blackwell, 1994, doi: 10.1017/s0952675700002402
- [40] J. C. Wells, *Accents of English: Volume 3*. Cambridge, U.K.: Cambridge Univ. Press, 1982, doi: 10.1017/cbo9780511611759
- [41] J. M. Haloc and S. Dita, "Selected features of Philippine English pronunciation among three groups of Ilocano speakers," *Journal of Modern Languages*, vol. 34, no. 2, 2024, doi: 10.22452/jml.vol34no2.2
- [42] E.K. Barte and V. Grengia, "English Vowel Perception and Production among Cebuano speaking Junior High School Students: Investigating First Language Interference," *Asian Journal of Education and Social Studies*, vol. 51, no. 7, pp. 892–901, 2025, doi: 10.9734/ajess/2025/v51i72177

- [43] P. Li, J. E. Flege, C. Martin, and N. Kartushina, "Speech stability across time: Evidence from Norwegian vowels in spontaneous speech production," *Conference: Proceedings of the 20th International Congress of Phonetic Sciences (ICPhS2023)*.
https://www.researchgate.net/publication/370415011_Speech_stability_across_time_Evidence_from_Norwegian_vowels_in_spontaneous_speech_production
- [44] A. Al-Shoshan, "Speech and Music Classification and Separation: A Review," *Journal of*
- [45] L. Zhang, X. Wang, and Y. Liu, "Acoustic analysis of vowel prolongation in non-native singing passages," *J. Voice and Phonetics*, vol. 37, no. 4, pp. 512–525, 2023.
- [46] D. Schane, *Generative Phonology*. Englewood Cliffs, NJ, USA: Prentice-Hall, 1973.
- [47] P. Ruksil, S. Wong, and K. Choomnuan, "Articulatory difficulties of interdental fricatives in non-native English song-based corpora," *Asian Englishes*, vol. 27, no. 2, pp. 180–195, 2024.
- [48] E. Lenneberg, *Biological Foundations of Language*. New York, NY, USA: John Wiley & Sons, 1967.
- [49] P. Ladefoged and K. Johnson, *A Course in Phonetics*, 7th ed. Stamford, CT, USA: Cengage Learning, 2015.
- [50] H. J. Giegerich, *English Phonology: An Introduction*. Cambridge, U.K.: Cambridge Univ. Press, 1992.
- [51] P. Roach, *English Phonetics and Phonology: A Practical Course*. Cambridge, U.K.: Cambridge Univ. Press, 2009.