

AN INTERDISCIPLINARY ANALYSIS OF TAYLOR SWIFT'S *FOLKLORE*: LEXICAL PROFILING, THEMATIC COMPLEXITY, AND PEDAGOGICAL APPLICATIONS IN THE CREATIVE ECONOMY

Nguyen Quoc Khanh¹, Vo Nhat Quang²,
To Thi Thuy Loan³, Le Thi Thuy Nhung⁴

Received date: May 07, 2026

Revised date: May 29, 2026; Accepted date: September 20, 2026

ABSTRACT

Integrating popular culture artifacts into applied linguistics and interdisciplinary arts research offers a useful way to examine language learning within the creative economy. This paper uses lexical profiling and qualitative thematic analysis to explore the pedagogical potential of Taylor Swift's eighth studio album, *folklore*, in advanced English as a Second or Foreign Language (ESL/EFL) contexts. Drawing on corpus linguistics methods through Lextutor VocabProfile, the study analyzes the lexical density and semantic richness of the 16-track corpus. The quantitative results show a Type-Token Ratio (TTR) of 0.19 and K1+K2 lexical coverage of 94.68%, indicating that the album closely approaches the minimum lexical coverage threshold, especially when Academic Word List (AWL) items are included. At the same time, the relatively high presence of AWL and off-list vocabulary adds to the album's qualitative depth. The thematic analysis reveals layered patterns of meaning, including temporal distortion, relational rupture, moral ambiguity, and cinematic reflexivity. By bridging quantitative lexical profiling with interdisciplinary literary analysis, this study addresses an important research gap. It shows how culturally resonant music can serve as both a pedagogical resource and a knowledge asset in the global creative economy.

Keywords: *Applied linguistics, corpus linguistics, creative economy, ESL/EFL pedagogy, Lexical profiling, Taylor Swift, Thematic analysis*

1. The University of Da Nang.
Email: av32.1b2cva1d047@oncce.udn.vn

2. Ho Chi Minh University of Banking

3. University of Social Sciences and Humanities, Vietnam National University HCMC

4. Ho Chi Minh University of Banking

PHÂN TÍCH LIÊN NGÀNH VỀ ALBUM *FOLKLORE* CỦA TAYLOR SWIFT: ĐẶC ĐIỂM TỪ VỰNG, TÍNH PHỨC TẠP CỦA CHỦ ĐỀ VÀ ỨNG DỤNG SỰ PHẠM TRONG NỀN KINH TẾ SÁNG TẠO

Nguyễn Quốc Khánh, Võ Nhật Quang,
Tô Thị Thùy Loan, Lê Thị Thùy Nhung

Ngày nhận bài: 07 tháng 5 năm 2026

Ngày nhận bài sửa: 29 tháng 5 năm 2026; Ngày duyệt đăng: 20 tháng 9 năm 2026

TÓM TẮT

Việc tích hợp các sản phẩm văn hóa đại chúng vào nghiên cứu ngôn ngữ học ứng dụng và nghệ thuật liên ngành mang lại một phương thức hữu ích để xem xét quá trình học ngôn ngữ trong nền kinh tế sáng tạo. Bài viết này sử dụng phương pháp phân tích đặc điểm từ vựng (lexical profiling) và phân tích chủ đề định tính để khám phá tiềm năng sự phạm của album phòng thu thứ tám của Taylor Swift, *Folklore*, trong bối cảnh giảng dạy tiếng Anh như ngôn ngữ thứ hai hoặc ngoại ngữ (ESL/EFL) ở trình độ nâng cao. Dựa trên các phương pháp ngôn ngữ học khối ngữ liệu (corpus linguistics) thông qua công cụ Lextutor VocabProfile, nghiên cứu tiến hành phân tích mật độ từ vựng và sự phong phú về ngữ nghĩa của khối ngữ liệu gồm 16 bài hát. Kết quả định lượng cho thấy tỉ lệ Loại từ - Lược từ (Type-Token Ratio - TTR) là 0,19 và độ phủ từ vựng K1+K2 đạt 94,68%, cho thấy album này bám sát ngưỡng độ phủ từ vựng tối thiểu, đặc biệt khi bao gồm cả các từ thuộc Danh sách Từ vựng Học thuật (Academic Word List - AWL). Đồng thời, sự xuất hiện tương đối cao của các từ vựng AWL và từ vựng ngoài danh sách (off list) đã góp phần bồi đắp thêm chiều sâu định tính cho album. Phân tích chủ đề hé lộ các tầng ý nghĩa đa dạng, bao gồm sự biến dạng của thời gian, sự đổ vỡ trong các mối quan hệ, sự mơ hồ về đạo đức và tính phản thân điện ảnh (cinematic reflexivity). Bằng cách kết nối phân tích định lượng đặc điểm từ vựng với phân tích văn học liên ngành, nghiên cứu này giải quyết một khoảng trống nghiên cứu quan trọng và cho thấy các tác phẩm âm nhạc có sức lan tỏa về mặt văn hóa có thể đóng vai trò vừa là nguồn tài nguyên sự phạm, vừa là tài sản tri thức trong nền kinh tế sáng tạo toàn cầu.

Từ khóa: Ngôn ngữ học ứng dụng, ngôn ngữ học khối ngữ liệu, nền kinh tế sáng tạo, phương pháp giảng dạy ESL/EFL, phân tích đặc điểm từ vựng, Taylor Swift, phân tích chủ đề.

1. INTRODUCTION

The shift toward a creative global economy has changed how knowledge, culture, and language-based artifacts are produced, circulated, consumed, and studied (The Work Foundation, 2007). Within this broader economic context, the cultural and creative

industries, including music, film, literature, and digital media, are no longer viewed only as entertainment sectors. They are increasingly recognized as important drivers of socio-economic resilience, human capital development, and interdisciplinary innovation (Ministry of Culture, Sports and Tourism &

Vietnam Institute of Culture and Arts Studies, 2013). At the same time, academic institutions such as Vietnam National University, Hanoi (VNU), particularly through units such as the School of Interdisciplinary Sciences and Arts, have shown growing interest in the links between linguistics, cultural policy, and creative entrepreneurship (Vietnam National University, Hanoi, 2025).

In today's music industry, Taylor Swift holds a particularly influential position, creating a cultural and economic phenomenon that scholars and media commentators often describe as "Swiftology" (Driessen, 2026). Her work has attracted wide academic attention across sociology, media studies, gender studies, and literary criticism, particularly in areas such as narrative identity, fandom studies, and the use of figurative language (Driessen, 2026; McCann et al., 2025). However, a clear gap remains in applied linguistics. Existing linguistic and discourse-analytic studies on Swift have often focused on stylistic devices and basic corpus patterns in pop lyrics, which are commonly described as repetitive and syntactically simple. Far fewer studies have used rigorous mixed-methods and interdisciplinary frameworks, specifically bridging quantitative lexical profiling with qualitative thematic analysis, to evaluate the pedagogical viability of more complex indie-folk hybrid albums within the creative economy (Kendong et al., 2023).

Written and produced during the global COVID-19 lockdowns, *folklore* marks a notable departure from Swift's earlier autobiographical pop anthems. The album presents a dense fictional storyworld shaped by literary and linguistic hybridity, drawing on techniques associated with epistolary fiction, folklore, and cinematic discourse (Murphy & Thompson, 2026).

This paper combines lexical profiling with qualitative thematic analysis to examine the pedagogical potential of Taylor Swift's *folklore*. It aims to show how the album may function as a valuable authentic material for ESL/EFL pedagogy while also highlighting the role of interdisciplinary arts within the creative economy. The following research questions guide the study:

RQ1: What is the quantitative lexical profile of the *folklore* corpus, and to what extent does its lexical coverage align with established comprehension thresholds for ESL/EFL learners?

RQ2: What are the main qualitative thematic structures in the lyrics, and how are advanced vocabulary items, including AWL and off-list words, used to construct these literary narratives?

RQ3: How do the lexical properties identified in RQ1, and the thematic structures identified in RQ2 support the integration of *folklore* as an authentic pedagogical material within the creative economy?

2. LITERATURE REVIEW

2.1. The Efficacy of Authentic Materials in Language Acquisition

In applied linguistics, "authentic materials" refer to written or spoken texts created for real-world communication rather than specifically designed for language teaching (Sadiqzade, 2025). These materials can include news broadcasts, podcasts, films, and popular music. Their use in language classrooms has been closely associated with stronger learner motivation, lower affective filters, and better intercultural communicative competence (Alamri, 2025).

Music is especially valuable because it offers several neurocognitive benefits. Studies suggest that musical and linguistic syntax are

processed through overlapping neural regions (Patel, 2003, 2008). The rhythm and repetition found in songs can also support memory encoding and incidental vocabulary learning (Vivone et al., 2025). In addition, popular songs expose learners to natural pronunciation patterns, colloquial expressions, and discourse markers that are often missing from textbook dialogues (Wang et al., 2026). Even so, many educators remain cautious about using pop music in the classroom. Their concerns often relate to cognitive overload, fast delivery, unfamiliar slang, and dense metaphorical language (Rahmat, 2025).

2.2. Lexical Coverage and Comprehension Thresholds

To evaluate the pedagogical suitability of an authentic text more objectively, linguists often use lexical coverage. This refers to the percentage of running words in a text that learners need to recognize to understand it (Nation, 2006). In a seminal study published in *The Canadian Modern Language Review*, Nation (2006) argued that unassisted comprehension of unsimplified written texts typically requires about 98% lexical coverage, corresponding to a vocabulary size of roughly 8,000 to 9,000 word families.

Building on this work, Laufer and Ravenhorst-Kalovski (2010), in their study published in *Reading in a Foreign Language*, proposed two thresholds for second language learners. The optimal threshold is consistent with Nation's 98% coverage level. However, they also identified a minimal threshold of 95% lexical coverage, which can support basic but adequate comprehension when learners know approximately 4,000 to 5,000 word families. Assessing the *folklore* corpus against these empirically supported thresholds is therefore essential for determining its suitability for learners at specific CEFR proficiency levels.

2.3. “Swiftology” and the Intersection of Linguistics and the Creative Economy

The link between Swift's artistry and the creative economy is significant. Journal studies show that Swift employs complex pragmatic strategies, including “disnarration,” to invite listeners into the meaning-making process of her narratives (Harrison & Ringrow, 2022). Other studies have examined her use of deixis, figurative language (metaphor, synecdoche, and paradox), and interpersonal speech functions (Gustsa & Laili, 2022; Siagian & Ma'shumah, 2025). By combining advanced literary techniques with broad commercial appeal, Swift's work reflects a central idea of the creative economy: human creativity and knowledge production are key economic resources (Cong, 2019). To provide a unified theoretical framework for this paper, we define the ‘creative economy’ not merely as a commercial sector but as an ecosystem in which popular artifacts can be transformed into ‘knowledge assets,’ resources with measurable pedagogical and intellectual value. Within this paradigm, the traditional boundary between commercial pop music and ‘high art’ dissolves. Analyzing her lyrics through this lens therefore offers two important benefits: it can support linguistic pedagogy and help explain how a globally influential cultural product creates meaning and value.

While previous research has provided valuable insights into isolated stylistic devices such as figurative language, deixis, and narrative strategies in pop lyrics (Gustsa & Laili, 2022; Siagian & Ma'shumah, 2025), these studies often treat the text strictly as an object of literary critique rather than an applied pedagogical resource. Conversely, while scholars emphasize the efficacy of authentic materials and lexical thresholds, they rarely test these frameworks against complex,

contemporary pop-cultural artifacts. The current research creates a dialogue between these distinct strands. By synthesizing objective lexical profiling, pedagogical interpretation, and qualitative thematic analysis within a single album corpus, this study shifts the focus from mere stylistic description to empirically evaluating the pedagogical affordances of such materials.

2.4. Contextualizing Song-Based Pedagogy in Vietnam

Within the Vietnamese academic context, the linguistic analysis of song lyrics has established a strong foundation. For instance, scholars have extensively analyzed the lyrical structures of Trịnh Công Sơn, demonstrating how his conceptual metaphors and lexical choices enrich the Vietnamese language and reflect deep philosophical concepts (Nguyen, 2015). In parallel, recent empirical studies in Vietnam have highlighted the practical benefits of integrating English pop music into EFL contexts. Researchers such as Lam and Nguyen (2022) and Le (2023) have demonstrated that utilizing English songs significantly enhances listening skills, vocabulary retention, and learning autonomy among Vietnamese EFL learners. Building on this local academic tradition, the current study expands the analytical scope by applying mixed-methods lexical profiling to a contemporary global pop album and assessing its viability for advanced EFL pedagogical frameworks.

3. METHODOLOGY

3.1. Corpus Compilation

The primary dataset consists of a self-compiled corpus of lyrics from the 16 standard tracks of Taylor Swift's *folklore* (2020). The lyrics were sourced from and checked against the official album booklets. Importantly, the corpus reflects the lyrics as they are sung, so repeated choruses and refrains were retained rather than removed. This choice was intentional,

as auditory repetition is central to how songs are experienced and can directly support incidental vocabulary learning. As a result, this natural repetition lowers the overall Type-Token Ratio (TTR). To improve the accuracy of the quantitative profiling, we systematically removed non-lexical paralinguistic markers, such as “ooh” and “ah”, as well as structural annotations, such as [Chorus].

3.2. Quantitative Profiling via Lextutor VocabProfile

The standardized corpus was analyzed using Lextutor VocabProfile, a computational tool developed by Cobb (n.d.). This tool classifies vocabulary into four frequency bands based on the British National Corpus (BNC) and the Academic Word List (AWL) (Coxhead, 2000):

K1 words: the 1,000 most frequent word families in English.

K2 words: the 1,001 to 2,000 most frequent word families.

AWL words: vocabulary items commonly used in academic discourse.

Off-list words: low-frequency words, proper nouns, such as *Peter* and *Wendy*, domain-specific terms, and creative neologisms that do not appear in the previous lists.

The software also calculates lexical density, the ratio of content words to total tokens, and the Type-Token Ratio (TTR), which measures lexical variation.

3.3. Qualitative Thematic Analysis

Rather than applying a *priori* macro-theoretical framework, this study adopted a data-driven, inductive approach to thematic analysis. The thematic categories emerged directly from the lyric corpus through close examination of micro-linguistic features, specifically pragmatics, spatial/temporal deixis, and metaphorical language. The finalized lyric corpus was systematically coded by the four researchers using the qualitative data analysis software QualCoder (Curtain, 2026) based on

an established codebook with four main themes: (1) Memory and Time, (2) Relational Rupture, (3) Complexity and Transgression, and (4) Cinematic Metaphors.

3.4. Data Reliability and Inter-rater Agreement (Fleiss’ Kappa)

To ensure methodological rigor, reproducibility, and objectivity in both corpus preparation and qualitative thematic analysis, inter-rater reliability (IRR) was calculated. Since four independent researchers categorized the data using nominal scales, Fleiss’ Kappa (κ) was selected as the appropriate statistical measure (Laerd Statistics, 2019). The resulting κ values were interpreted using the framework proposed by Landis and Koch (1977), where values between 0.61 and 0.80 indicate substantial agreement, and values between 0.81 and 1.00 indicate almost perfect agreement.

The IRR assessment was conducted in two phases:

Data cleaning phase: A representative sample of 30 lyrical segments was randomly selected from the corpus. The four raters independently classified each segment into three predetermined decision codes: (1) keep, (2) standardize, or (3) remove. The calculation produced a Fleiss’ Kappa value of $\kappa = 0.923$. Based on the Landis and Koch (1977) benchmarks, this indicates almost perfect agreement and confirms the reliability of the standardization protocol before the LexTutor analysis.

Qualitative thematic analysis phase: Based on the subset of 15 highly complex lyrical segments, the thematic coding produced a Fleiss’ Kappa value of $\kappa = 0.91$, which also indicates almost perfect agreement. Minor discrepancies, mainly related to overlapping metaphors, were resolved through collaborative discussion until full consensus was reached.

Table 1. Inter-rater reliability statistics for corpus cleaning and thematic coding

Assessment Phase	Number of Raters	Sample Size (Segments)	Fleiss’ Kappa (κ)	Landis & Koch (1977) Interpretation
Data Cleaning	4	30	0.923	Almost perfect agreement
Thematic Coding	4	15	0.91	Almost perfect agreement

For full transparency and methodological reproducibility, the detailed coding procedure, managed in QualCoder, along with the codebook and sample-coded segments, is provided in the Appendix.

4. QUANTITATIVE RESULTS AND DISCUSSION: LEXICAL PROFILE ANALYSIS

Addressing the first research question (RQ1), the Lextutor VocabProfile results provide a detailed quantitative profile of the *folklore* corpus. This analysis clarifies the linguistic features that shape the album’s accessibility to language learners, particularly

by showing how its lexical coverage aligns with established comprehension thresholds.

4.1. Global Metrics: Lexical Density and Type-Token Ratio

As shown in Table 2, the analysis included 5,032 tokens, 974 distinct word types, and 617 word families. The corpus has a Type-Token Ratio (TTR) of 0.19, with 5.17 tokens per type, indicating relatively high lexical repetition. This pattern is consistent with spoken discourse and popular song lyrics, where choruses, refrains, and melodic repetition naturally create repeated exposure to words (Pavia et al., 2019). From a pedagogical perspective, this recurrence

is particularly useful because repeated exposure to vocabulary in a natural, rhythmic context can support neurocognitive encoding and incidental vocabulary learning (Vivone et al., 2025).

Table 2. Overall lexical profile of the folklore corpus

Metric	Value
Total tokens	5,032
Different words/types	974
Word families	617
Type-token ratio	0.19
Tokens per type	5.17
Lexical density	0.44
On-list tokens	4,786
On-list types	807
Tokens per family	7.76
Types per family	1.31

The corpus has a lexical density of 0.44, meaning that about 44% of its words are lexical content words. This level is close to conversational English, providing EFL learners with familiar grammatical structures while still allowing them to engage with more complex vocabulary.

4.2. Frequency Bands and Comprehension Thresholds

The distribution of vocabulary across standardized frequency bands reveals critical insights into the text’s accessibility.

Table 3. Frequency-band distribution of the folklore corpus

Frequency band	Families	Types	Tokens	Percentage	Cumulative coverage
K1 words (1–1000)	470	643	4,516	89.75%	89.75%
K2 words (1001–2000)	132	148	248	4.93%	94.68%
AWL words	15	16	22	0.44%	95.12%
Off-list words	—	167	246	4.89%	100.00%
Total	617+	974	5,032	100.00%	—

Note. The profiling tool does not assign off-list word families; therefore, the total number of families is reported as 617+.

As detailed in Table 3, K1 words account for 4,516 tokens, or 89.75% of the corpus, while K2 words contribute an additional 248 tokens,

or 4.93%. Together, the 2,000 most frequent English word families provide 94.68% lexical coverage.

For ease of interpretation, Figure 1 summarizes the corpus’s lexical coverage across the four frequency bands.

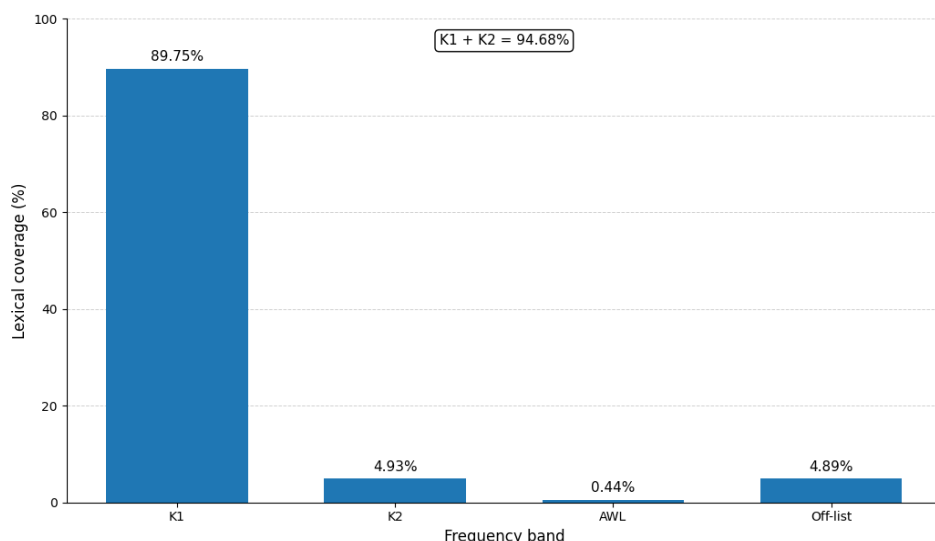


Figure 1. Lexical coverage by frequency band in the folklore corpus

Note. K1 = 1–1000 most frequent word families; K2 = 1001–2000 most frequent word families; AWL = Academic Word List.

As shown in Figure 1, the corpus is strongly dominated by high-frequency vocabulary, with K1 and K2 words together making up 94.68% of total lexical coverage. This pattern suggests the album is generally accessible to intermediate learners while still offering lexical depth through AWL and Off-list items. Laufer and Ravenhorst-Kalovski (2010) proposed that learners need at least 95% lexical coverage to achieve adequate independent comprehension. In this study, the corpus comes very close to that threshold. Therefore, it may suit B1+/B2 learners, particularly when teachers provide scaffolding, glossaries, and guided listening activities. The large proportion of familiar K1 and K2 vocabulary may also lower cognitive load, especially when learners are supported in interpreting metaphorical language and less frequent expressions (Alamri, 2025).

4.3. The Role of AWL and Off-List Vocabulary in Literary Elevation

While K1 and K2 words make the corpus accessible, the remaining 5.33%, which consists

of AWL and Off-list items, gives the album much of its lexical and literary richness. As Table 4 shows, these items are the main source of the corpus’s stylistic sophistication.

The corpus includes 22 Academic Word List tokens, accounting for 0.44% of the total. These include words such as *adjusting*, *assume*, *assumptions*, *couple*, *finally*, *injury*, *integrity*, *invisible*, *mental*, *obvious*, *persist*, *potential*, *seek*, *sphere*, *trace*, and *version*. The presence of academic vocabulary in a commercial pop album is notable, suggesting a deliberate attempt to add philosophical and psychological depth to the narrative. For example, words such as *integrity* and *potential* in songs about relational breakdown point to a level of self-reflection that is less common in mainstream chart music.

Off-list vocabulary makes up a larger share of the corpus, with 246 tokens, or 4.89% of the total. This category includes vivid nouns and creative adjectives that shape the album’s distinctive indie-folk style. Examples include *cardigan*, *ricochet*, *clandestine*, *marvelous*, *masquerade*, *epiphany*, *godforsaken*, *folklore*, *cobblestones*, *dwindle*, *mercurial*, and *sleight*. These low-frequency words serve as important

stylistic markers. From a pedagogical perspective, this 5% margin of unfamiliar but advanced vocabulary may create an appropriate “zone of proximal development.” Learners can draw on the strong contextual support provided

by the 95% familiar vocabulary, along with the music’s emotional tone, to infer and gradually acquire these more advanced lexical items (Vivone et al., 2025).

Table 4. *Selected AWL and off-list items from the folklore corpus*

Category	Selected lexical items	Pedagogical value
AWL items	<i>adjusting, assume, assumptions, integrity, invisible, mental, obvious, persist, potential, seek, sphere, trace, version</i>	<i>Useful for teaching abstract, academic, and reflective vocabulary</i>
Off-list items	<i>cardigan, ricochet, clandestine, marvelous, masquerade, epiphany, godforsaken, cobblestones, dwindle, mercurial, sleight</i>	<i>Useful for teaching literary imagery, metaphor, and low-frequency vocabulary</i>

5. QUALITATIVE RESULTS AND DISCUSSION: THEMATIC ANALYSIS

While quantitative profiling reveals the corpus’s structural accessibility, qualitative analysis is needed to assess its literary value. To address the second research question (RQ2), the thematic coding phase, facilitated by QualCoder, identified four main thematic clusters. Operating through the analytical frameworks of critical discourse analysis (CDA) and pragmatics, the thematic coding moves beyond subjective literary interpretation. The analysis reveals a high concentration of specific linguistic devices, namely conceptual metaphors (Lakoff & Johnson, 1980), spatial and temporal deixis, and complex psychological discourse (Siagian & Ma’shumah, 2025). This section also shows how advanced AWL and Off-list vocabulary are used in context to build these layered narratives.

5.1. Theme 1: Memory and the Passage of Time

The handling of time and memory serves as a central structural device throughout *folklore*. The album often builds its narratives around a contrast between the perceived

certainty of youth and the deeper disillusionment that comes with adulthood.

5.1.1. Youth and Innocence

The album frequently presents youth as a stage that adults tend to underestimate. In “cardigan,” the recurring refrain highlights adolescent marginalization, noting how older generations “assume you know nothing” (Swift, 2020). The generic pronoun “they” frames the adult world as distant, judgmental, and almost unified in dismissing young people. However, this assumption is immediately countered by the speaker’s retrospective certainty, as recalled through intimate visual markers such as “dancing in your Levi’s” and being “drunk under a streetlight” (Swift, 2020). The tension between what others assume and what the speaker knows from lived experience creates a powerful defense of adolescent emotional truth.

A similar return to childhood appears in “seven,” where memory is operationalized through sensory spatial deixis, inviting the listener to “picture me in the trees” and recalling being “too scared to jump in” the creek at age “seven” (Swift, 2020). Spatial deixis brings listeners into a specific, seemingly untouched

past. It highlights a form of innocence that exists before the child fully enters the rules and expectations of adult “civility.”

5.1.2. *Nostalgia and Regret*

As time passes, nostalgia in *folklore* often becomes mixed with regret. In “the 1,” the speaker evaluates unfulfilled counterfactual pasts through conditional phrasing, musing that it “would have been fun” if the partner had “been the one” (Swift, 2020). This structure points to the permanence of past choices. The condition can no longer be met, making the imagined outcome impossible.

In “August,” temporal transition is captured through metaphor, where the summer month dynamically “slipped away” and is poignantly “sipped away like a bottle of wine” to signify fleeting intimacy (Swift, 2020). By presenting the month as something that can “slip” and be “sipped” away, the lyrics turn time into an active force. This image conveys the fragile illusion of owning a summer romance, as well as the unavoidable loss of the present moment.

From a pedagogical perspective, this theme gives learners a useful way to explore how temporal deixis, conditional structures, and metaphorical personification create deeper narrative meaning.

5.2. Theme 2: Relational Rupture and Emotional Trauma

The second thematic cluster moves away from conventional pop breakup tropes. Instead, it uses more complex metaphorical frameworks, particularly geopolitical displacement, warfare, and haunting, to express the trauma of broken emotional bonds (Gustsa & Laili, 2022).

5.2.1. *Territorial Displacement and Exile*

In “exile,” the end of a relationship is framed through the conceptual metaphor of geopolitical displacement. By utilizing spatial

deixis to establish an ingroup/outgroup dynamic, the language of borders, belonging, and displacement is operationalized. The lyrics employ geopolitical terminology, shifting from the intimate spatial deixis of a “homeland” and “town” to the isolated state of “exile” (Swift, 2020). By comparing a romantic partner to a “homeland,” the song transforms private heartbreak into a crisis of identity and belonging. The speaker is left emotionally stateless, pushed out of the place that once offered safety and attachment.

This spatial alienation recurs in “my tears ricochet,” where the speaker possesses absolute freedom of movement “anywhere I want” yet remains structurally barred from returning “home” (Swift, 2020). Here, freedom of movement becomes painfully empty because the speaker is still barred from the one place that matters most. The line suggests that emotional loss is not only about distance, but also about being unable to return to a former sense of home.

5.2.2. *Anger and Retribution*

In contrast to the album’s more reflective moments, *folklore* also engages directly with feminine anger. In “mad woman,” the speaker challenges the patriarchal policing of female affect by subverting pejorative labeling, declaring that provocations regarding being “crazy” or “angry” only fuel the resistance of a “mad woman” (Swift, 2020). The lyrics take words such as “crazy” and “mad,” which are often used to dismiss women, and turn them into signs of resistance.

The imagery then shifts to an overtly hostile register, using combative metaphors in which speech is transformed into “breathe flames” and retaliation is externalized as “cannons all firing” (Swift, 2020). These metaphors present speech itself as an act of

attack and self-defense. A similar transformation occurs in “my tears ricochet,” where passive sorrow becomes an active force, forcing the antagonist to experience the emotional “ricochet” of the speaker’s tears (Swift, 2020). The tears are no longer passive signs of sadness. Instead, they become symbolic projectiles that return to wound the person responsible for the harm. In this way, the speaker resists silent victimhood and reclaims emotional power.

In the EFL classroom, this thematic cluster offers a rich context for examining combative metaphors and the pragmatic use of spatial deixis to express emotional distance.

5.3. Theme 3: Relational Complexity and Transgression

The corpus explores complex moral ambiguities, presenting relationships shaped by secrecy, fate, and the difficult process of making amends.

5.3.1. *Illicit and Secret Love*

“Illicit affairs” offers a sharp and unromantic portrayal of infidelity. The song uses specific, low-frequency vocabulary to show the slow decay caused by secrecy. The lyrics utilize precise imperative syntax to document the exhausting behavioral protocols of maintaining deception, instructing the subject to “keep your eyes down” and “leave no trace behind” to obscure the narrative (Swift, 2020). The repeated imperative verbs, such as “make sure,” “keep,” and “leave,” create pressure and control. They also show how emotionally tiring it is to maintain the deception.

Rather than romanticizing the transgression, the song depicts incremental emotional attrition through iterative syntax, emphasizing how the relationship structurally “dies... a million little times” (Swift, 2020). This phrasing quantifies emotional attrition. Rather than depicting a singular event,

the iterative syntax highlights a gradual psychological accumulation of distress.

5.3.2. *Invisible Connections and Fate*

By contrast, “invisible string” explores love through the idea of predestination. The text employs specific color symbolism and spatial markers to establish predestination, indexing historical loci such as “Centennial Park” and mundane items like a “teal” shirt to map out early parallel tracks (Swift, 2020). These details make fate feel concrete and personal, as if ordinary places and colors had always been quietly leading the speakers toward each other.

The narrative achieves cosmic inevitability through a rhetorical question that prompts the listener to contemplate a predetermined “invisible string” that is structurally “tying you to me” (Swift, 2020). The rhetorical question functions as a pragmatic marker of realization. Within this conceptual metaphor, romantic attachment is not presented as an active pursuit, but as a predetermined spatial alignment between the subjects.

5.3.3. *Effort and Atonement*

The album also examines the difficult reality of accountability. In “this is me trying,” the speaker mitigates interpersonal distance through a vulnerable declarative statement, simply asserting that “this is me trying” to acknowledge accountability (Swift, 2020). The plainness of the line is striking, especially when placed beside the album’s more elaborate metaphors. Its emotional weight comes from its directness.

A similar struggle appears in “betty,” where remorse is framed through verbal repetition, contrasting past transgression, “the worst thing that I ever did,” with the present desire to “make it up to you” (Swift, 2020). The repetition of “did” and “do” reflects the speaker’s circular thinking, as guilt turns into a desperate desire for repair.

From a teaching perspective, these tracks provide authentic examples of how imperative syntax, color symbolism, and rhetorical questions can express urgency, fate, and emotional accountability.

5.4. Theme 4: Artistic and Cinematic Metaphors

A key argument for the album's value within the creative economy lies in its strong intertextuality and reflexivity. The lyrics draw on the language of cinema and the performing arts to show how memory, identity, and social relationships are often shaped through performance.

5.4.1. Cinematic Projections

The album uses film metaphors throughout to convey emotional distance and fatalism. In "exile," the recurring meta-cinematic phrase "seen this film before" implies that the speaker experiences interpersonal conflict as a predetermined script with an unappealing "ending" (Swift, 2020), suggesting that the speaker experiences the failing relationship as if it were a scripted narrative rather than a fully controllable lived event. The patterns of conflict feel familiar, almost like recognizable plot devices, which makes the painful ending seem unavoidable.

In "the 1," unfulfilled romantic potential is conceptualized as an unproduced masterpiece, asserting that the "greatest films of all time were never made" (Swift, 2020). Here, the "film" becomes a metaphor for a life or relationship that could have become meaningful but never fully came into being. "This is me trying" extends this cinematic register through analog imagery, objectifying memory as an unyielding "flashback in a film reel" trapped on a single local screen (Swift, 2020). This metaphor presents memory as something

projected repeatedly before the speaker. It also suggests that the speaker is trapped as a captive viewer, forced to revisit past pain and failure.

5.4.2. Fictional Personas and Theatricality

The album also uses theatrical metaphors to examine the emotional strain of maintaining a public identity. In "mirrorball," the speaker adopts a highly malleable performative persona, metaphorically equating herself to a "mirrorball" that can "change everything" to achieve social assimilation (Swift, 2020). This image captures the pressure to constantly adapt to remain visible and accepted.

The narrative amplifies this performative precarity via theatrical metaphors, depicting continuous effort as balancing on a "trapeze" to maintain the audience's gaze (Swift, 2020). The trapeze suggests both performance and danger. Public affection becomes unstable, and the speaker's fear of being ignored or forgotten gives the image its emotional force.

In "cardigan," the text introduces intertextual depth through a classical literary allusion, framing the inevitability of separation as a futile attempt to alter the script of "Peter losing Wendy" (Swift, 2020). This allusion to *Peter Pan* frames the speaker's experience through a familiar cultural story of loss and separation. By trying to "change the ending," the speaker acknowledges a desire to rewrite a tragic narrative that already seems deeply embedded in cultural memory.

From a pedagogical perspective, this reflexivity allows advanced learners to engage with intertextuality and meta-narrative references. It also helps develop the critical literary competence needed in modern knowledge economies.

Table 5. *Thematic architecture and primary linguistic drivers within the folklore corpus.*

Thematic Category	Conceptual Framework	Key Linguistic / Pragmatic Devices	Representative Tracks
Memory & Time	Time as an active agent; childhood as epistemic clarity.	Temporal deixis, third conditionals, personification.	<i>August, seven, the 1</i>
Relational Rupture	Betrayal as geopolitical exile and physical warfare.	Spatial deixis, combative metaphors, AWL integration.	<i>exile, my tears ricochet</i>
Complexity & Transgression	Infidelity as decay; fate as an unseen tether.	Spatial deixis, combative metaphors, AWL integration.	<i>illicit affairs, invisible string</i>
Cinematic Metaphors	Memory as film projection; identity as circus performance.	Meta-narrative references, intertextuality, theatrical lexicon.	<i>mirrorball, hoax</i>

6. IMPLICATIONS, APPLICATIONS, AND FUTURE DIRECTIONS

6.1. Interdisciplinary Implications for the Creative Economy

Within this macroeconomic framework, *folklore* operates not merely as a commercial cultural product, but as a site of linguistic capital production (Bourdieu, 1991). Human Capital Theory (Florida, 2002) can unify the structural link between lexical complexity and the creative economy. In a knowledge-driven creative economy, advanced lexical profiling, characterized by high densities of Academic Word List (AWL) items and sophisticated off-list vocabulary, is not merely a stylistic trait; rather, it serves as a critical pedagogical vehicle for developing high-level cognitive flexibility, linguistic competence, and conceptual literacy among workforce learners. Repurposing a globally resonant popular artifact into an ESL/EFL classroom environment actively operationalizes this link. By transforming a commercial pop-culture text into an educational ‘knowledge asset,’ the curriculum enhances learners’ linguistic capital,

directly contributing to the human capital infrastructure that drives innovation in contemporary creative industries (Cong, 2019). When a globally influential pop artist moves beyond formulaic commercial writing and creates a dense literary-linguistic hybrid, it can reshape the cultural landscape in important ways (McCann et al., 2025; Sherman, 2024).

This repositioning of pop music as a form of high art can also generate wider economic and academic value. It encourages secondary academic activities, including conferences, peer-reviewed linguistic studies, and university courses, which further place the artist’s work within institutional settings (Driessen, 2026). From a policy perspective, artists such as Swift show how complex Off-list vocabulary, such as *clandestine* and *mercurial*, and sophisticated narrative devices, such as disnarration and deixis, can be woven into commercially successful and highly accessible cultural products. This suggests that broad market appeal does not require intellectual or linguistic simplification. Developing these high-value, interdisciplinary creative competencies therefore aligns closely with the mission of

institutions seeking to strengthen cultural industries (Vietnam National University, Hanoi, 2025).

6.2. Pedagogical Applications in Advanced ESL/EFL Contexts

Synthesizing the quantitative and qualitative findings in response to the third research question (RQ3), the relationship between lexical accessibility and thematic depth in *folklore* supports a shift in pedagogical practice. The findings suggest that culturally relevant authentic materials can be integrated into advanced language classrooms as effective, evidence-based tools for language learning within the wider creative economy (Alamri, 2025).

6.2.1. Incidental Vocabulary Acquisition through Rhythmic Repetition

The corpus has a Type-Token Ratio of 0.19, which indicates a high level of word recurrence. In an EFL context, repeated exposure to Off-list vocabulary and AWL terms within the emotional and melodic structure of songs may help lower learners' affective filter (Schwarz, 2013). Teachers can use tracks such as "illicit affairs" to introduce advanced vocabulary, while allowing students to infer meaning from the strong support provided by the surrounding 95% K1/K2 vocabulary (Laufer & Ravenhorst-Kalovski, 2010).

6.2.2. Pragmatics and Discourse Analysis

The album is also a useful resource for teaching advanced pragmatics, especially spatial and temporal deixis. Instructors can use songs such as "exile" and "seven" to show how pronouns, including *I*, *you*, and *they*, and locational markers, such as *homeland*, *town*, *trees*, and *creek*, help create psychological distance, intimacy, or alienation (Siagian & Ma'shumah, 2025). This moves learners beyond

memorized grammar and toward a deeper understanding of communicative intent.

6.2.3. Enhancing Literary and Cultural Competence

By examining cinematic and theatrical metaphors in songs such as "mirrorball" and "the 1," advanced learners can engage in literary analysis often associated with classical texts. These intertextual references also help develop intercultural awareness and critical thinking skills, which are increasingly important in the modern knowledge-based creative economy (Sadiqzade, 2025).

6.2.4. Proposed Task-Based Pedagogical Framework

To operate these theoretical concepts without inducing cognitive overload, educators can adapt these tracks into a Task-Based Language Teaching (TBLT) framework. For instance, using the track "exile":

Pre-listening (Lexical Scaffolding):

Instructors introduce the spatial deixis and AWL vocabulary (e.g., *exile*, *homeland*, *defending*) to establish the conceptual metaphor of territorial loss.

While-listening (Guided Inference):

Learners map the shifting pronouns (*I*, *you*) to track the psychological distance between the two speakers in the song.

Post-listening (Productive Application):

Advanced learners analyze how the geopolitical metaphors emphasize relational trauma and apply these structures in their own writing or debate tasks.

While this framework provides a concrete roadmap for classroom integration, subsequent empirical studies are necessary to quantify its impact on learner retention and engagement.

6.3. Limitations and Future Research

Although the proposed TBLT framework underscores the album's practical utility, these pedagogical applications must be evaluated against several inherent limitations of the current study. First, the dataset is deliberately limited to a single 16-track corpus. While this limits the statistical generalizability of the findings to contemporary pop music as a whole, it serves as an exploratory case study. In applied linguistics research employing detailed thematic analysis, highly contextualized micro-corpora are often necessary to achieve the analytical depth required to examine complex pragmatic markers and implicit metaphors (Milton, 2009). Second, although the qualitative thematic analysis was supported by a very high inter-rater reliability score ($\kappa = .91$), interpreting literary and cinematic metaphors still involves some degree of subjectivity. Finally, although the pedagogical recommendations are grounded in established lexical coverage thresholds for B1+ learners, they remain theoretical in this study.

Future research should address these limitations by extending the analysis to the artist's full discography through a diachronic corpus approach. This would help trace how linguistic complexity has developed across modern commercial songwriting over time. In addition, empirical studies are needed to test the pedagogical value of these materials in real classroom settings. For example, experimental research could measure incidental vocabulary acquisition and retention among advanced EFL learners, particularly those preparing for high-stakes standardized proficiency exams. Such studies would provide stronger evidence for the effectiveness of authentic pop music materials in language learning contexts.

7. CONCLUSION

This interdisciplinary analysis shows that Taylor Swift's *folklore* goes beyond the usual

linguistic patterns of mainstream popular music and can be understood as a sophisticated literary and linguistic corpus.

In response to RQ1, the study applied Corpus Linguistics methods, supported by an almost perfect inter-rater reliability score ($\kappa = .91$ to $.923$). The quantitative profiling showed that K1 and K2 vocabulary together accounted for 94.68% of the corpus. This level of lexical coverage closely approaches the minimal comprehension threshold proposed by Laufer and Ravenhorst-Kalovski (2010), making the album broadly accessible to B1+ learners.

For RQ2, the qualitative thematic analysis showed that the album's depth comes largely from its use of Academic Word List (AWL) items and a wide range of Off-list vocabulary. These advanced lexical choices help build complex narratives shaped by temporal distortion, geopolitical metaphors of exile, and cinematic reflexivity. Through these devices, the album explores memory, trauma, moral ambiguity, and emotional loss.

In response to RQ3, the findings support the pedagogical use of high-quality pop cultural texts in advanced ESL/EFL contexts. Materials such as *folklore* can create useful opportunities for incidental vocabulary acquisition and advanced pragmatic analysis. They also show how authentic cultural products can function as valuable educational resources. Overall, recognizing the linguistic complexity of selected pop music texts can expand the range of authentic materials available for advanced language instruction. This approach not only enriches ESL/EFL pedagogy but also highlights the practical and interdisciplinary value of artistic production within the global creative economy.

APPENDIX: DATA PREPROCESSING PROTOCOL AND THEMATIC CODEBOOK

Table A. Data Cleaning Protocol and Decision Matrix (Phase 1)

Decision Code	Operational Criteria for Raters	Raw Lyric Example	Processed Output
1. Keep	Valid English lexical tokens and standard syntactic structures.	<i>"I knew you dancing in your Levi's"</i>	<i>I knew you dancing in your levis</i>
2. Standardize	Remove punctuation, convert to lowercase, and resolve non-standard contractions to ensure accurate Lextutor categorization.	<i>"Cause I knew you'd step on the glass"</i>	<i>because I knew you would step on the glass</i>
3. Remove	Non-lexical paralinguistic vocalizations (e.g., ooh, ah) and structural song annotations.	<i>"[Chorus] Ooh, I thought I've seen this film before"</i>	<i>I thought I had seen this film before</i>

Note. This protocol standardized the raw lyric files before Lextutor processing to ensure valid tokenization.

Table B. Qualitative Thematic Codebook and Sample Applications (Phase 2)

Thematic Category	Operational Definition & Key Markers	Sample Lyrical Segment & Rater Rationale
1. Memory and Time	Focus: Contrast between childhood epistemic certainty and adult disillusionment; nostalgia. Markers: Temporal deixis, third conditionals, personification.	<i>"August slipped away into a moment in time... sipped away like a bottle of wine." (August)</i> Rationale: Employs personification to manipulate the passage of time and articulate regret.
2. Relational Rupture	Focus: Trauma of severed connections and feminine rage, avoiding standard pop clichés. Markers: Geopolitical displacement (exile, borders), warfare lexicon.	<i>"You are not my homeland anymore... now I am in exile, seeing you out." (exile)</i> Rationale: Utilizes spatial deixis to frame emotional trauma as physical, territorial displacement.
3. Complexity & Transgression	Focus: Morally ambiguous dynamics, infidelity mechanics, and cosmic predestination. Markers: High density of imperative verbs, color symbolism.	<i>"Make sure nobody sees you leave / Hood over your head... leave no trace behind." (illicit affairs)</i> Rationale: Imperative syntax simulates the exhausting, rigid mechanics of maintaining secrecy.
4. Cinematic Metaphors	Focus: Performative nature of identity and viewing personal history through fictional media. Markers: Theatrical/circus lexicon, analog film terminology.	<i>"You know the greatest films of all time were never made." (the 1)</i> Rationale: Uses meta-narrative film terminology to express unfulfilled relational potential.

Note. This codebook was used by four independent raters in the Qual Coder environment to categorize a 15-segment reliability sample, yielding a Fleiss' Kappa of $\kappa = 0.91$.

REFERENCES

- Alamri, W. (2025). Evaluating the benefits and challenges of using authentic materials in EFL context for listening purposes. *Frontiers in Education*, 10, Article 1611308. <https://doi.org/10.3389/feduc.2025.1611308>
- Bourdieu, P. (1991). *Language and symbolic power* (J. B. Thompson, Ed.; G. Raymond & M. Adamson, Trans.). Harvard University Press.
- Cobb, T. (n.d.). *The complete lexical tutor*. <https://www.lex tutor.ca/>
- Cong, P. T. (2019). Creative industries, investment trends and their impact on the economy and financial sectors: A study from Vietnam. *International Journal of Innovation, Creativity and Change*, 7(5), 31–49. <https://doi.org/10.53333/IJICC2013/07504>
- Coxhead, A. (2000). A new academic word list. *TESOL Quarterly*, 34(2), 213–238. <https://doi.org/10.2307/3587951>
- Curtain, C. (2026). *QualCoder* (Version 3.6) [Computer software]. GitHub. <https://github.com/ccbogel/QualCoder/releases>
- Driessen, S. (2026). Invisible string(s): Exploring conspiratorial participatory practices among Swifties. *Popular Communication*. Advance online publication. <https://doi.org/10.1080/15405702.2026.2617543>
- Florida, R. (2002). *The rise of the creative class: And how it is transforming work, leisure, community and everyday life*. Basic Books.
- Gustsa, F., & Laili, E. N. (2022). An analysis of figurative language in Taylor Swift's folklore album. *Teaching English as a Foreign Language, Literature and Linguistics*, 2(2), 11–26. <https://doi.org/10.33752/teflics.v2i2.3434>
- Harrison, C., & Ringrow, H. (2022). Disnarration and the performance of storytelling in Taylor Swift's folklore and evermore. *The International Journal of Literary Linguistics*, 11(1). <https://doi.org/10.15462/ijll.v11i1.140>
- Kendong, F. A., Daud, A. S., & Joharry, S. A. (2023). A corpus-driven analysis of Taylor Swift's song lyrics. *International Journal of Modern Languages and Applied Linguistics*, 7(2). <https://doi.org/10.24191/ijmal.v7i2.7489>
- Laerd Statistics. (2019). *Fleiss' kappa in SPSS Statistics*. <https://statistics.laerd.com/spss-tutorials/fleiss-kappa-in-spss-statistics.php>
- Lam, T. N. D., & Nguyen, D. K. (2022). English foreign language teachers' perceptions on utilizing songs for teaching English listening skills to young learners: A case study in Tra Vinh Province, Vietnam. *Tra Vinh University Journal of Science*, 47, 39–46.
- Landis, J. R., & Koch, G. G. (1977). The measurement of observer agreement for categorical data. *Biometrics*, 33(1), 159–174. <https://doi.org/10.2307/2529310>
- Laufer, B., & Ravenhorst-Kalovski, G. C. (2010). Lexical threshold revisited: Lexical text coverage, learners' vocabulary size and reading comprehension. *Reading in a Foreign Language*, 22(1), 15–30. <https://doi.org/10.64152/10125/66648>
- Le, T. (2023). Practice beyond the classroom: Vietnamese university students' perception of using songs to learn English. *FOSTER: Journal of English Language Teaching*, 4(4), 218–230. <https://doi.org/10.24256/foster-jelt.v4i4.145>
- McCann, H., Faichney, E., Trelease, R., & Whatman, E. (Eds.). (2025). *Taylor Swift: Culture, capital, and critique*. Routledge. <https://doi.org/10.4324/9781003509271>

- Milton, J. (2009). *Measuring second language vocabulary acquisition*. Multilingual Matters. <https://doi.org/10.21832/9781847692092>
- Ministry of Culture, Sports and Tourism, & Vietnam Institute of Culture and Arts Studies. (2013). *National strategy for the development of cultural industries in Vietnam to 2020, vision 2030: Draft version*. British Council.
- Murphy, I., & Thompson, A. P. (2026). ‘Call it what you want’: Ethical value in the lyrics of Taylor Swift and its educational application for betterment. *Journal of Moral Education*. Advance online publication. <https://doi.org/10.1080/03057240.2026.2630330>
- Nation, I. S. P. (2006). How large a vocabulary is needed for reading and listening? *The Canadian Modern Language Review*, 63(1), 59–82. <https://doi.org/10.3138/cmlr.63.1.59>
- Nguyen, T. B. H. (2015). *Ấn dụ tri nhận trong ca từ của Trịnh Công Sơn*. Nhà xuất bản Khoa học Xã hội.
- Patel, A. D. (2003). Language, music, syntax and the brain. *Nature Neuroscience*, 6(7), 674–681. <https://doi.org/10.1038/nn1082>
- Pavia, N., Webb, S., & Faez, F. (2019). Incidental vocabulary learning through listening to songs. *Studies in Second Language Acquisition*, 41(4), 745–768. <https://doi.org/10.1017/S0272263119000020>
- Rahmat, H. (2025). Pop songs as a creative learning tool to improve students’ listening abilities. *Journal of English Language Proficiency*, 2(2), 51–63. <https://doi.org/10.36312/jelap.v2i2.5686>
- Sadiqzade, Z. (2025). Melodies in the classroom: Using songs to enhance English language learning. *Acta Globalis Humanitatis et Linguarum*, 2(3), 126–133. <https://doi.org/10.69760/aghel.0250020018>
- Schwarz, M. (2013). Learning with Lady Gaga & Co.: Incidental EFL vocabulary acquisition from pop songs. *VIEWS*, 22, 17–48.
- Sherman, M. (2024, December 9). *Taylor Swift’s Eras Tour ends by shattering own record, grossing an estimated \$2.2B, Pollstar says*. Associated Press.
- Siagian, T. L., & Ma’shumah, N. K. (2025). Beyond the context: Exploring the use of deixis in Taylor Swift’s Folklore album. *Deskripsi Bahasa*, 8(2), 104–117. <https://doi.org/10.22146/db.19148>
- Swift, T. (2020). *Folklore* [Album]. Republic Records.
- The Work Foundation. (2007). *Staying ahead: The economic performance of the UK’s creative industries*. Department for Culture, Media and Sport.
- Vietnam National University, Hanoi. (2025, January 3). *VNU’s 2024 highlights*.
- Vivone, A., Rumlich, D., & Lehmann-Wermser, A. (2025). Maximizing linguistic and cognitive potential: The role of vocal training with pop songs in teaching English as a foreign language. *Learning: Research and Practice*, 11(1), 74–91. <https://doi.org/10.1080/23735082.2025.2501043>
- Wang, W., Zheng, L., & Zhang, J. (2026). Effect of online digital storytelling on the comprehension of authentic listening materials and engagement of junior high school EFL learners. *Scientific Reports*, 16, Article 6639. <https://doi.org/10.1038/s41598-026-36913-4>