

From Prompt to Platform: Agency in a Suno Workflow for Taigi Pop Music

Dong-Feng Guo

National Cheng Kung University, Institute of Creative
Industries Design
Tainan, Taiwan (R.O.C.)
l16111110@gs.ncku.edu.tw

Wei-Chi Chien

National Cheng Kung University, Institute of Creative
Industries Design
Tainan, Taiwan (R.O.C.)
chien@xtdesign.org

Abstract

Prompt-based AI music can appear to move directly from user text to model output. This paper asks how agency is formed across written requests, platform interfaces, generated songs, and culturally informed listening. Agency means making a difference within the workflow, not possessing intention. Drawing on music mediation, human-machine agency, platformization, and text-to-music research, the study examines four Suno 5.5 runs that produced eight Taigi Pop Music (Taiwanese-language pop music) songs. The case matters because Taigi Pop Music is recognized through language, vocality, arrangement, genre memory, social position, and local meaning rather than a language label alone. Comparisons of paired songs generated from the same visible settings show shared features alongside audible differences. They also show that complete lyrics can coexist with different surrounding requests and musical associations. Although these comparisons cannot identify internal causes, they reveal how musical requests are organized, made audible, compared, and interpreted across the platform workflow.

1 PROMPT TEXT IN A PLATFORM-BASED AI MUSIC WORKFLOW

When a creator attempts to make Taigi Pop Music through a system such as Suno, the creative act is not simply a matter of writing one prompt and receiving one song. Prompt text matters, but it is only one step in a workflow that also includes lyrics and style input areas, tags, metatags, interface defaults, generated audio, and the creator’s decisions to accept, reject, revise, or regenerate a song. The paper therefore asks how musical agency is formed across prompt text, platform input areas, generated sound, and listening judgment.

1.1 Suno as a Platform Workflow for Taigi Pop Music

Existing text-to-music and Suno-related studies make Suno a suitable case because they show that prompt-based generation can be studied as a platform workflow, not only as a technical problem of generating audio from text. The text-to-music literature gives this paper its baseline: natural-language descriptions can guide generated music, and generated audio can be evaluated for how well it fits those descriptions (Agostinelli et al., 2023). Copet et al.’s MusicGen paper adds a more specific point: what it calls controllability is the practical question of whether users can steer musical results through text descriptions and, in MusicGen’s case, melody inputs, rather than merely receive high-quality audio (Copet et al., 2023). For this study, however, the central object is not the model alone. The central object is the workflow in which a creator writes prompt text, works with lyrics and style input areas, uses tags or metatags, hears generated audio, and decides what counts as an acceptable musical result. The Suno-Caps dataset treats prompts, alignment comments, and emotion annotations as data for studying Suno-generated

songs (Civit et al., 2024). Casini et al. show that commercial text-to-music platforms organize creation through prompt text, lyrics, tags, metatags, and interface metadata (Casini et al., 2025). These are not cited here as examples of similar products; they justify the methodological choice to treat Suno as a concrete site where music mediation through a platform can be observed without claiming access to hidden training data. The case is therefore specific: what happens when this platform workflow is used to produce Taigi Pop Music, a musical formation whose identity cannot be reduced to a genre label or a language tag?

Taigi Pop Music is an especially strong case because it asks the platform to handle language, vocality, arrangement, genre memory, and local meaning at the same time. Tsai, Ho, and Jian’s introduction to *Made in Taiwan* frames Taiwanese popular music through language politics, industrial history, and multiple trajectories rather than as a single authentic sound (Tsai et al., 2019). Ho Tung-Hung’s chapter on postwar nativism connects popular music to changing social moments, cultural nativism, and identity politics without treating nativism as a fixed substance (Ho, 2019). Chen Pei-Feng’s work on Taigi songs foregrounds colonial and postcolonial histories, vocality, and social memory (Chen, 2019a). Chen Pei-Feng’s study of Japanese-period Taigi pop songs shows that the category is historically mediated rather than simply linguistic (Chen, 2017). Chen Pei-Feng’s work on “Taiwanese flavor” further shows how ideas of local sound are historically contested and reclassified (Chen, 2022a). The case is therefore more than a question of whether Suno can produce intelligible Taigi lyrics.

This is why the case matters beyond whether a single platform succeeds or fails. Commercial text-to-music

platforms organize musical requests through prompt text, labels, tags, metatags, and style descriptions. When such requests invoke musical memory, tradition, or local sound, researchers and music practitioners need ways to ask what those simplified terms preserve, what they reduce, and what materials are needed before such judgments can be supported. Because this study has no access to Suno's training data, these questions must be studied through observable case records, generated audio, public platform pages or screenshots, and listening notes rather than assumed in advance.

1.2 Agency as Making a Difference Across Prompt Text, Generated Songs, and Listening

In this paper, agency means the capacity to make a difference in the musical workflow. It does not require human-like intention. Bates makes this distinction when showing how musical objects and technical arrangements can change what happens in practice (Bates, 2018). Prompt text names musical directions. The interface organizes how lyrics, style description, and other settings are entered. Generated songs introduce vocal, instrumental, and formal features that were not fully settled in writing. The first author's listening notes identify which of those features can be compared and interpreted. Taigi Pop Music scholarship provides the historical and musical context for those judgments.

The paper's central claim is therefore:

> Prompt-based AI music should be studied as a workflow rather than a direct path from prompt to output. In this case, the analysis traces agency across the text and visible settings recorded for the four runs, Suno's input areas, the generated songs, and the first author's listening judgments. These parts do not exercise the same control: they make different kinds of difference as a musical request is written, organized, made audible, and interpreted. Music mediation, human-machine agency, platformization, text-to-music research, and Taigi Pop Music scholarship help explain those differences.

The literature supports this claim conceptually. Section 4 places written materials, paired songs, and listening notes together; Section 5 asks what they show about agency and Taigi recognizability.

2 MUSIC MEDIATION, CREATIVE INTERFACES, AND PLATFORM CONDITIONS

This section builds the conceptual path behind the claim in Section 1.2, moving from music mediation to AI music tools, platformization, and text-to-music systems. Together, these bodies of work justify studying prompt text, interface, generated songs, and listening as one workflow of musical action.

2.1 Prompt Text as a Mediator of Musical Intention

Music mediation theory helps prevent a common simplification in discussions of AI music: the idea that technology either merely carries human intention or replaces it. Born and Barry describe mediation as transmission, translation,

and transformation across relations such as sound, technologies, discourses, social relations, sites, and spaces (Born & Barry, 2018). They also orient this paper methodologically: trace which sonic, discursive (the words and categories used to describe music), embodied, material, and social mediations are present, then examine their relations without assuming that discourse is a full guide to practice (Born & Barry, 2018). This matters because prompt text is a discourse about music, while generated audio and the work around it are musical, technical, and listening practices that may not obey that discourse. Hennion adds that musical works appear through mediations such as gestures, bodies, habits, materials, spaces, languages, institutions, and listeners' attachments, not apart from them (Hennion, 2012). Latour makes this usable as an account of agency: actors are made to act by many others, and mediators transform, translate, distort, and modify what passes through them (Latour, 2005). Bates's organological extension of ANT brings the argument down to musical matter: instruments, technical objects, spaces, and users enter a changing give-and-take in which agency means making a difference, not possessing intention (Bates, 2018). Prior's discussion of glitch shows why contemporary music sociology needs this kind of account, since musical styles and broader musical settings are shaped through technologies, devices, techniques, and artefacts that solidify and transform practice (Prior, 2008).

For this paper, the practical implication is direct: prompt text is not a transparent vessel for intention. It is a mediator. It translates a musical desire into language, and then that language is translated again by interface input areas, platform categories, the audible features of generated songs, and the listener's evaluation. The resulting song is therefore not simply "what the user asked for" or "what the model made." It is the outcome of a chain of transformations. This is why the method keeps prompt records together with the separation of lyrics and style input areas, the tags and metatags supplied, the songs kept or rejected, and the listening frame used to judge the result.

2.2 AI Music Tools as Cycles of Writing, Listening, and Revision

AI music tool scholarship makes this account of agency in the workflow more concrete. Fiebrink and Caramiaux sharpen this point by treating machine learning algorithms as creative tools and interfaces, not only as engines to be measured by accuracy (Fiebrink & Caramiaux, 2018). Their examples of interactive machine learning show musicians training a system, trying it in real time, noticing unexpected behavior, and changing examples so the system better fits a developing musical idea (Fiebrink & Caramiaux, 2018). For a Suno study, this shifts attention to the user's cycle of writing prompt text, listening, selecting, revising, and regenerating. Deruty et al. make a parallel point from contemporary popular music production: in studio practice, composing is often intertwined with recording, editing, mixing, sound design, arrangement, and production judgment (Deruty et al., 2022). This helps explain why Suno-generated songs can be treated as audio objects that still have to be compared, accepted, rejected,

and revised in a production-like workflow. Xambó and Roma help specify agency questions for this interaction (Xambó & Roma, 2024). Their live-coding study does not map directly onto Suno, but it asks what is legible to users, what can be modified, how predictable the system is, and how many human and nonhuman agents are involved (Xambó & Roma, 2024). In a platform workflow, those questions become: what can the creator see, what can be changed, what unexpected features the generated song introduces, and which human, data, interface, and system elements participate in the action. Väkevä and Partti likewise show that generative AI music raises questions of agency within concrete educational and musical settings (Väkevä & Partti, 2025).

This matters for the Suno case because popular music is not only melody, harmony, or lyrics. It is also vocal delivery, production texture, arrangement, genre expectation, and selection. A creator working with Suno may take part in the music-making process not by directly composing every note, but by choosing descriptors, editing lyrics, listening for fit, rejecting generated songs, and deciding which audible features matter. These actions are not secondary to the music; they are part of the musical workflow. The analysis therefore asks not only whether Suno “understands” a prompt, but how a creator’s musical judgment works across interface inputs, generated audio, comparison, rejection, and the notes that explain those decisions.

2.3 Platformization as Conditions for Musical Production

Platformization research is necessary here because cultural production is increasingly organized by platforms. Poell et al. define platformization as a transformation of cultural production through platform infrastructures, datafication, governance principles, technological standards, and changing relations between producers, audiences, advertisers, and platform corporations (Poell et al., 2017). In more direct terms, this means that platforms shape cultural production through technical arrangements, data collection, governance rules, interface design, and economic relations. Their account matters here because a platform does not merely host cultural content; it reorganizes production and circulation through interface arrangements, metrics, defaults, and rules (Poell et al., 2017). Seaver’s work on knowing algorithms adds a methodological caution: cultural researchers should study algorithmic systems rather than imagine a single algorithm waiting to be revealed (Seaver, 2019a). Seaver also argues that cultural details are technical details when the object is an algorithmic system, which supports treating prompt input areas, platform categories, user habits, and the records kept while working with the system as part of the system rather than as external context (Seaver, 2019a). Seaver’s account of captivation further connects algorithmic systems to attention and cultural attachment (Seaver, 2019b). Prey and Esteve-Del-Valle’s study of Spotify shows that musicians perceive, understand, and experience algorithmically constructed peer networks (Prey & Esteve-Del-Valle, 2024). For a Suno workflow, this suggests treating how a creator writes prompt text, chooses interface inputs,

and uses tags and metatags as part of how they make sense of the platform’s algorithmic system, not merely as inputs to a model.

Prior makes this platform argument more concrete. His pairing of Bourdieu and ANT shows that music-making is shaped by positions and forms of capital, while his discussion of glitch shows how devices, techniques, and artefacts participate in what musical practice makes possible (Prior, 2008). In the Suno case, the platform is therefore not only infrastructure. It is also an organized setting in which interface inputs, prompt text, tags, metatags, and defaults shape what a creator can try, compare, revise, or circulate.

The scholarship reviewed here supports this conceptual move, but it does not replace case evidence about Suno’s current interface or rights policy. This study can therefore analyze a Suno workflow through platformization and music mediation, while exact claims about current Suno interface inputs, subscription tiers, commercial rights, or model version behavior have to be tied to public platform pages or screenshots from the time of the study. The methodological point is to describe the visible platform arrangement without pretending to know hidden training data or internal model operations.

2.4 Prompt, Tag, and Alignment Evidence in Text-to-Music Studies

Text-to-music research gives this paper an empirical vocabulary. MusicLM shows that natural-language descriptions can guide AI music generation (Agostinelli et al., 2023). Copet et al.’s MusicGen study makes control more explicit: it treats text descriptions and melody inputs as ways for users to guide the generated music, which is why this paper records not only generated songs but also the prompt records, tags, and interface inputs used to request them (Copet et al., 2023). The SunoCaps dataset is especially relevant because it treats prompts, expert alignment comments, and emotion annotations as data for Suno-generated songs (Civit et al., 2024). Casini et al.’s study of Suno and Udio is also relevant because it examines text-conditioned AI-generated music through prompt, lyric, tag, and metatag practices (Casini et al., 2025). Coelho’s work on artist-like traces clarifies the ethical risk of stylistic proximity, while also showing why this paper should avoid claims about artist-specific training signals without a dedicated audit protocol (Coelho, 2025).

These studies support a specific methodological premise: prompt-output relations can be studied through records. They do not imply that prompt text fully controls generated songs, nor do they allow this paper to claim what is inside Suno’s training data. This study therefore keeps the observable records together: prompt records, interface structure, tags, generated audio, and listening notes. Together, these records allow the study to examine how musical intention changes in a platform-based text-to-music workflow, without claiming that a commercial model has or lacks cultural understanding.

3 TAIGI POP MUSIC AS LANGUAGE, VOCALITY, ARRANGEMENT, AND LOCAL MEANING

This section gives the case its cultural specificity. Taigi Pop Music is not used here as an exotic example or merely as a non-English prompt. It is used because its recognizability depends on several layers at once: language, pronunciation, vocality, arrangement, genre memory, social position, local meaning, and historical mediation.

3.1 Language Labels and the Limits of “Taiwanese Hokkien Pop”

Taiwanese popular music scholarship shows why Taigi Pop Music cannot be reduced to the label “Taiwanese Hokkien pop.” Tsai, Ho, and Jian’s introduction to *Made in Taiwan* warns that Taiwanese popular music cannot be reduced to one authentic sound or one language category; it has to be read through language politics, industry, identity, and historical trajectories (Tsai et al., 2019). Chen Pei-Feng’s work on Taigi songs under continuous colonial conditions further shows that language is inseparable from social memory, translation, cover practices, and historically produced ways of singing (Chen, 2019a). Chen Pei-Feng’s study of Japanese-period Taigi popular songs adds that the category was formed through folk-song discourse and musical modernization, not through language alone (Chen, 2017).

For the Suno case, this means that Taigi Pop Music should be treated as the paper’s main English term, not just as a translation choice. The gloss “Taiwanese-language pop music” helps international readers understand the term, but it should not replace the main term. The alternative label Taiwanese Hokkien pop may be useful as a platform-recognition comparison, but it should not become the paper’s main term because it risks reducing the case to a linguistic label.

3.2 Vocality, Re-Japanization, and Structures of Feeling

The terms in this subsection come from Chen Pei-Feng’s historical framework for Taigi songs under continuous colonial conditions. Japanization, re-Japanization, self-Japanization, traditionalization, and canonization name processes through which vocal style, cover practice, market formation, political marginality, and later memory produced and reproduced what later listeners may hear as Taigi song tradition (Chen, 2019a).

Chen Pei-Feng’s work is central because it reframes Taigi songs as social history through sound. The key implication for this paper is that vocality is evidence. Singing style, ornament, phrasing, affective delivery, and the listener’s memory of such delivery all matter. A generated song should therefore not be evaluated only for whether its lyrics are in Taigi or whether its surface genre tag appears correct.

Chen Pei-Feng’s account of Japanization, re-Japanization, self-Japanization, traditionalization, and canonization is also methodologically useful (Chen, 2019a). It prevents the paper from treating enka-like vocality as an ancient

essence of Taigi Pop Music. Instead, enka-like sound can be understood as historically produced through colonial history, postwar cover practices, market formations, social marginality, and later canonization. This raises a question for the case: whether a platform turns this history into a timeless nostalgic marker.

The concept of structure of feeling is especially important. Taigi songs can carry historically formed affects such as longing, displacement, endurance, marginality, powerlessness, and self-rescue. If a platform prompt reduces these histories to generic mood words such as “sad,” “nostalgic,” or “emotional,” something culturally significant may be lost. This therefore becomes an analysis question for the Suno case record.

3.3 Arrangement, Vocal Rhetoric, and Genre Memory

The arrangement literature adds another layer. Chen Wan-Ling’s work on Japanese-period Taigi popular song arrangement shows that the sound of these songs cannot be understood only through lyrics or melody; introductions, interludes, Western-style arrangement, recording time limits, and arranger decisions all shape the musical object (Chen, 2016a). Chen Wan-Ling’s study of arrangement structure further clarifies why musical form and arrangement decisions matter to Taigi popular song analysis (Chen, 2016b). This is directly useful for the Suno case because generated music can be examined for structure, instrumentation, intro/interlude behavior, and arrangement logic, not only for language.

The following labels are treated as historical and analytical terms, not as self-evident style names. Chen Pei-Feng’s study of 1960s Taigi pop and social marginality connects genre memory to class, migration, and the postwar cultural setting (Chen, 2019b). Chen Pei-Feng’s first article on “Taiwanese flavor” connects the term to double decolonization and postwar musical politics (Chen, 2022a). The second continues the account through musical and cultural reclassification (Chen, 2022b). The third extends the debate (Chen, 2023). The fourth links musical identity to decolonization politics (Chen, 2024). Liao’s work on Jody Chiang adds a later case for thinking about vocal identity, rhetoric, and canonized popular memory (Liao, 2020).

For orientation, one published listening reference for the 1990s Taigi ballad memory discussed here is 江蕙 (Jody Chiang), 〈酒後的心聲〉 (1992; [YouTube](#)).

The literature therefore identifies several genre memories that generated songs may mobilize, avoid, or confuse: Japanese-period popular song arrangement, enka-like vocality, nakashi or electronic keyboard associations, post-1980s “Taiwanese flavor,” Mandopop ballad conventions, Taigi rock, and contemporary Taigi indie/pop.

3.4 Post-Martial-Law Taigi Rock and Contemporary Indie Authenticity

This study avoids freezing Taigi Pop Music as old, nostalgic, or automatically enka-like. Chen Ru-Shou’s discussion of Taigi rock shows that post-martial-law experimentation also belongs to the broader history of Taigi

popular music (Chen, 1996). Jhuang's study of Sorry Youth shows how contemporary Taigi indie practice can produce local authenticity through scene practice and everyday life (Jhuang, 2022). This matters because a Suno prompt for "Taigi Pop Music" might be judged against contemporary pop or indie criteria, not only against older recording histories.

For this study, authenticity should not mean purity. It should be treated through relations among language, place, everyday life, vocal attitude, arrangement, scene practice, fan culture, and the listening practices of people familiar with the music. That frame gives the analysis a way to avoid essentializing Taigi Pop Music while still taking cultural specificity seriously.

3.5 Listening Frame for Suno x Taigi Pop Music

Taken together, the studies cited above do not produce a checklist for judging generated songs. They give this study a listening frame for asking how, and with what uncertainty, a song becomes recognizable as Taigi Pop Music. The first question is linguistic and vocal: whether sung lines are recognizable as Taigi or Mandarin, mix languages, or use nonspecific or synthetic-sounding syllables, and whether vocal delivery suggests Taigi ballad phrasing, Mandopop delivery, indie vocal attitude, enka-like ornament, or an unclear synthetic-sounding voice.

The second question concerns musical memory. Arrangement, instrumentation, intro and interlude behavior, rhythm section, keyboard or guitar texture, and production style can call up older recording markers, studio-pop conventions, postwar Taigi ballad, Mandopop, Taigi rock, contemporary Taigi indie/pop, or other formations. These musical cues also shape affect. A generated song may suggest longing, displacement, endurance, marginality, powerlessness, or self-rescue, or it may reduce such histories to generic sadness or nostalgia.

The third question concerns social and local recognizability. The song may imply an out-migrant, worker, woman, performer, sailor, socially marginal figure, urban narrator, or generic romantic speaker. It may also make place, everydayness, fan culture, or local authenticity audible, or it may only add a generic "Taiwanese" decoration. The case analysis does not require every generated song to answer all of these questions. It uses this frame to place listening notes beside prompt text, generated lyrics, visible platform records, and audio.

4 LISTENING MEMORY, PROMPT PRACTICE, AND GENERATED SONGS

4.1 Listening Memory and Earlier Prompt Practice

The first author's understanding of Taigi Pop Music developed through television, popular songs, online listening, platform recommendations, deliberate searching, and reading about Taigi songs and Taiwanese popular music. This history informs the first author's judgments of pronunciation, vocal delivery, arrangement, affect, genre memory, and local meaning in the generated songs. The listening notes record those judgments from this stated

position rather than presenting them as a shared standard among all listeners.

Earlier use of Suno also shaped the formal case. In earlier versions with fewer visible options, requests that named only "Taigi Pop Music" or "Taiwanese-language pop music" did not consistently produce songs close to the first author's listening memory. Those attempts drew attention to the relation among language descriptions, lyrics, style descriptions, phonetic choices, and generated sound. The case begins with the materials preserved for the four Suno 5.5 runs.

4.2 Four Runs, Eight Songs, and Their Visible Records

The case contains four runs and eight songs. S1 Simple Ballad and S2 Simple Indie use Simple mode. A1 Advanced Ballad and A2 Advanced Indie use Advanced mode. Each run produced two songs: S1A/S1B, S2A/S2B, A1A/A1B, and A2A/A2B.

The written materials differ between the modes. In Simple mode, the first author entered one Song Description, and Suno generated the title and lyrics. The saved record also includes the Prompt / Style Summary shown with each song. In Advanced mode, the first author supplied the title, complete lyrics, style description, exclude text, and slider settings separately. The modes therefore differ both in how the writing is arranged and in which text is supplied by the first author or generated by Suno.

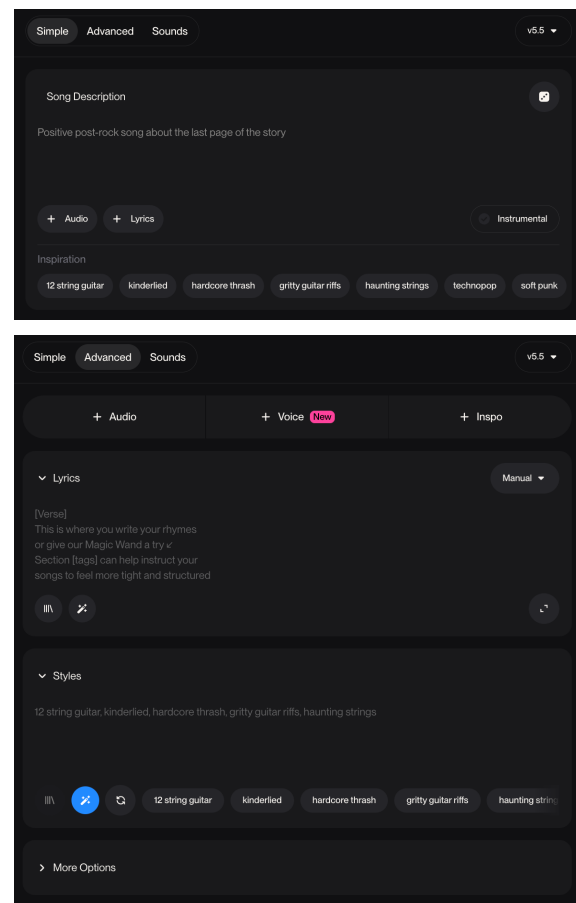


Figure 1. Suno 5.5 mode interfaces: Simple mode & Advanced mode

These records allow three comparisons. The interface comparison asks how the two modes arrange musical requests. The within-run comparison asks which features recur or differ between two songs generated from the same saved visible settings. The A1/A2 comparison keeps the complete lyrics unchanged while the title, style description, exclude text, sliders, and other settings differ. Because several settings change together, this comparison cannot assign an audible difference to one setting alone.

The listening frame in Section 3.5 guides the notes without becoming a score or checklist. Each song is described through the dimensions that can be heard clearly enough, while cultural interpretations are identified as judgments grounded in the first author’s listening context.

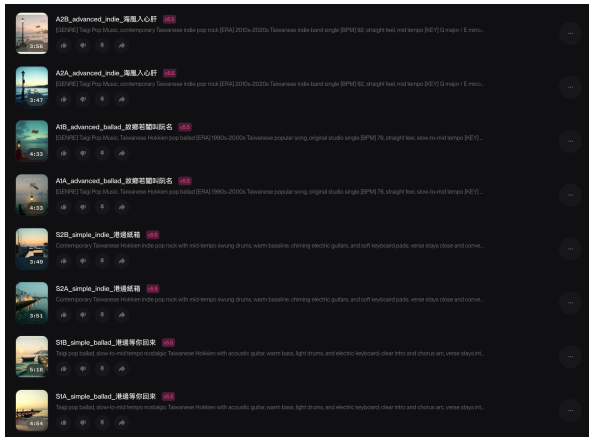


Figure 2. Suno 5.5 case record: eight generated songs.

4.3 Written Requests, Generated Songs, and Listening Across Four Runs

The saved S1 summary describes a nostalgic Taigi pop ballad with acoustic guitar, warm bass, light drums, electric keyboard, close verses, and a fuller chorus. It also ends with labels including rap, edm, ballad, and acoustic. Suno generated the title 〈港邊等你回來〉 and lyrics about waiting, family memory, and returning home. In S1A, pronunciation in the chorus was more recognizable as Taigi than in the verse, which leaned toward Mandarin. The vocal sounded raspy, projected, and synthetic. S1B remained in a similar ballad setting but had a smoother surface and shifts between male- and female-sounding vocal qualities. In both songs, the harbor setting and the speaker’s relation to home were clearer in the lyrics than in the audio alone.

The saved S2 summary describes contemporary Taiwanese Hokkien indie pop rock with mid-tempo drums, warm bass, chiming electric guitar, soft keyboard, and conversational verses. Suno generated the title 〈港邊紙箱〉 and lyrics about fish markets, factory work, leaving home, and returning to the harbor. S2A had a plain male vocal, clean electric guitar, warm bass, loose drums, and a gentle lift before the final refrain. S2B remained within an indie or pop-rock setting but included broader pop and more synthetic vocal qualities. In both songs, the band arrangement was easier to hear than the working life described in the lyrics.

The A1 materials request a 1990s-2000s Taigi pop ballad with acoustic guitar, electric keyboard, warm bass, light drums, and a warm, slightly raspy male narrator. Its lyrics repeatedly return to “故鄉若閣叫阮名” and connect departure, loneliness, and return. A1A contained recognizable Taigi pronunciation, but parts of its musical setting moved toward Mandopop convention. Its vocal shifted after the first chorus and included karaoke-like phrasing. A1B had a stronger minor-mode feeling and a clearer relation to 1990s-2000s Taigi-song memory. Its pronunciation was more consistently recognizable as Taigi, although the later vocal sounded more female. Both songs retained the ballad setting while differing in pronunciation, vocal identity, and musical association.

A2 uses the complete lyrics entered for A1 but changes the title and other settings. Its style description requests contemporary Taiwanese indie pop rock, clean electric guitar, warm bass, light drums, soft keyboard, and conversational vocal delivery. The shared lyrics still contain directions for acoustic-guitar arpeggio, restrained verses, an open chorus, a reduced-drums bridge, and a soft outro. The written materials therefore place ballad and contemporary indie directions together.

A2A combined recognizable Taigi pronunciation and conversational male vocals with clean electric guitar, warm bass, light drums, soft keyboard, and a mid-tempo guitar-band arrangement. The notes also record enka-like ornament in parts of the vocal. A2B remained in a similar band setting but carried stronger associations with Mandopop and ballad memory. Its chorus was easier to understand than its verse. The two songs shared an indie or pop-rock arrangement while differing in pronunciation clarity and in the balance among indie, Mandopop, and ballad associations.

The A1/A2 comparison does not identify which setting produced a particular feature. It shows that the same complete lyrics can remain present while the surrounding written materials and generated songs differ. Across all four runs, lyrics make narrative, place, and social position explicit, while audio makes pronunciation, vocal quality, arrangement, affect, and genre memory available for listening. The two forms of evidence overlap, but neither replaces the other. Section 5 begins from these relations rather than from any one prompt, setting, or song.

5 AGENCY AND TAIGI RECOGNIZABILITY ACROSS THE WORKFLOW

5.1 Paired Songs and Shared Lyrics Across Runs

Across the four runs, each A/B pair began from the same saved visible settings. Each pair repeated a broad musical setting while differing in pronunciation, vocal quality, or musical association, as detailed in Section 4.3. These comparisons do not isolate why those differences appeared. They establish a narrower point: the visible settings describe part, but not all, of the generated song.

A1 and A2 add a cross-run comparison. The complete lyrics remain the same, while the title, style description,

exclude text, sliders, and other settings change together. Because those settings changed together, the comparison cannot assign an audible feature to one setting. It shows that lyrics can provide a stable narrative and formal outline without fully determining musical identity.

5.2 How Writing, Interface, Generation, and Listening Make a Difference

These comparisons clarify how agency appears across the workflow. The first author chooses how to describe the music, which lyrics and settings to use, and which differences matter when listening. The interface places some requests together and keeps others separate; it also structures whether titles and lyrics are supplied by the first author or generated by Suno. Generation introduces audible features not fully settled in writing. Listening does not produce those features, but it makes them available for comparison, selection, and cultural interpretation.

These parts do not control the song in the same way. Bates's account of agency asks what makes a difference in a particular relation (Bates, 2018). Born and Barry describe mediation as transmission, translation, and transformation (Born & Barry, 2018). Hennion treats techniques, objects, practices, and listening as part of music-making (Hennion, 2012). Poell et al. place interface organization within cultural production (Poell et al., 2017). Simple and Advanced modes therefore matter here as ways of organizing textual work, not as a ranking of their quality or control.

5.3 Taigi Recognizability and the Workflow as the Unit of Analysis

Taigi recognizability explains why listening remains in the analysis. Lyrics make narratives of place, work, family memory, departure, and return explicit. Audio makes pronunciation, vocal delivery, arrangement, affect, and genre memory available for listening. A language label or complete lyrics cannot settle whether a song becomes recognizable as Taigi Pop Music. That judgment depends on text, sound, musical history, and the first author's stated listening context. It does not demonstrate agreement among Taigi Pop Music listeners.

This case shows why preserved writing, visible platform records, paired songs, and listening notes should be kept together while remaining distinct forms of evidence. It does not provide a general evaluation of Suno or a model of how all listeners recognize Taigi Pop Music. Independent listeners, repeated generations, and other platform versions could examine whether these relations recur beyond this case.

Within this study, the unit of analysis is the workflow rather than the prompt or generated song alone. Agency becomes visible by tracing what changes as a musical request is written, organized by the interface, made audible, and interpreted through listening. No single part completely explains the song. Following those relations moves the analysis from prompt to platform.

AUTHOR DECLARATIONS

Conflicts of interest: The authors report no conflicts of interest relevant to this submission.

Ethical considerations: This study involved no third-party participants and collected no personally identifiable data.

Use of AI: This study uses AI in two distinct ways. Suno 5.5 was used to generate songs. ChatGPT was used as an auxiliary tool to help draft, organize, and revise Suno prompt texts and related case materials, and to polish manuscript prose.

REFERENCES

- Agostinelli, A., Denk, T. I., Borsos, Z., Engel, J., Verzetti, M., Caillon, A., Huang, Q., Jansen, A., Roberts, A., Tagliasacchi, M., Sharifi, M., Zeghidour, N., & Frank, C. (2023). Musi-cLM: Generating music from text. *arXiv*. <https://doi.org/10.48550/arXiv.2301.11325>
- Bates, E. (2018). Actor-network theory and organology. *Journal of the American Musical Instrument Society*, 44, 41–51. https://amis.mircat.org/jamis/2018_041.pdf
- Born, G., & Barry, A. (2018). Music, mediation theories and actor-network theory. *Contemporary Music Review*, 37(5–6), 443–487. <https://doi.org/10.1080/07494467.2018.1578107>
- Casini, L., Cros Vila, L., Dalmazzo, D., Kaila, A.-K., & Sturm, B. L. T. (2025). Data-driven analysis of text-conditioned AI-generated music: A case study with Suno and Udio. *arXiv*. <https://doi.org/10.48550/arXiv.2509.11824>
- Chen, P.-F. (2017). 重新省思日治時期台語流行歌曲：以民謠觀的建立和音樂近代化作為觀點. *台灣文學研究學報*, 25, 9–58.
- Chen, P.-F. (2019b). 不同時空下的臺日天涯淪落人：1960年代臺語流行歌曲和社會邊緣人. *文化研究*, 28, 61–106. [https://doi.org/10.6752/JCS.201912_\(28\).0003](https://doi.org/10.6752/JCS.201912_(28).0003)
- Chen, P.-F. (2019a). 歌唱臺灣：連續殖民下臺語歌曲的變遷. 衛城出版.
- Chen, P.-F. (2022b). 雙重解殖與「臺灣味音樂」的形塑 (II)：臺語流行音樂之路 (陳和平篇 上). *臺灣文學研究集刊*, 28, 99–142. <https://doi.org/10.53106/181856492022080028004>
- Chen, P.-F. (2022a). 雙重解殖與「臺灣味音樂」的形塑 (I)：臺語流行音樂之路 (劉清池篇). *臺灣文學研究集刊*, 27, 107–146. <https://doi.org/10.53106/181856492022020027004>
- Chen, P.-F. (2023). 雙重解殖與「臺灣味音樂」的形塑 (III)：臺語流行音樂之路 (陳和平篇 中). *臺灣文學研究集刊*, 30, 91–152. <https://doi.org/10.53106/181856492023080030004>
- Chen, P.-F. (2024). 雙重解殖與「臺灣味音樂」的形塑 (IV)：臺語流行音樂之路 (陳和平篇 下). *臺灣文學研究集刊*, 31, 95–160. <https://doi.org/10.53106/181856492024020031004>
- Chen, R.-S. (1996). 台語搖滾初探. *中外文學*, 25(4), 134–142. [https://doi.org/10.6637/CWLQ.1996.25\(4\).134-142](https://doi.org/10.6637/CWLQ.1996.25(4).134-142)
- Chen, W.-L. (2016a). 文化混雜的聽覺經驗初探：日治時期臺語流行歌的西式編曲. *歷史臺灣：國立臺灣歷史博物館館刊*, 12, 7–51.
- Chen, W.-L. (2016b). 日治時期臺語流行歌曲之編曲研究：編曲樂段之結構與組成形式分析. *臺灣音樂研究*, 23, 55–99.
- Civit, M., Drai-Zerbib, V., Lizcano, D., & Escalona, M. J. (2024). SunoCaps: A novel dataset of text-prompt based AI-

- generated music with emotion annotations. *Data in Brief*, 55, 110743. <https://doi.org/10.1016/j.dib.2024.110743>
- Coelho, G. (2025). The artist is present: Traces of artists residing and spawning in text-to-audio AI. In *Proceedings of the 6th Conference on AI Music Creativity (AIMC 2025)*. Zenodo. <https://doi.org/10.5281/zenodo.16946623>
- Copet, J., Kreuk, F., Gat, I., Remez, T., Kant, D., Synnaeve, G., Adi, Y., & Défossez, A. (2023). Simple and controllable music generation. *arXiv*. <https://doi.org/10.48550/arXiv.2306.05284>
- Deruty, E., Grachten, M., Lattner, S., Nistal, J., & Aouameur, C. (2022). On the development and practice of AI technology for contemporary popular music production. *Transactions of the International Society for Music Information Retrieval*, 5(1), 35–49. <https://doi.org/10.5334/tismir.100>
- Fiebrink, R. A., & Caramiaux, B. (2018). The machine learning algorithm as creative musical tool. In R. T. Dean & A. McLean (Eds.), *The Oxford handbook of algorithmic music* (pp. 181–208). Oxford University Press. <https://doi.org/10.1093/oxfordhb/9780190226992.013.23>
- Hennion, A. (2012). Music and mediation: Toward a new sociology of music. In M. Clayton, T. Herbert, & R. Middleton (Eds.), *The cultural study of music: A critical introduction* (2nd ed., pp. 249–260). Routledge. <https://hal.science/hal-00771905>
- Ho, T.-H. (2019). Profiling a postwar trajectory of Taiwanese popular music: Nativism in metamorphosis and its alternatives. In E. Tsai, T.-H. Ho, & M. Jian (Eds.), *Made in Taiwan: Studies in popular music* (pp. 23–42). Routledge. <https://doi.org/10.4324/9781351119146-3>
- Jhuang, C.-S. (2022). 歹勢好勢擺是凡勢：台語獨立樂團拍謝少年音樂真誠性的實踐 [Master's thesis]. National Taiwan Normal University.
- Latour, B. (2005). *Reassembling the social: An introduction to actor-network-theory*. Oxford University Press. <https://doi.org/10.1093/oso/9780199256044.001.0001>
- Liao, Z.-K. (2020). 江蕙及其歌曲在台語流行音樂之發展探究與語藝分析 [Master's thesis]. Shih Hsin University.
- Poell, T., Nieborg, D., Duffy, B. E., Prey, R., & Cunningham, S. (2017). The platformization of cultural production. *Selected Papers of AOIR 2017: The 18th Annual Conference of the Association of Internet Researchers*, 1–19. https://pure.rug.nl/ws/portalfiles/portal/103169850/AoIR2017_Platformization_of_Cultural_Production.pdf
- Prey, R., & Esteve-Del-Valle, M. (2024). The algorithmic network imaginary: How music artists understand and experience their algorithmically constructed networks. *The Information Society*, 40(1), 18–31. <https://doi.org/10.1080/01972243.2023.2274070>
- Prior, N. (2008). Putting a glitch in the field: Bourdieu, actor network theory and contemporary music. *Cultural Sociology*, 2(3), 301–319. <https://doi.org/10.1177/1749975508095614>
- Seaver, N. (2019b). Captivating algorithms: Recommender systems as traps. *Journal of Material Culture*, 24(4), 421–436. <https://doi.org/10.1177/1359183518820366>
- Seaver, N. (2019a). Knowing algorithms. In J. Vertesi & D. Ribes (Eds.), *digitalSTS: A field guide for science & technology studies* (pp. 412–422). Princeton University Press. <https://doi.org/10.2307/j.ctvc77mp9.30>
- Tsai, E., Ho, T.-H., & Jian, M. (2019). Introduction: Problematizing and contextualizing Taiwanese popular music. In E. Tsai, T.-H. Ho, & M. Jian (Eds.), *Made in Taiwan: Studies in popular music* (pp. 1–19). Routledge. <https://doi.org/10.4324/9781351119146-1>
- Väkevää, L., & Partti, H. (2025). Generative AI as a collaborator in music education: An action-network theoretical approach to fostering musical creativities. *Action, Criticism, and Theory for Music Education*, 24(3), 16–52. <https://doi.org/10.22176/act24.3.16>
- Xambó, A., & Roma, G. (2024). Human-machine agencies in live coding for music performance. *Journal of New Music Research*, 53(1–2), 33–46. <https://doi.org/10.1080/09298215.2024.2442355>