

FROM RHYTHM TO REPRESENTATION: BTS'S "IDOL" AS A MULTIMODAL AUDIO SCORE

Veronika Mattová

Masaryk University

veronika.mattova@phil.muni.cz

Michal Schwarz

Masaryk University

schwarz@phil.muni.cz

ABSTRACT

Contemporary Korean music, especially through artists like group BTS, has reshaped global music by integrating traditional Korean elements with modern pop influences. This study examines how BTS's song "IDOL" functions as a multimodal audio score, blending traditional rhythmic structures like jangdan (cyclical rhythms) and dance forms such as bukcheong sajanori (lion mask dance) with contemporary K-pop production. By analyzing Jeongganbo, a traditional Korean time-based notation system, this paper demonstrates how these cultural elements are adapted and transformed into visual, rhythmic, and sonic components of "IDOL". The study highlights how rhythmic cycles, onomatopoeic syllables, and choreography are used to create a hybridized musical experience. Additionally, the research explores how "IDOL" exemplifies transmedia storytelling, where music, video, choreography, and fan interaction form a dynamic, layered narrative. This paper illustrates how BTS reinterprets traditional Korean musical practices within a global pop context, showing how multimodal audio scores and transmedia strategies allow for a deeper, more immersive engagement with music, bridging cultural and temporal gaps.

1. INTRODUCTION

K-pop, a genre of popular music coming from South Korea, a global cultural phenomenon, has revolutionized the way music is produced, consumed, and experienced worldwide. While often associated with contemporary pop music, K-pop also integrates elements of traditional Korean music (gugak, 국악), offering a hybridized musical experience (Kim et al., 2024). This research explores how BTS, one of the most successful K-pop groups, incorporates traditional Korean elements into their

work, specifically in the song "IDOL". The song and its accompanying performance serve as an excellent case study for understanding how traditional practices, such as jangdan (장단) (rhythmic structures) and bukcheong sajanori (북청사자놀이) (lion mask dance), are incorporated into modern K-pop performances, creating a unique multimodal audio score that bridges cultural and temporal gaps.

This study will analyze how multimodal representations, including audio, visual, and performative elements, function as an audio score in K-pop, particularly in Korean music band, BTS's music (Cake Corporation, 2024). By examining Jeongganbo (정간보), a traditional Korean time-based notation system, the study aims to show how elements of traditional notation are preserved, adapted, and transformed into visual and rhythmic representations in BTS's performance, ultimately creating a hybrid audio score. By situating this case study within broader frameworks of media convergence and transmedia storytelling, the paper also considers how musical meaning is distributed across artefacts, platforms, and participatory fan practices.

2. JEONGGANBO: TRADITIONAL KOREAN NOTATION AND ITS RELEVANCE TO MODERN MUSIC

Jeongganbo (정간보) is a traditional Korean musical notation system used primarily in gugak (국악) (Korean classical music) (Kim et al., 2024). Unlike Western music notation, which is pitch-based and uses a staff, Jeongganbo is organized around time segmentation and rhythmic representation (Tran, Park, & Jung, 2023). This system focuses on time-based grids rather than the representation of pitch, which makes it unique compared to Western systems of musical notation (Kim et al., 2024). Jangdan (장단), understood as a cyclical rhythmic framework, is central to Korean musical notation and performance. In their analysis of Jeongganbo, Tran, Park, and Jung (2023) demonstrate how rhythmic structure and time

Copyright: © 2024 First author et al. This is an open-access article distributed under the terms of the [Creative Commons Attribution License 3.0 Unported](https://creativecommons.org/licenses/by/3.0/), which permits unrestricted use, distribution, and reproduction in any medium, provided the original author and source are credited.

segmentation are encoded through repetitive, cycle-based patterns.

While Jeongganbo is not widely used today in mainstream music, its influence can still be felt in contemporary Korean musical practices, particularly in traditional performance and certain K-pop choreographies (Tran, Park, & Jung, 2023). When we look into the past again from the bigger perspective, not only Jeongganbo but also its music theory had in the Korean cultural context quite valid representation of social and cosmological order rather than a neutral technical record (Han, 1983). This higher meaning gives us another academic explanation, why is it still relevant, to this day.

3. THE ROLE OF MULTIMODAL AUDIO SCORES IN K-POP

In K-pop, music and visuals are inseparable. Music videos, lyrics, and live performances function not just as musical scores, but as multimodal representations of the music. The term multimodal audio score in context of this paper, refers to the idea that musical compositions can be represented through a combination of sound, text, image, and movement, while in the case of BTS's "IDOL", the music video (BTS, 2018) and its lyrics (Genius English Translations, 2018) serve as multimodal audio scores, where each medium adds layers of meaning to the music, allowing for a deeper, more nuanced understanding of the song.

For example, the BTS Lyrics Inside book, in its Global Edition, transforms the traditional lyrics into a hybrid score. Instead of simply presenting the text, the book incorporates rhythmic notation, onomatopoeic syllables, and diagrammatic representations of time, reflecting the structure of Jeongganbo (정간보) (Cake Corporation, 2024). The use of symbols and images to represent rhythm is an important feature of this multimodal score, as it allows the reader to engage with the music in a way that goes beyond the traditional means of music notation. The cross-modal translation between Korean onomatopoeia-creation of words that phonetically imitate or evoke sounds linked to specific actions or objects (Nissa et al., 2021) and Romanized English gloss (bum badum...) provides a rich, dynamic representation of the music, offering fans a unique way to interact with the song.

4. TRANSMEDIAL STORYTELLING ACROSS MEDIA

4.1. Media Convergence and Participatory Culture

The concept of transmedia storytelling is inseparable from Henry Jenkins' foundational work on media convergence. In *Convergence Culture: Where Old and New Media Collide*, Jenkins (2006) defines convergence not as a purely technological process, but as a profound cultural shift that reshapes how media content is produced, circulated, and consumed. Convergence refers to the flow of content across multiple media platforms, the

cooperation among different media industries, and the migratory behavior of audiences who actively seek meaningful narrative experiences across media boundaries.

Crucially, Jenkins rejects the notion of passive media consumption. Instead, he emphasizes participatory culture, where audiences function as active interpreters, collaborators, and, in some cases, co-creators of meaning. Media convergence, therefore, occurs "within the brains of individual consumers and through their social interactions with others" (Jenkins, 2006), highlighting the cognitive and social dimensions of contemporary media engagement. This participatory framework enables fans to construct personal and collective mythologies by connecting fragmented narrative elements dispersed across media environments.

Jenkins further introduces the concept of collective intelligence, adapted from Pierre Lévy, to describe how knowledge production becomes a shared social process and if we look at this topic from the convergence culture, no single individual possesses complete knowledge; rather, meaning emerges through collaboration, discussion, and shared interpretation (Lévy, 1997; Jenkins, 2006). This principle is particularly relevant to global fandoms, where interpretive labor is distributed across transnational communities and digital platforms.

4.2. Transmedia Storytelling Beyond a Buzzword

While Jenkins' work laid the theoretical groundwork for understanding convergence and participation, subsequent scholarship has critically examined how transmedia storytelling is operationalized in practice. Sousa, Martins, and Zagalo (2016) argue that transmedia storytelling has often been reduced to a fashionable buzzword, frequently misapplied to describe mere cross-media promotion or audience interaction. They contend that genuine transmedia storytelling requires the meaningful distribution of narrative elements across multiple media and artefacts, where each component contributes uniquely to the overall storyworld.

Importantly, Sousa et al. (2016) conceptualize transmedia storytelling as a dynamic process involving three principal stakeholder groups: the media industry, creators, and the public. Each group enters the transmedia ecosystem with distinct stakes. Media industries prioritize economic sustainability and brand expansion; creators seek artistic expression and narrative coherence; audiences pursue emotional engagement, identity work, and participatory belonging. These interests may diverge, but they can also converge productively when transmedia principles are authentically applied.

Rather than framing transmedia as a linear top-down strategy, Sousa et al. emphasize negotiation, reciprocity, and creative tension among participants. In this sense, transmedia storytelling becomes an ecosystem where creativity, participation, and fandom coexist alongside economic and institutional interests. This perspective extends Jenkins' participatory model by explicitly addressing power relations and the structural conditions shaping transmedia production.

4.3. K-pop as a Transmedia Ecosystem

K-pop represents a particularly illustrative case of transmedia storytelling in action. As a global cultural industry, K-pop operates through a highly integrated media environment that includes music, performance, visual aesthetics, social media, merchandise, documentaries, books, and interactive fan platforms. These elements do not merely replicate content across channels; they expand narrative meaning and deepen audience immersion. This is not some new practice. Imaginaries of all kinds can be seen in various Korean music contexts. For example, specifically instrument imagery grounds rhythmic notation in embodied Korean ensemble practice (Rockwell, 1983). This can be understood as an early form of storytelling that continues to inform imaginary practices today.

In the case of BTS, transmedia storytelling functions as a bridge between traditional Korean cultural elements and contemporary global media practices. Musical compositions incorporate traditional rhythms and symbolic references, while choreography, visual storytelling, and lyrical narratives extend these meanings across formats (Cook, 2001). This convergence between “old” and “new” media reflects Jenkins’ assertion that convergence culture does not eliminate older media forms but reconfigures their relationships within new contexts.

From the perspective proposed by Sousa et al. (2016), BTS’ transmedia universe exemplifies a relatively balanced alignment of stakeholder stakes. The media industry (HYBE- Company of BTS) strategically coordinates platforms and distribution; the artists actively participate in narrative construction and thematic continuity; and the fandom (ARMY) engages in interpretive, creative, and affective labor that sustains the storyworld across cultural and linguistic boundaries.

4.4. The BTS Lyrics Inside Book as a Transmedial Artefact

The BTS Lyrics Inside book can be understood as a transmedial artefact that crystallizes the convergence of music, text, and visual design. Rather than functioning solely as a paratextual supplement, the book recontextualizes lyrics as literary and cultural material, inviting close reading, translation, and reinterpretation. This artefact enables fans to engage with the music beyond auditory consumption, transforming listening into a multimodal interpretive practice.

Through this lens, fans become active participants in meaning-making processes, aligning with Jenkins’ model of participatory culture and Sousa et al.’s emphasis on fandom as a central transmedia actor. The Lyrics Inside book operates as a node within a larger narrative network, connecting albums, performances, interviews, and fan-generated content. It thus supports a shared narrative space where emotional symbolism, cultural identity, and collective intelligence intersect, far behind the music, itself.

5. DISCUSSION: CASE STUDY

5.1. BTS “IDOL” as a Multimodal Audio Score

The BTS Lyrics Inside book offers a compelling example of how elements of traditional Korean music are recontextualized through contemporary multimodal notation. Rather than functioning solely as a repository of lyrics, the book operates as a hybrid audio score that combines text, sound cues, visual layout, and performance imagery into a single interpretive space.

Through vocalized rhythmic patterns such as “덩기덕 쿵더러러” (deonggi-deok kung-deoreo, onomatopoeic syllables imitating drum sounds) (Cake Corporation, 2024), cross-modal translations, for example from Korean onomatopoeia to Romanized English glosses, and diagrammatic layouts, the act of reading lyrics is transformed into a multimodal experience (Boddy & Boddy, 2010). Images of performance further extend this experience by visually referencing bodily movement and rhythmic gesture, suggesting how sound, motion, and meaning coalesce in performance (Cake Corporation, 2024).

Importantly, this multimodal strategy does not merely illustrate the song but actively guides interpretation. The reader is positioned not as a passive consumer of lyrics but as an embodied participant who is invited to imagine rhythm, movement, and affect simultaneously.

The lyrics explicitly include traditional exclamations, for instance:

“열쌍 좋다 ... 지화자 좋다”

(yeolssang jota ... jiwaja jota,

“Oh yeah, it’s awesome ... Hooray, it’s so awesome”)

(Cake Corporation, 2024)

As explained in the BTS Lyrics Inside book, 열쌍 (yeolssang, a joyful exclamation) and 지화자 (jiwaja, a celebratory chant) are vocal expressions known as 추임새 (chuiimsae, supportive rhythmic exclamations), traditionally used during singing or dancing in Korean music. In “IDOL,” the phrase “You can’t stop me lovin’ myself” followed by these exclamations functions as an expression of heightened excitement embedded directly into the rhythmic structure of the song (Cake Corporation, 2024).

Beyond their explanatory role, these exclamations also signal a continuity between traditional performance practices and contemporary pop expression. Their inclusion suggests that affective participation, collective response, and vocal embodiment remain central to musical meaning-making across genres and historical periods (Catalano, 2024).

This multimodal audio score enables fans to engage with BTS’s music beyond listening alone, encouraging exploration of traditional jangdan as rhythmic cycle

principles that underlie the song. The integration of these rhythmic frameworks into a global pop format illustrates the hybrid nature of K-pop, where inherited musical logics coexist with contemporary production and global circulation.

This hybridity becomes especially visible in live performance contexts. In the 2018 Melon Music Awards (MMA) special performance of “IDOL,” the members of BTS incorporated traditional Korean dance forms as part of a large-scale pop spectacle. As documented in the lyrics book, a drum dance, a fan dance, and a mask dance, drawing on performance traditions such as pungmulnori (풍물) traditional percussion-based folk performance) and bukcheong sajanoreum (북청사자놀이), also known as alion mask dance ritual (Cake Corporation, 2024), used as a valid cultural component in nowadays vibrating performance.

Here, choreography functions as an extension of notation. Movement becomes another representational layer through which rhythm and cultural reference are communicated, reinforcing the idea of the song as a multimodal score that unfolds across media and performance contexts.

This logic extends into the music video itself. Beyond the lyrics, the choreography, costumes, and visual aesthetics incorporate recognizable elements of Korean cultural heritage. These include lion mask dancing from lion mask dance and also buchaechum (부채춤), in our context knowns as a fan dance- motifs, and visual references to Korean landscape painting. These elements are juxtaposed with both traditional and Euro-Asian architectural imagery, emphasizing the layered and hybrid visual language of “IDOL” (Cake Corporation, 2024).

Such visual strategies further underscore how cultural memory is activated through imagery. Rather than functioning as static references, these elements invite imaginative engagement, allowing viewers to connect rhythm, movement, and symbolism across historical and contemporary frames.

Finally, rhythmic mnemonics such as “덩기덕 쿵더러러” (deonggi-deok kung-deoreo, vocalized representation of drum strokes) (Figure 1) imitate the sounds of the 장구 (janggu, hourglass-shaped traditional Korean drum). As described in the lyrics book, these vocalizations correspond to specific drum strikes within traditional jangdan (rhythmic cycles) (Cake Corporation, 2024).

⊙		○	⋮
명 (deong)	기덕 (gi-deok)	쿵 (kung)	더러러러 (deo-reo-reo-reo)

Figure 1. “IDOL’s” “덩기덕 쿵더러러” (deonggi-deok kung-deoreo) as commonly used jangdan in traditional Korean music. This notation represents rhythmic patterns from Jeongganbo, a traditional Korean time-based system. Adapted from Cake Corporation. (2024). BTS Lyrics Inside (Global Edition). Big Hit Music Co., Ltd. ISBN 979-11-90996-80-8.

In this sense, “IDOL” (BTS, 2018) exemplifies how traditional Korean music notation and performance logic can be reimagined as a multimodal, performative score. Here, one can already observe early narrative and symbolic structures that continue to support diverse imaginary practices in contemporary popular culture.

6. LIMITATIONS AND FUTURE RESEARCH

This study is based on a single case study, BTS’s “IDOL,” which allows for a detailed analysis but may appear selective. While the song was chosen due to its explicit integration of traditional Korean rhythmic structures, performance practices, and visual symbolism, the findings should not be generalized to all K-pop productions.

Future research aims to expand this collaborative Korean music research to include multiple artists, groups, and songs that employ similar musical and performative strategies. Examining a broader corpus would enable comparative analysis and strengthen the generalizability of the proposed concept of multimodal audio scores.

Such expansion may also benefit from interdisciplinary approaches, combining musicology, media studies, and multimodal analysis, to further explore how traditional Korean musical frameworks continue to inform contemporary popular music.

7. CONCLUSION

The integration of traditional Korean music (gugak, 국악) into contemporary K-pop offers a compelling example of multimodal music representation, in which sound, text, image, notation, and performance interact across media. In this context, music emerges not merely as an auditory phenomenon but as a distributed representational system spanning multiple artefacts and platforms (Han, 2018).

Korean music notation, particularly Jeongganbo (정간보), provides an important conceptual background for this discussion. As a time-based notational system that prioritizes rhythm, duration, and embodied performance, Jeongganbo reflects a long-standing understanding of music as a performative and interpretive practice rather than a fixed textual object. This historical perspective highlights that notation in Korean musical contexts has traditionally functioned as a mediating framework between sound, body, and meaning.

Against this backdrop, the *BTS Lyrics Inside* book can be understood as a contemporary hybrid audio score. Through rhythmic mnemonics, visual layouts, and instrumental imagery, it renders rhythmic structure legible without relying on conventional staff notation. While these representations do not constitute formal notation systems, they operate according to analogous principles, translating rhythm, timing, and gesture into multimodal forms. In doing so, popular music artefacts implicitly echo earlier Korean notational logics while adapting them to global and transmedia environments.

From a transmedial perspective, musical meaning circulates across albums, lyrics books, music videos, merchandise, and live performances. These artefacts do not merely accompany sound but actively participate in its representation and interpretation, enabling fans to engage with music as a visual, textual, and embodied experience. This participatory dimension underscores the role of fans as interpreters within contemporary K-pop ecosystems.

These observations align with current research in Music Information Retrieval and music computing, which increasingly emphasizes the importance of multimodal representations for understanding musical content (Simonetta et al., 2019). Although this study does not propose computational models, it offers a culturally grounded case illustrating how multimodal music representation already operates in practice.

Looking ahead, hybrid artefacts such as lyrics books, annotated visual scores, and interactive media environments represent promising sites for future research at the intersection of Korean music notation, multimodal representation, and transmedia storytelling. Examining these forms may offer valuable insights into how music is not only heard but also seen, read, and embodied in the digital age, contributing to a broader understanding of sound as score.

Acknowledgments

This paper was prepared thanks to financial support within the project “Evolutionary Interferences of Religion and Governance in Inner Asia: Comparison of mutual impacts with tributary countries: Mongolia, Korea, and Vietnam” financed by the Czech Science Foundation at Masaryk University, Faculty of Arts (project code: GA23-06953S).

8. REFERENCES

1. Boddy, J., & Boddy, J. (2010). A multimodal way of experiencing and exploring music. *Interdisciplinary Science Reviews*, 35(2), 138–153.
2. BTS. (2018, August 24). *IDOL* [Music video]. YouTube. <https://www.youtube.com/watch?v=pBuZEGYXA6E>
3. Cake Corporation. (2024). *BTS lyrics inside* (Global edition). Big Hit Music Co., Ltd. ISBN 979-11-90996-80-8
4. Catalano, E. (2024). The colonial era and beyond: The past and present and its effect on music and K-pop in Korea. *Colorado Journal of Asian Studies*, 11(1). <https://doi.org/10.33011/cjas.v1i1.2963>
5. Cook, N. (2001). Theorizing musical meaning. *Music Theory Spectrum*, 23(2), 170–195. <https://doi.org/10.1525/mts.2001.23.2.170>
6. Genius English Translations. (2018, August 24). *BTS – IDOL* (English translation). Genius. <https://genius.com/Genius-english-translations-bts-idol-english-translation-lyrics>
7. Han, M.-Y. (1983). Religious origins of Korean music. In Korean National Commission for UNESCO (Ed.), *Traditional Korean music* (pp. 23–34). The Sisa-yong-sa Publishers.
8. Han, D., Gotham, M., Kim, D., Park, H., Lee, S., & Jeong, D. (2024). Six dragons fly again: Reviving 15th-century Korean court music with transformers and novel encoding [Preprint]. arXiv.
9. Han, D.-Y. (2018, March). Gugak and K-pop: Sharing the spotlight. Korean Cultural Center. <https://www.korean-culture.org/eng/webzine/201803/sub03.html>
10. Jenkins, H. (2006). *Convergence culture: Where old and new media collide*. New York University Press.
11. Jung, J., Kim, D., Lee, S., Cho, S., Soh, H., Bukey, I., Donahue, C., & Jeong, D. (2025). Unified cross-modal translation of score images, symbolic music, and performance audio (arXiv:2505.12863). arXiv. <https://doi.org/10.48550/arXiv.2505.12863>
12. Kim, D., Han, D., Jeong, D., & Valero-Mas, J. J. (2024). On the automatic recognition of Jeongganbo music notation: Dataset and approach (Version 1) [Preprint]. Research Square. <https://doi.org/10.21203/rs.3.rs-4668854/v1>
13. Lévy, P. (1997). *Collective intelligence: Mankind's emerging world in cyberspace* (R. Bononno, Trans.). Perseus Books.
14. Nissa, F. K., Herniwati, Triarisanti, R., & Azizah, A. (2021). Forms and translation procedures of Korean onomatopoeia and mimesis in the webtoon. In *Proceedings of the Fifth International Conference on Language, Literature, Culture, and Education*

(ICOLLITE 2021). Atlantis Press.
<https://doi.org/10.2991/assehr.k.211119.046>

15. Rockwell, C. J. (1983). Trends and developments in Korean traditional music today. In Korean National Commission for UNESCO (Ed.), *Traditional Korean music* (pp. 106–117). The Si-sa-yong-sa Publishers.
16. Simonetta, F., Ntalampiras, S., & Avanzini, F. (2019). Multimodal music information processing and retrieval: Survey and future challenges. In *Proceedings of the 2019 International Workshop on Multilayer Music Representation and Processing (MMRP)*.
<https://doi.org/10.1109/MMRP.2019.8665366>
17. Sousa, M. N. e, Martins, M. de L., & Zagalo, N. (2016). Transmedia storytelling: The roles and stakes of the different participants in the process of a convergent story, in divergent media and artefacts. In A. Lugmayr & C. Dal Zotto (Eds.), *Media convergence handbook* (Vol. 2, pp. 117–135). Springer. https://doi.org/10.1007/978-3-642-54487-3_7
18. Tang, M., Li, J., Yang, L., Zhang, Z., Tian, J., Li, Z., Zhang, L., & Wang, P. (2025). NOTA: Multimodal music notation understanding for visual large language model (arXiv:2502.14893). arXiv.
<https://doi.org/10.48550/arXiv.2505.12863>
19. Tran, M. L., Park, C., & Jung, J. H. (2023). Topological data analysis of Korean music in Jeongganbo: A cycle structure. *Journal of Mathematics and Music*, 17(3), 403–432.
<https://doi.org/10.1080/17459737.2022.2164626>