

Comparative Analysis of Compositions for Virtual Singers and Human Singers

Zeyun Chen^{1,a,*}

¹*Composition Department, Central Conservatory of Music, Beijing, China*

a. yuner050101@outlook.com

**corresponding author*

Abstract: This article discusses the similarities and differences between traditional vocal works and virtual singer works. It introduces the vocal characteristics of virtual singers, explores the differences between their performances and those of real singers, and analyzes the considerations that composers have for harmony when composing for virtual singers and real vocals, such as the way in which traditional functional harmonies and modern twelve-tone sequences are harmonic, and what the differences are, and composers' ideas for using different harmonies for different groups when composing. This article focuses on virtual singers Kafu and Hatsune Miku, as well as some representative works by the two virtual singers. Several songs written for real vocals will be analyzed, such as the Mass, Rossini's opera, and the vocal parts of Boulez's *Ownerless Mallet*. The impact of different vocal characteristics on the uniqueness of musical works, along with the technical differences and similarities between compositions for virtual and real singers, will be discussed.

Keywords: Virtual singers, vocal performance, harmony theory, serialism, twelve-tone technique.

1. Introduction

ACG (Anime, Comic, Game) culture, a prominent Internet subculture, has deeply influenced the lives of teenagers today [1]. Virtual singers, a component of this “two-dimensional culture” (Er Ci Yuan), have gained widespread attention. Virtual singers created by Japanese studios, characterized by their distinct vocal styles and appealing character designs, have introduced a new dimension to musical sound. Upon examining this phenomenon, the author observed that composers encounter both similarities and differences when composing for virtual singers compared to real vocalists. Likewise, the singing techniques of virtual singers often exhibit both resemblances and contrasts with those of real singers.

This paper uses Kafu, developed by Kami Mura Studio, and Hatsune Miku as case studies to explore these similarities and differences in vocal technique and compositional style. The discussion focuses on the unique vocal characteristics of virtual singers, the design of their vocal range, and the rhythmic elements in their performances. It also compares the harmonic functions of accompaniment in virtual singer works and traditional art songs.

2. Vocal Characteristics of Virtual Singers

Kafu, a virtual singer primarily based on advanced singing techniques, was developed through the Cevio AI platform, with her vocal identity derived from KAF. Her voice is characterized by high tension and expressiveness. In comparison to other virtual singers, Kafu's voice has a distinctive breathy quality, which lends greater realism and vocal tension to her performances. Moreover, her TA (Tension Adjustment) voice tuning creates a youthful and energetic tone, making her voice especially suited to a variety of musical styles. Her slightly hoarse timbre enhances the emotional impact of songs that express themes of love or laziness, making her performance style particularly unique.

Hatsune Miku, a virtual singer developed by Crypton Future Media and released on August 31, 2007, uses Yamaha's VOCALOID speech synthesis program. She has since become one of the most popular and widely recognized virtual singers. Hatsune Miku's popularity extends beyond music to a variety of cultural and creative products, and she is often portrayed as a fully developed character, adding to the immersive experience for users. VOCALOID 2.0 allows fans to compose their own songs and lyrics using Hatsune Miku's voice, encouraging creativity and imagination in the virtual singer community [2].

3. Differences Between Virtual Singers and Real Singers

The author has identified several key differences between virtual singers and real vocal performances. For instance, vibrato is a common technique in real vocal singing, but in virtual singers' performances, it is often difficult to achieve due to the mechanical nature of their sound production. Virtual singers are generally unable to naturally produce vibrato, as their vocalization is driven by algorithmic principles. For example, in certain vocal works that require emotional expression, real singers can easily convey feelings, whereas this is more challenging for virtual singers. In the song "Panis angelicus" (Figure 1), Zhang Xiangming suggests that a slower vibrato is necessary. The melody of this piece is sweet and delicate, expressing deep gratitude from ordinary people to the benevolent gods. Since the work is a lyrical, slow-paced church hymn, the singer is required to use controlled vibrato to convey these emotions effectively [3].



Figure 1: "Panis angelicus" is from the church hymn in the Soleri Mass.

In contrast, virtual singers do not produce such vibrato when holding long notes to express emotion. The hoarse sound quality of some virtual voices can even feel uncomfortable when paired with vibrato. For example, in "Asahi" (Figure 2), composed by Iori Kanzaki, when expressing love during the climax of the piece, vibrato is intentionally avoided. Instead, a smooth, sustained note is used. In the second measure, the note E reaches a high point of emotional expression, yet no vibrato is applied.



Figure 2: "Asahi".

When comparing real vocals and virtual singers, the author also found that composers approach the two differently in terms of vocal range. For virtual singers, composers often use extreme pitches and vocal ranges that are difficult or impossible for human singers to achieve. Thanks to machine-based production, these extreme ranges can add tension to the climax without causing strain or damage to the singer's vocal cords.

For example, in the piece "Elite" by Chinozo (Figure 3), the composer repeatedly loops the notes at the soprano's passaggio (voice transition point) and pushes the highest note to an unattainable register for real singers. A typical soprano experiences a transition around F2, and it is rare in real vocal music for the melody to hover around this note. Additionally, the work makes extensive use of the B2 note, which is highly challenging for real singers but easily performed by virtual singers.



Figure 3: "Elite".

In real vocal compositions, composers tend to avoid continuously oscillating around the vocal passaggio, instead keeping the melody below or just above the transition point. This prevents vocal fatigue. For example, in "Una Voce poco fa" (Figure 4), the melody uses modular scales, ornamentation, and triplets. The scales alternate, creating a smooth and manageable singing experience, requiring good breath control and clear diction [4].



Figure 4: "Una Voce poco fa" [5].

The author also found that virtual singers are often used to perform at speeds that are difficult for human singers to achieve. This contributes to a futuristic, mechanical sound that captivates audiences.

Virtual singers can execute extremely fast musical passages that would otherwise be exhausting for real singers, reducing rehearsal time and the potential for performance errors. For example, Hatsune Miku's "The Disappearance of Hatsune Miku" has become a classic virtual singer work. The tempo of this piece reaches 245 beats per minute (BPM), far exceeding the human limit of 208 BPM. This super-high-speed song cannot be performed by human singers. The rapid tempo evokes a sense of parting and sadness, with the song disappearing just as quickly as it is sung. This feeling of futurism and emotional detachment is achieved without the discomfort or vocal strain that would affect a real singer. The song reaches an instantaneous speed of 720 BPM, with 12 notes per second—far beyond the limits of human vocalization. Until 2018, it was considered the fastest "farewell song." This genre of ultra-fast electronic music has since become a hallmark of virtual singer performances [6, 7].

4. Harmonic Similarities and Differences Between Virtual Singers and Art Song Vocals

In comparing the vocal works of virtual singers with art songs, the author has identified some key similarities and differences in harmonic techniques. Most works written for virtual singers are intended to cater to popular tastes, featuring tonal music with clear harmonic progressions and resolutions. This creates a sense of anticipation for the listener, making the melodies catchy and widely appealing. For instance, in the song "Don't Say Goodbye" (Figure 5), the main melody transitions smoothly into the chorus through the use of traditional harmonic techniques. The use of a standard cadential formula to conclude the verse gives the audience a strong sense of expectation.

The image displays a musical score for the song "Don't Say Goodbye". It features two staves: a vocal staff (S.) and a piano staff (Pno.). The key signature is D major (two sharps) and the time signature is 4/4. The vocal line begins at measure 22 with the lyrics "明日 が来 ないように 祈 る そ れ だ け". The piano accompaniment consists of chords. Below the piano staff, the chord progression is labeled as TSVI, S, K, D7, and T.

Figure 5: "Don't Say Goodbye".

On the other hand, compositions for real vocals often lean towards avant-garde and experimental approaches. Some 20th-century composers, for instance, have favored a more logical, structured form of composition. Since Schoenberg introduced the twelve-tone technique, many vocal composers have adopted this method, including Pierre Boulez. In his work "The Hammer Without a Master", the contralto section completely abandons traditional harmonic systems. Boulez uses sequences to create music that emphasizes logical construction, contributing to the development of avant-garde music. The following is an excerpt from the contralto section of "The Hammer Without a Master" (Figure 6).

III

«l'artisanat furieux»

Modéré sans rigueur (♩ = 84)

Flûte en sol

Voix d'Alto

Fl. en sol

Voix

La rou - lot - te rou - ge au

bord du

clou Et

Figure 6: "The Hammer Without a Master".

Unlike traditional harmonic resolutions, "The Hammer Without a Master" employs a monolithic twelve-tone sequence (Figure 7), which rejects traditional tonal concepts. There is no final cadence, and the work resists the usual directional pull of classical harmony. This contrasts sharply with virtual singer works like "Don't Say Goodbye", which adhere to tonal harmony and melodic conventions.

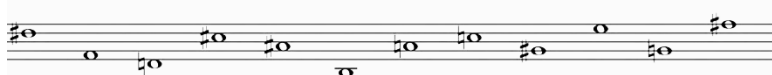


Figure 7: Twelve-tone sequence from "The Hammer Without a Master".

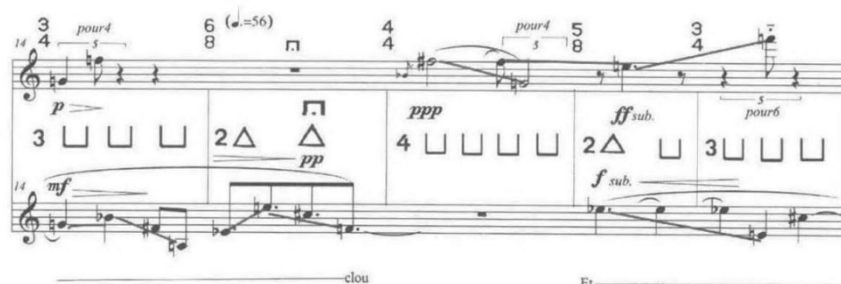


Figure 8: "The Hammer Without a Master", mm. 14 - 18.

In the third movement of "The Hammer Without a Master", Boulez designs a twelve-tone sequence for the mezzo-soprano section. In measures 6-8, we see a descending chromatic progression (G-sharp to F) in the mezzo-soprano voice. In measure 14, B-flat and A form a descending major seventh, followed by an ascending minor ninth between B-flat and E-flat in measure 15 (in Figure 8). Boulez manipulates pitch relationships systematically, utilizing sequences to create intricate pitch structures and connections [8].

Through this analysis, the author notes that Boulez selectively uses the twelve-tone sequence as a resource for pitch material while preserving the integrity of the sequence itself. In contrast, "Don't Say Goodbye" employs traditional harmonic language, designed to create a catchy melody that resonates with a broad audience. The latter is suitable for mass consumption at public events or for emotional comfort, providing listeners with an intuitive understanding of popular harmony. Boulez, on the other hand, writes melodies using twelve tones, reflecting a more mathematical and physical approach to pitch relationships. This method underscores the ongoing evolution of musical development and the attention to detail in connecting pitches. Such techniques have laid the foundation for the advancement of classical music.

5. Conclusions

With the rapid development of technology and society, virtual idols have introduced new possibilities for vocal performances. In the past, composers composing for real singers were constrained by the limitations of human physiology, particularly regarding vocal range and rhythmic complexity. As a result, some musical works could not meet the expectations of fans. However, the emergence of virtual idols has transformed the relationship between fans and performers, shifting from "I support what idols do" to "I decide what idols do". This change has expanded the scope of musical production, allowing virtual idols to exist as unique entities separate from their human counterparts.

By comparing the works of virtual singers with the traditional art song "The Hammer Without a Master" in terms of harmony, melody, and structure, the author observes that differences in vocalization methods and audience expectations lead to distinct compositional demands. Composers composing for virtual singers and those composing for real vocals approach their work with different perspectives, intentions, and cultural or educational backgrounds. These differences have created a clear distinction between virtual singer compositions and traditional art songs.

The author believes that with the continued growth of virtual singers, more traditional composers will explore composing for virtual vocals. Composers should strive to find the most appropriate approach to utilize the strengths of both real and virtual singers, depending on the specific needs of the work. An excellent composer can excel in both fields, contributing to the advancement of contemporary music while also preserving traditional compositional techniques. The author hopes that as society evolves, audiences will not only enjoy virtual singer music but also appreciate art songs, achieving a balanced musical landscape.

References

- [1] Ye, Runze, Xu Ganli. (2024) *The Two-Way Interaction Between Mainstream Culture and ACG (Anime, Comics, and Games) Subculture: A Case Study of the Virtual Singer Luo Tianyi*. *Youth Journal* (1), 103-111.
- [2] Yang, Yuxuan. (2023) *From Lynn Minmay and Hatsune Miku to A-SOUL: The Generational Evolution and Localization of Virtual Idols* [J]. *Audiovisual*, (2):124-128.
- [3] Zhang, Xiangming. (2012) *On the Role of "Vibrato" in Vocal Performance*. [D]. Fujian Normal University.
- [4] Zhao, Na. (2011) *A Study on the Vocal Characteristics of Coloratura Mezzo-Soprano*. [D]. Shaanxi Normal University.
- [5] Xu, Yihong. (2012) *Japan's New Idol—Hatsune Miku* [J]. *Japanese Knowledge*, 2012(12): 36-37. DOI:10.3969/j.issn.2095-4948.2012.12.021.
- [6] Dan Wei, (2018) *A Study on the Cultural Phenomenon and Communication Characteristics of Virtual Music Characters in the Context of New Media*. [D]. Shenzhen University.
- [7] Wu, Jixue. (2017) *The Transformation of Sequences—An Initial Exploration of the Pitch Organization in Pierre Boulez's Chamber Work "Without Owner's Hammer."* *New Voices in Music: Journal of Shenyang Conservatory of Music*, 35(1), 10.
- [8] Wang, Shuangfeng. (2019). *Why Are Virtual Idols So Popular?* [J]. *Shanghai Informatization* (10), 3.