

Comparative Analysis of Compositions for Virtual Singers and Human Singers

Zeyun Chen

*Central Conservatory of Music, Beijing, China
yuner050101@outlook.com*

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.

For this discussion, Kafu, developed by the Kami Mura Studio, and Hatsune Miku will be used as case studies. The similarities and differences in the singing techniques and composition styles of virtual singers will be examined. The distinct vocal styles of virtual singers, the structure of their vocal ranges, the rhythmic aspects of their vocals, as well as the harmonics of accompaniment in virtual-singer pieces and traditional art songs, will all be considered.

2. Vocal characteristics of virtual singers

Kafu, a virtual singer focusing on sophisticated vocals, was created on the Cevio AI platform, with her voice characteristics taken from the KAF. Her voice is high-tension. As a virtual singer, Kafu's breathy sound makes her voice more realistic. Also, her TA (Tension Adjustment) voice tuning has a youthful sound. The uniqueness of her voice stems from her hoarse quality, which makes her singing sound more emotional when she sings love or laziness lyrics. The uniqueness of her singing style is most exceptional.

Hatsune Miku, created by Crypton Future Media on August 31, 2007, employs Yamaha's 'VOCALOID' speech synthesis software. Hatsune Miku was soon recognised as one of the most famous virtual singers. Hatsune Miku has been famous not only as a singer but also as an element of various forms of creative products in Japanese culture. Hatsune Miku can also be regarded as a fully realised character, enhancing her users' experience. 'VOCALOID 2.0' enables users to create their own songs as well as lyrics with Hatsune Miku's voice, which promotes creativity among virtual singers [2].

3. Differences between virtual singers and real singers

The writer has identified some of the main differences between virtual singers' performances and real singing. For example, in real singing, vibrato is an important concept. However, in the case of virtual singers' performances, it is hard to produce this effect due to the mechanical sound-production process used by virtual singers. For this reason, it can be concluded that virtual singers cannot produce vibrato naturally. The main reason is that the production of sound in virtual singers' performances is dependent on mathematical formulas. For this reason, according to the song "Panis angelicus" (as illustrated in Figure 1), singers must produce vibrato slowly. The song is melodic, sweet, and slow. For this particular song, which acts as a church hymn, the singer must produce vibrato delicately [3].

The image displays a musical score for the piece "Panis angelicus" from the Soleri Mass. The score is written in G major and 4/4 time, with a tempo marking of "Andante con moto". It consists of piano accompaniment and vocal lines. The lyrics are provided in both Chinese and Latin. The Chinese lyrics are: "天神賜食糧，手芸芸眾生。" and "上天的賜予人類得以共享。". The Latin lyrics are: "Pa - nis an - ge - li - cus fit pa - nis ho - mi - num." and "Dat pa - nis coc - li - cus fi - gu - ris ter - mi - num". The score includes dynamic markings such as *mp*, *p*, *rit.*, *p*, *rit.*, *mf*, and *f*. The piano part features a steady accompaniment with some melodic lines, while the vocal part is more melodic and expressive. The score is numbered 1 at the bottom right.

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"

While comparing real vocals and virtual singers, the writer discovered that composers interact differently with both, particularly regarding vocal range. For virtual singers, composers tend to opt for extremely high or low pitches, which would be difficult or impossible for human singers. However, with machine production, it becomes easy to create tension during the climax of a song without harming the human singer’s vocal cords.

For instance, in the song “Elite” by Chinozo (Figure 3), the composer extends the passage of the soprano’s voice transition note by continually repeating it. The transition of a normal soprano voice happens at F2, while it is unusual to find a song revolving around this note. The song also utilises the B2 note extensively. The B2 note is difficult in reality, while it can be accomplished easily by a virtual singer.



Figure 3. "Elite"

However, in real-world vocals, the composer does not use an oscillating pattern of the vocal passaggio. The pattern below or slightly above the transition region is more common. The key is avoiding vocal fatigue. “Una Voce poco fa” (Refer to Figure 4) utilises modular scales, ornamentations, or triplets. The scales vary. This ensures smooth vocals, which require efficient breath control [4].

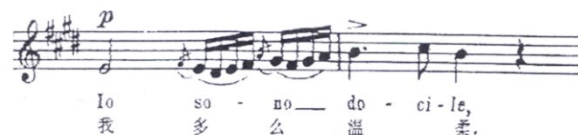


Figure 4. "Una Voce poco fa" [5]

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

While comparing the musical compositions of virtual singers with art songs, the writer has identified various similarities and divergences in the use of harmonies. Most musical pieces composed for virtual singers are aimed at appealing to society's prevailing musical interests, which often entail tonal pieces with defined harmonies. This always leaves the listener with a feeling of anticipation, making the musical pieces catchy. For example, in the musical piece "Do not Say Goodbye" (Figure 5), the key musical piece gradually shifts into the chorus section by using conventional harmonic techniques. The application of the conventional cadential formula at the end of the verse always leaves the listeners feeling there is an element of expectation.

The image shows a musical score for the song "Don't Say Goodbye". It consists of two staves: a vocal line (S.) and a piano accompaniment (Pno.). The key signature is three sharps (F#, C#, G#) and the time signature is 4/4. The score starts at measure 22. The vocal line has the lyrics: 明日 (Ashita) が来 (き)ない (ない)ように (ように) 祈 (いの)る (る) それ (それ) だけ (だけ). The piano accompaniment features chords labeled 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 (♩ = 83)

The score consists of four systems, each with a Flute (Fl. en sol) and Voice (Voix) part. The tempo is 'Modéré sans rigueur' with a metronome marking of ♩ = 83. The music features complex rhythmic patterns with various time signatures (3/8, 4/4, 4/8, 3/4, 2/4) and dynamic markings such as *mf*, *mp*, *pp*, *ppp*, *ff*, and *f sub.*. The lyrics are: 'La rou - lot - te rou - ge au bord du'. The score includes numerous rhythmic notations like 'pour 4', 'pour 5', and 'pour 6' with corresponding note values and rests.

Figure 6. "The Hammer Without a Master"

Unlike traditional harmonic resolutions, "The Hammer Without a Master" employs a monolithic twelve-tone sequence (Figure 6), 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.

A single line of musical notation showing a twelve-tone sequence. The notes are: G#4, A4, B4, C5, D5, E5, F5, G5, A5, B5, C6, D6. The sequence is presented in a non-linear, monolithic fashion, rejecting traditional tonal concepts.

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

Musical score for measures 14-18. The top staff is for the Flute (Fl. en sol) and the bottom staff is for the Voice. The tempo is 'Modéré sans rigueur' with a metronome marking of ♩ = 56. The score includes dynamic markings such as *p*, *pp*, *ppp*, *ff sub.*, and *f sub.*. The lyrics are: 'clou Et'. The score includes rhythmic notations like 'pour 4', 'pour 5', and 'pour 6' with corresponding note values and rests.

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].

From this analysis, the writer highlights how Boulez uses the twelve-tone sequence as a source of pitches while maintaining its integrity. However, “Do not Say Goodbye” uses conventional harmony, aiming to produce a catchy tune that appeals to many. The second piece is fit for consumption by the masses, preferably at events where listeners can gain insight into popular harmony. However, Boulez uses twelve tones in composing melodic pieces. The composer’s approach highlights the ever-evolving nature of developments in classical musical composition. The focus on pitches shows how musical efforts have helped create progress in the world of classical musical composition.

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 compositions of virtual singers with the traditional art song “The Hammer Without a Master” in terms of harmony, melody, and composition, the writer notes that the various approaches to vocalisation practices and listener expectations place different demands on musical composition. The composers of pieces performed by virtual singers, as well as those performed by real vocals, work with different perspectives. This has resulted in an observable distinction between compositions by virtual singers and traditional art songs.

The writer thinks that as the number of virtual singers continues to rise, more traditional composers will be interested in writing arrangements for them. Composers will develop strategies for making the most of both the benefits of virtual and human singers. A talented composer can do both well, thereby benefiting the world of modern music and traditional composition methods. The writer hopes that, as time passes, the world will enjoy not only virtual singer music but also art songs.

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.