

УДК 78.07
ББК 85.313(3)

DOI: 10.30628/1994-9529-2021-17.1-173-213
received 02.07.2020, accepted 31.03.2021

MELISSA-LILI ARENDT

Martin-Luther-University Halle-Wittenberg,
Halle (Saale), Germany
ResearcherID: AAV-7230-2020
ORCID: 0000-0003-0792-2780
e-mail: melissaliliarendt@gmx.de

NATALIA NOWACK

TU Dortmund University,
Dortmund, Germany
ResearcherID: AAV-4888-2020
ORCID: 0000-0002-2442-6167
e-mail: natalia.nowack@web.de

THE VIRTUAL ARTIST. *VOCALOID* AND AN EXPERIMENT ON ITS PERCEPTION

Abstract. Over the past decade the voice synthesiser *Vocaloid* has gained great popularity in Japan. Critics debate on whether *Vocaloid* can be called a new musical instrument and if its creation marks a new era in (Japanese) popular music. The unique characteristic of a *Vocaloid* is its Alter Ego, its “virtual shape”, which is illustrated like common anime or manga characters such as Sailor Moon or Kamikaze Kaito Jeanne, thus owning not just a name but also human-like features. It was due to the use of avatar images that *Vocaloid* managed to spread across the world via the internet. The most famous anthropomorphic singer is called Hatsune Miku and is known for her long turquoise-coloured hair tied up in two ponytails. Being the uncrowned princess of Nico Nico Douga—the Japanese equivalent of YouTube, Miku opened up the pathway for even more *Vocaloids* like Luka Megurine or Rin and Len Kagamine. The contiguity of the new creating type to the “screen arts” is maintained by its own existence in the sphere of digital media. This new phenomenon is not very known in Europe, so the question on the principles of its existence is a vital one to discuss. Furthermore, we shall test

how its music and appearance are perceived by people who do not belong to the fandom. How does a Non-Japanese listener react to a musical performance done by a computer program? One of the first answers to this question can be found in a study, which was conducted at the Martin-Luther-University in Halle (Saale), Germany. It focused on the reactions of the listeners. Although Europeans do not show as much interest in artificial intelligence as the Japanese, the test subjects showed great sympathy towards the singing program.

The essay's content is divided into four parts. It begins with a contemplation of Vocaloids sociological aspects (1), followed by the introduction of a self-produced classification of its performances (2) and continues with an explanation on how the experimental research was conducted (3). The last part contains a summarised presentation of the results and a perspective on future research (4). The authors claim this research to be one of the very first tries to shed light on how the popularity of this new musical phenomenon can be explained.

Keywords: Vocaloid, synthesiser, Hatsune Miku, Japanese pop music, virtual idol, avatar, perception

МЕЛИССА-ЛИЛИ АРЕНДТ

Галле-Виттенбергский университет имени Мартина Лютера,
Галле, Федеративная Республика Германия
ResearcherID: AAV-7230-2020
ORCID: 0000-0003-0792-2780
e-mail: melissaliliarendt@gmx.de

НАТАЛИЯ НОВАК

Дортмундский технический университет
Дортмунд, Федеративная Республика Германия
ResearcherID: AAV-4888-2020
ORCID: 0000-0002-2442-6167
e-mail: natalia.nowack@web.de

ВИРТУАЛЬНЫЙ АРТИСТ. VOCALOID И ЭКСПЕРИМЕНТ ПО ЕГО ВОСПРИЯТИЮ*

Аннотация. Уже свыше десяти лет в Японии пользуется большой популярностью компьютерная программа *Vocaloid* («Вокалоид»), синтезирующая

* Русский текст статьи оказался несколько подробнее, чем английская версия.

певческий голос. Критики спорят о том, можно ли называть *Vocaloid* новым музыкальным инструментом и считать его создание началом новой эры в (японской) поп-музыке. Уникальной особенностью вокалоида является его Alter Ego, его визуальный образ: синтезируемые голоса получают не только имена, но и внешность в стиле персонажей аниме и манга, наподобие Сейлор Мун (*Sailor Moon*) или Неуловимой воровки Жанны (*Kamikaze Kaito Jeanne*), и, таким образом, одушевляются.

Именно благодаря аватарам программа *Vocaloid* распространилась по всему миру через интернет. Самую популярную человекоподобную исполнительницу, девушку с длинными бирюзовыми волосами, собранными в два хвоста, зовут Мику Хацунэ (*Hatsune Miku*). Некоронованная принцесса видеохостинга Nico Nico Douga — японского аналога YouTube, Мику открыла путь, по которому последовали другие вокалоиды, например, Мегурине Лука (*Luka Megurine*), Рин и Лен Кагамине (*Rin and Len Kagamine*). Близость к искусствам экрана обеспечивается самим бытованием новой разновидности творчества в сфере цифровых медиа.

В Европе новый феномен известен пока относительно мало. Тем интереснее оказался вопрос о принципах его бытования. Кроме того, не стоит упускать возможности отследить его восприятие в кругах, в фэндоме еще не втянутых. Как реагирует слушатель не из Японии на музыкальное выступление, сгенерированное компьютерной программой? Одним из первых на эти вопросы отвечает исследование, проведенное университетом имени Мартина Лютера в Галле (*Halle*), Германия. Хотя европейцы не проявляют к искусственному интеллекту такого сильного интереса, как японцы, испытуемые продемонстрировали выраженную симпатию к программе, имитирующей пение.

Статья состоит из четырех частей. После рассмотрения некоторых особенностей феномена с социологической точки зрения (1) следует описание ряда его признаков на основе специально предпринятой классификации (2). Затем описывается процедура экспериментального исследования (3). В заключительной части статьи обобщаются выводы и обозначаются перспективы на будущее (4). Авторы рассматривают данную работу как первую ступень в ряду исследований, призванных пролить свет на загадку успеха нового явления.

Ключевые слова: вокалоид, синтезатор, Хацунэ Мику, популярная музыка Японии, виртуальный кумир, аватар, восприятие музыки

The voice synthesiser software *Vocaloid*, which was developed at the Pompeu Fabra University Barcelona together with Yamaha Corporation, was publicly presented in the year of 2003.¹ After just one year, the first version of the software was quickly brought to the market. It is remarkable that

¹ Speech synthesis is more complex compared to sound synthesis.

the first version is based on the English language, albeit that the speech synthesis of the Japanese language is much easier to conduct.² Since there was not much interest in such a unique product at the time the sales figures stayed at a relatively low level.³ Yamaha then decided to involve third parties, which used anime avatars for a better marketing strategy. This strategy led to resounding success: the character Hatsune Miku was found. Developed by *Crypton Future Media* and released in 2007, “she” managed to gain enough popularity to be called the first virtual idol of popular music history. In contrast to the previous Vocaloid models Miku was sold 30 000 times in the first half year, which is a great sales figure for any kind of synthesiser. Next to the Japanese and English voice banks, the Vocaloids have now also “learned” to sing in Korean, Spanish and Chinese.

There are two kinds of Vocaloid performances. One being the “classic” live concert as a hologram projection on screen, which dances and sings for its fans. Most likely, the whole audience is equipped with glow sticks to dance together in unison with the music and the virtual artist. The second kind of performance can be watched on social media, for example YouTube. There you can find the original story-telling music videos, which were animated either by fans or actual illustrators. Not only does the archetype of a hero encourage the fantasy of the composer, but it also limits itself to one simple motif or idea. Vocaloid-Songs without an optical component are rarely ever heard of, therefore one can describe music productions with Vocaloid predominantly as a synthesis of auditory and optical aspects.

Although being a media-effective phenomenon, there has not been much research activity on the topic just yet. Most of the available studies are from Japan or South Korea, with the main focus on technological or medial aspects of Vocaloid (H. Kenmochi, T. Shiba, P.W. Galbraith & J.G. Karlin, and a few others) [1, 2, 3]. The first US-American and European publications view Vocaloid in the context of recent cultural development and ethical issues posed by this new technology (P. Davison, N. Eidsheim) [4, 5 and 6, pp. 115–150]. However, aspects

² The Japanese language is based on syllables; therefore, it is much easier to record a singer’s library. All voice synthesisers are optimised in one certain language. For example, if one uses English lyrics and combines it with a Japanese voice bank one can definitely hear a bit of an “accent”. The first synthesiser, which was released with a Japanese voice bank, was MEIKO.

³ Remarkably, one of the first commercially available Vocaloid albums was recorded in Russia. It contains tracks in Russian and English: “M’esto pod solntsem [A Place in the Sun]: Aleksey Ustinov, Aleksey Kosolapov feat. LëLëk [LEON]”, Label Virartech, Novosibirsk 2006.

of speech synthesis already were subjects of many international studies for a long time now (e.g. P. Shapiro) [7]. Yet, their significance exceeds the field of interests of music historians.

The purpose of this essay is to give a general characteristic of Vocaloid as a new musical event, including a classification of the compositions for this synthesiser. In addition, the authors present data, which was collected after conducting an experimental study in terms of music perception. The reactions of the test persons were examined in an area, which was not involved with the “culture” of Vocaloid.

New instrument or new genre?

The influence of Vocaloid and its most popular character Hatsune Miku shows great impact on the Japanese music industry. The Japanese society generally demonstrates a high affinity towards artificial intelligence and robotics. Furthermore, there may be observed a certain fondness for symbols and mascots. Miraitowa and Someity, the mascots for the Olympic Games 2021 in Tokyo, can be mentioned as an example of combinations between robotics and mascots. Not only did they start out as a cartoonish illustration, they also spread across the nation in the shape of a robot [8]. Meanwhile, South Korea has also been working on virtual artists and uses almost the same kind of technology to commercialise their internationally known K-Pop-Stars and even video games, such as *League of Legends* [9, 10].

The development of Vocaloid seems promising for music sociologists as well, because of the gradual emergence of a new musical instrument as a current level of artistic development, and the answer of the music industry for its demands. The latter is more or less obvious enough, since the industry needs new innovations, which can be easily accepted by our youth. Today, our youth lives in its own digital dimension [11, p. 552].⁴ In this digital dimension one can find content designed in a joking manner, such as personality tests to find out which Vocaloid resembles your persona the most.⁵ But one is also able to digitally follow the next musical events, which were already executed

⁴ Some sociologists and cultural scholars think that digitality is now spreading across society [12, 13].

⁵ URL: <https://www.testedich.de/quiz57/quiz/1541174128/Welcher-Vocaliod-bist-du> (accessed 13.4.2020).

outside of Japan, for example, in Toronto, Los Angeles, Dallas, even rarer in Paris, London, Cologne and ... Moscow. Remarkably, all tickets in Moscow were sold out instantly making it impossible to find any leftover tickets.⁶ The last facts provide serious evidence for the “export” of the new culture. One begins to consider the question of the educational use of the phenomenon [14].

Whether on stage or on screen, the Vocaloid music becomes a new kind of mass art. The role of representation ought to be the most significant when comparing it to common music products [15, p. 188]. The virtual artist seems to need a much greater amount of personality traits, more than an actual human singer. The paths of communications are separating. Firstly, we have the effect of music, and (at the same time) the effect of the underlying character. This corresponds to the five basic axioms of Paul Wazlaweck’s theory of communication. Many years ago, he stated, that it is not possible “not to communicate” [16, p. 51].⁷

Nevertheless, Vocaloid is only a computer software with specific technical features. The use of this software varies between many aspects: composing with specific effects on the main back vocals, which is executed by professional musicians, or simple tryout songs, which are written by fans, who may just want to enhance the universe around Hatsune Miku. Also, this software was already used by artificial intelligence. *Honda’s* humanoid robot ASIMO uses this software to sing popular songs.

The marketing strategy is remarkably interesting, as the voice of Hatsune Miku and all other Vocaloids, that are being sold, are publicly available. The new division of roles in terms of music production (the composer is an artist; however, the virtual singer is not), shows its connection to the present. The copyright of songs, which were produced with the software, belong to the composer alone. Therefore, the software is equal to every other synthesiser. The shape of Miku (Fig. 1) belongs to its creator *Crypton Future Media*, who published her image in the year of 2012 under the free license (CC-BY-NC).

⁶ Moscow, 25.03.2012. Also, there was a public screening of “Opera AOI”, a performance combining Vocaloid songs with Japanese puppet theatre, also known as Bunraku.

⁷ Today Wazlaweck’s theory is still prominent, even in the field of computer science.



Fig. 1. Miku⁸



Fig. 2. KAITO⁹

The program is easy to use. In principle, the user types music notes and song texts into the software and obtains a sounding product. Alternatively, one can also use a MIDI keyboard to insert music notes. The file output can be either .wav-file or VST. An unlimited amount of interested composers¹⁰ get the opportunity to use an own artist. The expression “individual art industry” (E. Dukov) describes perfectly how this new model functions [17, S. 176]. The dream of finding a good musician, who is not just talented, but also provides good characteristics, finally seems to have come true. This might just be the biggest advantage for music production with artificial voices, because the composer and musician will not be able to argue against each other (reminiscent of *The Beatles* or *The Strokes*). By now, there are over 100 Vocaloid products and even more fan-created voice banks, which run by the software name of UTAU. Users of UTAU are able to decide on how the character should look like and how the singer’s voice shall be altered.

From a sociological point of view, this new phenomenon seems very promising due to its social advantages. Being an instrument and not just a

⁸ Hatsune Miku, © Crypton Future Media, Inc. 2007, licensed under a CC BY-NC: URL: <http://creativecommons.org/licenses/by-nc/3.0/>. See the image source: URL: https://ec.crypton.co.jp/pages/prod/virtualsinger/cv01_us (accessed 21.03.2021).

⁹ KAITO, © Crypton Future Media, Inc. 2006, licensed under a CC BY-NC: URL: <http://creativecommons.org/licenses/by-nc/3.0/>. See the image source: URL: <https://vocaloid.fandom.com/de/wiki/KAITO> (accessed 21.03.2021).

¹⁰ According to Crypton, there were around 100 000 published songs, 170 000 videos and one million illustrations five years after Miku’s release. URL: https://ec.crypton.co.jp/pages/prod/vocaloid/cv01_us (accessed 21.03.2021).

synthesiser, Vocaloid means a lot more, since it does offer a persona. This persona has a voice with its very own timbre (or even many timbres), character traits and even its own repertoire. Although the companies try to develop very individual traits for each Vocaloid, its artificial music has one aspect in common with “real” music. Its development is being led by the beauty of the human voice. As for Japanese music, male and female voices are generally very high-pitched.

Calling the entirety of Vocaloid music a synergetic music genre would lead to the suspicion, that one cannot find any stylistic homogeneity. Many productions do differ a lot in terms of lyrics and music, and the general atmosphere does not play a major role. Although the majority of productions has a rather playful sound, there are also songs with serious content, and thus provoke thinking. The song *Hello, how are you?* depicts the inner struggle of a timid and strongly introverted person. While the tonal structure emphasises the verbal expression by using a lot of pauses and constantly changing directions (this can be heard very well in the intro, not included in ex. 1), recreating the feeling of insecurity, the sound of Miku’s voice and her piano accompaniment suggest a very lively, happy mood. This contradicts the lyrics strongly, which tell us about the protagonist’s insurmountable problems. The optical presentation of the music video helps to understand the content better, as it shows a crying girl.



Example 1. Hoehoe-P, *Hello, how are you?* Verse, self-made transcription of the singing voice

“Hello.” I opened the window, and softly whispered out. “How are you?” —

There’s no one here, I am alone in the room. Morning. The morning’s come, a morning of pouring rain. Tick-tock, somebody wind up my coil? [...] Why are you hiding away? Are you afraid of being laughed at? Don’t you want to meet someone? Do you really want that? Trapped in a sea of ambiguity, it’s too painful for me to even breathe. I just wanted to hear some voices.

I am really weak.

Text: Hello, how are you? Verse (until [...]), Chorus, translated from Japanese by M.-L. Arendt

The relation between lyrics, music and storyline is a very different one in the next Song, *Lost One's Weeping*. Here we can find a stronger unity between music and lyrics. Constant return of the melody to the same tone (*F sharp*) symbolically shows the impossibility of overcoming certain frameworks. That emphasises the underlying concept of the song — the dissatisfaction with the hard (Japanese) school system. The pressure on learning and perfectionism as a life principle are confronted with real emotions of a student. These emotions adopt the shape of a *doppelgänger* of the student. The conflict can also be seen in the animated music video.



Example 2. Neru, *Lost One's Weeping*, Chorus, self-made transcription

Can you read the Kanji on the board? Can you read that child's imagination?
 Tell me, who dyed his heart black? Who was it, who was it?
 Can you solve this equation on the abacus?
 Can you loosen the rope around that child's neck?
 Is it okay, if we remain like this?

Text: *Lost One's Weeping*, Chorus, translation from Japanese by M.-L. Arendt

These two examples suggest the thought, that an equation of optical design, sound effect and verbal content is not possible. However, artistic solutions allow a great amount of individuality. This model is not new, in fact, it is already known in film music. The song *Hello, how are you?* uses a contradictory relation between music and text, whereas the conflict in *Lost One's Weeping* is shared in the storyline.

Should these differences in terms with the artistic solution be proof enough to deny Vocaloid-Productions its right to call itself an own genre of music? Obviously not. One can find similar constellations throughout the history of music: hybrid-cultures such as *fusion*, *video game music* or *gothic*. It is only about the determination of one musical or non-musical feature, which can be used to generalize its existence. So, why not choose the direct path and use the term of the sociological kind—vocal-instrumental music “from user”,

without the medium of an artist? It would be interesting to clarify who this user is, what is the secret of the collective intelligence, which has in this case an enviable productivity.

From the authors' perspective, one of the arguments in favor of the new genre is the fact, that one is able to classify the music produced with Vocaloid.

Classification

Due to the fact, that Vocaloid has not yet been studied thoroughly, Melissa-Lili Arendt suggests the first model of classification, to give a better insight on the music. The classification does not focus on available musical categories. Instead it focuses on the content- and stylistic qualities of the text, which determine the overall design of the song.¹¹ In this suggested model there are five classes: **entertaining**, **epic**, **lyrical**, **socially critical** and **mixed** (Image 3). The first two classes differ a lot from each other. The earlier discussed songs belong into the lyrical and socially critical types.

Entertaining	Epic	Lyrical	Socially critical	Mixed
Party songs (EDM), suggesting a fun atmosphere	Songs show structure of a fairytale or legend	High focus on emotions (e. g. heartbreak)	Explicit social critical point of view (e. g. school system)	Mix of two classes
e.g.: One-Cup P, <i>Koneko no Payapaya</i>	e.g.: Mothy, <i>Daughter of Evil</i>	e.g.: Hoehoe-P, <i>Hello, how are you?</i>	e.g.: Neru, <i>Lost One's Weeping</i>	e.g.: Utata-P, <i>This is the Happiness and Peace of Mind Committee</i>

Fig. 3. Classification of titles, which are performed by Vocaloids

The type “entertaining” contains all songs, which fit the genre of Electronic Dance Music (short: EDM). These songs last no more than five minutes, display a variety of cute and childish content, and encourage the listeners to dance. It might just be the most important class of this music genre, since these songs spread more easily via popular social media sites, which can lead to an important stepping stone for composers. Most of these

¹¹ In the process of creating the classification, 70 Songs were analysed in total.

songs do occur in Japanese charts and are sung quite frequently in karaoke bars [cf.: 18]. These titles show a fairly homogenous design with their fast rhythms, which accompany a simple melody. The text is usually very amusing, sometimes a fond reminiscence of nursery rhymes. The visuals are clear and easily understood.

The type “epic” is very much complex, because all songs are structured like fairytales, that correlate with each other. The meaning of the lyrics is significantly important, since it also contains formulations, which are only used in fairytales. The story will be split into several songs, which are sung by different Vocaloids. This allows a different point of view to the respective content.¹² Telling a story in a fairytale style does take time, therefore, the songs last for a maximum of eight minutes. The design of this kind of composition is very different compared to other classes. Songs with contents of Japanese folklore are usually combined with a set of synthesised traditional Japanese music instruments. The instruments and the vocal style change respectively when the plot is set in a western world. The optical design changes as well; the singing Vocaloids wear historic garments, depending in which area of the world the plot is set.

The class “lyrical” describes all situations, in which emotions play a major role, such as heartbreak or loss of a loved one. The genres vary widely, and the optical design uses allegories. The profound lyrics are one of the most important components, because they lead to a large scope of interpretation. Because of the ambivalent interaction between lyrics, music and visuals emotions are not explicitly disclosed. Therefore, recipients are forced to put themselves into the protagonist’s place.

The elements of “socially critical” are very similar to those belonging to the class of “lyrical”, except there is a focus on socio-critical claims, often exaggerating certain conflicts, forcing the recipient to reflect critically on the conflict. The reflection, which, at first, occurs in the sphere of art wanders to the sphere of communication, thus leading to many open discussions on social media. The solutions to the conflicts merge with the lyrics into the visuals, while the musical structure maintains its neutrality. Music videos often use strong dramaturgical means.¹³

¹² A good example for a cyclical solution is the series *Nanatsu no Taizai (The seven deadly sins)* created by composer Mothy.

¹³ Cf. *Hana Ichi Monme* by Team Kamiuta: URL: <https://www.youtube.com/watch?v=b2ieL0gjmfs> (accessed 21.03.2021).

The last class, “hybrid” shows a mixture between two classes, such as “entertaining” and “lyrical” or “lyrical” and “socially critical”.

Study

Against the background of the situation in Japan and South Korea, it is important to research whether inhabitants of other countries are just as open towards this phenomenon. It might just be a matter of time until the great rapture of Vocaloid takes over in western countries. The purpose of the study ($N = 75$) was to examine the perception of the recipients listening to selected examples of Vocaloid songs. This exploratory study chose its participants using simplified principles.¹⁴

The selection of examples relies on the following thoughts. As already presented above, music, lyrics and visuals are not equal. The audio examples are remarkably similar to Japanese popular music, so it might be difficult for listeners with a different socialization background to perceive this music genre in the first place. Because of this, the study required examples without any counterpoint design structures and characteristics, which do not overwhelm the unprepared recipient.

The study was conducted online, on *Sosci-Survey*. The stimuli were 4 short fragments, which, at first, were only audio files, and later replayed with visuals as a video, respectively (cf. fig. 4). The length of each fragment was about 40 seconds long. Two songs were sung by Hatsune Miku, the remaining ones were sung by KAITO and Rin Kagamine. Two videos captured live shows, while the other two showed fragments of the original music videos. The sociodemographic questions examined the degree of music education of the participants, their musical preferences, and the curiosity towards Japanese culture.

¹⁴ This fact does not reduce the value of the results, however, the results gained from this study require further research.

Song	Vocaloid	Composer	Views on YouTube
Crescent Moon	KAITO	Kurouza-P	1,9 Mio. views ¹⁵
Deep Sea Girl	Hatsune Miku	Yuuyu-P	7 Mio. views ¹⁶
Meltdown	Rin Kagamine	Iroha	12 Mio. views ¹⁷
This is the Happiness and Peace of Mind Committee	Hatsune Miku	UtataP	5 Mio. views ¹⁸

Fig. 4: Table of used music examples

It had to be established whether the reactions of the recipients grew stronger when watching the videos (1), whether the evaluations of the stimuli were influenced by verbally mentioned interest to the different culture (2), and lastly, if the degree of music education influences the recipients' evaluations (3).

Since not all of the recipients speak Japanese, the authors focused on the perception of sound and visualization only. The semantic differential was applied to measure the affective reactions, using verbal profile, which consisted of 19 bipolar adjective pairs with six levels. The questionnaire was answered by 75 persons, of which 80% showed a high affinity towards non-European and especially Japanese culture. This is explained by the fact that the survey was shared in the German-Japanese-Societies e.V. and a Vocaloid radio station called *Vocalnexus*.

This study was statistically analyzed on SPSS using ANOVA, t-test, and cluster analysis. The results show that under the same conditions, whether audio or video recording, object-specific representations could only be demonstrated for very few items. The recipients were not able to establish rules for orientation purposes in the short time of the experiment.

¹⁵ Cf. OccasionalSubs (2013), *Kaito—Crescent Moon (上弦の月)*, Youtube-Video: URL: <https://www.youtube.com/watch?v=eYMAJfq7UB0> (accessed 22.04.2020).

¹⁶ Cf. Crownrecord (2010), *ゆうゆ feat.初音ミク「深海少女」*, Youtube-Video: URL: <https://www.youtube.com/watch?v=2CwBFr-Eoxg> (accessed 22.04.2020).

¹⁷ Cf. Googoo888 (2012), *[60fps Full風] Meltdown 炉心融解—Kagamine Rin 鏡音リン Project DIVA English Romaji Dreamy theater ドリーミーシアター*, Youtube-Video: URL: <https://www.youtube.com/watch?v=jrldXNpoaac> (accessed 22.04.2020).

¹⁸ Cf. うたたP / UtataP Official Channel (2012), *【Hatsune Miku/UtataP】こちら、幸福安心委員会です。【Original】*, Youtube-Video: URL: <https://www.youtube.com/watch?v=kzUmM-uQVOM> (accessed 22.04.2020).

Participants

Country	Gender		Age	Plays an instrument		Speaks Japanese		Knows Vocaloid	
	female	male	Average	yes	no	yes	no	yes	no
	Quantity	Quantity		Quantity	Quantity	Quantity	Quantity	Quantity	Quantity
Bulgaria	1	0	26	0	1	0	1	1	0
Germany	24	35	35	26	34	24	36	39	21
France	1	1	25	1	1	0	2	2	0
United Kingdom	0	1	24	1	0	0	1	0	1
Netherlands	2	0	21	1	1	0	2	2	0
Austria	2	1	18	1	2	1	2	2	1
Romania	0	1	46	1	0	0	1	1	0
Slovakia	0	1	13	0	1	0	1	0	1
USA	1	1	35	1	1	0	2	2	0
Not specified	1	1	22	2	0	0	2	2	0

Fig. 5. Selected data of the participants

Let us look at one part of the statistical analysis. The mean values (arithmetical mean and median) in relation of each individual example (audio and video stimuli) are almost identical. Especially the melodic songs *Crescent Moon* and *Deep Sea Girl* show almost identical reactions. This is also true for *This is the Happiness and Peace of Mind Committee*. However, there are a few adjective pairs, such as “aggressive—friendly” [aggressiv—friedlich], that do show a significant difference (cf. fig. 6, showing the arithmetic means).

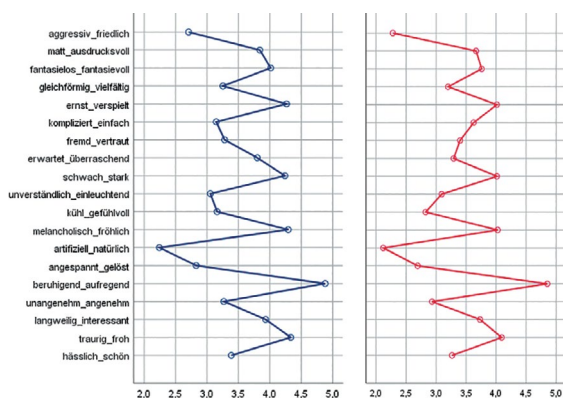


Fig. 6. Means audio (left) / video, Hatsune Miku:
This is the Happiness and Peace of Mind Committee

The decision to use different visuals proved to be advantageous. The recordings of a live show, which showed the Vocaloid on stage, led to more positive reactions of the recipients. The most significant difference was found between the audio and video fragments of the song *Meltdown* (cf. Fig. 7). T-test calculations proved significant differences in the following adjective pairs: “aggressive—friendly”, “artificial—natural” [artifizuell—natürlich] and “ugly—beautiful” [hässlich—schön]. Values for “ugly—beautiful”: $t(74) = -6.17$, $p < .001$.

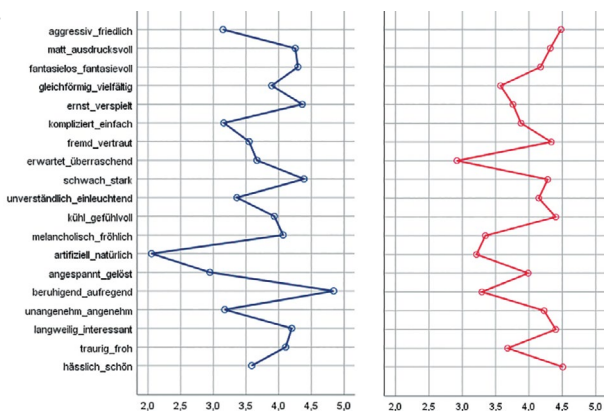


Fig. 7. Means audio (left) / video, Rin Kagamine: *Meltdown*

However, there were no further significant differences between audio and video fragments in the vast majority of calculations with other music examples. The exceptional results of *Meltdown* stand out clearly. When comparing figure 6 with 7, one can tell that the “vertical” audio (in blue) shows great similarities with *This is the Happiness and Peace of Mind Committee*. It seems obvious that the evaluation of the video example, which was taken from a live concert, had to be conspicuous. The following boxplots show the qualities of the adjective pair “aggressive—friendly” (fig. 8). As it turned out, this was an adjective pair, which showed the biggest differences under the ratings for the individual terms in all evaluations.

The result of mostly having little to no changes in the reactions between audio and video fragments can be viewed as a surprise, considering the fact that it has been observed for many years, that visuals do lead to a more positive reaction [19, 20]. A plausible explanation can be delivered when considering further results. So, let us discuss the other two research aspects first.

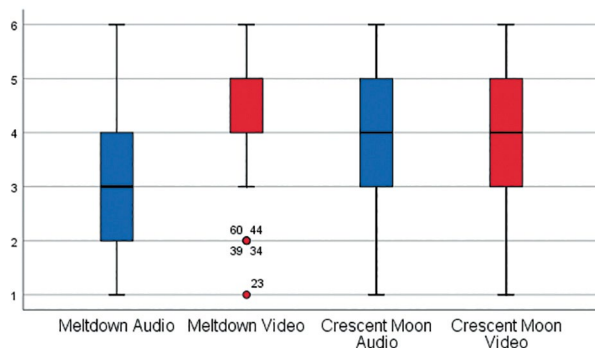


Fig. 8. Evaluations of audio and video examples in "aggressive—friendly", Title *Meltdown* and *Crescent Moon*

It was noticeable that almost all recipients had positive evaluations in the aspects of valence¹⁹ (ugly—beautiful, boring—interesting, unpleasant—pleasant). The cluster analysis of the participants judging behavior²⁰ sorted all cases in two uneven clusters, of which the cluster with negative evaluations consisted of three people. The ANOVA calculations of every adjective pair in relation to the music examples confirm the validity of the assumption that there are only few object-related representations. Only the video fragment of *Meltdown* shows significant differences in various adjective pairs. Among the other music examples were no objective-specific representations to be observed (except for "aggressive—friendly" and "ugly—beautiful").²¹ It is conceivable that the cultural background of Vocaloid music leads to more "polite" evaluations.

The data of the participants was divided into two groups, in which the first group confirmed their interest in Japanese culture. We understand this as knowing about the elements of this culture. The second group did not show any interest in Japanese culture therefore we calculated the means of both groups. It is worth noting that the first group was significantly bigger than the

¹⁹ For the meaning of valence, see P. Kuppens et al. [21]

²⁰ Ward-Method, audio and video evaluations (viewed as one) of "ugly—beautiful", "boring—interesting" or "unpleasant—pleasant".

²¹ The analysis of variance in dependency of the evaluations to the music examples of a) the audio stimuli or b) the audio and video stimuli only proves to be significant for "aggressive — friendly" and "ugly—beautiful". E.g. audio examples in "aggressive—friendly": $F(3, 296) = 50.144, p < .001$.

second one. However, we did not manage to find a connection between the participants' choices and their interest in Japan (s. Fig. 9).²²

Scale "aggressive—peaceful"

		<i>Crescent Moon</i> Audio	<i>Crescent Moon</i> Video	<i>Deep Sea</i> Audio	<i>Deep Sea</i> Video	<i>Meltdown</i> Audio	<i>Meltdown</i> Video	<i>Peace of Mind</i> Audio	<i>Peace of Mind</i> Video
		Mean Value	Mean Value	Mean Value	Mean Value	Mean Value	Mean Value	Mean Value	Mean Value
Interested in Japanese culture	Yes	4,18	3,73	4,66	4,56	3,08	4,53	2,63	2,16
	No	3,77	4,69	4,85	4,08	3,46	4,23	3,08	2,85

Fig. 9. Means of scale "aggressive—peaceful" of all 8 stimuli

Comparing the already discussed similarities of the participants choices in relation to various examples the lack of connection between the judgments and interest in Japanese culture, it is likely that the participants must have been uncertain about the genuine musical characteristics of the used materials, which could not be "compensated" by the visual stimuli, hence the lack of difference between the reactions of audio and video fragments.

The degree of music education of the participants—the last question of the present investigation—had little influence on the evaluation of the music examples, too. The means in relation to music education do differ slightly, however, the differences are not significant (fig. 10).

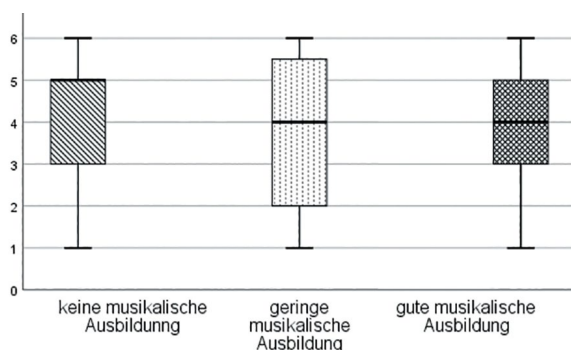


Fig. 10. Evaluations of all 4 video fragments of "ugly—beautiful" in relation to degree of music education.

Left to right: no music education / little musical education / good musical education

²² Although the comparison of the scale "ugly—beautiful" for *Crescent Moon* shows a higher resonance within the first group: $t(73) = 2.18, p < .032$.

The variance analyses are negative. The data of figure 10 shows: $F(2, 72) = 0.075, p < .928$. A lack of orientation in the presented materials could not be compensated by more content or general musical experience. This perspective also explains the missing change between the reactions of audio and video. The evaluations of the participant group, which shows a lesser degree of musical experience, does differ a lot more in comparison to the group with professional musical experience. This could be a tendency, which requires further research with a much higher number of subjects. Finally, the database contains very similar values that come from an insufficient heterogenous sample.

Results and outlook

The phenomenon of Vocaloid brings new forms of interaction between the author and its audience, and between the author and the musical material in Japanese, South Korean and, more recently, American popular music. The primary goal of the present essay was to trace the existence of this phenomenon. Vocaloid is a new occurrence of popular (youth) culture with its own form of material design and sometimes with socio-critical potential. The occurrence is not very known in Europe, yet.²³ This explorative study, which was conducted at the Martin-Luther-University, shows that recipients can respond favourably to this new phenomenon. With all the seemingly obvious results, the study shows an important conclusion: the sympathy of the listener seems to be created on a sociological level—the development of a new musical instrument with individual features—rather than, as expected, the creation of the product. The next step could be to conduct another study which defines the object of its assessment.

The current homepage of Yamaha Corporation advertises its new product *Vocaloid™5*, which is available as standard and premium versions, with 4 or 8 voices banks and 100 preinstalled singing styles at the price range of 250\$ to 400\$. The instrument is affordable and well equipped. It is worth continuing further research in its existence, its effects on the cultural industry and its perception. The authors contemplate a new research design using recordings of live concerts, rather than the original music videos, which follow the

²³ Vocaloid also appears in internationally shown anime series. URL: https://mikupa.ru/hatsune_miku_v_anime (accessed 22.04.2020).

design of typical Japanese anime movies. It would be interesting to research the perception towards compositions, which are not aligned with Japanese popular music as well. Furthermore, one would need music examples, which are similar to Vocaloid music, but performed by real singers. Using English lyrics could also help to examine the complete interaction between music, lyrics and visuals. Research in terms of intercultural comparison would also be highly informative.

What matters is the procurement of information on Vocaloid music in a traditional genre. Breaking traditions, the singing synthesiser fuses together with the academic and traditional art. By now, there are symphonic concerts and even theatrical performances, such as kabuki or bunraku featuring Vocaloid. Since 2016 there has been a yearly kabuki performance with Hatsune Miku,²⁴ which uses the holographic technology on the actors, too. This new collaboration breaks the rules of traditional Japanese theatre, adding communication between actors and audience. Therefore, the contribution of virtual reality leads to a bigger number of recipients and a much younger audience.

* * *

Программное обеспечение для синтеза речи *Vocaloid* (вокалоид), разработанное в Барселонском университете Помпеу Фабра совместно с корпорацией *Yamaha*, было представлено публике в 2003 году¹, и уже через год на рынке появилась первая версия программы. Примечательно, что она была на английском языке, хотя синтезировать японскую речь намного проще². В тот момент интерес к столь уникальному продукту был невелик, и показатели продаж держались на относительно низком уровне³. Тогда *Yamaha* решила привлечь третьих лиц, которые в целях совершенствования маркетинговой стратегии придумали аниме-аватары. Эта стратегия привела к оглушительному успеху. Мику Хацунэ, персонаж, раз-

²⁴ The kabuki performance was hosted in the Minamiza theatre, Kyoto. URL: <https://chokabuki.jp/2016/en/> (accessed 22.04.2020).

¹ По сравнению с синтезом звука синтез речи более сложен.

² Японский язык основан на слоговом принципе, поэтому на нем записывать голосовую библиотеку быстрее. Все синтезаторы речи оптимизируются для имитации конкретного языка. Например, при вводе текста английской песни в японский голосовой банк можно услышать «акцент». Первым синтезатором с японским голосовым банком был *MEIKO*.

³ Стоит отметить, что один из первых коммерчески доступных альбомов *Vocaloid* был записан в России. Он содержит треки на русском и английском языках: «Место под солнцем»: Алексей Устинов, Алексей Косолапов feat. *LëLik [LEON]*», лейбл — *Virartech*, Новосибирск, 2006.

работанный компанией *Crypton Future Media* и выпущенный в 2007 году, стал настолько популярным, что «ее» называют первым виртуальным кумиром в истории популярной музыки. В отличие от предыдущих моделей вокалоида, уже за первое полугодие Мику была продана 30 000 раз, а это считается отличным показателем для любого синтезатора. Наряду с японским и английским голосовыми банками вокалоиды теперь «научились» петь на корейском, испанском и китайском языках.

Выступления вокалоидов делятся на два типа. Один из них — «традиционный» живой концерт, когда голограмма, проецируемая на экран, танцует и поет для своих фанатов. Чаще всего слушатели, держа в руках светящиеся палочки, танцуют под музыку вместе с виртуальным артистом. Второй тип — выступления на экране компьютера, в социальных сетях, например, на YouTube. Это оригинальные музыкальные клипы, анимированные либо фанатами, либо профессиональными иллюстраторами. Архетип героя направляет и, в то же время, очевидно ограничивает фантазию композитора. У вокалоида фактически нет песен без визуальной составляющей, поэтому создание композиций в этой программе можно описать как синтез звукового и визуального аспектов.

Несмотря на то, что речь идет о феномене, эффективно использующем всевозможные медиа, до настоящего времени он мало изучен. Большая часть доступных исследований проводилась в Японии или Южной Корее и фокусировалась, в основном, на особенностях технологии или медийных аспектах программы (Х. Кенмоши, Т. Шибя, П.В. Гэлбрэйт, Д.Г. Карлин и некоторые другие) [1, 2, 3]. Американские и европейские публикации рассматривают вокалоид в контексте развития современной культуры, а также этических вопросов, возникающих в связи с этой новой технологией (П. Дэвисон, Н. Эйдсхейм) [4, 5 и 6, с. 115–150]. Аспекты, связанные с синтезом речи, напротив, уже давно пользуются популярностью как объект исследования (например, П. Шапиро) [7]. Их значение, однако, выходит за рамки интересов музыковедения.

Цель данной статьи — дать общую характеристику вокалоида как нового музыкального явления, в том числе предложить классификацию песенных композиций для виртуального синтезатора. А также представить результаты экспериментального исследования по целостному восприятию произведений для вокалоида с учетом того, что реакции реципиентов изучались в кругах, не вовлеченных (по крайней мере пока) в новую культуру.

Новый инструмент или новое направление в музыке?

Влияние вокалоида вообще и, в частности, его наиболее популярного персонажа Мику Хацунэ на музыкальную индустрию Японии огромно. Японское общество в целом весьма предрасположено к явлениям искусственного интеллекта, а также демонстрирует особое пристрастие к разнообразным символическим изображениям. В качестве примера можно привести талисманы для Олимпийских игр 1921 года в Токио: Мирайтова и Сомэйти получили не только аватаров в стиле аниме, но и распространились по всей стране в качестве маскотов и роботов [8]. Между тем, Южная Корея также работает над созданием виртуальных артистов, используя аналогичные технологии для коммерциализации всемирно известных звезд кей-поп⁴ и даже видеоигр, например, *League of Legends* [9, 10].

Становление вокалоида представляет интерес для социологии музыки как в смысле появления нового музыкального инструмента на текущем этапе художественного развития, так и в плане ответа музыкального бизнеса на потребности рынка. С последним все достаточно просто. Рынку необходимы новые раздражители, которые будут с легкостью приняты молодежью. А молодежь живет сегодня в своем, цифровом измерении [11, с. 552]⁵. В этом цифровом пространстве можно найти забавный контент, например, тесты на определение типа личности с целью выяснить, кто из вокалоидов похож на тебя⁶. Но также там можно узнать о будущих музыкальных событиях концертного формата, в том числе далеко за пределами Японии — в Торонто, Лос-Анжелесе, Далласе, иногда в Париже, Лондоне, Кельне и даже один раз в Москве⁷, где достать билет было практически невозможно. А это уже серьезное свидетельство «экспорта» новой культуры. Начинает рассматриваться вопрос и о педагогическом использовании феномена [14].

⁴ Кей-поп (K-pop, Korean pop) — жанр поп-музыки, возникший в Южной Корее и превратившийся в масштабную музыкальную субкультуру с миллионами поклонников во всём мире.

⁵ С точки зрения некоторых социологов и культурологов цифровая система координат распространилась уже на все общество [12, 13].

⁶ URL: <https://www.testedich.de/quiz57/quiz/1541174128/Welcher-Vocaliod-bist-du> (дата обращения: 13.04.2020).

⁷ 25.03.2012. В столице имел место даже официальный показ видеозаписи Оперы «Аой», в которой вокалоидная музыка совмещена с театром кукол бунраку.

На сцене или на экране, исполненная вокалоидом музыка становится новой разновидностью массового искусства. При этом репрезентация играет, очевидно, более важную роль, чем в традиционном искусстве [15, с. 188]. Похоже, виртуальном артисту необходимо большее число индивидуальных качеств, чем живому певцу. Пути коммуникации раздваиваются. В первую очередь, мы испытываем воздействие музыки, параллельно на нас воздействует образ персонажа. Это примерно то, о чем говорилось уже много лет тому назад, в теории коммуникации Пауля Вацлавика, с ее положением о невозможности не общаться [16, с. 51]⁸.

И все-таки, вокалоид — это прежде всего программное обеспечение с конкретными техническими возможностями. Использоваться программа может по-разному: как музыкантами-профессионалами, для создания композиций с наложением специальных эффектов на вокал, так и фанатами-любителями, сочиняющими свои первые песни с единственной целью присоединиться к вселенной, возникшей вокруг Мику Хацунэ. В области искусственного интеллекта данное ПО уже тоже нашло себе применение. С помощью этой программы робот-гуманоид *ASIMO*, разработанный фирмой *Honda*, напевает популярные мелодии.



Рис. 1. Мику⁹



Рис. 2. Кайто¹⁰

⁸ В области информатики теория Вацлавика играет и по сей день очень важную роль.

⁹ *Hatsune Miku*, © *Crypton Future Media, Inc.* 2007, лицензия CC BY-NC. URL: <http://creativecommons.org/licenses/by-nc/3.0/>. Источник изображения см.: URL: https://ec.crypton.co.jp/pages/prod/virtualsinger/cv01_us (дата обращения: 21.03.2021).

¹⁰ *KAITO*, © *Crypton Future Media, Inc.* 2006, лицензия CC BY-NC. URL: <http://creativecommons.org/licenses/by-nc/3.0/>. Источник изображения см.: URL: <https://vocaloid.fandom.com/de/wiki/KAITO> (дата обращения: 21.03.2021).

Маркетинговая стратегия удивительно интересна, поскольку голоса Мику Хацунэ и всех других вокалоидов, которые продаются, общедоступны. Новое соотношение ролей в музыкальном творчестве (композитор является артистом, а виртуальный певец — нет) оказалось созвучным современности. Авторские права на песни, сочиненные с помощью данного программного обеспечения, принадлежат исключительно композитору: в этом смысле вокалоид аналогичен любому другому синтезатору. Что касается образа Мику (рис. 1), он принадлежит ее создателям, компании *Crypton Future Media*, которая в 2012 году опубликовала изображение на условиях свободной лицензии (*CC-BY-NC*).

Принцип пользования синтезатором достаточно прост. Пользователь загружает в программу ноты и текст песни и получает готовый звуковой продукт. Для ввода нот можно также использовать MIDI-клавиатуру. Готовый файл выводится в формате .wav или VST. Таким образом, своего собственного певца-исполнителя получает неограниченное число заинтересованных лиц¹¹. Формулировка новейшего времени «индивидуальная художественная индустрия» (Е. Дуков) подходит к новой модели идеально [17, с. 176]. Но сбывается и давняя мечта композитора — найти исполнителя, обладающего не только талантом, но и покладистым характером. Если вспомнить о многочисленных конфликтах в среде музыкантов, например, между членами группы *The Beatles* или *The Strokes*, недооценивать преимущества от использования искусственного голосового аппарата невозможно. Стоит порадоваться тому, что на сегодняшний день существует более 100 продуктов от *Vocaloid*, а также множество созданных фанатами голосовых библиотек, которые работают в программе под названием *UTAU*. Пользователи *UTAU* могут сами решать, как будет выглядеть певец и как изменить его голос.

С социологической точки зрения, новое явление кажется наделенным многообещающими социальными преимуществами. Как инструмент, вокалоид представляет собой нечто иное, качественно большее, чем синтезатор — он обладает чертами личности. Этой личности соответствует свой собственный тембр голоса (или даже несколько тембров), свой тип характера, включая безупречную биографию, свой репертуар. При том, что каждому вокалоиду придаются индивидуальные черты, с музыкальной точки зрения синтезаторы голоса имеют на сегодняшний день один несомненный общий

¹¹ По свидетельству корпорации *Crypton* только за пять лет с момента выпуска Мику с ее помощью было сочинено более 100 000 произведений, записано 170 000 видео и создано около миллиона иллюстраций. URL: https://ec.crypton.co.jp/pages/prod/vocaloid/cv01_us (дата обращения: 21.03.2021).

признак. В разработках доминирует представление о красоте человеческого голоса, распространенное в многовековой культуре Японии. А значит, голоса высокого, вне зависимости от того, женский он или мужской.

Если рассматривать музыку для вокалоида как разновидность синергетического жанра, можно, конечно, предположить, что в ней отсутствует стилистическая однородность. Как в плане текста, так и в плане музыки, и даже в отношении зрительного решения разные произведения отличаются друг от друга иногда очень явно. Несмотря на то, что большая часть музыки для вокалоида несет в себе веселую, жизнерадостную ноту, мы предпочтем сосредоточиться на сочинениях более глубокого плана, тех, что призывают задуматься. К примеру, песня *Hello, how are you?* посвящена застенчивости и неуверенному поведению явно выраженного интроверта. В то время как нотный текст, насыщенный паузами и сменой направления движения (прежде всего во вступлении, за пределами нотного примера), формально следует за словом, общее впечатление от голоса, сопровождаемого только фортепиано, скорее, беззаботное, живое, даже веселое. Это впечатление находится в полном противоречии к тексту, в котором повествуется о непреодолимых проблемах протагониста. Экранное решение рисованной мультипликации вносит толику ясности — мы видим плачущую девушку-подростка, которой все не под силу.



Нотный пример 1. Ноехоэ-Р, *Hello, how are you?*, куплет, транскрибированная партия голоса

«Привет!» — я открыла окно и тихо прошептала в него: «Как дела?» —
Здесь никого нет, я одна в комнате. Утро. Пришло утро, утро с проливным дождем.

Тик-так, будто кто-то заводит механизм. [...]

Почему ты прячешься? Ты боишься, что над тобой будут смеяться?
Разве ты не хочешь с кем-нибудь познакомиться? Ты действительно этого хочешь? —

В море двусмысленности мне слишком больно дышать.
Я только хотела услышать человеческий голос. Я очень слаба.

Текст: Ноехоэ-Р, *Hello, how are you?*, куплет (до [...]), рефрен,
перевод с японского М.-Л. Арендт, перевод на русский Н. Новак

По-другому сведены воедино лирика, звуковое оформление и действие в примере *Lost One's Weeping*. В этом примере наблюдается большее единство средств между музыкой и текстом. Постоянное возвращение мелодии к звуку *fis* символизирует невозможность преодолеть некие рамки, что связано с основной идеей песни — критикой жесткой системы японского образования. Перфекционизм как насаждаемый жизненный принцип осмысливается методом противопоставления (идеального) ученика и его реальных эмоций. Последние принимают аллегорическую форму двойника школьника. Конфликт переносится в область сценического решения.



Нотный пример 2. Neru, *Lost One's Weeping*, рефрен, транскрибированная партия голоса

Можешь ли ты прочесть кандзи на доске?
 Можешь ли увидеть внутренний мир этого ребенка?
 Скажи мне, кто окрасил его сердце в черный цвет?
 Кто это сделал, кто это сделал?
 Можешь ли ты решить это уравнение на счетах?
 Можешь ли ты ослабить веревку, сдавливающую ребенка шею?
 А что, если мы останемся такими навсегда?

Текст: Neru, *Lost One's Weeping*, рефрен,
 перевод с японского М.-Л. Арендт, перевод на русский Н. Новак

Приведенные примеры наводят на мысль, что исходить от стандартных соотношений между словом, музыкой и дизайном музыкального клипа не приходится. Художественные решения допускают большую степень индивидуальности. Эта модель не нова, она знакома, к примеру, по музыке кино. Пользуясь определениями из области кинематографа, можно в случае *Hello, how are you?* — в зависимости от интерпретационного подхода — говорить о поляризации или о контрапунктном взаимоотношении между музыкой и текстом, а в случае *Lost One's Weeping* о конфликте внутри сценического решения.

Можно ли рассматривать наметившиеся различия в творческих решениях как основание не считать музыкальную продукцию с участием

вокалоида единым музыкальным направлением? Похоже, что нет. В истории музыки известно много подобных соотношений — от гибридной культуры джаз-рока или музыки видеоигр и до готики (*Gothic*). Речь идет, очевидно «только» о том, чтобы определить тот музыкальный или же внемузыкальный признак, который можно будет принять за критерий направления. Тогда почему бы не дать определение социологического порядка — вокально-инструментальная музыка «от пользователя», без посредничества живого исполнителя? Интересно было бы в дальнейшем попытаться ответить на вопрос, каков он, этот пользователь, в чем секрет коллективного разума, отличающегося в данном случае завидной плодовитостью.

С точки зрения авторов этой статьи одним из аргументов в пользу выделения нового направления является еще и тот факт, что музыка для вокалоида поддается классификации.

Классификация

Так как до настоящего времени не предпринималось попыток систематизировать произведения для голосового синтезатора, Мелисса-Лили Арендт предлагает первую модель классификации¹². В ее основу положены не музыкальные критерии, а содержание и стилистические особенности текстов, определяющих общий стиль композиции. В предлагаемой модели пять типов: **развлекательные, эпические, лирические, социально-критические и смешанные** (рис. 3). Первые два типа отличаются друг от друга принципиально. Образцы, о которых говорилось выше, относятся к группам *лирических* и *социально-критических*.

К типу *развлекательных* относятся песни, соответствующие жанру электронной танцевальной музыки (*Electronic Dance Music*, или *EDM*). Их максимальная продолжительность не превышает пяти минут, а содержанием их является, кроме разнообразной детской и трогательной тематики, приглашение к движению или непосредственно к танцу. Речь в данном случае идет, по-видимому, о центральной разновидности этой музыки, т. к. благодаря распространению через социальные сети она получает особый резонанс. Последнее служит многим композиторам путевкой к дальнейшей карьере в музыкальном бизнесе. Очень многие номера

¹² В процессе осмысления классификации было проанализировано более 70 произведений.

оказываются в японских чартах и становятся излюбленным материалом для караоке [для ср.: 18]. Принцип дизайна произведений этого типа единообразен: демонстративно-искусственный паттерн-ритм сопровождает легко запоминающуюся мелодию, текст чаще всего забавен, стилизован под детскую песню или ею же и является. Визуальное решение достаточно однозначно.

Развлекательные	Эпические	Лирические	Социально-критические	Смешанные
Стиль песен для вечеринок (EDM), ассоциации с игрой и весельем	Построение по принципу саг, легенд или сказок	Изображение эмоций (например, любовная тоска)	Выражение социально-критической точки зрения	Сочетание двух типов классификации
Пример: One-Cup P, <i>Koneko no Paupapa</i>	Пример: Mothy, <i>Daughter of Evil</i>	Пример: Hоеhое-P, <i>Hello, how are you?</i>	Пример: Neru, <i>Lost One's Weeping</i>	Пример: Utata-P, <i>This is the Happiness and Peace of Mind Committee</i>

Рис. 3. Классификация произведений, исполняемых при помощи вокалоида

Эпический тип оформлен принципиально сложнее, песни выстраиваются по принципу саги или же легенды, сказки, им свойственна преемственность. Значение текста достаточно велико, при этом используется множество типично сказочных формул. Часто один сюжет разбит на несколько песен, исполняемых разными вокалоидами, что дает возможность представить разные точки зрения¹³. Так как для полноценного рассказа нужно определенное время, песни этого типа намного длиннее группы развлекательных, они длятся по 8 и более минут. Оформление треков этой группы принципиально иное — зрелищное и сюжетное. Так, при озвучивании содержания, связанного с Японией, используются по преимуществу инструменты национальной традиции, если же обыгрывается сюжет западный, инструменты, а иногда и стиль вокала, меняются. Для визуального оформления характерно стремление к реалистичности — например, используются исторические или национальные костюмы.

¹³ Примером цикличного решения служит серия композитора Mothy *Nanatsu no Taizai* (по-русски: «Семь смертных грехов»).

Группа *лирических* песен демонстрирует всевозможные ситуации, в которых доминирует сфера эмоций, например, грусть, переживание потерь, ярость. Треки, представленные в этой группе, крайне разнообразны. Одной из ведущих составляющих является глубоко продуманный (прочувствованный) текст, оставляющий, как правило, место для индивидуальной интерпретации. Взаимосвязь между текстом, музыкой и изображением решается неоднозначно, эмоции не лежат на поверхности. Слушатель приглашается заглянуть во внутренний мир протагониста, погрузиться в эмоциональное переживание.

Характеристики песен *социально-критического типа* похожи на те, что представляют особенности лирических, за исключением того, что они сосредоточены на социальных моментах и заостряют внимание на конфликтах, в результате чего реципиент подводится к критической оценке событий. Обсуждение последних переходит из сферы искусства в сферу коммуникации, дискуссии ведутся на интернет форумах и в социальных сетях. Воплощение конфликтов часто переносится в область экрана, они представлены зрелищно, в то время как музыкальная составляющая сохраняет нейтральность. При создании видео нередко используют сильные драматургические приемы¹⁴.

Последняя группа включает в себя песни, основанные на смешении признаков тех типов, о которых говорилось выше, к примеру, *развлекательных* и *лирических* или же *лирических* и *социально-критических*.

Исследование

На фоне необычайного распространения вокалоида, прежде всего в Японии и Южной Корее, возникает вопрос, в той ли же мере открыты на встречу новому явлению жители других стран. Возможно, это только вопрос времени, в какой момент восторженное отношение распространится по всему миру. Задачей исследования было зафиксировать восприятие отобранных фрагментов песен вокалоидов. Дизайн исследования ($N = 75$) был поисковый, в первую очередь по той причине, что выборка участников велась по упрощенной схеме¹⁵.

¹⁴ Например, в треке *Hana Ichi Monme* от Team Kamiuta: URL: <https://www.youtube.com/watch?v=b2iel0gjmfs> (дата обращения: 21.03.2021).

¹⁵ Этот факт не умаляет значение выводов исследования, поскольку речь шла о первичной постановке вопроса о восприятии музыки, созданной с применением вокалоида, но делает необходимой дальнейшую проверку полученных результатов.

Подбор материала для исследования предварили следующие сообщения. Как было показано выше, музыка и текст или же музыка, текст и видео соотносятся каждый раз по-разному. Так как аудио составляющая в данном случае ориентирована на японскую популярную музыку, ее восприятие слушателем иной социализации затруднено в принципе. Следовательно, было необходимо найти примеры, которые не ошеломили бы неподготовленного слушателя сложным контрапунктическим сочетанием аудио и видеоряда.

Проводилось исследование онлайн. В качестве раздражителей использовали 4 коротких фрагмента, сначала только как аудио записи, а после их оценки, с добавлением видео (см. рис. 4). Длительность каждого фрагмента составляла около 40 секунд. Кроме Хацунэ Мику использовались треки двух других персонажей — Кайто и Рин Кагамине. В двух видео показывались записи с концертов, в двух других отрывки из музыкальных роликов. Социо-демографические вопросы включали, в первую очередь те, что уточняли степень музыкальной подготовки участников, их музыкальные предпочтения и наличие интереса к японской культуре.

Проверке подлежало: усиливаются ли реакции участников за счет зрительных раздражителей (1), влияет ли на оценку стимулов вербально зафиксированный интерес к чужой культуре (2) и, наконец, играет ли роль музыкальное образование испытуемых (3).

Трек	Вокалоид	Композитор	Данные по YouTube
<i>Crescent Moon</i>	КАИТО	Kurousa-P	1,9 Mio. views ¹⁶
<i>Deep Sea Girl</i>	Hatsune Miku	Yuuyu-P	7 Mio. views ¹⁷
<i>Meltdown</i>	Rin Kagamine	Iroha	12 Mio. views ¹⁸
<i>This is the Happiness and Peace of Mind Committee</i>	Hatsune Miku	UtataP	5 Mio. views ¹⁹

Рис. 4. Таблица использованных в эксперименте примеров

¹⁶ Ср.: OccasionalSubs (2013), *Kaito—Crescent Moon (上弦の月)*, Youtube-Video. URL: <https://www.youtube.com/watch?v=eYMAjfq7UB0> (дата обращения: 22.04.2020).

¹⁷ Ср.: Crownrecord (2010), *ゆうゆう feat. 初音ミク「深海少女*, Youtube-Video. URL: <https://www.youtube.com/watch?v=2CwBFr-Eoxg> (дата обращения: 22.04.2020).

¹⁸ Ср.: Googoo888 (2012), *[60fps Full風] Meltdown 炉心融解—Kagamine Rin 鏡音リン Project DIVA English Romaji Dreamy theater ドリーミーシアター*, in: URL: <https://www.youtube.com/watch?v=jrldXNpоаас> (дата обращения: 22.04.2020).

¹⁹ Ср.: うたたP / UtataP Official Channel (2012), *[Hatsune Miku/UtataP]こちら、幸福安心委*

Исходить из того, что все участники знают японский язык, не приходилось, поэтому рассматривалось восприятие только звукового или же звукового и визуального оформления фрагментов. Измерения проводились с помощью семантического дифференциала — вербального профиля, состоявшего из 19 пар противоположных по смыслу определений, среди них как оценочные понятия, так и понятия-категории. Шкала состояла из 6 делений. Полностью ответили на вопросы исследования 75 человек, причем группа участников оказалась состоящей из лиц с большим интересом к неевропейской, и в частности, к японской культуре (более 80 %). Последний факт объясняется тем, что приглашение принять участие в исследовании поддержали (и распространили среди своих членов или слушателей) общества немецко-японской культуры и специализирующаяся на вокалоидах радиостанция *VocalNexus*.

Участники исследования

Страна	Пол		Возраст	Plays an instrument		Speaks Japanese		Knows Vocaloid	
	ж	м		да	нет	да	нет	да	нет
	Кол-во	Кол-во	Средний	Кол-во	Кол-во	Кол-во	Кол-во	Кол-во	Кол-во
Болгария	1	0	26	0	1	0	1	1	0
Германия	24	35	35	26	34	24	36	39	21
Франция	1	1	25	1	1	0	2	2	0
Англия	0	1	24	1	0	0	1	0	1
Нидерланды	2	0	21	1	1	0	2	2	0
Австрия	2	1	18	1	2	1	2	2	1
Румыния	0	1	46	1	0	0	1	1	0
Словакия	0	1	13	0	1	0	1	0	1
США	1	1	35	1	1	0	2	2	0
НЕ указана	1	1	22	2	0	0	2	2	0

Рис. 5. Избранные данные участников эксперимента

Из статистических операций просчитывались в первую очередь анализы отклонений, t-тесты и кластерный анализ. Полученные данные показывают, что при одних и тех же условиях, вне зависимости от того, демонстрировалась ли запись аудио или видео, только на очень ограниченное количество вербальных категорий пришлось оценки, которым

員会です。【Original】, in: URL: <https://www.youtube.com/watch?v=kzUmM-uQVOM> (дата обращения: 22.04.2020).

можно приписать значение данных от объекта. По-видимому, в ходе эксперимента испытуемые не успевали выработать для себя принципа ориентации в предложенном материале.

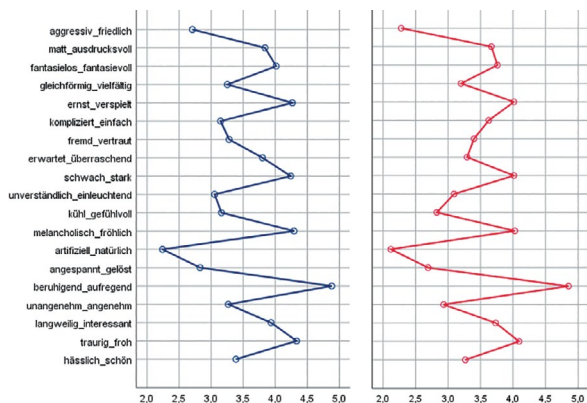


Рис. 6. Средние значения: аудио (слева) и видео, *Hatsune Miku: This is the Happiness and Peace of Mind Committee*

Рассмотрим часть статистических расчетов по порядку. Средние значения показаний (как среднее арифметическое, так и медиан), полученные по отдельно взятым стимулам, почти идентичны у многих категорий как в аудио- так и в видеозаписи. В первую очередь у мелодичных напевов *Crescent Moon* и *Deep Sea Girl* средние показания восприятия практически равны друг другу, но и у *This is the Happiness* сколько-либо значительные отклонения наблюдаются лишь в связи с немногими парами определений, к примеру в определениях «агрессивный — мирный» [*aggressiv—friedlich*], рис. 6, среднее арифметическое значений).

Решение в пользу использования видеоматериала различного происхождения оказалось удачным. В обоих случаях, где были задействованы «живые» выступления вокалистов, оценка фрагмента на видео оказалась выше той, что давалась тому же примеру в звукозаписи. Четко выраженными различия между примерами аудио и видео были тем не менее только в одном случае — трек *Meltdown* (ср. рис. 7). Т-тест подтвердил статистически значимые различия в определениях «агрессивный — мирный», «искусственный — естественный» [*artifiziel—natürlich*] и др., например, «отвратительный — прекрасный» [*hässlich—schön*]: $t(74) = -6.17, p < .001$.

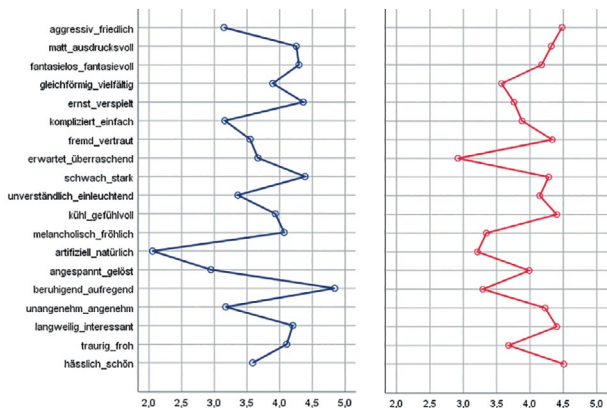


Рис. 7. Средние значения: аудио (слева) и видео, Rin Kagamine: *Meltdown*

При большинстве расчетов оценок других музыкальных примеров разницу средних величин подтвердить статистически не удалось. Необычность результатов по стимулу *Meltdown* видится по уже имеющимся графикам. Даже при поверхностном сравнении изображений на рис. 6 и 7 становится ясно, что вертикаль «аудио» у *This is the Happiness* и у *Meltdown* (синий цвет) выстраивается аналогично. Очевидно, что оценки видеозаписи концерта с *Meltdown* выпадают из общего русла. При помощи диаграмм Voxplot на рис. 8 можно дополнительно отследить общую структуру результатов, касающихся определений «агрессивный — мирный». Как выяснилось, этой паре определений соответствует наибольший разбой среди показаний по отдельным понятиям всего исследования.

Тот результат, что оценка фрагментов не изменялась регулярно при переходе от аудио к видео, можно рассматривать как неожиданность эксперимента, т. к. уже в течение многих лет ученые наблюдают заметный прирост положительных оценок, как только к звучанию присоединяется оптический раздражитель [19, 20]. Разумное объяснение станет возможным при дальнейшем рассмотрении результатов эксперимента. Обратимся же прежде к двум другим аспектам исследования.

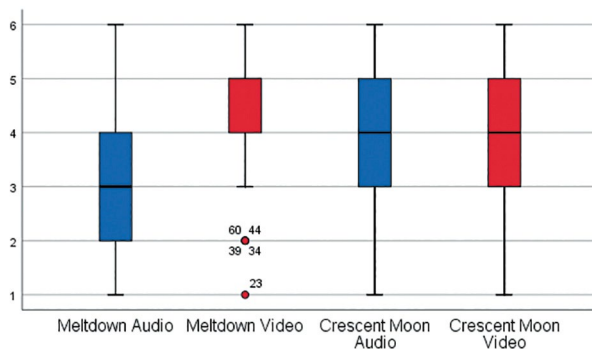


Рис. 8. Оценки примеров аудио- и видео, полученные по шкале «агрессивный — мирный», треки *Meltdown* и *Crescent Moon*

В процессе работы с числами обратил на себя внимание тот факт, что при наличии сугубо валентных понятий²⁰ («отвратительный — прекрасный», «скучный — интересный», «неприятный — приятный» и пр.) почти все участники эксперимента примеры оценили положительно. Проведенный кластерный анализ оценочного поведения испытуемых²¹ разбивал группу участников неизменно на 2 неравных класса, причем в группе с «плохими» показателями оставалось не более 3 человек.

С использованием *ANOVA* в отношении всех пар понятий по очереди удалось подтвердить предположение, что оценок от объекта среди результатов окажется мало. Лишь уже обративший на себя внимание видефрагмент *Meltdown* отличается в отношении многих пар от других стимулов. Среди остальных раздражителей оценочное поведение в зависимости от объекта оценки (за исключением пар «агрессивный — мирный» и «отвратительный — прекрасный») не наблюдалось²². Возможно, что культурный фон музыки для вокалоида вызывает оценочное поведение с применением «бонуса вежливости». Используя данные по интересу участников к японской культуре, который мы понимаем как знакомство с

²⁰ О значении валентности см.: П. Куппенс [21].

²¹ Метод Варда, оценки примеров аудио и видео (взятые вместе) с помощью шкал «*hässlich — schön*», «*langweilig — interessant*» и «*unangenehm — angenehm*» («отвратительный — прекрасный», «скучный — интересный» и «неприятный — приятный»).

²² Анализы отклонений оценки в зависимости от данных по стимулу в пределах а) только звуковых стимулов или б) звуковых и оптических стимулов оказались статистически значимыми только в отношении пар «агрессивный — мирный» и «отвратительный — прекрасный». Примеры аудио на основе «агрессивный — мирный»: $F(3, 296) = 50.144$, $p < .001$.

элементами этой культуры, были просчитаны результаты по группам с интересом и без него. Надо, конечно, учитывать, что большая часть испытуемых оказалась в группе с наличием интереса. Определенной зависимости оценок от интереса выявить, однако, не удалось (см. рис. 9)²³.

Шкала «агрессивный — мирный»

		<i>Crescent Moon</i> аудио	<i>Crescent Moon</i> видео	<i>Deep Sea</i> аудио	<i>Deep Sea</i> видео	<i>Meltdown</i> аудио	<i>Meltdown</i> видео	<i>Peace of Mind</i> аудио	<i>Peace of Mind</i> видео
		Средний показател	Средний показател	Средний показател	Средний показател	Средний показател	Средний показател	Средний показател	Средний показател
Интерес к японской культуре	Yes	4,18	3,73	4,66	4,56	3,08	4,53	2,63	2,16
	No	3,77	4,69	4,85	4,08	3,46	4,23	3,08	2,85

Рис. 9. Средние арифметические оценок по шкале «агрессивный — мирный» по всем 8 примерам

Сопоставляя «похожесть» оценок различных примеров, о которой говорилось выше, и отсутствие взаимосвязи между результатами и наличием интереса к культуре Японии, можно, пожалуй, исходить из вызванного у участников эксперимента чувства неуверенности по отношению к сугубо музыкальным признакам предложенного материала, «компенсировать» которое визуальные раздражители были неспособны. Таково, на наш взгляд, первое объяснение того, что аудио-видео эффект проявился в очень малой степени.

Степень музыкальной подготовки участников — последний вопрос исследования — имела тоже очень незначительное влияние на оценку примеров. Хотя средние показатели и различаются в зависимости от того, имели ли испытуемые музыкальное образование или нет, различия эти более или менее символические.

²³ Хотя при сравнении по шкале «отвратительный — прекрасный» оценки *Crescent Moon* без видео были, например, выше у респондентов, отмечающих у себя интерес к японской культуре: $t(73) = 2.18, p < .032$.

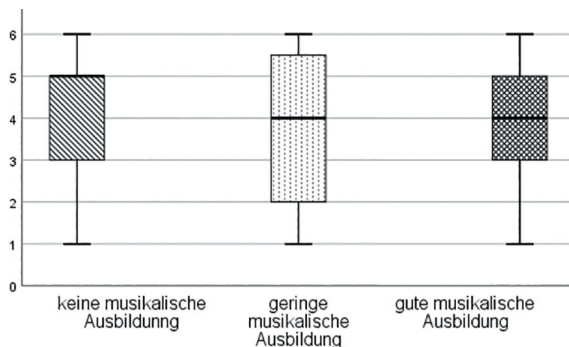


Рис. 10. Оценки всех 4 примеров видео по шкале «отвратительный — прекрасный» в зависимости от музыкального образования испытуемых. Слева направо: без музыкального образования / начальное музыкальное образование / хорошее музыкальное образование

Анализ отклонений остается отрицательным. Показатели, соответствующие изображению на рис. 10, равны, например: $F(2, 72) = 0.075$, $p < .928$. Отсутствие ориентации в предоставленном материале не восполняется, таким образом, не только добавлением сюжетного образа, но и наличием профессионального музыкального образования. Таким образом, и с этой точки зрения отсутствие регулярного эффекта сдвига оценки в связи с добавлением зрелищного элемента объяснимо. При этом эффект среди участников эксперимента с отсутствием музыкальной подготовки наблюдался все же чуть чаще, чем среди людей с профессиональным музыкальным образованием. Это уже тенденция, хотя для ее проверки необходимо повысить число испытуемых.

Результаты и перспективы

Феномен *Vocaloid* привносит в японскую, южнокорейскую, а с недавних пор и в американскую популярную музыку новые виды взаимодействия между автором и публикой, а также между автором и музыкальным материалом. Первичной задачей данной статьи было обрисовать бытование феномена. Вокалоид — это новое явление популярной (молодежной) культуры, обладающее своими собственными схемами оформления материала и по временам несущее в себе социально-критический

потенциал. В Европе это явление еще не распространилось²⁴. В поисковом эксперименте, проведенном на базе университета им. Мартина Лютера, было видно, что новое явление пробуждает у реципиента по крайней мере доброжелательное отношение. При всей кажущейся очевидности результата напрашивается немаловажный вывод: симпатия слушателей вызвана, скорее, явлением социологического порядка, таким как появление необычного, с индивидуальными параметрами музыкального инструмента, чем самим полученным в результате использования этого инструмента продуктом. На следующем этапе необходимо исследование, которое позволит уточнить объект оценки.

Актуальная интернет-страница корпорации *Yamaha* рекламирует продукт *Vocaloid™5* в базовом и улучшенном вариантах, с 4 или же с 8 библиотеками голосовых данных и со 100 настроенными стилями вокала при стоимости в пределах от 250\$ до 400\$. Таким образом, инструмент вполне доступен и хорошо оснащен. Имеет смысл планировать дальнейшие исследования его бытования, воздействия на музыкальный бизнес и восприятие. Авторы статьи могут себе представить дизайн исследования с использованием записей концертов, поскольку последние производят более сильное впечатление на зрителя-слушателя, чем ролики клипов. Было бы интересно проследить восприятие тех немногих работ с вокалоидом, которые не отмечены характерными признаками японской музыки. Кроме того, необходимы примеры с похожими данными, исполняемые не программой. В случае использования примеров с английским текстом предоставится возможность отследить восприятие всего комплекса художественных мер, находящихся, как уже было показано, в сложных взаимоотношениях. Любопытно было бы также провести сравнительное исследование в различных культурах.

Очень важна, по-видимому, информация о формах бытования вокалоида в пределах традиционных жанров. Вопреки представлениям о строго соблюдаемой чистоте традиции, синтезатор вокала шаг за шагом завоевывает себе место в рамках классического искусства. Как в области симфонической музыки, так и в традиционном японском театре — *кабуки* и *бунраку* — ветви развития традиции и моды переплетаются. С 2016 года ежегодно проводятся представления *кабуки* с участием

²⁴ В то же время в российском интернете понятие «вокалоид» демонстрирует кое-где и обратное: см. данные клуба приверженцев вокалоида: URL: https://mikupa.ru/hatsune_miku_v_anime (дата обращения: 20.04.2020).

артистов-голограмм²⁵, причем новая технология распространяется и на актеров традиционного жанра. Новый вид театрального действия изменяет структуру традиционных спектаклей, в которых не было предусмотрено общение актеров с публикой. Но виртуальность несет с собой преимущества — она ведет к расширению круга зрителей, а заодно и к омолаживанию публики.

REFERENCES

1. Kenmochi H. Singing synthesis system VOCALOID and Hatsune MIKU. *Digital Content Association of Japan* (DCAJ). 2011, pp. 33–49.
2. Shiba T. *Hatsune Miku ha naze sekai wo kaeta no ka?* (初音ミクはなぜ世界を変えたのか?) [Why did Hatsune Miku change the world?], Tokio: Ohta Publishing, 2014. 301 p.
3. Galbraith P.W., Karlin J.G. (Eds.) *Media Convergence in Japan*. Ann Arbor: Kinema Club, 2016. 296 p.
4. Davison P. The language of internet memes. In M. Mandiberg (Ed.), *The Social Media Reader* (pp. 120–134). New York: NY University Press, 2012.
5. Eidsheim N. Synthezing race: Towards an analysis of the performativity of vocal timbre. *TRANS*. 2009. 13. URL: <https://www.sibetrans.com/trans/articulo/57/synthesizing-race-towards-an-analysis-of-the-performativity-of-vocal-timbre> (accessed 22.04.2020).
6. Eidsheim N. *The race of sound: Listening, timbre, and vocality in African American music (refiguring American music)*. Durham: Duke University Press, 2019. 268 p.
7. Schapiro P. Introduction and timeline. In P. Schapiro (Ed.), *Modulations. A history of electronic music: Throbbing words on sound* (pp. viii–7). New York: DAP, 2000.
8. The Tokyo Organizing Committee of the Olympic and Paralympic Games, 2020. URL: <https://tokyo2020.org/en/news/new-robots-unveiled-for-tokyo-2020-games> (accessed 22.04.2020).
9. Jackson J., Sun-Young L. When K-pop meets technology. *The Korea Herald*, 2015. URL: <http://www.koreaherald.com/view.php?ud=20150126000821> (accessed 22.04.2020).
10. League of Legends (2018), *POP/STARS—Opening ceremony presented by Mastercard | Finals | 2018 World Championship*, URL: <https://www.youtube.com/watch?v=p9oDlvOV3qs> (accessed 22.04.2020).
11. Rhein S., Müller R. Musikalische Selbstsozialisation Jugendlicher: Theoretische Perspektiven und Forschungsergebnisse [Musical selfsocialisation of Adolescents: Theoretical perspectives and Empirical Findings]. *Diskurs Kindheits- und Jugendforschung*. 2006. 4, S. 551–568.

²⁵ Спектакли были организованы театром Минамидза (*Minamiza theatre*), Киото. URL: <https://chokabuki.jp/2016/en/> (дата обращения: 22.04.2020).

12. Nassehi A. *Muster: Theorie der digitalen Gesellschaft* [Patterns: Theory of the digital society]. München: C.H. Beck, 2019. 352 p.
13. Haar R. *Simulation und virtuelle Welten. Theorie, Technik und mediale Darstellung von Virtualität in der Postmoderne*, Bielefeld: Transcript, 2019. 385 S.
14. Thibeault M.D., Matsunobu K. Learning from Japanese vocaloid Hatsune Miku. In J.L. Waldron, S. Horsley, and K.K. Veblen (Eds.). 2020. DOI: 10.1093/oxfordhb/9780190660772.013.30.
15. Flöter L.S. *Der Avatar—die Schatten-Identität, Ästhetische Inszenierung von Identitätsarbeit im phantastischen Rollenspiel* [The avatar—the shadow identity: Aesthetic staging of identity work in a fantastic role play game]. Baden-Baden: Tectum, 2018. 473 S.
16. Watzlawick P., Beavin J.H., Jackson D.D. *Pragmatics of human communication. A study of interactional patterns, pathologies, and paradoxes*. New York: W.W. Norton & Company, 1967. 296 p.
17. Dukov E. *Set': Publika i iskusstvo*. [The net: The audience and the art]. Moscow: GII 2016. 212 p.
18. Search results for Hatsune Miku in a karaoke platform. *JoySound*, 2020. URL: <https://www.joysound.com/web/search/cross?match=1&keyword=%E5%88%9D%E9%9F%B3%E3%83%9F%E3%82%AF> (accessed 22.04.2020).
19. Kopiez R., Platz F. Der audio-visuelle Interpret: Wie die Wechselwirkung von Sehen und Hören unser Gefallensurteil beeinflusst [When the eye listens: A Meta-analysis of how audio-visual presentation enhances the appreciation of music performance]. In A. Jacobshagen (Ed.), *Perspektiven musikalischer Interpretation* (S. 25–41). Würzburg: Königshausen & Neumann, 2016.
20. Griffiths N.K. The effects of concert dress and physical appearance on perception of female solo performers. *Musicae Scientiae*. 2008. 12, pp. 273–290. DOI: 10.1177/102986490801200205.
21. Kuppens P., Tuerlinckx F., Yik M., Koval P., Coosemans J., Zeng K.J., Russel J.A. The relation between valence and arousal in subjective experience varies with personality and culture. *Journal of Personality*. 2017. 85 (4), pp. 530–542. DOI: 10.1111/jopy.12258.

ЛИТЕРАТУРА

1. Kenmochi H. Singing Synthesis System VOCALOID and Hatsune MIKU // Digital Content Association of Japan (DCAJ). 2011. P. 33–49.
2. Shiba T. Hatsune Miku ha nazesekai wo kaeta no ka [Почему Мики Хацунэ изменила весь мир?]. Tokio: Ohta Publishing, 2014. 301 p. (яп.).
3. Galbraith P.W., Karlin J.G. *Media Convergence in Japan*. Ann Arbor: Kinema Club, 2016. 296 p.

4. Davison P. The language of internet memes // *The Social Media Reader* / ed. by M. Mandiberg. New York: NY University Press, 2012. P. 120–134.
5. Eidsheim N. Synthesizing Race: Towards an Analysis of the Performativity of Vocal Timbre. *TRANS*. 2009. Vol. 13. URL: <https://www.sibetrans.com/trans/articulo/57/synthesizing-race-towards-an-analysis-of-the-performativity-of-vocal-timbre> (дата обращения: 22.04.2020).
6. Eidsheim N. *The Race of Sound: Listening, Timbre, and Vocality in African American Music (Refiguring American Music)*. Durham: Duke University Press, 2019. 268 p.
7. Schapiro P. Introduction and Timeline // *Modulations. A history of electronic Music: throbbing words on sound* / ed. by P. Schapiro. New York: DAP, 2000. P. viii–7.
8. The Tokyo Organising Committee of the Olympic and Paralympic Games, 2020. URL: <https://tokyo2020.org/en/news/new-robots-unveiled-for-tokyo-2020-games> (дата обращения: 22.04.2020).
9. Jackson J., Sun-Young L. When K-pop meets technology // *The Korea Herald*. 2015. Jan. 26. URL: <http://www.koreaherald.com/view.php?ud=20150126000821> (дата обращения: 22.04.2020).
10. League of Legends (2018), *POP/STARS—Opening Ceremony Presented by Mastercard | Finals | 2018 World Championship*, URL: <https://www.youtube.com/watch?v=p9oDlvOV3qs> (дата обращения: 22.04.2020).
11. Rhein S., Müller R. Musikalische Selbst Sozialisation Jugendlicher: Theoretische Perspektiven und Forschungsergebnisse [Музыкальная самосоциализация молодежи: теоретические перспективы и результаты исследований] // *Diskurs Kindheits- und Jugend Forschung*. 2006. Vol. 4. S. 551–568.
12. Nassehi A. *Muster. Theorie der digitalen Gesellschaft [Структуры. Теория электронно-цифрового общества]*. München: C. H. Beck, 2019. 352 S.
13. Haar R. *Simulation und virtuelle Welten. Theorie, Technik und mediale Darstellung von Virtualität in der Postmoderne [Моделирование и виртуальные миры. Теория, техника и медиальное отображение виртуальности в эпоху постмодерна]*. Bielefeld: Transcript, 2019. 385 S.
14. Thibeault M.D., Matsunobu K. Learning from Japanese Vocaloid Hatsune Miku // *The Oxford Handbook of Social Media and Music Learning* / ed. by J. L. Waldron, S. Horsley, K. K. Veblen. 2020. DOI: 10.1093/oxfordhb/9780190660772.013.30.
15. Flöter L.S. *Der Avatar—die Schatten-Identität, Ästhetische Inszenierung von Identitätsarbeit im phantastischen Rollenspiel [Аватар как теневая идентичность: Эстетическая инсценировка работы над созданием идентичности в рамках фантастической ролевой игры]*. Baden-Baden: Tectum, 2018. 473 S.
16. Watzlawick P., Beavin J.H., Jackson D.D. *Pragmatics of Human Communication. A Study of Interactional Patterns, Pathologies, and Paradoxes*. New York: W. W. Norton & Company, 1967. 296 p.

17. Дуков Е. Сеть. Публика и искусство. М.: ГИИ. 2016. 212 с.

18. Search results for Hatsune Miku in a karaoke platform // Joysound. 2020. URL: <https://www.joysound.com/web/search/cross?match=1&keyword=%E5%88%9D%E9%9F%B3%E3%83%9F%E3%82%AF> (дата обращения: 22.04.2020).

19. Kopiez R., Platz F. Der audio-visuelle Interpret: Wie die Wechselwirkung von Sehen und Hören unser Gefallensurteil beeinflusst [Аудиовизуальный исполнитель. О том, как взаимодействие между видимым и слышимым влияет на оценку позиции «нравится»] // *Perspektiven musikalischer Interpretation* / ed. by A. Jacobshagen. Würzburg: Königshausen & Neumann, 2016. S. 25–41.

20. Griffiths N.K. The effects of concert dress and physical appearance on perception of female solo performers // *Musicae Scientiae*. 2008. Vol. 12. P. 273–290. DOI: 10.1177/102986490801200205.

21. Kuppens P. et al. The Relation Between Valence and Arousal in Subjective Experience Varies with Personality and Culture // *Journal of Personality*. 2017. Vol. 85. № 4. P. 530–542. DOI: 10.1111/jopy.12258.

ABOUT THE AUTHORS

MELISSA-LILI ARENDT

Post-graduate student,
Department of Musicology,
the Martin-Luther-University Halle-Wittenberg,
Kleine Marktstr. 7,
06108, Halle (Saale)

ResearcherID: AAV-7230-2020

ORCID: 0000-0003-0792-2780

e-mail: melissaliliarendt@gmx.de

NATALIA NOWACK

Substitute professorship,
Department of Music & Musicology,
the TU Dortmund University,
Emil-Figge-Str. 50,
44227, Dortmund

ResearcherID: AAV-4888-2020

ORCID: 0000-0002-2442-6167

e-mail: natalia.nowack@web.de

СВЕДЕНИЯ ОБ АВТОРАХ

МЕЛИССА-ЛИЛИ АРЕНДТ

Магистр искусств, аспирант
кафедры музыковедения,
Галле-Виттенбергский университет
имени Мартина Лютера,
Кляйне Марктштр. 7,
06108, Галле

ResearcherID: AAV-7230-2020

ORCID: 0000-0003-0792-2780

e-mail: melissaliliarendt@gmx.de

НАТАЛИЯ НОВАК

И.о. профессора
кафедры музыки и музыковедения,
Дортмундский технический университет,
Эмиль-Фигге-Штр. 50,
44227, Дортмунд

ResearcherID: AAV-4888-2020

ORCID: 0000-0002-2442-6167

e-mail: natalia.nowack@web.de