

Limelight of the Idols: Voice, Virtuality, VTubers

Stefan Greenfield-Casas

Figure 1 shows a solo idol standing on a stage, singing a cover of Yoasobi's »Gunjou«. The ornate stage is set in grayscale, though what looks like purple ink flows in the background. Staircases flank both sides, as do giant floating orbs. Above, spotlights fan out; below, a mirrored stage floor reflects the topside world. In the background, a composite jumbotron made of smaller hexagonal screens frames the idol. The camera cuts at the pre-chorus, and an enormous crowd is displayed – mostly hidden in the shadows – but their penlights (light sticks) illuminate both the scale of the venue as well as the size of the audience. The camera cuts again to the idol. Generally speaking, she looks like a traditional Japanese idol, with notably long flowing hair and a mic set amplifying her voice... but she has little flaps on the top of her head (reminiscent of the dumbo octopus), bangs that appear to be tentacles, a floating halo behind her... and also appears to be animated! Regardless, she moves with decidedly human motions, anchoring her to the corporeal

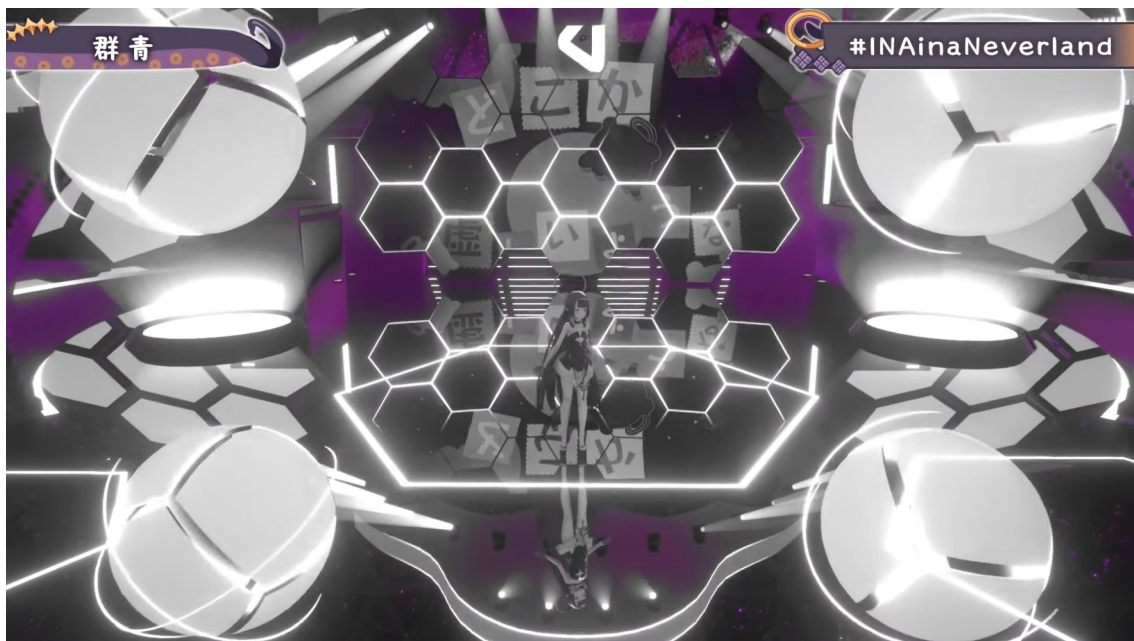


Figure 1: Beginning of Ina's »Gunjou« performance during her 2023 birthday stream. Screen capture by author

world even if she does not present as materially ›real‹. Figure 2 shows the ending pre-chorus-turned-refrain in which the stage suddenly alights with a full spectrum of flashing colors. As the song comes to an end, the jumbotron screen displays other idols with fantastical designs and costumes as the main idol dances back and forth.¹



Figure 2: Ending of Ina's »Gunjou« performance (now accompanied with her friends qua chorus) during her 2023 birthday stream. Screen capture by author

The preceding paragraph offers a description of a celebratory event that took place in December 2023. An event that took place *online*, however, and which we might describe as *virtual*, rather than actual.² Here we have the VTuber Ninomae Ina'nis, simply called ›Ina‹ by her fans and colleagues, singing and performing as part of her birthday celebration in 2023 during a live stream.³ She is a VTuber (a portmanteau of ›virtual‹ and ›YouTuber‹), an *actual* person who dons a simulated, *virtual* persona (Lo 2025) in order to stream. Furthermore, she herself exists in a liminal space between the virtual persona of

¹ Rayzorexe, Ninomae Ina'nis – 群青/Gunjou (Feat. HoloEN) 【Ina's BIRTHDAY LIVE 2023】 , YouTube, <https://youtu.be/i2S4Jz6GK7M> (09.12.2023); original stream: Ninomae Ina'nis, 【3D LIVE】 『Neverland Paradox』 Ninomae Ina'nis Birthday Live 2023 #INAINaNeverland, YouTube, <https://youtu.be/yT2ety9Rflo> (09.12.2023).

² I mean this in two senses of the word: both the contemporary understanding of the word as having to do with simulated spaces, as well as the Deleuzian sense of ›real, but not actual‹. My thoughts here are especially indebted to Paul Sanden (2019) and his theorization of ›virtual liveness‹. For a Massumian analysis of VTubers, the virtual, and the actual, see Gwillim-Thomas (2023).

³ While it appears parts of the stream were pre-recorded, it was premiered live – that is, had a set premiere date in which fans/viewers joined together to watch it when it premiered.

her idol and the actual person she really is – while there may be overlap between the two, the performative nature of being an idol should not go without comment.

From this opening and initial figures, I wish to draw attention to three things:

1. The mechanics of VTubing, called ›Live 2D rigging‹. Here, a virtual model is designed by an artist, and is then ›rigged‹ such that when the streamer moves, their model responds accordingly using motion capture technology. This is essentially virtual puppeteering and, as such, many of my thoughts going forward are indebted to Hayley Fenn's (2022) writings on puppets and music. Because these VTubers can stream as whatever they want to be, many, perhaps even most, use fantastical designs for their models. Ina's official ›lore‹ describes her as a priestess of the ›Ancient Ones‹, of which we can perhaps see some elements in her fantastical outfit. (Indeed, while not visible here, this includes tentacles she can summon from the void. I will return to this point below.) Even if a VTuber chooses to remain humanoid, however, part of the appeal of streaming as a VTuber is the ability to hide one's face and instead don a persona as a *naka no hito* (the person within). Akiko Sugawa-Shimada (2023b) (glossing Nozawa Shunsuke) notes that the concept of the *naka no hito* began with *seiyū*, Japanese *voice* actors who brought the 2D anime characters they were voicing closer to the 3D world – leading to what has since been called the 2.5D. This leads me to my second point.
2. Sound and, in particular, *voice*. If using a VTuber model is in part a way of maintaining anonymity, as musicologist Yun-Kiu Lo (2023; 2025) has convincingly argued, then the streamer's voice is the exception. While some VTubers use various ways to disguise their voice, many leave their voice unfiltered – or relatively so, at least, as they are inherently mediated through microphone, internet, and speaker. And while voice is in large part based in speaking, here we see this also comes to encompass singing. This leads me to my third point.
3. Ina is a member of Hololive Productions, the most popular VTuber agency.⁴

⁴ As the company officially stylizes it, ›hololive‹ is usually written in all lowercase. For clarity and formatting's sake, I will treat it as proper noun and capitalize the company's name across this article. Hololive is a *jimusho* or management office (talent agency) and thus maintains some level of control over these streamers and what activities they can (and cannot) do. Yet its brand is strong enough that

Though the company is based in Japan, it has English and Indonesian-speaking branches as well.⁵ While there are other prominent VTuber agencies, part of what initially separated Hololive is that it advertises its talents as *idols*, following the tradition of the Japanese idol industry (Aoyagi 2005; Lo 2025). Because of this, music and karaoke streams, known as *utawaku* (live singing), have been a critical element of Hololive since its inception.

In this article I argue that part of the reason for Hololive's success is based around its talents' *voices* and that their branding as idols highlighted the voice as a fetish object that tethers the ›live 2D‹ virtual model to the material ›real‹ world. In particular, I interrogate how fans remix and rearrange these streamers' music and voice clips to deepen their characters' narrative ›lore‹, constructed around both the streamer's real identity and their character's supposed backstory.

Furthering Hayley Fenn's (2022) claim that puppets afford a kind of ›inherent musicality‹, I draw from theories of voice, idol culture, and Japanese media to show how voice becomes an important element in bringing these virtual puppets to life. This vocal focus extends beyond the VTuber avatar, however, to construct the ›media mix‹ that fans buy into and in which they participate (Sugawa-Shimada 2023c; Steinberg 2012).⁶ Participation here includes fans purchasing merchandise, attending virtual karaoke sessions, and even using the talents' voices or music in their own creations. The music and voices of these VTubers are especially important here, as they are usually used as the basis of fan creations, ranging from animated shorts to indie games. I use these fan creations to consider not only how official Hololive music is rearranged and voices are used for these games and shorts, but also how Hololive streamers themselves engage with these fan creations.

Across this article, I will use the aforementioned VTuber Ina as my throughline case study, though I will refer to additional VTubers as well. Because these VTubers will often

many VTubers are willing to sacrifice their complete agency for the support and publicity it otherwise offers.

⁵ More specifically Ina is a member of ›Hololive Myth‹, the first generation of Hololive's English (speaking) branch. She is, however, trilingual, and can speak English, Japanese, and Korean.

⁶ As a generalization, many (if not most) VTuber fans are those also interested in games, anime, manga, and other media related to Japanese subcultures. Because of this, they are thus primed to participate within, buy into, and *consume* this media economy.

stream for hours at a time multiple times a week (resulting in hundreds – if not thousands – of hours of streaming footage), I focus on Ina as the VTuber whose streams I have viewed the most.⁷ Yet, another way fans engage with VTubers (and streamers in general) is through ›clipping‹ channels, fan channels who will curate and ›clip‹ specific moments from streams. These clips range in length and focus, sometimes highlighting something humorous a streamer said or a moment of virtuosity (or failure), sometimes translating streamers into other languages, and sometimes even editing multiple streams or points-of-view down into one cohesive narrative. Clipping itself can be an art form, with VTuber clippers especially often adding transitions, references, or memes as a way of adding a creative element to the clip. To this end, I will frequently cite clips in footnotes as a way of providing easily accessible audiovisual examples.

Finally, before moving into the body of this article, I also wish to give some insight into my own positionality here. I am not a VTuber myself (no, not even in secret) – my understanding of this phenomenon is thus rooted in theory and fandom rather than practice. I also wish to explicitly acknowledge the Japanese origins of VTubers, which will inform, but not govern, my article.⁸ Following Lo, I view this phenomenon as transnational, akin to what Rose Bridges (2017, 2023) similarly suggests of anime (here also noting that both of these media products serve Japan's soft power [Iwabuchi 2015]).

Voice and VTubing

Though VTubers have existed since at least 2016 with the self-proclaimed A.I. Kizuna AI, they became a global phenomenon in 2020. On October 20, 2020, Hololive Production's Gawr Gura became the first VTuber in Hololive (and the third VTuber overall) to reach one million subscribers on YouTube (Thomas 2020). As of the time of writing, she is

⁷ As with many people, I started watching streams during the pandemic lockdown of 2020. Ina debuted in that same year on September 12.

⁸ While a historiography examining VTubers and their relation to Japanese theater exceeds the scope of this article, the use of masks in *noh* and *kyōgen* theater, the use of puppets in *bunraku* theater, and especially the almost complete focus on voice in the oral tradition of *yose* (perhaps most famously in *rakugo*) should not go unmentioned. Indeed, one Hololive member, Juufuutei Raden, is an apprentice in the *rakugo* tradition. Raden Is My Kamioshi Clip Ch., *Raden Reveals the Origin of Her Name and Why She's Not Allowed to Perform Rakugo*. 【HOLOJP】 【ENG SUBS】 , YouTube, <https://youtu.be/Jogyqb7nu7k> (18.06.2025).

currently the most subscribed-to VTuber, with over 4.5 million subscribers (Hololive 2022).⁹ Outside of Hololive, on February 16, 2022, the VTuber Ironmouse became the most subscribed-to streamer across the entire streaming service Twitch, with over 95,000 concurrent *paying* subscribers. On September 30, 2024, she broke the all-time record for the most paying subscribers on all of Twitch, with over 320,000 concurrent subscribers.¹⁰ While I give this brief historical gloss, this article will not provide extensive consideration to the history of VTubers – my point here is largely to draw attention to the prominence VTubers have in online streaming spaces.¹¹

As I mentioned in the introduction, while these VTubers are, as with most streamers, well-known for playing games on stream, Hololive members in particular are known for another reason as well: their status as ›idols‹, and the songs they produce and sing.¹² Indeed, previous research on VTubers has also focused on the voice and, in particular, singing. Jose Garza (2022) has, for example, examined the original character songs that Hololive members commission and/or produce, arguing that they use these songs as a way of contradictorily both strengthening their persona (and thus distancing who they ›really‹ are), as well as creating a close bond between fan and streamer. Yun-Kiu Lo (2023; 2025) instead focuses on VTuber karaoke sessions, or *utawaku*, that are – as she argues – »the VTubers' signature activity« (2023). While I will address singing to some extent in what follows, I am more interested in how the voice extends beyond just singing as a way of forming (virtual) identity.

For the philosopher Mladen Dolar, the voice functions as a kind of meeting point between a neo-Cartesian split between body and, for him, *language*. It simultaneously functions as

⁹ At the time of writing, now 45 of Hololive's 88 members have over one million subscribers on YouTube.

¹⁰ Part of Ironmouse's success here was likely due to her raising money for the Immune Deficiency Foundation. She herself cannot leave her house because of her compromised immune system, and so found solace in streaming (Diaz 2024).

¹¹ For those interested in an in-depth history of VTubers, see Kovacic (2021) and especially Lo (2025).

¹² Idols (transliterated as *aidoru* in Japanese) are entertainers (›talents‹) who are trained and marketed not only as singers but also as dancers, actors, models, and overall media personalities. While the term did not have the religious connotations one might associate with idolatry, recent scholarship has started to consider otherwise (Aoyagi, Galbraith, and Kovacic 2021). This quasi-religious fanaticism is perhaps best represented with the manga/anime *Oshi no Ko* and especially the anime's opening »Idol« as performed by YOASOBI.

a break from, as well as an indication of, material presence (Dolar 2006, 59–60). In the context of VTubers, this relates to Sugawa-Shimada's theorization of 2.5D, a »space between the two-dimensional (fictional space where our imagination and fantasy work) and the three-dimensional (reality where we physically exist)« (2023c, 47).¹³ Here, because voice is the only ›real‹ element of the streamer that fans experience, VTuber voices become fetish-objects, a Lacanian ›objet petit a‹ to close the distance between the streamer and fan. Voice here is porous, materially bridging the virtual and real, offering 3D audio to an otherwise 2D screen, a sonic manifestation of Sugawa-Shimada's theorization of the 2.5D.¹⁴ This obsession with voice is hardly without precedent. I am reminded of Proust and his Grandmother's voice (Žižek 2012), the film *DIVA* (France 1981, Jean-Jacques Beineix) in which a postman secretly records the voice of a soprano who has refused to have her voice recorded (cf. Poizat 1992), and, more recently, Robert Harper playing almost two hundred hours of the game *FINAL FANTASY XIV* just to hear his late friend's voice one last time in the game (Williams 2022).

There is no singular way VTubers approach their respective voices. Some follow the Japanese *kawaii* (›cute‹) aesthetic, performing an overly high voice that is not their natural speaking voice, or by even manufacturing vocal tics (such as Hololive's Usada Pekora ending most sentences with »peko«).¹⁵ Others instead use their natural voices, but have what Barthes (1977) would identify as a grain in their voice based on their inherent vocal

¹³ To be clear, Sugawa-Shimada did not invent this term, though she has theorized the 2.5D more than any other English-language scholar (see also Galbraith 2020). For a brief history of the term see Sugawa-Shimada (2023a); other representative writings on the 2.5D can be found in the special issue to which this article serves as introduction.

¹⁴ I would further contend that the VTuber is the ultimate (or at least most balanced) realization of the 2.5D. Other 2.5D examples such as cosplayers or materially actual idols lean much closer to the 3D and thus cannot truly fulfill the fantasy of the 2D as well as the VTuber. On the flip side, a character like Hatsune Miku is too rooted in the 2D, even though ›she‹ materially exists as a 3D vocal synthesis program (Kovacic 2021: 95-96). And yet, as Sugawa-Shimada argues: »it is not the case that fans are incapable of distinguishing between fantasy and reality. More precisely, they have a pleasurable interplay with the ambiguous boundary between fantasy and reality within 2.5-D space« (2023b: 131).

¹⁵ As Reddit user deoxix (2020) elucidates: »In anime some characters have some finish particles to their phrases like »-desu« or »-aru« that never happens in real life but for Japanese natives it makes them sound cuter. For Pekora the particle is »-peko«. In her lore (please don't take it seriously) this is because she comes from the country of Pekoland. There they speak Pekolandish in which somehow every word is »peko«. Pekora uses this expression often to laugh before doing some mischievous plan or to mock them after they lose to her. Sometimes wrote with arrows [PE↗ KO↘ PE↗ KO↘ PE↗ KO↘ PE↗ KO↘] to represent the very characteristic changes of pitch she does while saying it. Once you hear it you will never forget it.«

qualities or perhaps even a regional dialect. The now-retired VTuber Shizukou, for instance, has a strong British accent which she leans into for some of her short skits.¹⁶ A handful of VTubers even explicitly disguise their voices, whether by using voice modulators or, as the indie VTuber Zentreya uses, text-to-speech to communicate.¹⁷ Some are well-known for their singing (and some are even professional voice actors),¹⁸ but others are known specifically for their voice superseding language: namely, their laughs, screams, or sneezes.¹⁹

The importance of VTubers' voices is perhaps made most obvious by the VTuber Ophelia Midnight's ›voice study‹ streams, where she chooses a popular VTuber and analyses and replicates their voices.²⁰ Another famous case of VTuber vocality deals with the VTuber Snuffy, who initially used (performed) a high, *kawaii* voice for her streams, but then eventually revealed that her natural voice was much lower. Somewhat surprisingly, this revelation was generally well-received, with fans and other VTubers praising Snuffy for her honesty.²¹

There are also many examples of streamers reflecting on their own voice, as well as the voices of others, on stream.²² Ina – with whom I began my article – has discussed how she

¹⁶ Shizuko, *What people expect me to sound like when I say I'm from England*, TikTok, www.tiktok.com/t/ZP88cPEep/ (10.06.2022).

¹⁷ Cooksie, *Zen's TTS had a voice crack*, YouTube, <https://youtu.be/o7O0-QRnhrs> (17.09.2021).

¹⁸ Hoshimachi Suisei is perhaps the most notable VTuber singer/idol (she is also a member of Hololive). This is in many ways encapsulated in her being the first (and so far only) VTuber to perform on *The First Take: THE FIRST TAKE*, *Hoshimachi Suisei – Stellar Stellar / THE FIRST TAKE*, YouTube, <https://youtu.be/AAsRtnbDs-0> (20.01.2023). As another example, ›Monarch‹ is the VTuber persona of the professional singer and voice actress AmaLee.

¹⁹ VTubers famous for their laughs include Dokibird, Henya the Genius, and the aforementioned Pekora (see footnote 15). Uruha Rushia was especially known for her ›death metal‹ scream before she stopped streaming. As for sneezes, fans will often type ›TSKR‹ (an abbreviation for ›Tasakaru‹ which is a kind of blessing in Japanese) after a VTuber sneezes while streaming. This is doubly coded as something akin to a ›bless you‹, as well as thanks for this ›blessing‹ (i.e., hearing an idol, a supposedly ›perfect‹ being, doing something un-idol-like; see footnote 12).

²⁰ Though she is no longer active, the playlist of her voice studies is available at Ophelia Midnight Ch. [VReverie], *Vocal Study* (playlist), YouTube, www.youtube.com/playlist?list=PLnWQAsQI24TJoMhgxyua4DcmEaQqJh2V9 (27.08.2025). Another YouTuber who provides VTuber voice analyses is Mirii; her playlist of these analyses is available at Mirii, *Vtuber Voice Analysis* (playlist), YouTube, www.youtube.com/playlist?list=PL77ipalFoor8vRiCn976X5agiBUd-XkxE&si=qEfNbu6VhxrxtkaR (27.08.2025).

²¹ Venandr, *how revealing her REAL VOICE blew this VTuber up (Snuffy) #Shorts*, YouTube, <https://youtu.be/LW53zbRAHMc> (20.05.2022).

²² Holofan Clips, *Matusri's Fake Voice And Holomems Who Uses Their Real Voice 【Hololive】*, YouTube, <https://youtu.be/6sHX0MypaIw> (25.04.2025); see also this extended compilation focusing on the ›Real Voice‹ of various Hololive members: Jihuu Translation, *Hololive: Real Voice*,

herself questions the difference between how her voice currently sounds while she streams versus how it sounded when she first started.²³ Another Hololive VTuber, Inugami Korone, has further reflected that depending on what virtual medium she is streaming, she unintentionally uses different voices (e.g., on YouTube she projects more than she does when she is hosting a ›Space‹ on X [formerly known as Twitter]).²⁴ Perhaps the most explicit reflection on voice comes from Amanogawa Shiina's TikTok addressing her ›real voice‹. As she states in response to accusations of not using her ›real‹ voice: »Don't act like you don't have different tones of voice for when you're doing customer service versus when you're just with your closest friends. And no, it is not a voice changer, it is my *real* voice.«²⁵ Here, Shiina draws attention to the fact that voices change for *everyone* and should not be considered immutable (cf. Sterne 2003).²⁶ For Ina, this was over the course of many years, perhaps in part because of becoming more relaxed as a streamer, perhaps because of a new microphone, perhaps because of a physiological change. For Shiina, however, our voices are always in flux, varying not only on a daily basis, but even from one event to another.

With these virtual vocalities, I am reminded of Michel Chion's theorization of vococentricism in film. As he states of a film's soundtrack, »there are not *all the sounds including the human voice*. *There are voices, and then everything else*« (1999, 5–6, emphasis in original). Here, however, we have moved from beyond the classical screen-based model of the projected film to the contemporary practice of streaming on the screens of computers and smartphones. *And yet this focus on voice remains*. To be clear, the attention to voice is not completely unique to VTubers – generally speaking, most streamers will have some emphasis on their voice when streaming. This could range from communications with teammates, to talking to chat (i.e., those watching and commenting on the stream in

YouTube, <https://youtu.be/cVFdLAf6xSg> (20.08.2024).

²³ Vaan Ch., *Ina can't imitate Ninomae Ina'nis voice*, YouTube, https://youtu.be/g_gdHjSp21U (30.06.2024).

²⁴ OtakMori Vtuber Clips, *Korone Shows the Differences Between Her Real Voice and Stream Voice* **【Hololive | Eng Sub】**, YouTube, <https://youtu.be/OfaRTGmyXgc> (29.05.2023).

²⁵ Amanogawa Shiina, *Is This VTuber's Voice Fake?*, TikTok, www.tiktok.com/t/ZP88V4beL/ (05.05.2023).

²⁶ The VTuber Nimi Nightmare comes to a similar conclusion in her YouTube short considering VTuber voices: Nimi Nightmare, *Why Vtubers sound different than their debuts*, YouTube, www.youtube.com/shorts/RvrGhNbtLOo (27.01.2025).

real time [Anderson 2017]). As Stephen Tatlow writes regarding *player* voices in video games:

Voice will always symbolize part of the real identity of the player. Vocalizations may allow players to identify – or at least assume – each other’s identity[,] gender, ethnicity, geographical or cultural upbringing, and more. As a result, it may create tensions when connected to a specific in-game character that the player controls. In these situations, voice becomes part of the character identity, introducing physicality to virtuality. (2020, 17)

There is a similar vocal play here with VTubers, but at an intensified level; here, voice becomes existential. Note my framing: Sartrean existentialist, not essentialist. That is, rather than voice essentializing one being and its essence, it is one way of bringing *into* existence the VTuber model, the virtual puppet as an instrument. While it would be easy to say that the *naka no hito* breathe life into their VTuber models, I will instead quote Hayley Fenn in her discussion of marionettes and ›Marionettenspieler‹ (puppeteers):

Marionettenspieler do not breathe life into the puppet, but rather breathe life *with* the puppet. In this sense, and through the vibrancy of its strings, the marionette is both dangly appendage to and vitally independent from the Marionettenspieler, which, to the audience, reads as authorial, in the moment. (2022, 227)

In other words, this is a symbiotic relationship between puppet and puppeteer. And, if the Live 2D models reveal the flatness of these posthuman beings (and here I am reminded of Takashi Murakami’s theory of the ›superflat‹ in Japanese art [Favell 2011]), then the voice fleshes them out, creating a hyperreal (Baudrillard 1994) identity. VTubers are hardly unaware of this fact and will not infrequently draw intertextual connections that highlight their post or augmented humanity. Ina, for instance, has created official art for Hololive which she titled ›Do VTubers Dream of Colorful Tomorrows?‹ The name is, of course, in part a reference to Philip K. Dick’s classic science fiction novel *Do Androids Dream of Electric Sheep* (1966), which was adapted into the cult cyberpunk film *BLADE RUNNER* (USA 1982, Ridley Scott).²⁷ In a more musical and vocal vein, we might consider

²⁷ Ninomae Ina’nis, ›Do Vtubers Dream of Colorful Tomorrows?‹, X [Twitter], <https://x.com/ninomaeinanis/status/1637835283273244673?lang=en> (20.03.2023). Including ›tomorrow‹ in the title is a reference to Ina’s most famous stream clip, which I will discuss below (see footnote 38). And her emphasis on ›colorful‹ is perhaps also reflected in the ending of her ›Gunjou‹ performance (Figure 2), with Hololive and its members here literally brightening her world.

Gawr Gura singing »The Moon Song« from the movie HER (USA 2013, Spike Jonze), a film about a sentient AI, Samantha, whose existence is purely vocal.²⁸

Fenn builds a theory of the ›inherent musicality‹ of puppets, a »form of musicality [that] traverses sensory planes [...] and is ultimately made possible by their fundamental muteness, a muteness that allows them to appear as both vessel and creator, listener and singer« (2022, 263). Her theorization here is informed by Actor Network Theory, and thus elides essentialist formations of identity, and instead attends to the constituent parts (or actants) of what she calls the ›performance network‹:

When expressive agency traverses an entire network of performing bodies, technologies, and objects, synchronization effects are not fixed but emergent, constructing diverse and flexible relationships between sound and image. My concept of a »poetics of synchronization« captures the relationships of simultaneity and alignment inherent in audio-visual media, while accounting for the multiple perspectives and shifts of agential concentration engendered by puppet performance. (2022, 172)

This poetics of synchronization here is further mediated by virtual time stretched across both streaming latency (Gagen and Cook 2016), but also the temporal displacement of how and when fans hear these voices.²⁹ Thinking of these elements in the context of a network leads me to consider how voices exist beyond their streams: namely, the media mix.

Media Mix and Sekaikan

Sugawa-Shimada's theorization of the 2.5D is explicitly based on Japanese pop culture, where she draws particular attention to the Japanese theory of the media mix (*media mikkusu*), a »media franchise marketing strategy in which cultural contents in multiple works are produced or distributed across multiple media platforms« (2023c, 47). The media mix is a two-pronged theoretical construct, what Marc Steinberg (2012) identifies as the *marketing* media mix, and the *anime* media mix. The former is the actual Japanese marketing theories that construct the latter, the assemblage of (generally physical) media

²⁸ Pardun?, *Gawr Gura cried while singing The Moon Song - Full part 1 and 2*, YouTube, https://youtu.be/qIU5_GCnL8Q (02.11.2020).

²⁹ Beyond streaming latency, we might also consider a fan going back to (re)watch the video-on-demand of a stream that has already occurred.

products for a particular anime or game series, e.g., toys, DVDs, trading cards, foodstuffs, and home goods. Writing on the media mix as related to 2.5D theatrical performances, Sugawa-Shimada suggests that »participation of fans is vital to create the 2.5-D space. Fans show their affection for the characters/actors they love by purchasing all the goods related to them. This practice is called *ofuse* (›donation to gods‹) and to gift the goods or rent DVDs to their friends who are unfamiliar with the 2.5-D theatrical performances is called *fukyô* (›missionary work‹), as it is expanding their network« (2023b, 133).

In terms of Hololive, we can see this represented in the various goods the company and its talents will sell to fans. While some of these may be somewhat expected (e.g., figures and posters of the talents, CDs and vinyl of their music), others are perhaps less so (e.g., a laundry hamper, voice packs, etc.).³⁰ Yet the media mix ties into another Japanese media theory, what cultural theorist Ôtsuka Eiji (2010) theorizes as the ›*sekaikan*‹.³¹ The *sekaikan* is the ›grand narrative‹ that encapsulates all stories, experiences, and objects related to an overarching franchise, group, event, etc. Relating this back to Actor Network Theory, the *sekaikan* is the overarching network, built up from its composite elements, and the media mix constitutes the (nonhuman) actants within the network. As Matthew Richardson writes while examining Japanese idols, »the media mix permits access to an intangible *sekaikan* mediated by the products. The question of an original is wholly irrelevant, because the media mix facilitates ongoing engagement with a set of affects experienced through time« (2016, 153).

For Ôtsuka, the *sekaikan* is born in large part from the *otaku* subcultures in Japan and their obsessive attempts to consume the *sekaikan* in its entirety. In his frequently cited essay on the phenomenon, Ôtsuka gives examples ranging from the Bikkuriman Chocolates and the highly prized stickers that came with them, to the MOBILE SUIT GUNDAM anime. The *sekaikan*, then, is constructed from not only the ›lore‹ of these characters, but also their daily (give or take) streaming and anything that contributes to the media mix. Thus, a

³⁰ Geek Jack, [20220521 - 20220627] »Ninomae Ina'nis Birthday Celebration 2022« »Laundry Ina Bag«, *Geek Jack*, <https://shop.geekjack.net/products/ninomae-inanis-birthday-celebration-2022-laundry-ina-bag> (04.05.2025); Geek Jack, hololive Summer 2022 Episode Voice Complete Set, *Geek Jack*, <https://shop.geekjack.net/collections/hololive-summer-2022/products/hololive-summer-2022-episode-voice-set> (04.05.2025).

³¹ While *sekaikan* translates literally to ›worldview‹ or ›Weltanschauung‹, it might more productively be thought of as a kind of *worldbuilding*.

clothes hamper is just as important as a voice pack a fan might purchase, specifically because both ultimately contribute to and extend the world of Hololive in a fansumer's daily life.

But the *sekaikan* does not stop there. While English theories of transmedia tend to privilege one canonical narrative (Jenkins 2008, 2010), the *sekaikan* is agnostic to such rigid narratives, to such a point that even *fan creations* are seen as just as important in the construction of this world.³² While fan art is almost certainly the most common form of fan creation, perhaps because Hololive's streamers are *gaming* idol streamers, fan-produced games – often created at high levels of production – are not uncommon. Notable examples include the platformer SMOL AME (2021), the horror game EVIL GOD KORONE (2021), the ROGUELITE and VAMPIRE SURVIVORS homage HOLOCURE (2022), and the fighting game IDOL SHOWDOWN (2023). The latter two games even broke out of the Hololive fandom and became well regarded games by those with no knowledge of Hololive and its streamers. Other games, such as AQUARIUM (2023) and DAYS WITH OLLIE (2024) were direct collaborations between Hololive members and fan development teams. Here, the *sekaikan* legitimized the prosumer's creation with Hololive's development of ›Holo Indie‹, Hololive's game publishing division.³³

In what follows, I will consider Ina's role in Besto Game Team's IDOL SHOWDOWN, published by Holo Indie and released in 2023. While the game allows players to fight using their favorite Hololive talents with special attacks based on their character's lore and/or inside jokes from their streams, notably, it also includes their *voices*, initially clipped directly from their streams. ›Initially‹ is the operative word here, because some of the talents themselves eventually offered to (re)record their voices for the developers of the game. In one notable case, Shishiro Botan was streaming IDOL SHOWDOWN for her audience, playing as herself in the game. After one battle, she comments that the developers of the game must have clipped her voice from previous streams, but that she was (playfully) unhappy with the end result. To better represent herself, she instead paused her

³² This is not to diminish the role fandom and, e.g., fan fiction play outside of Japan – but the emphasis and building upon specific IPs is where I argue the difference exists (see Greenfield-Casas and Mc Glynn 2023; cf. Sugawa-Shimada 2023b: 127).

³³ *holo Indie*, CCMC, <https://ccmc-corp.com/en/holoindie/> (accessed 04.05.2025).

playthrough of the game and (on stream in real-time) provided voice clips she would approve of within the game. She begins with her sing-song catchphrase »La-Lion«, directly imploring that the developers use it as her greeting. From there, she records more dynamic sounds – the imagined sounds of taking a hit, as well as of combating others.³⁴ Here, voice works both to strengthen the fidelity of the game, but also as a ›fiduciary medium‹, to borrow a recent term from Matthew Mendez (2023). In other words, it is a live vocal signature of endorsement for this new world *qua* game.

IDOL SHOWDOWN Case Study

Though IDOL SHOWDOWN was initially released in 2023, Ina was only added into the game in May of 2024. Her attacks, like all playable characters in the game, are based on one of five reference types, what I will call preexisting, character, personal, stream, and postexisting references. Preexisting references are based on hypotexts that predate Ina as a streamer, for instance Kuroboshi's (the artist who illustrated Ina's character) love of octopuses and Lovecraftian mythos. Character references are based on the ›lore‹ of Ina as a character, for example, her being a ›Priestess of the Ancient Ones‹ (and thus able to summon tentacles from ›The Void‹) set on world domination. Personal references are based on traits or facts about the actual streamer outside of their streams. Ina (or the *naka no hito* who plays [as] Ina) is a professional illustrator outside of her streams, so in the game she primarily fights with a paintbrush. Stream references include inside jokes or references that have come about from streams, such as Ina threatening to ›bonk‹ chat with a crowbar if they misbehave, and so she also sometimes fights with a crowbar.³⁵ Finally, postexisting references (which I borrow as a term from Jonathan Godsall (2019), though with a different meaning here) are based on fan extensions of VTubers and their lore – the *sekaikan*. The example par excellence for Hololive is animator mazu's 12-minute »Myth's Bad Ending« animation depicting a possessed Ina (shown in Figure 3) fighting

³⁴ Sushi Clips, *Botan records and provides her voices to the game developer during her stream [Hololive/Eng sub]*, YouTube (06.05.2023), https://youtu.be/-QdVxqaWp_4; original stream: Botan Ch. 獅白ぼたん, **【Idol Showdown】最強のアイドルは誰だ...！ホロライブ格闘ゲームやるぞい！【獅白ぼたん/ホロライブ】**, YouTube, <https://youtu.be/LEgrV2vwORQ> (06.05.2023).

³⁵ Holopaca, *What happens when we squish Inas hair?*, YouTube, <https://youtu.be/dbPMVH6jtGY> (31.10.2020).

against the other four members of Hololive Myth (her friends).³⁶ Figure 4 shows Ina's ultimate attack in the game, ›Forbidden *INK*antation‹, which is based on this animation.³⁷



Figure 3: A possessed Ina fighting against her friends in »Myth's Bad Ending«.
Screen capture by author

While the game was originally released with voice clips selected by the developers, I have already shown how some talents instead re-recorded (on stream) voice clips they would prefer used in the game. Because Ina's character was added essentially a year after the initial release, however, she instead collaborated with the developers to voice most of her lines (with the most notable exception being a preexisting clip of Ina's voice cracking in excitement while exclaiming »tomorrow!« that has garnered 1.5 million views).³⁸ As she stated of this voice clip on stream while playing IDOL SHOWDOWN: »That, I did not rerecord! That is a clip. *laughs* I'm glad they added it, though.« She continues on to try

³⁶ Mazumaro, *Myth's Bad Ending* | Hololive fan-animation, YouTube, <https://youtu.be/euPfAilSpuU> (29.10.2022).

³⁷ Idol Showdown | Besto Game Team, *Idol Showdown Next Fes: Ninomae Ina'nis Debut Trailer*, YouTube, <https://youtu.be/mNtDthjAs9Y> (28.04.2024). It is worth noting that Ina is quite fond of puns. Because of this, all of her special attack names are puns.

³⁸ Random Clipper, »Tomorrow«, *Ina said calmly...*, YouTube, <https://youtu.be/sMcfFmR0MmA> (10.02.2022); original stream: *Ninomae Ina'nis, Nintendo Direct Feb 2022 Watchalong! ✖Not a Mirror Stream*, YouTube, <https://youtu.be/EHpxi7khHb0> (09.02.2022).

and replicate the iconic »tomorrow!« but fails, saying with a laugh, »I can't voice crack on command!«³⁹

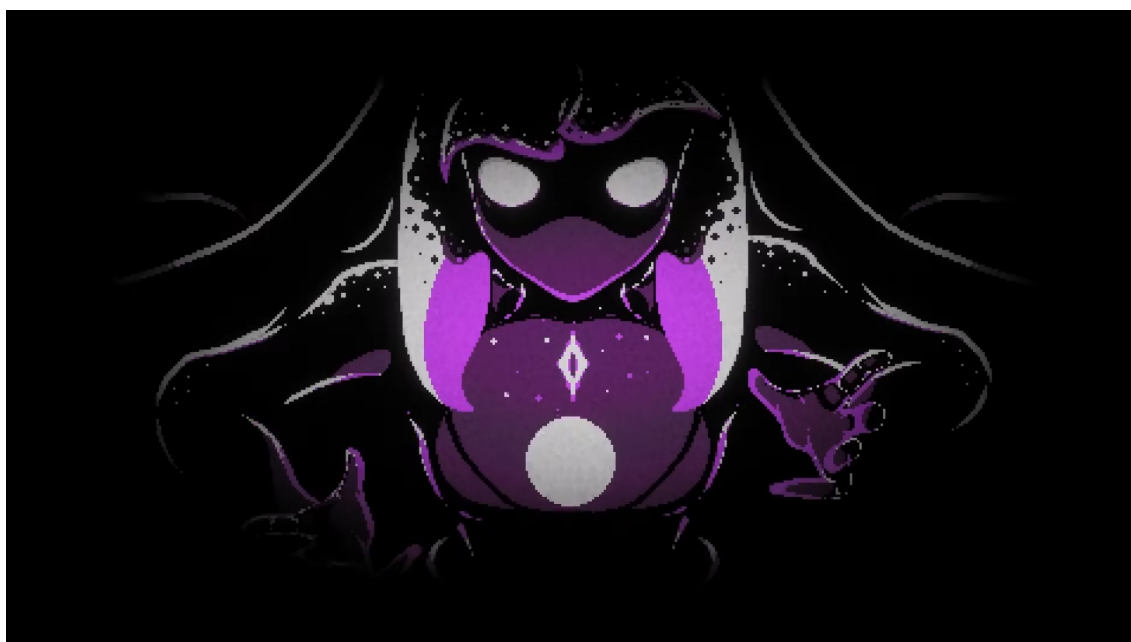


Figure 4: »Forbidden *INKantation*« in *IDOL SHOWDOWN*. Screen capture by author

Finally, beyond Ina's voice-*qua*-voice, she is musically represented in the game with two musical arrangements: the first is »WAHoy!«,⁴⁰ a transformative arrangement of her first original song »Violet« blended with »Ahoy!« a song by Hololive's Houshou Marine.⁴¹ In this arrangement, a choir chants Ina's name, harkening to her status as a Priestess of the Ancient Ones. Also sampled, however, is Ina's own voice: from saying her signature »Wah!« (with which she opens all streams) and her own name, to her laugh and some dialogue as well.

The second piece is »Violet Blossom«,⁴² an electronica mashup of »Kanalumi«, a piece Ina commissioned from one of her favorite composers, Camellia, and »Meconopsis«,

³⁹ Vaan Ch., *The Ninomae Ina'nis finally appears in Idol Showdown*, YouTube, <https://youtu.be/eSRevWnPj5M> (07.05.2024); original stream: Ninomae Ina'nis, *Idol Showdown* INA INA INAAAAA, YouTube, <https://youtu.be/XYL0ruvuu28> (06.05.2024).

⁴⁰ Idol Showdown | Besto Game Team, *Idol Showdown OST Track 30 – »WAHoy!«* *【VIOLET / Ahoy!! 我ら宝鐘海賊団☆】*, YouTube, <https://youtu.be/Q1yBM6BRU4g> (13.08.2023).

⁴¹ Marine is also one of Ina's unitmates in UmiSea, an idol subunit within Hololive; at the time of writing, she is the second-most subscribed-to VTuber (behind Hololive's Gawr Gura), with just shy of 4 million subscribers.

⁴² Idol Show | Besto Game Team, *Idol Showdown OST Track 37 – Violet Blossom* *【Kanalumi, Meconopsis】*, YouTube, <https://youtu.be/jEuaFsH9YKA> (06.05.2024).

Ina's second original song, composed by another one of her favorite composers, M2U.⁴³ While Ina's *voice* is not sampled in this piece, gamers who played as Ina when she first became available in the game would likely already have her »Meconopsis« melody in their ears, as she officially released this song within just a few days of becoming a playable character in the game. The attention Ina is given in the game – referentially, musically, vocally – is indicative of the care the development team put into *all* elements of the game. Indeed, while IDOL SHOWDOWN's 52-track OST does feature original music, most of its score is based on arrangements of Hololive's catalogue of over 500 original songs.

Conclusion

In this article, I have considered the role *voice* plays for VTuber – and especially Hololive – streams. Drawing on Sugawa-Shimada's theorization of the 2.5D, I have argued that the voice is one instrumental way of bridging the gap between the virtual 2D world of the screen-and-stream and the material-and-real 3D world. Furthermore, I have examined how VTubers' voices are used beyond their stream by fans – *prosumers* – as one way of extending the all-consuming narrative of the *sekaikan*. But this is an ouroboric parasocial relationship (Hoffner and Bond 2022), where the streamers themselves then play these games for fans and (if the game is ultimately endorsed) even work directly with the developers to continue to build this worldview.

In the conclusion to her recent paper on VTuber anonymity and singing, Yun-Kiu Lo (2023) noted that »Many of my interviewers responded that they are concerned about their voice control and feel the *utawaku* and music activities are necessary in VTuber streaming, but they often cannot come up with a solid reason [as to why].« While she continues on to suggest that karaoke streams offer a way for VTubers to emotionally connect with their audience over music, I offer another, perhaps too obvious answer: that is, because the most popular VTuber agency built its success around singing, other VTubers felt they

⁴³ She mentioned both of these composers during her debut stream: Ninomae Ina'nis, **【DEBUT STREAM】 Hi! #hololiveEnglish #holoMyth**, YouTube, 12 September 2020), <https://youtu.be/ujCxiHpVYmg?t=1300> (27.08.2025). The given link accounts for the specific time, but she begins talking about these composers at 00:21:40.

must follow suit. While voice is obviously an important element of VTubing and streaming in general, that should not necessitate the need *to sing*. And yet, even Hololive's weakest singers still perform karaoke (Lo 2025), even when they have vocal talents elsewhere: whether through ASMR streams, voice impressions, or perhaps even a general grain to their voice. Hololive did not create the first VTubers, but in becoming the most popular company and one centered around voice *and especially singing*, their talents set the stage for others to follow suit.

Bibliography

- Aoyagi, Hiroshi (2005): *Islands of Eight Million Smiles: Idol Performance and Symbolic Production in Contemporary Japan*. Cambridge: Harvard University Press.
- Aoyagi, Hiroshi/Galbraith, Patrick W./Kovacic, Mateja (eds.) (2021): *Idology in Transcultural Perspective: Anthropological Investigations of Popular Idolatry*. Cham, Switzerland: Palgrave Macmillan.
- Anderson, Sky LaRell (2017): Watching People Is Not a Game: Interactive Online Corporeality, Twitch.tv and Videogame Streams. In: *Game Studies: The International Journal of Computer Game Research* 17/1, <https://gamestudies.org/1701/articles/anderson> (26.08.2025).
- Barthes, Roland (1977): The Grain of the Voice. In: Heath, Stephen (ed. and trans.): *Image Music Text*. London: Fontana Press, pp. 179–189.
- Baudrillard, Jean (1994): *Simulacra and Simulation*. Translated by Sheila Faria Glaser. Ann Arbor: The University of Michigan Press.
- Bridges, Rose (2017): *Cowboy Bebop Soundtrack*. New York: Bloomsbury Academic (33 1/3 Japan).
- Bridges, Rose (2023): *Constructions of Gender, Sexuality, and Identity in Contemporary Japanese Anime Soundtracks*. PhD diss., University of Texas at Austin.
- Chion, Michel (1999): *The Voice in Cinema*. Translated by Claudia Gorbman. New York: Columbia University Press.
- Deoxix (2020): The Basic Hololive JP dictionary for newcomers! In: *Reddit*, November 27, www.reddit.com/r/Hololive/comments/k2dfz8/the_basic_hololive_jp_dictionary_for_newcomers/ (03.05.2025).
- Diaz, Ana (2024): Ironmouse is now the most subscribed-to channel on Twitch ever. In: *Polygon*, 30 September, www.polygon.com/twitch/458345/ironmouse-breaks-subs-record-twitch-subathon (04.05.2025).
- Dolar, Mladen (2006): *A Voice and Nothing More*. Cambridge: MIT Press.
- Favell, Adrian (2011): *Before and After Superflat: A Short History of Japanese Contemporary Art 1990–2011*. Hong Kong: Blue Kingfisher Limited.

- Fenn, Hayley (2022): *Breath, Gravity, Giants, and Death: Puppet-Musical Encounters from Die Zauberflöte to Today*. PhD diss., Harvard University.
- Gagen, Justin/Cook, Nicholas (2016): Performing Live in Second Life. In: Whiteley, Sheila / Rambarran, Shara (eds.): *The Oxford Handbook of Music and Virtuality*. Oxford and New York: Oxford University Press, pp. 191–209.
- Galbraith, Patrick W. (2020): 2.5-Dimensional. In: *Japanese Media and Popular Culture*, 1 May, <https://jmpc-utokyo.com/keyword/2-5-dimensional/> (22.07.2025).
- Garza, Jose (2022): Voice of a New Idol: The VTuber Original Song as a Form of Agency. Presentation at *Music and the Moving Image*. 26–29 May.
- Godsall, Jonathan (2019): *Reeled In: Pre-existing Music in Narrative Film*. New York: Routledge.
- Greenfield-Casas, Stefan/Mc Glynn, James Denis (2023): »We Are What They Grow Beyond«: Visions (and Sounds) of a Transnational *Star Wars*. In: *Musicology Now*, 18 December, <https://musicologynow.org/we-are-what-they-grow-beyond-visions-and-sounds-of-a-transnational-star-wars/> (03.05.2025).
- Gwillim-Thomas, Patrick (2023): The Actualizing Platform: 2.5 Dimensions in the VTuber Media Ecology. In: Sugawa-Shimada, Akiko (ed.): »2.5D Cultures« Special Issue, *Mechademia: Second Arc* 15/2, pp. 49–69.
- Hoffner, Cynthia A./Bond, Bradley J. (2022): Parasocial Relationships, Social Media, & Well-being. In: *Current Opinions in Psychology* 45, PMID: 101306.
- Hololive (2022): Gawr Gura Has Achieved 4 Million YouTube Subscribers. In: *Hololive*, 16 June, <https://hololive.hololivepro.com/en/news/20220616-02-2/> (04.05.2025).
- Iwabuchi, Koichi (2015): Pop-culture Diplomacy in Japan: Soft Power, Nation Branding and the Question of »International Cultural Exchange«. In: *International Journal of Cultural Policy* 21/4, pp. 419–432.
- Jenkins, Henry (2008): *Convergence Culture: Where Old and New Media Collide*. Revised edition. New York: New York University Press.
- Jenkins, Henry (2010): Transmedia Storytelling and Entertainment: An annotated syllabus. In: *Continuum: Journal of Media & Cultural Studies* 24/6, pp. 943–958.
- Kovacik, Mateja (2021): Cyborg in Idology Studies: Symbiosis of Animating Humans and Machines. In Hiroshi Aoyagi/Galbraith, Patrick W./Kovacik, Mateja (eds.): *Idology in Transcultural Perspective: Anthropological Investigations of Popular Idolatry*. Cham, Switzerland: Palgrave Macmillan, pp. 91–131.
- Lo, Yun-Kiu (2023): The Authentic Anonymous Idols: Virtual Sound, Music and Play on VTuber's Identity. Paper presented at the »Co-op« *Conference of the AMS Ludomusicology Study Group*. Online. 28 January.
- Lo, Yun-Kiu (2025): Idols Living in the Virtual: The Authentication of VTuber's Persona Through Sound and Music in Digital Realm. In: *Journal of Sound and Music in Games* 6/2, pp. 96–121.
- Mendez, Matthew (2023): Fiduciary Media: On Music's Contractual Organologies. Paper presented at *Instruments, Interfaces, Infrastructures: An Interdisciplinary Conference on Musical Media*. Harvard University, 11–13 May.

- Ōtsuka, Eiji/Steinberg, Marc (trans. and ed.) (2010): World and Variation: The Reproduction and Consumption of Narrative. *Mechademia: Second Arc* 5 »Fanthropologies«, pp. 99–116.
- Poizat, Michel (1992): *The Angel's Cry: Beyond the Pleasure Principle in Opera*. Translated by Arthur Denner. Ithaca: Cornell University Press.
- Richardson, Matthew (2016): *Marketing Affect in Japanese Idol Music*. PhD diss., Northwestern University.
- Sanden, Paul (2019): Rethinking Liveness in the Digital Age. In: Cook, Nicholas/Ingalls, Monique M./David Trippett (eds.): *The Cambridge Companion to Music in Digital Culture*. Cambridge: Cambridge University Press, pp. 178–192.
- Steinberg, Marc (2012): *Anime's Media Mix: Franchising Toys and Characters in Japan*. Minneapolis: University of Minnesota Press.
- Sterne, Jonathan (2003): *The Audible Past: Cultural Origins of Sound Reproduction*. Durham: Duke University Press.
- Sugawa-Shimada, Akiko (2023a): Introduction. In: Sugawa-Shimada, Akiko (ed.): »2.5D Cultures« Special Issue, *Mechademia: Second Arc* 15/2, pp. 1–8.
- Sugawa-Shimada, Akiko/Nishimura, Keiko (trans.) (2023): Emerging ›2.5-dimensional‹ Culture: Character-oriented Cultural Practices and ›Community of Preferences‹ as a New Fandom in Japan and Beyond. In: Sugawa-Shimada, Akiko (ed.): »2.5D Cultures« Special Issue, *Mechademia: Second Arc* 15/2, pp. 124–139.
- Sugawa-Shimada, Akiko (2023c): A 2.5D Approach to the Media Mix: The Potentialities of Fans' Produsage. In: *Mechademia* 16/1, pp. 47–55.
- Tatlow, Stephen (2020): Everyone in Space Wants to Hear You Scream: Toward a Framework for Understanding Player Voice in Virtual Worlds. In: *Journal of Sound and Music in Games* 1/3, pp. 15–34.
- Thomas, Miles (2020): VTuber Gawr Gura Hits Milestone of 1 Million Subscribers on YouTube: Chumbuds are 1 million strong today. In: *Crunchyroll*, 22 October, www.crunchyroll.com/news/latest/2020/10/22/vtuber-gawr-gura-hits-milestone-of-1-million-subscribers-on-youtube (04.05.2025).
- Williams, Mike (2022): Voice Actor Plays 194 Hours of FFXIV Just To Hear His Friend Again: Poet Robert Harper dives into Eorzea to hear the voice of friend Stephen Critchlow once again. In: *Fanbyte*, 7 September, www.fanbyte.com/ffxiv/news/voice-actor-plays-194-hours-of-ffxiv-just-to-hear-his-friend-again (03.05.2025).
- Žižek, Slavoj (2012): *Less Than Nothing: Hegel and the Shadow of Dialectical Materialism*. London: Verso.

Empfohlene Zitierweise

Greenfield-Casas, Stefan: Limelight of the Idols: Voice, Virtuality, VTubers. In: *Kieler Beiträge zur Filmmusikforschung* 19 (2025), S. 59–79, DOI: <https://doi.org/10.59056/kbzf.2025.19.p59-79>.

Kieler Beiträge zur Filmmusikforschung (ISSN 1866-4768)

Die Inhalte dieses Werks werden unter der Lizenz CC BY 4.0 Creative Commons Namensnennung 4.0 zur Verfügung gestellt. Hiervon ausgenommen ist das Bildmaterial, das abweichenden, in den Bildlegenden spezifizierten Bestimmungen unterliegt.