Artificial Intelligence & Popular Music:
SKYGGE, Flow Machines, and the Audio Uncanny Valley

Melissa Avdeeff

Coventry University

ad1250@coventry.ac.uk

Abstract: This article presents an overview of the first AI-human collaborated album, *Hello World*, by SKYGGE, which utilizes Sony's Flow Machines technologies. This case study is situated within an overview of current and emerging uses of AI in popular music production, and connects those uses with myths and fears that have circulated in discourses concerning the use of AI in general, but also in music specifically. By proposing the concept of an audio uncanny valley in relation to AIPM (artificial intelligence popular music), this article offers a lens through which to examine the more novel and unusual melodies and harmonisations made possible through AI music generation, and questions how this content relates to wider speculations about posthumanism, sincerity, and authenticity in both popular music, and more wider assumptions of anthropocentric creativity. In its documentation of the emergence of a new era of popular music, the AI era, this article surveys: (1) the current landscape of artificial intelligence popular music focusing on the use of Markov models for generative purposes; (2) posthumanist creativity and the potential for an audio uncanny valley; and (3) issues of perceived authenticity in the technologically mediated "voice".

Keywords: artificial intelligence, popular music, creativity, uncanny valley, SKYGGE

Artificial Intelligence & Popular Music:
SKYGGE, Flow Machines, and the Audio Uncanny Valley

Melissa Avdeeff, Coventry University

"It is often said that television has altered our world. In the same way, people speak of a new world, a new society, a new phase of history, being created – 'brought about' – by this or that new technology: the steam engine, the automobile, the atomic bomb. Most of us know what is generally implied when such things are said. But this may be the central difficulty: that we have gotten so used to statements of this general kind, in our most ordinary discussions, that we can fail to realize their specific meanings" – Raymond Williams [1].

***

To echo the words of Raymond Williams, and to begin so boldly: we are on the edge of a new phase of popular music production, and by extension, a new phase of music consumption made possible by artificial intelligence. With the proliferation of AI and its developing practices in computational creativity, the debates surrounding "technology as opposed to art"[2] have not ceased; rather, they have extended into all aspects of creative expression. The questions surrounding the aura in the age of mechanical reproduction [3] have become heightened, authorial meaning convoluted by the creation of content based on various AI and algorithmic production techniques. The longstanding practice of defining creativity as something innately human is challenged by new technologies and ways of being with both technology and creative expression. While the questions over intellectual ownership, authorship, and anthropocentric notions of creativity run throughout the myriad forms of creative expression that have been impacted by artificial intelligence, in this article I explore the new human-computer collaborations in popular music, made possible by recent developments in the use of Markov models and other machine learning strategies for music composition and production.

There is a much longer history of artificial intelligence music in the art music tradition [4-6], but its use in popular music is currently predominantly that of novelty, experimentation, and largely as a tool for collaboration. The purpose of this article, therefore, is to document the present practice and understanding of the use of artificial intelligence in popular music, focusing on the first human-AI collaborated album, *Hello World*, by SKYGGE, utilizing the Flow Machines technologies and led by François Pachet and Benoît Carré. Using SKYGGE as a case study creates space to discuss some of the current perceptions, fears, and benefits of AI pop music, and builds on the theoretical discussions surrounding AIM (artificial intelligence music) that have occurred with the more Avant Garde uses of these technologies. As such, in this article I survey (1) the

current landscape of artificial intelligence popular music focusing on the use of Markov models for generative purposes; (2) posthumanist creativity and the potential for an audio uncanny valley; and (3) issues of perceived authenticity in the technologically mediated "voice". These themes are explored through a case study of two tracks from *Hello World*: "In the House of Poetry" featuring Kyrie Kristmanson and "Magic Man".

Considering the relative newness of AIPM (artificial intelligence pop music), there is limited research on the subject, which this article addresses. By building on work from the AIM tradition, AIPM marks a separation of not only technique and audience reception/engagement, but also concept. Whereas AIM has largely been concerned with pushing the boundaries of *possibility* in composition, AIPM is currently being used more to speed up the production process, challenge expectations of creative expression, and will inevitably mark a key shift in music production eras: analog, electric, digital, and AI. In many ways, these differences run parallel to other distinctions that arguably still exist between the art and popular worlds: academic/mainstream, art/commerce. The use of these technologies creates shared practice in the surrounding discourse, but their use also marks differences in approaches to the creative process.

AIPM has not had its "breakthrough" moment yet, and it's likely that it never occurs in such a momentous fashion as Auto-Tune's re-branding as an instrument of voice modulation in Cher's "Believe" (1996), subsequent shift into hip hop via T-Pain, and "legitimization" through other artists such as Kanye West and Bon Iver [7-8]. These key moments in popular music history have become mythologized, marking the rise of Auto-Tune as not only a shift in production method, but also our relationship to, and understanding of, the human voice as authorial expression, sparking some of the first debates about posthumanistic music production and cyborg theory in musicology [9-11].

In 2019 Auto-Tune is mundane[1]. Following the cycle of most new technologies – the synthesizer, DAWs, electronic/digital drum machines – Auto-Tune had its moment of disruption, its upheaval period before settling into normative practice [12]. It is yet to be seen how AIPM also navigates this process and how, like those that came before it, the technology evolves to incorporate unintended uses within the music industry. It is those moments of unintentionality that are often the most captivating and push the industry forward, not unlike Auto-Tune and its now ubiquitous presence as a vocal manipulator, but of course the most well-cited case, that of hip hop and sampling as an unintended use of turntables. It is often these initial moments of novelty that spurn new genres, practices, and trajectories in popular music. This article therefore also

---

[1] The "death of auto-tune" has come and gone as Jay-Z embraces the technology on "Apesh*t" (2018).

documents these initial moments of AIPM, before it inevitably reaches the next stage of use, and extends into unanticipated uses.

CREATIVE AI: SHIFTING PRAXIS
An in-depth overview of the varied production techniques of AIPM is outside the scope of this article, but a certain degree of historical and practical summary is contextually necessary. I want to suggest that AIPM marks a distinct shift in the continued development of popular music production, one that will certainly spark new discourses regarding posthumanist creativity, but in order to do so, it must be clear what these production methods actually entail.

A lot of misperception surrounds the use of AI in the production of artistic outputs, music included. Many of these misperceptions in the public arena extend from generalized fears and myths of artificial intelligence, and the role of supercomputers in a speculative near future. As Simone Natale and Andrea Ballatore have found, in their discursive analysis of AI in scientific trade magazines, many of these myths and fears solidified during the rise of AI in the 1950s – 70s, and continue to inform public opinion and perception of the role and potential of AI technologies. These fears, centred on depicting new computing processes as "thinking machines," foster scepticism and criticism of AI capacities, as spread and perpetuated through narrative tropes. Generally speaking, the AI technology myths tend to humanize the technology, often connecting AI to ideas of superhuman or supernatural powers [13]. Terms such as "thinking" and "intelligence" imply a human universality in the definition of consciousness, and AI technologies are thus perceived to occupy similar forms of intelligence, and by proxy rational thought. These apprehensions about AI music are often bound to the uncanniness of non-human creative production.

Defining human and computational intelligence is further complicated within creative endeavours. Arguably, the anthropocentricism of creativity runs deeper than notions of human intelligence, especially when taking into account the supposed exceptionality of music. Computational creativity not only challenges the idea of human-defined creativity, it further breaks down long-held distinctions between amateur and professional musicians. Building on the era of the "bedroom producer," AI music production may push the industry into novel capacities, "amplifying the capabilities of today's artists and elevating amateurs to the level of seasoned pros" [14]. Sandra Upson warns that the potential to deceive may also become amplified, intensifying current questions about the *authenticity* of computer-altered outputs (like the voice and Auto-Tune, or bodies and Photoshop), towards questions about the role of human in the output, if any [14].

Many of the technological myths or fears surrounding computational creativity, and AIM and AIPM in particular, also tend to be encapsulated in issues of labour and the apparent value of *highly visible* forms of labour. The same discourses accompanied various electronic and digital technologies at their inception, notably the drum machine, as discussed above, and the DJ when the turntable was largely superseded by DAWs. This connection to manual forms of labour at the point of performance is highlighted by the inclination of many Top-40 DJs to "look busy" while playing pre-recorded sets. This is not to say that live mixing is not common, but that the desire to demonstrate one's productivity is linked to the perceived value of that performance by the audience [15]. It is yet to be seen how live AI music production might challenge these existing notions of labour: how might AI production be made visible? Or, will the visibility of labour in music performance be subsumed into wider neoliberal understandings of the invisibility of labour in digital, and online, production?

Changes in the perception of labour have shifted within postdigital culture [16]. Contemporary youth, those who are highly involved in online production and have grown up in a post-social media environment, are more likely to recognize the labour involved in digital outputs; for example, the curated social media accounts, the online branded aesthetics, the YouTube influencers, and, to extend this further, new celebrity models made possible by postdigitality such as the Kardashians. The invisible, and often gendered [17-18], labour of online media may be extended to the function of AI in artistic productions, but this is yet to be confirmed.

By challenging established hierarchies, and perceptions of labour, postdigital production and AI collaboration connect the "supernatural" and "magic" connotations of *intelligent* technology to neoliberal notions of labour and production. The lack of generalized public understanding of how AI works lends it to be perceived as a form of "magic," or Black Box, and something that could transition into a runaway even with implications beyond human control. As will be discussed, as the boundaries between human and technology become increasingly negotiated, the fears that the intelligence made possible through AI could surpass human intelligence are intensified. It is only once their use becomes available to the public en masse that they will shift into the mundane, as another tool for creative expression.

As the accessibility of AIPM software becomes increasingly available for public use, it will most likely become another platform that encourages "everyday creativity." Apps, such as TikTok, Instagram, Snapchat, and so on, increase the invisibility[2] of digital production, by affording a form of everyday creativity whereby users are involved in a cycle of consuming, producing, and

---

[2] Prior to the rise of the term postdigital, and during the evolution of the internet as exceptional to indispensable [19], I had described the internet, and corresponding technologies, as invisible [20]. This invisibility is now "observable" in the ways in which youth use participatory digital media.

remixing. These technologies are expanding the creative process, and the concept of the prosumer, to something that is central to everyday life, especially in youth practices. Like the proliferation of DAWs, it is anticipated that the widespread proliferation of AIPM production methods will encourage more forms of "everyday music" production, making it easier for the general user to enter into the arena of music making. And, like the technologies that have come before, further challenge amateur/professional distinctions, while also encouraging new forms of innovation and multi-modal creativity.

AIPM: ARTIFICIAL INTELLIGENCE POPULAR MUSIC

The history of AIPM extends from the history of AIM, overviews of that history have been addressed elsewhere [21]. While SKYGGE's *Hello World* is documented as the first full album to be a true collaboration between AI and human production techniques [22], there are prior instances in popular music history that anticipated this album. For example, David Bowie's use of the Verbasizer on *Outside* (1995). Although not a tool for music production, the Verbasizer is a text ramdomizer program that automates the literary cut-up technique, in order to alter the meaning of pre-written text through random juxtapositions of materials [23]. In addition, since 2014, Logic has had the option to auto-populate drum tracks for users, and more recently they have improved upon this feature by adding quantization that mimics a more "human" sense of timing based on minute variations vs. strict adherences to tempo. And most notably, Amper, the "world's first AI music composer," released its beta software in 2014 and functions through human and AI collaboration to create new music based on mood and style. At the time of writing, the beta platform has been taken offline and the enterprise version, Amper Score, is soon to be released. However, the main market for Amper Score are companies/individuals who require royalty-free music to accompany branded content, such as podcasts, promotional videos, and other content that would normally use copyrighted stock music. Similar capabilities can be seen in Jukedeck, also founded in 2012 Current uses of Amper and Jukedeck are to speed up the process of creating royalty-free stock music, or even Muzak, and they are yet to make an impact on the production of mainstream popular music. They are generally used to produce music that is meant to accompany other audio-visual content, and therefore does not need to be highly emotionally engaging, or even, arguably, incredibly musically interesting. That being said, if the added benefit of using AI to produce background content ultimately results in the increased quality of said music, I see no drawbacks.

Schedl, Yang, and Herrera-Boyer define the canon of music systems and applications that are utilized to automate certain music production or music selection process for online stores and streaming services as intelligent technologies, while also acknowledging the problems in defining something as intelligent [5]. As there is no definitive definition for what intelligence entails, the word is used more often as a marketing technique than a description of what a

product/software/platform can do, functionally speaking. As they note, "musical intelligence could either be a capability to build interconnections among the different layers of the music work…to generate sequences that other systems consider as 'music,' or to anticipate musical features or events while listening to a musical performance, or all of them at the same time" [5]. They therefore build a typology of examples of intelligent music systems, which includes: systems for automatic music composition and creation, such as CHORAL, Iamus, Coninuator, OMax; Systems to predict music listening, such as Just-For-Me, and Mobile Music Genius; systems for music discovers, such as Musicream, and algorithms to curate music based on mood/emotions. Further tools used in current AI assisted popular music production include: IBM Watson, Magenta, NSynth, AIVA, and Sony Flow Machines.

AI music production has taken a variety of forms, from algorithmic to the use of neural networks and Markov chains. Fundamental to all methods, though, is the use of statistical analysis, or a set of rules and probabilities, which are then use to predict and create new materials based on what has been created before. Algorithmic music, which has a much longer history and use in Western Art Music, is less relevant to a discussion of AIPM, but it should be noted that there has long been an intimate connection between music and algorithms, as many musical formats are also bound to particular sets of rules, or algorithms. Rules-based compositions, such as Theme and Variations, or 12-bar blues, can be defined as algorithmic musics, because of their dependence on specific rules as the basis for composition, although aleotoric, or chance music, has more commonly been associated with algorithmic production.

Markov chains and other forms of machine learning composition rely on a corpus of existing music, scores, and/or audio, that are used as the basis for new compositions. Algorithmic compositions differ from machine learning composition (of which the use of Markov chains falls under) in that algorithmic music accomplishes a predefined goal, whereas generative modelling is not necessarily accomplished by means of a pre-determined algorithm. These terms necessarily overlap, but differ in their aims. Algorithmic music is music generated by algorithm; Markov chain produces musical elements based on an analysis of a defined corpus. According to Charles Ames, a Markov chain "models the behaviour of a sequence of *events*, each of which can assume any one of a fixed range of *states*. Changes in state between consecutive events commonly are referred to as *transitions*" (author's emphasis) [24]. It should be noted that, although there is an act of "creation" made possible through AI software that creation is nevertheless dependent on the music of human origin, which is used as the basis for analysis. Google DeepMind's explorations of iterations on AI iterations for visual content hint at a possibility of entirely AI-driven audiovisual content. It will be interesting to see how these processes may be taken up in music production, as well. I question whether iterations on an AI-collaborated musical corpus will

fail to engage human tastes in meaningful ways, and/or push human creativity and listening practices into new directions.

Amongst others, Curtis Roads traces the use of Markov chains for generative modelling of music to early instances in the 1950s – 1960s [25]. Most notably, Hiller's fourth movement of the Illiac Suite (1956). Many of these early uses of computers in music composition build on the use of indeterminacy in avant-garde music. Sandred et al. note that computers were able to speed up this process, functioning as a "tool that could take random decisions very quickly" [26]. While the first three movements of the *Illiac* suite were composed algorithmically, based on traditional rules of tonal harmony, the fourth movement was based on probability tables, or Markov chains. They utilized the ILLIAC computer to produce intervals for each instruments, constrained by the "rules" or expectations of tonal harmony [24]. Following the first performance of the suite in 1956, news sources derided the piece, claiming the audience was "resentful" of being made to engage with an "electronic brain," with one listener warning that it "presaged a future devoid of human creativity" [27]. Hiller, in response, emphasized the conceptual nature of the composition, but not without comment on the human capacity for creativity and the socialized constraints of human tastes and perception; in his words: "These correspondences suggest that if the structure of a composition exceeds a certain degree of complexity, it may overstep the perceptual capacities of the human ear and mind" [27]. The discomfort felt by these audience members in experiencing something created by an "electronic brain" or "intellectual machine" was most likely triggered by a certain amount of fear surrounding the role of human in creative endeavours. The reactions align with some of the fears and myths surrounding other AI discourses at the time, during the rise of algorithmic and AI discourse in the mainstream conscious. The knowledge that this music was not created entirely by human minds seems to evoke a degree of unease, not unlike the response seen in the uncanny valley of robotics and computer generated human images. Even though the *Illiac* suite is not considered AIPM, the allusions to the uncanny, the unfamiliar, and the, at times, resentment of computer-assisted composition holds true across genre distinctions.

Some have noted that the *Illiac* suite lacks the "journey" necessary for emotional engagement – that there is no melodic drive towards a climax, or overarching sense of purpose. It is difficult to determine whether those critiques are filtered through the knowledge of computer involvement, or if the same could be said of other human-created atonal or aleatoric musical works. Because of the way in which Markov chains function, in that they make predictions from instance to instance, or transition to transition, without the need for extensive information of preceding data, the music created in the *Illiac* suite could be seen as a reflection of that form of creation – and thus lacking a "story," something more akin to a string of independent variables. The music in the fourth movement moves from thought to thought, but does its abstractness belie a sonic

narrative – and how necessary is that narrative to audience engagement? Avant-garde music continues to push the boundaries of conventional aesthetics, and, in many ways, the addition of the use of a computer is no different.

The inclusion of the computer, which ultimately operates in binary and therefore lacks a certain degree of nuance afforded to humans, makes tangible the processes that are considered more autonomic in humans. For example, similar questions are often considered in response to the advancement of self-driving cars; ethical questions concerning risk assessment, decision making, and the consequences of these decisions are being debated, not because humans have not been capable of causing harm with cars previously, but once these decisions must be programmed ahead of time, the metaphorical trolley problem becomes reality – and the outcomes become pre-meditated as opposed to instinctive acts in moments of crises. While the stakes are considerably lower in music composition, by drawing attention to the algorithmic or statistical process of music production, it lays bare the ways in which authors are not singular entities, or "genius figures," but synthesizers and translators of all that one has heard before, often filtered through collaborations, inspiration, and the like. AI collaboration takes Roland Barthes' theory of the death of the author to its logical conclusions through the act of deliberately identifying the corpus necessary for production thereby automating the process of musical socialization that requires years of experience in a human body.

By unmasking the influences necessary for the creative act, AIPM is not entirely different from the "human" creative process, technically speaking. Where it may differ is in the notion that human *experience* is translated into creativity and artistic outputs. Can the same be said of computational creativity? Are these more philosophical questions even necessary, or do they perpetuate the trope of the tortured genius, and previously explored connections between visible labour, in this case "paying one's dues," and perceptions of value? At the core, though, these dilemmas question what it means to *create*, and draw attention to the creative process, human or otherwise. AI software is not creating, in a holistic sense, but neither is it replicating. The act of synthesizing is probably the most apt term to apply to current forms of AIPM, as a form of incremental creativity. As a technological aid, most existing AIPM software can suggest melodies, orchestrations, instrumentations, based on constraints provided; it is a tool for music production, a collaborator, as opposed to an independent creator. As Benoît Carré has noted – it's like having another person in the studio to bounce ideas off [28]. Some of those ideas just happen to push our ears into unfamiliar sonic realms.

Current capabilities of AI and generative modelling techniques in popular music are limited to those of tools, collaborative tools to aid in the discovery/production of novel sounds, melodies, and harmonisations. The technology has not yet reached the point of holistic music composition

whereby a fully cohesive and engaging popular music song can be created without human involvement[3]. However, the problem-solving AIM technique has been developed into something more analogous to true human-computer collaboration.

FLOW MACHINES: BIRTH OF SKYGGE

In 2016, the Sony CSL Research Laboratory, where Flow Machines is located, was credited with creating the first complete AI popular music track, "Daddy's Car," inspired by the early music of The Beatles. The track, while interesting, did not spark much mainstream interest, beyond those interested in the process. The software used, Flow Machines, is also responsible for the SKYGGE *Hello World* album, a much more musically engaging work, and one meant to blur the boundaries of conceptual and commercialized pop music.

From the Flow Machines website:

> The main part of the Flow Machines project is Flow Machines Professional. Flow Machines Professional is an AI assisted music composing system. Using this system, creators can compose melody in many different styles which he wants to achieve, based on its own music rules created by various music analysis. Creators can generate melody, chord, and base by operating Flow Machines. Furthermore, their own ideas inspired by Flow Machines. From here, the process will be the same as the regular music production. Arrange them with DAW, put lyrics, recording, mixing, mastering etc… Flow Machines cannot create a song automatically by themselves. It is a tool for a creator to get inspiration and ideas to have their creativity greatly augmented [29].

The use of the term "augmented creativity" is intriguing, as it centres the idea of the software as collaborative tool, as opposed to entirely AI-created works.

Flow Machines is an evolving project, based on the initial research made possible through an ESRC grant to study potential uses of AI in music production. François Pachet, Pierre Roy, and their team have made incredible strides in the use of various AI systems, most notably Markov models, to create software for AI-human pop music collaborations. Before turning to an analysis of selected tracks from the *Hello World* album, I first briefly outline the evolution of Flow Machines, from The Continuator, a device to encourage "flow" in music creation, to Flow Composer, a constraints-based lead sheet generator, to Flow Machines and the first AI-human

---

[3] Although not discussed here, Taryn Southern's *I AM AI* (2018) album is notable for being the first to use Amper on a large-scale production. She learned on Amper, and then incorporated IBM Watson and Google Magenta, as well. Her album differs from AIVA and Flow Machines in that the latter developed software specifically for their music production whereas Southern drew on the existing technologies available, exploring the possibilities of cross-platform production and the varying levels of user engagement that each offer.

AIPM collaboration. While the evolution of these applications and techniques can be found in a series of publications by Pachet, Roy, and members of their team, a succinct summary of such development has not yet been available.

The basis for all three iterations of the technology involves a triangulation between an exploration into the use of machine learning in popular music production, Mihály Csíkszentmihályi's concept of *flow* [30-31], and the manipulation and development of individual style as an innovative marker of creativity. As such, The Continuator (circa 2004) was one of the first interactive intelligent music systems intended for mainstream users, including children. As Pachet notes, "Unfortunately, purely interactive systems are never intelligent, and conversely, intelligent music generators are never interactive. The Continuator is an attempt to combine both worlds: interactivity and intelligent generation of music material in a single environment" [32]. The device, essentially a synthesizer with MIDI input and output, continues to play after the user inputs (plays) a musical phrase. The Continuator outputs in the style of the last phrase inputted by the user, and thus can be used as a tool for music generation. For example, one could play a riff, and allow the Continuator to continue to play in that style, while one is then able to play with/alongside the new materials. It may help to develop improvisatory skills, or function like an intelligent looping system. Pachet's empirical tests found that children, especially, connected quite easily with the technology, often exhibiting "Aha" effects much quicker and more easily than adult users, and could more often be observed entering into flow states. The device demonstrates intriguing uses for live performance, but also great potential for music education through the potential for scaffolding complexity of tasks.

Flow Composer (circa 2014) functions as a "constraints-based lead sheet generation tool," [33] whereby the application serves three purposes: autonomous generation of lead sheets, harmonization by inferring chords based on a given melody, and interactive composition. Combining Markov modelling with constraints based on meter solves issues in previous Markov generative models in that it allowed for more control over the music generation process, and more aesthetically pleasing results. By creating models based on a series of constraints (meter, melody, etc), Flow Composer and Flow Harmonise are able to generate more successful outputs. Within the autonomous generation capacities, lead sheets are created based on the style of an indicated corpus by first training the Markov+Meter models for chord and melody generation, and applying parameters such as number of chord changes/notes [34].

Flow Machines (circa 2016 – present) are a series of "new generation authoring tools aiming to fosters users' creativity." Like Flow Composer, they are interactive programs that encourage manipulation and "play" with musical styles. Flow Machines builds on the Harmonizer and Composer, improving upon the generative models used, including text and audio generation.

Users are encouraged to "manipulate styles as computational objects,"[34] thereby making transparent both the building blocks of creativity, as well as the music composition process itself. The statistical properties of music are arranged into probability tables, as the basis for new generation within certain style parameters. As noted above, this is, for the most part, no different than human music generation, but in human composition these processes occur more often in a subconscious level of perception.

Most popular music is created in adherence, or at least in allusion to, certain constraints or structures, and creativity emerges based on the development of an individual or authorial *style* within those constraints. Flow Machines draws attention to the function of style in music generation. As Pachet, Ghedini, and Roy note, they "consider style as malleable texture that can be applied to a structure, defined by arbitrary constraints," and "applying style to well-chosen structure may lead to creative objects" [35]. By inserting *style* into Csíkszentmihályi's influential model of creativity, they argue that creativity doesn't occur in individual works of art, but across a series of applications of an individual's style. Style – as an extension of skills acquisition – and flow therefore function in tandem to demonstrate the historiometry of creativity, as opposed to its singularity. Just as this proposed model of creativity recognizes the development of style as central to creativity, the Flow Machines technologies explore the application of style to musical structures and constraints, as a production of creativity. This begs the question of whether *style* is something that emerges from lived experience, and therefore indicative of authorial expression, and/or skills development, and therefore a product of probability.

The addition of intelligent musical systems adds a posthuman dimension to creativity, which will be explored in future work. What should be reiterated, however, is that this posthuman creativity is not solely the product of computational models; there remains a human element, both in the corpus used, and the collaboration between human and computer. It's posthumanistic in its extension of what creativity entails, and the anthropocentric notions of creativity that endure. It challenges the notion that music and creativity are both exceptional and "human," and deconstructs traditional understandings of these concepts. As previously noted, prior instances of AI and AIPM performances have elicited mixed response, but always with a theme of apprehension and fears about the loss of humanity with the introduction of computers, under the assumption that extending notions of humanity into the posthuman is inherently negative. Using AI as a collaborative tool may help redefine models of creativity, and the expectations of familiarity in popular music production

Pachet and Roy's work addresses a lack of holistic approaches to intelligent music generation. Building on "Daddy's Car," *Hello World* represents the first viable application of the Flow Machines applications, and potentially commercially viable AIPM human-computer collaborative

album. Beyond its notoriety as the first full AIPM album, *Hello World* is distinct in its combination of novelty and unfamiliar sounds, within Top-40 pop music conventions/constraints. The following section examines two tracks from the album: "In the House of Poetry" and "Magic Man" as examples of the uncanny sonic valley and authenticity of *voice*.

AUDIO UNCANNEY VALLEY: NOVELTY & POSTHUMAN SINCERITY

Because of the way in which AIPM is generated, there is great potential for unfamiliar, novel, and at times uncanny, sounds to be created. Human ears, and aesthetic expectations do not influence AIPM generations, but rather they are formed from probabilities, code, and constraints. While it is the prerogative of the human collaborator to pick and choose which musical elements are ultimately utilized, and whether those elements should be manually manipulated, SKYGGE often dwells in that sonic unfamiliarity, using it to their aesthetic advantages. I propose using the *audio uncanny valley* as framework for discussing these moments.

Introduced by Masahiro Mori [36], the uncanny valley, as a theory concerning robots and virtual character design whereby a closeness to human likeness evokes a negative human reaction [37], is contested and lacks empirical validation. Arguments have been made suggesting that there is a generational element in regards to how people react to these uncanny humanlike depictions, marking a difference between those who grew up immersed in virtual characters in comparison to those who did not [38]. The same can, or will likely be, said about any notion of an audio uncanny valley. The validity of concept is of less importance than a method for describing the reaction that often occurs in response to AIM and AIPM whereby listeners express discomfort at the addition of a computer "voice," or the lack of "soul," or other anthropocentric elements of human creativity.

These reactions mirror those found by the AI myths and fears studies, other early AIM performances, and other outputs produced by Pachet and collaborating musicians. For example, in response to a 2018 concert of experiments in machine learning for music creation [39] amongst the positive reviews were those who expressed concerns about the new technologies and their effects. For example, one audience member remarked "the computer generated pieces 'miss' something – would we call this 'spirit,' emotion, or passion?", with another adding "I think the science is fascinating and it's important to explore and push boundaries, but I'm concerned about the cultural impact and the loss of human beauty and understanding of music" [39]. There is a lack of available reactions to the *Hello World* album to make similar comparisons, and further research should empirically explore responses to AIPM, without listeners having prior knowledge of the mode of production, as that knowledge filters one's expectations, often looking to "best" the computer in some way or another. Similar reactions have been found to other posthuman forms of musical performance, such as holographic performances of the deceased alongside

13

corporeal musicians, such as the 2012 Tupac Shakur Coachella Music Festival event. Ken McLeod notes that these holographic performers often "evoke a sense of Freud's notion of the uncanny in that they manifest a virtual co-presence with both a visual and an aural trace of a larger creative power" [40]. He finds that "audiences are paradoxically attracted to the familiarity of the hologram while simultaneously being repulsed by its seemingly artificial, trans- or post-human unfamiliarity" [40]. In both cases, a mind/body or human/machine duality is upheld as "normal," familiar, and desirable, not unlike Freud's original notion of the uncanny as an interrogation of the alive/dead binary, through one's reaction to seeing something alien in a familiar setting, such as a prosthetic limb [41]. The uncanny, when embodied in monstrous form, represents an "abomination who exists in liminal realm between the living and the dead, simultaneously provoking sympathy and disgust" [41]. The uncanny, in sonic form, blurs the boundary between human and machine production, at times provoking wider fears about the future of human-technology relationships.

Whether or not music can be said to have a "soul" or "spirit" that manifests through human creativity is ultimately beside the point, as audio uncanniness (as a model) can exist whether produced humanly or computationally. The concept of an uncanny audio valley was first suggested by Francis Ramsey to articulate responses to simulated sound, and the perceived "naturalness" of said sounds. The closer simulated sound comes to natural sound, the more human ears pick up on the subtle differences that mark it as *strange*, or *uncanny*. As Winifred Phillips echoes, "It sounds almost real...but something about it is strange. It's just wrong, it doesn't add up" [42]. Comparatively, Mark Grimshaw takes up the notion of an audio uncanny valley as a positive aim for certain formats, particularly when provoking fear in horror games [43]. He suggests that the defamiliarization that occurs through distortion of sound whereby it still retains elements of naturalness can be exploited to evoke desired emotions. This can be observed not only in video games, but also horror films, contextualising visual elements. Brenton, Gillies, Ballin, and Chatting note, however, that context and presence are important; the uncanny audio succeeds through the perception of co-presence. The uncanny valley may be "analogous to a strong sense of presence clashing with cues indicating falsehood" [41] in virtual environments; similarly, the audio uncanny valley in video games and horror films formats exploits those falsehoods to deliberately incite fear and dread. AIPM, however, extends this theory into the purely auditory: audio cues that lack "presence" in an embodied sense, where the unease emerges from unfamiliarity and unexpected compositional techniques. In some ways this harkens back to the unease experienced in the first instances of schizophonia with the introduction of recorded sound.

In a video describing her SKYGGE collaboration, Kyrie Kristmanson notes: "For me, it is like a folk song, but virtual, digital. It doesn't feel like a human creation. It is a machine folk song tradition.

I find that most interesting." And that "You can somehow feel the melody was not composed by a human" [44]. The idea that a machine could have a folk song tradition is fascinating, as it places the machine within the lineage of the corpus from which it generated new compositional materials, yet she finds it to be distinct from such traditions. The eeriness of the track is undeniable, but it also owes much to the otherworldly quality of Kristmasnson's singing voice, itself. In this track, the piano, strings, and some voice stems are generated by SKYGGE, all instruments are played by SKYGGE (except some of Kirstmanson's vocals), and is generated based on a corpus of folk ballads and jazz tunes [45].
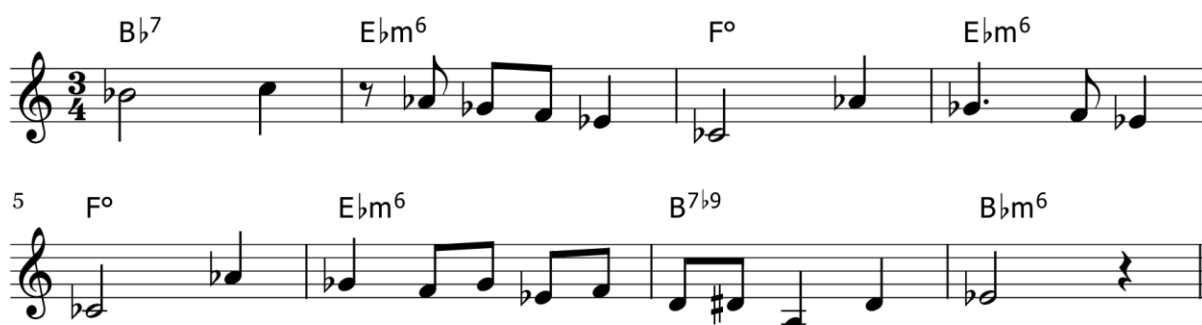

Figure 1. Chorus: "In the House of Poetry" by SKYGGE and Flow Machines, featuring Kyrie Kristmanson [46].

"In the House of Poetry" is split into two sections – verse and chorus – with the chorus evoking the "enchanting charm of ancient folks melodies" and the verse aligning more with jazz conventions. The *Hello World* website notes that Flow Machines (SKYGGE) proposed an "unconventional and audacious harmonic modulation" in the chorus within "an ascending melody illuminating the song" [45]. The harmonic modulation is definitely unique within Top-40 standards, but notably, so is the melody (see Figure 1). It is uncommon for mainstream pop tracks to have such a wide range in the melodic contour, adding to its aesthetic of unfamiliarity. What is not immediately obvious to the ear, is that in the second section of the track, the vocal line is generated by SKYGGE from recordings of Kristmanson's voice [45]. For myself, these moments are the ones that are most likely to cause unease (or even excitement), due to the knowledge of their production. That vocals can be generated by SKYGGE that are largely indistinguishable from Kristmanson's recorded vocals is remarkable, and hints at future possibilities of generating vocals that are outside the bounds of human possibility, yet sound almost "natural" enough to reach the audio uncanny valley state. Currently, this can be accomplished to some degree by synthesizing the human voice, but the added dimension of autonomous generation may further push these aesthetics into the unfamiliar and uncanny. Through Auto-Tune, and other voice manipulation software, listeners have become familiarized to posthuman voices.

Similarly, "Magic Man" is generated based on a corpus of 1980s French pop music and many of the uncanny elements come about through the nonsensical lyrics generated by SKYGGE whereby the generated vocal line comes close to, but is not quite English. Benoît Carré describes the production process:

> The melody and the chords were generated from a French pop song. The melody came very fast. Then, I chose to make Flow-Machines generate a folk guitar and many voices. Flow-Machines adapted the folk guitar track to the chords of my new song and the voices were adapted to the melody. It can be like grafting the coat of a bear on the back of an elephant – but sometimes it gives very nice results! All the generated voices gave a very nice disco mood, using the recurrent word, "Magic Man". I kept these vocals as the machine generated them, and then added real vocals to get a more precise sound [47].

Of the lyrics, Carré notes, "The machine generated all the unusual and strange lyrics. Listening, it felt a bit like going back to childhood, when you don't understand all the lyrics of the songs you listen to" [47]. The fact that SKYGGE generates a vocal line approximating human language is quite uncanny/exciting. The repetition of the phrase "Magic Man" draws attention to the lyrics at each repetition, while the remainder of the syllables fade into more subconscious levels of listening. It is easy to casually listen to this track and not immediately realize that the majority of the lyrics are not, in fact, "real". Just as mainstream music listeners have become more accustomed to posthuman voices, these generated lyrics follow a similar trajectory in regards to the increased familiarity that Western audiences have with other-than-English lyrics. For example, the current rise in popularity of K-Pop in the West has broken through previously held barriers to entry based on language. It is therefore not difficult to conceive of a commercially successful track based on random syllabic generation that approximates the English language.

We know someone through their voice – the connection between voice and embodiment is well established – yet "Magic Man" disrupts that connection. It's a voice without an origin, a language of only signifiers. I question how that might impact the reception of *voice* and the perceived authenticity or sincerity of autonomously generated voices. A new form of posthuman sincerity? Scruton suggests that a person demonstrates their understanding of language through its use, in the same way that a musician shows their understanding of the semantics of music by playing [48]. In "Magic Man," SKYGGE's lyrics come close enough to human language that, on first listen, the differences are not immediately clear. The uncanniness emerges in repeat listens, as the differences between English and Not-Quite-English become highlighted, and one's mind wanders to the further potential of autonomous language generation.

As noted, familiarity with disembodied voices is not new, as they occur not only in recorded sound, but also loudspeakers, phones, radios, and so on. In addition, a disembodied voice alludes to the supernatural or mystical, connecting back to those myths and fears of "thinking" machines.

This can be taken together with the fact that in pop music, a completely unaltered "natural" voice is quite rare. In "Magic Man" we have a disembodied voice, an uncanny simulation of voice and language, and perhaps a form of sincerity that emerged in the confusion between disembodied voices as supernatural/authoritative.

CONCLUSION

To conclude with a bit of futurology, I anticipate that in the coming months and years, the use of AI in music production will become mainstream and routine. The ways in which these technologies can speed up production, mixing, and even composing, make them invaluable to the production of music, and especially pop music with its culture of ephemera. They can help push human creativity into novel and at times uncanny audio spheres, and this begs the question of how, or if, tastes will adapt. As a tool for music production, AI will no doubt be commonplace soon, but as a *creator* of pop music melodies, structures, and instrumentations, it is less clear how audiences will approach that content if it sounds too far removed from expectation.

In the documented myths and fears surrounding the uncanniness of AI-produced creative content, for many, the idea that a computer can create music challenges beliefs in the exceptionalism of human creativity, and the *authenticity* of music composition. The discourses surrounding these issues reflect ongoing discussions about the perceived value of single-authored content, and the patriarchal understandings of the "genius figure." In making the composition process more transparent through a explicit corpus of "inspiration," it draws attention to the fact that all musicians draw inspiration from others, and that creativity is a product of all that one has encountered before, most often in collaboration with the experiences and processes of others. While the addition of AI makes that process more explicit, it could be argued that posthuman creativity has long existed, by means of challenging the very existence of authorial expression tied to the individual.

I am most excited for what becomes of the unexpected uses of AI content production. I anticipate their incorporation into mobile phone apps, whereby one can easily create custom music to accompany videos, photos, etc. Or infinite "chill" playlists in that are entirely AI produced and platform-owned. The future of creativity and humans' role in that future is speculative at best. Regardless of the "intelligence" of AI, it is principally human-driven and consumed, and, as such, it will be human agents who ultimately guide its use and progress. SKYGGE's *Hello World* is a product of these new forms of production and consumption, and functions as a pivot moment in the understanding and value of human-computer collaborations. The album is aptly named, as it alludes to the first words any new programmer uses when learning to code, as well as serving as an introduction to new AI-human collaborative practices. Hello, World, welcome to the new era of popular music.

REFERENCES

1. Raymond Williams, *Television: Technology and Cultural Form* (New York: Schocken, 1974).
2. Simon Frith. "Art versus Technology: The Strange Case of Popular Music." *Media, Culture and Society* 8 (1986): 263 – 279.
3. Walter Benjamin. "The Work of Art in the Age of Mechanical Reproduction," in *Illuminations*, ed. Hannah Arendt (New York: Schocken Books, 1969): 1 – 26.
4. Eduardo Miranda, ed., *Readings in Music and Artificial Intelligence* (Amsterdam: Harwood Academic, 2000);
5. Markus Schedl, Yi-Hsuan Yang, and Perfecto Herrera-Boyer. "Introduction to Intelligent Music Systems and Applications." *ACM Transactions on Intelligent Systems and Technology* 8, no.2 (2016): 17.1 – 8.;
6. C. Roads. "Artificial Intelligence and Music." *Computer Music Journal* 4, no.2 (1980): 13 – 25.
7. Simon Reynolds, "How Auto-Tune Revolutionized the Sound of Popular Music," *Pitchfork*, September 17, 2018, https://pitchfork.com/features/article/how-auto-tune-revolutionized-the-sound-of-popular-music/;
8. Yanto Browning, "Auto-Tune, and Why We Shouldn't Be Surprised Britney Can't Sing," *The Conversation*, July 17, 2014, https://theconversation.com/auto-tune-and-why-we-shouldnt-be-surprised-britney-cant-sing-29167
9. Keren Omry. "Bodies and Digital Discontinuities: Posthumanism, Fractals, and Popular Music in the Digital Age." *Science Fiction Studies* 43, no.1, (2016): 104 – 122.;
10. Joseph Auner. "'Sing It For Me': Posthuman Ventriloquism in Recent Popular Music." *Journal of Royal Musical Association* 128, no.1 (2003): 98 – 122.;
11. Robin James. "'Robo-Diva R&B': Aesthetics, Politics, and Black Female Robots in Contemporary Popular Music." *Journal of Popular Music Studies*, 20, no.4 (2008): 402 – 423.
12. Marshall McLuhan, *Understanding Media* (New York: Signet, 1966).
13. Simone Natale and Andrea Ballatore. "Imagine the Thinking Machine: Technological Myths and the Rise of Artificial Intelligence." *Convergence* 23 (2017): 1 – 16.
14. Sandra Upson, "Artificial Intelligence is Killing the Uncanny Valley and Our Grasp on Reality," *Wired*, December 16, 2017, https://www.wired.com/story/future-of-artificial-intelligence-2018/.
15. Ed Montano. "'How Do You Know He's Not Playing Pac-Man While He's Supposed to be DJing?': Technology, Formats and the Digital Future of DJ Culture." *Popular Music* 29, no.3, (2010): 397 – 416.

16. Christian Fuchs. "Digital Prosumption Labour on Social Media in the Context of the Capitalist Regime of Time." *Time & Society* 23, no.1 (2014): 97 – 123.

17. Kylie Jarrett. "The Relevance of 'Women's Work': Social Reproduction and Immaterial Labor in Digital Media." *Television & New Media* 15, no.1 (2014): 14 – 29.;

18. Brooke Duffy. "Gendering the Labor of Social Media Production." *Feminist Media Studies* 14, no.4 (2015): 710 – 714.

19. Donna Hoffman, Thomas Novak, and Alladi Venkatesh. "Has the Internet Become Indispensable?" *Communications of the ACM* 47, no.7 (2004): 37 – 42.

20. [Redacted for anonymity]

21. Eduardo Miranda and John Biles, eds., *Evolutionary Computer Music* (London: Springer, 2007).

22. Melissa Avdeeff, "AI's First Pop Album Ushers in a New Musical Era," *The Conversation*, October 2, 2018, https://theconversation.com/ais-first-pop-album-ushers-in-a-new-musical-era-100876

23. Matthew Braga, "The Verbasizer was David Bowie's 1995 Lyric-Writing Music App," *Motherboard*, January 11, 2016, https://motherboard.vice.com/en_us/article/xygxpn/the-verbasizer-was-david-bowies-1995-lyric-writing-mac-app

24. Charles Ames. "The Markov Process as a Compositional Model: A Survey and Tutorial." *Leonardo* 22, no.2 (1989): 175.

25. Curtis Roads. "Research in Music and Artificial Intelligence." *Computing Surveys* 17, no.2 (1985): 163 – 190.

26. Örjan Sandred, Mikael Laurson, and Mika Kuuskankare, "Revisiting the Illiac Suite: A Rule-Based Approach to Stochastic Processes," *Sonic Ideas/Ideas Sonicas* 2 (2009): 42 – 46.

27. Tiffany Funk, "A Musical Suite Composed by an Electronic Brain: Reexamining the *Illiac Suite* and the Legace of Lejaren A. Hiller Jr.," *Leonardo Music Journal* 28 (2018): 19 – 24.

28. Alex Marshall. "Is this the world's first good robot album?" *BBC Culture*, 21 Jan 2018. http://www.bbc.com/culture/story/20180112-is-this-the-worlds-first-good-robot-album

29. "Flow Machine – AI music-making," *Flow Machines*, https://www.flow-machines.com/

30. Mihály Csíkszentmihályi, *Flow and the Psychology of Discovery and Invention* (New York: Harper Collins, 1996);

31. Mihály Csíkszentmihályi, *Flow: The Psychology of Optimal Experience* (New York: Harper Perennial, 2008).

32. François Pachet. "On the Design of a Musical Flow Machine," in *A Learning Zone of One's Own*, ed. Mario Tokoro & Luc Steels, (Amsterdam: IOS Press, 2004), 3.

33. Alexandre Papadopoulos, Pierre Roy, and François Pachet, "Assisted Lead Sheet Composition Using FlowComposer," in *International Conference on Principles and Practice of Constraint Programming*, ed. M. Rueher (Springer, 2016).

34. Pachet, François, Pierre Roy and Fiammetta Ghedini. "Creativity through Style Manipulation: the Flow Machines project." (2013).

35. Fiammetta Ghedini, François Pachet, and Pierre Roy, "Creating Music and Texts with Flow Machines," in *Multidisciplinary Contributions to the Science of Creative Thinking*, ed. Giovanni Corazza and Sergio Agnoli, (Springer: 2015), 334.

36. Masahiro Mori, "The Uncanny Valley," *Energy*, 7, no.4 (1970): 33 – 35.

37. Edward Schneider, Yifan Wang, and Shanshan Yang, "Exploring the Uncanny Valley with Japanese Video Game Characters," in *Situated Play – Proceedings of DiGRA 2007 Conference* (2007).

38. Annalee Newitz, "Is the 'Uncanny Valley' a Myth," *iO9*, March 9, 2013, https://io9.gizmodo.com/is-the-uncanny-valley-a-myth-1239355550

39. Bob Sturm, Oded Ben-Tal, Úna Monaghan, Nick Collins, Dorien Herremans, Elaine Chew, Gaëtan Hadjeres, Emmanual Deruty and François Pachet, "Machine Learning Research that Matters for Music Creation: A Case Study," *Journal of New Music Research* 48, no.1 (2018): 36 – 55.

40. Ken McLeod, "Living in the Immaterial World: Holograms and Spirituality in Recent Popular Music," *Popular Music and Society*, 39, no.5 (2016): 510.

41. Harry Brenton, Marco Gillies, Davis Ballin, and David Chatting, "The Uncanny Valley: Does it Exist and is it Related to Presence?" *Presence-Connect*, 8, no.1 (2005).

42. Winifred Phillips, "Virtual Reality in the Uncanny Aural Valley," *Gamasutra*, April 8, 2015, http://www.gamasutra.com/blogs/WinifredPhillips/20150804/250439/Virtual_Reality_i n_the_Uncanny_Aural_Valley.php

43. Mark Grimshaw, "The Audio Uncanny Valley: Sound, Fear and the Horror Game," *Games Computing and Creative Technologies: Conference Papers* (2009).

44. SKYGGE MUSIC, "Kyrie Kristmanson Talking About Her Collaboration with SKYGGE," *YouTube*, December 7, 2017, https://www.youtube.com/watch?v=yxTF-UFvoHU

45. "Credits: Track by Track," *Hello World Album*, https://www.helloworldalbum.net/track-by-track/

46. SKYGGE MUSIC. "Flow Machines for 'Hello World Album' by SKYGGE," *YouTube* 1 Feb 2018. https://www.youtube.com/watch?v=jPp0jIJvDQs

47. Phillip Williams, "Can Artificial Intelligence Write a Great Pop Song?" *RedBull*, February 16, 2018, https://www.redbull.com/gb-en/SKYGGE-artificial-intelligence-making-pop-songs

48. Paul Barker, Christopher Newell, and George Newell. "Can a computer-generated voice be sincere? A case study combining music and synthetic speech." *Logopedics Phoniatrics Vocology*, 2013: 1 – 9.