



International Journal of Music

Journal homepage: <https://trpub.org/journal.php?id=MGxgAuFEdO>

TRPub

IJM

International Journal of
Music

Vol. 1 Issue 1 2026

Perspective

Open Access

Generative Audio Models Force Conservatories to Rethink What Composition Teaching Is For

Freya Venkatesan^{1*}, Manon Harutyunyan², Diego Mahlangu³

¹Department of Music, Yangtze University Centre for Music, Ireland

²Department of Music, Rhine Center for Music, Ghana

³Department of Music, Parana Graduate Institute of Music, Chile

Article Info

Volume 1, Issue 1, 2026

[doi: 10.62715/ijm.2026.v1.i1.pp25-29](https://doi.org/10.62715/ijm.2026.v1.i1.pp25-29)

Received: 02 October 2025

Accepted: 19 November 2025

Published: 25 January 2026

Abstract

Systems that generate finished audio from a line of text have moved from research papers to consumer products within a few years. They can produce a plausible song, cue or pastiche in seconds, with no notation and no performer. This Perspective argues that their arrival poses a question conservatories have been able to postpone: what is composition teaching for? Much of that teaching still rests on two assumptions, that producing stylistically competent music is hard and therefore worth certifying, and that a portfolio of finished pieces is reliable evidence of what a student can do. Generative audio weakens both. The article sets out the argument, then takes seriously three objections: that automated composition is old news, that the models cannot write for live performers or sustain long forms, and that traditional craft remains the only route to musical judgement. It concludes that craft exercises keep their value as training of the ear and should be defended on those grounds, while assessment should shift from the product to the reasons a student can give for it, and that work with performers, critical listening and an understanding of how the models are built and trained belong at the centre of the curriculum. The position is offered for debate and reports no empirical study.

Keywords: *Generative audio, Composition pedagogy, Conservatoires, Music and artificial intelligence, Assessment, Musical craft*

© 2026 Freya Venkatesan et al. This is an open access article under the CC BY license (<https://creativecommons.org/licenses/by/4.0/>), which permits unrestricted use, distribution, and reproduction in any medium, provided you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons license, and indicate if changes were made.

1. Introduction

A first-year composition student can now type a description of a piece, wait less than a minute, and play back a recording that many listeners would take for the work of a competent professional. The student has written no notes, hired no players and made no decisions about harmony or orchestration. Whatever one thinks of the result, the event alters the ground on which composition is taught, because the thing a conservatory has traditionally certified, the ability to produce idiomatic and well-made music, has become cheap to imitate.

This article takes the view that the right response is neither prohibition nor enthusiasm. Bans cannot be enforced outside the practice room, and the claim that the tools simply make everyone a composer

* Corresponding author: Freya Venkatesan, Department of Music, Yangtze University Centre for Music, Ireland. E-mail: freya.venkatesan@yangtze.edu

© 2026 Freya Venkatesan et al. This is an open access article distributed under the Creative Commons Attribution License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.

mistakes output for understanding. The more useful question is the one the technology forces: what composition teaching is meant to give a student that a generated file does not contain. Answering it requires conservatories to say aloud what has long been assumed.

2. What the Models Do

The line of development is short. WaveNet showed that a neural network could generate raw audio sample by sample (van den Oord et al., 2016). Jukebox applied related ideas to music with singing, conditioned on artist, genre and lyrics (Dhariwal et al., 2020). MusicLM and MusicGen then produced music from free text descriptions, with the second released openly for others to build on (Agostinelli et al., 2023; Copet et al., 2023). Commercial services built on similar principles now return whole songs with vocals to anyone who asks.

Two features distinguish these systems from earlier music software. They work in sound, not in symbols, so their output arrives performed, mixed and complete. And they are trained on very large collections of existing recordings, which means that what they return is a statistical rendering of how music of a requested kind has usually gone. They are at their most convincing where a style is well represented and conventions are strong, which is also where pastiche exercises and much commercial writing sit.

3. What Composition Teaching Has Assumed

Conservatory composition teaching combines two inheritances. One is craft passed on through graded exercises: species counterpoint after Fux, chorale harmonisation, fugue, orchestration from models (Fux, 1965). The other is the individual lesson, in which a student brings work in progress and a teacher responds. Kingsbury's ethnography of an American conservatory describes how such institutions treat talent as a property of persons that the teacher discerns and the institution ratifies (Kingsbury, 1988). Behind both stands what Goehr calls the work-concept, the idea that music exists as discrete, authored, finished works (Goehr, 1992).

Assessment follows from this. A student is judged on a portfolio of scores and recordings, and the portfolio is read as evidence of the skills that produced it. The inference from product to skill was always imperfect, since teachers edit, performers rescue and notation software flatters. It held well enough while making a polished piece demanded most of the abilities being certified. Table 1 sets out the main aims that composition teaching has pursued, what generative audio does to each, and the adjustment argued for in the remainder of this article.

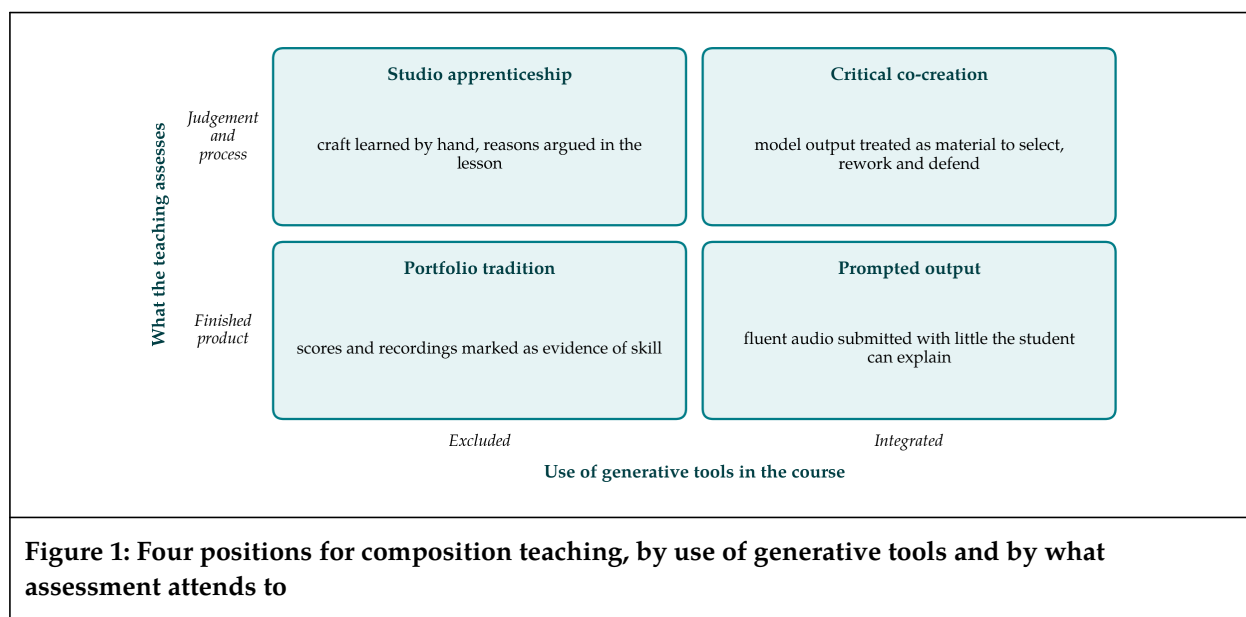
Table 1: Traditional aims of conservatory composition teaching, the effect of generative audio on each and a proposed reorientation		
Aim of teaching	What generative audio changes	Proposed reorientation
Stylistic fluency	Plausible pastiche is available on request	Teach pastiche as ear training, not as output
A portfolio of works	A finished recording no longer shows who decided what	Assess reasons, drafts and spoken defence
Technical craft	Surface polish is separated from understanding	Keep exercises; justify them as training of judgement
An individual voice	Models return the average of a style	Ask what a piece does that the average would not
Writing for performers	Largely untouched: models emit audio, not parts	Move collaboration with players to the centre
Professional readiness	Routine commissions face automated competition	Add data, rights and tool literacy to the course

4. The Argument: From Product to Judgement

If a stylistically competent product can be had without the skills, then the product can no longer be the main thing taught or the main thing marked. What remains scarce is judgement: hearing what a passage is doing, knowing why one continuation is better than another for this piece and these players, and being able to change it on purpose. Boden's account of creativity is helpful here. She separates exploring an established conceptual space from transforming it, and observes that recognising the value of a new idea is as much part of creativity as generating it (Boden, 2004). A model can generate without end. Selection and evaluation fall to whoever is listening.

Seen this way, the old exercises look different without ceasing to matter. Nobody needs a student's fourth-species counterpoint as music. Its use is that writing it slowly, by hand, against rules, builds an inner hearing of voice leading that later lets a composer notice what is wrong with a bar. Sennett describes craft in general as a dialogue between hand and head in which problems are found as well as solved (Sennett, 2008), and Schön describes professional knowledge as reflection carried on in the midst of action (Schön, 1983). Both accounts locate the value of practice in what it does to the practitioner. Conservatories should defend counterpoint and orchestration on exactly those grounds and stop implying that the exercises are valuable as artefacts.

Figure 1 maps the choice facing a department along two lines: whether generative tools are kept out of the course or brought into it, and whether assessment attends to the finished product or to the student's judgement and process. The lower right-hand corner, where tools are permitted and only the product is marked, is the position an institution reaches by default if it changes nothing. The upper row is defensible with or without the tools. The argument of this article is that the vertical move matters more than the horizontal one.



There is a second part to the argument. Small insisted that music is something people do together, and that its meanings lie in the relationships a performance sets up (Small, 1998). A composer who writes for a string quartet enters such relationships: with four players, their instruments, a rehearsal, a room. Generated audio bypasses all of them. The part of composition that consists in imagining what particular people can do, and then learning from what they do in fact, is the part the models leave alone, and it deserves more of the timetable than it usually receives.

5. Three Objections

The first objection is that none of this is new. Hiller and Isaacson programmed a computer to compose a string quartet in the 1950s (Hiller and Isaacson, 1959), and Cope's software wrote pastiche Bach and

Chopin that unsettled audiences decades ago (Cope, 2001). Recording, the synthesiser and the sequencer each prompted predictions that musicianship would wither, and Katz has shown how musicians absorbed and redirected such technologies (Katz, 2010). The precedent is real, but the earlier systems needed expert operators and produced notation or simple synthesis. The difference now is access and finish: any student can obtain complete audio, in any common style, without knowing anything.

The second objection is that the models are irrelevant to what conservatories teach. They do not write parts that a bassoonist can play, and their grasp of long-range form is weak. That is fair as a description of the present, and it is the reason this article treats writing for performers as the secure centre of the subject. It would be unwise to rest a curriculum on a technical limitation, however. Many graduates will also earn part of their living from media, library and commercial music, where delivered audio is the product and where the pressure will be felt first.

The third objection comes from the other side: that making room for the tools abandons craft and lets students skip the slow work on which judgement depends. The risk is genuine. Participants in a songwriting contest that required the use of machine learning reported that steering the models and sifting their output was laborious and that musical decisions still had to be made by people (Huang et al., 2020), which suggests that the tools reward those who already know what they want. The answer is sequencing. Students should meet generative systems after, and alongside, training that gives them something to judge with.

6. Implications for Curriculum and Assessment

Several practical changes follow. Assessment should include what a file cannot fake: sketches and drafts, an oral defence in which the student explains choices and responds to alternatives, and workshop sessions where a piece is revised in front of players. These practices already exist in many departments and need to carry more of the mark. Calls to reform music degrees around creativity and integration of skills predate the present technology and point the same way (Sarath et al., 2017).

Students also need to understand the tools as made things. That includes, at a non-specialist level, how a model is trained and why it returns the typical; whose recordings were used and on what terms, a matter on which copyright law remains unsettled (Sturm et al., 2019); and what the alternatives look like. Lewis's *Voyager*, an improvising program conceived as a partner with its own behaviour and not as a generator of products, shows that a musical machine can embody an aesthetic and a politics chosen by its maker (Lewis, 2000). A composer who can read systems in that way is better placed to use, refuse or redesign them.

Finally, institutions should say plainly why they teach composition. Three answers survive the arrival of generative audio: to train musical judgement, to prepare people to make music with and for others, and to sustain traditions of thought about sound that are worth having whether or not a market pays for them. Certifying the ability to turn out competent music in a given style is no longer among them.

7. Conclusion

Generative audio does not make composers redundant, and it does not leave composition teaching where it was. It removes the scarcity of the plausible product, on which a good deal of teaching and most assessment have quietly depended. What it cannot supply is a reason for one musical choice over another, or the experience of working something out with other musicians in a room. A conservatory that organises its teaching around those two things will have an answer when students ask why they should spend years learning what a machine appears to do in seconds. The answer is that they are learning to hear, to judge and to collaborate, and that the writing was always a means to that.

Conflict of interest

To be declared by the author(s) before publication.

Funding

To be declared by the author(s) before publication.

References

- Agostinelli, A., Denk, T.I., Borsos, Z., Engel, J., Verzetti, M., Caillon, A., Huang, Q., Jansen, A., Roberts, A., Tagliasacchi, M., Sharifi, M., Zeghidour, N. and Frank, C. (2023). [MusicLM: Generating music from text](#). arXiv preprint arXiv:2301.11325.
- Boden, M.A. (2004). [The Creative Mind: Myths and Mechanisms](#) (2nd ed.). Routledge, London.
- Cope, D. (2001). [Virtual Music: Computer Synthesis of Musical Style](#). MIT Press, Cambridge, MA.
- Copet, J., Kreuk, F., Gat, I., Remez, T., Kant, D., Synnaeve, G., Adi, Y. and Défossez, A. (2023). [Simple and controllable music generation](#). *Advances in Neural Information Processing Systems*, 36.
- Dhariwal, P., Jun, H., Payne, C., Kim, J.W., Radford, A. and Sutskever, I. (2020). [Jukebox: A generative model for music](#). arXiv preprint arXiv:2005.00341.
- Fux, J.J. (1965). [The Study of Counterpoint: From Johann Joseph Fux's Gradus ad Parnassum](#) (rev. ed., trans. and ed. Mann, A.). W.W. Norton, New York.
- Goehr, L. (1992). [The Imaginary Museum of Musical Works: An Essay in the Philosophy of Music](#). Clarendon Press, Oxford.
- Hiller, L.A. and Isaacson, L.M. (1959). [Experimental Music: Composition with an Electronic Computer](#). McGraw-Hill, New York.
- Huang, C.-Z.A., Koops, H.V., Newton-Rex, E., Dinculescu, M. and Cai, C.J. (2020). [AI Song Contest: Human-AI co-creation in songwriting](#). *Proceedings of the 21st International Society for Music Information Retrieval Conference*.
- Katz, M. (2010). [Capturing Sound: How Technology Has Changed Music](#) (rev. ed.). University of California Press, Berkeley.
- Kingsbury, H. (1988). [Music, Talent, and Performance: A Conservatory Cultural System](#). Temple University Press, Philadelphia.
- Lewis, G.E. (2000). [Too many notes: Computers, complexity and culture in Voyager](#). *Leonardo Music Journal*, 10: 33-39. <https://doi.org/10.1162/096112100570585>
- Sarath, E.W., Myers, D.E. and Campbell, P.S. (2017). [Redefining Music Studies in an Age of Change: Creativity, Diversity, and Integration](#). Routledge, New York.
- Schön, D.A. (1983). [The Reflective Practitioner: How Professionals Think in Action](#). Basic Books, New York.
- Sennett, R. (2008). [The Craftsman](#). Yale University Press, New Haven.
- Small, C. (1998). [Musicking: The Meanings of Performing and Listening](#). Wesleyan University Press, Hanover, NH.
- Sturm, B.L.T., Iglesias, M., Ben-Tal, O., Miron, M. and Gómez, E. (2019). [Artificial intelligence and music: Open questions of copyright law and engineering praxis](#). *Arts*, 8(3): 115. <https://doi.org/10.3390/arts8030115>
- van den Oord, A., Dieleman, S., Zen, H., Simonyan, K., Vinyals, O., Graves, A., Kalchbrenner, N., Senior, A. and Kavukcuoglu, K. (2016). [WaveNet: A generative model for raw audio](#). arXiv preprint arXiv:1609.03499.

Cite this article as: Venkatesan F., Harutyunyan M. and Mahlangu D. (2026). [Generative Audio Models Force Conservatories to Rethink What Composition Teaching Is For](#). *International Journal of Music*. 1(1), 25-29. doi: 10.62715/ijm.2026.v1.i1.pp25-29