

Did Melody Become a Schrödinger Cat?

Commentary on Clark & Arthur

KLAUS FRIELER[1]

Max Planck Institute for Empirical Aesthetics

ABSTRACT: I comment on Clark and Arthur's response to a YouTuber's claim of the death of melody for which they used corpus analysis and statistical methods of computational musicology. While I basically appreciate the effort, I also will discuss three pertinent problems I see ingrained here: whether such claims can be substantiated in any form in the first place, and how to react to dubious claims disseminated from YouTube musicology, and I will also shortly discuss some methodological issues.

Submitted 2023 July 9; accepted 2023 August 9.

Published 2023 November 9; <https://doi.org/10.18061/emr.v17i2.9548>

KEYWORDS: *corpus studies, dead melodies, quantum mechanics, cats*

THE basic intention of the present paper by Clark & Arthur is a very much honorable one. In 2019, “Inside the Score”, one of the many musicology YouTubers, published a video claiming the “death of melody”, while providing some anecdotal evidence for this rather bold statement [2]. The video explicitly disclaims that the diagnosed death of melody does not imply any aesthetical dismissal of modern pop music with its “dead melodies” (examples from Hans Zimmer to Billie Eilish are given), but one can suspect that this is only paying lip service. The lurid title of the video is probably implying otherwise as does the follow-up video, by the same author, entitled “What makes good melody?” [3] implying that “dead melodies” are, in fact, not “good melodies”. There is clearly a certain tinge of cultural pessimism to the whole affair. The video has as of now (2023) racked up 2,2 Mio clicks – an audience probably several orders of magnitude larger than the article by Clark & Arthur (or this comment, for that matter) will ever have. As the article by Clark & Arthur tries to somehow test the hypotheses of the “Death of Melody” using more substantial evidence in form of a corpus of 1.500 pop songs, it belongs to the genre of “fact checking”. which is very important, but, unfortunately, often also pointless, if it occurs after the fact once the cat is out of the box

The questions I like to address here are threefold. First, if the claim is well-defined, second, if the present paper does the job of testing the original claim, and, finally, if it should have been done in the first place.

SCHRÖDINGER MELODIES

Erwin Schrödinger presented his famous cat paradox in 1935 to highlight some conceptual problems with quantum mechanics [4]. It has become since then a widely known colloquial trope, a staple of pop and high culture alike [5]. It describes the undefined state of a cat in a closed box subjected to a devilish mechanism involving poison triggered by a random nuclear decay process. This setup traps the poor cat in a superposition state of being dead and alive at the same time, until somebody opens the box, in which instant the cat's quantum wave function will collapse and settle on a defined state – dead or alive. The same can be said of melodies, as long as nobody is looking, they are dead and alive at the same time.

According to Inside the Score, melodies are dead now; according to the present paper by Clark and Arthur, melodies are very much still alive. In contrast to Schrödinger's cat, the state of melody is not tied to some intrinsically random quantum process but it similarly hinges on a process of measurement employed – apart from the fact that none of the involved parties actually defines what a melody is in the first place or what the scope of their claim is. (Does a bass line or a rap verse count as melody? Are melodies in contemporary jazz also dead?) Clark and Arthur, at least, implicitly define melodies as the outcome of a certain algorithm, *Melodia* (Salamon, 2014), run on a collection of pop songs taken from the Billboard Hot 100. Both parties use a more or less well-defined set of features to measure the liveness of melodies, implicitly



in the case of *Inside the Score* and explicitly in the present paper. Both set of features seem to be very ad-hoc. In the case of the YouTube video, affairs are much worse as only a few examples are given to illustrate what is meant with “dead melodies.” Clark and Arthur are much clearer about their methods and rationales. Nevertheless, they use five different features without discussing the point, how a single universally valid measure of “aliveness” could be defined over a multidimensional feature space. A melody could become a Schrödinger cat: alive in regard to some aspects and dead in regard to others. Of course, I can imagine a psychological approach by asking humans to judge the liveliness of a melody (letting aside problems of enculturation and representativeness) with a subsequent modeling these judgments with a large set of melodic features.

To be fair, as Clark and Arthur mainly set out to debunk the notion of the death of melody, they only need to show that for a range of reasonable features no trend in the claimed direction can be found, and this is what they achieve ultimately. Only one of five features shows a trend compatible with a recent death of melody. Thus, the matter seems readily be settled, right?

WHEN IS MESSY TOO MESSY?

The method Clark and Arthur are forced to employ in order to achieve their goal, using a melody extraction algorithm to create a corpus of pop song melodies, is symptomatic and also somewhat problematic.

First of all, copyright will prohibit the existence of publicly available and historically relevant corpus for pop music for a foreseeable future. This forces researchers in the field of computational musicology or music information retrieval (MIR) to either use copyright-free melodies or to create their own private corpus. Streaming services like Spotify and Apple are of course here in a privileged position. The first solution is mostly not an option for musicological or historically-oriented work, such as the present study, for obvious reasons, as the historically relevant music is (still) copyrighted. The second solution is also a bit problematic, as the (former) state-of-art melody extraction algorithms produced rather unreliable results (Clark & Arthur report accuracies of about .16 to .21).

A “messy” approach, i.e., compensating very noisy measurements with very large sample sizes, might work in principle for a certain set of applications, but nobody has ever shown this to be the case. There are several theoretical and practical arguments that could be put forward here to cast doubt on the validity of this approach. At least, an estimation of the necessary samples size to compensate for noisy measurements would be nice to have. Additionally, a sensitivity analysis of the used features to transcription errors by way of simulation might give more (or less) confidence. At the moment, I am a bit in the dark here and only the final results, which seem to make sense, give some evidence, because they display some trends and difference, which is, however, a cyclic argument. As the time range of music used here spans 50 years, some trends might be just due to change in recording technology or other confounders.

Another issue is that the intricate interconnectedness of musical dimensions can have detrimental effects on derived features, e.g., pitch intervals. For instance, if every second pitch is correctly estimated by an algorithm, which is three times better than the *Melodia* results, than this still implies that on average nearly 63% of all semitone intervals are wrong. Measuring rhythmic features such as nPVI, which is based on duration ratios which are based on unreliably determined onsets and offsets might be plagued by similar problems.

But there is hope. Modern source separation and melody extraction algorithms based on deep learning are much better than the older algorithms. Still, error propagation might be a problem (cf. Frieler et al., 2019). Repeating the corpus creation with modern SOTA tools, e.g., a source separation tool such as Spleeter by Deezer [5] or Music Source Separation by ByteDance [6] followed by a melodic transcription tool such as pYIN (Mauch & Dixon, 2014, [7]) or CREPE (Kim et al, 2018, [8]), might be a fruitful starting point.

A CRUCIAL QUESTION

Disregarding the methodological intricacies, the main question is this: does this type of research – or the use of flashy claims like in the original video, on which the present study inevitably capitalizes on – do (computational) musicology a disservice? The world-wide headlines and the social media ripples stirred in the past by comparable studies, e.g., Mauch et al. (2015), in which grand claims were based on questionable methods (at least from an etic/emic perspective [9]), did put musicology and MIR on the public table, no

doubt. This might help in creating a brand and generating public interest for our discipline, and, hopefully, in the long run, also research funding. But it might also drive musicology dangerously close to being perceived as a purely entertaining game that only superficially still resembles science. And, in fact, there is genuine musicological interest in questions that these studies claim to tackle, e.g., tracking stylistic changes in context of larger social, historical, and technological developments, but these intricate matters require much more sensitive methods and genuinely interdisciplinary efforts spiced with a large portion of modesty – at least in my view.

However, to leave the field to YouTubers with their, at times, dubious claims (even if light-handedly employed to generate clicks and thus income), is also not an option [10]. In this sense the present article by Clark and Arthur is an honorable attempt in setting the record straight. The main problem is that the original claim was already dead on arrival, and could have been more appropriately answered by actually not answering it. Since, to paraphrase Wolfgang Pauli, another father of quantum mechanics, the claim of the “death of melody” is not only not correct, it is not even wrong.

ACKNOWLEDGEMENTS

This article was copyedited by Matthew Moore and layout edited by Diana Kayser.

NOTES

[1] Klaus Frieler, Max Planck Institute for Empirical Aesthetics, Grüneburgweg 14, 60322 Frankfurt/Main, Germany, Klaus.frieler@ae.mpg.de

[2] <https://www.youtube.com/watch?v=K0Vn9V-tRCo&t=0s> [24.05.2023]

[3] <https://www.youtube.com/watch?v=wQWgsr88GQI&t=330s> [24.05.2023]

[4] https://en.wikipedia.org/wiki/Schr%C3%B6dinger%27s_cat [24.05.023]

[5] https://en.wikipedia.org/wiki/Schr%C3%B6dinger%27s_cat_in_popular_culture [24.05.023]

[6] https://github.com/bytedance/music_source_separation [24.05.2023]

[7] <https://github.com/deezer/spleeter> [24.05.2023]

[8] <https://github.com/marl/crepe>

[9] With the emic-etic perspective, I mean here the “etic” (outsider’s) use of audio features to make claims about “emic” (inside) developments in the field of popular music, where the mapping between the two is basically unclear or fuzzy. The distinctions emic/etic stems from the difference between phonetics and phonemics. The first deals with acoustically measurable speech sounds and the latter with the meaningful distinctions between these sounds made by speakers of a language. The emics/etics topic lies at the heart of computational and corpus musicology, as the necessary simplifications and formalizations that have to be made in order to employ statistical methods are in conflict with richness of emic verbal descriptions used by practitioners and listeners as well as with researchers with a social science and humanities background. To use another quantum mechanical metaphor: there seem to be a Heisenberg uncertainty relationship between emics and etics, between specificity and universality, between reliability and validity, as one cannot have both at the same time without sacrificing some aspect of the other.

[10] It should be, however, noted that there are some YouTubers, such as Adam Neely [<https://www.youtube.com/c/adamneely>, 24.05.2023], who create entertaining and scientifically sound musicological content, and who provide likely an important service for the popularization of our field.

REFERENCES

Frieler, K., Başaran, D., Höger, F., Crayencour, H.-C., Peeters, G., Dixon, S. (2019) Don't hide in the frames: Note- and pattern-based evaluation of automated melody extraction algorithms. In: *Proceedings of the 6th International Conference on Digital Libraries for Musicology (DLfM '19)*, November 9, 2019, Den Haag, Niederlande. <https://doi.org/10.1145/3358664.3358672>

Mauch, M. & Dixon, S. (2014). pYIN: A fundamental frequency estimator using probabilistic threshold distributions. In: 2014 IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP), Florence, Italy, 659–663. <https://doi.org/10.1109/ICASSP.2014.6853678>

Mauch, M., MacCallum, R. M., Levy, M. & Leroi, A.M. (2015). The evolution of popular music: USA 1960–2010, *Royal Society Open Science*, 2(5). <https://doi.org/10.1098/rsos.150081>

Salamon, J., & Gomez, E. (2012). Melody Extraction From Polyphonic Music Signals Using Pitch Contour Characteristics. *IEEE Transactions on Audio, Speech, and Language Processing*, 20(6), 1759–1770. <https://doi.org/10.1109/TASL.2012.2188515>

Kim, J. W., Salamon, J., Li, P., & Bello, J. P. (2018). CREPE: A Convolutional Representation for Pitch Estimation. Proceedings of the IEEE International Conference on Acoustics, Speech, and Signal Processing (ICASSP), 2018, <https://doi.org/10.1109/ICASSP.2018.8461329>