Statistical Variance and Transpositional Analysis: A Commentary on Plazak (2016)

KRIS SHAFFER
University of Colorado-Boulder

ABSTRACT: Joseph Plazak argues for a more nuanced representation of the environment in which listeners encounter music — in particular, questioning the assumption that listeners who engage with music repeatedly do so at the same absolute pitch level. While I agree with Plazak's problematization of this assumption, as well as his advocacy of further investigation into it, Plazak's corpus study has two key methodological flaws. However, a follow-up study in line with Plazak's research could both address these flaws and the more general question that Plazak raises, to the benefit of our understanding of a number of foundational cognitive issues.

JOSEPH Plazak's study of user-posted YouTube lyric videos leads him to several key conclusions:

1. A large proportion of YouTube users alter pitch and/or tempo of a popular music recording before posting to YouTube.
2. This phenomenon changes the listening landscape in which popular songs are
3. "learned" by listeners, which in turn "problematizes a basic assumption about music perception."
4. The tools currently available for conducting automated pitch and tempo analysis are usable, but still inaccurate enough to require manual verification of automated results, which in turn limits corpus size for audio-based studies like this one.
5. This phenomenon is disappearing, so researchers should download data now before the videos are deleted.

Plazak also advocates further research on the effect of multiple transpositions on musical learning and absolute pitch, as well as the relative merits of music information retrieval (MIR) tools.

I agree with Plazak about the need to verify pitch-analysis results from existing tools like Sonic Visualiser/Chordino, and the additional workload it can create. I also agree that this change in the listening landscape necessitates new research in musical learning. As recent research in absolute pitch (AP) and relative pitch (RP) problematizes the AP/RP binary (Ross, Gore, & Marks, 2005), Plazak's study likewise demonstrates that not all recorded music is fixed in absolute pitch. Variance in what we generally have assumed to be absolute — the pitch level and tempo of recordings engaged by listeners — certainly affects our understanding of musical learning, AP, RP, and heightened tonal memory (HTM). However, I take issue with two of Plazak's lines of reasoning.

The first is statistical. Plazak found that 42% of the examples in his 2011 corpus contained pitch and/or tempo alterations, but only 24% of his 2015 corpus contained those alterations. He concludes that the phenomenon is disappearing and speculates on why that might be. While the phenomenon may in fact be disappearing, Plazak's data does not demonstrate that — at least not in this form. He does not account for chance variance as a possible reason for the difference in those two percentages. The small size of the two corpora in question make chance variance a highly probable explanation for this difference, and an important null hypothesis to disprove before concluding an effect and inquiring into the reason for that effect.

The second line of reasoning with which I take issue is both statistical and cognitive. In Plazak's 2011 corpus, roughly one third of the user-generated videos were pitch-altered. For Plazak, this represents a significant change in the musical learning environment ― one which should lead us to reframe our research into listeners' abilities to recall music *at pitch*. After all, if *at pitch* is no longer a single pitch level, it calls the quantitative data of studies like Halpern (1989) and Levitin (1994) into question, and therefore any conclusions about musical memory, AP, and HTM.

However, all of the pitch alterations in Plazak's 2011 corpus and most of the pitch alterations in his 2015 corpus involve alterations of one or two semitones. These one-to-two-semitone alterations are within the window of "good" AP memory in the cognitive studies Plazak references; that is, listeners who recalled a song one or two semitones away from the official recording's pitch level were considered to have good pitch memory. These overlapping ranges do not "problematize a basic assumption about music perception" so much as reveal the need for greater precision of measurement and greater sophistication of calculating relationships between data from cognition experiments and what "at pitch" means for "the" recording. YouTube can provide an excellent starting point, as users have access to rudimentary analytics (number of plays) for various recordings. Play counts of the official recordings on YouTube, Spotify, and other streaming services can be compared and combined to ascertain a more nuanced understanding of the listening environment that forms the background of those ― and new ― studies on pitch recall.

While I disagree with Plazak about the extent to which his YouTube data problematizes the conclusions of those cognitive studies, I agree that the phenomenon Plazak has documented should lead us to explore with greater nuance the complexities of the online listening environment and their effect on musical learning and our understanding of the formation of pitch memories. Such a line of inquiry has tremendous potential for expanding our understanding of online listening habits, pitch memory development, and the natures of absolute pitch, relative pitch, and heightened tonal memory.

REFERENCES

Ross, D.A., Gore, J.C., & Marks, L.E.(2005). Absolute pitch: Music and beyond. *Epilepsy & Behavior,* (*7*)4, 578–601. http://dx.doi.org/10.1016/j.yebeh.2005.05.019