

Commentary on “Transpositions within user-posted YouTube lyric videos: A corpus study”

GARY YIM[1]
Ohio State University

ABSTRACT: Responding to Plazak’s corpus study on pitch and tempo transpositions in user-created music videos on YouTube, I suggest that such transpositions may have musically communicative intentions. Besides the nominal changes designed to defeat copyright infringement detection algorithms, I propose that pitch and tempo changes may be used for listener enjoyment. Some comments on the methodology of the experiment are included.

Submitted 2015 December 15; accepted 2016 April 11.

KEYWORDS: *music communication model, pitch transposition, tempo, song versions*

IN “Transpositions within user-posted YouTube lyric videos: A corpus study”, Plazak investigates one particular way in which Internet users may customize their music listening experiences. When YouTube users upload unofficial videos of current pop songs (with the words overlaid, hence “lyric videos”), the music itself is often altered to be at a pitch level or tempo that is different from the original commercial recordings. By compiling and examining two corpora of unofficial music videos on YouTube, one from 2011 and one from 2015, Plazak reports that a substantial proportion of user-uploaded music has been shifted in either pitch or tempo.

Why is the music altered in this way? Plazak identifies the likely motivation for these pitch and tempo shifts, but in my response, I consider some other broader goals. In doing so, I hope to add to the author’s comments on the musical communication model (Kendall & Carterette, 1990). I will conclude by commenting on some methodological caveats that may be pertinent for YouTube corpus studies in general.

RECONSIDERING COMMUNICATION

The article by Plazak is framed as a study on the role of the listener, according to the music communication model of Kendall and Carterette (1990). It is suggested that by editing the music, even if only by shifting the pitch or tempo, the technologically empowered “listener” is no longer merely a passive receiver at the end of the communication chain. Well, certainly they are not – after all, the video is made publically available on YouTube so that others can subsequently experience it. So perhaps they’re not really the “listeners” at all, and instead, the YouTube poster is best seen as an additional link in the music communication chain: much like a DJ or a producer, they have the opportunity to recode the musical message before it reaches others.

What then is being communicated or expressed when a YouTube poster alters the pitch or tempo of the music? Or what is the intent? The only explicit discussion in the article on the motivation for these alterations relate to defeating YouTube’s algorithms which detect copyright infringement. Another possibility to consider is customizing a recording so that someone can play or sing along with it in a comfortable key or tempo (including the kinds of “karaoke” videos that were explicitly excluded in this study). Both of these motivations lead to solely pragmatic pitch or tempo shifts that assumedly don’t add much to the musical message.

What about the “chipmunk” and “Minion” videos? Although they were also excluded from this study, I would contend – doing my best to keep a straight face – that they are actually important examples of the music communication model. In a sense, the YouTube poster’s version subverts the original message of the song to convey one of their own, at least qualifying it as a “reinterpretation” of sorts. But just so we can discuss something else, here’s a question: besides parody, do people use such transpositions of pitch

and tempo to make serious musical contributions? That is, if YouTube posters are links in the musical communication chain, do they use this opportunity to make a meaningful impact on the listener?

Pitch Transposition

Searching YouTube for “pitch shift songs”, I was only able to find a handful of pitch transposition examples that could fit the criteria described above (being “serious” and not pragmatic or parodies). At first, I found one fourteen-song playlist containing Korean pop and Japanese anime songs that were pitch shifted in order to change the gender of the singers. For example, the song “Hot Summer” performed by the all-girl K-pop group “f(x)” was pitch shifted down, so that it sounds like a boy band. This inspired me to search instead for “male version songs” (and “female”, “boy”, and “girl version songs”), which yielded a trove of “gender transposed” songs. Some of the search results were cover versions recorded by singers of a different gender than the original artist, but there were also numerous pitch-shifted versions achieving the same effect. As far as I can tell, the gender changes are not for karaoke or copyright reasons, and (probably) not for parody either. To me, the ostensible purpose for sharing these alternate gender versions is for fans of the original music to enjoy them.

Tempo Transpositions

It was easier to find examples of videos using “serious” tempo transpositions that were not just pragmatic or parody. Searching for “slowed down music”, one can find many reduced tempo versions of songs, either with or without a corresponding pitch shift down. My initial impression is that these songs are created as “chill” or “laid back” versions of the originals, which may be appropriate for a variety of recreational activities. Subjectively, the tempi for these versions seem to be reduced on average by about 50%, as a first guess. (As an aside, I should mention that it is unfortunate that the magnitude of tempo transpositions in Plazak’s study was not reported.) For example, you might be able to find a version of Michael Jackson’s “The Way You Make Me Feel” that lasts just over eleven minutes. Again, the ostensible purpose for these videos is as alternative versions of songs that can be enjoyed by fans of the originals.

However, besides these kinds of tempo reductions (which I would call “plausible”) there are also some tempo reductions that I might call “excessive.” For example, a top hit returned was the song “Call Me Maybe” by Carly Rae Jepsen, but played “1000% slower.” Unfortunately, the original song is hardly recognizable, but the transposition does provide a somewhat pleasing ethereal ambiance.

Searching for “sped up music” (or “speed up music”) does return a lot of “chipmunk” versions as expected. However, I was pleased to discover that many increased-tempo videos are additionally labeled as “no chipmunk.” Thus, the search for “sped up no chipmunk” returns a plethora of songs that presumably are intended for authentic enjoyment (rather than parody). For example, I found a sped-up version of “Black Widow” by Iggy Azalea that sounds suitable for certain kinds of dancing or exercise, or if you were just in the mood to hear a slightly more pumped up version of it. Although there is a range of tempo increases for fast versions, it seems to me that you can’t increase it all that much before the song sounds silly. I would say that the “plausible” tempo increases were definitely less than 50% and maybe closer to 20%.

Some Research Questions

Corpus-based studies may be used quantify the kinds of transpositions described here. How much are the tempi of songs actually increased (for “non-chipmunk” versions) or decreased on average? Which songs are more likely to be slowed down? (R&B and alternative rock songs?) Which are more likely to be sped up? (Pop songs?) Which gender swaps are most common in songs? (Does this depend on the user’s gender?) I suspect that more exploratory and descriptive research in music YouTubeology will be necessary before hypothesis-testing questions (such as those alluded to in the article) are feasible. For example, studying slowed down versions of songs on YouTube may give some insight into how musical “grooviness” is produced and perceived. But in order to do this, one would probably want to first find out which genres tend to be tempo-reduced, what the typical tempo-reductions are, whether they are commonly combined with pitch-reductions, and what non-groove reasons there may be to reduce tempo (at least) – and you’d have to do your research quickly.

RESEARCH IN MUSIC YOUTUBEOLGY

As Plazak has noted, it is important to collect this data before it disappears. It was reported that the proportion of transposed content in videos in 2011 was less than in 2015, but it might be worth clarifying that due to the method of sampling, this is likely limited to the nominal shifts designed to avoid detection of copyright infringement. Thus, it might be true that it's harder now to find evidence regarding the relationship of implicit AP and nominal YouTube transpositions (as hypothesized by Jakubowski & Müllensiefen, 2013). But other kinds of pitch or tempo shifts, especially those intended for listeners' enjoyment as alternate versions, perhaps may not be declining at all. My final two comments may have some bearing on future YouTube research in this vein or in general.

Using Search Results

In Plazak's study, inclusion in the corpus was not by random selection. Rather, it depended on the video's rank by YouTube's search algorithm. As it was not explicitly mentioned in the article, I assume that the default search options were used. Although this supposedly means "by relevance", I would surmise that the ranking algorithm (the same used as Google's search engines) also takes video popularity into account (ratings, page views). My concern is that popularity may be confounded with the prevalence of the transpositions under study. For example, are videos more popular *because* they aren't shifted? (Some videos advertise "no pitch shift" in their titles.) Or, since a popular video might be more likely to catch the attention of copyright holders, do popular videos use more shifts to avoid detection? Either way, using randomized selection of videos, rather than selecting by rank, may avoid such potential confounds with video popularity.

A secondary point concerns the method of selection: the official videos were always skipped. Actual users are not likely to do this, which means that the corpora formed using this method may not characterize typical user experiences. In the context of Plazak's article, this may weaken some claims related to learning AP. I would be curious to know how many videos were skipped before selecting the ten videos for each song in the corpus. Also, note that search results for YouTube are customized per user, accounting for things like search history, location, and languages. In general, I would promote either selecting all the top results returned by YouTube, or selecting them randomly, if possible.

Validating Automated Tools

I appreciate the validation step in the methodology to guard against false positives from the automated tools, and the report on the performance of the tools is informative. But it's not entirely surprising that they were imperfect at detecting the pitch and tempo shifts – even the YouTube filters don't catch everything. On the other hand, I wonder if some validation to guard against false negatives would have been helpful too – perhaps there were transpositions that were not detected. Admittedly, this seems far less likely: there is a false negative only when a shift occurs that is measured incorrectly in such a way that it matches the original. In general, Plazak's advice to compare a variety of audio analysis toolkits is sound. However, some sampling of the negative matches (or perhaps using a separate data set) would have made the validation process even more reassuring.

ACKNOWLEDGMENTS

I would like to thank Joe Plazak for conducting this innovative study. Although there are not yet many YouTube-based music studies, exploratory and descriptive research such as this are crucial for paving the way. I also thank the editor for giving me not only this opportunity to respond, but also an excuse to spend way too much time on YouTube.

NOTES

[1] Correspondence can be addressed to: Gary Yim, School of Music, 1866 N College Rd., Columbus, OH 43210, yim.28@osu.edu.

REFERENCES

Jakubowski, K., & Müllensiefen, D. (2013). The influence of music-elicited emotions and relative pitch on absolute pitch memory for familiar melodies. *The Quarterly Journal of Experimental Psychology*, *66*(7), 1259-1267. <http://dx.doi.org/10.1080/17470218.2013.803136>

Kendall, R. A., & Carterette, E. C. (1990). The communication of musical expression. *Music Perception* *8*(2), 129-163. <http://dx.doi.org/10.2307/40285493>