Encoding and Analyzing the Timbre in Popular Song (TiPS) Corpus

Lindsey Reymore
Arizona State University
United States
lreymore@asu.edu

Ben Duinker
McGill University
Canada
benjamin.duinker@mail.mcgill.ca

Matthew Zeller
Musical Instrument Museum
United States
matthew.zeller@mim.org

Nicholas Shea
Arizona State University
United States
njshea@asu.edu

Leigh VanHandel
University of British Columbia
Canada
leigh.vanhandel@ubc.ca

Christopher William White
University of Massachusetts Amherst
United States
cwmwhite@umass.edu

Jade Roth
McGill University
Canada
jade.roth@mail.mcgill.ca

Jeremy Tatar
McGill University
Canada
jeremy.tatar@mail.mcgill.ca

Nicole Biamonte
McGill University
Canada
nicole.biamonte@mcgill.ca

Abstract

Timbre and texture are important and perceptually salient stylistic and structural parameters in popular music, yet their specific functional roles in this repertoire have not been theorized. This report describes the construction and encoding of a new popular-music corpus, Timbre in Popular Song (TiPS). The corpus comprises 400 songs, including 100 songs each from four disparate genres: country, pop, heavy metal, and hip hop. Song selection in TiPS balances genre typicality with considerations of gender and racial diversity as well as chronological representation; details related to timbre, texture, and form for each song are being encoded and will be analyzed by genre to identify normative timbral and textural combinations, as well as typical differences among genres.

Introduction

Popular music is well known for the wide timbral variety afforded by the use of both electronic and acoustic instruments, as well as the performance tradition of using gritty-sounding timbres for expressive effect—e.g., distorted electric guitar, the raspy or growly quality of vocal fry, or the "dirty," gravelly saxophone sound created by humming. Formal sections in popular music are usually more clearly defined than in classical music and are often marked by contrasts in timbre and texture. Yet, despite the importance of these parameters in musical expression and structure, and their centrality to music listening, their functional roles in this repertoire have not been closely examined or theorized.

Previous popular music corpora have been encoded primarily with respect to form and pitch (e.g., Burgoyne et al. 2011, De Clercq & Temperley 2011); our Timbre in Popular Song (TiPS) database is the first multi-genre corpus to be constructed for encoding timbral and textural information (the hip-hop corpus in Duinker & Martin 2017 includes annotations for texture and, to some degree, timbre). The TiPS database encodes 400 songs: 100 each from the contrasting genres of country, pop, heavy metal, and hip hop, with ten songs per year from 1990–1999. Since our focus is to make comparisons across genres, we limited this initial stage to songs from a single decade. Our approach to song selection takes into account artists' gender, race, and ethnicity, to help counter the systemic biases often propagated in music corpora and better reflect real-world diversity.

1. Building the Corpus

We compiled a "parent" corpus of 150–225 songs for each genre, based on pre-existing sources (e.g., *Billboard* charts), and encoded demographic variables of gender, race, and ethnicity for each artist. This information was used to create a "child" corpus for each genre. The child corpora each consist of 100 songs, balanced for artist representation (limited to 5 songs per artist) and distributed evenly across the decade.

1.1 Parent Corpora

Parent corpora for each genre were derived from various sources, as no single source provides parallel lists across genres. The pop music parent corpus was drawn from *Billboard's* "Top Songs of the 90s," which includes 500 songs in a variety of genres. Song titles were filtered through iTunes to identify genre; 182 songs that iTunes labelled as "pop" constitute the pop parent corpus. The hip-hop parent corpus includes songs ranked #1 between 1990 and 1999 on *Billboard*'s "Hot Rap Songs" chart, totalling 163 unique songs. The country parent corpus consists of the top 20 "Hot Country Songs" from the *Billboard* Year-End charts for each year from 1990 to 1999, totalling 200 songs. Unlike the other three genres, there are no metal *Billboard* charts, so we derived the metal parent corpus of 225 songs from four lists based on critical acclaim (e.g., *Metal Hammer* 2018, Podoshen 2017). As these lists included songs that might be classified as hard rock or alternative rock, we developed a genre-filtering method similar to that used for the pop parent corpus, but with additional input from expert metal listeners.

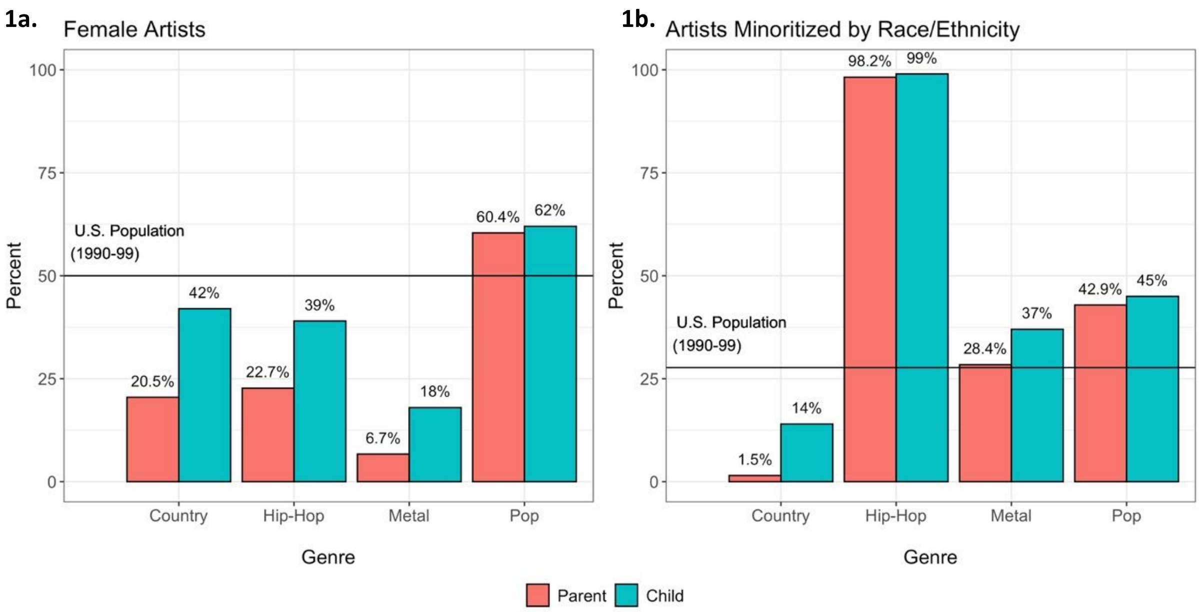
1.2 Demographic Encoding

Next, we encoded demographic information for all artists across the four parent corpora. Our demographic encoding system, developed from work by Shea (Shea 2022, Shea et al. 2024), records whether the lead artist or any band members belong to a marginalized racial, ethnic, or gender group. We adopted the racial and ethnic categories used by The Music Coalition of the Annenberg Inclusion Initiative at the University of Southern California (Smith et al. 2018). With respect to gender, we considered any gender identity other than male as subject to marginalization. We also recorded a variable called "primary" that reflects the agency of marginalized artists within their ensembles. This variable is encoded as positive if either the title/lead member of the ensemble or more than half of the group's members are racial, ethnic, or gender minorities. For each entry, demographic encoders included their sources of information and commented on their decision-making processes as needed.

1.3 Child Corpora

Based on our demographic encoding, we calculated the proportions of marginalized artists in each parent corpus by gender and race/ethnicity. We then set target proportions for gender and race/ethnicity for the child corpora, based on the geometric mean between the proportions in the parent corpora and the relevant U.S. Census data for that decade. In constructing the child corpora, we retained all songs by marginalized artists from the parent corpora. Next, in cases where the target proportion was below the population proportion, we supplemented the corpus via other sources in order to increase representation. Such adjustments for gender were made to the country, hip hop, and metal corpora, and adjustments for race/ethnicity were made to the country corpus. Figure 1a shows differences in gender representation by the proportion between parent and child corpora for each genre; Figure 1b shows the same comparison for racial/ethnic representation.

We recognize that there are multiple parameters of diversity, and any sampling method is necessarily imperfect. Our approach is not a solution to deeply ingrained issues of inequity, but our anti-discriminatory sampling framework represents a step toward more inclusive music scholarship.



Figures 1a-1b: Proportions of marginalized artists in parent and child corpora compared to proportions in the U.S. population for 1990-1999 based on census data (horizontal line).

2. Encoding Timbral and Textural Information

Encoded metadata includes artist(s), song and album title, *Billboard* rank, release year, and a link to the source recording used. Information about timbre and texture is encoded by formal section; each row in the dataset corresponds to one formal section of a song, with variables in columns. Repeating sections (e.g., verse, chorus) are numbered in the order they occur (e.g., the first and second choruses are labelled "C1" and "C2," respectively). Variables related to timbre and texture were recorded for each formal section, as shown in Table 1.

Category	Name	Definition	Format
Form	Formal Unit	Formal module type	Categorical-S
	Timestamp	Time at beginning of formal unit	Time (min:sec)
	Timbral Marker	Whether a salient timbral event begins unit	Binary
	Change in Timbral Profile	Whether there is a salient change in overall timbre from previous unit	Binary
Instrumentation	Explicit Beat Layer	Instruments articulating pattern of beats	Categorical-M
	Functional Bass Layer	Instruments playing lowest pitched layer	Categorical-M
	Harmonic Fill Layer	Instruments filling the registral space between bass and melody	Categorical-M
	Melodic Layer	Instruments carrying a melody	Categorical-M
	Novelty Layer	Instruments that provide coloristic effects	Categorical-M
Texture	Texture Category	Describes general texture type	Categorical-S
	Variation	Describes how timbre and/or texture changes within a formal unit	Categorical-S
	Texture Rating	General textural density	Integer
Voice	Number of Voices	Number of distinct voices singing melodically	Integer
	Vocal Production	Most prominent vocal qualities of the lead vocalist(s)	Categorical-M
	Speak/Sing	Mode of delivery by lead vocalist(s)	Categorical-M
	Backup Vocals	Textural function of the background vocals	Categorical-M

S = single variable selection; M = multiple possible selections

Table 1: Timbre and texture variables encoded in the TiPS corpus

These variables were created and refined collaboratively by the authors. After compiling an initial set of proposed variables, several authors applied them to sample songs. We discussed the results, revising variables and definitions in response. We then completed a second round of encoding using the refined variables and a second round of collaborative revision. Next, research assistants (RAs) were trained to encode variables; each RA completed five songs, which were carefully reviewed by one of the authors. This resulted in a third stage of collaborative refinements to the variables and definitions, after which the RAs continued encoding. Encoding for a sample song is shown in Table 2.

3. Project Status

Song encoding is complete, and we are currently cross-checking the annotations. Once finished, the database will be made available open-access in CSV format. Our first analytical goal is to identify normative timbral, textural, and formal characteristics of each genre: for example, we will identify which timbral vocal qualities (e.g., vocal fry, breathiness, growling, belting) are most typical of each genre, and how the function of background vocals varies. We will investigate how timbre and texture interact with form and genre, and examine correlations among the variables, such as the close relationship between instrumentation and texture. From these analyses, we will develop new theoretical models for timbre and texture in popular music.

formal unit	timestamp	timb. marker	timb. change	explicit beat	funct. bass	harm. filler	melody	novelty	texture cat.	variation	texture rating	# of voices	vocal prod.	delivery	backup voc.	comments
I	0:01	0	increase	drums	e bass:clean	e gtr:clean, synth:dark, synth:strings		sample:wind chime, lead voc, sec voc	varied		1	0			other	Spoken & processed voc
C1	0:34	1	same	drums	e bass:clean	e gtr:clean, synth:strings, synth:dark		lead voc, sec voc	no mel	1	3	0			other	Secondary vocal spoken
C2	0:52	1	same	drums	e bass:clean	e gtr:clean, synth:strings, synth:dark, bg voc	lead voc	lead voc, sec voc, pno	mel/accomp	1	4	M	airy/breathy	sung	harmony, other	timbral marker: piano
V1	1:13	1	same	drums	e bass:clean	e gtr:clean, synth:dark	lead voc		mel/accomp	1	2	1	airy/breathy	sung		
P1	1:33	1	same	drums	e bass:clean	e gtr:clean, pno, synth:dark, synth:strings, vibes	lead voc		mel/accomp	1	3	1	airy/breathy	sung		
C2	1:41	1	same	drums	e bass:clean	e gtr:clean, pno, synth:dark, synth:strings, bg voc	lead voc	sample:wind chime, lead voc, sec voc	mel/accomp	1	4	M	airy/breathy	sung	harmony, other	pno unison w/synth
V2	2:02	1	same	drums	e bass:clean	e gtr:clean, synth:bright, synth:dark	lead voc	sample:wind chime	mel/accomp	1	3	1	airy/breathy,creaky/fry	sung		Synth high countermel
V3	2:21	1	same	drums	e bass:clean	e gtr:clean, synth:bright, synth:dark, vibes	lead voc	secondary vocal	mel/accomp	1	3	1	airy/breathy,creaky/fry	sung	other	Spoken & processed voc
P2	2:42	1	same	drums	e bass:clean	e gtr:clean, pno, synth:dark, synth:strings, vibes	lead voc		mel/accomp	1	3	1	airy/breathy,belt/chest	sung		
C3	2:50	1	same	drums	e bass:clean	e gtr:clean; pno, synth:dark, synth:strings, vibes, bg voc	lead voc	sample:wind chime, sec voc	mel/accomp	1	5	M	airy/breathy,falsetto/head	sung	harm, countermel, other	Deeper hit on downbeat
C4	3:21	1	same	drums	e bass:clean	e gtr:clean; pno, synth:dark, synth:strings, vibes, bg voc	lead voc	secondary vocal	mel/accomp	1	4	M	airy/breathy	mostl	melody,other	
C5	3:40	0	decrease	drums	e bass:clean	e gtr: clean, synth:bright, synth:strings, vibes	lead voc	sample:wind chime, sec voc	varied	1	3	1	airy/breathy	sung	other	Progressive drop-out
C6	3:59	1	decrease	drums	e bass:clean	e gtr:clean; pno, synth:dark, synth:strings, vibes, bg voc	lead voc	sample:wind chime, sec voc	varied	1	4	M	airy/breathy	sung	harmony,other	Progressive drop-out

Table 2: Encoding for Janet Jackson, “That’s the Way”

Conclusion

Our study offers a model for creating a more equitable corpus of popular songs using a new demographically adjusted sampling method, as well as a new encoding system for timbre and texture, which are treated as structural parameters for the first time in a corpus of popular songs. The next phase of our project will expand the repertoire beyond our initial four genres of pop, country, metal, and hip hop to add three more important genres: rhythm & blues, rock, and electronic dance music. We will also expand the corpus beyond the 1990s to encompass all decades in the sixty-year span between the 1960s and 2010s.

Acknowledgements

This research has been supported by funding from the Analysis, Creation, and Teaching of Orchestration (ACTOR) Project, the Social Sciences and Humanities Research Council of Canada (SSHRC), and the Fonds de recherche du Québec - Société et culture (FRQSC).

References

Burgoyne, J.A., Wild, J., & Fujinaga, I. (2011). An expert ground truth set for audio chord recognition and music analysis. In *Proceedings of the 12th International Society for Music Information Retrieval Conference*, 633–38. https://ismir2011.ismir.net/papers/OS8-1.pdf

De Clercq, T. & Temperley, D. (2011). A corpus analysis of rock harmony. *Popular Music* 30(1), 47–70. doi:10.1017/S026114301000067X https://www.cambridge.org/core/journals/popular-music/article/abs/corpus-analysis-of-rock-harmony/C5210A8EC985DDF170B53124F4464DA4

Duinker, B. & Martin, D. (2017). In search of the golden age hip-hop sound (1986–1996). *Empirical Musicology Review* 12(1-2), 80–100. doi: 10.18061/emr.v12i1-2.5410 https://emusicology.org/article/view/5410

Metal Hammer. (2018). The 100 best metal songs of the 90s. October 5, 2018 (orig. August 9, 2017). https://www.loudersound.com/features/the-100-best-metal-songs-of-the-1990s

Podoshen, J. (2017). Here's an alternate 'Best Metal Songs of the 90s' list. *Metal Insider*. https://metalinsider.net/lists/heres-an-alternate-best-metal-songs-of-the-90s-list

Shea, N. (2022). A demographic sampling model and database for addressing racial, ethnic, and gender bias in popular-music empirical research. *Empirical Musicology Review* 17(1): 49-58. Forthcoming.

Shea, N., L. Reymore, C. W. White, B. Duinker, L. VanHandel, M. Zeller, & N. Biamonte (2024). Diversity in music corpus studies. *Music Theory Online* 30.1. Forthcoming.

Smith, S. L., Choueiti, M., & Pieper, K. (2018). Inclusion in the recording studio? gender and race/ethnicity of artists, songwriters, and producers across 600 popular songs from 2012-2017. Annenberg Inclusion Initiative. https://assets.uscannenberg.org/docs/inclusion-in-the-recording-studio.pdf