

# **Consonant & Dissonant Musical Sonorities: A Data Science Approach**

**STATS/CSE 780  
Project Proposal**

Konrad Swierczek - 001423065

3/12/23

## Introduction

Considerable debate has been devoted to musical consonance and dissonance, or the pleasantness of a sound. The study of this phenomenon goes as far back as Pythagoras, but the 20th and 21st centuries, and our understanding of acoustics, auditory perception, and mathematics, led to experimental and modelling solutions. Theories ranging from “roughness” (interactions of tones on the basilar membrane) (Plomp and Levelt 1965; Kameoka and Kuriyagawa 1969a, 1969b; Hutchinson and Knopoff 1978, 1979; Mashinter 2006), “harmonicity” or “periodicity” (how similar is a sound to a harmonic series) (Parncutt 1988; Milne, Laney, and Sharp 2016; Harrison and Pearce 2018), and familiarity Lahdelma, Eerola, and Armitage (2022). Harrison and Pearce (2020) and Eerola and Lahdelma (2021) set out to compare state-of-the-art models of consonance and dissonance to human behaviour data. The results of these studies found a combination of these models may be necessary to explain the underlying mechanisms of consonance and dissonance. Here, a re-examination of the datasets explored in Harrison and Pearce (2020) and Eerola and Lahdelma (2021) will attempt to develop solutions for prediction and classification of consonant and dissonant sounds. Despite the progress made, it is still unclear if there is a line between these two percepts in a binary sense. Unsupervised learning methods can aid in determining if a boundary exists between consonant and dissonant sounds, while supervised learning can help use the underlying features to generate a predictive model for determining how new sounds will be classified. In addition, feature selection will be used to narrow the scope of any predictive model to statistically salient features.

<sup>1</sup>

## Methods

The Anatomy of Consonance dataset (Eerola and Lahdelma 2021) is a collection of 9 datasets with 617 observations from human behaviour experiments on consonance/dissonance perception. It includes 33 feature continuous feature, three of which will be excluded due to lack of clarity in documentation (“TDL”, “TDL1”, “neg\_log\_prob”), 7 metadata properties (only the categorical variable “dataset” referring to the source of the observation is of interest here), and 3 normalized rating variables which correspond to the mean, standard deviation and standard error of human participant ratings for a particular sound (observation). The features include various roughness,

---

<sup>1</sup>All materials and reproducible code used are available at <https://github.com/konradswierczek/STATS780>

harmonicity, periodicity, spectral profile, and familiarity models as well as features describing the numerosity, or amount of pitches, in a sound.

First, linear regression using forward stepwise subset selection and cross validation will be performed on the dataset to create a model that predicts rating while also selecting only the most relevant features out of the large pool. However, the above approach does not guarantee the subset of features selected will represent the most suitable group of features, much less an approximation of human perception since all of these features have in their own way been implicated in perceptual processes. A second linear regression will be performed on components generated using principal component analysis. The performance of these two different approaches can be compared directly on a hold-out dataset. Alternatively, an unsupervised learning approach not relying on ratings will use k-means clustering ( $k=2$ ) to find consonant and dissonant clusters. The result of this analysis will be used with logistic regression to predict consonant and dissonant sounds.

These two approaches use different underlying assumptions to develop a model for consonance and dissonance: one assumes humans rate sounds in a roughly binary fashion (i.e., something is either dissonant or consonant to varying degrees), or dissonance is simply a continuous variable akin to “cold” being an absence of heat. This analysis will help shed light on this issue.

All analyses will be performed in R and Python (R Core Team 2020; Van Rossum and Drake 2009)

## Exploratory Results

Results of correlation analysis. Outlier exclusion is not necessary for the rating variable as it is already an average of multiple participants and has been normalized by Eerola and Lahdelma (2021). Outliers will not be excluded for the feature variables numerosity and pitch range due to these sounds being outside of common musical practice. None of the feature variables or rating variable have missing values, but the standard deviation and standard error for the ratings are missing values. However, these values will likely not be used in the analysis. Some of the features are highly correlated and variable (Figure 2, Figure 3), making this dataset a ideal candidate for principal component analysis (Gareth et al. 2013). Figure 1 shows the distribution of rating values, which appears to be a skewed or bimodal distribution consistent with the idea that chords

are grouped into consonant or dissonant (the peak at a lower value is also consistent with the western music general preference for consonant sounds).

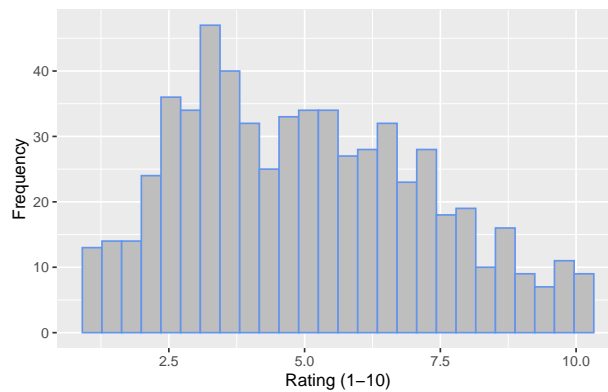


Figure 1: Distribution of ratings across 617 observations. Each observation is the average of all participants in the experiment.

## Next Steps

Model training and data analysis, and interpretation of the results will be completed by the end of March, 2023. Presentation of the results to the STATS/CSE780 class will be held in the first half of April, pending scheduling. The final report will be completed by April 17, 2023.

The results of this study should not only further our understanding of consonance and dissonance, but also create a more simple modelling approach for applied tasks including music perception experiments, content-based music classification and recommendation systems, and other auditory perception challenges such as developing hearing aids enhanced for music.

## References

- Eerola, Tuomas, and Imre Lahdelma. 2021. "The Anatomy of Consonance/Dissonance: Evaluating Acoustic and Cultural Predictors Across Multiple Datasets with Chords." *Music & Science* 4: 20592043211030471.
- Gareth, James, Witten Daniela, Hastie Trevor, and Tibshirani Robert. 2013. *An Introduction to Statistical Learning: With Applications in R*. Springer.
- Harrison, Peter, and Marcus T Pearce. 2018. "An Energy-Based Generative Sequence Model for Testing Sensory Theories of Western Harmony." *arXiv Preprint arXiv:1807.00790*.
- . 2020. "Simultaneous Consonance in Music Perception and Composition." *Psychological Review* 127 (2): 216.
- Hutchinson, William, and Leon Knopoff. 1978. "The Acoustic Component of Western Consonance." *Journal of New Music Research* 7 (1): 1–29.
- . 1979. "The Significance of the Acoustic Component of Consonance in Western Triads." *Journal of Musicological Research* 3 (1-2): 5–22.
- Johnson-Laird, Phil N, Olivia E Kang, and Yuan Chang Leong. 2012. "On Musical Dissonance." *Music Perception: An Interdisciplinary Journal* 30 (1): 19–35.
- Kameoka, Akio, and Mamoru Kuriyagawa. 1969a. "Consonance Theory Part i: Consonance of Dyads." *The Journal of the Acoustical Society of America* 45 (6): 1451–59.
- . 1969b. "Consonance Theory Part II: Consonance of Complex Tones and Its Calculation Method." *The Journal of the Acoustical Society of America* 45 (6): 1460–69.
- Lahdelma, Imre, Tuomas Eerola, and James Armitage. 2022. "Is Harmonicity a Misnomer for Cultural Familiarity in Consonance Preferences?" *Frontiers in Psychology* 13: 124.
- Mashinter, Keith. 2006. "Calculating Sensory Dissonance: Some Discrepancies Arising from the Models of Kameoka & Kuriyagawa, and Hutchinson & Knopoff."
- McDermott, Josh H, Alan F Schultz, Eduardo A Undurraga, and Ricardo A Godoy. 2016. "Indifference to Dissonance in Native Amazonians Reveals Cultural Variation in Music Perception." *Nature* 535 (7613): 547–50.
- Milne, Andrew J, Robin Laney, and David B Sharp. 2016. "Testing a Spectral Model of Tonal Affinity with Microtonal Melodies and Inharmonic Spectra." *Musicae Scientiae* 20 (4): 465–94.
- Parncutt, Richard. 1988. "Revision of Terhardt's Psychoacoustical Model of the Root (s) of a

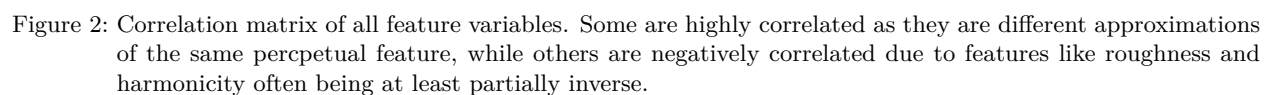
- Musical Chord.” *Music Perception* 6 (1): 65–93.
- Parncutt, Richard, and Graham Hair. 2011. “Consonance and Dissonance in Music Theory and Psychology: Disentangling Dissonant Dichotomies.” *Journal of Interdisciplinary Music Studies* 5 (2).
- Plomp, Reinier, and Willem Johannes Maria Levelt. 1965. “Tonal Consonance and Critical Bandwidth.” *The Journal of the Acoustical Society of America* 38 (4): 548–60.
- R Core Team. 2020. *R: A Language and Environment for Statistical Computing*. Vienna, Austria: R Foundation for Statistical Computing. <https://www.R-project.org/>.
- Van Rossum, Guido, and Fred L. Drake. 2009. *Python 3 Reference Manual*. Scotts Valley, CA: CreateSpace.
- Zajonc, Robert B. 2001. “Mere Exposure: A Gateway to the Subliminal.” *Current Directions in Psychological Science* 10 (6): 224–28.

## Supplementary Materials

::: {.cell}

```
# knitr setup
knitr::opts_chunk$set(echo = FALSE, warning = FALSE, message = FALSE)
knitr::opts_chunk$set(fig.pos = "H", out.extra = "")
#####
# Imports
packages <- c("dplyr", "readr", "ggplot2", "knitr", "ggcorrplot", "tidyr")
lapply(packages, library, character.only = TRUE)
#####
set.seed(75)

# Anatomy of Consonance Dataset
# https://github.com/tuomaseerola/anatomy-of-consonance
# Pull Data
temp <- paste(tempfile(), ".zip", sep = "")
options(timeout = 60 * 10)
"https://github.com/tuomaseerola/DCD/archive/refs/heads/master.zip" %>%
download.file(temp)
experiment1 <- unz(temp, "DCD-master/data/DCD_predictors.csv") %>%
  read_csv()
"github.com/tuomaseerola/anatomy-of-consonance/archive/refs/heads/main.zip" %>%
download.file(temp)
link <- "anatomy-of-consonance-main/data/experiment2_data.csv"
experiment2 <- unz(temp, link) %>%
  read_csv()
link <- "anatomy-of-consonance-main/data/experiment3_data.csv"
experiment3 <- unz(temp, link) %>%
  read_csv()
vars <- c(7, 11:27, 27:34, 37:39, 42)
# Missing values for feature variables>
```





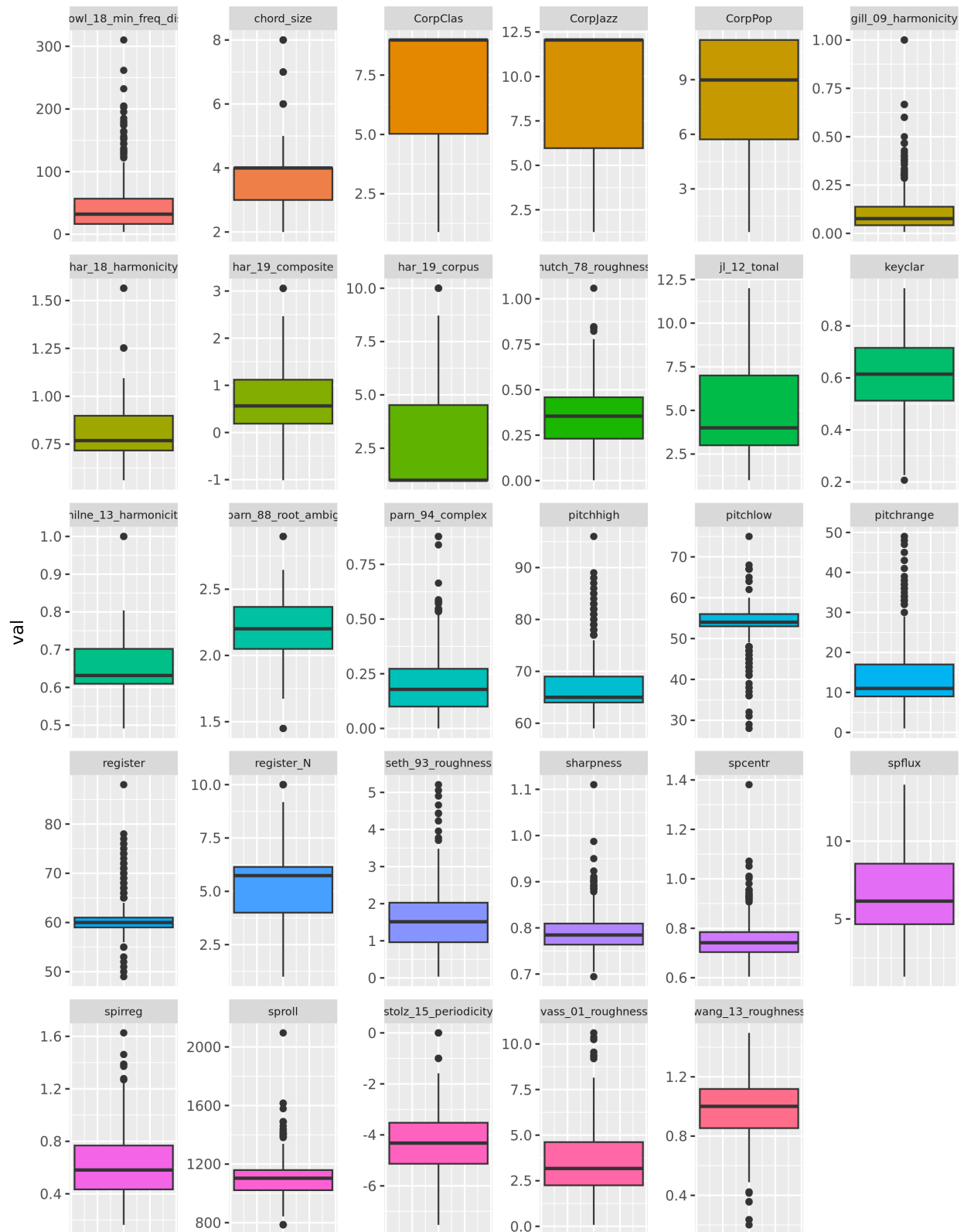


Figure 3: Distribution of continuous feature variables. As all the variables are sourced from different publications and represent various constructs, they appear to be fairly different from one another.

```
table(is.na(experiment3[, vars]))
# Missing values for ratings
table(is.na(experiment3[, 7]))
# Missing values for rating SD/SE
table(is.na(experiment3[, 8:9]))
# Distribution of ratings
binwidth <- function(col) ((max(col)-min(col))/sqrt(length(col)))
ggplot(experiment3, aes(x = rating)) +
  geom_histogram(binwidth = binwidth(experiment3$rating),
                 colour = "cornflowerblue", fill = "gray") +
  xlab("Rating (1-10)") +
  ylab("Frequency")
# Correlation matrix for features
cor_mat <- round(cor(experiment3[, vars]), 3)
ggsave("img/cor_matrix.png", width = 8.5, height = 11,
        ggcorrplot(cor_mat, insig = "blank", type = "lower")
)

ggsave("img/feature_dist.png", width = 8.5, height = 11,
        ggplot(gather(experiment3[, vars], feature, val),
               aes(y = val, fill = feature)) +
        geom_boxplot() +
        facet_wrap(~feature, scales = "free") +
        theme(axis.title.x=element_blank(), axis.text.x=element_blank(),
              axis.ticks.x=element_blank(), legend.position = "none",
              strip.text.x = element_text(size = 7))
)
```

...