# Billboard Melodic Music Dataset (BiMMuDa) Transcription Strategy

Madeline Hamilton

December 13, 2023

## 1 Introduction

The purpose of this document is to describe in detail the exact data creation process employed when compiling the Billboard Melodic Music Dataset. Music transcription is complicated, and lots of decisions were made regarding how exactly the melodies were to be translated into music notation. For a less detailed overview, see our TISMIR paper entitled *The Billboard Melodic Music Dataset (BiMMuDa)*.

Please note that the structure of this document borrows heavily from that of the transcription guide for the CoCoPops project, headed by Claire Arthur and Nat Condit-Schultz.[1] We thank Claire and Nat for permitting us to use this structure.

## 2 People

This section lists the individuals involved in the BiMMuDa project and describes their roles.

### 2.1 Annotators

Three people were directly involved in the creation of the BiMMuDa: Madeline Hamilton, Edward Hall and Ana Clemente. Below are each annotator's qualifications and role in the creation process.

Madeline Hamilton - Madeline has 10 years of formal music education, which focused on piano performance but also included music theory and auditory skills. All manual transcription was undertaken by Madeline. Transcriptions that she was not completely confident in were transferred to either Ana or Edward for checking and correction. This occurred for

[1] https://github.com/Computational-Cognitive-Musicology-Lab/CoCoPops-Billboard-legacy/blob/master/MelodicTranscriptions/TranscriptionInstructions_2019.pdf

about one third of the songs. More recently, Madeline also performed a round of checks and corrections for every transcription.

Edward Hall - Edward has both a Bachelors and Masters degree in Music Studies. As part of these programs, he studied music theory, tonal and post-tonal analysis, techniques of composition, performance (viola), musicology, and music perception. Edward checked and corrected half of the transcriptions Madeline was not confident in.

Ana Clemente - Ana has 16 years of formal music education which covered topics in music theory, composition, analysis, auditory skills, solo performance, chamber music, and orchestra. She also has over 20 years of experience as a professional musician and conductor, performing in ensembles and orchestras. Ana checked and corrected half of the the transcriptions Madeline was not confident in.

### 2.2 Other

Marcus Pearce played a supervisory role during the creation of the dataset and provided guidance and editing for the TISMIR paper.

## 3 General

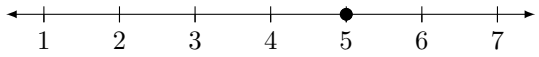
Here we outline the general transcription strategy.

### 3.1 Song Version

Many songs represented in the BiMMuDa have multiple versions, e.g., an album version, a single version, remixes, etc. For every song, the version that is transcribed is the exact version that appeared on the *Billboard* year-end singles chart, which is the single version.

### 3.2 Degree of Interpretation/Musicality

Transcriptions can differ based on their degree of musicality. Some transcriptions are more literal, attempting to capture exact onset/offset times and the exact frequencies of sung pitches. Others are more musical and try instead to capture artists' "intended" pitches and rhythms. The most musical transcriptions are lead sheets, where only very simplified versions of the melody is notated. We can envision this phenomenon on a continuous scale, with "1" meaning "Most Literal Transcription" and "7" meaning "Most Musical Transcription".




The BiMMuDa transcriptions are intended to be a "5" on this scale. They include a fair bit of detail more than a lead sheet, transcribing a particular version/performance of a song rather than providing sheet music for the song in general. However, the transcriptions are musical enough to be somewhat easily translated into musical notation, and the transcriptions infer artists' intentions to some degree.

### 3.3 Main Melody Identification

In the BiMMuDa, only the lead/main melody is transcribed. "Main melody identification" refers to the process of determining which instrument is carrying the main melody. Most of the time, the lead vocalist obviously carries the main melody. In more recent pop songs, a lead vocalist and one or more featured vocalists take turns singing the main melody (see, for example, "Meant to Be" by Bebe Rexha featuring Florida Georgia Line, 2018_03 in the dataset). However, there are some cases where the main melody is not apparent. The procedure for each of these cases is listed below:

- **Duets -** Some songs, such as "Simple Melody" by Gary and Bing Crosby (1950_05 in the dataset), feature two vocalists who sing equally salient melodies, which may overlap or be sung simultaneously. The annotators' best judgement was used to decide which of the two melodies to transcribe for a given bar. Normally, the transcribed melody alternates between the two vocal lines in a way the annotators see fit.

- **Overlapping Melodies -** Often, main melodies will overlap slightly. The most common case is that the final note of the main melody from one section overlaps with the first note of the main melody of the following section. In the full transcriptions, the latter is prioritized, but the transcriptions of individual sections include any notes that were clipped in the full sheet music.

- **Songs with No Vocals -** In this dataset, the songs with no vocals are usually orchestral pieces, such as "The Song from Moulin Rouge" by Percy Faith (1953_01 in the dataset) (see also the list of songs in the dataset of non-vocal main melodies, on the front page of the GitHub repository). In such pieces, one melody line usually the loudest and most perceptually salient. This is the melody that was transcribed.

- **Harmonies Close in Volume to the Main Melody -** The main melody can be difficult to identify in some songs due to the presence of several harmonies which are very close in volume to the lead melody. Some good examples of this are "I Get Around" by the Beach Boys and "Bad Guy" by Billie Eilish (1964_05 and 2019_04 in the dataset, respectively). These are the most difficult main melody identification cases. To

choose a melody line to transcribe, the annotators listened and tried to sing along to the song several times. The melody line most often sung was transcribed.

- **Ad-libs -** Some genres of music often feature ad-libs, like "yeah" or "uh" sounds, sung in between the primary lyrics. An example would be "War" by Edwin Starr (1970_05 in the dataset). Whether or not these are included in the transcription depends on their perceived importance and volume relative to the vocals singing the primary lyrics, with the annotators using their best judgement. Non-pitched ad-libs are not transcribed.

### 3.4 Repetition of Sections

As explained in the TISMIR paper, pop songs often consist of repeated sections. It is important to note that minor variations between repetitions of a section, for example, a repetition of a verse or chorus melody), are *not* captured in the transcriptions.

### 3.5 Non-melodic Sections

In popular music, the lead melody does not usually begin in the first bar, and it does not always play throughout the entire song. Songs often have an introductory 4, 8, or 16 bars before the verse begins, and sometimes bridge sections with no lead vocals. These gaps are accurately represented in the transcriptions with blank bars.

### 3.6 Key and Tempo

Key and tempo changes are included in the full transcriptions. Only major and minor keys are used; when the song utilizes mixolydian and/or dorian modes, the closest major or minor key is used in the sheet music and MIDIs, with the appropriate accidentals. The metadata uses only major and minor keys.

## 4 Pitch

In this section, transcription guidelines regarding pitch are outlined.

### 4.1 Range

The transcriptions attempt to capture the absolute pitches being sung or played. This means notating the pitches in the correct octave as well as with the correct note names.

### 4.2 Vocal Runs / Glide

Some styles of pop music feature long vocal runs, where vocalists sing several notes in the same syllable. Examples include songs by Mariah Carey (1994_05, 1996_02 in the dataset) Boyz II Men (1994_03, 1992_01 in the dataset), and Whitney Houston (1993_01 in the dataset). Vocal runs are transcribed in a moderate amount of detail, with a focus on accuracy. Glides are generally not transcribed.

### 4.3 Out-of-Key Singing

When the vocalist sings out of key (that is, slightly flat or sharp), annotators infer the intended pitch. In general, the transcriptions are averse to using accidentals unless the accidental is clearly intended.

### 4.4 Non-pitched Vocals

Many songs feature non-pitched vocals, such as rapping, in conjunction with singing. These are not transcribed; any bars with non-pitched vocals are left blank. Sometimes, vocals are semi-pitched: for example, autotune effects can introduce a pitch-like quality to non-pitched vocals (see "The Box" by Roddy Ricch (2020_03)). Vocals are generally not transcribed unless they are clearly pitched.

## 5 Rhythm

This section details the transcription guidelines regarding rhythm.

## 6 Expressive Timing and Non-Rhythmic Singing

Vocalists will often be slightly behind or ahead of the beat, either unintentionally or for an intentional emotional or stylistic effect. These minor deviations are not captured by the transcriptions; metronomic timing is used instead.

Other songs have very weak or nonexistent metrical elements, and vocalists sing freely, only loosely being guided by meter or not at all. Examples include "Mona Lisa" by Nat King Cole (1950_02) and "Because of You" by Tony Bennett (1951_02). This makes transcription difficult because the MuseScore and MIDI formats require a strict meter. In these cases, annotators tried their best to encode onset times such that the sheet music is easy to read (that is, they tried not use very complicated rhythms while transcribing), but the onset times still somewhat align with the audio.

# 7 Swing

When songs feature swing timing, swing timing is encoded directly into the corresponding MuseScore and MIDI files.