

DEMONSTRATION OF THE MEERTENS TUNE COLLECTIONS

Peter van Kranenburg, Martine de Bruin, Louis P. Grijp
Meertens Institute, Amsterdam

Frans Wiering
Utrecht University

{peter.van.kranenburg,martine.de.bruin,louis.grijp}@meertens.knaw.nl

F.Wiering@uu.nl

With the Meertens Tune Collections (MTC) the Meertens Institute provides a rich set of collections of musical data for research purposes, such as musicological investigations or music information retrieval tasks. The collections contain thousands of items. Over the past decades, these data have been collected in the Database of Dutch Songs.¹ The online interface of the Database of Dutch songs provides access at the level of individual records through extensive search and browse functionality. With the MTC, several collections are provided as a whole. The melodies are provided in ****kern** format. Furthermore, pdf and png files of the scores are provided, as well as midi files, lilypond sources, and separate files with the (syllabized) lyrics. MTC currently consist of the following collections:

MTC-OGLAUDIO

Contents Collection *Onder de groene linde*: **7178 audio recordings** collected by Dutch field workers during the 1950s–1980s.

Data types mp3

MTC-OGGLSCANS

Contents Scans of **3754 transcriptions** of recordings from *Onder de groene linde* as made during the 1950s–1980s. The music is hand-written, the lyrics are typed.

Data types jpg

MTC-FS

Contents **4120 folk songs**, digitally encoded both from *Onder de groene linde* (2503) and from various related written sources (1617).

Data types ****kern**, midi, lilypond, png, pdf, txt

MTC-INST

Contents **2368 instrumental tunes**, digitally encoded from 18th-century Dutch manuscripts and printed scores.

¹<http://www.liederenbank.nl>



© Peter van Kranenburg, Martine de Bruin, Louis P. Grijp, Frans Wiering.

Licensed under a Creative Commons Attribution 4.0 International License (CC BY 4.0). **Attribution:** Peter van Kranenburg, Martine de Bruin, Louis P. Grijp, Frans Wiering. “Demonstration of The Meertens Tune Collections”, 15th International Society for Music Information Retrieval Conference, 2014.

Data types ****kern**, midi, lilypond, png, pdf

MTC-ANN

Contents Annotated Corpus: 360 melodies used in various publications. With annotations on motif occurrences and melodic similarity at phrase level.

Data types ****kern**

MTC-LC

Contents Large Corpus: 4830 melodies used in various publications.

Data types ****kern**

The collections come with a wealth of meta data. Some highlights: most melodies have a tune family label, which connects variants of the same tune. Several levels of membership are indicated. For the audio recordings, we provide date and location (longitude and latitude) of recording and date of birth and place of birth of the singer(s). For all folk song melodies, phrase endings are indicated. Meta data is provided in comma-separated text files (csv) in UTF8 encoding.

Examples of Music Information Retrieval tasks that could be addressed with this data set:

- Melodic Segmentation - using the phrase markers
- Melodic Similarity - using the tune family labels
- Geographic Clustering - using geographic information of singers and recordings
- Audio-Score alignment - using mp3 and scans
- Singer classification - using singer names
- Optical Music Recognition - using the scans
- Pattern discovery - using the motif annotations
- Lyric-Audio alignment
- And many more...

A concise description of the MTC is available in the following report:

Van Kranenburg, Peter & Martine de Bruin & Louis P. Grijp & Frans Wiering. “The Meertens Tune Collections”. *Meertens Online Reports*, No. 1. Amsterdam: Meertens Institute, 2014.

If you use this data for a research paper, please cite this report.

The MTC are available from:

<http://www.liederenbank.nl/mtc>.