HKB
Hochschule der Künste Bern
Haute école des arts de Berne
Bern University of the Arts



Catalogue of the Ehrenreich Collection, Version 1.0

Documentation

Beat Estermann, Bern University of Applied Sciences

Version 1.0
7 July 2019

Berner Fachhochschule
Haute école spécialisée bernoise
Bern University of Applied Sciences

Contents

1	The Ehrenreich Collection	3
	1.1 Years covered and uniqueness of the recordings	3
	1.2 Digitization and preservation	3
	1.3 Copyright and access policy	4
2	The Catalogue	5
	2.1 Data quality	5
	2.2 Complementary meta data	5
	2.3 Data fields	6
	2.4 Tape identifiers	10
3	Contact	11

Cover photo:
One of the 2260 reel-to-reel tapes in Ehrenreich's main collection, complete with the box and the documentation it contains. Photo: Beat Estermann, March 2018, CC-0.

This document is made available for re-use under the
Creative Commons Attribution 4.0 International (CC BY 4.0) License.

The actual collection metadata is made available free of any copyright restrictions.
The terms of the CC0 1.0 Universal Public Domain Dedication apply.

1 The Ehrenreich Collection

This document contains the documentation of **Version 1.0 of the Catalogue of the Ehrenreich Collection**, which is maintained by the Bern University of Arts, to whom the collection has been bequeathed in 2017.

The Catalogue pertains to the main collection assembled by Leroy Allan Ehrenreich (1929-2016), a New York based opera aficionado, who collected recordings of opera performances and recitals throughout his lifetime. His main collection comprises 2260 reel-to-reel tapes holding up to 14 hours each, totaling up to 30'000 hours of music recordings. Out of these, roughly one third are live recordings which Ehrenreich, a regular concert and opera goer, mostly made himself; roughly another third are bootlegs of radio broadcasts, while the last third is made up of recordings that have been copied over from commercial sound carriers (from LPs and magnetic tapes, and in later years also from CDs).

In addition to the main collection, Ehrenreich's legacy in the possession of the Bern University of Arts comprises a series of further items that so far have only been partially inventoried:

- Reel-to-reel tapes (so-called "master tapes") and Digital Audio Tapes (DAT), which Ehrenreich used in the recording process.
- Commercial sound carriers (magnetic tapes and CDs).
- Video tapes.
- A reference library, mainly focusing on opera (monographs, periodicals, annual programs, etc.)
- Ehrenreich's recording equipment (several generations of equipment, covering several decades).
- A few items of personal correspondence as well as a photograph collection covering Ehrenreich's travels (unrelated to opera).

1.1 Years covered and uniqueness of the recordings

Most of the live bootleg recordings have been made between 1965 and 2010, with some exceptions predating this period. Even if some of the performances may in addition have been recorded officially with the aid of professional equipment, these recordings are unique from a research perspective as they have not undergone any post-processing.

Most of the radio bootlegs are from the period 1935 to 2010. Many of them were made from live broadcasts. It has not been researched yet to what extent these recordings are being preserved and potentially made available by the institutions in charge of the respective radio archives.

Commercial releases represented in the collection cover the period 1895 to 2010. Most of them can still be acquired commercially. It can be assumed that these recordings are preserved elsewhere. The early recordings in the collection stem from collectors' LPs and CDs; to our knowledge, Ehrenreich's collection does not contain any data carriers predating the 1950s.

1.2 Digitization and preservation

Digitization of the live bootleg recordings has started in 2018, a process that will span several years. Measures to ensure long-term preservation of both the digitized versions and the original tapes are currently evaluated and are expected to be in place by 2020.

Quite unsurprisingly, material deterioration of the reel-to-reel tapes, which have an expected life span of approx. 30 years, has already set in. For this reason, digitization of the most valuable tapes is currently given the highest priority.

1.3 Copyright and access policy

Most of the live recordings and radio broadcasts are covered by copyright. Under the United States fair use doctrine, they may be used for educational and research purposes. Tapes that have already been digitized are thus being made available to researchers and educators. Recordings that are not available commercially and that have not yet been digitized may be digitized on demand.

2 The Catalogue

Ehrenreich himself has kept two catalogues: the first one organized according to **composers**, and the second one according to **performers** (singers and conductors). Both catalogues were originally maintained in a word processing format, the first one comprising 4816 pages, and the second one 11581 pages. In the meanwhile, the performers' catalogue has been made available in a spreadsheet format (approx. 115'000 lines) and serves as the main catalogue of the collection. Transformation of the composers' catalogue is not envisaged, as it does not contain any additional data.

Both original catalogues as well as the catalogue in spreadsheet format have been made available online, free of copyright for anyone to use[1]. The catalogue in spreadsheet format is available in two versions:

- **A full version in Excel format.** This version corresponds to the master version that is maintained by the Bern University of Arts. It contains not only the (partially) cleansed data, but also the original data as it was extracted from Ehrenreich's catalogue as well as some error reports generated during data transformation.

- **A working file in Excel and CSV format.** This version should be appropriate for most users of the collection. It contains only the current version of the data.

2.1 Data quality

Note that the data has only been partially cleansed and extracted into the appropriate columns. There are at least three sources of error in the data:

- When entering data in his catalogues, Ehrenreich has made a number of typos and mistakes – which is not very surprizing as, working with a word-processor, he had hardly any possibility to systematically check the catalogue entries for consistency.

- While Ehrenreich maintained his catalogues in a layout that was rather easy to transform to a spreadsheet format, some ambiguities and inconsistencies remained, resulting in some data not being properly transformed.

- After transforming the data to the spreadsheet format, we have started to cleanse the data and to extract some information that was originally provided in one text string into separate fields. On one hand, this job has not been completed yet (some further data cleansing and data extraction could still be carried out mainly by triangulating between the various entries in the original catalogue); on the other hand, we may also have introduced some of our own errors in the process.

A further possible source of error may stem from inaccurate data in the original sources used by Ehrenreich. So far, we have no indication that this might be a major issue in the case of this collection; in fact, it can be assumed that Ehrenreich triangulated between various sources when entering the data. There are however some indications in the catalogue that would suggest that in a few cases he was not entirely sure where exactly the recordings on the tapes were from. Our impression so far is that he has worked quite carefully and that problematic cases have been highlighted in the catalogue whenever he had any doubts.

2.2 Complementary meta data

Each of the 2260 tape boxes belonging to the main collection contains not only a reel-to-reel tape, but also further information concerning the recordings on the tape: comprehensive contents and cast

[1] The files are available at: https://old.datahub.io/dataset/ehrenreich-collection-database

lists, (excerpts of) programme booklets and paper clippings (see the cover photo of this report for an example). Along with the two largely redundant catalogues, this additional material ensures quite comprehensive documentation of the recordings in Ehrenreich's main collection. By triangulating between these different data sources and by comparing them to further sources available on the Internet, the data quality and completeness of the catalogue will be gradually improved.

To facilitate this process, the additional material contained in the tape boxes is being digitized and made available alongside the recordings. Unfortunately, due to copyright restrictions, we are unable to make digitized copies of the additional material freely available on the Internet. To limit possible damage to the tapes through additional handling, the additional material is digitized at the same time as the recordings.

2.3 Data fields

The table below contains all the data fields of the catalogue (the fields with a grey background are only included in the full version of the catalogue).

Label	Description	Comment
lnb	line number	When sorted on this field, the catalogue entries are in exactly the same order as in Ehrenreich's original performers' catalogue. *N.B.: In some rare cases, extra lines had to be added to the catalogue after the data transformation; in this case, decimals were used to place the additional lines into the right order.*
_ide	tape identifier (original entry)	This field contains the unmodified, original entry for the tape identifier.
ide	tape identifier	This field contains the tape identifier. In the case of catalogue entries with only one or several cross-references instead of a tape identifier, the first cross-reference was used as tape identifier (whether this practice is justified in all cases still needs to be established). *N.B.: The digitized recordings from Ehrenreich's main collection are currently stored in folders named according to this tape identifier.*
ide_?	unclear entry for type identifier	A question mark in this field signals an unclear entry for the type identifier (in the original file).
ide_LetterCode	tape identifier (letter code)	This field is used to sort the catalogue on the tape identifier field; it does not contain any additional information.
ide_Num	tape identifier (number)	This field is used to sort the catalogue on the tape identifier field; it does not contain any additional information.
ide_Letter	tape identifier (final letter)	This field is used to sort the catalogue on the tape identifier field; it does not contain any additional information.
ref	cross-reference	This field contains cross-references to other tapes as provided by Ehrenreich.
digi	digitization status	This field contains the string "digitized" if the given tape has already been digitized.
track	track and channel	This field contains the information on which track and channel the recording has been stored on the reel-to-reel tape. The digital files are named accordingly. The digitization process typically results in 4 files per tape, as each of the two tracks has two channels that were written to.

		This information is not available yet within the Excel file. The information is typically apparent from the image of the tape and/or the documentation that goes with the tape. The respective scans are made available alongside the digitized recordings.
position	start position	This field contains the start position (hh:mm:ss) of the recording within the digital file. *This information has not been provided yet.*
length	length	This field contains the length (hh:mm:ss) of the recording. *This information has not been provided yet.*
art	artist name	This field contains the name of the artist, mostly in the format "FAMILY_NAME, FIRST_NAME". Some entries contain the name of an artists' collective.
voc	voice / performer type	This field contains the original entry for voice type / performer type. It serves documentation purposes; the information has been extracted to other data fields.
voice_type	voice type	This field contains singers' voice type.
gender	gender	This field contains singers' gender (derived from the voice type indication).
pft	performer type	This field indicates the artist's role in the performance (e.g. conductor, singer, musician, narrator). In some rare cases, the role given is "composer".
arm	artist remark	Some 4000 entries contain a remark regarding the artist (e.g. pseudonym used; voice type different from the artist's usual voice type; singer acting at the same time as conductor).
com	composer name	This field contains the name of the composer, mostly in the format "FAMILY_NAME". Sometimes, the first name or the initials are given for disambiguation purposes (e.g. "J.S. Bach"). Where the composer is unknown, this is also indicated. In this case, the entry may point to the origin of the composition (e.g. "AMERICAN", "AMERICAN FOLK", "AMERICAN POPS", "ANDEAN FOLK", "SPANISH CHURCH", "SPANISH RENAISSANCE", etc.). Some entries contain the names of two composers.
opr	opera	This field contains the name of the opera, the song cycle, or some other work comprising multiple parts. In some cases, it may also contain the name of an individual song if no complex work is indicated. If the name of an opera or other work with multiple parts is given in CAPITAL LETTERS, this indicates that there is at least one recording of the full work with the artist in question in the Ehrenreich collection. If the catalogue is put in its original order, the recordings of full works are always listed before the recordings of the parts of the work in question (e.g. individual arias, songs, etc.).
rol	role	This field contains the name of the character role. If a singer sang several roles, the different roles are separated by a semicolon. Sometimes, the number of roles is indicated without naming them, or there is an indication that the singer in question filled-in several roles.
act	act	This field contains information regarding the act(s) and sometimes the scenes that were included in the performance / recording.

sng	song or aria	This field contains the title of the song(s) or aria(s) included in the performance / recording. Sometimes, reference is made to scenes, without naming any specific arias. Individual songs / arias are delimited by a semicolon.
sng_LEN	length of the sng field	This variable can be used to sort the entries by the length of the entry for song or aria.
city	city	The city of the performance, as given in the performance field. Extraction of the data has not been completed yet.
venue	venue	The venue of the performance, as given in the performance field. Extraction of the data has not been completed yet.
ensemble	ensemble	The ensemble involved in the performance, as given in the performance field. Extraction of the data has not been completed yet.
event	event	The performance event, as given in the performance field. Extraction of the data has not been completed yet.
broadcaster	broadcaster	The broadcaster involved in the recording/broadcasting of the event, as given in the performance field. Extraction of the data has not been completed yet.
programme	programme	The radio programme on which the performance was broadcasted, as given in the performance field. Extraction of the data has not been completed yet.
per	performance	This field contains the further specification of the performance, the recording, and/or the broadcast.
nbr	number of performances	This field contains the number of performances of the same work by the same artist listed in the catalogue.
lan	language of performance	This field contains the language of the performance / recording.
wit	co-performer	This field contains the co-performers involved in the performance / recording.
liv	live recording	This field contains the string “Live” if there is an indication in the performance field that the recording was made during a live performance or from a live broadcast.
cdr	recording from CD	This field contains the string “CD” if there is an indication in the performance field that the recording has been copied from a commercial CD.
lpr	recording from a phonograph record	This field contains the string “LP” if there is an indication in the performance field that the recording has been copied from a commercial disk record (LP, acetate, etc.). It contains the string “Cylinder” if the original recording had been distributed on a phonograph cylinder.
vid	recording from Video	This field contains the string “Video” if there is an indication in the performance field that the recording has been copied from a commercial video tape.
bca	recording from a broadcast	This field contains the string “BC” if there is an indication in the performance field that the recording has been made from a broadcast.
pfr	pf recording	This field contains the string “PF” if there is an indication in the performance field that the singers were accompanied by a pianoforte.
yer	year of recording	This field contains the year in which the recording was made. In case the recording stretches over several years, this field contains the first year.
yer_end	last year of recording	In case the recording stretches over several years, this field contains the last year. If the year indication is

		imprecise (e.g. “1950s”), this field contains the last year of the interval (e.g. “1959”).
yer_release	year of release	This field contains the year in which the recording was commercially released.
datMM	month of recording	This field contains the “month” element of the date on which the recording was made.
datDD	day of recording	This field contains the “day” element of the date on which the recording was made.
dat	date of recording	This field contains the date of the recording in YYYY-MM-DD format, with varying precision (year, month, or day). In case the recording stretches over several dates, this field contains the first date. The field assembles the information from the “yer”, “dataMM”, and “datDD” fields; it does not contain any additional information.
dat_end	last date of recording	In case the recording stretches over several dates, this field contains the last date. If the date indication is imprecise (e.g. “January or February 1949”), this field contains the end date of the interval (e.g. “1949-02”).
dat_bdcst	date of broadcast	Date of the broadcast.
Helvetica	helvetica	This field contains the last names of Swiss artists who were involved in the recording (as composer, conductor, or singer).
not	note	This field contains additional notes regarding the entry.
era	error artist	remainders of the artist field (generated during the data transformation process)
erc	error composer	remainders of the composer field (generated during the data transformation process)
ero	error opera	remainders of the opera field (generated during the data transformation process)
erp	error performance	remainders of the performance field (generated during the data transformation process)
ers	error song	remainder of the song or aria field (generated during the data transformation process)
erx	error on remaining	remainders of the remaining fields (generated during the data transformation process)
ora	original of artist	original artist field (extracted automatically during the data transformation process)
orc	original of composer	original composer field (extracted automatically during the data transformation process)
oro	original of opera	original opera field (extracted automatically during the data transformation process)
ors	original of song	original song or aria field (extracted automatically during the data transformation process)
orp	original of performance	original performance field (extracted automatically during the data transformation process)

Table 1: Data fields

2.4 Tape identifiers

Identifiers of reel-to-reel tapes are made up of three parts: A letter code, an integer, and optionally a single letter added for disambiguation purposes (example: OP-46B). There is a small number of catalogue entries that refer to CDs, LPs, or cassettes that do not follow this pattern. The table below contains a list of the letter codes used by Ehrenreich, along with the number of tapes per letter code. It gives an idea how the collection has been organized physically and how large the different sections are compared to one another.

Tape Letter Code	Label	Number of tapes
Bar-	Barock	127
CL- *(KL-)*	Carl Loewe / Karl Loewe	1
Country- *(CM-)*	Country Music	3
EAO-	English & American Opera	81
EIO-	Early Italian Opera	14
FC- *(FAS-)*	French Chanson / French Art Songs	9
FO-	French Opera	201
Folk- *(F&P-)*	Folk Music (Folk & Pop)	17
fs-	Franz Schubert	30
G&S-	Gilbert & Sullivan	19
GD-	Gaetano Donizetti	88
gm-	Gustav Mahler	15
GO-	German Opera	94
GP-	Giacomo Puccini	113
GR-	Gioachino Rossini	110
GV-	Giuseppe Verdi	240
hw-	Hugo Wolf	6
Int-	(?)	22
jb-	Johannes Brahms	8
LR-	Lieder Recitals	117
LVB-	Ludwig van Beethoven	3
MC-	Musical Comedy	16
MGL-	Miscellaneous German Lieder	6
MGO-	Modern German Opera	36
ML-	Miscellaneous Lieder	23
OP-	Opera Potpurri	108
Pops-	(?)	17
RO-	Russian Opera	71
RS-	Richard Strauss *(see also Robert Schumann)*	60
rsc-	Robert Schumann *(see also RS-)*	6
RW-	Richard Wagner	161
SLO-	Slavic Opera	33
SO-	Spanish Opera	15
VB-	Vincenzo Bellini	49
Ver-	Verismo	79
VO-	Viennese Opera	24
VR-	Voice Recitals	171
WAM-	Wolfgang Amadeus Mozart	107
Total		**2261**

Table 2: Tape letter codes

3 Contact

Bern University of the Arts

Laura Moeckli
Email: laura.moeckli@hkb.bfh.ch