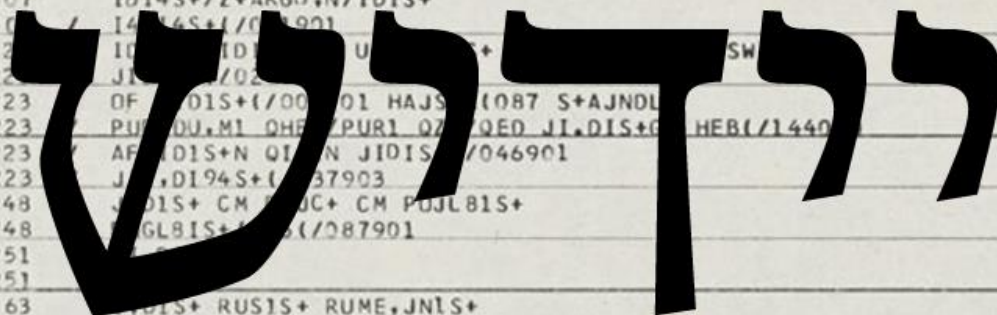


PAGQUND INTVN R *** RESPONSES ***

Digitizing 140,000 Pages of Yiddish Linguistics: Strategies and Lessons Learned



Michelle Chesner
Norman E. Alexander Librarian for Jewish Studies
Columbia University

Association of Jewish Libraries Proceedings

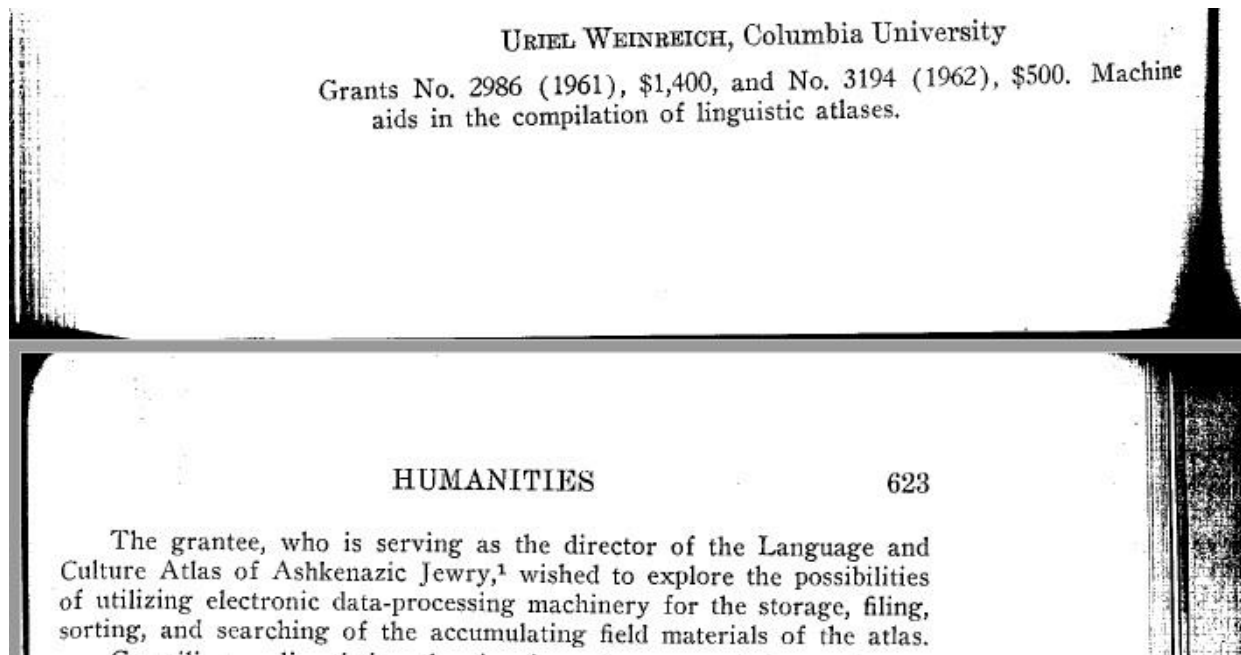
June 2019

0010200	47197	/	JID+		
0010200	47266	/	JID3S+		
0010200	47266	/	JID3S+		
0010200	47287	/	E53R DT G3HE5RT(/017100		
0010200	47287	/	ID 01 JID(/0010131302		
0010201	47287	/	ID1S+(/197903		
0010201	47287	/	ID1S+(/185903		
0010201	47287	/	DFID1S+(/169900		
0010201	47287	/	DFI,DIS+(/133901		
0010200	47301	/	9 ID14S+ CM G0JIS+(/001040		
0010201	47301	/	*(/141020(/001050		
0010200	47307	/	ID14S+/Z+ARGO,N/IDIS+		
0010201	48223	/	1414S+(/001901		
0010200	48223	/	ID1S+(/001901		
0010201	48223	/	JID1S+/02		
0010201	48223	/	DF ID1S+(/00101 HAJSD (087 S+AJNDL		
0010201	48223	/	PUR DU,M1 QHE PURL 02 QED JI,DIS+0 HEB(/1440		
0010203	48223	/	AF ID1S+N QIN N JIDIS (/046901		
0010203	48223	/	JID194S+(/037903		
0010200	48248	/	ID1S+ CM GUC+ CM P0JL81S+		
0010203	48248	/	GL81S+ (/087901		
0010200	48251	/			
0010200	48251	/			
0010200	48263	/	ID1S+ RUS1S+ RUME,JN1S+		
0010200	48272	/	Z+3RGO,N(051 QLIT/JID1S+ QCOLLOQ		
0010201	48272	/	PAS+1 QN1S+K81NZ1RGON1S+ VORT(861(/145010		
0010201	48272	/	ID1S+(/122902		
0010201	48272	/	ID1S+(/175910		
0010201	48272	/	ID1S+(/006901		
0010202	48272	/	Z+3RGO,N		
0010200	48274	/	I,DIS+/RUS1S+ (/0010101S/001 G0JIS		
0010200	48276	/	NOR ID1S+		
0010200	48284	/	JID3S+ (/IRE,IS+ QINT31GENCJ3		
0010200	48286	/	Z+3RGO,N		
0010201	48286	/	141,014S+(/044909		
0010201	48286	/	RUS+IS+(/044907		
0010202	48286	/	MIRT8IJ1NUNJEN1M(/0010101S		
0010200	48288	/	ID1S+		
0010201	48288	/	VESL QF LB-ODK3(208 QNIT JID1S+(/005050		
0010200	48296	/	ID1S+ (/KBA,N381,01S+		
0010201	48296	/	ID1S+ (/0010101S		
0010202	48296	/	AJID2(002(/004902(302		
0010200	48309	/	ID1S+ CM G0JIS+(/001040		
0010200	48311	/	ID1S+		
0010201	48311	/	*ANU,R1M3N8 ID3S+ (256 (/0010101 (/129050(898//IN		
0010202	48311	/	1NAZA, ID1S+OJZ(/055009(/035902(/111080		

Background

"Data ages like wine, software ages like fish"
(attributed to Andrew Hankinson)

American Philosophical Society Grant:



Digital Humanities, circa 1961

Description of Project

Project: Language and Culture Atlas of Ashkenazic Jewry

Proposer: Professor Uriel Weinreich, Chairman, Department
of Linguistics

Planner: Uriel Weinreich and Marvin I. Herzog
(research assistant)

Computer: Mendel Hoffman

Dates: ~~beginning~~ February 27, 1961 to January 31, 1963

Magnitude: Number of cards - 200,000

Machines used: 026, 402, 519, 082, 077, 650

Number of hours: about 500

Results: An atlas of Yiddish dialects and of Jewish
folk-culture patterns of Central and Eastern
Europe, for correlations with language and
culture boundaries in the non-Jewish societies
of the same territories

Publications: The Atlas itself and numerous reports
and monographs on special subjects.

Method: Frequency distributions, symbol tables,
profile system tabulations.

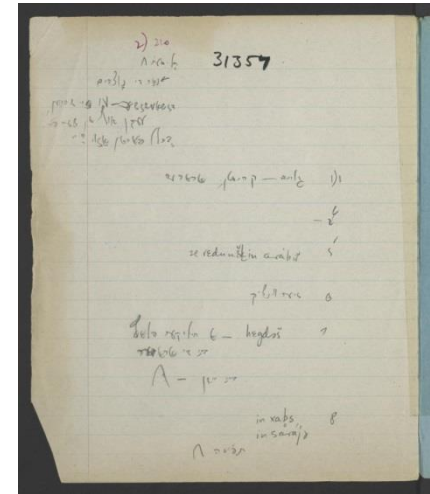
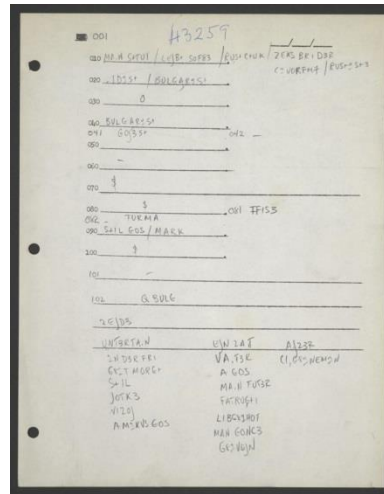
Data Collection Formats

Sound:

- Reel to reel audio

Transcription:

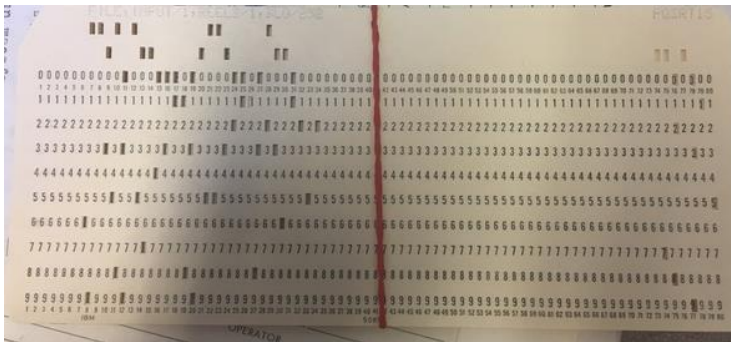
- Blue books
- Answer sheets
- Printouts



PAGE 721
LANGUAGE AND CULTURE ATLAS OF ASHKENAZIC
DATE OF

PAR	LOCRF	PAGQUNO	INTVN	R	***	RESPONSES
0400300		46338				PRIC+A,TK1S QRUSS/ QHENKIC+3S HENC+
0400300		46338				*C+3--...
0400300		47197				HENC+DX
0400300		47266				E44,NIC+IS(080
0400300		47266				E44,NIC+IS(080
0400300		47268				E44,NI4C+I4S
0400300		47287				Q1 ENC+K81S
0400300		47307				9 ENIC+KI4S*
0400300		47307				PRXXC+A,TKI4S Q RUS/ 1E,NI4C+KI4S

Reverse engineering old media



And reprocessing a ~300 box archival collection



Decisions, decisions

- “Scope creep”
- Fitting within internal infrastructure
 - Data v. image as data

Integrated into existing image viewer

COLUMBIA UNIVERSITY LIBRARIES

Log In Help

THE LANGUAGE AND CULTURE ARCHIVE OF ASHKENAZIC JEWRY

Start a new search...

BACKGROUND MATERIAL ABOUT THE PROJECT USER GUIDE MAP SEARCH

ABOUT THE LANGUAGE AND CULTURE ARCHIVE OF ASHKENAZIC JEWRY

The Language and Culture Archive of Ashkenazic Jewry (LCAAJ), an extraordinary resource for research in Yiddish studies, consists of 5,755 hours of audio tape field interviews with Yiddish speaking informants collected between 1959 and 1972 and ca. 100,000 pages of accompanying linguistic field notes. The Archive does not include transcriptions of the interviews.

lc.library.columbia.edu/lcaaj#

BROWSE THE COLLECTION

NAME

REGION

Decisions, decisions

- “Scope creep”
- Fitting within internal infrastructure
 - Data v. image as data
- Information overload (no “custom” option)

Dedicated Libguide for all the additional details

LCAAJ Digital Archive User Guide: Introduction

This is a guide to using the Digital Archive to the Language and Culture Atlas of Ashkenazic Jewry (dlc.library.columbia.edu/lcaaj)

Introduction	Overview of the Collection	Questionnaire	Interviewee (Location) Data	Answer Data	Audio tapes and recordings	Using the Data	Appendices
Glossary	Bibliography of Works citing the LCAAJ	Acknowledgements	Copyright Statement				

Introduction

This is a supplementary guide to the [Digitized data for the Language and Culture Atlas](#) of Ashkenazic Jewry, which makes available online the linguistic and ethnographic information about Yiddish, gathered by the LCAAJ team so that researchers could map or otherwise systematically analyze it. (A history and fuller description of the LCAAJ project can be accessed by [clicking here](#).)

For information on how to search the Digitized LCAAJ, see [Searching for Answers to Questions](#).

Before beginning your search, we strongly suggest that you look at the page providing an [overview of the collection](#), and look at the important [scope note](#) below.

What the digitized LCAAJ website does:

- Provides access to the images of the pages with answers given in interviews in the form of [Blue Books](#), [Answer Sheets](#), and [Printouts](#)
 - Further information about these materials can be found in this guide under [Answer Data](#)
- Provides *filters* to allow one to limit their search by format or location/region
- Provides detailed metadata for those answers to allow ease of searching
- Provides persistent URLs and IIIF manifests to allow use of the images in other contexts

What the digitized LCAAJ website **DOES NOT** do:

- Provide an interface for digitally mapping data (for that, we recommend taking a look at the [LCAAJResearch](#) page, as well as joining our community of LCAAJ scholars by sending us an email at lcaaj@library.columbia.edu)
- Provide text-searchable answers to the questions (for that, see [LCAAJResearch](#) for access to downloadable PDFs of the printouts with OCR, and to read more about our plans on gathering that information as a community).

PAGE_ID	INTV	R	RESPONSES
001000	0000	0000	0000
001001	0001	0001	0001
001002	0002	0002	0002
001003	0003	0003	0003
001004	0004	0004	0004
001005	0005	0005	0005
001006	0006	0006	0006
001007	0007	0007	0007
001008	0008	0008	0008
001009	0009	0009	0009
001010	0010	0010	0010
001011	0011	0011	0011
001012	0012	0012	0012
001013	0013	0013	0013
001014	0014	0014	0014
001015	0015	0015	0015
001016	0016	0016	0016
001017	0017	0017	0017
001018	0018	0018	0018
001019	0019	0019	0019
001020	0020	0020	0020
001021	0021	0021	0021
001022	0022	0022	0022
001023	0023	0023	0023
001024	0024	0024	0024
001025	0025	0025	0025
001026	0026	0026	0026
001027	0027	0027	0027
001028	0028	0028	0028
001029	0029	0029	0029
001030	0030	0030	0030
001031	0031	0031	0031
001032	0032	0032	0032
001033	0033	0033	0033
001034	0034	0034	0034
001035	0035	0035	0035
001036	0036	0036	0036
001037	0037	0037	0037
001038	0038	0038	0038
001039	0039	0039	0039
001040	0040	0040	0040
001041	0041	0041	0041
001042	0042	0042	0042
001043	0043	0043	0043
001044	0044	0044	0044
001045	0045	0045	0045
001046	0046	0046	0046
001047	0047	0047	0047
001048	0048	0048	0048
001049	0049	0049	0049
001050	0050	0050	0050

Decisions, decisions

- “Scope creep”
- Fitting within internal infrastructure
 - Data v. image as data
- Information overload (no “custom” option)
- Ongoing work
 - Divisional priorities (DLST v. DS)

Future Plans (pending..)

- Time-based media
- [Creative reuse](#) (and feedback)
- Matching audio to transcript
 - Time-based media project
- Archives as window into the world of early Yiddish in the US
- Classroom potential
- Exhibit

