

Apparat: A Computer Cataloguing System for Sound Recordings

Fiona Campbell

Department of Sound Records,
Imperial War Museum, London

In the last issue of *Archives and Manuscripts*, David Lance described the archive holdings of the Department of Sound Records in the Imperial War Museum. This article describes the computer system which has been adopted to assist in the preparation of finding aids to the archive's oral history collection. The nature of sound recordings inhibit the 'browsing' possible among more visual reference materials such as documents, photographs and particularly books and, therefore, to be objective, their finding aids must provide more detail — particularly concerning the contents of the recordings — than is required for other sources.

The finding aids which were developed in the first few years of the Department's existence consisted of catalogues and indexes. For each oral history interview carried out, a catalogue card was filled in with details of the informant, the circumstances of the interview and a synopsis of its contents; the completed cards remained within each informant's file. At the conclusion of each project, the synopses of all the interviews within it were gathered together to form a small catalogue which was printed and made available for sale.

The catalogued information was complemented by indexes of various topics mentioned in the interview. These fell into two categories: indexes of proper names — persons, places, units, aircraft and ships — and a subject index. Each entry comprised the keyword, interview accession number and a short summary of the context of the topic.

In 1976 it was decided to automate the preparation of finding aids for several reasons. Firstly, it was intended that the Department's holdings would be made fully available to the public in July 1977 and it was realised that the existing finding aids would not be adequate for the resulting needs. The catalogue cards could not be shown to the public as they contained private information concerning the informants and they

were sometimes completed inconsistently, were not always up-to-date and, as loose records, were prone to loss and damage. The printed catalogues were extremely time-consuming to prepare and expensive to produce in the short runs necessitated by our small number of customers and there was no central authority where information required quickly — such as the recordings' contractual restrictions — could be found. The new service which the Department intended to offer demanded that an up-to-date catalogue of the collection would be available for departmental use and selections of it available for public consultation.

Secondly, it was felt that just as our archive master tapes are stored away from the Museum as a precaution against the possibility of fire, flood or even a bomb attack (which the Museum has suffered from in recent years) so physical security should be provided for the finding aids which facilitate access to these recordings. This could be achieved by the creation of a database of all relevant information for permanent storage in a computer located away from the Museum.

The third reason — and, given the financial stringencies governing the Museum's work, by no means the least — was the availability of a suitable and relatively cheap computer system. In 1974 the Museum's Department of Information Retrieval — which is responsible for providing assistance in the cataloguing and indexing work of the collecting departments — had, in association with systems analysts, set about designing a computer programme to monitor the complicated conservation needs of the Museum's archival collection of films. As a result APPARAT (Archive Preservation Programme and Retrieval by Automated Techniques) was developed to run on an ICL 1906A computer located in the Department of Education and Science's ADP unit in Darlington, County Durham. APPARAT comprised two sub-systems: the first concerned with the collection's conservation needs and the second with its subject content; the latter could be easily adapted to meet the needs of a collection of sound recordings. As APPARAT was a purpose-built system, tailored to meet the needs of a specialised archival collection with its large free text areas and generous field-lengths, its adoption did not involve the compromises in detail and quality which have been forced on many archival institutions attempting to adapt bibliographic computer programmes to their needs.

However, not only was there a computer system available but also several features of our oral history collection made its finding aids particularly suitable for conversion to a computer format. These included the planned nature of the collection, which is organised on a project basis, the technical similarities of the medium and the standardised language. Further, the unit for computerisation — the interview — is easily identifiable. These features must be the envy of archivists currently involved in projects to computerise the finding aids to their more traditional material and undoubtedly struggling with

APPARAT		INTERVIEW DATE: 1997	
APPROVAL NUMBER	00079441		
series or project title	(50) 2491	BATTAL MANAGEMENT IN THE SPANISH CIVIL WAR	
purpose of interview	(50) 2491	BATTAL MANAGEMENT	
type	(550) 2491	OPERATIONS	
reference or initials	(550) 2491	1st B	
supplementary title	(160) 2491		
revision number	(550) 2491	ONE	
date of birth and/or death	(2012) 1897	1 11 1907	
highest rank in period covered	(542) 1898	Battalion Commander	
nationality	(515) 1897	BRITISH	
history of interview	(5165) 1897	British volunteer serving with the 10th International Brigade - in the Spanish Civil War as a Battalion commander 1937-38	
dates covered	(19) 1897	1937-1938	
interviewer	(530) 1897	TIA Kootenai, P.T.C.	
original medium	(20) 1897	Tape 7-50 1897	
reference	(4) 1897	TWA	
country of origin	(520) 1897	GB	
date recording completed (dd-mm-yy)	(8) 1897	01 01 97	
additional recording dates	(14) 1897	01 26 97	
location recorded	(575) 1897	London	
birth language	(525) 1897	English	
secondary language	(515) 1897	RO2, EN	
evaluating quality	(530) 1897	100%	
nationality information	(5200) 1897	100%	
contacting company	(5150) 1897	100%	
number of reels	(2) 1897	1	
track control format (ms, frames, stereo etc)	(6) 1897	110	
running time	(4) 1897	02	

INTERVIEW DATE: 1997
APPROVAL NUMBER: 00084441

series or project title	(50) 2491	BATTAL MANAGEMENT IN THE SPANISH CIVIL WAR	
purpose of interview	(50) 2491	BATTAL MANAGEMENT	
type	(550) 2491	OPERATIONS	
reference or initials	(550) 2491	1st B	
supplementary title	(160) 2491		
revision number	(550) 2491	ONE	
date of birth and/or death	(2012) 1897	1 11 1907	
highest rank in period covered	(542) 1898	Battalion Commander	
nationality	(515) 1897	BRITISH	
history of interview	(5165) 1897	British volunteer serving with the 10th International Brigade - in the Spanish Civil War as a Battalion commander 1937-38	
dates covered	(19) 1897	1937-1938	
interviewer	(530) 1897	TIA Kootenai, P.T.C.	
original medium	(20) 1897	Tape 7-50 1897	
reference	(4) 1897	TWA	
country of origin	(520) 1897	GB	
date recording completed (dd-mm-yy)	(8) 1897	01 01 97	
additional recording dates	(14) 1897	01 26 97	
location recorded	(575) 1897	London	
birth language	(525) 1897	English	
secondary language	(515) 1897	RO2, EN	
evaluating quality	(530) 1897	100%	
nationality information	(5200) 1897	100%	
contacting company	(5150) 1897	100%	
number of reels	(2) 1897	1	
track control format (ms, frames, stereo etc)	(6) 1897	110	
running time	(4) 1897	02	

INTERVIEW DATE: 1997	APPROVAL NUMBER: 00084441
series or project title	(50) 2491
purpose of interview	(50) 2491
type	(550) 2491
reference or initials	(550) 2491
supplementary title	(160) 2491
revision number	(550) 2491
date of birth and/or death	(2012) 1897
highest rank in period covered	(542) 1898
nationality	(515) 1897
history of interview	(5165) 1897
dates covered	(19) 1897
interviewer	(530) 1897
original medium	(20) 1897
reference	(4) 1897
country of origin	(520) 1897
date recording completed (dd-mm-yy)	(8) 1897
additional recording dates	(14) 1897
location recorded	(575) 1897
birth language	(525) 1897
secondary language	(515) 1897
evaluating quality	(530) 1897
nationality information	(5200) 1897
contacting company	(5150) 1897
number of reels	(2) 1897
track control format (ms, frames, stereo etc)	(6) 1897
running time	(4) 1897

problems such as provenance, non-standard language and inequalities in the length, size and content of documents.

APPARAT

The distinctive feature of APPARAT is that it produces several finding aids from a single database. It is an off-line system because the relatively small use of our collection does not warrant the expense of a direct link with the computer.

Input

The database comprises the total number of catalogue records within it. A catalogue record is created for each interview unit by inputting information on the three pre-printed catalogue sheets reproduced in Figure 1. The amount of this information can include all the fields or only a few of them — providing the accession number is included — thereby permitting the creation of skeleton records for housekeeping purposes which can be completed at a later date.

The fields provide details on five categories of information:

Interviewee: name, nationality, decorations, dates of birth and, where applicable, death and a short free text career record. Additional information about the interviewee can be entered in the "Notes" field.

Interview: accession number, project title, language, recording dates and interviewer. Technical details include duration of recording, number of reels and public service listening copies transcript, original medium, track configuration and recording quality.

Housekeeping: copyright and contractual restrictions, sponsoring and production organisation, associated material, cataloguer and date catalogued.

Synopsis: free text area summarising the contents of the interview. There is no fixed field limit imposed by the computer as additional sheets may be used although, in order to produce an effective finding aid, we restrict the word limit to between 50-75 words for a 30 minute reel. The synopsis — which is intended to contain objective information only — is supplemented by the "Remarks" field where interviewers or cataloguers are able to give vent to their personal opinions of the interviewee or the interview.

Index terms: dates, areas, topics. This facility is not currently being used.

The completed sheets are sent to the OCR bureau at the ADP unit in Darlington where the information is processed into a machine-readable form on paper tape and fed into the computer.

Output

APPARAT has the capacity to produce the following outputs: a comprehensive catalogue, ad hoc lists, indexes and printed catalogues.

Information can be output in the following formats: upper and lower case line print, computer typeset and Computer Output Microfiche (COM).

Catalogue

This is the most important output and comprises all the information on sound records so far entered into the system. It is arranged in accession number order and produced on COM. A specimen catalogue entry is reproduced in Figure 2. The same catalogue format with the omission of information of a confidential nature is also produced on COM for the use of public visitors.

Listings

APPARAT is able to compile ad hoc lists by conducting retrospective searches. These are carried out by asking questions positively or negatively such as "Nationality must be British" or "Production company must not be BBC". In numerical fields, information can be requested by specifying minimum or maximum values such as "Date recording completed must be later than 01 0 1977". Up to five targets can be listed in a single search and can either be given as cumulative stipulations, as acceptable alternatives or as a combination. Thus "Interviews must have access unrestricted and English language and cover the 1914-18 period" or "Highest rank must be Field Marshal or General". The resulting information can then be sorted according to criteria which need not have anything to do with the search requests.

Indexes

The subject indexing facility on the APPARAT system is not used by our Department as it is unable to provide the kind of detail essential for a sound collection. At the moment APPARAT is only able to produce a book-type index: a key-word followed by a list of accession numbers thus leaving it up to the user to discover if the context of the keyword is of interest to him. The improvement of APPARAT's indexing system would be a costly venture and so it is probable that the present manual system — which is a much more useful finding aid — will remain in operation.

Printed Catalogues

APPARAT is capable of producing printed catalogues from the database by using computer typesetting and it was hoped that these would negate the need for the expensive printed catalogues we currently produce. Unfortunately, a protracted dispute in the British printing industry prevented the early realisation of this goal. As this obstacle has now been removed, the Museum is again investigating the question of typeset output from APPARAT.

000375 0 ARCHIVE ID:A000187 I D6RLS S/TR:MONO R/T:009UMINS ACCESS: CASS:03 VETTED:Y
 WESTERNFRONTLIFEANDO SERIES:WESTERN FRONT: LIFE AND OPERATIONS 1914-1918

BAIN

Andrew Ramsay

colonel

COUNTRY OF ORIGIN:GB
 LOCATION:
 LANGUAGE 1 :English
 IWM CATALOGUE REFS: 001:WF 14/18
 1914-18

DATE:22 05 74 ADDIT. DATES:

TRANSCRIPT:F CASSETTE:C
 IWM CATALOGUE USE:Y
 REC QUALITY:

CREDITS 1 SPONSOR:IWM
 2 PRODUCTION COMPANY:IWM

INTERVIEWEE HISTORY:British company commander with infantry, worked as a staff officer 1916-18
 DECORATIONS/AWARDS:MC, DL BORN:3 XI 1892 DIED:19 V 1976
 HIGHEST RANK IN PERIOD DESCRIBED:Captain NATIONALITY:British
 DATES COVERED:1915-1918 INTERVIEWER:D G Lance ORIG:TAPE 7.5 IPS FULL M TRANSCRIPT:FULL

DETAILED SYNOPSIS DESCRIBING CONTENTS IN 015-MINUTE UNITS (1 CASSETTE HOLDS 2 UNITS)

- #01 1915 routine of trench warfare at Beaumont Hamel. Rum ration - type of meals in front line. Badly constructed British trenches - French Army using dead bodies for parapets.
- #02 Routine in reserve areas: cleaning equipment, training, recreation. Trench digging contest between 7th Argylls and Royal Engineers. Recounts verses of songs popular with Highlanders.
- #03 Comparison of Territorial with regular troops. Duties as assistant adjutant. Intelligence work. Sniping parties - success of illegal armour-piercing bullets. 1916 minelaying operations in La byrinth.
- #04 Dangers of tunnelling. Problems of mud and cold winter 1916-17. German machine gun barrage at Poelcapelle wiping out 9th Royal Scots. Staff work 1916-18. Organising battalion supplies in 154th Infantry Brigade.
- #05 Supply and salvage work. Allocation of reinforcements, troop movements. Organising troop billets during Champagne offensive. Responsibilities as commander of 152nd Brigade. Unexpected visit of Prince Albert.
- #06 Prince Albert riding pony. 1918 - British reaction to German tanks and flame-throwers.

CATLGR:LK DATE CATLGD:1075 SOURCE:
 CONTRACTUAL:UNRESTRICTED

COPYRIGHT:IWM

000376 0 ARCHIVE ID:A000376 I D4RLS S/TR:MONO

R/T:006UMINS ACCESS: CASS:02 VETTED:Y

 *J 03

Error Correction

Although errors in some fields are thrown up in the validation routine at the input stage, other errors can only be spotted by keen proof-reading of the computer print out. Errors are amended in two ways — in the large fields such as synopses, the computer can be instructed to replace single characters or symbols while in smaller fields corrections are effected by re-writing the entire entry.

Some problems have arisen in the operation of APPARAT. Initially, for example, with the intention of saving space both in computer files and on printout, the Museum opted for a non-standard convention for shift characters (to move text into and out of upper case) and for a 160-character line printout. When the Museum began using COM, it was found that these procedures caused technical difficulties and produced results of poor visual quality, so printouts were reformatted to industry standard (132 characters) and the shift convention was translated. Another problem arose from the expectation, when APPARAT was introduced for oral history cataloguing, that no interview would extend beyond three hours, for which it was thought a synopsis of 2500 character maximum length would provide adequate description. In effect the Department has conducted much longer interviews (the longest to date is 12 hours) and found that more text is sometimes required even for short interviews. A solution has been evolved — using up to 9 continuation synopses (each of the same 2500 character length) — but it is not totally satisfactory.

The conversion to an automated system has not resulted in the diminishing of our work load for, although a computer may be capable of conducting operations speedily and accurately, it does not have the intellectual capacity to recognise errors in the database. The resulting demand for a high degree of accuracy both at the input and proof-reading stages places an additional burden on the cataloguer. This problem was compounded in our case by our initial mistake in delegating to temporary staff the routine task of converting our manual records to computer input forms. Their unfamiliarity with the nature of our material resulted in many errors and their work had to be carefully checked and corrected before it could be entered into the computer. An additional burden of work has also risen at the proof-reading stage because the ICL 1906A is not a lower case machine and shift characters must be used to indicate shifts out of and into capital letters. Cataloguers are only aware of these characters at the proof-reading stage as the OCR bureau is responsible for the conversion of upper and lower case for computer input.

It is hoped that in the not too distant future the full potential of APPARAT will be realised by the production of printed catalogues by computer typesetting, the creation of an acceptable subject index and possibly the inclusion of a technical file giving additional information

about the recording. It is also intended that the input forms will be adapted for use with other parts of the department's holdings such as music and sound effects.

However, even without these features, the operation of APPARAT has resulted in several tangible benefits to the Department which justifies its expense in terms of both money and effort. It has provided us with the means to process information about the collection both quickly and accurately, with physical security to safeguard this information and with new finding aids to assist in the wider dissemination of our material. APPARAT's main success, however, has been that these benefits have not been gained at the expense of the finding aids themselves. The careful design of the system with its inclusion of generous field lengths and large free text areas has enabled us to continue providing the detailed information and descriptive synopses which characterised our manual system and are essential for effective finding aids to sound collections. The use of APPARAT has improved on our existing finding aids in terms of accuracy, consistency, comprehensiveness, speed and security without impairing the intellectual quality of their content.