

Title: The *Barcelona Historical Marriages Database (BHMD)*: a new research infrastructure for Historical Demography

Anna Cabré
Joana Maria Pujadas-Mora

Centre d'Estudis Demogràfics (CED)
Universitat Autònoma de Barcelona (UAB)

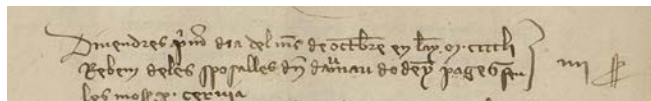
The *Barcelona Historical Marriage Database (BHMD)* is the main product of the project "*Five Centuries of Marriages*" - 5CofM (ERC-2010-Advanced Grant-269796, IP: Anna Cabré), launched in May 2011 and coming soon to an end. The project has been carried jointly by researchers of the Autonomous University of Barcelona (UAB), the Centre for Demographic Studies (CED) and the Centre for Computer Vision (CVC).

The BHMD compiles the information contained in 291 books of marriage licenses (*Llibres d'Esposalles*) conserved at the Barcelona Cathedral, that is, some 615.000 marriages celebrated between 1451 and 1905 in the 250 parishes of the Diocese, from the most rural villages in its periphery to the urban core of the city. Because the *Llibres d'Esposalles* were initially created as the instrument of a tax-collecting system on marriages, with different fee tiers according the socio-economic standing of the couple, they provide an exceptional insight on the social structures of the past defined by the own criteria of each period.

Names and family names of grooms (from 1876 two family names), names of brides (the inclusion of the bride in the licenses started in 1485) and their marital status (mostly for the brides and for the widowed grooms) are the common items of each record (see Figure 1). The brides' surnames were not registered until 1643 since women were not esteemed as individuals in their own right. Single brides were related to their fathers and widows to their late husbands. Parents' information (names) was registered quite often except for the period 1643 – 1750 and until 1643 whether they were alive or not at the time of their children's wedding. From 1715 onwards geographical location was recorded about the parish where the marriage was celebrated.

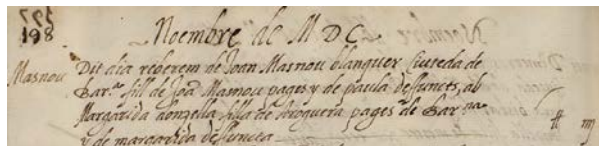
Figure 1: Examples of marriage licences.

1451



Friday, first of the month of October in the year MCCCCLI / It was received the marriage license of Dodeny Dalmau, peasant / priest Peter Cerny did it. IIII shillings.

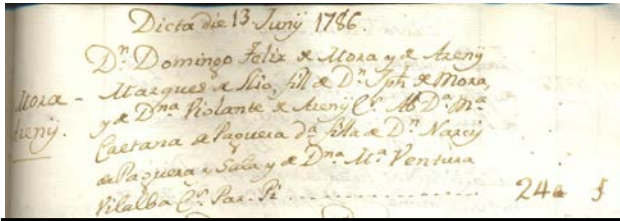
1600



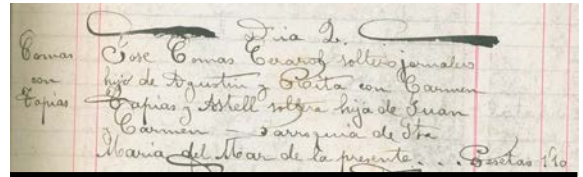
The same day, it was received from Joan Masnou, tanner citizen of / Barcelona son of Joan Masnou, peasant, and Paula, both were dead / with unmarried Margarida, daughter of Anguera, peasant from Barcelona / and Margarida, deaths. IIII shillings.

1786

1903



On said date of the 13 of June 1786
 Don Domingo Felix de Mora y de Areny / Marques of Llio son of
 Don Joseph de Mora / and of Dona Violante de Areny with Dona
 Maria / Caetana de Paguera daughter of Don Narcis / de
 Paguera y Sala and of Dona Maria Ventura / Vilalba Pi Parish /
 24 pounds

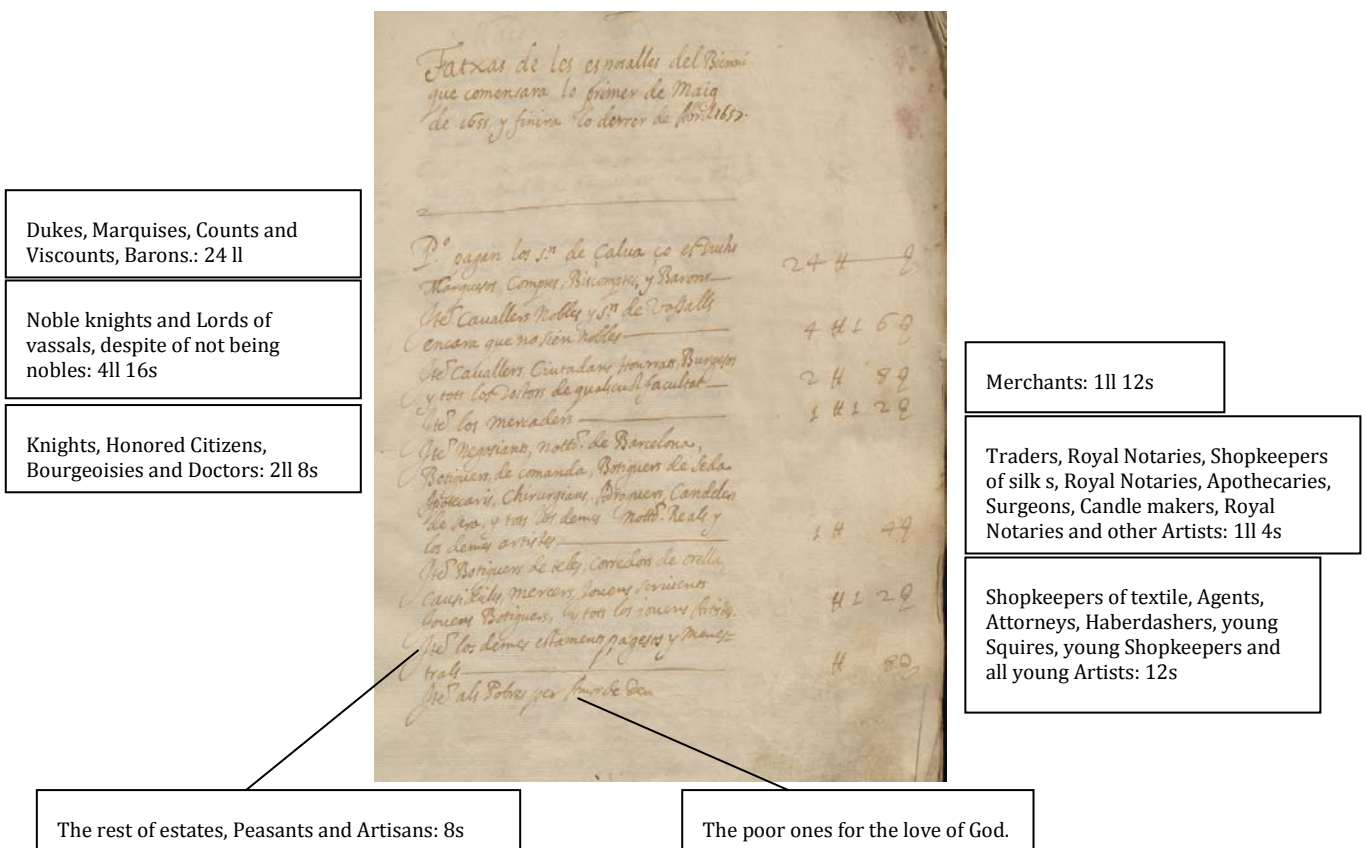


Day 2
 Jose Comas Gerard unmarried day laborer / son of
 Agustín and Rita with Carmen / Tapias y Astell
 unmarried daughter of Juan and Carmen - Santa / Maria
 del Mar Parish of this city / Pesetas 1'10

The fees mirror the whole social structure, from the nobility with the highest sum to those declared poor and exempt from tax (see Figure 2). They were fixed in a seven-tiered scale from 1575 until 1649: 1) Nobility, 2) Military Citizens, 3) Honoured Citizens (those who could hold public office), 4) Merchants, Lawyers, Physicians, 5) Master of Guilds, 6) Farmers and small artisans and 7) The poor.

From 1649 to 1857 an eight-tiered scale was found due to the inclusion of a level reserved for the merchants. After 1857 a seven-tiered scale re-emerged.

Figure 2: Example of fees (1655).



During five years, a research team combining historians, demographers and computer scientists have developed methods and instruments enabling the construction of the database by over 150 remote transcribers using a web-based crowdsourcing platform. Besides the experimentation on the collaborative transcription paradigm the platform main advantages are the centralized management of tasks and monitoring of work, assisting error detection and correction. For this reason the platform has two modules: administration module and transcription module.

The administration module allows us first to define the images (pages or volumes) to be transcribed, to register users and assigns tasks to users. Additionally, the administrator can keep track of the progress of the tasks and get statistics on the number of completed pages, the profile of the work activity per user. In the transcription interface the information is displayed in two panels (see Figure 3). In the top panel the page image is presented. This is an image browser panel, so the user can zoom in, zoom out, scroll down and scroll up to easily see the register that is currently transcribed. The user can browse pages to see previous or next ones in the book without losing the data already filled in. In the bottom panel the system displays two forms with the relevant information corresponding to a marriage.

Figure 3: Web-based crowdsourcing platform 5CofM project. Transcription module.

The screenshot displays the transcription interface for the 5CofM project. At the top, there is a navigation menu with options: EDIT, HELP, VOLUMES, USERS, STATISTICS, CONFIGURATION, PERSON, and EXIT. Below the menu is a toolbar with zoom and window adjustment options. The main area shows a scanned page of a handwritten marriage record. The text on the page is in Spanish and includes names like 'Juan', 'Bernat', 'Catherina', 'Magdalena', 'Pere Juan', and 'Gabriel Oliver'. The bottom panel contains a data entry form with the following sections:

- GROOM:** Form of address, Name (juan), Surname 1 (framura), Surname 2, Marital Status (dropdown), Occupation (pescador), Origin, Residence (barcelona), Widower of.
- GROOM'S FATHER:** Form of address, Name (bernat), Surname 1, Surname 2, Marital Status (dropdown), Occupation, Origin, Residence, Deceased (No).
- GROOM'S MOTHER:** Form of address, Name (catherina), Surname 1, Surname 2, Marital Status (dropdown), Occupation, Origin, Residence, Deceased (No).
- BRIDE:** Form of address, Name (magdalena), Surname 1, Surname 2, Marital Status (Donzella), Occupation, Origin, Residence, Widow of, Occupation, Residence.
- BRIDE'S FATHER:** Form of address, Name (pere juan), Surname 1 (montserrat), Surname 2, Marital Status (dropdown), Occupation (pescador), Origin, Residence, Deceased (No).
- BRIDE'S MOTHER:** Form of address, Name (elisabeth), Surname 1, Surname 2, Marital Status (dropdown), Occupation, Origin, Residence, Deceased (No).

At the bottom of the form, there are fields for 'Day' (agost), 'Year' (1617), 'Liquidation', 'Observations', and 'See Other pages'. The footer shows 'Volume: 69', 'Page: 13r', 'Numbers: 4', 'Parish: ', 'Tax: 4e (6)', and buttons for 'Save', 'Cancel', and 'Delete'.

The platform also includes functionalities of standardization and harmonization for names, surnames, occupations and geographical locations and record linkage. The data is transferred into a digital format compatible with spreadsheets that are also downloaded from this platform.

Presently undergoing a thorough revision, the BHMD is expected to be ready for use by the scientific community at the end 2016. At the time of the

celebration of the EPC-2016 at Mainz, in September, the authors will certainly be in the position of giving the practical details for the access and use of this new and unique research infrastructure.

Publications referring the BHMD

-Fornés, A.; Lladós, J.; Mas, J.; Pujadas-Mora, J.M.; Cabré, A. (2014) “A Bimodal Crowdsourcing Platform for Demographic Historical Manuscripts”. *DATeCH 2014*, Madrid, 103-108.

- Thorvaldsen, G; Pujadas-Mora, J. M.; Andersen, T.; Eikvil, L.; Lladós, J.; Fornés, A.; Cabré, A. (2015) “A Tale of two Transcriptions. Machine-assisted transcription of historical sources”. *Historical Life Course Studie*,2: 1-19.

-Villavicencio, F; Jordà, J. P.; Pujadas-Mora, J. M. (2015) “Reconstructing lifespans using historical marriage records of Catalonia from the 16th and 17th centuries”. In Bloothoof et al. (eds.) *Population reconstruction*. Berlin: *Springer*, 199 – 216.