

Viaplana, Joaquim; Maria-Rosa Lloret, Maria-Pilar Perea i Esteve Clua (2007):
COD. Corpus Oral Dialectal. Barcelona: Promociones y Publicaciones
Universitarias (PPU). [CD-ROM]
ISBN: 978-84-477-0990-8

Presentation_English

Aquest document conté el text Presentation_English, una introducció en anglès al CD del Corpus Oral Dialectal (COD). El COD és un component del Corpus de Català Contemporani de la Universitat de Barcelona (CCCUB), un arxiu de corpus de llengua catalana oral contemporània que ha estat confegit pel grup de recerca Grup d'Estudi de la Variació (GEV) amb la finalitat de contribuir a l'estudi de la variació dialectal, social i funcional en la llengua catalana.

Aquest i altres materials del CCCUB són accessibles directament al Dipòsit Digital de la UB (<http://diposit.ub.edu>) o a través del web del CCUB (<http://www.ub.edu/ccub>).

Aquesta obra està subjecta a la llicència de:
Reconeixement–NoComercial–SenseObraDerivada



<http://creativecommons.org/licenses/by-nc-nd/3.0/es/deed.ca>

Presentation

Since 1991 the Department of Catalan at the University of Barcelona has compiled and organized in databases a *Corpus Oral Dialectal* (COD) of contemporary Catalan. The corpus consists of phonetic and morphological materials obtained through a 600-item questionnaire and a set of spontaneous texts. The survey was conducted in the main county (*comarca*) capitals –or equivalent towns– of the Catalan linguistic domain with middle-age informants between 30-45 years old at the time of the study (between 1994 and 1996). These informants also had to fulfil the following requirements: they had not attended school in Catalan, belonged to the middle class, had no higher education, and had not spent lengthy periods away from their home. The criteria used in the selection of the locations and informants were chosen with the aim of compiling a dialectal corpus unlike the ones already existing for Catalan in traditional geolinguistics (e.g. *Atles del Domini Lingüístic del Català*, directed by Joan Veny and Lídia Pons), a corpus which should gather information on the current state of the language and its future perspectives.

This CD-ROM presents the results obtained from the questionnaire in eight databases (© Microsoft Access), each of which has a structure adapted to its contents: relevant phonetic aspects, regular verbal morphology, pronominal clitics, articles, possessives, strong personal pronouns, demonstratives and locatives. The CD-ROM can be used for teaching, and it can also be a useful tool in dialect research.

The project was possible over the years thanks to the research grants awarded by the Spanish Ministry of Education and Science (PB90-0505, PB94-0951, PB97-0889, BFF2001-3798 and HUM2004-01504, the last one with the additional support of the ERDF, *European Regional Development Fund*). The publishers of the present data have participated in all of them from the beginning.

The first phase of the project design took place within the framework of project PB90-0505 (“Variation in language. Current Catalan”), directed by Antoni M. Badia i Margarit. At this

stage the questionnaire was designed, the locations and the informant profile were selected, and several pilot tests of the survey were carried out to improve the questionnaire. Within the framework of project PB94-0951 (“Variation in language. Current Catalan”), directed by Joaquim Viaplana, the complete survey was carried out, part of the materials were transcribed, and the data from the questionnaire were entered into databases. Within the framework of project PB97-0889 (“Variation in language. Current Catalan”), also directed by Joaquim Viaplana i Lleonart, the transcription of the questionnaire data was completed, work proceeded on the creation of interrelated databases, and the transcription of some spontaneous texts was begun.¹ Within the framework of project BFF2001-3798 (“Analysis and interpretation of the dialectal linguistic variation of Catalan through the exploitation of a speech corpus”), directed by Joaquim Viaplana, databases were completed, and a first complete revision of the questionnaire materials was carried out for their future publication. Simultaneously, the transcription of the spontaneous texts continued, some of which were published for teaching purposes.² Finally, within the framework of the present project, where the questionnaire data are published, HUM2004-01504 (“Exploitation of a dialectal speech corpus: analysis of linguistic variation and development of computer applications for automatic transcription”), directed by Maria-Rosa Lloret, a second complete revision of the data has taken place, and the unification of the databases for the present publication has been finished. Work has continued with the transcription of the spontaneous texts.³

Since 1995, the above-mentioned projects have had the additional support of the *Generalitat de Catalunya* through the grants awarded such as the Established Research Group “Study group of variation, *Grup d’estudi de la variació (GEV)*” (1995SGR-00292,

¹ For a complete overview of the project at this stage, see the following two articles: Alturo, Núria, Emili Boix and Maria Pilar Perea (2002), “*Corpus de Català Contemporani de la Universitat de Barcelona. A General Presentation*”, in Pusch, C. D. and W. Raible (ed.), *Romanistische Korpuslingistik*, 155-170, Gunter Narr Verlag, Tübingen; Lloret, Maria-Rosa and M. Pilar Perea (2002), “A Report on the *Corpus Oral Dialectal del Català Actual (COD)*”, *Dialectologia et Geolinguística* 10, 59-76.

² Viaplana, Joaquim and Maria Pilar Perea (ed.) (2003), *Textos orals dialectals del català sincronitzats. Una selecció* [with CD-ROM], PPU, Barcelona.

³ For a complete overview of the project at this stage, see the following article: Clua, Esteve and Maria-Rosa Lloret (2006), “New tendencies in geographical dialectology. The Catalan *Corpus Oral Dialectal (COD)*”, in Montreuil, Jean-Pierre (ed.), *New Perspectives on Romance Linguistics. Volume II: Phonetics, Phonology and Dialectology*, 31-45, John Benjamins, Amsterdam/Philadelphia.

1996SGR-104, 1998SGR-0004, 12000SGR-00001, 2001SGR-00004 and 2005SGR-01046), directed by Joaquim Viaplana.⁴

Data transcription also had the participation of Clàudia Pons (1999-2002), as FPI scholarship holder within the framework of project PB97-0889; Pere Grimalt (2001-2004), Elisenda Campmany (2003-2006) and Laia Querol, as FPI scholarship holders within the framework of project BFF2001-3798; and Laia Rosàs (2004-2007), as FPI scholarship holder within the framework of project HUM2004-01504. There was also the occasional collaboration of David Esplugas, Marc Esplugas and Esteve Valls.

We hope this publication promotes research in the field of dialectology, phonetics, phonology and inflectional morphology in our country. We also hope it raises interest in our language among foreign scholars doing research in these areas.

Barcelona, October 2006

Joaquim Viaplana

Maria-Rosa Lloret

Maria Pilar Perea

University of Barcelona

Esteve Clua

Pompeu Fabra University

<<http://www.ub.edu/lincat>>

⁴ For more information on project HUM2004-01504, the following website can be consulted: <<http://www.ub.edu/lincat>>, and on the Established Research Group <<http://www.ub.edu/filcat>> (Recerca > Lingüística tabs).