

RESEÑA
MATERIALS FOR THE SOCIOLINGUISTIC DESCRIPTION AND CORPUS-
BASED STUDY OF SPANISH IN BARCELONA
(2009, R. E. VANN)

IZASKUN ELORZA
UNIVERSIDAD DE SALAMANCA

En la última década hemos podido observar un creciente interés por los estudios de corpus del español hablado a través de la cada vez más numerosa publicación de los resultados de proyectos centrados en la investigación sociolingüística. Se trata por lo general de trabajos que constituyen descripciones de algunos rasgos distintivos de una u otra variedad geográfica y social. Sirva de botón de muestra el trabajo de análisis del español hablado en Valencia (Gómez Molina 2002), desarrollado actualmente como parte del Proyecto para el Estudio Sociolingüístico del Español de España y América (PRESEEA), que agrupa más de 30 proyectos bajo la dirección de Francisco Moreno Fernández desde la Universidad de Alcalá (cf. Moreno Fernández 2006). Otros proyectos de enfoque sociolingüístico que se han desarrollado también en esta última década son los que han permitido compilar otros corpus, como el ALCORE y el COVJA (Azorín Fernández y Jiménez Ruiz 1999; Azorín Fernández 2002; 2005), que en este caso recogen muestras de la variedad hablada en Alicante. Mucho más recientemente, se suma a este fructífero panorama descriptivo el trabajo de Robert E. Vann publicado en 2009 por la editorial *Edwin Mellen Press*.

Apoyándose en la concepción de Bourdieu (2001) de que todas las prácticas lingüísticas de una comunidad son valoradas tomando como baremo las prácticas dominantes y legitimadas por la propia comunidad, Vann desarrolla su estudio a partir de una concepción de la realidad lingüística española en la que el uso de la lengua por parte de los hablantes es aceptado o rechazado de acuerdo con su mayor o menor proximidad o coincidencia con las características del castellano, es decir, con la variedad dominante del español peninsular, definida como “*the dialect of Spanish native to Castile*” (pág. 13).

El estudio de Vann se centra en la variedad de contacto del español hablado en Cataluña. Esta variedad del español peninsular no dispone aún de una descripción exhaustiva, por lo que el trabajo de Vann es de gran interés para los dialectólogos del español ya desde su propio planteamiento inicial. Para describir la variedad elegida, Vann no sólo ha estudiado los escasos trabajos previos disponibles, sino que además ha llevado a cabo un trabajo de campo recopilando datos del español hablado en Barcelona, mediante la grabación de conversaciones coloquiales entre sujetos pertenecientes a dos redes sociales de diferentes zonas de la ciudad (el distrito de Les Corts y el de Sant Andreu respectivamente), parte de cuyas transcripciones se presentan en este libro.

La obra de Vann está dividida en dos partes. La primera parte recoge una introducción crítica al estudio del español de Cataluña, cuyo objetivo es contextualizar de forma apropiada el corpus compilado por Vann, así como los datos presentados en la segunda parte del libro. Esta introducción se extiende a lo largo de seis capítulos. Los dos primeros capítulos abordan la descripción del español de Cataluña desde una perspectiva sociohistórica. El Capítulo 1 (“*Brief history of the political economy of language in Catalonia*”) introduce, desde la aproximación de la Economía Política del

Revista Electrónica de Lingüística Aplicada (ISSN 1885-9089)

2009, Número 8, páginas 276-280

Recibido: 19/12/2009

Aceptación comunicada: 01/02/2010

Lenguaje desarrollada por Bourdieu (1991) y otros autores, el marco teórico a partir del cual Vann analiza la situación lingüística de Cataluña. Para ello utiliza el concepto de *habitus*, que se refiere en este contexto a las sensaciones y reacciones viscerales aprendidas mediante la experiencia vital que orientan al individuo en sus acciones. El *habitus* se considera homogéneo entre individuos con similares conocimientos y experiencia vital, puesto que es el resultado de las experiencias vividas y la historia compartida por una generación tras otra. Lo que Vann pretende mediante esta aproximación es revelar el conjunto de valores, creencias y actitudes relativos a la realidad sociolingüística del contexto estudiado, tanto en relación con las prácticas lingüísticas (*linguistic habitus*) de los sujetos que participan en las conversaciones, como con las ideologías y discursos políticos (*political habitus*) que se revelan en las interacciones. Vann logra este objetivo mediante un cuidadoso diseño en las características de los dos grupos elegidos para la recopilación de datos, lo cuales difieren socialmente en cuanto a la lengua que utilizan habitualmente para la interacción social (una red utiliza habitualmente el español, mientras que la otra suele utilizar el catalán). En el Capítulo 2 (“*Language ideologies in academic treatment of Catalan Spanish*”) se presenta un interesante y detallado análisis del tratamiento que ha recibido el español de Cataluña en la investigación lingüística del siglo XX, poniendo de relieve que –al igual que ha sucedido con las demás variedades de contacto del español peninsular– el español de Cataluña ha sido tradicionalmente ignorado por la dialectología española desde un punto de vista descriptivo. Las razones que Vann señala para esta situación atienden tanto a cuestiones de tipo ideológico e histórico, específicamente al afán de la política de Franco de presentar la realidad del país como un todo uniforme que incluiría también la lengua, así como a un desarrollo de la dialectología del español peninsular desde una perspectiva esencialmente diacrónica y prescriptivista. Los argumentos presentados explican por qué el español de Cataluña ha sido reconocido por los lingüistas como una variedad del español peninsular sólo a partir de los años 80 y, en consecuencia, la escasez de estudios de tipo descriptivo disponibles hasta el momento.

Los dos capítulos siguientes están dedicados a la descripción de la investigación llevada a cabo hasta la fecha sobre el español hablado en Cataluña. El Capítulo 3 (“*Spoken language corpora of Spanish in Catalonia*”) introduce el concepto de corpus utilizado en el estudio, así como una descripción de la metodología aplicada, contrastando ambos con los estudios previos llevados a cabo sobre corpus orales del español de Cataluña, y muy especialmente con los trabajos de Vila Pujol (2001) y de Sinner (2001; 2002; 2004), de los cuales se diferencia en el tipo de registro que recoge (lengua coloquial de conversaciones espontáneas en grupo). Precisamente uno de los puntos fuertes del trabajo publicado por Vann consiste en la metodología de compilación de corpus que utiliza, basada en las conversaciones desarrolladas en grupos sociales en los que él se había integrado previamente como un participante más, lo que le permitió acceder a un registro coloquial de la lengua oral que no estaba documentado hasta el momento. Por este motivo, puesto que el corpus recopilado es único, el libro de Vann constituye una obligada fuente de referencia para los estudiosos del español hablado en Cataluña. El Capítulo 4 (“*Corpus-based linguistic descriptions of Spanish in Catalonia*”) presenta una descripción de las características documentadas hasta el momento mediante la compilación de datos reales del dialecto estudiado. Se recogen aquí datos relativos a la pronunciación, al vocabulario y a la gramática con ejemplos provenientes de diferentes corpus, incluyendo el del propio autor.

Revista Electrónica de Lingüística Aplicada (ISSN 1885-9089)

2009, Número 8, páginas 276-280

Recibido: 19/12/2009

Aceptación comunicada: 01/02/2010

Los últimos dos capítulos de la primera parte se dedican a describir el trabajo de investigación llevado a cabo por Vann. El Capítulo 5 (“*Methods of data Collection, digitization and transcription*”) presenta una relación pormenorizada del diseño de la investigación llevada a cabo y la metodología utilizada para la recopilación de datos y el tratamiento de los mismos para su transcripción. Los investigadores que trabajan en este campo encontrarán un excelente nivel de detalle en la descripción presentada por Vann, por ejemplo en cuanto al proceso que siguió para digitalizar las 21.5 horas de grabaciones de conversaciones en grupo que había recopilado en 1995, parte de las cuales se presentan en el libro, por lo que se trata de información muy útil para la replicación del diseño y la metodología en investigaciones posteriores de la misma naturaleza. El Capítulo 6 (“*Network profiles and selected speaker data for the Vann corpus*”) recoge de forma resumida pero suficiente los perfiles de los sujetos que participan en las conversaciones transcritas en la segunda parte del libro. Estos perfiles provienen de dos fuentes: la observación directa cualitativa llevada a cabo por el investigador durante un periodo de observación previo a la recogida de datos y los cuestionarios que completaron los sujetos que participaron en las conversaciones reproducidas.

Descripción y niveles de representación del corpus Vann

Además del indudable interés para los dialectólogos del español peninsular, la obra de Vann está fundamentalmente dirigida a un público más amplio. Por este motivo, la descripción es amena y carece de la densidad metalingüística y estadística que se puede esperar en trabajos de tipo académico orientados a un lector experto en la materia. Así, además de su indudable valor para los dialectólogos, el libro es igualmente apto e interesante para estudiantes de nivel universitario y para cualquier persona del ámbito de la comunicación intercultural, además de los propios hablantes de la variedad estudiada.

La orientación del libro hacia un público amplio es asimismo utilizada por Vann para justificar el nivel de representación de los datos transcritos en la segunda parte del libro. Como señala Llisterri (1999: 54), “un corpus puede concebirse como un conjunto estructurado de materiales lingüísticos en el que se distinguen diversos niveles de representación correspondientes a diferentes grados de elaboración de los datos que los constituyen”. En cuanto a la elaboración de los datos registrados, el corpus Vann presenta un único nivel de representación, que corresponde con el nivel de transcripción ortográfica. Tratándose de un corpus oral, la utilización de un sistema de transcripción poco detallado supone en la práctica ciertas limitaciones en cuanto a las posibilidades de explotación y análisis de los datos recogidos en él. Así pues, aunque Vann declara que su interés en observar la máxima fidelidad, usabilidad y potencial accesibilidad a los datos en la transcripción se corresponde con el objetivo de crear un recurso lingüístico disponible para todos, incluyendo también a aquellas personas que carecen de conocimientos lingüísticos (pág. 97), se echan en falta otros niveles de representación de elementos segmentales y suprasegmentales que permitirían a los investigadores realizar estudios más completos de la variante de contacto representada, por ejemplo de la pronunciación y la entonación, puesto que se trata de rasgos pertinentes y característicos de dicho dialecto. Asimismo, aunque es cierto, como Vann propone, que

las posibles anotaciones del corpus pueden ser realizadas posteriormente por los investigadores que lo utilicen en función de los objetivos concretos de cada investigación (pág. 97), habría sido de gran utilidad que la edición del trabajo incluyera el acceso o una copia en formato digital, además de la transcripción en papel, para facilitar la introducción de metadatos en el corpus.

Por otra parte, es precisamente la transcripción ortográfica del corpus lo que presenta un punto débil en el trabajo de Vann que, aunque no hace que el trabajo pierda valor, sí plantea algunos problemas a la hora de trabajar con los datos aportados. Tal como el propio autor describe, la transcripción de corpus orales presenta una variada problemática, que incluye en este caso particular la identificación de más de una lengua en las conversaciones (en ocasiones se utiliza el catalán en mitad de la conversación sin transiciones marcadas) o la identificación de qué participante está interviniendo en cada ocasión, además de las complejidades de tipo técnico o la elección de qué convenciones se van a seguir en la transcripción. Parte de las transcripciones del corpus Vann fueron hechas a partir de la grabación analógica y corregidas cuando se llevó a cabo el proceso de digitalización. No obstante, y a pesar de que no es posible cotejar la transcripción presentada en la segunda parte del libro, una atenta lectura revela algunos casos en los que parece que la transcripción es inconsistente o incoherente. Algunos ejemplos son la transcripción “Ya, me echo yo aquí, en la cultura aquí y yo” (pág. 168), en lugar de “[...] me he hecho [...]”, la transcripción “mal educada” en “Eres una mal educada” (pág. 220), o también la transcripción “Cruiz” refiriéndose (presumiblemente, por el contexto) a Cruyff (pág. 238).

En resumen, el trabajo de Vann constituye una obra de referencia imprescindible para los dialectólogos del español peninsular, tanto por su singularidad en la metodología utilizada como por la transparencia de la propia investigación y de los datos recogidos, lo que permitirá a los investigadores una replicación de los mismos. No obstante, sería deseable que en sucesivas ediciones la transcripción del corpus presentado se revisara para pulir las pequeñas inconsistencias o incoherencias que presenta, con el fin de no dificultar la posibilidad de cotejar los datos con los de otros corpus similares. También sería deseable que la transcripción recogida en la segunda parte del libro fuera accesible en un formato digital que permitiera incluir metadatos para añadir otros niveles de representación y análisis. Por último, hay que reseñar asimismo que lo interesante del tema tratado, así como el ameno estilo narrativo hacen que la lectura de este libro sea también de gran interés para cualquier persona que desee aumentar su conocimiento sobre los aspectos históricos y socioculturales del español actual de Cataluña y ver ejemplos reales de dicho dialecto.

Referencias bibliográficas

- Azorín Fernández, D., ed. 2002: *El proyecto @LCORE: Alicante-corpus oral del español*. Alicante, Servicio de Publicaciones de la Universidad de Alicante. (CD-Rom)
- Azorín Fernández, D. 2005. Corpus oral para el estudio del lenguaje juvenil y del español hablado en Alicante: el corpus ALCORE y COVJA. *Oralia: Análisis del discurso oral* 8, págs. 265-288. ISSN 1575-1430

- Azorín Fernández, D. y J. L. Jiménez Ruiz. 1999. *Corpus oral de la variedad juvenil universitaria del español hablado en Alicante*. Alicante, Instituto Alicantino de Cultura Juan Gil-Albert.
- Bourdieu, P. 2001 (1985). *¿Qué significa hablar?*. Madrid: Akal.
- Gómez Molina, J. R. 2002. El “corpus” del español hablado de Valencia: Proyecto para el Estudio Sociolingüístico del Español de España y América (PRESEEA). En J. L. Blas, M. Porcar Miralles, S. Fortuño Llorens y M. Casanova (coords.). *Estudios sobre lengua y sociedad*. 217-235. Castellón de la Plana: Publicacions de la Universitat Jaume I.
- Llisterri, J. 1999. Transcripción, etiquetado y codificación de corpus orales. *Revista Española de Lingüística Aplicada*. Volumen Monográfico “Panorama de la Investigación en Lingüística Informática”, 53-82.
- Moreno Fernández, F. 2006. El español en su variación geográfica y social: informe sobre el corpus “PRESEEA”. En A. M. Cestero Mancera, I. Molina Martos y F. Paredes García (coords.). *Estudios sociolingüísticos del español de España y América*. 15-34. Madrid: Arco Libros.
- Sinner, C. 2001. Corpus oral de profesionales de la lengua castellana en Barcelona. [Disponible en <http://www.carstensinner.de/castellano/corpusorales/index.html>]
- Sinner, C. 2002. Análisis contrastivo de un corpus oral de diferentes variedades del castellano: Aspectos cuantitativos y cualitativos. En C. D. Pusch y W. Raible (eds.). *Romanistische Korpuslinguistik – Korpora und gesprochene Sprache / Romance corpus linguistics – corpora and spoken language*. 279-292. Tübinga: Narr.
- Sinner, C. 2004. *El castellano de Cataluña*. Tübinga: Niemeyer.
- Vila Pujol, M. R. 2001. *Corpus del español conversacional de Barcelona y su área metropolitana*. Barcelona: Edicions Universitat de Barcelona.