

Oralia

Análisis del discurso oral

8

2005

Handwritten text in a cursive script, likely representing a spoken discourse analysis. The text is arranged in approximately 15 lines, with some lines starting with a capital letter. The script is fluid and connected, typical of cursive handwriting.

ilse

Grupo de investigación
Universidad de Almería



ARCO/LIBROS, S.L.

ilse

Grupo de investigación
Universidad de Almería

CORPUS PARA EL ESTUDIO DE LAS HABLAS ANDALUZAS I.
EL CORPUS DEL HABLA DE SEVILLA Y EL CORPUS DEL
HABLA DE ALMERÍA

PEDRO CARBONERO CANO
(Universidad de Sevilla)

LUIS CORTÉS RODRÍGUEZ
ANTONIO M. BAÑÓN HERNÁNDEZ
(Universidad de Almería)

1. INTRODUCCIÓN

Creemos que fue el trabajo de W. Labov (1966) sobre el inglés de la ciudad de Nueva York uno de los motivos que ayudó a desplazar el centro de estudio de las hablas rurales al de las variedades urbanas, en las que se suponía que la norma lingüística habría de presentar un carácter más heterogéneo. Dicha derivación hacia el estudio de este tipo de hablas se vio reforzada en el mundo hispánico por el famoso *Proyecto de estudio coordinado de la norma lingüística culta de las principales ciudades de Iberoamérica*, ampliado más tarde a la Península Ibérica. Precisamente, la metodología que se habría de aplicar en este Proyecto exigía la elaboración de grandes corpus orales grabados. En la incorporación, aunque algo tardía, de la ciudad de Sevilla a dicho Proyecto, tuvo su origen el primero de los corpus aquí presentados. La tardanza aludida, sin embargo, iba a hacer factible el enriquecimiento del corpus sevillano con respecto a los ya iniciados en las grandes ciudades, pues en tanto que estos sólo se ocuparon de informantes de un nivel sociocultural alto, el corpus sevillano incorporó también en su muestra los niveles medio y bajo. El corpus del habla de Sevilla, junto al de Madrid, tuvo el privilegio de ser el primero de los llevados a cabo en nuestro país.

Los trabajos que van surgiendo a partir de estos corpus pioneros confirman cómo la ciudad realmente sustenta los diversos intentos de diversificación lingüística en función de los diferentes estratos (socioculturales, de edad, de sexo, etc.) y de las situaciones contextuales; por otro lado, también se descubre que dicha ciudad obliga a una nivelación para facilitar la comprensión dentro de la dispersión impuesta por variados tipos de vida, de intereses o de razas, de donde se deduce que el modo más directo de investigar la variedad urbana es a partir de un gran número de informantes, jerárquicamente divididos, pero también de un número variado de situa-

ciones discursivas (entrevista, conversaciones ocultas, telefónicas, programas radiofónicos, televisivos, debates, etc.), lo que nos va a permitir acercarnos con mayores posibilidades al estudio de las más diferentes cuestiones discursivas: de los marcadores de conexión e interacción a ciertas actitudes de los hablantes ante su propia lengua; de las unidades de segmentación a cuestiones propias del análisis crítico del discurso. Y con esta intención se va a llevar a cabo el segundo corpus que presentamos en este artículo, el *Corpus del habla de Almería*.

2. EL CORPUS LINGÜÍSTICO DEL HABLA DE SEVILLA

El corpus para el estudio del habla urbana de Sevilla se elaboró inicialmente con el propósito de participar en el Proyecto de estudio coordinado de la norma lingüística culta de las principales ciudades del mundo hispánico (hoy conocido como *Proyecto de la norma culta hispánica Juan M. Lope Blanch*)¹. Para ello se realizaron entrevistas grabadas a hablantes sevillanos cultos, con una duración aproximada de treinta minutos cada una. Posteriormente, con el objetivo de proceder a los oportunos contrastes sociolingüísticos, se realizaron encuestas del mismo tipo a hablantes de otros dos niveles socioculturales: medio y popular.

En consecuencia, el corpus actualmente disponible consta de un total de 72 entrevistas, realizadas a otros tantos informantes, distribuidas en tres grupos de 24 grabaciones cada uno, que representan, respectivamente, a tres niveles: *culto*, *medio* y *popular*.

Las 24 entrevistas de cada nivel están distribuidas por igual atendiendo al sexo (hombres y mujeres) y a la edad de los informantes (1ª, 2ª y 3ª generación). Por tanto, hay cuatro entrevistas para cada grupo social, según la siguiente distribución:

Nº de encuestas realizadas en cada nivel sociocultural:

	1ª gen.	2ª gen.	3ª gen.
Hombres	4	4	4
Mujeres	4	4	4
Total:	24		

¹ Las ciudades vinculadas con este Proyecto son: México, Santiago de Chile, Bogotá, Caracas, San José de Costa Rica, La Habana, La Paz, Lima, San Juan de Puerto Rico, Buenos Aires, Córdoba (Argentina), Las Palmas de Gran Canaria, Madrid, Granada y Sevilla. Para más información, véase en este mismo volumen el trabajo de J. A. Samper.

2.1. Objetivos

Con el corpus del nivel culto se ha pretendido colaborar en el amplio proyecto del mundo hispánico, para conocer la norma culta oral del español actual, contrastando diferencias y percibiendo semejanzas, atendiendo a las diferentes parcelas de la lengua: la fonética, el léxico y la morfosintaxis. Asimismo, con el análisis de estas muestras de habla sevillana se busca una aproximación al conocimiento de la norma culta andaluza.

Por otra parte, el corpus completo constituye un conjunto de materiales que permite el estudio en profundidad del habla de Sevilla, en sus dimensiones sociolingüísticas y en el contexto de las hablas andaluzas.

Los análisis sobre el discurso oral también pueden adquirir dimensiones aplicadas a la didáctica de la lengua y al uso de la misma en los medios de comunicación social.

2.2. Variables sociológicas y selección de informantes

Atendiendo a los requerimientos del proyecto, todos los informantes que fueron seleccionados eran sevillanos: personas que han nacido en Sevilla y que han vivido la mayor parte de su vida en esta ciudad. Se cuidó que los encuestados pertenecieran a los distintos barrios y distritos. Para el muestreo se tuvieron en cuenta las variables sociológicas de nivel sociocultural, edad y sexo.

Se establecieron tres niveles socioculturales: culto, medio y popular. Los informantes seleccionados para el nivel *culto* son personas con estudios universitarios terminados, que declaran ser lectores habituales, han viajado al extranjero y poseen conocimientos de una segunda lengua. Pertenecen a profesiones como médicos, abogados, licenciados, profesores, jueces, ingenieros, religiosos y escritores. En el nivel *popular* la mayor parte de los encuestados se inscriben entre los sevillanos que han realizado los estudios primarios y algunos se sitúan en zonas culturales colindantes, como los que tienen estudios primarios incompletos y otros con estudios de bachillerato o formación profesional. Están representados diversos empleos o actividades laborales: camarero, agente de ventas, mecánico, contable, pastelero, dependiente, carnicero, albañil, vendedor de prensa, secretaria, peluquera, ama de casa y desempleado. Para representar al nivel *medio* se seleccionaron hablantes cuya formación académica fuese, al menos, de bachillerato o tuviesen una titulación profesional de nivel equivalente. Los encuestados ejercen alguna de las siguientes profesiones: administrativo, comerciante, ajustador, estudiante, ama de casa, aparejador, visitador médico, ayudante técnico sanitario, cuidador-educador, secretaria, funcionario, maestra o auxiliar de consulta médica.

En lo que respecta a la variable edad, se establecieron tres generaciones de hablantes delimitadas así: menores de 30 años (*primera generación*), entre 30 y 45 años (*segunda generación*) y mayores de 45 años (*tercera generación*). Por otra parte, se utilizó la variable sexo, que sirve para agrupar a los informantes en *hombres* y *mujeres*.

2.3. Diseño y realización de las encuestas

Las encuestas se diseñaron para propiciar una conversación semidirigida que fuera lo más libre y espontánea posible. El encuestador llevaba un guión previo pero abierto, con temas propuestos que giraban en torno a la vida del informante, su entorno familiar y profesional, Sevilla y su fisonomía, sus costumbres, sus manifestaciones culturales, los medios de comunicación, etc. Pero se trataba de hacer olvidar al informante, en la medida de lo posible, que la conversación se estaba grabando. Se procuró siempre no forzar al hablante en una dirección temática no deseada por él. De esta manera, el Grupo de Investigación "Sociolingüística Andaluza" de la Universidad de Sevilla dispone de un archivo sonoro, en el que las 72 encuestas realizadas se encuentran almacenadas en cintas magnetofónicas y están siendo copiadas para su conservación en discos compactos.

2.4. Transcripción

Tras una cuidadosa y repetida audición de las cintas, se procedió a la transcripción de las conversaciones obtenidas con las encuestas. Se trata de una transcripción ortográfica: las palabras están escritas según la ortografía académica y los signos de puntuación atienden tanto a las pausas naturales como a los criterios razonables de la configuración sintáctica de los textos. Ello facilita los posteriores análisis léxicos, morfosintácticos y discursivos, aunque no refleja la peculiar fonética de los informantes sevillanos. Lógicamente para los análisis fonéticos se hace necesaria la audición directa de las grabaciones. Sólo excepcionalmente se optó por transcribir con una ortografía peculiar ciertos términos que claramente exigen un tratamiento gráfico diferente para no perder su valor expresivo (p. ej.: *jartá* 'hartada', *mijita* 'migajita', *cantaor* 'cantador'...).

2.5. Formato y soporte

El material grabado y transcrito constituye un corpus de aproximadamente 350.000 palabras, correspondientes a unas 35 horas de grabación. Para

facilitar su consulta y ponerlo a disposición de los investigadores interesados, las transcripciones fueron publicadas, en primera instancia, en formato de libros. Se encuentran en tres volúmenes, cada uno de los cuales agrupa un nivel sociocultural:

- Nivel culto: en *Sociolingüística Andaluza*, 2 (Lamíquiz y Pineda, eds., 1983).
- Nivel popular: en *Sociolingüística Andaluza*, 4 (Lamíquiz y Roperio, eds., 1987).
- Nivel medio: en *Sociolingüística Andaluza*, 6 (Ollero y Pineda, eds., 1992).

Cada uno de los mencionados volúmenes, además de la transcripción de las encuestas, contiene en su parte introductoria los criterios utilizados para su realización.

Una selección de las encuestas del nivel culto forma parte del Macrocorpus de la norma lingüística culta de las principales ciudades del mundo hispánico (Samper, Hernández Cabrera y Troya, eds., 1998).

Las concordancias e índices del corpus de los tres niveles fueron publicados en forma de microfichas en el volumen *Sociolingüística Andaluza*, 7 (Pineda y Ollero, 1993).

Además de ello, el grupo de investigación responsable del proyecto del habla de Sevilla tiene informatizado el texto de todas las encuestas en formato Word y puede facilitar copias a cualquier investigador que las solicite.

Aparte de este corpus, actualmente disponible, hay trabajos en vías de desarrollo para la recopilación y puesta a punto de otros materiales. Entre ellos está la realización de nuevas encuestas grabadas, que van a complementar y actualizar las que ahora existen, atendiendo a diferentes situaciones discursivas (diálogo abierto entre varios informantes, entrevistas, discursos individuales en situaciones formales). También se está elaborando el léxico del habla culta de Sevilla, siguiendo el cuestionario y las pautas metodológicas marcadas en el proyecto internacional de estudio coordinado de la norma culta hispánica.

2.6. Identificación

Para la identificación de cada encuesta se estableció una etiqueta de cuatro dígitos:

- El primero se refiere al nivel sociocultural: C (culto), M (medio), P (popular).

- El segundo alude a la edad del informante: 1 (primera generación), 2 (segunda generación), 3 (tercera generación).
- El tercero concierne al sexo: V (varón), H (hembra).
- El último es un número personal de identificación de cada informante.

Atendiendo a estos criterios, si una encuesta, por ejemplo, está identificada como C2V3, sabremos que estamos ante un hablante del nivel culto, de la segunda generación, varón, y es el hablante nº 3 de este grupo sociolingüístico. Otro ejemplo: la encuesta P3H1 es la que corresponde a un hablante del nivel popular, que pertenece a la tercera generación, que es mujer y que es el informante nº 1 de su grupo.

Aparte de eso, en cada uno de los volúmenes publicados, antes de la transcripción de cada encuesta aparecen los datos específicos del informante: edad, estudios, profesión, etc.

2.7. Bibliografía producida

Los materiales para el estudio del habla urbana de Sevilla han sido y siguen siendo objeto de diversos estudios y análisis que abordan múltiples aspectos de los usos lingüísticos: fonéticos, léxicos, morfosintácticos y discursivos. Muchos de estos estudios se han publicado en los volúmenes de la serie titulada *Sociolingüística Andaluza* (Publicaciones de la Universidad de Sevilla), de la que hasta el momento presente han aparecido trece números y que es el vehículo fundamental para la difusión de los resultados de las investigaciones del Grupo. Pero hay trabajos que han sido publicados en otros medios científicos (revistas, actas de congresos, etc.), tanto por los propios miembros del Grupo como por otros investigadores que se han interesado por estos materiales. Además, nuestro corpus ha sido una fuente documental para investigaciones académicas realizadas en forma de Tesis de Licenciatura y de Doctorado.

La bibliografía científica producida como resultado de todas estas investigaciones sobre el corpus del habla de Sevilla es la que aparece referida a continuación:

- AGUDO, J. A. (1985): "Uso discursivo de *ir* en el habla urbana culta de Sevilla", en V. Lamíquiz y F. Rodríguez-Izquierdo (eds.), *Sociolingüística Andaluza*, 3: *El discurso sociolingüístico*, Sevilla, Publ. Universidad, págs. 121-135.
- ALCAIDE, E. (1990): "La oración de predicado no verbal como medio de cohesión textual. Estudio contrastivo en las encuestas del habla urbana de Sevilla (niveles popular y culto)", en P. Carbonero y M^a T. Palet (eds.), *Sociolingüística Andaluza*, 5: *Habla de Sevilla y hablas americanas*, Sevilla, Publ. Universidad, págs. 125-135.

- ALCAIDE, E. (1993): "Anotaciones sobre algunos usos de la interjección en el habla urbana de Sevilla (nivel popular)", en P. Carbonero y C. Fuentes (eds.), *Sociolingüística Andaluza*, 8: *Estudios sobre el enunciado oral*, Sevilla, Publ. Universidad, págs. 215-235.
- ANDÚJAR, A. (1993): "Comportamiento discursivo de *todavía* en el habla urbana culta de Sevilla", en P. Carbonero y C. Fuentes (eds.), *Sociolingüística Andaluza*, 8: *Estudios sobre el enunciado oral*, Sevilla, Publ. Universidad, págs. 161-182.
- ASENCIO, A.M^a (1996): *Comportamiento discursivo de los verbos de voluntad en el habla urbana culta de Sevilla y el habla urbana popular de Sevilla*, Tesis Doctoral, Universidad de Sevilla.
- CARAVEDO, R. (1999): "Estudio de los nexos temporales. Síntesis de los trabajos presentados", Memoria de la VI Reunión de Trabajo de la Comisión Ejecutiva para el Estudio coordinado de la Norma Lingüística Culta Hispánica, en J. Matluck y C. Solé (eds.), *Actas del Simposio Internacional de la Lengua Española: Pasado, Presente y Futuro*, The University of Texas at Austin, págs. 45-55.
- CARBONERO, P. (1985a): "Norma estándar y actitud sociolingüística. Sobre la aceptación y uso de algunos rasgos lingüísticos en hablantes sevillanos", en V. Lamíquiz y P. Carbonero (eds.), *Sociolingüística Andaluza*, 1: *Metodología y estudios*, Sevilla, Publ. Universidad, págs. 141-150.
- CARBONERO, P. (1985b): "Sobre ciertas construcciones de relativo en el habla urbana de Sevilla", en V. Lamíquiz y F. Rodríguez-Izquierdo (eds.), *Sociolingüística Andaluza*, 3: *El discurso sociolingüístico*, Sevilla, Publ. Universidad, págs. 65-85.
- CARBONERO, P. (1990): "Usos de las formas verbales en *-ra* y en *-se* en el habla de Sevilla (nivel popular)", en P. Carbonero y M^a T. Palet (eds.), *Sociolingüística Andaluza*, 5: *Habla de Sevilla y hablas americanas*, Sevilla, Publ. Universidad, págs. 45-58.
- CARBONERO, P. (1992): "Queísmo y dequeísmo en el habla culta de Sevilla: Análisis contrastado con otras hablas peninsulares y americanas", en E. Luna Traill (coord.), *Scripta Philologica in honorem Juan M. Lope Blanch*, México, UNAM, II, págs. 43-63.
- CARBONERO, P. (1993): "La expresión de la relación concesiva en el habla urbana culta de Sevilla", en P. Carbonero y C. Fuentes (eds.), *Sociolingüística Andaluza*, 8: *Estudios sobre el enunciado oral*, Sevilla, Publ. Universidad, págs. 35-60.
- CARBONERO, P. (2000a): "Construcciones partitivas en los usos orales (un estudio sobre el habla de Sevilla)", en P. Carbonero, M. Casado y P. Gómez Manzano (eds.), *Lengua y discurso. Estudios dedicados al Prof. Vidal Lamíquiz*, Madrid, Arco/Libros, págs. 173-182.
- CARBONERO, P. (2000b): "Norma culta y actitudes lingüísticas de los andaluces", en A. Valencia y L. Piña (eds.), *Actas del Congreso Internacional: El Español Culto en el Mundo Hispánico*, Santiago de Chile, Universidad Bolivariana, págs. 25-32.
- CARBONERO, P. (2001a): "Las construcciones con *lo que* y su uso en el habla

- de Sevilla", en P. Carbonero y R. Guillén (eds.), *Sociolingüística Andaluza*, 12: *Identidad lingüística y comportamientos discursivos*, Sevilla, Publ. Universidad, págs. 63-98.
- CARBONERO, P. (2001b): "Cuestiones de sintaxis oral: causales, explicativas y justificativas", en E. Méndez, J. Mendoza y Y. Congosto (eds.), *Indagaciones sobre la Lengua. Estudios de Filología y Lingüística Españolas en Memoria de Emilio Alarcos*, Sevilla, Publ. Universidad, págs. 313-322.
- FUENTES, C. (1985): "Sobre las oraciones consecutivas en el habla urbana de Sevilla", en V. Lamíquiz y F. Rodríguez-Izquierdo (eds.), *Sociolingüística Andaluza*, 3: *El discurso sociolingüístico*, Sevilla, Publ. Universidad, págs. 87-103.
- FUENTES, C. (1990a): "Procedimientos intradiscursivos: *decir* y los explicativos", en P. Carbonero y M^a T. Palet (eds.), *Sociolingüística Andaluza*, 5: *Habla de Sevilla y hablas americanas*, Sevilla, Publ. Universidad, págs. 103-123.
- FUENTES, C. (1990b): "Algunos operadores de función fática", en P. Carbonero y M^a T. Palet (eds.), *Sociolingüística Andaluza*, 5: *Habla de Sevilla y hablas americanas*, Sevilla, Publ. Universidad, págs. 137-170.
- FUENTES, C. (1990c): "Apéndices con valor apelativo", en P. Carbonero y M^a T. Palet (eds.), *Sociolingüística Andaluza*, 5: *Habla de Sevilla y hablas americanas*, Sevilla, Publ. Universidad, 171-196.
- FUENTES, C. (1993a): "Claro: modalización y conexión", en P. Carbonero y C. Fuentes (eds.), *Sociolingüística Andaluza*, 8: *Estudios sobre el enunciado oral*, Sevilla, Publ. Universidad, págs. 99-126.
- FUENTES, C. (1993b): "Desde luego, por supuesto, naturalmente", en P. Carbonero y C. Fuentes (eds.), *Sociolingüística Andaluza*, 8: *Estudios sobre el enunciado oral*, Sevilla, Publ. Universidad, págs. 127-159.
- FUENTES, C. (1997): "Sintaxis coloquial andaluza: la cortesía en el habla urbana de Sevilla", en A. Narbona y M. Roperó (eds.), *El Habla Andaluza* (Actas del Congreso del Habla Andaluza), Sevilla, págs. 461-471.
- FUENTES, C. y ALCAIDE, E. (1996): *La expresión de la modalidad en el habla de Sevilla*, Sevilla, Publ. Ayuntamiento.
- GUILLÉN, R. (1987): *El habla culta de Sevilla: estudio léxico*, Sevilla, Alfar.
- GUILLÉN, R. (1988): *La expresión de la opinión subjetiva en el comportamiento discursivo del habla urbana de Sevilla*, Tesis Doctoral, Universidad de Sevilla.
- GUILLÉN, R. (1990): "Uso discursivo de *creer* y *pensar* en el habla urbana de Sevilla (nivel popular)", en P. Carbonero y M^a T. Palet (eds.), *Sociolingüística Andaluza*, 5: *Habla de Sevilla y hablas americanas*, Sevilla, Publ. Universidad, págs. 85-102.
- GUILLÉN, R. (1993): "La elipsis y sus repercusiones sintácticas en el discurso", en P. Carbonero y C. Fuentes (eds.), *Sociolingüística Andaluza*, 8: *Estudios sobre el enunciado oral*, Sevilla, Publ. Universidad, págs. 87-97.
- GUILLÉN, R. (1997): "Estudio de las expresiones lexicalizadas en el habla urbana de Sevilla", en A. Narbona y M. Roperó (eds.), *El Habla Andaluza* (Actas del Congreso del Habla Andaluza), Sevilla, págs. 553-562.
- GUILLÉN, R. (2000): "Uso discursivo de determinadas formas verbales en el habla urbana de Sevilla", en P. Carbonero, M. Casado y P. Gómez Man-

- zano (eds.): *Lengua y discurso. Estudios dedicados al Prof. Vidal Lamíquiz*, Madrid, Arco/Libros, págs. 445-454.
- GUILLÉN, R. (2001a): "Variación funcional de los llamados verbos de habla", en P. Carbonero y R. Guillén (eds.), *Sociolingüística Andaluza*, 12: *Identidad lingüística y comportamientos discursivos*, Sevilla, Publ. Universidad, págs. 99-118.
- GUILLÉN, R. (2001b): "Las formas verbales en español: norma académica y uso discursivo", en E. Méndez, J. Mendoza y Y. Congosto (eds.), *Indagaciones sobre la Lengua. Estudios de Filología y Lingüística Españolas en Memoria de Emilio Alarcos*, Sevilla, Publ. Universidad, págs. 379-392.
- LAGUILLO, J. (1990): "El español de Andalucía y América. Una aproximación léxica", en P. Carbonero y M^a T. Palet (eds.), *Sociolingüística Andaluza*, 5: *Habla de Sevilla y hablas americanas*, Sevilla, Publ. Universidad, págs. 243-278.
- LAMÍQUIZ, V. (1985a): "Tipología sintáctica de *coger* y *tomar*", en V. Lamíquiz y F. Rodríguez-Izquierdo (eds.), *Sociolingüística Andaluza*, 3: *El discurso sociolingüístico*, Sevilla, Publ. Universidad, págs. 51-63.
- LAMÍQUIZ, V. (1985b): "El sistema verbal idealizado y su comportamiento discursivo", en V. Lamíquiz y F. Rodríguez-Izquierdo (eds.), *Sociolingüística Andaluza*, 3: *El discurso sociolingüístico*, Sevilla, Publ. Universidad, págs. 105-120.
- LAMÍQUIZ, V. (1993): "Conexión conmutadora entre enunciados", en P. Carbonero y C. Fuentes (eds.), *Sociolingüística Andaluza*, 8: *Estudios sobre el enunciado oral*, Sevilla, Publ. Universidad, págs. 11-34.
- LAMÍQUIZ, V. y CARBONERO, P. (1987): *Perfil sociolingüístico del sevillano culto*, Sevilla, Publ. Instituto de Desarrollo Regional.
- LAMÍQUIZ, V. y PINEDA, M.A., eds. (1983): *Sociolingüística Andaluza*, 2: *Encuestas del habla urbana de Sevilla -nivel culto-*, Sevilla, Publ. Universidad.
- LAMÍQUIZ, V. y ROPERÓ, M., eds. (1987): *Sociolingüística Andaluza*, 4: *Encuestas del habla urbana de Sevilla -nivel popular-*, Sevilla, Publ. Universidad.
- MARTÍN CID, M. (2001): "Lo dicho y su reflejo: los interpretadores del discurso en el habla culta de Sevilla", en P. Carbonero y R. Guillén (eds.), *Sociolingüística Andaluza*, 12: *Identidad lingüística y comportamientos discursivos*, Sevilla, Publ. Universidad, págs. 187-219.
- MIRÓ, R. y PINEDA, M.A. (1990a): "Determinación sociolingüística de la presencia / ausencia del pronombre personal sujeto", en P. Carbonero y M^a T. Palet (eds.), *Sociolingüística Andaluza*, 5: *Habla de Sevilla y hablas americanas*, Sevilla, Publ. Universidad, págs. 37-44.
- MIRÓ, R. y PINEDA, M.A. (1990b): "Perífrasis de infinitivo en el habla urbana de Sevilla", en P. Carbonero y M^a T. Palet (eds.), *Sociolingüística Andaluza*, 5: *Habla de Sevilla y hablas americanas*, Sevilla, Publ. Universidad, págs. 59-83.
- MORENO, M^a J. (1993): "Oraciones condicionales 'incompletas'. Estudio contrastivo entre el habla urbana de la ciudad de México y el habla urbana de Sevilla, nivel culto", en P. Carbonero y C. Fuentes (eds.), *Sociolingüística Andaluza*, 8: *Estudios sobre el enunciado oral*, Sevilla, Publ. Universidad, págs. 61-85.

- OLLERO, M. (1993): "Léxicos excluyentes y léxicos preferentes en el habla popular de Sevilla", en P. Carbonero y C. Fuentes (eds.), *Sociolingüística Andaluza*, 8: *Estudios sobre el enunciado oral*, Sevilla, Publ. Universidad, págs. 269-307.
- OLLERO, M. y PINEDA, M.A., eds. (1992): *Sociolingüística Andaluza*, 6: *Encuestas del habla urbana de Sevilla -nivel medio-*, Sevilla, Publ. Universidad.
- PALET, M^a T. (1990): "El diminutivo en el habla urbana de Sevilla (nivel popular)", en P. Carbonero y M^a T. Palet (eds.), *Sociolingüística Andaluza*, 5: *Habla de Sevilla y hablas americanas*, Sevilla, Publ. Universidad, págs. 25-36.
- PINEDA, M.A. (1995): *Sociolingüística Andaluza*, 10: *Texto informatizado del habla urbana de Sevilla*, Sevilla, Publ. Universidad.
- PINEDA, M.A. y OLLERO, M. (1993): *Sociolingüística Andaluza*, 7: *Concordancias e índices del habla urbana de Sevilla*, Sevilla, Publ. Universidad.
- RAMOS, M^a M. (1993): "La intensificación del adjetivo y el adverbio en el discurso (sintaxis oral)", en P. Carbonero y C. Fuentes (eds.), *Sociolingüística Andaluza*, 8: *Estudios sobre el enunciado oral*, Sevilla, Publ. Universidad, págs. 183-213.
- RODRÍGUEZ-IZQUIERDO, F. (1985a): "Procedimientos de topicalización en el habla culta de Sevilla", en V. Lamíquiz y F. Rodríguez-Izquierdo (eds.), *Sociolingüística Andaluza*, 3: *El discurso sociolingüístico*, Sevilla, Publ. Universidad, págs. 31-49.
- RODRÍGUEZ-IZQUIERDO, F. (1985b): "Procedimientos de negación en el habla de Sevilla", en V. Lamíquiz y F. Rodríguez-Izquierdo (eds.), *Sociolingüística Andaluza*, 3: *El discurso sociolingüístico*, Sevilla, Publ. Universidad, págs. 137-154.
- RODRÍGUEZ-IZQUIERDO, F. (1990): "Aportaciones léxicas del habla popular de Sevilla al Diccionario de la Real Academia Española", en P. Carbonero y M^a T. Palet (eds.), *Sociolingüística Andaluza*, 5: *Habla de Sevilla y hablas americanas*, Sevilla, Publ. Universidad, págs. 197-214.
- RODRÍGUEZ-IZQUIERDO, F. (1993): "Aportaciones léxicas al DRAE del habla culta de Sevilla", en P. Carbonero y C. Fuentes (eds.), *Sociolingüística Andaluza*, 8: *Estudios sobre el enunciado oral*, Sevilla, Publ. Universidad, págs. 257-267.
- ROMERO HEREDIA, J. (1996): *Estructuras semántico-sintácticas de los verbos saber, conocer, comprender y entender en español. Estudio en el habla urbana de Sevilla*, Tesis Doctoral, Universidad de Sevilla.
- ROMERO HEREDIA, J. (2001): "De alguna diferencia entre construcciones con *sé* y construcciones con *conozco*", en P. Carbonero y R. Guillén (eds.), *Sociolingüística Andaluza*, 12: *Identidad lingüística y comportamientos discursivos*, Sevilla, Publ. Universidad, págs. 141-148.
- SAMPER, J.A., HERNÁNDEZ CABRERA, C. y TROYA, M., eds. (1998): *Macrocorpus de la norma lingüística culta de las principales ciudades del mundo hispánico*, edición en CD-ROM, Las Palmas de Gran Canaria, Publ. Universidad.
- SANTANA, J. (1998): "Si como marcador discursivo: estudio en el habla urbana de Sevilla", *Interlingüística*, 7 (Actas del XI Encuentro de la Asociación de Jóvenes Lingüistas), págs. 211-216.

- SANTANA, J. (1999): "Las perifrasis verbales de gerundio en el habla urbana culta de Sevilla: caracterización y uso discursivo", *Philologia Hispalenses*, 13, págs. 175-195.
- SANTANA, J. (2001a): "Usos del gerundio en el habla urbana de Sevilla (nivel popular)", en P. Carbonero y R. Guillén (eds.), *Sociolingüística Andaluza*, 12: *Identidad lingüística y comportamientos discursivos*, Sevilla, Publ. Universidad, págs. 119-140.
- SANTANA, J. (2001b): "Análisis de la forma verbal *cantaré* en la lengua hablada de Sevilla", en E. Méndez, J. Mendoza y Y. Congosto (eds.), *Indagaciones sobre la Lengua. Estudios de Filología y Lingüística Españolas en Memoria de Emilio Alarcos*, Sevilla, Publ. Universidad, págs. 139-153.
- SANTANA, J. (2003a): *Sociolingüística Andaluza*, 13: *Las oraciones condicionales: estudio en la lengua hablada*, Sevilla, Publ. Universidad.
- SANTANA, J. (2003b): "La alternancia *cantaré* / *voy a cantar* en el habla de Sevilla", *Actas del XXIII Congreso Internacional de Lingüística y Filología Románicas*, II/1, págs. 319-330.
- SANTANA, J. (2004): "Preposición + (artículo) + que relativo: análisis en la norma lingüística culta panhispánica", *Boletín de Lingüística*, 21, Caracas, págs. 66-91.
- SIBÓN, T. (1993a): *La coordinación asimétrica. Análisis del habla urbana de Sevilla*, Tesis Doctoral, Universidad de Sevilla.
- SIBÓN, T. (1993b): "Estudio lexemático sintáctico del sistema verbal. *Sentir, oír, escuchar* en el habla de Sevilla (niveles popular y culto)", en P. Carbonero y C. Fuentes (eds.), *Sociolingüística Andaluza*, 8: *Estudios sobre el enunciado oral*, Sevilla, Publ. Universidad, págs. 237-255.

2.8. Participantes en el Proyecto

La elaboración del corpus del habla de Sevilla y las investigaciones realizadas sobre el mismo han sido y son posibles gracias a la participación de algunas instituciones y de muchas personas que han intervenido en las diferentes fases del proyecto: los profesores miembros del Grupo de Investigación "Sociolingüística Andaluza: Estudio Sociolingüístico del Habla de Sevilla" (reconocido dentro del Plan Andaluz de Investigación de la Junta de Andalucía, con el código HUM-141), los licenciados que generosamente han participado como colaboradores honorarios adscritos al grupo, los alumnos internos del Departamento de Lengua Española, Lingüística y Teoría de la Literatura de la Universidad de Sevilla, los informantes sevillanos que han accedido pacientemente a la realización de las encuestas, el Secretariado de Publicaciones de la Universidad de Sevilla, el Vicerrectorado de Investigación, el Ministerio de Educación y Ciencia, que a través de la Dirección General de Investigación Científica y Técnica subvencionó el Proyecto PB87-0924, la Dirección General de Universidades de la Junta de Andalucía con sus Ayudas periódicas para apoyar a los Grupos de Investigación y Desarrollo.

llo Tecnológico Andaluces, y todos aquellos que de una u otra forma han dedicado atención y apoyo a este proyecto.

Dirección del Proyecto:

Pedro Carbonero Cano
Departamento de Lengua Española, Lingüística y Teoría de la Literatura
Facultad de Filología
Universidad de Sevilla
c/ Palos de la Frontera, s/n
41004 SEVILLA
pcarbonero@siff.us.es
Tfno.: 954 55 15 34

Página web del Grupo de Investigación "Sociolingüística Andaluza":

<http://www.sociolinguiticaandaluza.us.es>

3. EL CORPUS LINGÜÍSTICO DEL HABLA DE ALMERÍA

El corpus del habla de Almería ha sido el objetivo más ambicioso del grupo de investigación ILSE desde su fundación. Tres eran nuestros objetivos iniciales: 1) conseguir unos documentos útiles tanto para quienes estuvieran interesados en *aspectos gramático-textuales* relativos al uso de ciertos elementos discursivos en contextos orales, como para aquellos otros investigadores que prefirieran detenerse en la aceptabilidad de ciertas construcciones y expresiones lingüísticas, o en la posibilidad de seleccionar, entre las varias formas que circularan, las que fueran preferibles para determinados fines; 2) ofrecer unos textos rentables para quienes se aproximasen a las grabaciones con un objetivo más próximo al *análisis crítico del discurso*; por ejemplo, la representación discursiva de los inmigrantes africanos o las estrategias lingüísticas utilizadas a la hora de opinar sobre la droga o sobre la drogadicción; y 3) acceder a datos rentables para el estudio de la *competencia metacomunicativa de los almerienses*, especialmente mediante la petición a los entrevistados de valoraciones sobre su habla.

Desde el primer momento, además, quienes formábamos parte del Grupo nos dimos cuenta de que para cumplir esos objetivos, se hacía necesario llegar a un acuerdo en torno, como mínimo, a los siguientes aspectos: 1) la utilización de una *metodología* lo suficientemente rigurosa como para poder hablar de auténtica representatividad de la muestra de testigos seleccionados; 2) la realización de un intenso proceso de reflexión en torno a los géneros discursivos y a los modelos de interacción que utilizaríamos para completar el corpus, comenzando, como es natural, por la entrevista; 3) el procedimiento más

adecuado para ofrecer *grabaciones* de calidad, y *transcripciones* útiles de todos los textos elegidos para formar parte del corpus del habla de Almería; 4) el *almacenamiento* de toda la documentación derivada de este corpus; y 5) la incorporación de la *historia* como un elemento determinante a la hora de tomar decisiones en torno a la división espacial de la ciudad.

3.1. Los tipos de interacción

Queríamos que nuestro corpus recogiera distintas situaciones comunicativas tanto bidireccionales como unidireccionales; para el tipo de interacciones nos valimos, aunque con cifras de tiempos diferentes, de la clasificación ya establecida por los autores del Corpus del habla de Málaga (Ávila, 1999: 23); en cuanto a las primeras, las bidireccionales, establecimos las siguientes situaciones: a) conversaciones en presencia, con conocimiento de los informantes –entrevista–, pero también sin conocimiento de los interlocutores –conversación oculta; b) comunicaciones a distancia, y dentro de éstas, bien «libres» –conversación telefónica familiar y en radio y TV– bien «condicionado» –entrevista telefónica. En cuanto a las unidireccionales, también separamos: a) la situación cara a cara conferencia, clases, sermones– y b) a distancia –TV y radio.

BIDIRECCIONAL				UNIDIRECCIONAL	
En presencia		A distancia		En presencia	A distancia
Libre	Condicionado	Libre	Condicionado	Libre	Libre
A)	B)	C)	D)	E)	F)
Conversaciones en casa	Entrevista	Conversación telefónica familiar y en Radio y TV	Entrevista telefónica	Conferencia Clases	TV Radio
Grupo A= 20 per. 4hs. 30'	Grupo B= 100 per. 45 hs. 37'	Grupo C = 60 per. 25 hs. 30'	Grupo D= 20 6 hs. 43'	Grupo E = 10 per. 5hs. 10'	Grupo F= 10 per. 5 hs. 32'

Tabla 1

Grupo A: a) conversaciones en casa; b) conversaciones en otros lugares (bares, centros, etc.)
Grupo B: a) en casa de los entrevistados; b) en ambientes escolares o universitarios, etc.
Grupo C): a) grabaciones telefónicas familiares y amistosas; b) solicitando servicios; c) a medios de comunicación; d) registradas en contestadores automáticos.
Grupo D): a) programas de opinión en radio y televisión; b) entrevistas radiofónicas; c) entrevistas en medios de comunicación a través del teléfono.
Grupo E): a) conferencias; b) ponencias; c) retransmisiones de los medios de comunicación.

En la actualidad, tenemos un corpus de 220 personas, de las cuales 100 son las entrevistadas directamente y el resto, 120, participa en los diferentes tipos de situaciones contextuales, con lo que podemos dar cabida a los distintos parámetros: 1) *+planificado* (discursos unidireccionales, entrevistas radiofónicas, entrevistas semidirigidas cara a cara; *-planificado* (entrevistas semidirigidas cara a cara, conversaciones telefónicas familiares y grabaciones inadvertidas); 2) *+contenido interactivo* (grabaciones inadvertidas, entrevistas semidirigidas) *+contenido de información* (entrevistas dirigidas, discursos unidireccionales conferencias, informativos); *+narrativos* (discursos unidireccionales, conversaciones inadvertidas); *-narrativos* (conversaciones telefónicas, entrevistas dirigidas); *ambos interlocutores activos -situación de interlocución-* (entrevistas semidirigidas, conversaciones familiares, etc.); *receptor pasivo situación de monolocución-* (programa de radio y televisión, informativos). El tiempo total de grabaciones es de 96 horas y 2 minutos, de las que 47 hs. y 37' corresponden a las entrevistas (situaciones bidireccionales, cara a cara y no libre), en tanto que el resto, 48 hs. y 25' se distribuyen entre las restantes situaciones comunicativas.

A modo de ejemplo, y debido a las limitaciones espaciales con las que contamos, vamos a centrarnos únicamente en el proceso seguido para la consecución de datos relacionados con el género bidireccional condicionado (la entrevista semidirigida).

3.2. Caracterización cuantitativa de la ciudad de Almería

El Municipio de Almería se encuentra dividido administrativamente en ocho distritos, de los cuales los siete primeros corresponden a la capital, y el octavo a los Anejos (Cabo de Gata, Cuevas-Retamar); y aunque La Cañada y El Alquíán administrativamente pertenecen al distrito séptimo, para nuestro trabajo los hemos agrupado con el resto de Anejos. Cada uno de estos ocho distritos se encuentra, a su vez, dividido en secciones, unidades de población tan irregulares que pueden tanto referirse a unas pocas manzanas de viviendas como a todo un barrio; los ocho distritos de Almería se distribuyen en 107 secciones.

Los datos de población recogidos en los padrones municipales se refieren a secciones y distritos que son unidades administrativas y, por tanto, algo convencionales, de modo que no suelen coincidir con agrupaciones naturales.

Por ello, y dadas las características del estudio que pretendíamos llevar a cabo, procedimos al establecimiento de unidades de población más acordes con el concepto de barrio, entendiendo como tal toda agrupación tradicional y natural de población que así es considerada por los propios habitantes, con algunos rasgos diferenciales en relación con barrios o zonas limi-

trofes. Debe entenderse, sin embargo, que el establecimiento de límites rígidos no es posible más que en algunos casos.

De este modo, partiendo de las 107 secciones administrativas, teniendo en cuenta el proceso de desarrollo de la ciudad y atendiendo a la realidad sociocultural de la misma, hemos establecido cuarenta barrios que constituyen agrupamientos mucho más naturales que los ocho distritos y que las 107 secciones. Son éstos: Pescadería, Chanca, Almedina, San Cristóbal, Catedral, Paseo Viejo, Paseo Nuevo, Las Huertas, Alfareros, Regocijos, Quemadero, Barrio Esperanza, Los Ángeles, Amatisteros, Araceli, Piedras Redondas, Los Almendros, Alhadra, San Luis, Los Molinos, El Puche, Barrio Alto, Regiones, El Diezmo, Altamira, Nueva Andalucía, Gachas Colorás, Mediterráneo, Estación, Tagarete, Cortijo Grande, Ciudad Jardín, 500 Viviendas, Zapillo, Jaúl, Nueva Almería, La Cañada, El Alquíán, Cuevas-Retamar y Cabo de Gata.

Había que empezar, en consecuencia, por la caracterización cuantitativa de dichos barrios. Ello nos obligó a la recopilación de una serie de datos que presentamos a través de unas cuantas tablas (I, II, III). Pero como estas no reflejaban a primera vista las peculiaridades de cada zona, trasladamos algunos de los datos allí recogidos a mapas en los que se podrían apreciar mucho mejor determinadas diferencias y distribuciones².

Con estos datos, establecimos un 'nivel' para cada barrio en una escala de 0 a 10, que se determinó asignando a cada uno de los niveles la siguiente puntuación: estudios universitarios de grado superior, 10 puntos; estudios universitarios de grado medio, 8.5; estudios de Bachillerato o Formación Profesional de 2º grado, 7; estudios de Educación General Básica o Formación Profesional de 1º grado, 5.5; estudios primarios, 4; sin estudios pero no analfabetos absolutos, 2.5; analfabetos, 1. Exactamente igual ocurre con los tramos de edad, según los grupos que previamente habíamos establecido (18-35; 35-55 y +55).

También dispusimos de datos, en sus correspondientes mapas, de la densidad de población según los barrios y las zonas.

3.3. Establecimiento de diez zonas para el estudio sociolingüístico de Almería

Para el mejor tratamiento de los datos conseguidos y para procurar que la selección de informantes se ajustase lo más posible a la realidad poblacional

² Los datos los hemos obtenido del Padrón municipal de 1993; los relativos a la distribución por sexos del de 1994. Hay que hacer constar, por tanto, que se trata de datos absolutos, no de una muestra, por lo que aunque es evidente que puede haber errores en la interpretación posterior de algunos datos (p.e., una persona que declara tener solamente estudios primarios puede ser tan culta como un universitario, aunque la hayamos incluido en el grupo de nivel cultural bajo), estos han de ser por fuerza insignificantes.

de la ciudad, de modo que no se concentrasen en este o aquel barrio, en este o aquel nivel de estudios o tramo de edad, procedimos del modo siguiente:

En primer lugar, agrupamos esos cuarenta barrios en diez zonas para el estudio lingüístico, numeramos esas zonas de acuerdo con su antigüedad y después trasladamos los datos de las tablas I, II y III a las tablas Ib, IIb y IIIb que representan esas diez zonas y a las tablas Ic, IIc y IIIc que representan esas mismas diez zonas pero distribuidas a su vez en una serie de subzonas:

Zona 1: *Pescadería*: En ella incluimos los barrios de La Chanca, Pescadería y Almedina. Se trata, pues, del núcleo sobre el que se forma la ciudad de Almería.

Zona 2: *Centro*: que formamos con los barrios que se fueron generando a partir de ese núcleo inicial, fundamentalmente en dos momentos: zona incluida en el amurallamiento de Jayrán del siglo XI, a saber, Catedral, Paseo Viejo, Paseo Nuevo; y la expansión al noreste de esa muralla, con los barrios de Las Huertas, a lo largo de los caminos de salida hacia Pechina y Granada, (zona de San Sebastián-Plaza de Santa Rita); y al norte, Alfareros y Regocijos.

Zona 3: *Barrio Alto*: En esta zona reunimos el Barrio Alto, Regiones, Diezmo y Amatisteros, esto es, los que se formaron a lo largo de las otras grandes vías de salida de la ciudad primitiva hacia Pechina, fuentes de Alhadra y la Vega; se trata de las actuales Calle de Murcia-Real del Barrio Alto-Carretera de Níjar y Alcalde Muñoz-San Juan Bosco-Carrera del Perú.

Zona 4: *Estación*: Integramos en ella los barrios de Estación, Ciudad Jardín, Altamira, Mediterráneo y Nueva Andalucía, que se fueron formando entre los años cuarenta (Ciudad Jardín) y setenta.

Zona 5: *Esperanza*: se conformó con los barrios de Santa Isabel (Esperanza), Quemadero y Los Ángeles. Si bien los barrios del Quemadero y Santa Isabel tienen mayor antigüedad que los de la zona 4, el núcleo fundamental por número de habitantes de esta zona es el barrio de los Ángeles, que empieza a formarse en los primeros años de la posguerra, con algunas viviendas sociales, y continúa su expansión desde los años 1965-1970.

Zona 6: *Vega*: Es una agrupación de barrios algo más heterogénea que las anteriores, pero los que agrupamos en ella son todas expansiones de la ciudad hacia la vega de Levante. Los más antiguos son El Tagarete y las 500 Viviendas; y aunque el origen del Zapillo hay que situarlo en antiguas casas de pescadores, su desarrollo ha sido posterior. Parte de Gachas Colorás-La Goleta, Cortijo Grande y Nueva Almería son las más recientes expansiones hasta el punto de que aún están en desarrollo.

Zona 7: *Los Almendros*: La formamos con los barrios de Los Almendros, Piedras Redondas y Araceli.

Zona 8: *El Puche*: Se corresponde con el barrio del mismo nombre; es de características tan diferenciadas que no permite su agrupamiento con otros barrios.

Zona 9: *Alhadra*: incluimos en esta zona los barrios de Alhadra, San Luis, y Los Molinos.

Finalmente, agrupábamos en la que hemos denominado Zona 10 todos los *Anejos*, aunque, como es de esperar, las características de cada uno de ellos difieran de las del resto.

3.4. Selección de informantes

En cuanto al proceso de selección de los informantes, se suele aludir constantemente a Labov (1966) para quien si una comunidad dada se distribuye de acuerdo con factores extralingüísticos como los que hemos tenido en cuenta, es suficiente una muestra del 0.025% de informantes para que sea proporcionalmente representativa. De acuerdo con este criterio, del total de 165.913 habitantes de Almería la muestra debería componerse de 42 informantes. Sin embargo, para favorecer una mayor comprensión, hemos preferido ampliar considerablemente ese porcentaje y proponer una muestra muy superior: 100 entrevistados directamente, lo que nos lleva a un índice del 0.060 del total de la población de la ciudad, más 120, que participan, como antes indicábamos, en los diferentes tipos de situaciones contextuales.

Para proceder a la selección de los 100 entrevistados, establecimos tres tramos de edad:

- 1: de 18 a 35 años
- 2: de 36 a 55
- 3: + de 55

y tres niveles socioculturales, según unas pautas señaladas con amplitud en nuestro Proyecto:

- A: nivel alto
- B: nivel medio
- C: nivel bajo

Para determinar de qué zonas deben seleccionarse y cuántos de esos informantes deben corresponder a cada uno de esos tramos de edad y niveles

socioculturales, subdividimos las subzonas en 84 núcleos menores de población, barriadas, si se quiere, pero con entidad propia reconocida de una manera más o menos general como tales por el sentimiento ciudadano.

De esas barriadas hemos seleccionado a los informantes de acuerdo con el porcentaje que corresponde a cada tramo de edad y nivel de estudios, de manera proporcional al número de habitantes adultos. Para ello hemos aplicado la siguiente fórmula:

$$\frac{T_i * T_n * 100}{\text{Total habitantes adultos de la ciudad} * \text{Total de habitantes adultos de la Zona}}$$

Siendo: T_i : Habitantes de la Zona por tramo de edad (1,2,ó 3)

T_n : Habitantes de la Zona por nivel de estudios (A,B o C).

De este modo, obtenemos una tabla en la cual se nos indica cuántos informantes se deben seleccionar de cada zona en la columna *total*, y en el subtotal de las columnas 1-A, 1-B, etc. se nos marca el número de informantes que se deben seleccionar en cada uno de esos tramos-niveles:

TABLA 2. Perfil de las personas seleccionadas según porcentaje de población por zona, tramos de edad y nivel de estudios

	1-A	1-B	1-C	2-A	2-B	2-C	3-A	3-B	3-C	Total
Zona 1										
Subtotal	0.09	0.90	2.24	0.11	0.90	2.15	0.07	0.67	1.68	8.82
Test.		1	2		1	2		1	2	9
Zona 2										
Subtotal	0.88	2.25	2.66	1.01	2.57	3.01	0.99	2.58	3.10	19.05
Test.	1	2	3	1	2	3	1	3	3	19
Zona 3										
Subtotal	0.14	1.17	1.98	0.15	1.31	2.19	0.12	1.01	1.75	9.81
Test.		1	2	1	1	2		1	2	10
Zona 4										
Subtotal	1.21	3.16	2.51	1.70	4.36	3.45	0.86	2.20	1.74	21.19
Test.	1	3	2	2	4	4	1	2	2	21
Zona 5										
Subtotal	0.10	1.15	2.28	0.12	1.31	2.58	0.07	0.77	1.51	9.89
Test.		1	2		1	3		1	2	10

TABLA 2. Perfil de las personas seleccionadas según porcentaje de población por zona, tramos de edad y nivel de estudios (continuación)

	1-A	1-B	1-C	2-A	2-B	2-C	3-A	3-B	3-C	Total
Zona 6										
Subtotal	0.30	1.73	1.92	0.39	2.05	2.21	0.22	1.24	1.39	11.45
Test.		2	2		2	3		1	1	11
Zona 7										
Subtotal	0.01	0.29	0.94	0.01	0.28	0.83	0.00	0.17	0.49	3
Test.			1			1			1	3
Zona 8										
Subtotal	0.00	0.22	1.04	0.00	0.19	0.88	0.00	0.05	0.25	2.63
Test.			2			1				3
Zona 9										
Subtotal	0.07	0.71	1.29	0.07	0.79	1.44	0.04	0.45	0.84	5.69
Test.		1	1		1	2			1	6
Zona 10										
Subtotal	0.09	1.02	1.83	0.10	1.15	2.05	0.06	0.76	1.38	8.44
Test.		1	2		1	2		1	1	8
Total	2.90	12.60	18.68	3.67	14.91	20.77	2.45	9.90	14.13	99.97
Test.	2	12	19	4	13	23	2	10	15	100

Zonas: 1=Pescadería; 2=Centro; 3=Barrio Alto; 4=Estación; 5=Esperanza; 6=Vega; 7=Almendros; 8=Puche; 9=Alhadra y 10=Anejos.

Tramos de edad: 1=de 16 a 30 años; 2=de 31 a 55 años; 3=>de 55.

Nivel de estudios: A=alto; B=medio; C=bajo.

De cada zona establecimos previamente un cuadro; como ejemplo presentamos a continuación el de la Zona 1 (Pescadería):

TABLA 3. Personas que deben ser entrevistadas en cada zona según tramos de edad (1=de 18 a 35; 2=de 36 a 55 y 3=>de 55) y nivel de estudios (A=alto; B=medio, y C=bajo)

ZONA 1: PESCADERÍA										
		1-A	1-B	1-C	2-A	2-B	2-C	3-A	3-B	3-C Total
01a	Alcazaba	0.00	0,01	0.03	0.00	0.01	0.03	0.00	0.01	0.04 0.11
01a	Cerro del Hambre	0.00	0,07	0.32	0.00	0.06	0.28	0.00	0.05	0.21 0.98

TABLA 3. Personas que deben ser entrevistadas en cada zona según tramos de edad (1=de 18 a 35; 2=de 36 a 55 y 3= >de 55) y nivel de estudios (A=alto; B=medio, y C=bajo) (continuación)

ZONA 1: PESCADERÍA											
		1-A	1-B	1-C	2-A	2-B	2-C	3-A	3-B	3-C	Total
01a	Cuevas-Chamberí	0.00	0.05	0.20	0.00	0.04	0.17	0.00	0.02	0.10	0.58 :13.29:2
01a	Hoya	0.00	0.00	0.03	0.00	0.00	0.04	0.00	0.00	0.05	0.18
01a	Las Piedras	0.01	0.05	0.09	0.01	0.04	0.09	0.01	0.05	0.10	0.44
01b	Ctra. Málaga	0.01	0.08	0.22	0.01	0.08	0.20	0.00	0.05	0.14	0.78
01b	Huerta Ruano	0.00	0.16	0.38	0.00	0.14	0.35	0.00	0.11	0.26	1.40 :3.76:4
01b	San Roque	0.01	0.06	0.14	0.01	0.06	0.13	0.01	0.05	0.12	0.58
01c	Almedina	0.02	0.12	0.22	0.02	0.12	0.22	0.02	0.12	0.24	1.09
01c	Parque-Atarazanas	0.03	0.09	0.10	0.05	0.14	0.16	0.02	0.05	0.06	0.69
01c	Parque-Fondeadero	0.01	0.06	0.08	0.01	0.06	0.07	0.00	0.04	0.05	0.37 :3.76:4
01c	Plaza Pavía	0.01	0.07	0.11	0.01	0.07	0.11	0.01	0.05	0.07	0.53
01c	Reducto	0.00	0.10	0.32	0.00	0.09	0.29	0.00	0.08	0.25	1.08
	Subtotal	0.09	0.90	2.24	0.11	0.90	2.15	0.07	0.67	1.68	8.82 9

Se nos indica que debíamos seleccionar 8.82 informantes, que redondeábamos en 9. En las columnas de subtotales se nos señala que hemos de optar por 2 del tramo-nivel 1-C (índice 2.24) y otros dos del tramo-nivel 2-C (índice 2.15). También parecía claro que podíamos seleccionar dos más del tramo 3-C (índice 1.68). Nos quedaban tres informantes por seleccionar: nos decidimos por uno del tramo 1-B, otro del tramo 2-B y finalmente, otro del tramo 3-B, porque así nos lo marcaba el índice del subtotal; restaba únicamente decidir de qué parte de esa zona serían esos informantes. En general, nos decidimos por la que marcaba el índice más alto. Pero no siempre fue así, pues si ya se había seleccionado uno de la misma zona, se optaba por un informante de otra aunque en ella el índice fuese menor. Puede observarse que entre los informantes que se deben seleccionar en la Zona de Pescadería, que hemos descrito como ejemplo, la fórmula aplicada no determinaba ninguno del nivel A en ninguno de los tramos de edad; en cambio, en los niveles C los índices establecían la selección de dos informantes. Ello concordaba rigurosamente con el nivel de estudios de los habitantes de la zona.

3.5. La entrevista y el cuestionario

La entrevista es el género más utilizado en los corpus urbanos. Nos hemos inclinado, para el nuestro, por entrevistas poco estructuradas pero dirigidas y con la particularidad de que, además, en el tramo último de las mismas la actitud de los propios hablantes ante su lengua se constituye también en tema de conversación. Naturalmente, hay opiniones a favor y en contra de su uso para la obtención de datos a partir de los cuales efectuar análisis del discurso oral e incluso para el análisis propiamente sociolingüístico³. En todo caso, no se puede dudar de la licitud de la elección, siempre y cuando la entrevista se contextualice como tal y no necesariamente como una muestra de conversación espontánea⁴. Es curioso, sin embargo, contrastar cómo la conversación grabada sin conocimiento de los entrevistados no tiene tal tipo de críticas, como si la única habla que hubiera que estudiar fuera la coloquial, o no hubiera más posibilidad de conseguirla que con este procedimiento⁵.

Nuestras cien entrevistas se hicieron siguiendo este cuestionario:

Primera parte: cuestiones varias

- I. Sin grabación. Toma de contacto. Comentario sobre el tiempo, noticia actual, temas de Almería, etc.
- II. Grabación:
 - A. 1) ¿A qué se dedica usted? ¿Trabaja, estudia...?
 - 2) ¿Cómo se desarrolla un día normal de su vida?
 - 3) ¿De cuánto tiempo libre dispone? ¿En qué lo emplea?
 - 4) ¿Qué hace los fines de semana y festivos?
 - 5) ¿Tiene alguna afición particular?

³ Entre las críticas positivas sobre la entrevista compartimos, por encima de otras, la de Pilleux (1995: 98) al defender su riqueza expresiva con fines semiolingüísticos: «en una entrevista se puede encontrar el discurso *narrativo*, esto es, una secuencia de eventos que hace mención al pasado; el discurso *expositivo*, estructurado en torno a un tema determinado; el discurso *argumentativo*, en el que la secuencia puede mostrar razones a favor o en contra de una determinada posición u opinión, y el discurso *descriptivo*, usado para detallar o describir a personas, lugares o cosas. En efecto, otras fórmulas para la obtención de información (pensamos sobre todo en la encuesta) no aseguran tal riqueza estilística.

⁴ Nancy Vázquez ya reflexionó muy atinadamente sobre este asunto en la comunicación que presentó al XXV Simposio de la Sociedad Española de Lingüística (Zaragoza, 11-14 de diciembre de 1995) y que se tituló «Corpus de lengua hablada en la ciudad de A Coruña: el rol del entrevistador en la conversación semidirigida».

⁵ Nos referimos a opiniones ya antiguas que han venido repitiéndose en distintas épocas y en distintas disciplinas; se basan en las dificultades que proporciona la búsqueda de estilo de habla espontáneo e informal, porque se considera que es este el más deseado.

- 6) ¿Le gustaría dedicar su tiempo libre a otro tipo de actividades?
- 7) ¿Cómo son sus vacaciones: viaja, se queda en casa...?
- 8) ¿Cómo serían sus vacaciones ideales?
- B. 1) ¿Desde cuándo vive usted en este barrio? ¿Qué opina de él? ¿Dónde vivía anteriormente?
- 2) ¿Existe alguna asociación que se preocupa por el mismo?
- 3) ¿Qué mejoras cree que son necesarias?
- 4) ¿Qué es lo más típico de su barrio: fiestas, monumentos, ambiente...?
- C. 1) ¿Qué opina sobre los inmigrantes africanos en Almería?
- 2) ¿Qué opina sobre el problema de la droga?
- 3) ¿Cree que se deberían legalizar las drogas blandas?
- 4) ¿Qué piensa de la juventud actual?
- 5) Cuénteme una anécdota curiosa o graciosa que le haya ocurrido en su vida.

Segunda parte: actitudes lingüísticas

- I. A) Cuando ha salido fuera de Almería, ¿le han reconocido por su modo de hablar de dónde es?
- B) ¿En qué rasgos? Entonces, ¿cree que Almería tiene sus rasgos lingüísticos propios?; por ejemplo, ¿reconoce un tonillo particular? ¿Cómo lo valora? ¿Conoce usted algún otro aspecto del habla de Almería (muletillas, palabras...)?
- C) ¿Considera la forma de hablar de Almería más próxima a la de Murcia o a la del resto de Andalucía?
- D) ¿Qué nombre da usted a la lengua o al modo de hablar que utiliza?
- E) ¿Dónde cree usted que se habla mejor el español?
- F) ¿A usted le gustaría hablar como ellos? ¿Por qué?
- II. A) ¿Cree que en Almería capital habla todo el mundo con la misma corrección? ¿Hay algún barrio con características peculiares en su forma de hablar?
- B) A su juicio, ¿hay diferencias entre el habla de los hombres y el habla de las mujeres? ¿Por qué?
- C) ¿Quiénes hablan mejor, los viejos o los jóvenes? ¿Por qué?
- D) ¿Dónde cree que se habla mejor: en Almería capital o en la provincia?
- III. A) Según su opinión, ¿ayuda la televisión a hablar mejor o peor? ¿Por qué?
- B) ¿En qué tipo de programa aparecen los andaluces en televisión (informativos, variedades, humor...)?
- C) ¿Se siente identificado con su forma de hablar?
- D) ¿Cree usted que a los niños se les debe corregir para que hablen como en Madrid?

En cualquier momento, el cuestionario se podía alterar por parte del entrevistador, en favor de una mayor fluidez. Efectuamos su diseño de manera que fuera útil para los futuros estudiosos interesados en investigaciones sociolingüísticas (aspectos fónicos, morfosintácticos, léxicos, discursivos, de actitudes, etc.), también para quienes pretendan realizar estudios sobre el discurso oral en general (siempre en el marco de la entrevista como género), o, finalmente, para los que quieran ocuparse del análisis crítico del discurso. De esta forma, el cuestionario participaba de los objetivos comentados al comienzo de la descripción del Corpus del habla de Almería.

Tanto la selección de bloques temáticos como la selección de preguntas concretas, así como el orden en el que aquellos y esas se presentan, no son producto del azar, naturalmente, sino que obedecen a objetivos muy concretos. La primera parte, denominada genéricamente «cuestiones varias», está dividida en tres bloques temáticos. En el primero de ellos, se tratan aspectos relacionados con la vida cotidiana de los testigos, tales como el trabajo, el tiempo libre, las aficiones o las vacaciones. De esta forma, aseguramos la paulatina implicación del entrevistado en la dinámica propia de la entrevista, con temas sencillos, de fácil respuesta y alejados por completo de asuntos socialmente conflictivos. El segundo bloque, aun manteniéndose dentro de lo cotidiano y sencillo, supone un pequeño paso en lo que significa solicitar información sobre un tema algo más comprometido: la situación del barrio, sin que ello exija asunción de responsabilidad, por lo que el discurso no ha de perder fluidez alguna. El tercer bloque de esta primera parte del cuestionario, en el que se supone que el recelo hacia el entrevistador, de haber existido, se habrá aliviado o desaparecido, se divide en dos partes, la más conflictiva y la más sencilla. El tema de las drogas, que, si para algunos es un tema del que pueden hablar de manera inhibida, para otros es fuente de incomodidad y obliga a un prudente distanciamiento discursivo; hemos de contar con esta diversidad de conductas. Los últimos dos temas son muy sencillos (la juventud actual y la anécdota) buscando que el entrevistado muestre el registro más informal o al menos la actitud más relajada.

En la segunda parte del cuestionario se solicita información sobre aspectos lingüísticos relacionados con Almería. Lo planteamos, eso sí, más como conversación que como preguntas breves, porque estábamos convencidos de que de esta forma, dejando hablar todo lo que se deseara, extraeríamos información no prevista pero de sumo interés. Además, de esta manera, era más fácil identificar los tipos de discurso sobre actitudes, valoraciones, etc. de la lengua que caracterizan (aunque no de manera exclusiva) a un determinado testigo (Bañón, 1993).

De todas formas, hemos ofrecido a los entrevistadores absoluta libertad para la incorporación de otros temas de interés, sobre todo si se ha observado algún apunte en esta dirección por parte del entrevistado (incluyendo, como es natural, los minutos previos a la grabación); a la par se le ha hecho

una serie de consideraciones que han de seguir. Esta permisividad fue posible porque nos aseguramos de que los entrevistadores fuesen buenos conocedores de la estructura y de la complejidad pragmalingüística del género. Nos pareció interesante, pues, realizar algunos seminarios en los que se iniciase un proceso de reflexión global en el que era necesario contemplar y valorar también las intervenciones de los entrevistadores, y no sólo de los entrevistados, aun a sabiendas de que lo que será objeto de estudio preferente y de cuantificación es el discurso de estos últimos. Esta reflexión, por lo demás, no debía limitarse a los momentos previos al inicio de la grabación de las cintas, sino que con la llegada de cada una de las grabaciones debía iniciarse también un proceso suplementario de análisis y observación que constatare nuestras impresiones iniciales o las completase con nuevos datos que revertirían favorablemente en las restantes. Tanto de la detección de errores como de la observación de aciertos se aprende. En todo caso, una transgresión de lo más adecuado en lo que debiera ser la función socio-discursiva del entrevistador no invalida necesariamente una grabación, dado que siempre hay que contar con la reacción del entrevistado, y es su discurso —como decíamos— el que más nos interesará. No obstante, al estar ante una actividad dialógica e interdependiente, parece claro que una intervención se entenderá sólo en el contexto proporcionado por la intervención anterior. Si la actitud es colaboradora, poco puede llegar a importar una equivocación o una reacción inoportuna del entrevistador, puesto que, siendo así, es difícil que se resienta la fluidez conversacional. Por otro lado, es normal que estas situaciones se produzcan a veces incluso en entrevistadores expertos. Es cierto, por último, que los sentidos de adecuado e inadecuado en términos conversacionales difieren según sea un tipo u otro de entrevistados y según sea una u otra su relación con los entrevistadores.

Los asuntos abordados preferentemente durante esos seminarios fueron los siguientes: 1. Los límites entre la entrevista y otros géneros que puedan solaparse con ella voluntaria o involuntariamente. 2. El cuidado de los márgenes de la estructura global del género: preentrevista y presentación, despedida y postentrevista. 3. Simetría y asimetría en el desarrollo semiótico y pragmático de la entrevista. 4. La duración de la entrevista y la extensión de los turnos. 5. La petición de información. 6. La promoción de respuestas. 7. El cuidado en las transiciones y en los cambios de tema. 8. La comunicación no verbal en la entrevista con fines semiolingüísticos.

3.6. *La grabación, la transcripción y el almacenamiento del corpus*

La etapa de transcripción y almacenamiento se llevó a cabo siguiendo, en buena parte, los criterios metodológicos del PRESEEA (Moreno Fernán-

dez, 1996, 2000) para los materiales del «Proyecto para el estudio sociolingüístico del español de España y América»; así se han transcrito íntegramente las grabaciones en ortografía ordinaria, utilizando el procesador de textos «Word TM». Los archivos correspondientes a cada grabación han sido almacenados en PC, y en discos compactos como copia de seguridad. Cada transcripción dispuesta para el análisis cuenta con dos correcciones realizadas por personas diferentes. Las convenciones utilizadas para transcribir los materiales están basadas en el sistema SGML y siguen las normas internacionales de la TEI (*Text Encoding Initiative*). Se trata de una serie de marcas o etiquetas que reflejan los más diversos rasgos de los textos transcritos; se han utilizado las mismas etiquetas que aparecen en Moreno Fernández (1996, 2000).

3.7. *Producción científica desarrollada a partir del corpus ILSE*

Las investigaciones microestructurales en el contexto del discurso oral realizadas a partir de esta muestra se han efectuado, siempre que ha sido posible, en los márgenes establecidos por la superestructura y por la macroestructura del texto. Pensamos, por ejemplo, en la importancia del mencionado cruce de jerarquías (la funcional y la social) en el estudio de la atenuación, o en las diferencias que es posible encontrar en este mismo proceso semiolingüístico a la hora de hablar de un tema socialmente relevante o de otro con menor relevancia social y, por tanto, sobre el que no cuesta tanto asumir responsabilidades enunciativas.

Los documentos orales incluidos en el Corpus del habla de Almería han servido a los miembros del grupo ILSE no sólo para el estudio del comportamiento discursivo de los hablantes, sino que también han propiciado la reflexión teórica sobre la oralidad y el establecimiento de propuestas tipológicas relativas, por ejemplo, a las unidades de segmentación y a los marcadores del discurso (Cortés y Camacho, 2005). Camacho (2003a, 2003b y en prensa) también ha estudiado minuciosamente la repetición en el discurso oral. La reflexión sobre los géneros que han sido utilizados a partir de ejemplos concretos del corpus ILSE han dado lugar igualmente a algún trabajo, por ejemplo, sobre la dinámica interlocutiva de la entrevista con fines semiolingüísticos (Bañón, 1997b). Como hemos indicado en varias ocasiones, el análisis crítico del discurso era otro de los campos en los que los miembros del grupo de investigación se mostraron interesados. Fruto de ese interés son los trabajos de Bañón sobre el uso de la expresión *a lo mejor* (1999) y, sobre todo, su investigación a propósito de la representación discriminatoria de los inmigrantes africanos en el discurso oral de los testigos almerienses entrevistados (1997a). Como ejemplo del interés por el estudio de

las actitudes lingüísticas de los almerienses y de su competencia metacomunicativa, podríamos mencionar el trabajo de Torres Montes (1997).

A partir del Corpus del habla de Almería se llevará a cabo, además, en los próximos años un proyecto titulado *Repetición y estructuras de iteración en el discurso oral en español: Uso de los mecanismos, tipología y estratificación social*, subvencionado por el Ministerio de Educación y Ciencia (HUM2004-01551 FILO). Igualmente, está en proceso de realización una tesis doctoral sobre la interacción en contextos jurídicos, cuya autora es Susana Ridao.

3.8. Personas de contacto para la consulta del corpus o para la obtención de datos complementarios

Luis Cortés Rodríguez

lcortes@ual.es

Antonio M. Bañón Hernández

amhernan@ual.es

Universidad de Almería

Departamento de Filología

Ctra. Sacramento s/n

La Cañada de San Urbano

04120 Almería

Página electrónica del grupo ILSE:

www.grupoilse.org

BIBLIOGRAFÍA

- ÁVILA, A.M. (1999): *Léxico de frecuencia del español hablado en la ciudad de Málaga*, Málaga, Universidad.
- BAÑÓN, A.M. (1993): «Adolescencia, variación lingüística, competencia metacomunicativa y enseñanza de la lengua», *Estudios de lingüística de la Universidad de Alicante*, 9, 1993, págs.253-285.
- BAÑÓN, A.M. (1997a): «La representación discriminativa de los inmigrantes africanos en el discurso oral», *Discurso. Teoría y Análisis*, 21/22, págs.103-132.
- BAÑÓN, A.M. (1997b): «Reflexiones sobre la dinámica interlocutiva en la entrevista con fines semiolingüísticos», *Revista de Investigación Lingüística*, 1, págs.7-36.
- BAÑÓN, A.M. (1999): «Análisis sintáctico-semántico de la locución *a lo mejor* en español hablado», en *Actas del XI Congreso Internacional de la Asociación Española de Lingüística y Filología de la América Latina. Las Palmas de Gran Canaria (1996)*, José Antonio Samper y Magnolia Troya (coords.), Universidad de Las Palmas/Librería Nogal, Las Palmas, tomo I, págs.137-145.
- CAMACHO, M^a. M. (2003a): *La repetición en el discurso oral*, Tesis doctoral, Almería, Universidad.
- CAMACHO, M^a. M. (2003b): «Algunos oficios interactivos de la repetición en el discurso oral: funciones eulógicas y dislógicas», *Oralia*, 6, págs.119-146.
- CAMACHO, M^a. M. (en prensa): «La repetición como metadiscurso», en *Aspectos meta-discursivos del español*, Navarra, Universidad.
- CORTÉS, L. (1992): «Materiales para un proyecto de estudio sociolingüístico del habla de Almería», en *Estudios de español hablado (Aspectos teóricos y sintáctico cuantitativos)*, Almería, IEA./Campus Universitario de Almería, págs. 29-50.
- CORTÉS, L. y M. CAMACHO (en prensa): *Unidades de segmentación y marcadores del discurso: elementos esenciales en el procesamiento discursivo oral*, Madrid, Arco/Libros.
- LABOV, W. (1966): *The social Stratification of English in New York City*, Washington, D.C., Center for Applied Linguistics.
- MORENO, F. (1996): «Metodología del «Proyecto para el estudio sociolingüístico del español de España y América (PRESEEA)», *Lingüística*, 8, págs. 257-287.
- MORENO, F. et alii (2000): «La sociolingüística de Alcalá de Henares en el «Proyecto para el estudio sociolingüístico del español de España y América (PRESEEA)», *Oralia*, 3, págs. 149-168.
- PILLEUX, M. (1995): «El análisis 'tópico del discurso' en una entrevista», *Revista de Lingüística Teórica y Aplicada*, 33, págs. 5-14.
- TORRES MONTES, F. (1997): «Actitudes lingüísticas en la ciudad de Almería», en A. Narbona y M. Roperio (eds.), *El habla andaluza*, Sevilla, Seminario Permanente del Habla Andaluza, págs. 635-647.