

*Corpus sociolingüístico de la ciudad de México*

*Materiales de PRESEEA – MÉXICO*

Volumen I

Hablantes de instrucción superior

Pedro Martín Butragueño      Yolanda Lastra

Coordinadores

El Colegio de México

*Corpus sociolingüístico de la ciudad de México*

*Materiales de PRESEEA – MÉXICO*

Volumen I

Hablantes de instrucción superior

Coordinación

Pedro Martín Butragueño y Yolanda Lastra

Entrevistas

Mario Chávez (4), Alejandra Espinosa (9), Tania García Torres Uhartt (5),  
Norohella Huerta (5), Yolanda Lastra (1), Paola Quintanar (6), Doriam Reyes  
(1), Leonor Rosado (3), Adriana Salas (1), Julio Serrano (1)

Transcripciones

Mario Chávez (2), Alejandra Espinosa (7), Tania García Torres Uhartt (16),  
Karla Herrera (1), Paola Quintanar (1), Nury Quiroz (4), Doriam Reyes (1),  
Adriana Salas (1), Eugenia Silva (4)

Revisiones

Alejandra Espinosa (9), Tania García Torres Uhartt (50), Carlos Ivanhoe Gil  
Burgoin (36), Pedro Martín Butragueño (36), Nury Quiroz (15)

Edición

Pedro Martín Butragueño

## Índice

Introducción.....pp. v-lxvi

## Corpus PRESEEA – MÉXICO

## Entrevistas con hablantes de instrucción superior

Entrevista 1.....	pp. 1-24
Entrevista 2.....	pp. 25-62
Entrevista 3.....	pp. 63-89
Entrevista 4.....	pp. 90-115
Entrevista 5.....	pp. 116-139
Entrevista 6.....	pp. 140-174
Entrevista 7.....	pp. 175-206
Entrevista 8.....	pp. 207-241
Entrevista 9.....	pp. 242-262
Entrevista 10.....	pp. 263-289
Entrevista 11.....	pp. 290-310
Entrevista 12.....	pp. 311-331
Entrevista 13.....	pp. 332-348
Entrevista 14.....	pp. 349-370
Entrevista 15.....	pp. 371-384
Entrevista 16.....	pp. 385-404
Entrevista 17.....	pp. 405-433
Entrevista 18.....	pp. 434-458
Entrevista 19.....	pp. 459-484
Entrevista 20.....	pp. 485-501
Entrevista 21.....	pp. 502-538
Entrevista 22.....	pp. 539-563
Entrevista 23.....	pp. 564-588

Entrevista 24.....	pp. 589-652
Entrevista 25.....	pp. 653-682
Entrevista 26.....	pp. 683-704
Entrevista 27.....	pp. 705-727
Entrevista 28.....	pp. 728-764
Entrevista 29.....	pp. 765-799
Entrevista 30.....	pp. 800-823
Entrevista 31.....	pp. 824-857
Entrevista 32.....	pp. 858-885
Entrevista 33.....	pp. 886-922
Entrevista 34.....	pp. 923-951
Entrevista 35.....	pp. 952-984
Entrevista 36.....	pp. 985-1006

## INTRODUCCIÓN

Pedro Martín Butragueño

EL COLEGIO DE MÉXICO

Yolanda Lastra

UNIVERSIDAD NACIONAL AUTÓNOMA DE MÉXICO y

EL COLEGIO DE MÉXICO

La ciudad de México es la mayor área urbana hispanohablante del mundo, además de una de las ciudades más grandes del planeta. Aunque tamaño y diversidad no son exactamente lo mismo, la variedad social y demográfica que encierra, así como las dificultades prácticas derivadas de sus dimensiones, desafían la posibilidad de llevar a cabo un estudio lingüístico mínimamente representativo. Las transcripciones ahora presentadas constituyen un fragmento de un proyecto sociolingüístico más amplio, uno de cuyos principales frutos es precisamente la formación del *Corpus sociolingüístico de la ciudad de México* (CSCM).

## ANTECEDENTES

Hacia 1993 habíamos empezado ya a pensar en la necesidad de un estudio sociolingüístico que se ocupara de al menos algunos aspectos de la desbordante complejidad de la ciudad de México. Existía ya una tradición de estudios de la ciudad, cuyos principales hitos eran la tesis de Marden de 1896, recogida en el volumen preparado por Henríquez Ureña en 1938; las propias observaciones de Henríquez Ureña en ese mismo volumen; la tesis de Matluck de 1951 sobre el Valle de México; los materiales recogidos sobre el habla culta y popular de la ciudad y los trabajos redactados a partir de ellos (*cf.* Lope Blanch 1986); la tesis de Perissinotto, publicada en 1975; y los materiales del centro del país del *Atlas Lingüístico de México*<sup>1</sup>. Pero faltaba, a nuestro juicio, un estudio contemporáneo que, beneficiándose de la

---

<sup>1</sup> Véase en especial la reinterpretación de algunos mapas en el libro de Moreno de Alba de 1994, así como también Lope Blanch (1983, 1990).

comparación en tiempo real con las descripciones y los materiales del pasado, acogiera algunas dimensiones sociolingüísticas consideradas sólo parcialmente hasta ese momento<sup>2</sup>.

Aquella inquietud vino a coincidir con el ambicioso propósito de organizar un “Proyecto para el estudio sociolingüístico del español de España y de América” (Moreno Fernández 1993, 1997), PRESEEA. Dado el interés del proyecto compartido, que garantizaba la posibilidad de comparar materiales y de dialogar con las experiencias de otros equipos, nos interesamos pronto en sumarnos a él. Nos pareció que PRESEEA solicitaba unos mínimos razonables, pero que además el proyecto local debía cubrir otros aspectos propios de la realidad de la ciudad de México, por lo que procedimos a desarrollar una metodología adicional específica, tanto en los grupos de informantes que habían de cubrirse como en el diseño y condiciones de la encuesta. Después de diferentes consideraciones, realizamos las primeras entrevistas en abril de 1997<sup>3</sup>.

## EL ESTUDIO SOCIOLINGÜÍSTICO DE LA CIUDAD DE MÉXICO

### *La ciudad de México*

La primera tarea consistió en delimitar qué se iba a entender por ciudad de México, por lo menos como punto de partida para el trabajo, pues a la larga habría que repensar su carácter o no como comunidad de habla, precisamente a la vista de los materiales y en especial de los datos de evaluación subjetiva (como los obtenidos en algunas de las pruebas descritas *infra*).

La llamada Zona Metropolitana de la ciudad de México (ZMCM) es un conglomerado urbano en el que habitan cerca de veinte millones de personas<sup>4</sup>. El *Censo* del año 2000, en concreto, señala una población de 18 396 677 individuos, repartidos entre el Distrito Federal (8 605 239) y su zona conurbada (9 791 438). Está formada por 75 entidades, de las cuales 16 son delegaciones del Distrito Federal y el resto municipios

---

<sup>2</sup> Habían aparecido sólo algunas propuestas de proyectos sociolingüísticos (*cf.* Zimmermann 1982, Lara y Zimmermann 1988). Véase también, por ejemplo, Rodríguez y Williamson (1981), Lastra (1992a y b).

<sup>3</sup> La investigación ha recibido apoyo de El Colegio de México y de CONACYT; en cuanto a este último, hemos recibido financiamiento para diferentes tramos de la investigación a través de dos proyectos: “Estructura fónica de la diversidad lingüística en México” (1999-2002, 27598H) y “Diversidad y variación fónica en las lenguas de México” (U44731H, 2004-2007).

<sup>4</sup> Un recuento detallado de los habitantes por entidad y de las lenguas habladas en ellas puede verse en <http://lef.colmex.mx>, en la sección de “Datos demolingüísticos” y en Martín Butragueño (2009).

conurbados del Estado de México (además de uno del Estado de Hidalgo), que circundan, salvo por el sur, al Distrito Federal<sup>5</sup>.

La figura 1 muestra la proyección geográfica de la ZMCM:

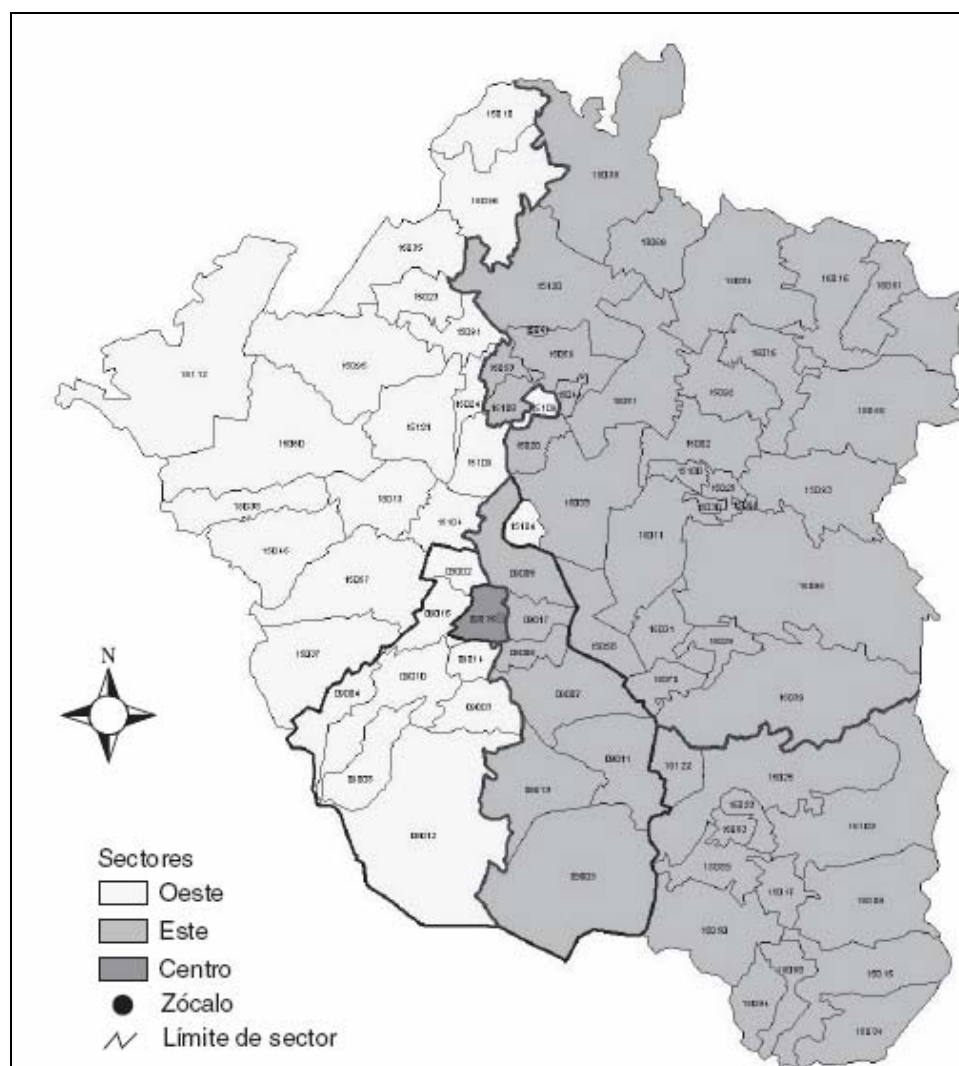


Figura 1. Zona Metropolitana de la ciudad de México.

Fuente: Graizbord y Acuña (2007, p. 301, mapa 1).

Como puede observarse, la ZMCM puede subdividirse de varias maneras. Por una parte, como se ha dicho, cabe diferenciar entre el Distrito Federal o ciudad de México

<sup>5</sup> Para una revisión de los escenarios demográficos y urbanos de la ZMCM entre 1990 y 2010, puede verse el documento de CONAPO incluido en <http://www.conapo.gob.mx/publicaciones/zmcm.htm#> (consultado el 3 de septiembre de 2008).

propiamente dicha, que en el mapa corresponde a la amplia sección en forma de gota al centro-sur de la mancha urbana, frente a la restante zona conurbada. Otra gran división se produce al diferenciar entre oriente y poniente de la ciudad y de la zona metropolitana. El grado de integración urbana, por otra parte, es diverso. En la ZMCM se incluyen grandes áreas residenciales y de servicios, pero también zonas industriales e incluso semirurales; existen también pueblos originarios que han quedado incluidos dentro de la ciudad, los cuales conservan parte de sus tradiciones y de sus redes sociales (*cf.* Mora Vázquez 2007, Ortiz Elizondo 2007).

A la hora de planear nuestro estudio, decidimos dejar de lado una parte de la Zona Metropolitana, por dos motivos: en primer lugar, por el carácter casi inabarcable de la ciudad para un pequeño equipo de trabajo; en segundo término, por considerar que la integración de los diferentes municipios conurbados responde a una realidad gradual en términos de servicios, infraestructura y, sobre todo, modos de vida urbanos. En ese sentido, decidimos considerar sólo los municipios que ya estuvieran conurbados y considerados dentro de la ZMCM para 1970, pero no a los sumados en momentos posteriores. De esa manera, se disponía del tiempo aproximado de una generación completa (considérese que el proyecto se estaba planeando a mediados de los años noventa), lo cual, a nuestro juicio, permitía defender la idea de una integración notoria tanto de las personas como de las áreas en la estructura de bienes y servicios propia de la ciudad de México. Tal decisión suponía considerar como frontera para el levantamiento de datos el área formada por todo el Distrito Federal más 11 municipios del Estado de México: Atizapán de Zaragoza, Coacalco, Cuautitlán, Chimalhuacán, Ecatepec, Huixquilucan, Naucalpan, Nezahualcóyotl, La Paz, Tlalnepantla y Tultitlán. En total, la Zona Pertinente (ZP) para nuestro estudio incluye 27 entidades, entre delegaciones y municipios. Hay que aclarar en seguida que no hemos pretendido llevar a cabo una estratificación areal de los informantes, pero sí definir unos límites geográficos y procurar un razonable reparto de las zonas consideradas.

La figura 2 muestra la llamada Zona Pertinente de la ZMCM, diferenciando en dos tonos el Distrito Federal y los municipios considerados en el Estado de México. Las delegaciones se anotan con números romanos y los municipios con números arábigos.



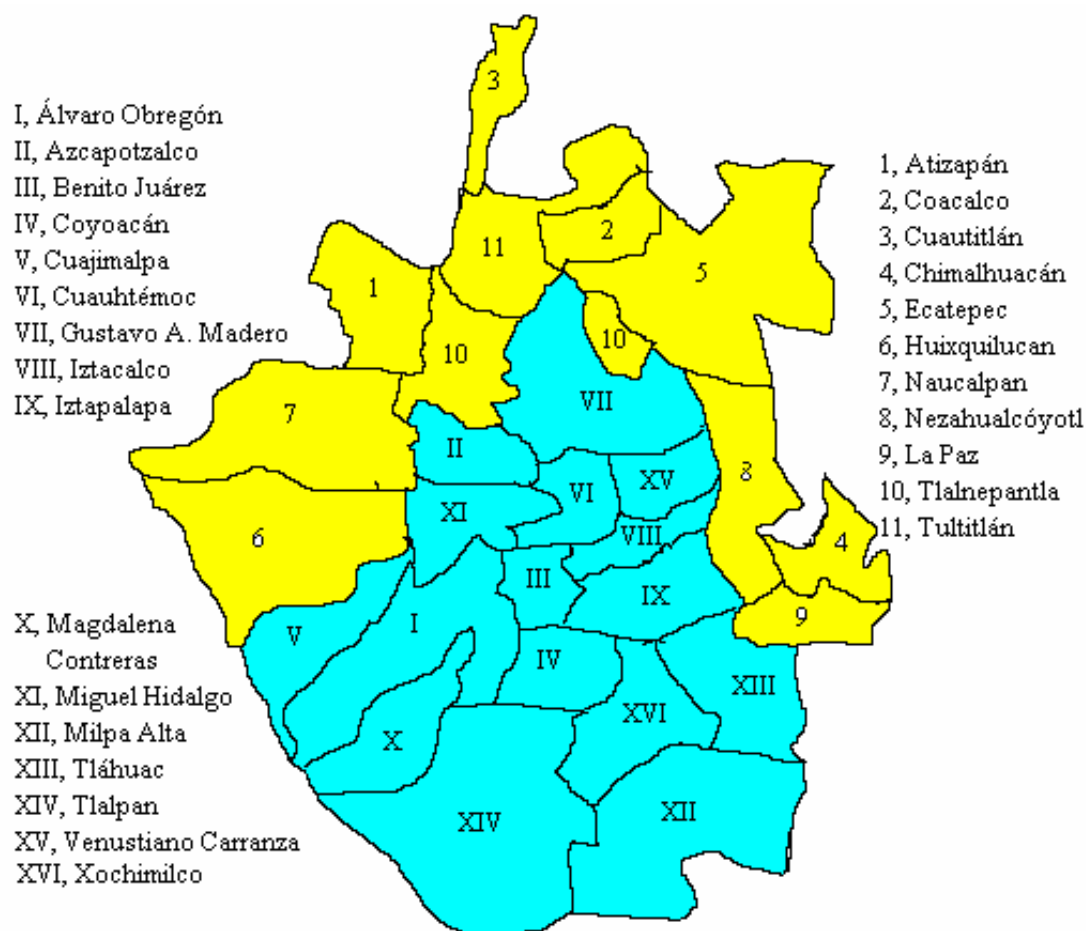


Figura 2. La llamada «Zona Pertinente» para el CSCM

En esta sección de la ciudad de México vivían en el año 2000, según el *XII Censo General de Población y Vivienda*, 15 159 386 personas, de ellas 8 605 239 en el Distrito Federal, y 6 554 239 en la zona conurbada del Estado de México<sup>6</sup>.

En el momento de diseñar el proyecto, era obvio que existían muchas diferencias internas en la ZP. La población económicamente activa (PEA) estaba formada en 1990 por 4 532 817 personas. Los ingresos de esta PEA formaban una pirámide de amplia base y vértice muy agudo, pues un 64.1% de ella percibía menos de 2 salarios mínimos<sup>7</sup>, el 26.2%

<sup>6</sup> En el momento de planear el estudio, teniendo a la vista los datos del *Censo* de 1990, el área propia de la ZP sumaba un total de 13 470 973 personas, repartidas en 8 235 744 en el Distrito Federal, y 5 235 229 en la fracción del Estado de México considerada.

<sup>7</sup> El salario mínimo sirve en México como unidad para calcular tanto los ingresos como diferentes tipos de servicios y de pagos (por ejemplo, las multas, establecidas en número de salarios mínimos). En el momento

de 2 a 5 salarios mínimos, y el 9.7% más de 5 salarios mínimos. Además, existía una distribución geográfica del ingreso. El tercio de entidades con mayor concentración de PEA que percibiera menos de dos salarios mínimos se encontraba al sur y al este de la ZP: Milpa Alta (81.8% de su población), Tláhuac (77.1%), Chimalhuacán (72.5%), Xochimilco (71.6%), Iztapalapa (70.7%), Ecatepec (69.5%), Cuajimalpa (69.4%) Magdalena Contreras (67.7%) y La Paz (66.9%). En contraste, la mayor concentración de ingresos se daba básicamente al oeste de la ZP. Algo parecido cabe decir con respecto a los estudios, pues la mayor concentración de población con estudios superiores se daba al centro y al oeste de la ZP, mientras que la menor concentración se proyectaba en el este. En cuanto a actividades productivas, las no muchas personas dedicadas al sector primario se concentraban al sur y al este, las dedicadas al sector secundario en el norte y al sector terciario al sur y en el centro<sup>8</sup>.

Otro aspecto de gran importancia a la hora de diseñar una investigación sociolingüística en la ZP es la importante presencia de grandes flujos migratorios, no sólo para describir la presencia lingüística misma del inmigrante, sino por el papel fundamental que el contacto dialectal y lingüístico tiene en el entendimiento de los mecanismos de cambio y variación lingüística. Para 1990, aproximadamente la cuarta parte (24.8%) de la población del Distrito Federal (2 046 064 personas) era de origen inmigrante; y en la fracción del Estado de México considerada en la ZP más de la mitad de los individuos (57.8%, 3 026 684 personas) tenían ese origen. Hoy día los flujos migratorios se han estabilizado en el Distrito Federal, pero no en el Estado de México<sup>9</sup>. Un estudio sociolingüístico competente tiene que tomar en cuenta estas realidades<sup>10</sup>.

---

actual (diciembre de 2009), el salario mínimo equivale a unos 1 650 pesos mensuales, es decir, a unos 125 dólares estadounidenses.

<sup>8</sup> Para más datos, véase Martín Butragueño (2010).

<sup>9</sup> El Distrito Federal muestra el saldo neto migratorio más bajo de todo el país, mientras que el Estado de México muestra el tercero más grande (INEGI 2004, p. 11).

<sup>10</sup> Las entidades con mayor componente rural son las que tienen menos inmigrantes: Milpa Alta (10%), Xochimilco (18.9%), Tláhuac (20.5%). Buena parte de los municipios mexiquenses, en contraste, están mayoritariamente poblados por personas venidas de fuera: Coacalco (71.5%), Ecatepec (64.9%), Nezahualcóyotl (59.5%). Hay que considerar, además, que los hijos de inmigrantes nacidos en la nueva entidad ya no figurarán con esta calificación.

*El corpus sociolingüístico de la ciudad de México (CSCM)*

En el marco del CSCM se han recogido hasta el momento muestras de habla de poco más de 300 personas, habiéndose acopiado —aproximadamente— unas 500 horas de grabación. El CSCM está concebido en cinco módulos: (i) el *módulo nuclear*; (ii) el módulo de *inmigrantes*; (iii) registros de *niños y adolescentes*; (iv) grabaciones de *marginados*; (v) entrevistas en *grupo*.

El *módulo nuclear* incluye registros de unas 180 personas de muy variadas características, localizadas muchas de ellas a través del cuerpo de las redes sociales de que forman parte. Todas las personas tienen más de 20 años; todas han nacido en la que hemos llamado Zona Pertinente, o bien han llegado a muy temprana edad de estados cercanos. Se trata en lo básico de entrevistas individuales —aunque no faltan ejemplos con intervenciones de otros participantes—, que en bastante casos desembocan en estructuras caracterizables como conversaciones grabadas (*infra*) y que han sido en su mayoría llevadas a cabo en el ámbito del informante (trabajo o domicilio). Una gran parte de estos materiales conforma la selección escogida para el *subcorpus PRESEEA*, que describimos *infra*: se trata de las entrevistas semiinformales de 108 individuos, a razón de 36 por cada uno de tres grupos de nivel educativo. Las 36 transcripciones correspondientes al nivel educativo alto constituyen los documentos que ahora publicamos.

El *subcorpus* de inmigrantes está planeado para incluir dos tipos diferentes de inmigrantes. Por un lado, aquellos que provienen de estados limítrofes del centro del país, y que han aportado el contingente principal de migración a la ciudad de México<sup>11</sup>. Por otro lado, se están considerando también varios subgrupos de inmigrantes provenientes de áreas más alejadas del país; se trata de movimientos menos cuantiosos en términos demográficos, pero sumamente interesantes en su dimensión cualitativa, por lo mucho que implican de procesos interesantes para el estudio del contacto dialectal<sup>12</sup>. Hasta ahora, la base de datos del CSCM incluye unas 75 personas adscritas a este subcorpus.

---

<sup>11</sup> Para una revisión de los flujos migratorios en México en diferentes momentos del siglo XX, véase Martín Butragueño (2010).

<sup>12</sup> Por supuesto, son de gran importancia también los procesos de contacto dialectal asociados a la presencia de hispanohablantes procedentes de otros países (véase al respecto Pesqueira 2008), así como los vinculados a la presencia de hablantes de lenguas indígenas; para una panorámica demolingüística en este sentido, véase Martín Butragueño (2009), y para el caso de los inmigrantes otomíes, Guerrero Galván (en prensa).

Los registros de niños y adolescentes incluyen en este momento unos 20 informantes más, en una secuencia que va de los 3 a los 19 años. Por lo pronto, se trata sólo de un grupo de control, pero el interés de proyectar los fenómenos de variación y cambio hacia edades más tempranas, bajo la hipótesis de que los aspectos variables del lenguaje se adquieren al tiempo que los categóricos, sugiere la necesidad de acrecentar la muestra bajo parámetros específicos<sup>13</sup>.

Un cuarto módulo está formado por grabaciones de otra veintena de personas etiquetadas como *marginados*. Se trata de personas carentes de los más mínimos recursos, y con pocas perspectivas de obtenerlos<sup>14</sup>. Sobra señalar la importancia social y sociolingüística de considerar un grupo de estas características, normalmente excluido también de los muestreos urbanos. En el futuro, el trabajo con ciertas redes debería aportar bastante más información de la obtenida hasta el momento.

Por fin, un quinto grupo de grabaciones focalizan la interacción de hablantes en grupos de diferente tamaño, bajo diferentes circunstancias —conversaciones informales, grupos de trabajo, reuniones planeadas, entrevistas colectivas, etcétera— que exigen el diseño de metodologías especiales tanto para su tratamiento como para su análisis<sup>15</sup>.

El grado de avance con respecto a cada módulo, debemos reconocer, es diverso, y el futuro a medio y largo plazo requerirá esfuerzos renovados, en especial si en algún momento puede planearse un proyecto de minorías lingüísticas, dentro o al margen del propio CSCM.

Aunque nos ha parecido conveniente exponer este marco general del sentido del CSCM, ciertamente lo más pertinente en este momento es referirse en más detalle, como hacemos a continuación, al subcorpus de PRESEEA.

### *El subcorpus PRESEEA – México*

---

<sup>13</sup> Considérese, además, la gran proporción de población juvenil todavía hoy, a pesar de que el patrón esté cambiando y en México empieza a darse un proceso de envejecimiento de la población. Para 1990, el 38.3% de la población del país tenía todavía menos de 15 años, y la mitad de la población tenía menos de 20 años.

<sup>14</sup> Algunos de ellos son o han sido niños criados en la calle, otros son pequeños delincuentes. En el mismo grupo podrían sumarse, por ejemplo, ciertos grupos de inmigrantes con fuertes carencias.

<sup>15</sup> Uno de los resultados más palpables asociados a este tipo de materiales es la tesis de Musselman (2002), sobre la construcción de turnos de habla, quien recogió parte de estos materiales.

Las directrices empleadas para la construcción del subcorpus PRESEEA-México responden a las expuestas a propósito del proyecto general de PRESEEA, o “Proyecto para el estudio sociolingüístico del español de España y de América”<sup>16</sup>.

Las condiciones generales de las que parte el Proyecto para incluir una comunidad de habla son básicamente tres: que se trate de un núcleo urbano hispanohablante (monolingüe o bilingüe); que esté formado por una población que al menos en parte haya estado asentada de forma tradicional en el lugar; y que esta población presente cierta heterogeneidad sociológica. Las razones que subyacen a este planteamiento son el propósito de crear un corpus del español hablado urbano contemporáneo; se necesita, además, que exista una cierta conciencia de comunidad de habla con una determinada configuración socioestilística reconocida por los hablantes; y, naturalmente, que las diferencias internas sean lo bastante abultadas como para que se amerite el esfuerzo de investigación. No parece necesario argumentar que la ciudad de México (entendida en los márgenes de la Zona Pertinente) cumple ampliamente los requisitos iniciales.

El interés de PRESEEA radica esencialmente en el español *de* las ciudades que abarca, y no las soluciones lingüísticas que se documentan *en* cada ciudad. Precisamente, y bajo ese entendido, el CSCM recoge diferentes tipos de materiales, y no sólo los solicitados por el Proyecto compartido, como se ha descrito someramente *supra*.

A continuación se describen diferentes aspectos de la metodología empleada, en términos de la selección de informantes, la estructura de la entrevista, el respaldo del material y la transcripción y etiquetado de los materiales, procurando discernir los aspectos nucleares de PRESEEA de los más generales establecidos para el CSCM.

## METODOLOGÍA

### *Selección de informantes*

---

<sup>16</sup> Los dos documentos principales en los que se describe la metodología del proyecto general son Moreno Fernández (1997 y 2003). El primero aparece en una publicación que recoge los trabajos presentados en una reunión efectuada en julio de 1996, auspiciada por la Universidad de Alcalá de Henares; en cuanto a la versión revisada y fechada en octubre de 2003, puede consultarse en <http://www.linguas.net/portalpreseea>. Aunque naturalmente nuestro equipo ha tomado en consideración el documento de 2003, la guía esencial de trabajo, dadas las fechas en que se habían levantado buena parte de las entrevistas, es la anunciada en el texto inicial. Aunque las diferencias no son muy notorias, decidimos de todos modos repetir cierto número de entrevistas para acomodarlas lo mejor posible a la nueva metodología.

Dada la variedad de ciudades estudiadas en el marco de PRESEEA, la llamada muestra-tipo tiene como objetivo principal partir de unos parámetros comparables sociológica y estilísticamente. Tal muestra tiene como objetivo primordial ofrecer universos relativos comunes, más allá de las disposiciones específicas adicionales adoptadas por cada equipo para dar cuenta más cabal de la realidad que estudian (como es el caso).

En ese sentido, PRESEEA trabaja con muestras establecidas por cuotas con afijación uniforme, las cuales fraccionan el universo relativo estudiado según la estratificación establecida a partir de cierto número de variables sociales y toman el mismo número de informantes para cada cuota. Dos ventajas asociadas sobre el muestreo aleatorio, señaladas desde la metodología del Proyecto, son la más fácil comparación entre las cuotas de la misma muestra y entre muestras diferentes, y la obligación de buscar personas fuera de los círculos más accesibles.

PRESEEA considera tres variables sociales con valor estratificadorio a la hora de construir la muestra: el género, la edad y el grado de instrucción. Por supuesto, ello no impide recoger información social adicional, que actúa con valor post-estratificadorio; es decir, que se puede aplicar a posteriori sobre los datos para buscar patrones en ellos. Entre este segundo grupo de variables se encuentran en particular la profesión, los ingresos económicos, las condiciones de alojamiento y el modo de vida. Es posible, además, combinar algunas de ellas para proyectar índices mixtos; tal es el caso del llamado nivel socio-cultural. La información necesaria para caracterizar las diferentes variables sociales se obtiene de un cuestionario sociológico aplicado al final de la entrevista, además de la propia información que pueda desprenderse de ella. El CSCM ha respetado estas solicitudes, además de añadir más información social complementaria.

La muestra-tipo de PRESEEA está formado entonces por 18 cuotas o casillas, fruto del cruce de 2 géneros  $\times$  3 grupos de edad  $\times$  3 niveles de instrucción. De hecho, se propone, en principio, grabar 3 informantes por casilla, lo cual daría un total de 54 personas (18 casillas  $\times$  3 informantes por casilla). Una muestra mayor, de 4 informantes por casilla, es decir, de 72 individuos en total, otorgaría una proporción de 1 informante por cada 25 000 en una ciudad de unos dos millones de habitantes. Queda a criterio de los investigadores, en las ciudades más populosas, estudiar sólo ciertas áreas o aumentar progresivamente el

número de informantes, hasta un máximo de 6 por casilla, lo que da un total de 108 individuos. Esta última es precisamente la decisión adoptada para la ciudad de México.

La mayoría de los informantes se han localizado por medio del conocido procedimiento de amigo de un amigo. Buena parte del proceso de ubicación y la secuencia de levantamiento de materiales se ha llevado a cabo por medio del trabajo con un número relativamente grande de redes.

Las variables sociales, por otra parte, se interpretan de la siguiente manera:

(i) *Género*. Hombres y mujeres.

(ii) *Edad*. Se consideran tres generaciones. La primera, o generación 1, está formada por las personas que tienen entre 20 y 34 años. La segunda, o generación 2, comprende individuos de entre 35 y 54 años. Por fin, la tercera, o generación 3, supone tener en cuenta informantes de 55 o más años<sup>17</sup>.

(iii) *Grado de instrucción*. Se distinguen tres segmentos. En el más bajo, o grupo 1, se considera a las personas que han recibido, como máximo, educación primaria, de modo que tuvieron acceso a la enseñanza formal como mucho hasta aproximadamente los 11-12 años, tras haber podido pasar por seis años de escolarización. Se incluye aquí también, por tanto, a los analfabetas y a quienes no terminaron la primaria. El grupo intermedio, o grupo 2, hace referencia a las personas que cursaron como máximo la enseñanza media, sea hasta la escuela secundaria, sea hasta el bachillerato. Es decir, como máximo habrán estudiado hasta los 17-18 años tras pasar por unos 12 años de escolaridad. En lo que toca a la mayor escolaridad, involucra a los informantes con enseñanza superior, sea universitaria, sea técnica superior. Los individuos típicos, en ese sentido, habrán estudiado hasta los 21-22 años de edad, tras unos 16 años de escolaridad, aunque los años de estudio pueden superar este umbral<sup>18</sup>.

Las variables de postestratificación propuestas son las siguientes<sup>19</sup>:

(iv) *Nivel de ingresos*. La recomendación de PRESEEA, que considera que esta variable debe definirse localmente, es establecer cinco variantes o categorías. Provisionalmente, podría pensarse para la ciudad de México en una escala derivada del

<sup>17</sup> Se ha anotado también, desde luego, la edad precisa de cada persona.

<sup>18</sup> Asimismo se ha anotado el nivel preciso, los años de estudio y el carácter de los estudios efectuados.

<sup>19</sup> Los datos específicos para cada una de las variables siguientes en lo que toca a cada uno de los informantes en concreto, tanto para los 36 de este volumen, como para los 36 de nivel medio y para los 36 de nivel bajo, aparecerán de forma conjunta en el tercer volumen.

número de salarios mínimos. Una posibilidad sería establecer lo siguiente: nivel 1, hasta 2 salarios mínimos; nivel 2, de 3 a 5 salarios; nivel 3, de 6 a 8; nivel 4, de 9 a 12; nivel 5, de 13 en adelante<sup>20</sup>.

(v) *Condiciones de alojamiento*. El nivel más bajo, el 1, incluye las viviendas sin comodidades sanitarias y de difícil acceso; el intermedio, o 2, tiene en cuenta a las personas que viven en casas o departamentos modestos; el superior, que es el 3, alberga las casas o departamentos elegantes y espaciosos, con muchas comodidades.

(vi) *Profesión*. PRESEEA establece una división en cinco grupos con respecto a la profesión. Se trata de los siguientes segmentos: 1, chachareros, ropavejeros y vendedores ambulantes, obreros no especializados urbanos, obreros campesinos, servicio doméstico, servicios no especializados; 2, pequeños comerciantes, secretarios y oficinistas, obreros especializados, artesanos, mecánicos, vendedores en tiendas, cobradores, ayudantes técnicos, policías y guardias, soldados; 3, profesionales universitarios, personal docente de educación media y primaria, pequeños empresarios y productores, mandos intermedios, técnicos, supervisores; 4, profesionales universitarios de libre ejercicio, gerentes medios del sector público y privado, militares con graduación, medianos empresarios y productores, docentes universitarios; 5, altos funcionarios del poder ejecutivo, legislativo y judicial, altos oficiales del ejército, grandes empresarios privados, grandes hacendados, altos ejecutivos del sector público y privado.

(vii) *Modo de vida*. Como es sabido, el concepto de modo de vida permite relacionar estructuras de pequeño tamaño, como las redes sociales, con entidades de mayor tamaño, como es la comunidad de habla; el modo de vida descansa en los parámetros de trabajo, familia y tiempo libre. Se distinguen tres modos. En el modo de vida 1, los individuos se caracterizan por pertenecer a una unidad primaria de producción (agricultura, pesca, pequeños servicios). Existen relaciones cooperativas entre compañeros de profesión y la familia entera se encuentra implicada en la producción. Suele producirse el autoempleo. Y se tiene escaso tiempo libre: cuanto más se trabaja, más se gana. Es propio de redes sociales

---

<sup>20</sup> Ofreceremos una propuesta para PRESEEA-México a la vista de los materiales completos también en el tercer volumen, a partir de la información recopilada en la base de datos. Hay que considerar que, además de los ingresos propiamente dichos, no siempre fácilmente declarados por los informantes, o no siempre con exactitud, puede obtenerse información al respecto a partir de estimaciones indirectas: tipo de trabajo, de alojamiento, viajes realizados, escuelas a las que ha acudido la familia, número y tipo de automóviles, quizá entre otros datos.



estrechas. En el modo 2, las personas tienen empleo en un sistema de producción que no es controlado por los trabajadores. Se trabaja para ganar un sueldo y poder disfrutar de períodos de tiempo libre. Las relaciones laborales se encuentran separadas del ámbito familiar. Se da cierta movilidad laboral. Se establecen redes estrechas de solidaridad con los compañeros y los vecinos. Los hablantes adscritos al modo de vida 3 poseen profesión cualificada, capaz de controlar la producción y de dirigir los trabajos de otras personas. El tiempo de vacaciones está dedicado al trabajo. Se trabaja para ascender en la jerarquía y adquirir más poder. Se mantiene una actitud competitiva con los colegas<sup>21</sup>.

(viii) *Origen*. PRESEEA señala, como orientación de carácter general, que los hablantes candidatos a informantes deberían *a*) haber nacido en la ciudad; *b*) o bien haber llegado a ella antes de cumplir los diez años; *c*) o bien llevar viviendo allí más de veinte años, siempre y cuando su origen lingüístico no sea marcadamente diferente. En el subcorpus PRESEEA-México se ha procurado no tener casos de *c*), por los problemas que entraña la definición lingüística de pertenencia a una comunidad de habla. Todos los datos pertinentes para cada persona (lugar de nacimiento, fecha de llegada, edad de llegada, origen de los padres y del cónyuge) se han anotado. En el marco más general del CSCM sí se ha considerado, explícitamente, a diferentes grupos de inmigrantes, tanto de regiones cercanas y limítrofes, como de lugares más alejados geográfica y lingüísticamente.

Las seis últimas variables se han establecido sólo desde la perspectiva del CSCM.

(ix) *Ascenso social*. Se han considerado tres situaciones: 1, nivel social igual al de los padres; 2, nivel social por encima del de los padres; 3, nivel social por debajo del de los padres.

(x) *Grupo étnico*. Se pueden dar cinco posibilidades: 0, no se pertenece a un grupo étnico minoritario; 1, se trata de un bilingüe activo; 2, es un bilingüe pasivo; 3, es hijo de hablantes de otra lengua; 4, es miembro cultural de alguna comunidad donde se habla otra lengua.

---

<sup>21</sup> PRESEEA deja abierta la puerta para trabajar con modos de vida adicionales. En otro trabajo (Lastra y Martín 2000), hemos propuesto distinguir varios submodos de vida con respecto a los datos del CSCM. Se trata de los siguientes. Dentro del modo 1, los submodos A (tradición semirural), B (empleo informal) y C (pequeños comerciantes); en el modo 2, los submodos D (subempleados), E (obreros) y F (burócratas); y dentro del modo 3, los submodos G (o cuellos blancos) y H (profesiones liberales) (p. 32).

(xi) *Ubicación geográfica*. Para cada informante se ha anotado el domicilio donde reside; de esa forma la colonia, la entidad (delegación o municipio) o el área de la ZP pueden servir como elementos de postestratificación.

(xii) *Viajes y lecturas*. También de cada persona se han anotado los viajes realizados, dentro y fuera de México y el tipo de lecturas que suelen llevarse a cabo (libros, periódicos, revistas). También se ha registrado la clase de programas de radio y televisión que se escuchan y se ven<sup>22</sup>.

(xiii) *Estilo de habla*. Además de las distinciones internas que pueden llevarse a cabo dentro de la entrevista individual (por ejemplo, según su carácter general, de entrevista semiinformal o de conversación grabada; por los temas abordados; por los segmentos temporales, comparando el comienzo con el final), disponibles en todos los casos, sea que la grabación quede incorporada en el subcorpus de PRESEEA o no, un buen número de los materiales recopilados para el CSCM permiten la consideración de al menos otros dos estilos de habla diferentes: las respuestas a preguntas referidas a palabras aisladas y la lectura de un texto (*infra*).

(xiv) *Evaluación subjetiva*. Los datos de evaluación subjetiva proceden básicamente de dos de las pruebas descritas *infra*: la prueba de seguridad lingüística y la de actitudes y creencias.

### *La entrevista*

La entrevista llevada a cabo para el CSCM consta en general de tres partes: la entrevista o conversación grabada; una serie de pruebas lingüísticas; un cuestionario en el que se completa la información requerida para organizar las variables sociales descritas a grandes rasgos en el apartado anterior. La grabación se ha llevado a cabo por medio de grabadoras MINIDISC y, en especial, por medio de grabadoras DAT, que permiten la colocación de dos micrófonos. Siempre se han empleado micrófonos de solapa.

La inmensa mayoría de las entrevistas se han llevado a cabo en ámbitos propios del informante (domicilio, trabajo, escuela); existen dos buenos motivos para haber procedido de esa manera. La primera razón es de orden práctico: las inmensas distancias que hay que

---

<sup>22</sup> Junto con otros parámetros, como el grado de escolaridad, estas anotaciones pueden servir para establecer otro tipo de variables, como la de contacto con la norma.

recorrer en la ciudad de México hubieran hecho muy difícil pedir a los informantes que se desplazaran a un lugar específico para levantar las grabaciones; de esta manera, se ha podido dar más variedad geográfica a las áreas de residencia de los individuos registrados en el CSCM, tanto en el Distrito Federal como en el Estado de México, y aun cuando nunca hemos pretendido llevar a cabo una estratificación geográfica, sí es posible postestratificar la zona de residencia; el sesgo geográfico hubiera sido mucho mayor de haber tenido forzosamente que pedir a los informantes que acudieran a un punto determinado, o a varios puntos. La segunda razón tiene más que ver con el propósito general de una colecta sociolingüística: dado el objetivo último, la documentación del habla cotidiana en entornos más o menos naturales, parece más coherente acudir a los lugares de actividad ordinaria de los informantes. Dentro de este procedimiento, se ha procurado seleccionar los espacios que ofrecían mejores condiciones acústicas, al tiempo que se escogía el equipo más adecuado para la ocasión (por ejemplo, más o menos voluminoso según las dificultades materiales o de seguridad del lugar de grabación; micrófonos omnidireccionales o unidireccionales según las condiciones sonoras, etc.).

A continuación describimos las tres secciones de la grabación.

*a) Conversación grabada.*

La metodología de PRESEEA establece que las entrevistas deben tener una duración mínima de 45 minutos. Salvo muy contadas excepciones, las entrevistas individuales del CSCM son al menos de una hora de grabación, aunque no faltan los casos en que la conversación con el informante se ha prolongado por más tiempo; de hecho, junto con la grabación propia de las pruebas lingüísticas, llevada a cabo siempre después de la conversación, no pocas de las grabaciones llegan a durar dos horas.

La mayoría de las entrevistas llevadas a cabo para el CSCM y todas las del subcorpus de PRESEEA son individuales, es decir, con un solo informante. Tal perspectiva general no impide que en algunos casos haya un participante secundario —que salvo alguna excepción sólo interviene esporádicamente—, que haya a veces algunas interrupciones derivadas del entorno (algún familiar, una llamada telefónica, la voz de un cliente o de un colaborador, etc.) y que llegue a haber alguna persona más presente en la grabación, aun cuando no intervenga. No sólo no nos parece problemático, sino que, en coincidencia con la mejor tradición de

trabajo en sociolingüística, consideramos que esta consecuencia normal de trabajar en los entornos del informante colabora en la naturalidad de la situación (de hecho, permitiría incluso el análisis de fragmentos específicos como un estilo de habla diferenciado).

Dentro de cierta libertad y flexibilidad, los temas que se han inducido con más frecuencia dentro de las entrevistas corresponden precisamente a los tres parámetros esenciales cuya resolución caracteriza de forma cualitativa a los modos de vida. Se trata de la vida familiar (la infancia y el núcleo paterno; los juegos infantiles; los hijos y su crianza; el contraste generacional; el origen geográfico de las personas y de las costumbres); el trabajo y sus problemas (los estudios realizados y la dificultad para encontrar trabajo; las tensiones que se dan en el ámbito laboral; las relaciones con los compañeros de trabajo y con los jefes; el dinero y las perspectivas y limitaciones derivadas); el tiempo libre y otras dimensiones de la vida individual (la vida en la ciudad de México, su historia y sus dificultades; las festividades y las costumbres capitalinas, pretéritas y actuales; la televisión, el cine, los libros; los viajes; las comidas, etc.). Desde luego, se ha pretendido que los temas se encadenen de manera natural, y no producir un repertorio de intervenciones organizadas como pregunta y respuesta. Un requisito básico ha sido otorgar prioridad a lo que estuviera contando el informante. Ciertamente, no todos estos temas han surgido en todas y cada una de las entrevistas, pero sí hay bastantes recurrencias, y puede decirse que han sido los temas principales en el conjunto de todas las entrevistas del CSCM en general y en las 108 de PRESEEA en particular. En el cuadro 5 (*infra*) se sintetizan los temas más característicos de las 36 entrevistas presentadas ahora.

En su conjunto, nos parece que puede decirse que los materiales incluyen suficiente variedad discursiva, en términos de ofrecer fragmentos expositivos, narrativos, argumentativos, descriptivos y conversacionales.

#### *b) Pruebas lingüísticas.*

En más o menos las dos terceras partes de las entrevistas individuales realizadas para CSCM se han aplicado cuatro pruebas específicas adicionales. Estas pruebas no forman parte de PRESEEA, y no se incluyen en las transcripciones que presentamos ahora. Las mencionamos a grandes rasgos, de todas maneras. Se trata de (i) una lista de palabras preguntadas de forma aislada, a la manera de un breve cuestionario dialectal, exponiendo el concepto y solicitando

la voz; la finalidad es obtener datos de fonética segmental en un estilo mucho más atento al habla, pero oral y no de lectura. En segundo término, (ii) se ha levantado una sencilla prueba de seguridad lingüística. En tercer lugar, (iii) se ha procedido a una prueba de lectura con un texto cuya finalidad es documentar un estilo mucho más atento a la ejecución lingüística; el texto se preparó en especial para documentar patrones entonativos. Por fin, en cuarto y último término, (iv) se ha efectuado una encuesta de actitudes y creencias lingüísticas formada por diecisiete preguntas.

(i) Lista de palabras. Las respuestas, que se obtienen por el método de pregunta indirecta, buscan documentar un estilo muy atento a lo expresado con respecto a tres variables vocálicas y tres variables consonánticas, tal como se anota a continuación. La inmensa mayoría de las palabras aparecen en el *Atlas Lingüístico de México*, tenido en cuenta para poder ampliar el análisis por medio de la comparación geolingüística, en caso de considerarse interesante. Las palabras no se preguntan en el orden en que aparecen a continuación, sino de forma alternada; además, pueden servir para documentar el uso de más de un sonido:

Vocales. Debilitamiento: *golpes, franceses, pistola, chiste, policía, pesos, kilos, suspirar*. Cierre vocálico: *pecho, noche, mayo, pollo, maestro, teatro, cuerpo, perro*. Hiatos y diptongos: *peor, maestro, voltear, teatro, maíz, raíz, cohete, almohada, se ahoga, reüma, baúl, aúlla*.

Consonantes. En ataque silábico: *cajón, lejos, ceja, mojado, cansado, chiste, pecho, noche, mayo, pollo, aire, ropa, perro, madrina, tres, tlacoyo, atleta*. Implosivas patrimoniales: *verdad, pared, ciudad, admisión, platicar, comer, suspirar, sudor, cuerpo, viernes, marzo, regar, el revés*. Implosivas cultas: *cápsula, óptica, septiembre, concepción, submarino, obstáculo, observar, futbol, ritmo, atmósfera, admiración, doctor, lección, examen, extranjero, octubre, magnesia, ignorante, persignarse, magna, Iztacihuatl, Popocatépetl*.

(ii) Prueba de seguridad lingüística. La seguridad se mide al comparar las respuestas a las preguntas “¿cómo dice usted?, ¿cómo dicen los demás?, ¿cuál es la más correcta?” con respecto a una lista de pares léxicos y gramaticales; a mayores coincidencias, mayor seguridad. La lista de pares se anota a continuación, considerando que se preguntan de forma intercalada.

Seguridad léxica: *gelatina – jaletina; bilet – lápiz labial; kinder – jardín de niños; mofle – escape; ring – cuadrilátero; show – espectáculo; biho – tecolote; zipper – cierre.*

Seguridad gramatical: *ayer vinimos – ayer venimos; desvariar – disvariar; la abogado – la abogada; la sartén – el sartén; mis hermanos son medio tontos – mis hermanos son medios tontos; el azúcar – la azúcar; dijiste – dijistes; piecito – piececito; pancito – panecito; un poquito de agua – una poquita de agua.*

(iii) Texto leído. Como se ha señalado, la principal finalidad del texto es documentar varios patrones entonativos de baja frecuencia; puede tener otros usos. Por otra parte, son bien conocidas las limitaciones de una prueba leída para obtener materiales de utilidad sociolingüística. El texto incluido es el siguiente:

Mira, hay cosas que contadas no se sabe si son verdad o son mentira. Lo mejor sería olvidarse hoy de lo que pasó ayer.

Mi suegra, qué crees, decía que yo no tenía dónde caerme muerto. Decía, la muy ladina, que yo no era nada para su hija. De menso no me bajaba. ¡Quién la viera y quién la ve! Y mi suegro tantito peor. "A mi hija, no te la llevas", me decía. Claro que luego las cosas fueron cambiando. Pero entonces no me querían ni nadita.

Así que me la robé. "¿Te vendrás conmigo o te quedarás acá?" "¿Qué te quieres traer?" "Tráete tu ropa, tu maleta, tus cartas y lo que quieras". "¡Pero que no sea mucho, porque todo lo tenemos que cargar!" "Vente conmigo y no lo pienses más".

Ya que se vino no fue fácil al principio. Tuvimos que irnos a otro pueblo. Y se la pasaba llorar y llorar. "¿Otra vez llorando?", le decía. "Pues, ¿qué te traes?" "¿Te estás acordando de tu mamá, verdad?" "¡Ya deja de llorar, de penar, de gimotear!"

Ya cuando nació el primer hijo las cosas cambiaron. Como que a mis suegros se les aflojó el corazón. "Miren, miren a nuestro nieto", le decían a todo mundo. Y es que está reguapo el chaparro, con esos ojotes, que voltea a todas partes cuando le hablan.

(iv) Prueba de actitudes y creencias. En esta sección se han alternado las secciones generales con las preguntas personales para evitar, en lo posible, el condicionar unas respuestas con otras; las preguntas pertenecen a cuatro grandes ámbitos: México como país, el informante mismo, la ciudad de México y la norma lingüística. En algunas preguntas se ha empleado una tabla de puntuaciones con una

escala de siete puntos; ciertamente, también son interesantes los comentarios abiertos llevados a cabo por los informantes.

México como país.

- 1, ¿Qué lengua es más importante en México, el inglés, el español o alguna lengua indígena? ¿Cuál es más útil? ¿Cuál es más bonita? ¿Por qué?
- 2, ¿Hablamos bien el español acá en México (país)? ¿Por qué?
- 3, ¿Dónde se habla mejor español en México (país)?
- 4, ¿Quiénes hablan mejor, y en qué orden: los yucatecos, los veracruzanos, los norteros, los del D.F.? ¿Por qué?

Centradas en el informante.

- 5, ¿Quién habla mejor en su casa? ¿Por qué?
- 6, ¿Sus jefes o superiores hablan mejor que usted?
- 7, ¿En la televisión hablan mejor que usted?
- 8, ¿Qué es más importante para que los niños aprendan a hablar bien, la escuela o la casa?
- 9, ¿Quiénes hablan mejor, los jóvenes, los padres o los abuelos?
- 10, ¿Quiénes hablan mejor, los hombres o las mujeres?

La ciudad.

- 11, ¿Cuándo se hablaba mejor, en el D.F. de hace 20 o 30 años o en el de ahora? ¿Por qué?
- 12, ¿Cree que se ha perdido esa forma tradicional de hablar en cada colonia?
- 13, ¿Quiénes hablan mejor, los venidos a México de provincia o los de México de siempre?
- 14, ¿Quiénes hablan mejor y en qué orden: los de Tepito, los de Xochimilco, los de Coyoacán, los de Satélite?

Norma lingüística.

- 15, ¿Qué es eso de "hablar mejor" de que hemos estado hablando?
- 16, Elija una de estas cuatro:
  - a) Hablar bien es respetar las reglas de las gramáticas y los diccionarios.
  - b) Hablar bien es hablar como lo hacen los escritores, profesores, periodistas, políticos.
  - c) Hablar bien es hablar con naturalidad, como nuestros padres, tal como siempre se habló.
  - d) Hablar bien es adaptarse a las circunstancias, ponerse a la altura de quien tengamos delante.
- 17, ¿Qué tal habla usted?

c) Cuestionario sociológico.

En el cuestionario sociológico se reúnen diferentes datos de la entrevista (como la fecha, el lugar, el entrevistador, observaciones referentes al lugar, ruidos, incidencias varias, etc.); del informante (nombre y apellidos, sexo, lugar y fecha de nacimiento, fecha de llegada a la ciudad si procede, domicilio, condiciones del alojamiento, lugar del nacimiento del padre y de la madre, ascenso social, modo de vida y su descripción, grado de instrucción indicando los años y si se completaron los estudios, profesión, ingresos, escuelas, automóviles, viajes, lecturas, programas de radio y televisión que escucha y ve, idiomas que habla, grupo étnico si procede, observaciones varias, etc.); y de otros participantes u oyentes (nombre y apellidos, sexo, lugar y fecha de nacimiento, fecha de llegada a la ciudad si procede, modo de vida, grado de instrucción, profesión, relación con el informante y otros oyentes, observaciones varias, etc.).

#### *Digitalización y almacenamiento del material*

Todos los materiales sonoros se encuentran digitalizados en archivos wav a 44100 Hz, 16 bits, y forman parte del acervo del Laboratorio de Estudios Fónicos de El Colegio de México.

#### *Transcripción y etiquetado*

Con respecto a los criterios de transcripción y etiquetado empleados en esta publicación, se exponen a continuación *a)* los criterios generales; *b)* el formato de encabezado de las transcripciones; *c)* las convenciones empleadas para etiquetar los diferentes tipos de unidades; *d)* las marcas referidas a aspectos prosódicos; *e)* las señales para los llamados ruidos; *f)* la forma en que se han resuelto ciertos problemas gráficos y fonéticos; *g)* las convenciones empleadas para diferentes tipos de observaciones; *h)* la manera en que se presentan ciertas notaciones especiales; *i)* el tipo de soluciones empleadas para otro tipo de problemas, como los nombres propios y en general el respeto hacia el anonimato de los informantes<sup>23</sup>.

---

<sup>23</sup> Es importante señalar que se está preparando una versión complementaria de las transcripciones según otro sistema de etiquetado, ceñida al documento “Marcas y etiquetas mínimas obligatorias”, establecida por consenso entre los equipos participantes en el proyecto PRESEEA; este documento puede descargarse de la página <http://www.linguas.net/preseca>, en su forma definitiva establecida en febrero de 2008.



a) Criterios generales.

La transcripción y el etiquetado presentados ahora parten de una serie de criterios generales<sup>24</sup>:

(i) *Ortografía ordinaria*. Debe emplearse ortografía ordinaria en general, y no transcripción fonética. Pese al interés que pudiera tener transcribir fonéticamente el material grabado, su extensión hace casi impensable pensar en ello de forma masiva, cuando menos en este momento. Por otra parte, las búsquedas lingüísticas de secuencias en el material transcrito se simplifican en la medida en que se ha empleado ortografía ordinaria.

(ii) *Procesador de textos convencional*. Se utiliza un procesador de textos convencional, de forma que la manipulación de las transcripciones sea lo más simple posible; en la versión final se ofrece el material en diferentes formatos (texto procesado, texto sin formato y *pdf* o *portable document format*).

(iii) *Código de caracteres*. Los caracteres empleados deben pertenecer al ASCII (o *American Standard Code for Information Interchange*) ampliado, conocido como ISO 8859-1, que incluye el ASCII en su rango inicial de 128 caracteres, más otros 128 adicionales, apropiados para representar particularidades gráficas de diferentes lenguas occidentales, entre ellas el español<sup>25</sup>, de forma que la manipulación de los archivos presente o futura genere, en la medida de lo posible, pocos problemas.

(iv) *Corrección por un mínimo de dos personas*. El criterio de partida se ha superado ampliamente, pues todas las entrevistas han sido revisadas cuatro veces. Aunque todo trabajo es perfectible —y es muy claro que la transcripción de un texto oral siempre puede mejorarse— creemos que pueden considerarse documentos lingüísticos válidos.

(v) *Tres tipos de etiquetas iniciales*. Se describen *infra*.

(vi) *Marcas de transcripción*. Se describen en detalle a continuación de este apartado; se emplea por lo pronto un sistema no muy complejo, bastante semejante al empleado hasta ahora en las publicaciones de algunos de los equipos de PRESEEA<sup>26</sup>.

---

<sup>24</sup> Para el establecimiento de los diferentes criterios de transcripción, tanto generales como particulares, nos fueron especialmente útiles en su momento los trabajos de Edwards (1993) y Du Bois *et al.* (1993).

<sup>25</sup> El estándar ISO 8859-1 codifica el alfabeto latino, y sus 256 caracteres son los iniciales, a su vez, del estándar ISO 10646, desarrollado en conjunción con el llamado estándar *Unicode*.

<sup>26</sup> Como se ha comentado en una nota anterior, en una versión complementaria de estas transcripciones se ofrecerá el sistema de marcado establecido en el documento de PRESEEA de 2008 sobre marcas y etiquetas.

(vii) *Respeto a una serie de principios*. Al diseñar las categorías, éstas han de ser discriminables, exhaustivas y contrastivas; el material ha de privilegiar la legibilidad, a través de la prominencia visual y la distribución espacial, otorgando proximidad a los fenómenos relacionados, separabilidad a lo diferente, proporcionando en lo posible iconicidad espacio-temporal, prioridad lógica, marcado mnemotécnico, eficiencia y abreviación; por fin, ha de buscarse sistematicidad y predictibilidad en el tratamiento computacional.

(viii) *Organización espacial de los turnos de habla*. Se presentan de modo vertical.

(ix) *Organización espacial de los comentarios*. Aparecen en forma de texto corrido.

(x) *Notación de las palabras*. Véase (i).

(xi) *Unidad de análisis*. La unidad de análisis primordial adoptada es el grupo fónico o entonativo, el cual queda marcado por barras a lo largo de la transcripción. Tales grupos tienen un papel fundamental a la hora de estructurar el discurso hablado, y constituyen un puente fundamental para delimitar las unidades pragmáticas y las unidades sintácticas.

(xii) *Notación de la prosodia*. Como se acaba de decir, el grupo entonativo es la unidad fundamental; también se marcan alargamientos, reducciones y algunos elementos enfatizados<sup>27</sup>.

(xiii) *Turnos de habla*. Se establece el participante, la conclusión de los enunciados, los traslapes e interrupciones y la longitud de las pausas.

(xiv) *Material no verbal*. En la medida de lo posible, aparece como una lista cerrada.

(xv) *Concurrencia de marcas*. En caso de que aparezcan varias marcas en un mismo punto, el orden de las marcas será el mismo que el establecido en el catálogo aquí descrito.

(xvi) *Numeración de los turnos*. Los turnos de habla deben ir numerados.

## b) Encabezado.

---

Como se asienta allí, “el posterior tratamiento informático de los textos etiquetados se hará utilizando el lenguaje de codificación XML, versión adaptada y simplificada de SGML, que facilitará el intercambio de datos y la recuperación selectiva de información” (p. 2).

<sup>27</sup> Por razones prácticas, se ha prescindido de la marcación de la melodía, aunque hubiera sido de gran interés anotar cuando menos la resolución tonemática de cada grupo entonativo.

El encabezado de las transcripciones de las encuestas publicadas ahora consta de tres partes:

(i) El número correlativo de encuesta dentro de esta publicación. Por ejemplo, “Entrevista 7” o “Entrevista 12”.

(ii) La etiqueta de identificación dentro de la base de datos del Corpus sociolingüístico de la ciudad de México. Esta etiqueta aparece como “ME-252-31M-05”, “ME-107-31M-00”, etcétera, y ha de leerse de la siguiente forma:

*ME* vale por ‘materiales de la ciudad de México’, en contraste con otros materiales procedentes de otras comunidades de habla; todas las encuestas de esta publicación llevan, por tanto, la clave ME.

252, 107, y cualquier otro número, es el que corresponde de forma correlativa dentro de la base de datos.

La tercera sección de la etiqueta hace referencia, en este orden, al nivel de instrucción: 1, o ‘bajo’; 2, o ‘medio’; 3, o ‘alto’; al grupo de edad: 1, o ‘grupo joven’; 2, o ‘grupo intermedio’; 3, o ‘grupo de más edad’; y al género, valiendo *H* por ‘hombre’ y *M* por ‘mujer’, todo ello conforme a las categorías sociales descritas *supra* y según las normas de etiquetado expuestas en Moreno Fernández (1997, p. 151)<sup>28</sup>.

Los dos últimos dígitos refieren al año de levantamiento de la encuesta, de modo que 97 vale por ‘1997’, 00 por ‘2000’ o 05 por ‘2005’.

(iii) Se incluye también un resumen de las principales características de la entrevista y del entrevistado: siglas relativas al informante, el entrevistador y otros participantes; estudios, grupo de edad, edad precisa, género, tipo de estudios y ocupación del informante; soporte y fecha de la grabación; nombre del transcriptor y los revisores; lugar donde se realizó la entrevista; principales temas abordados. En (1) puede verse un ejemplo.

(1) Entrevista 1, encabezado.

Informante: I. — Nivel alto, joven, 25 años, hombre; estudios de arquitectura; arquitecto. —

Grabado en DAT, estéreo, en junio de 1999. — Entrevistador: E. — Transcripción: N.

<sup>28</sup> “Cada código debe ir precedido de dos caracteres identificadores de la comunidad de habla y de tres dígitos correspondientes al número del informante (desde el 001); a su vez, irá seguido de dos números, que reflejen el año de recogida de los materiales (p.e. MA00111H96: Madrid, informante 001, grado de instrucción 1, generación 1, hombre, año 1996)” (p. 151, n. 49). Los códigos base, tal como se consignan en el texto de la misma página, son 21H por instrucción 2, generación 1, hombre; 23M por instrucción 2, generación 3, mujer, etcétera.

Huerta. — Revisiones: T. García-Torres (*bis*), P. Martín, C. Gil. — Casa del informante. — No hay más participantes ni audiencia, salvo en el turno 33: P. — Experiencia de trabajo, arquitectura, construcción.

c) Unidades.

Se marcan tres tipos de unidades en la transcripción: el turno de habla, el traslape y la palabra trunca (además, obviamente, de la palabra gráfica).

(i) *Turno*. Para introducir el turno, se anota la inicial del participante, siendo en general *I* el informante, *E* el entrevistador y *P* otro participante. El turno empieza siempre con minúscula, y en su final no se anota nada en especial<sup>29</sup>. Los turnos van numerados, sin importar si se trata de turnos de realimentación o de turnos normales.

(ii) *Traslapes*. Los corchetes, [ ], se emplean para todo tipo de superposiciones, con independencia de que se trate de traslapes propiamente dichos o de diferentes tipos de interrupciones<sup>30</sup>. La formulación de los traslapes da lugar en no pocas ocasiones a estructuras bastante complejas, como la expuesta en (2).

(2) Entrevista 21, turno 432 y ss.

432 I: el martes dije “no/ me voy a desayunar con mis amigas?”/ y me dormí/ de las once a la una/ descolgué el teléfono

433 E: [qué rico]

434 P: [bien a gusto]

435 E: [ay/ qué rico]

436 I: [y me dormí]

437 E: huy qué bien/ qué buena onda darse esos lujos/ de pronto/ [¿no?]

438 I: [así es]

<sup>29</sup> No se anota el tipo de pausa precisamente por su carácter de final de turno. Sí se podía haber marcado el tipo de conclusión: por ejemplo, si el hablante ha terminado de hablar plenamente o si tiene el propósito de proseguir, y en caso de que tenga tal propósito si realmente tiene la oportunidad de retomar su turno o no.

<sup>30</sup> Estamos conscientes del interés de haber marcado con mayor detalle la taxonomía de traslapes; de hecho, en las primeras etapas de transcripción y etiquetado se trabajaba con un conjunto más amplio de etiquetas, a saber: *Turno trunco por traslape*: [TEXTO] —B inicia su turno cuando A se acerca a un punto de transición. Se marca en A y en B—; *Turno trunco por interrupción*: [---TEXTO---] —B inicia su turno antes de que A se acerque a un punto de transición. Se marca en A y en B—; *Turnos simultáneos*: [=TEXTO=] —A y B inician su turno simultáneamente. Se marca en A y en B. Si B y C interrumpen simultáneamente a A, sólo se marcará la interrupción entre A y B, y el turno simultáneo entre B y C. B quedará marcado [---[=TEXTO=]---]—; *Falso turno*: <ft> TEXTO </ft> —señal de realimentación por parte de B—. La extensión del material provocó que tuviéramos que sacrificar algunas marcas y simplificar el proceso de etiquetado.

439 P: [mh]

440 E: y que lo permita/ el trabajo/ ¿no?/ porque a veces no lo permite

En el ejemplo, E y P se traslapan en 433 y 434, produciendo turnos simultáneos como reacción al comentario de I en 432; después E e I entran en competencia en 435 y 436; y, por fin, E, I y P hablan a la vez en 437, 438 y 439.

(iii) *Palabra trunca*. La presencia de palabras trucas se anota por medio de un guión breve, como en *pala-*, por *palabra*.

d) Prosodia.

Se han considerado ocho marcas prosódicas básicas, procurando acercar la notación lo más posible a la ortografía ordinaria.

(i) *Énfasis fuerte*: ¡ ¡ !!, como en ¡ ¡cuidado!!

(ii) *Énfasis moderado*: ¡ ! , tal como se muestra en ¡más despacio!

(iii) *Interrogación*: ¿ ?, de forma que se marca ¿no llegó a comer?

(iv) *Suspensión voluntaria*: ..., como en el siguiente ejemplo:

(3) Entrevista 16, turno 14.

14 I: [¿sí?] o sea esa es mi función/ aunque dirás/ pues <~pus> está un poco fuera/ de lo que [yo/ ¿sí?/ pero **pues...**]

(v) *Alargamiento*: Los casos de alargamiento se anotan inmediatamente después del sonido alargado, sea vocal o consonante, pero no dentro del texto ordinario, sino como parte de las marcas de pronunciación aproximada, tal como ocurre en el ejemplo siguiente con *no*.

(4) Entrevista 16, turno 426.

426 I: [y/ ¿tú crees?/ ¡no <~no:~>! E]/ no creas que ese muchacho se las masca así tan fácilmente/ es abusadísimo/ luego luego empieza a ver por qué/ se asoma/ “oye papá/ ¿y por qué los de enfrente tienen luz y nosotros no?”

Se distinguen tres niveles de pausa: breve, media y larga.

(vi) *Breve: /*; este primer nivel corresponde apenas a un inciso; en ocasiones, puede consistir en un valle melódico, más que en una pausa propiamente dicha. Suele tener valor fónico.

(vii) *Media: //*; la pausa media corresponde sin duda a un silencio acústico, aunque su duración es variable y la interpretación de una pausa media es relativa, por diferencia con la breve y la larga. Suele tener valor sintáctico.

(viii) *Larga: ///*; la pausa larga lo es notoriamente más con respecto a la media. Suele tener valor discursivo.

Al final de turno no se marca pausa.

Se pueden emplear también las observaciones de (silencio) y de (lapso) para pausas progresiva y notoriamente más largas. No se han medido acústicamente las pausas. Propondríamos, como base de discusión, partir de la siguiente clasificación:

(5)

- a. Pausas breves: menos de 400 ms.
- b. Pausas medias: entre 401 y 800 ms.
- c. Pausas largas: entre 801 y 1200 ms.
- d. Silencios discursivos: entre 1201 y 2000 ms.
- e. Lapsos: más de 2001 ms.

e) Ruidos.

Los ruidos con significado comunicativo se anotan de la misma forma que los comentarios del analista (*infra*). En el repertorio de los llamados ruidos aparecen diferentes marcas, como (risa), (clic), (ruido), (siseo), etc. En alguna ocasión se ha introducido alguna etiqueta referida a circunstancias que no necesariamente tienen un valor comunicativo, pero que contribuye a la mejor inteligencia del desarrollo de la entrevista; es lo que ocurre con formas como (carraspeo) o (tos). Las risas en particular tienen una gran riqueza de funciones. Véase por ejemplo el siguiente fragmento:

(6) Entrevista 21, turno 782 y ss.

782 I: está guapísimo/ aparte es como osito de peluche/ entre gordito y/ o sea está sabrosito el muchacho/ [(risa) qué pena (interrupción) ¿verdad? <~verdá>/ no no/ mejor bórralo]

783 P: [(risa)]

784 E: [(risa)]/ no importa

785 P: ahí le ponemos “pi <~pi::>”/ [(risa)]

786 E: [no sí]/ subí todo el volumen para que se oiga bien

787 I: no/ fijate que sí/ yo tengo ahí una desviación hacia los ositos de peluche/ ¿eh? (risa)

788 E: o cualquier cosa que se les parezca

789 I: [este <~este:> mm <~mm:>]

790 E: [(risa)]

791 I: no trabajo una línea precisa/ soy diversificada/ pero los eh/ me parece guapísimo ese muchacho/ aunque es/ es muy amarillista de [repente]

792 E: [<...>]

793 I: pero no importa

La broma surgida durante la entrevista, a propósito de un conocido periodista televisivo, provoca risas de complicidad entre los otros participantes; una de las funciones de la risa es precisamente la de llevar a cabo tareas de realimentación que faciliten la prosecución del turno de habla del interlocutor. No es raro, en ese sentido, encontrar la marca de (risa) precisamente en el entrevistador.

#### f) Cuestiones gráficas y fonéticas.

(i) *Ortografía ordinaria*. Aunque la decisión general ha sido emplear ortografía ordinaria, ha sido necesario tomar no pocas decisiones de detalle, considerando que, por tratarse de formas procedentes de la lengua hablada, muchas de ellas no tienen un asiento gráfico claro, o bien su representación gráfica es todavía inestable. Para tomar decisiones ortográficas, se ha seguido la siguiente ruta en los casos de duda. En primer lugar, intentar documentar la palabra en el *Diccionario del español usual en México* (<http://intranet.colmex.mx/dem/>), en el *Diccionario de la lengua española* o en el *Diccionario panhispánico de dudas* (Real Academia Española, <http://www.rae.es/rae.html>). En segundo lugar, cuando no se ha podido localizar la voz en una de estas fuentes, se ha acudido a los corpus CREA y CORDE (también en la página de la RAE). Por fin, cuando ninguna de estas vías ha resultado fructífera, se ha rastreado la forma en buscadores de

Internet, por lo general acotando el entorno en páginas producidas en México, y seleccionando por lo regular la solución gráfica que ha resultado más frecuente.

A continuación se comentan algunos casos ilustrativos de las decisiones adoptadas, sin ninguna pretensión de exhaustividad.

Ciertas voces ligadas a la cultura urbana contemporánea no parecen tener todavía adaptaciones estables al español, y circulan como extranjerismos (básicamente anglicismos) crudos. Un par de ejemplos en ese sentido son *rave* y *hardcore*:

(7) Entrevista 2, turno 235 y ss.

235 E: ¿a ti te late todo esto del/ del **rave** <~reib>?/

236 I: pues más/ sí/ sí sí sí/ la música

237 E: mh/ ¿es/ la música que oyes?// ¿o qué otros tipos de música oyes?

238 I: eh/ pues sí escucho/ o sea/ escucho/ este/ música/ electrónica

239 E: mh

240 I: y también me gusta/ el/ **hardcore** <~járcor> y el

241 E: órale

242 I: y el punk// **hardcore** <~járcor>

La palabra *rave* hace referencia a una *rave party*, fiesta nocturna y prolongada de música y baile para los aficionados a la cultura electrónica. En cuanto a *hardcore*, o *hardcore punk*, es un subgénero musical perteneciente al *punk*. Más adaptadas, en cambio, son *punqueto* y *darqueto*, referidos a personas pertenecientes a algunas de las llamadas culturas urbanas.

(8) Entrevista 2, turno 389 y ss.

389 E: oye y/ ¿cómo es/ cómo es la gente que viene aquí?/ o sea uno// pues <~pus> está allá adentro o afuera como que/ a veces// parece que viene como puro **punqueto** o puro// no sé **darqueto**/ ¿no?

390 I: [ajá]

391 E: [pero] en general tú que/ o sea sí estás bien clavado en esto/ ¿qué tipo de gente viene?



Se ha procurado adoptar en general la ortografía académica con los préstamos, aunque no necesariamente es la más habitual en ciertos casos; es lo que ocurre con *zum*, quizá menos común que *zoom*:

(9) Entrevista 36, turno 109

109 I: pues no se trata del ver el panorama/ sino de hacer un acercamiento/ ¿no?/ un zum para// para concentrarse en unos cuantos/ poetas/ o narradores

Ciertas palabras habituales en la lengua hablada tienen ortografía dudosa o variable. Una de las más comunes es *okee*, que aparece innumerables veces en nuestros materiales y para la cual seguimos la forma registrada en el DEM. La voz *quesque* (en entrevistas 7 y 13, por ejemplo) se documenta de esa forma en el CREA. El DPD autoriza las grafías *zonzo* y *sonso*, pero hemos preferido la segunda. Para *folklore*, siguiendo al DRAE y al DPD, regularizamos como *folclore* y *folclórico*. La forma *jacuzzi* aparece de esa manera en el DRAE, y otro tanto ocurre con *spray*.

Mención aparte merecen las siglas. Cuando se pronuncian en una sola palabra las tratamos como un nombre propio: *Unam*. Cuando se pronuncian silabeando separamos cada una de las letras: *V I H*, *C U*, y se acompañan de una marca de pronunciación aproximada: <~ve i hache>, <~ce u>:

(10) Entrevista 23, turno 28

28 I: sí/ co-/ yo estudié en el **C C H** <~ce ce hache>/ entonces/ ahí como que/ era pura lectura// era pura lectura/ y realmente// sí/ sí me/ se me hizo el hábito en ese entonces// pero actualmente ya no/// ya no es// tan fácil por la// el ritmo de vida que llevo/ [quisiera]

(ii) *Pronunciación*. Se adjuntan marcas de pronunciación aproximada cuando es necesario. Entre las más abundantes están las que se refieren a *pues* <~pos> y a *entonces* <~tons>, pero se han incluido muchas otras, entre las que se encuentran los alargamientos, cambios de acentuación, epéntesis, cambios de timbre, etcétera.

g) Observaciones.

Se incluyen tres tipos principales de observaciones: comentarios del analista, fragmentos ininteligibles y fragmentos de transcripción insegura.

(i) *Comentarios del analista*. Los comentarios del analista, que funcionan como aclaraciones de diferente índole, se han anotado entre paréntesis. Algunos ejemplos de estos comentarios son: (interrupción) —de la grabación—; (risas de todos); (vacilación); (sic) —cuando lo transcrito pudiera causar extrañeza o parecer un error (ejemplo 11); (silencio) —para intervalos de 1.2 a 2 segundos—; (lapso) —para intervalos de más de 2 segundos—, etc. En general, se ha procurado reducir al mínimo el número de comentarios, tanto en lo que se refiere a su inclusión en las transcripciones como a la variedad de etiquetas consideradas. En ocasiones, pueden llegar a convertirse en acotaciones que aclaran algún aspecto que termina teniendo repercusiones en la entrevista, como ocurre con el comentario (sonido de celular) en el ejemplo (12).

(11) Entrevista 19, turnos 432-433.

432 I: necesita de un suelo de donde estar este proviéndose (**sic**) de ma- de minerales// y precisamente por esa// ahora <~ora> sí que esa alimentación que toma de los/ y los minerales que toma del suelo/ ahí es cuando hay que que estarle cambiando de tierra [de repente]

433 E: [claro]

En el ejemplo anterior, (sic) sirve para constatar que en efecto *proviéndose* fue lo expresado por I; recuérdese el criterio de respetar la variación morfológica y léxica presente en los datos.

(12) Entrevista 6, turno 252 y ss.

252 I: entonces <~tos>/ que está/ tiene un subsidio/ que nos dio el gobierno/ que nada <~na> más pagas el veinte por ciento/ pero todavía <~toavía> cincuenta mil pesos/ ahorita <~orita> para <~pa> nosotros/ ya estamos muy apretados porque no tenemos dinero (**sonido de celular**)

253 E: ¿eso [lo tienes]

254 I: [¿sí?]

255 E: que pagar tú?/ <...>

256 I: a ver permítame/ me están hablando/// bueno

257 P: ¿qué pasó?

258 I: ¿qué onda cómo estas?

Como puede observarse, en el ejemplo (12) es importante anotar la observación (sonido de celular), pues la irrupción del aparato telefónico provoca la interrupción de la entrevista ante la prioridad que se concede a la llamada, la aparición de un participante no presente e inesperado, P, a quien se alcanza a oír en la grabación, y las consecuentes intervenciones de I en respuesta a la llamada.

En ocasiones, la interrupción se produce por alguna falla técnica momentánea, tal como parece ocurrir en el ejemplo anotado a continuación:

(13) Entrevista 3, turno 137 y ss.

137 E: claro/ y y ¿desde chiquito sabes dibujar?

138 I: ah/ yo cometí un error

139 E: ¿por qué? (risa)

140 I: a mí desde chiquito me gustó dibujar (**interrupción de la grabación**) o...

141 E: perdón/ ya/ ajá

142 I: profundizar

143 E: mh

144 I: eh/ lo que hice fue/ me metí en otras cosas/ o sea <~sea>/ no sé/ son errores de// de la juventud/ [o sea]

145 E: [claro]

Más interesante es la circunstancia expuesta en el ejemplo (14), también marcada con una (interrupción de la grabación). En la entrevista 14 el tema principal es el cultivo de bonsáis, que es precisamente la actividad desarrollada por I. En cierto momento (turno 133), E insta al informante a que le enseñe las plantas, lo que provoca una interrupción.

(14) Entrevista 14: turno 132 y ss.

132 I: bueno/ eso es una/ muy buena pregunta// eh/ cada persona// tiene/ un carácter diferente/ tiene una personalidad diferente/ que es lo que la diferencia (sic) y la define// entre las demás/// entonces/// yo soy I/ y todo el mundo// que me conoce// sabe/ más o menos cómo soy/ porque soy igual con todos// entonces/ mi carácter// mi// mi emotividad// la/ la vuelco// sobre mis árboles// y todos mis árboles/ cada uno// tienen mi estilo// que si tú conoces// mi trabajo// lo vas a poder

identificar/ en cualquier exposición/ vas a poder decir/ “ese es de I”

133 E: y qué le parece si me enseña/ desconectamos aquí/ y me enseña/ su... **(interrupción de la grabación)**

134 I: es un pino/ no es un pino es un junco/ eh// excesivamente torcido// este/ estos árboles de troncos muy retorcidos/ son árboles que/ se supone que en la naturaleza/ están en lugares muy agrestes// muy agresivos/ eh/ que el clima es muy extremo/ que hay condiciones de un frío/ extremo/ a condiciones de un calor extremo/// y/ estas variaciones// en la temperatura/ en el clima// etcétera/ es lo que produce los retorcimientos en los árboles/ de forma natural/ entonces nosotros reproducimos esos retorcimientos/ y en algunos casos elegimos árboles que ya tienen retorcimientos/ en forma natural// lo único que hacemos es// exagerar un poco esos retorcimientos/ ¿no?/ y eso lo hacemos mediante el alambre// el alambre/ nos sirve/ para corregir/ y para guiar/ algunas de las/ de las ramas/ por ejemplo// mira aquí en/ en este árbol

135 E: sí

(ii) *Fragmento ininteligible*. Los casos en que resulta ininteligible lo expresado en un fragmento de la entrevista se han marcado por medio de ángulos con puntos suspensivos intermedios, <...>.

(iii) *Texto inseguro*. Por fin, cuando parece entenderse lo dicho, pero no existe plena seguridad acerca de la transcripción, o bien ha habido continuas discrepancias entre los diferentes revisores, el fragmento queda anotado entre ángulos, como en <voy a estar allá>.

#### h) Notaciones especiales.

Las notaciones especiales incluyen tres tipos de marcas.

(i) *Marginalia*. Aparecen en las transcripciones diferentes marcas verbales, como *ah, eh, oh, uh, ajá, mm, mh, um*, entre otras. La función de la mayor parte de ellas es servir como turnos de realimentación a lo expresado por el interlocutor, aunque en ocasiones sirven también como recursos para mantener el turno propio. Se ha procurado buscar un término medio entre resolver las vacilaciones gráficas que tales verbalizaciones suelen acarrear y representar con cierta fidelidad el tipo de material sonoro producido por el hablante.

(ii) *Sonidos inspirados*. Se emplea una doble *hh* para hacer referencia a este tipo de expresiones verbales. Por lo regular aportan un matiz de asombro.

(15) Entrevista 20, turno 214 y ss.

214 I: eh <~eh:>/ fijate <~fiate> un día me encontré a una señora// increíblemente// ¿cuánto crees que le daba su marido para que hiciera el gasto?

215 E: ¿a la semana?

216 I: ¡no/ en el día!

217 E: pues <~pus> no sé unos/ treinta pesos/ cuarenta pesos

218 I: veinte pesos

219 E: **ihh**

220 I: ¡veinte pesos!/ ¡y eran siete!/ párale de contar

En (15), E manifiesta en el turno 219 su asombro ante el escaso dinero recibido por la señora de quien se está hablando a través de un sonido inspirado de timbre vocálico relativamente prolongado.

(iii) *Citas y estilo directo*. Para el estilo directo y las citas se ha adoptado la convención ortográfica normal de abrir y cerrar el enunciado referido por medio de comillas dobles: “ ”. Este tipo de materiales parece de gran productividad para el estudio de diferentes aspectos, en especial los vinculados con secciones dialogadas de fragmentos narrativos, descriptivos o argumentativos, como ocurre en el ejemplo siguiente:

(16) Entrevista 24, turnos 215-216.

215 I: o sea dices/ **“gracias”**// yo sí/ a la mejor estoy mal/ pero yo sí/ llego a un espectáculo de danza y digo **“yo puedo hacer eso/ no estás bailando/ discúlpame/ pero no estás bailando”**/ yo/ que tengo/ medianamente entrenamiento en danza contemporánea este digo/ **“ay no/ si yo puedo hacer eso/ gracias/ no me digas que bailas”**/ pues <~pus>

216 E: claro

Muchos de los casos pertinentes, como el anterior, van introducidos por un *verba dicendi*. La viveza del texto viene propiciada por el movimiento constante entre el presente

del diálogo sostenido en la entrevista, y el presente referido y ejecutado al dar voz al momento pasado.

*i) Otras cuestiones.*

Nos referimos aquí en especial a problemas vinculados con la identificación apropiada de los participantes en la interacción grabada, así como de otras personas, instituciones o circunstancias mencionadas durante el intercambio. El conflicto se establece ante la necesidad de establecer con claridad quién está hablando y a quién o a qué se está refiriendo, pero sin que tal propósito comprometa el mantenimiento del anonimato.

(i) *Nombres propios.* El criterio general aplicado es que los nombres de persona se reemplazan por una letra mayúscula. El informante queda identificado por medio de una *I*, y el entrevistador por una *E*<sup>31</sup>; otros participantes esporádicos se marcan con *P* el primero, *Q* el segundo, *R* el tercero y así sucesivamente. Tales iniciales se aplican tanto al comienzo del turno —para identificar quién habla— como dentro de los turnos, reemplazando los nombres de los interlocutores cuando se llegan a mencionar.

Si en la entrevista se alude a terceras personas, el nombre se reemplaza igualmente por la inicial, de modo que un *Felipe* o una *Ana* serán, normalmente<sup>32</sup>, *F* y *A*, como en el ejemplo siguiente, en que se menciona a *H* y a *J*:

(17) Entrevista 1, turno 43.

43 I: y este// ya llegué y// y pues <~pus> toqué y te digo que salió y me dijo que no/ que era en otra puerta/ y ya en eso llegó el **H** en la camioneta y ya// me presenté con él/ le dije que venía d-/ de fulanita de tal/ que me había mandado/ y no sabía quién era/ y ya después se acordó y me dijo/ “ah pues <~pus> pásale”// y ya// entré y me dijo// me estuvo platicando/ me preguntó que dónde había estudiado/ que qué había hecho/ le dije que no// que nunca había trabajado/ que era mi primer trabajo// bueno/ mi primer trabajo en relación de la carrera// y ya/ me dijo que estaba muy bien/ que// me iba a poner supuestamente como a prueba// y le dije que sí/ que <no había ningún> problema// y ya/ me/ me presentó con un compañero que se llamaba **J**/ que estaba encargado de una obra ahí en// en/ (carraspeo) en Plutarco/ un Bital/ que estaban remodelando/ y ya/ me fui con él el primer día

<sup>31</sup> Se marca por medio de una «F» algún caso aislado en que ha habido un entrevistador secundario.

<sup>32</sup> Si se sospecha que incluso la inicial de un nombre podría propiciar la identificación, ésta se cambia por cualquier otra letra.

Ha habido ocasiones, con todo, en que ha resultado imposible encubrir la referencia a un nombre, como ocurre en (18)<sup>33</sup>.

(18) Entrevista 19, turno 538 y ss.

538 I: en el papelito la hojita/ le dejas que/ a que toque el agua y ahí mismo solita/ [ya cuando la ves ya]

539 E: [pero **las violetas** sí son bien] sangronas/ ¿verdad?

540 I: sí/ son/ bien especiales

541 E: para darse

542 I: yo/ ahora <~ora> que así **queriendo hacerle honor a mi nombre**/ que no lo he pensado en ese sentido pero/ alguien me dijo/ “ay sí/ **tú le quieres hacer honor a tu nombre**”/ (risa) las he querido cultivar en casa y no he podido

Los nombres propios de personajes públicos, sin embargo, se dejan íntegros, tal como hayan sido citados por los informantes, sea un *Octavio Paz* o un *Fox*. Así, *Diego Rivera* y *Frida Kahlo* aparecen mencionados durante la entrevista 3; el grupo *Niche* y el cantante *Homero* surgen en la entrevista 8; *Alberto Vázquez* en la 20, etc. Anotamos a continuación un ejemplo de mención de un personaje público:

(19) Entrevista 8, turno 904 y ss.

904 I: hace creo una semana o dos semanas

905 E: mh

906 I: vino **Dolores Padierna**/ aquí al barrio

907 E: mh

908 I: y según esto/ esta mujer/ ya ves que es delegada ahí en Cuauhtémoc// según ella/ sus planes son// quitar el/ el/ el campo/ ¿sí sabes que hay un campo de fútbol?

Los nombres propios de lugares e instituciones se conservan también, en principio, salvo que comprometan de alguna forma el principio de anonimato, sobre todo cuando el carácter confidencial de la entrevista sea especialmente delicado (por ejemplo al formular

---

<sup>33</sup> Situaciones como las presentadas en el ejemplo son realmente excepcionales; confiamos en que no se consideren infracciones graves al requisito de anonimato.

críticas sobre aspectos relacionados con el trabajo del informante). Por supuesto, se conservan cuando no parecen entrañar mayor compromiso. Así, por citar uno entre muchos ejemplos, en la entrevista 8 hay numerosos ejemplos de nombres de calles: *Circunvalación*, *Ferrocarril de Cintura*, *Zapateros*... Los mismos criterios se han aplicado a las siglas, en general conservadas salvo cuando dejan en entredicho el anonimato, tal como ocurre en (20):

(20) Entrevista 36, turno 110 y ss.

110 E: esa idea de [las parejas]

111 I: [esa] idea de las parejas/ ¿no?/ y bueno con eso podríamos entrar en/ e-/ ¡ya que hablamos del plan de estudios!/ pues hablar de la U en donde estamos/ ¿no?

112 E: sí [sí]

113 I: [y <~y::>] este/ que era/ creo que uno de tus/ ¡intereses!/ eh mm/ yo tengo ¡todos los años de la U!/ aquí/ y <~y::>/ pues tengo una// parte <de> la historia de mi vida// eh mm/// pues aquí en este <~este:>/ ¡en este recinto!/ ¿no?

(ii) *Otros problemas vinculados con la confidencialidad.* No siempre es fácil establecer qué materiales procedentes de entrevistas lingüísticas pueden o no publicarse, aun cuando el propósito de su análisis sea el trabajo científico. Considérese, además, que los medios digitales de transmisión de la información hacen muy difícil controlar empleos imprevistos de los datos. El usuario y el analista tienen también que considerar sus propias responsabilidades. En cuanto a las transcripciones mismas, la aplicación de criterios completamente rigurosos puede ser un sinsentido —por ejemplo, eliminar toda referencia a los lugares que anclan la historia de vida de un informante—, pero en el otro extremo se corre el riesgo de tomar decisiones subjetivas o casuísticas. En ocasiones se ha suprimido por completo algún breve fragmento que nos parecía podía entrar en conflicto con el principio de anonimato.

Los ejemplos que siguen muestran algunos de los problemas a los que nos referimos.

(21) Entrevista 25, turno 581 y ss.

581 I: entonces digo/ no creo que tan fácil los vaya a dejar perder/ ¿verdad? (silencio)



582 E: mh

583 I: ahora <~ora> yo este/ no sé con esto aquí con/ mi suegra/ bueno/ **mejor no platicamos de eso/ porque/ si eso se va a quedar un día se puede/ puede caer en**

584 E: mm/ **esto es/ información totalmente confidencial/** eh

585 I: [mm]

586 E: **[no] sale de/ de aquí porque/ es material que/ que yo uso/ y/ y que voy a/ vamos a usar de otra manera/ y aparte es gente que/ totalmente ajena aquí**

587 I: mm

588 E: incluso puede/ yo le puedo decir que/ pues es un señor y ya (silencio) pero bueno/ como [quiera]

589 I: [sí]/ no es que/// **ya mejor no platicar de eso**

590 E: no no no/ [claro/ claro]

591 I: [mm]

592 E: [mejor así]

593 I: [sí]/ pero yo por todo lo demás/ estoy contento aquí

Previo al fragmento de la entrevista 25 anotado como ejemplo (21), el informante, I, se ha venido refiriendo a una serie de contratiempos familiares relacionados con la propiedad de un terreno. Al llegar a la altura del turno 583, I cae en la cuenta de que la grabación y, en consecuencia, el registro de lo que está diciendo podría ser motivo de un futuro problema. Entre 584 y 588, E intenta tranquilizar a I, recordando el papel ajeno de los usuarios futuros de la grabación; ciertamente, más que referirse al carácter *confidencial* de la entrevista, E debió referirse al compromiso de anonimato. Los razonamientos de E, sin embargo, no parecen convencer a I, que en 589 señala que *ya mejor no platicar de eso*. De forma paradójica, en los turnos posteriores —no copiados arriba, pero incluidos en la transcripción— el informante sí sigue hablando del tema.

El ejemplo siguiente es un caso más extremo de la situación anterior, pues en (22) el informante, un médico, no se limita a una petición de abandonar un tema, sino que ante la irrupción de un compañero (en el turno 50) pide explícitamente que se apague la grabadora (t. 52); la grabación se reanuda desde el t. 54, y E retoma la entrevista en el t. 57.

(22) Entrevista 27, turno 45 y ss.

45 I: yo soy ortopedista

- 46 E: ah okey
- 47 I: siempre he dado/ siempre he dado// ortopedia/ en// en mil novecientos// eh/ di/ eh  
introducción a la cirugía
- 48 E: ah sí verdad/ es que vine el lunes y estaba operando
- 49 I: ajá
- 50 P: **perdón// ¿nos vamos a repartir?/ eh/ ¿los pacientes?// ¿cuál quiere?**
- 51 E: **(risa)**
- 52 I: **apáguelo**
- 53 E: **sí (interrupción)**
- 54 I: ¿de cuál?
- 55 E: los de antes/ ya son como/ disquitos
- 56 I: ah sí
- 57 E: entonces <~entós> me estaba diciendo que// que/ lo que/ lo que importa es dejar huella/  
¿no?
- 58 I: dejar/ huella positiva

Más común en general ha sido el tener que insistir en el carácter reservado de los materiales, a veces introduciéndose una leve broma por parte del entrevistador, como ocurre en (23), para vencer las reticencias del informante, quizá ante la incertidumbre por los temas que se van a abordar durante la entrevista.

(23) Entrevista 32, turno 155 y ss.

- 155 I: sí/ toda mi vida/ **quieres que [te platique toda mi vida]**
- 156 E: [sí/ está bien/ me] parece perfecto// ¿no?
- 157 I: no voy a terminar nunca
- 158 E: no le hace <...>/ pues <~pus> está bien/ así me entero
- 159 I: así te enteras
- 160 E: (risa)
- 161 I: **y luego/ confidencial**
- 162 E: no/ sí/ pero por eso es-/ **no te preocupes/ porque queda estrictamente confidencial/ ¿no?/**  
es material [que no]
- 163 I: [no sí este]
- 164 E: **no te preocupes/ no se los voy a poner a ninguno de tus grupos (risa)**

165 I: a ninguno de mis alumnos

166 E: sí/ a los niños

Por fin, el último ejemplo anotado en esta sección refiere una curiosa circunstancia —se está hablando del tarot y de numerología— que ha impedido ocultar el nombre del informante (*cf.* el turno 513). El pasaje se podía haber suprimido, pero nos ha parecido una licencia leve no hacerlo.

(24) Entrevista 30, turno 508 y ss.

508 E: ¿cómo saca eso?/ ¿por el número?

509 I: por numerología/ **I**/ que/ es decir/ hay una tablita que tú pones del uno al// [nueve/ me parece]

510 E: [nueve/ ¿no?/ sí]

511 I: uno/ dos/ tres/ cuatro/ cinco/ seis/ siete/ ocho/ nueve// y/ [cada letra]

512 E: [a cada letra]

513 I: no sé cómo es la correspondencia/ no recuerdo esto lo leí hace mucho ya// y// hay una correspondencia/ entonces/ tú vas sacando/ por ejemplo <...> **la jota tiene un número/ la o/ la ese y la e**

514 E: mh

515 I: te da cuatro números/ que después/ tienes que/ sumarlos// y te van a arrojar un número

#### *Tipos de archivos incluidos en el CD*

El CD que acompaña a estas páginas incluye cuatro tipos de archivos:

- a) Las transcripciones en formato de procesador de textos; aparecen en 36 archivos independientes.
- b) Las mismas transcripciones, volcadas ahora en *pdf*, también en 36 archivos independientes. Este mismo material se encuentra en la página <http://lef.colmex.mx>.
- c) Las transcripciones en formato texto, también en 36 archivos. En esta versión se ha suprimido el encabezado de cada entrevista y se ha añadido una etiqueta especial antes de cada turno de habla, la cual permite identificar su procedencia, tal como se

muestra en el ejemplo (25), en el que la etiqueta incluye entre ángulos la identificación de la entrevista:

(25) Etiquetas que establecen la entrevista a que pertenece cada turno de habla (Entrevista 28, turnos 878-888).

<ME-245-33H-05> 878 E: y esas obras que/ que escribiste/ ¿cómo te llegó la inspiración?

<ME-245-33H-05> 879 I: pues mira/ la primera que escribí

<ME-245-33H-05> 880 E: mh

<ME-245-33H-05> 881 I: fue un monólogo

<ME-245-33H-05> 882 E: mh

<ME-245-33H-05> 883 I: que se llama Herodes

<ME-245-33H-05> 884 E: ¡ah claro!/ divertidísimo

<ME-245-33H-05> 885 I: ah/ pues <~pus> tú la viste

<ME-245-33H-05> 886 E: sí/ yo la vi

<ME-245-33H-05> 887 P: ¿sí?

<ME-245-33H-05> 888 E: sí

d) Por fin, se han reunido las 36 transcripciones en formato texto, cada una con sus respectivas etiquetas identificadoras, según el procedimiento que se acaba de describir, en un solo archivo, formado por cerca de 450 000 palabras, distribuidas en algo más de 25 000 turnos de habla. El interés de este material es que permite llevar a cabo búsquedas conjuntas en muy poco tiempo, sea a través de las opciones proporcionadas desde cualquier procesador de textos, sea a través de un buscador de concordancias. En cuanto a estos últimos, existen varios disponibles en el mercado, y también otros gratuitos en la red. En principio, este archivo debería ser legible por cualquier buscador, dado su formato. Nosotros hemos empleado en algunos trabajos el software *Simple Concordance Program (SPC)*, diseñado por Alan Reed, el cual puede descargarse gratuitamente de la página <http://www.textworld.com/scp/> — incluye versiones para PC y para Mac—; resultados semejantes, no obstante, deberían poder obtenerse por medio de otros programas equivalentes. A continuación se anotan algunos ejemplos de las búsquedas que es posible llevar a cabo:

(26) Ejemplos extraídos de la lista proporcionada por *SPC* para *entonces* por medio de la opción *line* (en el subcorpus de PRESEEA-México, nivel alto, resulta haber 2480 ejemplos de *entonces*, distribuidos en 1873 líneas o turnos).

### **entonces**

7388 <ME-248-31M-05> 10 I: y el área específica donde quiero estar/ es donde te vas a desen-/ desenvolver con varones// donde hay situaciones de asesinatos/ de narcotráfico y cuestiones así// **entonces** <~tos> como que sí es un poquito difícil// porque <~porque:>/ aunque mi pareja no me limita/ no me margina/ no me pone trabas// me apoya y me lo ha demostrado/ sí me preocupa un poquito/ no tanto por él// sino por la presión que ejercen sobre él// de que <~que:>/ es tradicional/ ¿no?/ [o sea]

7390 <ME-248-31M-05> 12 I: económicamente/ yo sé que están muy bien/ me pueden brindar todo// necesidad de trabajar no tengo// pero/ esa como que no/ no es mi estilo de vida/ no quiero eso para mí/ **entonces**/ quiero desarrollarme// quiero aportar como lo he hecho// en mis posibilidades// y/ y me gusta brindar// a mi pareja/ a mi familia/ y si al rato me va más// o sea/ más puedo dar/ pero no quiero depender ni emocional ni económica ni/ moral ni nada// o sea/ quiero decir/ como que mi idea es/ “es mi pareja/ la amo// pero no soy de su propiedad ni me pertenece”// o sea/ que sienta cierta parte de independencia/ y cierta libertad// que no se sienta así como que obligado a estar conmigo

7392 <ME-248-31M-05> 14 I: y que si algo quiere hacer eh/ necesita nada más decírmelo/ yo nunca voy a ser un obstáculo// y así **entonces** igual y conmigo/ así como que quiero que sea esa misma situación/ o sea estar porque/ me nace estar/ no me siento obligada/ no me siento presionada/ y todo lo que hago hacia esa persona/ hacia mi familia/ hacia mi pareja/ hacia mis amistades/ es porque me nace hacerlo// no/ no lo hago así como que// por complacer/ por quedar bien/ o por ganármelos/ no/ no// pero sí me preocupa esa parte un poco porque/ con él no tengo problema// yo de hecho hasta le he comentado/ “mira/ ¿sabes qué? amorcito/ yo quiero trabajar en esto/// y y voy a luchar hasta llegar a donde esté/ y yo sé que a lo mejor voy a tener/ tener que enfrentarme a miles de cosas// y sin embargo/ no me preocupa/ porque/ yo sé que si tú de verdad me amas/ me vas a apoyar y no me vas a condicionar// y ahí vas a estar// y eso a mí es lo que me va a hacer sentir bien/ y conmigo no vas a tener ningún problema// porque no/ o sea/ no vas a tener que estar vigilando/ porque ya voy a andar/ de amante/ o acostándome con fulano/ [o con zutano”]

7394 <ME-248-31M-05> 16 I: le digo/ no// **entonces**/ pero sí te digo/ sí me preocupa un poquito la parte eh/ familiar/ no porque es muy tradicional// como que es muy remarcado de donde ellos vienen/ ¿no?/ es un lugar muy machista// económicamente/ están muy bien todos sus familiares/ por

parte de su papá y todo// son buenos proveedores/ pero sí como que tienen muy remarcado eso// de que cómo la mujer va a trabajar/ [cómo]

7400 <ME-248-31M-05> 22 I: ya sabes/ **entonces**// en/ esa es la parte que me p-/ que me preocupa un poquito/ porque sí ejercen cierta// presión hacia él// él no lo toma en cuenta/ porque <...> dice “yo soy l- el que decido

7402 <ME-248-31M-05> 24 I: [<porque>] yo soy el que estoy contigo/ convivo contigo/ y si tenemos hijos o no tenemos hijos/ y si yo te apoyo o no te apoyo/ es mi problema// ellos no van a decidir/ y si la solución es irnos pues <~ps>/ o sea// lo vamos a hacer”// ¿sí me entiendes?/ **entonces**/ te digo/ eso a él no le importa/ y a él no le provoca conflicto// pero tampoco yo quisiera que/ que él estuviera en medio/ porque sí es algo muy difícil/ ellos son muy unidos

7404 <ME-248-31M-05> 26 I: o sea/ muy a pesar de todo/ son muy unidos// **entonces** <~tons> yo no quería se-/ yo no querría ser así como que el motivo de/ de que hubiera ese// ¿no?/ esa separación// **entonces**/ o sea/ yo sé que nadie está exento de/ de algún conflicto// y muy a pesar de que estoy bien/ de que estoy tranquila/ como que/ no descarto la parte/ ¿no?/ [de que]

7406 <ME-248-31M-05> 28 I: si llega a haber algún problema/ “¿sabes qué?/ te quiero mucho esto lo otro pero/ mira/ mejor prefiero que// si tenemos un hijo o sea/ eres su papá/ vas a poder llevarlo/ traerlo/ comprarle// pero no quiero ni que te provoquen a ti problemas// ni <~ni:>/ ni quiero yo disgustarme// porque yo tengo mi carácter y/ y tú sabes que no me sé dejar// porque no me gusta que me presionen/ que me obliguen/ que// que me impongan// sé ceder/ pero sí yo sé que/ que no es lo correcto/ porque/ el tiempo cambia/ tienes que cambiar tu forma de pensar/ tus expectativas/ laborales/ familiares/ económicas”// **entonces** <~tons> le/ le comento a él/ “yo no quiero que me limiten// y si ese va a ser un problema o sea/ que tu familia// te/ haga sentir mal/ que te esté molestando/ que te esté molestando

7408 <ME-248-31M-05> 30 I: **entonces** <~entóns>/ ¿sabes qué?/ yo no quiero que tú estés en medio”/ **entonces** <~entóns> he llegado a pensar eso// yo sé que no es lo correcto// pero/ hasta cierto punto/ sería lo más sano

Obsérvese cómo en (26) cada línea del formato texto corresponde a un turno de habla de los que han sido marcados en la transcripción de las encuestas; así, si tomamos por ejemplo el segundo dato de la lista que se acaba de incluir,

(27)

7390 <ME-248-31M-05> 12 I: económicamente/ yo sé que están muy bien/ me pueden brindar todo// necesidad de trabajar no tengo// pero/ esa como que no/ no es mi estilo de vida/ no quiero eso

para mí/ **entonces**/ quiero desarrollarme// quiero aportar como lo he hecho// en mis posibilidades// y/ y me gusta brindar// a mi pareja/ a mi familia/ y si al rato me va más// o sea/ más puedo dar/ pero no quiero depender ni emocional ni económica ni/ moral ni nada// o sea/ quiero decir/ como que mi idea es/ “es mi pareja/ la amo// pero no soy de su propiedad ni me pertenece”// o sea/ que sienta cierta parte de independencia/ y cierta libertad// que no se sienta así como que obligado a estar conmigo

lo que se está indicando es que el ejemplo corresponde a la línea (o turno) 7390 de las 26319 incluidas en el conjunto de las 36 entrevistas del CSCM-PRESEEA-nivel alto, la cual pertenece a la entrevista <ME-248-31M-05>; de hecho, se trata del turno 12 de esa entrevista, que fue realizado por la informante (I). Por supuesto, si se precisa un contexto todavía más amplio, es posible acudir con estos datos a la entrevista específica (véase el cuadro 1 para la correspondencia entre la etiqueta de la entrevista y su número seriado).

Otra opción posible es obtener lo que podría llamarse un listado rápido con un contexto más breve. El ejemplo (28) podría ser parte de una búsqueda inicial destinada a establecer la productividad de ciertos prefijos en el español de México.

(28) Lista de las palabras que empiezan con *anti* (27 ejemplos en 23 líneas o turnos), en formato *kwic* (*keyword in context*).

**anti**

1470 es/ pues <~ps> es más bien es **antihigiénico**/ ¿no?/<ME-105-31H  
6318 // eran// cuartos/ de esos **antiguos**// muy altos/ algunas casas  
7058 todo]/<ME-181-31M-01> 750 I: [**antigüedades**]/ hace algunos años  
7064 1-31M-01> 756 I: **antigüedades**/ **antigüedades**/ libros viejos// cos  
7064 ]/<ME-181-31M-01> 756 I: **antigüedades**/ **antigüedades**/ libros  
8286 más la/ la pastilla de **anticoncepción** de emergencia/ y  
8286 pastilla de emergencia// el **anticonceptivo** de emergencia”/  
8297 de que/ cada año es/ y un **anticonceptivo**/ y un...”/ o sea/  
8414 este pues <~pus> ya muy **antiguo**/ y entonces cuando el terremoto  
8560 de de todos los sistemas **anticontaminantes** de la Chrysler  
9405 mh/ **antihemofílico**/ el factor **antihemofílico**/<ME-138-32H-01>  
9405 /<ME-138-32H-01> 188 I: mh/ **antihemofílico**/ el factor **antihemofílico**  
10650 / quizá un poco un estilo **anticuado**// esos señores les importa  
10897 / hasta recientemente/ ayer o **antier**/ recibí un correo/ de un  
11976 >// un medicamento hiper- este **antihipertensivo** de los más comunes  
11976 lo que hay en existencia”// **antibióticos**/ a veces hay uno/  
11985 me enfermé/ y me recetaron un **antibiótico** bastante fuerte/ y  
13490 muy grande/ ¿no?// ayer/ **antier**/ me habló/ apenas/ la persona  
14853 ahí en San Ildefonso// en el **antiguo** palacio de// de San Ildefonso

16671 /<ME-140-33H-01> 64 I: de **antiético** es// falsear// es decir  
 16741 los letreros de las películas **antiguas**/<<ME-140-33H-01> 135 E  
 19091 / ya es una historia muy/ muy **antigua**/ y luego la especialidad  
 21323 -02> 664 I: [recordar historia **antigua**]/<ME-220-33M-02> 665 E  
 22520 manejara una tabla periódica/ **anticuada?**"/ que si ya van/ ciento  
 24860 de la universidad es más **antigua** toda- <...>/ eh/ mi papá  
 25148 33M-03> 206 I: fui mamá de las **antiguas**// quedé como en la transición  
 25150 trabajaban/ y entre las mamás **antiguas**/ que eran de mi generación

Aunque en un buscador de concordancias convencional es tarea del analista establecer cuándo se trata en realidad de un prefijo y cuándo no, es posible al menos en algunos de ellos —lo es en *SPC*— establecer filtros que permitan, por ejemplo, la selección de *antiético* (línea 16671), pero el rechazo de formas como *antiguo*, *antigua*, *antiguas*.

Es posible realizar búsquedas análogas a propósito de cualquiera de las observaciones incluidas en la transcripción. (29), por ejemplo, reproduce uno de entre los 2066 casos distribuidos en 1932 turnos en que aparece (risa):

(29) Ejemplo de (risa); entrevista 22, turno 122.

13498 <ME-253-32M-05> 122 I: y que por ejemplo con T/ la persona que me habló/ y la persona con la que conviví varios años/ que vivimos juntas/ porque vivía en mi casa/ cuidaba a mis niñas// porque ella me enseñó otro trato con los niños/ otra forma de educarlos/ eh/// ahí es donde digo bueno es que/ es T/// y y aparte de tener una cultura/ compartir una cultura otomí/ pero es T// y es/ y ahí donde dices/ “es que soy yo”/ así/ que tengo todo esto/ ¿no?/ ahí es donde/ donde más lo aterricé/ ¿no?/ donde más lo comprendí// y sí digo/ bueno/ en términos de ba-/ ahorita <~orita> de explicar/ pues sí digo/ “es otomí”// pero ella siempre T/ ¿no?/ y mis niños T// y la querían hasta como mamá/ entonces/ imagínate por ejemplo mis hijos/ que tuvieron una formación/// de de dos personas/ distintas en cuestión cultural/// y que están nutridos de esto/ ¿no?/ por ejemplo tengo una niña que es la que/ ella se viste con el estilo de colores/ con los gustos/ con todo esto/ y y lo anhelaba tanto/ ella decía por ejemplo/ una tortilla hecha a mano/ decía “una tortilla de T”// porque ella las hacía/ y le daba/ o sea la cargaba acá atrás/ y le daba su/ prueba de su taquito/ ¿no?/ entonces ella una/ una tortilla hecha a mano/ decía “una tortilla hecha por T”/ y decía “¡ah!/ es que quiero una tortilla de T”/ ¿no?/ (**risa**)/ entonces <~entós> imagínate toda esa formación para ella// sus primeros tres años de vida// todavía a ella le gustan colores más vivos// este/ no sé/ es un/ una personalidad/ un híbrido bien chistoso /¿no?/ pero que/ que fue compartida/ ¿no?/ fue fue bien especial (**risa**)/ ella sí como ha sido el/ una experiencia bien especial en cuanto a/ esa formación// ya



al niño ya no le tocó / M tuvo más de mí/ porque la cuidé/ en sus primeros años de vida/ ¿no?/ o sea está más/ es muy muy/ a mi estilo/ digamos/ hasta en gustos/ en todo// pero la otra niña no// es este/ de los gustos de la señora/ ¿no?/ T (**risa**)/ estuvo muy// bueno/ se me hace// una/// pues fue una oportunidad muy padre/ ¿no?

Otra de las posibilidades disponibles para el investigador es la consideración de la frecuencia de las formas léxicas. El ejemplo (30) anota la lista de las palabras documentadas más de 1000 veces en PRESEEA-nivel alto:

(30) Palabras (y frecuencias absolutas) que aparecen más de 1000 veces (por orden descendente de izquierda a derecha y de arriba hacia abajo).

30519 me	15210 que	13619 no
13360 y	11905 de	8744 la
8617 a	7533 sí	6412 en
5853 el	5703 es	4638 pues
4033 lo	3826 o	3799 pero
3702 yo	3688 se	3588 un
3366 ya	3000 como	2959 los
2907 este	2751 por	2718 una
2533 con	2480 entonces	2476 porque
2319 te	2264 qué	2263 muy
2180 sea	2169 más	1984 las
1837 para	1785 bueno	1782 eso
1736 le	1664 así	1510 todo
1482 mi	1405 ahí	1329 era
1312 si	1230 claro	1207 está
1179 también	1141 hay	1138 del
1077 bien	1066 al	1038 ajá

De la misma manera, el ejemplo (31) incluye algunas de las muchas formas que fueron documentadas una sola vez:

(31)

1 simbolismo	1 simbología	1 simón
1 simpatizaba	1 simpáticas	1 simplifica
1 simplotas	1 simpósium	1 simulacro
1 simultáneamente	1 simultáneas	1 sincera
1 sindicalizadas	1 sindicato	1 siniestro
1 sinnúmero	1 sinodal	1 sintaxis
1 sintético	1 sintiéndome	1 sintiéndote
1 sintiéramos	1 siqueiros	1 sirena
1 sirvieran	1 sistematicidad	1 sistematizada

#### ASPECTOS LINGÜÍSTICOS MÁS INTERESANTES<sup>34</sup>

Los aspectos que se pueden trabajar por medio de un corpus de lengua hablada son casi tan variados como los analistas que se acerquen a él. Se anotan a continuación, no obstante, algunos de los temas que podrían abordarse a la vista de los materiales del CSCM; por supuesto, muchos de ellos han sido estudiados ya de una manera u otra en otros trabajos. Considérese, por otra parte, la posibilidad de llevar a cabo trabajos comparativos en tiempo real.

#### *Cuestiones fónicas*

Algunos de los hechos fónicos que merece la pena seguir considerando son los siguientes. En cuanto a la variación vocálica, parece haber bastante margen de estudio en cuanto a problemas clásicos, como el llamado debilitamiento, tanto en términos de su naturaleza acústica precisa como en cuanto a su distribución social y estilística; otro tanto puede decirse del cierre vocálico. En general, la caracterización acústica incluso de las vocales estándar precisa de mucho más trabajo. Los temas relacionados con consonantes son numerosos. Cabe recordar ahora el problema de la asibilación de (r) y (r), la articulación de obstruyentes oclusivas en posición intervocálica, la resolución de las codas silábicas que

---

<sup>34</sup> Esta sección retoma algunos de los comentarios incluidos en Martín Butragueño (2006).

forman parte de los llamados grupos cultos, la dorsalización de la (f), las características de la llamada (s) tensa, la formulación de grupos como (tl), la presencia de sonidos ( $\overline{ts}$ ) y (j) vinculados a léxico de ciertos orígenes; la articulación de ( $\overline{tj}$ ) y de (j) en términos tanto de modo como de lugar articulatorio, etc. Por supuesto, los problemas fónicos y sociolingüísticos se vuelven todavía más interesantes cuando se considera la presencia de personas de varios orígenes puestas en contacto en la ciudad de México.

En términos suprasegmentales, son también copiosas las cuestiones llamativas, empezando por la estructura silábica y acentual y por su proyección léxica en cuerpos de palabras específicas. Quisiéramos destacar en este terreno la importancia del estudio de la entonación para el conocimiento sociolingüístico de la ciudad de México, tanto en sí misma, a través de patrones claramente correlacionados con la distribución social, como la llamada entonación circunfleja, como en lo que toca a sus relaciones con la sintaxis y con el discurso.

#### *Aspectos gramaticales*

También son abundantes las dimensiones morfosintácticas que merecen interés a partir de los materiales derivados de un corpus oral. En general, la variación morfológica ha sido poco trabajada hasta el momento en su proyección social, así que la resolución morfológica de los paradigmas flexivos verbales y nominales en diferentes grupos sociales espera todavía un estudio detallado; otro tanto puede decirse de los procesos de derivación y de composición.

En el terreno puramente sintáctico, parece haber todavía bastante que decir acerca de la resolución social de los paradigmas verbales, en términos de considerar la organización de las relaciones de tiempo, aspecto y modo, no en el estándar sino en grupos sociales específicos. Parece también necesario continuar trabajando alrededor de la dimensión de cambio en las soluciones formales para muy diferentes necesidades discursivas, como puede ser la expresión del pasado, el presente y el futuro. Otro problema clásico es la relación entre indicativo y subjuntivo. Son numerosas también las cuestiones relacionadas con clíticos, con concordancia, con la expresión de relaciones condicionales, concesivas y causales, entre otras. Hay bastante que decir todavía en relación con la

exploración de materiales orales en relación con las hipótesis formuladas alrededor de los llamados canales de gramaticalización como vías de definición de contextos por los que transitaría el desarrollo de variables sintácticas, como se ha apuntado, por ejemplo, para relaciones entre el espacio, el tiempo y la causalidad, o para la expresión del perfecto y el perfectivo. La cuestión, en fin, de la posible distribución social de algún tipo de medida relacionada con la complejidad sintáctica, parece todavía lejos de resolverse.

A medio camino entre la sintaxis y el discurso, el material oral arroja luces sobre el problema de la distribución de la estructura informativa, asociada al orden de constituyentes y plasmada a través de diferentes categorías, como los temas extrapredicativos, los tópicos insertos en la predicación, los diferentes tipos de focos y los apéndices, entre otros materiales; tales estructuras se ven beneficiadas a su vez por medio de la consideración de la proyección melódica.

### *Léxico y discurso*

Nos parece que un corpus oral de esta naturaleza puede ofrecer ciertas perspectivas para entender mejor la estratificación social de la variación léxica, tanto en términos cuantitativos como en términos cualitativos. Más allá de esta utilidad central, existen muchas derivadas, de tipo documental, en la medida en que se ofrecen imágenes del léxico oral actual, y se aportan contextos propicios para el estudio de diferentes cuerpos de palabras, sea que se vinculen por características morfosintácticas, referenciales o, en especial, por su origen. En ese sentido, un corpus de este tipo puede ser una contribución a la comprensión del dinamismo léxico, en especial en términos de difusión y desplazamiento de cierto tipo de materiales; esta perspectiva se alcanza más cabalmente a través del trabajo comparado, sea en tiempo real con otras fuentes documentales del centro de México, sea en el tiempo aparente fruto del contraste con variedades lingüísticas de otras latitudes.

La dimensión discursiva puede abordarse también desde diferentes ángulos a partir de los materiales inscritos en el CSCM. Entre ellos se cuentan los siguientes: los marcadores discursivos como piezas fundamentales de construcción del texto, dotados a su vez de un rico abanico de variación significativa; la expresión de las modalidades enunciativas clásicas (declaración, interrogación, volición y el cruce de éstas con la exclamación, vinculadas todas a su vez con diferentes resoluciones entonativas); la

organización de narraciones, argumentaciones y descripciones; la estructura de los turnos de habla, en lo que toca a los mecanismos de mantenimiento, toma, cesión y realimentación, y el tipo de actos que cabe en las diferentes intervenciones; un sinnúmero de cuestiones vinculados a las repeticiones y las reformulaciones; el problema de las formas de tratamiento y en general la resolución de los principios de cortesía.

Éstos son, en suma, sólo algunos de los aspectos que se podría o que convendría estudiar.

#### ESTUDIOS REALIZADOS O EN CURSO

Una selección de los trabajos realizados en el entorno más cercano al *CSCM* aparece en la página web del Laboratorio de Estudios Fónicos de El Colegio de México: <http://lef.colmex.mx>.

#### LOS MATERIALES DEL VOLUMEN I: HABLANTES DE INSTRUCCIÓN SUPERIOR

En el cuadro 1 se anota la correspondencia entre el número de entrevista y algunos de los rasgos sociales del informante incluido en la grabación.

Cuadro 1. *Entrevistas y rasgos principales de los informantes*

<i>Número</i>	<i>Etiqueta</i>	<i>Nivel</i>	<i>Generación</i>	<i>Género</i>
1	ME-042-31H-99	alto	joven	hombre
2	ME-105-31H-00	alto	joven	hombre
3	ME-154-31H-01	alto	joven	hombre
4	ME-155-31H-01	alto	joven	hombre
5	ME-190-31H-05	alto	joven	hombre
6	ME-197-31H-01	alto	joven	hombre
7	ME-107-31M-00	alto	joven	mujer
8	ME-181-31M-01	alto	joven	mujer
9	ME-248-31M-05	alto	joven	mujer
10	ME-250-31M-05	alto	joven	mujer

11	ME-251-31M-05	alto	joven	mujer
12	ME-252-31M-05	alto	joven	mujer
13	ME-006-32H-97	alto	adulto	hombre
14	ME-056-32H-99	alto	adulto	hombre
15	ME-138-32H-01	alto	adulto	hombre
16	ME-139-32H-01	alto	adulto	hombre
17	ME-254-32H-05	alto	adulto	hombre
18	ME-257-32H-05	alto	adulto	hombre
19	ME-055-32M-99	alto	adulto	mujer
20	ME-137-32M-01	alto	adulto	mujer
21	ME-247-32M-05	alto	adulto	mujer
22	ME-253-32M-05	alto	adulto	mujer
23	ME-255-32M-05	alto	adulto	mujer
24	ME-259-32M-05	alto	adulto	mujer
25	ME-009-33H-97	alto	mayor	hombre
26	ME-084-33H-00	alto	mayor	hombre
27	ME-140-33H-01	alto	mayor	hombre
28	ME-245-33H-05	alto	mayor	hombre
29	ME-249-33H-05	alto	mayor	hombre
30	ME-294-33H-07	alto	mayor	hombre
31	ME-220-33M-02	alto	mayor	mujer
32	ME-221-33M-02	alto	mayor	mujer
33	ME-227-33M-03	alto	mayor	mujer
34	ME-229-33M-03	alto	mayor	mujer
35	ME-230-33M-03	alto	mayor	mujer
36	ME-264-33M-05	alto	mayor	mujer

En cuanto a las fechas de grabación, el cuadro 2 anota el momento preciso en que se recogió cada una de las grabaciones que forma parte de la serie de materiales transcritos.

Cuadro 2. *Fechas de realización de las entrevistas*

<i>Número</i>	<i>Etiqueta</i>	<i>Fecha</i>
1	ME-042-31H-99	junio de 1999
2	ME-105-31H-00	julio de 2000
3	ME-154-31H-01	mayo de 2001
4	ME-155-31H-01	mayo de 2001
5	ME-190-31H-05	febrero de 2005
6	ME-197-31H-01	noviembre de 2001
7	ME-107-31M-00	julio de 2000
8	ME-181-31M-01	abril de 2001
9	ME-248-31M-05	enero de 2005
10	ME-250-31M-05	enero de 2005
11	ME-251-31M-05	enero de 2005
12	ME-252-31M-05	enero de 2005
13	ME-006-32H-97	agosto de 1997
14	ME-056-32H-99	octubre de 1999
15	ME-138-32H-01	mayo de 2001
16	ME-139-32H-01	mayo de 2001
17	ME-254-32H-05	mayo de 2005
18	ME-257-32H-05	junio de 2005
19	ME-055-32M-99	octubre de 1999
20	ME-137-32M-01	mayo de 2001
21	ME-247-32M-05	febrero de 2005
22	ME-253-32M-05	enero de 2005
23	ME-255-32M-05	junio de 2005
24	ME-259-32M-05	junio de 2005
25	ME-009-33H-97	agosto de 1997
26	ME-084-33H-00	febrero de 2000
27	ME-140-33H-01	mayo de 2001
28	ME-245-33H-05	enero de 2005
29	ME-249-33H-05	enero de 2005

30	ME-294-33H-07	enero de 2007
31	ME-220-33M-02	mayo de 2002
32	ME-221-33M-02	mayo de 2002
33	ME-227-33M-03	mayo de 2003
34	ME-229-33M-03	mayo de 2003
35	ME-230-33M-03	abril de 2003
36	ME-264-33M-05	octubre de 2005

Como puede apreciarse la grabación más antigua se recogió en agosto de 1997, y la más reciente en enero de 2007, lo cual proporciona un arco temporal de diez años al conjunto de la colecta. El grueso de los materiales ahora presentados, sin embargo, se recogió entre los años 1999 y 2005 (las 34 entrevistas restantes, de hecho), segmento que otorga en realidad un grado mayor de concentración al corte temporal efectuado.

El cuadro 3 presenta información adicional pertinente precisamente para la consideración del tiempo aparente vinculado a las grabaciones, entre otros datos:

Cuadro 3. *Otros rasgos de los entrevistados*

<i>Número</i>	<i>Etiqueta</i>	<i>Edad</i>	<i>Año de nacimiento</i> <sup>35</sup>	<i>Estudios</i>	<i>Ocupación</i>
1	ME-042-31H-99	25	1974	arquitectura	arquitecto
2	ME-105-31H-00	25	1975	contaduría	perforador
3	ME-154-31H-01	30	1971	diseño	pasante
4	ME-155-31H-01	26	1975	arquitectura	estudiante
5	ME-190-31H-05	26	1978	ingeniería	ingeniero
6	ME-197-31H-01	29	1972	arquitectura	arquitecto y dueño de constructora
7	ME-107-31M-00	25	1974	letras	estudiante
8	ME-181-31M-01	29	1972	letras	pasante

<sup>35</sup> Considérese, en algunas aparentes discordancias, la relación entre el mes de nacimiento y el mes en que se realizó la entrevista.



9	ME-248-31M-05	30	1975	pedagogía	no trabaja
10	ME-250-31M-05	30	1974	odontología	dentista
11	ME-251-31M-05	21	1983	arqueología	estudiante
12	ME-252-31M-05	21	1983	medicina	estudiante
13	ME-006-32H-97	37	1959	letras	profesor
14	ME-056-32H-99	51	1948	biología	profesor
15	ME-138-32H-01	44	1957	medicina	pediatra
16	ME-139-32H-01	49	1952	odontología	dentista
17	ME-254-32H-05	39	1965	musicología	investigador
18	ME-257-32H-05	45	1960	biología	investigador
19	ME-055-32M-99	38	1961	diseño	comerciante
20	ME-137-32M-01	53	1947	medicina	médico general
21	ME-247-32M-05	43	1961	economía	profesora
22	ME-253-32M-05	37	1967	psicología	no trabaja
23	ME-255-32M-05	41	1963	odontología	dentista
24	ME-259-32M-05	40	1964	historia y danza	bailarina
25	ME-009-33H-97	56	1941	licenciatura inconclusa	contratista
26	ME-084-33H-00	55	1944	economía	floricultor
27	ME-140-33H-01	69	1932	medicina	ortopedista y cirujano
28	ME-245-33H-05	63	1941	historia e ingeniería química	ingeniero
29	ME-249-33H-05	64	1940	medicina y biología	investigador
30	ME-294-33H-07	72	1934	teatro	músico y actor

31	ME-220-33M-02	65	1936	arquitectura	arquitecta y profesora
32	ME-221-33M-02	57	1945	química	profesora
33	ME-227-33M-03	56	1947	licenciatura inconclusa	no trabaja
34	ME-229-33M-03	62	1941	psicología	profesora
35	ME-230-33M-03	62	1941	arquitectura	profesora jubilada
36	ME-264-33M-05	58	1947	letras	profesora

Los informantes más jóvenes incluidos en esta muestra tienen 21 años, y el de más edad alcanza los 72; tal distancia implica un recorrido aparente en la sección sincrónica de aproximadamente medio siglo. En cuanto a los años de nacimiento, las personas nacidas más tarde son de 1983, mientras que el informante nacido más tempranamente es de 1932. Dadas las fechas de recolección del material, puede decirse que la muestra aspira a representar el habla adulta de la segunda mitad del siglo XX.

En cuanto al nivel de estudios, la inmensa mayoría de los entrevistados había terminado los cursos propios de las carreras que estudiaron; sólo en un par de casos sus estudios quedaron inconclusos (ambos en el grupo de edad más avanzada, donde suele ser más difícil encontrar personas con estudios superiores). Además, aunque no se ha anotado en el cuadro, algunos de los entrevistados poseen posgrados. El grupo de nivel alto de estudios presentado está por tanto claramente definido en su conjunto.

La mayor parte de los colaboradores posee ocupaciones afines a los campos de estudio de que proceden. Se incluyen también, por otra parte, algunas personas que no trabajan en la actualidad, algún jubilado y varios estudiantes en etapas finales de sus carreras o pasantes.

Es también pertinente sintetizar los lugares donde se llevaron a cabo las grabaciones:

*Cuadro 4. Lugares donde se llevó a cabo la grabación de las entrevistas*

<i>Lugar</i>	<i>Número de entrevista</i>
Casa del informante	1, 5, 8, 9, 12, 14, 22, 25, 28, 33, 34, 35 <sup>36</sup>
Trabajo del informante	2, 10, 13, 15, 16, 17, 18, 19, 20, 21, 23, 24, 26, 27, 29, 30, 31, 32, 36
Escuela del informante	3, 4, 11
Casa del entrevistador	6
Laboratorio	7

La inmensa mayoría se levantaron en ámbitos propios del informante; sólo en dos casos se trabajó en el entorno ordinario del entrevistador (casa y laboratorio). Esto es importante, pues se pretendía un acceso que respetara las condiciones de vida del hablante y no lo intimidara, dado el propósito general de recoger fragmentos de habla relativamente cotidianos. La mayoría de las entrevistas se efectuaron en el trabajo del entrevistado (diecinueve), y a estas se podrían sumar las tres llevadas a cabo en la escuela del colaborador. Por fin, una docena de grabaciones, la tercera parte, se registraron en la casa misma de los informantes; en el cuadro 4 se anota el número de entrevista correspondiente a cada uno de estos ámbitos.

El último cuadro, por fin, anota los temas abordados en cada una de las grabaciones, aproximadamente en el orden en el que fueron surgiendo.

Cuadro 5. *Temas abordados en las entrevistas*

<i>Número</i>	<i>Etiqueta</i>	<i>Temas generales (en orden aproximado de aparición)</i>
1	ME-042-31H-99	Experiencia de trabajo, arquitectura, construcción
2	ME-105-31H-00	Perforaciones, tatuajes, forma de vida y trabajo
3	ME-154-31H-01	Estudios, amigos, anhelos, trabajo
4	ME-155-31H-01	Educación, amigos internacionales, becas al extranjero, proyectos futuros
5	ME-190-31H-05	Estudios, literatura, política

<sup>36</sup> En el caso de esta entrevista, la grabación se realizó en casa de una amiga de la informante.

6	ME-197-31H-01	Trabajo, familia, matrimonio, amigos comunes
7	ME-107-31M-00	Familia, género, matrimonio, vacaciones
8	ME-181-31M-01	Historia, costumbres y problemas del barrio de Tepito
9	ME-248-31M-05	Situación familiar, hijos, matrimonio, vida académica
10	ME-250-31M-05	Vida escolar, experiencias profesionales, relaciones escolares, viajes
11	ME-251-31M-05	Experiencias en la universidad, antropología, posibilidades de trabajo
12	ME-252-31M-05	Experiencias de trabajo, casos revisados y planes para el futuro
13	ME-006-32H-97	Niñez, terremoto, familia
14	ME-056-32H-99	Cultivo de bonsáis
15	ME-138-32H-01	Experiencias de trabajo, pacientes con casos poco comunes
16	ME-139-32H-01	Trabajo, hijos, vida personal
17	ME-254-32H-05	Teatro, comida, historia de trabajo, viajes
18	ME-257-32H-05	Historia de vida y proyectos de trabajo
19	ME-055-32M-99	Jardinería, bonsáis, plantas y árboles en general
20	ME-137-32M-01	Vida personal, trabajo
21	ME-247-32M-05	Familia, trabajo, gustos y pasatiempos, historia personal, educación de los hijos
22	ME-253-32M-05	Experiencia en la universidad, profesión, trabajo, hijos
23	ME-255-32M-05	Experiencias de vida, infancia, familia, matrimonio, hijos
24	ME-259-32M-05	Historia de vida, carrera, educación, danza, experiencia profesional
25	ME-009-33H-97	Sismo de 1985, llegada a la colonia, construcción de

		la casa, transporte público
26	ME-084-33H-00	Bonsáis, cultura de los mexicanos, leyendas prehispánicas, horticultura
27	ME-140-33H-01	Experiencia como estudiante, académica, laboral y personal
28	ME-245-33H-05	Viajes, idiomas, pasatiempos, política
29	ME-249-33H-05	Familia, matrimonio, hijos, medicina, trabajo
30	ME-294-33H-07	Ciudad de México antes, trabajo, tarot, religión
31	ME-220-33M-02	Historia de vida, estudios, hijos
32	ME-221-33M-02	Familia, profesión, trabajo, escuela, educación de los hijos
33	ME-227-33M-03	Hijos, estudios, historia de vida, viajes, otras culturas
34	ME-229-33M-03	Psicología, escuela de sus hijos, educación, amistades
35	ME-230-33M-03	Educación, trabajo, matrimonio, intereses, experiencias personales
36	ME-264-33M-05	Literatura, experiencias profesionales

Aunque existe gran variedad de temas, sobre todo en los más menudos, es posible establecer algunas generalidades al respecto. Una parte importante de las grabaciones abordan de una manera o de otra la historia de vida del informante, sea de forma general, sea a través de tres dimensiones fundamentales, a saber, el trabajo, la familia y el tiempo libre. Esto es de suma importancia, pues precisamente estas tres dimensiones permiten caracterizar el modo de vida a que se adscribe cada persona (*supra*). La mayor parte de las entrevistas permiten tener un rico panorama de las actividades, la formación, las expectativas, los intereses y las preferencias de los informantes. Aunque esto puede no ser cierto en todas y cada una de las grabaciones, el conjunto ofrece una variada red temática, con numerosas connivencias en cuanto a su organización; creemos que las suficientes para defender la unidad global del material. El tipo de situación nos parece bastar para promover de todos modos unos temas relativamente parecidos y relativamente comparables. Tal decisión nos parece ventajosa sobre procedimientos que corren el riesgo de la artificiosidad

e incluso, si no son bien aplicados, de parecerse más a la resolución de un cuestionario temático que a la realización de una entrevista semiinformal que aspira a llegar a ser en algunos de sus fragmentos una conversación grabada.

En ese sentido, el estilo empleado a lo largo de las entrevistas osciló entre dos extremos. Por una parte, del lado más formal, tenemos un estilo que podría llamarse de *entrevista semiinformal*; por lo regular, se caracteriza por estructuras dialógicas de pregunta-respuesta, turnos cortos por parte del informante y soluciones formales para las diferentes variables lingüísticas. En el otro extremo, tenemos lo que se ha llamado *conversación grabada* (cf. Silva-Corvalán 2001, p. 52 y ss.), que en sus expresiones más logradas llega a mostrar pocas diferencias con las conversaciones ordinarias, simétricas y de tema libre entre conocidos<sup>37</sup>. Es común en algunas de las entrevistas que comiencen por presentar una estructura de entrevista y que a lo largo de su desarrollo terminen por convertirse en conversaciones.

#### REFERENCIAS BIBLIOGRÁFICAS

CONAPO. Consejo Nacional de Población. 2008. “Escenarios demográficos y urbanos de la Zona Metropolitana de la ciudad de México, 1990-2010”. <<http://www.conapo.gob.mx/publicaciones/zmcm.htm#>> [consultado el 3 de septiembre de 2008].

*Cuestionario para la delimitación de las zonas dialectales de México*. 1970. México: El Colegio de México.

---

<sup>37</sup> En algún momento consideramos la posibilidad de introducir al comienzo de la entrevista un comentario general que evaluara el carácter predominante de la grabación, fuera como entrevista o como conversación, pero lo descartamos finalmente por la simplificación que hubiera entrañado. Un procedimiento más adecuado hubiera exigido marcar ciertas partes de las transcripciones con etiquetas más específicas, en las que se anotara, quizá entre otros, aspectos puntuales vinculados al surgimiento y abandono de los temas de conversación; el papel y nivel de participación de los diferentes interlocutores involucrados; partes de la entrevista vinculadas a otros tipos de circunstancias, como velocidad de habla, surgimiento de acontecimientos ajenos al ámbito de la entrevista misma, etc. Como con otras familias de etiquetas, hubiera sido posible, en definitiva, desarrollar un sistema detallado, aunque ciertamente muy interpretativo. No descartamos hacerlo en el futuro. Por otra parte, ya hemos comentado que en otras secciones del trabajo con los informantes —no incluidas en las transcripciones presentadas ahora— se recogen otros estilos de habla, tanto de lectura como de palabras preguntadas aisladamente.

- DU BOIS, JOHN W., *et al.* 1993. "Outline of discourse transcription", en *Talking Data. Transcription and Coding in Discourse Research*. Ed. J. A. Edwards y M. D. Lampert. Hilldale: Lawrence Erlbaum, pp. 45-90.
- EDWARDS, JANE A. 1993. "Principles and contrasting systems of discourse transcription", en *Talking Data. Transcription and Coding in Discourse Research*. Ed. J. A. Edwards y M. D. Lampert. Hilldale: Lawrence Erlbaum, pp. 3-31.
- GRAIZBORD, BORIS, y BEATRIZ ACUÑA 2007. "Movilidad residencial en la ciudad de México", *Estudios Demográficos y Urbanos*, 22, pp. 291-335.
- GUERRERO GALVÁN, ALONSO en prensa. "Otho 'bui. Migrantes otomíes en la ciudad de México", *Lengua y Migración / Language and Migration*.
- HENRÍQUEZ UREÑA, PEDRO 1938. "Datos sobre el habla popular de Méjico", y "Mutaciones articulatorias en el habla popular", en *El español en Méjico, los Estados Unidos y la América Central*. Ed. P. Henríquez Ureña. Buenos Aires: Universidad de Buenos Aires, pp. 277-324 y 329-379.
- INEGI. Instituto Nacional de Estadística, Geografía e Informática. 2004. *Módulo sobre migración 2002. Encuesta nacional de empleo*. México: INEGI.
- LARA, LUIS FERNANDO (dir.). *Diccionario del español usual en México*. <<http://intranet.colmex.mx/dem/>> [consultado en diferentes momentos].
- , y KLAUS ZIMMERMANN 1988. "Mexico", en *Sociolinguistics. An International Handbook of the Science of Language and Society*. Ed. U. Ammon, N. Dittmar y K. J. Mattheier. Berlin - New York: Walter de Gruyter, 2, pp. 1341-1347.
- LASTRA, YOLANDA 1992a. *Sociolingüística para hispanoamericanos: una introducción*. México: El Colegio de México.
- (ed.) 1992b. *Sociolinguistics in Mexico = International Journal of the Sociology of Language*, 96.
- , y PEDRO MARTÍN BUTRAGUEÑO 2000. "El modo de vida como variable sociolingüística en el estudio de la ciudad de México", en *Estructuras en contexto. Estudios de variación lingüística*. Ed. P. Martín. El Colegio de México, México, pp. 13-43.
- LEF. Laboratorio de Estudios Fónicos 2009. "Datos demolingüísticos". <<http://lef.colmex.mx/>>.

- LOPE BLANCH, JUAN M. (dir.) 1971. *El habla de la ciudad de México. Materiales para su estudio*. México: UNAM.
- (dir.) 1976. *El habla popular de la ciudad de México. Materiales para su estudio*. México: UNAM.
- 1983. *Estudios sobre el español de México*. 2a. ed. México: UNAM.
- 1990. *Investigaciones sobre dialectología mexicana*. 2a. ed. México: UNAM.
- 1986. *El estudio de la norma culta. Historia de un proyecto*. México: UNAM.
- (dir.) 1990-2000. *Atlas Lingüístico de México*. México: El Colegio de México – UNAM – FCE, 6 vols.
- MARDEN, C. CARROLL 1938. “La fonología del español en la ciudad de Méjico”, en *El español en Méjico, los Estados Unidos y la América Central*. Ed. P. Henríquez Ureña. Buenos Aires: Universidad de Buenos Aires, pp. 87-187. [Original de 1896].
- MARTÍN BUTRAGUEÑO, PEDRO 2006. “Características variables del español de la ciudad de México”, en *Estudios sociolingüísticos del español de España y América*. Ed. Ana Ma. Cestero Mancera, Isabel Molina Martos y Florentino Paredes García. Madrid: Arco Libros, pp. 153-157.
- 2009. “Inmigración lingüística en la ciudad de México”, *Lengua y Migración / Language and Migration*, 1, 1, pp. 9-37.
- 2010. “El proceso de urbanización: consecuencias lingüísticas”, en *Historia sociolingüística de México*. Ed. R. Barriga y P. Martín. México: El Colegio de México, vol. 2.
- MATLUCK, JOSEPH H. 1951. *La pronunciación del español en el Valle de México*. México: Edición de autor.
- 1952. “La pronunciación en el español del Valle de México”, *Nueva Revista de Filología Hispánica*, 6, pp. 109-120. [También en *Antología conmemorativa. Nueva Revista de Filología Hispánica. Cincuenta tomos*. Ed. A. Rivas e Y. Rodríguez. México: El Colegio de México, 2003, pp. 385-397].
- MORA VÁZQUEZ, TERESA 2007. “Los pueblos originarios en los albores del siglo XXI”, en *Los pueblos originarios de la ciudad de México. Atlas etnográfico*. Coord. T. Mora. México: Instituto Nacional de Antropología e Historia – Gobierno del Distrito Federal, pp. 23-41.



- MORENO DE ALBA, JOSÉ G. 1994. *La pronunciación del español en México*. México: El Colegio de México.
- MORENO FERNÁNDEZ, FRANCISCO 1993. "Proyecto para el estudio sociolingüístico del español de España y América (PRESEEA)", *Lingüística*, 5, pp. 268-271.
- 1997. "Metodología del Proyecto Sociolingüístico para el Estudio del Español de España y de América (PRESEEA)", en *Trabajos de sociolingüística hispánica*. Ed. F. Moreno. Alcalá de Henares: Universidad de Alcalá, pp. 137-167.
- 2003. "Metodología del Proyecto Sociolingüístico para el Estudio del Español de España y de América (PRESEEA) (versión revisada, octubre 2003)". <<http://www.linguas.net/portalsee/Preseea/Metodologia/>> [consultado en diferentes momentos].
- MUSSELMAN SHANK, REGINA 2002. *La construcción del turno de habla en la conversación ordinaria. Datos del español de la ciudad de México*. Tesis doctoral. México: El Colegio de México.
- ORTIZ ELIZONDO, HÉCTOR 2007. "Los pueblos originarios y el inexorable avance de la mancha urbana", en *Los pueblos originarios de la ciudad de México. Atlas etnográfico*. Coord. T. Mora. México: Instituto Nacional de Antropología e Historia – Gobierno del Distrito Federal, pp. 59-71.
- PERISSINOTTO, GIORGIO 1975. *Fonología del español hablado en la ciudad de México: ensayo de un método sociolingüístico*. Trad. Raúl Ávila. México: El Colegio de México.
- PESQUEIRA, DINORAH 2008. "Cambio fónico en situaciones de contacto dialectal: el caso de los inmigrantes bonaerenses en la ciudad de México", en *Fonología instrumental: patrones fónicos y variación*. Ed. E. Herrera y P. Martín. México: El Colegio de México, pp. 171-189.
- PRESEEA. Proyecto para el estudio sociolingüístico del español de España y de América. 2008. "Marcas y etiquetas mínimas obligatorias". <<http://www.linguas.net/preseea/>> [consultado en noviembre de 2009].
- REAL ACADEMIA ESPAÑOLA. Banco de datos (CORDE) [en línea]. *Corpus diacrónico del español*. <<http://www.rae.es>> [consultado en diferentes momentos].

- . Banco de datos (CREA) [en línea]. *Corpus de referencia del español actual*. <<http://www.rae.es>> [consultado en diferentes momentos].
- . *Diccionario de la lengua española*. 22<sup>a</sup>. ed. <<http://www.rae.es/rae.html>> [consultado en diferentes momentos].
- . *Diccionario panhispánico de dudas*. 1<sup>a</sup>. ed. <<http://www.rae.es/rae.html>> [consultado en diferentes momentos].
- REED, ALAN 2008. *Simple Concordance Program*. <<http://www.textworld.com/scp/>> [consultado en diferentes momentos].
- RODRÍGUEZ, ORALIA, y RODNEY WILLIAMSON 1981. “Diferencias sociales en el lenguaje: el caso de las narraciones de niños mexicanos de seis años”, en *Actas del VI Congreso Internacional de Hispanistas*. Toronto: Toronto Press, pp. 606-610.
- SILVA-CORVALÁN, CARMEN 2001. *Sociolingüística y pragmática del español*. Washington: Georgetown University Press.
- ZIMMERMANN, KLAUS 1982. “Perspectivas de la sociolingüística urbana en México”, *Boletín de Antropología Americana*, 6, pp. 105-117.