



CENTRO DE ESTUDIOS LINGÜÍSTICOS Y LITERARIOS

**PROCESOS SOCIOLINGÜÍSTICOS EN ESPAÑOL DE LA CIUDAD
DE MÉXICO. ESTUDIO EN TIEMPO REAL**

TESIS QUE PARA OBTENER EL GRADO DE
DOCTOR EN LINGÜÍSTICA

PRESENTA

JULIO CÉSAR SERRANO MORALES

ASESOR: DR. PEDRO MARTÍN BUTRAGUEÑO

MÉXICO, D. F.

SEPTIEMBRE DE 2014

Esta investigación fue posible gracias al apoyo del Consejo Nacional de Ciencia y Tecnología (CONACYT).

A Eugenia y Juliana

A mis padres: Ana Guadalupe Morales y Jesús Abel Serrano

A mis hermanos: Abel, Carlos, Servando, Claudia y Myrna

AGRADECIMIENTOS

A mis hermosas compañeras de vida: mi Eugenia del alma y mi hermosa Juliana, que me han apoyado tanto en este trayecto. Muchas gracias por todo el impulso, la paciencia, los cuidados y por tanto, tanto amor que me han dado todos estos años.

A Doña Celia y Julieta, que me dieron casa, familia y cariño en 1992-1993.

A todos mis profesores en la Escuela Nacional de Antropología e Historia y El Colegio de México.

A todos mis amigos en la Universidad Nacional Autónoma de México, en especial a Cecilia Rojas, Sergio Ibáñez, Francisco Arellanes, Sabine Pflieger (con Felipe incluido), María Ángeles Soler, Beatriz Arias y Fernando Rodríguez.

Particularmente debo agradecer el conocimiento, la paciencia y generosidad que me ha brindado todos estos años mi profesor y amigo Pedro Martín Butragueño. Muchas gracias por todo, Pedro. Con personas tan excepcionales como tú, las palabras siempre se quedan cortas.

A mis lectores de tesis, muchísimas gracias: Rebeca Barriga Villanueva, Francisco Arellanes y Francisco Moreno Fernández. Por supuesto, todos los errores y omisiones que restan son de mi total responsabilidad.

A mis amigos queridos que me sostienen todo el tiempo, en especial a Miroslava Cruz, Niktelol Palacios, Sergio Rodríguez, Irene (Sirenita) Hernández, Pedro Reygadas, Josefina Guzmán, Telmo Yescas, Diwenny Díaz, Felipe Durán, Eduardo Olivares, Samuel Herrera, Karina Verdín, Maribel Delgado, Francisco Zuno, Carmen Morúa, Jason Vázquez y Raúl Sandoval.

A mis alumnos, porque me dejan hablarles de mis inquietudes y sin protestar, en especial a Eréndira Cervantes, Sara (Aponte) Barrett, Maurel Sánchez de la Barquera y Paulina Landeros.

A mis papás, Lupita y Jesús Abel; a mis hermanos Abel, Carlos, Servando, Claudia y Myrna; a todos mis tíos, en especial a mi tío Servando y mi tía Rosy, y a todos mis primos.

ÍNDICE

	Pág.
INTRODUCCIÓN	IX
CAPÍTULO 1. ASPECTOS TEÓRICOS Y ANTECEDENTES DESCRIPTIVOS	1
1.1. Necesidad histórica de una lingüística empírica y socialmente realista: avances del programa de investigación de Weinreich, Labov & Herzog (1968)	3
1.1.1. Los problemas de la innovación (restricciones) y la transmisión	11
1.1.2. Los problemas del incrustamiento y la evaluación	13
1.1.3. Problema del accionamiento (<i>actuation</i>)	14
1.1.4. Tratando de encontrar respuestas a los problemas fundamentales del cambio	15
1.1.5. Los factores sociales y el problema del cambio	20

Índice

1.2. El problema de la formalización del cambio lingüístico: reglas variables, Teoría de Optimidad y Teoría de Optimidad Estocástica	29
1.3. Antecedentes de investigación	36
1.3.1. Las posibilidades de estudiar la variación y el cambio lingüístico en español de la ciudad de México: corpus disponibles	45
CAPÍTULO 2. MÉTODOS	51
Introducción: la estructura lingüística vista desde el variacionismo	51
2.1. En busca del habla espontánea: la entrevista como fuente de datos sociolingüísticos	53
2.2. Los corpus utilizados	57
2.2.1. Breve historia de los corpus utilizados	58
2.2.2. Características sociales de los informantes en 1970 y 2000	62
2.3. Evaluación de la calidad de los datos y procesos de cuantificación en esta tesis	65
2.3.1. Criterios de transcripción fonética y léxica	69
2.3.2. Métodos cuantitativos y programas utilizados	70
2.4. El análisis en tiempo aparente y tiempo real	72
CAPÍTULO 3. VARIABILIDAD DE LAS CONSONANTES RÓTICAS	81
3.1. Representación fonológica de las róticas del español	81
3.2. El proceso fonético de asibilación de las consonantes róticas en	87

español	
3.3. Las róticas en español de la ciudad de México: variantes consideradas	90
3.3.1. Variante 1: rótica canónica	90
3.3.2. Variante 2: aproximante [ɹ]	92
3.3.3. Variante 3: asibilada [ʒ]	93
3.3.4. Variante 4: realización “corta” o “alargada” de la consonante	99
rótica	
3.4. Análisis estadístico descriptivo de las róticas en español de ciudad de México	100
3.5. Análisis cuantitativo de la variabilidad de la rótica percusiva ante pausa	106
3.6. Análisis cuantitativo de la variabilidad de la rótica vibrante (r)	118
3.7. Formalización de la variabilidad de las róticas en Teoría de Optimidad	129
3.7.1. Formalización en TOE de la variabilidad de la (r) ante pausa	132
3.7.2. Formalización en TOE de la variabilidad de la rótica vibrante (r)	134
3.8. Discusión	135
3.8.1. Movimientos en la jerarquía de restricciones de las róticas en el tiempo real	135
3.8.2. Sociolingüística de las róticas en el tiempo real	139
3.8.3. Hacia una explicación de la retracción de la asibilación	140

CAPÍTULO 4. DEBILITAMIENTO DE VOCALES ÁTONAS	145
4.1. Debilitamiento vocálico en español de la ciudad de México: antecedentes descriptivos	146
4.2. Fonética y fonología del debilitamiento vocálico en español de ciudad de México	152
4.2.1. Variante relajada (o reducida) [v]	153
4.2.2. Variantes sordas [ɣ]	156
4.2.3. Variante elidida	159
4.3. Variables internas en el análisis del debilitamiento vocálico	164
4.4. Criterios de transcripción auditiva	166
4.5. El corpus total de vocales debilitadas: un primer análisis	168
4.6. Peso de los factores lingüísticos y sociales sobre el debilitamiento vocálico	175
4.6.1. Variante relajada	175
4.6.2. Variante sorda	176
4.6.3. Elisión vocálica	179
4.7. En busca de los líderes del debilitamiento vocálico	181
4.8. Análisis en Teoría de Optimidad Estocástica del debilitamiento vocálico	184
4.9. Conclusiones	187

CAPÍTULO 5. LOS LÍMITES DE LOS CORPUS ORALES: UNA REVISIÓN DE LA HIPÓTESIS DE LA DIFUSIÓN LÉXICA	191
5.1. El cambio neogramático y la hipótesis de la difusión léxica	193
5.2. Morfología culta I: el caso del morfema -dad	198
5.3. Morfología culta II: el caso de los adverbios terminados en el morfema -mente	203
5.4. El caso de los grupos consonánticos cultos	207
5.4.1. Los grupos consonánticos cultos en los hablantes de altos estudios	210
5.4.2. Los grupos consonánticos cultos en los hablantes de bajos estudios	212
5.4.3. Panorama general de los grupos consonánticos cultos en ECM	215
5.5. Acercamientos lexicoestadísticos a la diferenciación sociolingüística	217
5.5.1. Volumen y características textuales de los corpora utilizados	220
5.5.2. Regularización del texto original de las muestras	223
5.5.3. Palabras y tipos totales en el corpus	227
5.5.4. Estudio I: listas de vocablos sustantivos más frecuentes	228
5.5.4.1. Un método alternativo para comparar el acervo léxico	231
5.5.4.2. Evaluación del método alternativo de comparación de vocabularios	239
5.5.5. Estudio II: Un análisis alternativo de la variación léxica. El caso de los sustantivos	240

Índice

5.5.6. Estudio III: Campos referenciales en sustantivos	250
5.5.7. Estudio IV: el uso de la relación type/token para el análisis del léxico, p. 250	254
5.5.7.1. RTT en los distintos grupos de edad	255
5.5.7.2. Valores de RTT individuales	257
5.5.7.2.1. Informantes con mayor y menor densidad léxica en 1970	263
5.5.7.2.2. Informantes con mayor y menor densidad léxica en 2000	265
5.5.8. Conclusiones sobre el nivel lexicoestadístico	268
5.6. “Cada palabra tiene su propia historia”. Procesos de difusión, innovación y retracción léxica	270
5.6.1. Índices de mortandad y vitalidad léxica y el estudio del léxico de la ciudad de México	271
5.6.2. Búsqueda de palabras en los corpus originales	276
5.6.3. En busca de otras fuentes	280
5.6.4. “¡Qué buena onda!”	281
5.6.5. La palabra género	284
5.6.6. Celular (teléfono, telefonía)	288
5.6.7. La palabra súper	290
5.6.8. La palabra chido	292
5.6.9. La expresión adjetiva con padre/padrísimo	294

5.6.10. La palabra güey	295
5.6.11. Síntesis de resultados del análisis de palabras aisladas	298
5.7. Discusión: la viabilidad de los corpus orales basados en entrevistas para el análisis sociolingüístico del léxico	303
5.8. Conclusiones	309
CAPÍTULO 6. RECAPITULACIÓN: PROCESOS SOCIOLINGÜÍSTICOS DE CONVERGENCIA Y NIVELACIÓN EN ESPAÑOL DE LA CIUDAD DE MÉXICO	313
6.1. Procesos de nivelación sociolingüística en el nivel fonético, morfológico y léxico	314
6.1.1. Róticas asibiladas	315
6.1.2. Elisión de vocales átonas	317
6.1.3. Los morfemas -dad y -mente como testigos de nivelación sociolingüística	320
6.1.4. Sociolingüística de los grupos consonánticos cultos y el léxico sustantivo	321
6.2. Procesos de convergencia sociolingüística en ECM: en busca de explicaciones	323
6.2.1. Hipótesis sobre las diferencias en el levantamiento de los datos	325
6.2.1.1. La situación de entrevista en 1967-1975	325
6.2.1.2. La situación de entrevista en 1997-2007	326
6.2.1.3. Evaluación de la comparabilidad de los muestreos de 1970 y	327

Índice

2000	
6.2.2. Hipótesis sobre el cambio en el diseño de la audiencia	328
6.2.3. La hipótesis histórica: los cambios lingüísticos documentados reflejan un cambio social más general	332
6.3. Ensayo de geolingüística urbana: ubicación de los líderes de la variación	341
6.4. Procesos sociolingüísticos en la ciudad de México	345
BIBLIOGRAFÍA	349
APÉNDICE. CARACTERÍSTICAS SOCIALES DE LOS 48 INFORMANTES BASE	387

INTRODUCCIÓN

Esta tesis trata sobre algunos procesos de variación y cambio en curso en español de la ciudad de México en el tiempo real. El período comprendido se encuentra alrededor de los años 1970 y 2000. El planteamiento teórico y metodológico es sociolingüístico y variacionista, en la tradición consolidada principalmente por W. Labov y decenas de investigadores en todo el mundo, quienes parten de la premisa de que no es posible entender el cambio lingüístico sin una aproximación realista a los datos, lo que implica poner atención a factores internos (lingüísticos) y “externos” a la lengua (sociales y estilísticos). Los procedimientos inductivos, por lo tanto, son los preferidos para hacer afirmaciones más generales sobre los procesos de innovación, transmisión, difusión, evaluación y activación del cambio, problemas fundamentales que plantearon

Introducción

Weinreich, Labov & Herzog en 1968 como parte de un programa empírico para el estudio del cambio lingüístico.

Esta investigación se concentra en variables lingüísticas de distintos niveles estructurales: fonético, morfológico y léxico. Se basa en dos grupos de entrevistas con hablantes de español capitalino: el primero, levantado entre 1967 y 1975 en dos proyectos coordinados por Juan M. Lope Blanch en la Universidad Nacional Autónoma de México: “Norma culta” y “Habla popular de la ciudad de México”; el segundo, forma parte de las entrevistas del *Corpus Sociolingüístico de la Ciudad de México*, coordinado por Pedro Martín Butragueño y Yolanda Lastra en El Colegio de México y que iniciara el levantamiento de entrevistas en 1997 y concluyera en 2007. De esta manera, el lapso entre una muestra y otra es de alrededor de 30 años.

La idea original era concentrarse sólo en aspectos fónicos –por el evidente valor de los materiales de los 60-70, relativamente ignorados para la investigación de este tipo--; sin embargo, se volvió evidente que estos materiales podían aprovecharse para rastrear variables lingüísticas de distintos niveles estructurales y cuestionar así la ruta de difusión y cambio que seguían unas y otras.

La tesis se organiza de la siguiente manera. En el capítulo 1 se presentan los conceptos fundamentales sobre variación y cambio, una discusión sobre los avances del programa empírico para el estudio del cambio lingüístico, sobre Teoría de Optimidad y Teoría de Optimidad Estocástica (la herramienta formal de los capítulos fónicos) y los antecedentes de esta investigación. En el capítulo 2 se presenta el planteamiento metodológico general que incluye una descripción pormenorizada de

los corpus utilizados. El capítulo 3 trata el problema de la variación en las consonantes róticas; esta variable es de especial interés porque investigaciones en los años setenta señalaban la variante asibilada de (r) y (r) como un posible cambio en curso. El debilitamiento vocálico se aborda en el capítulo 4. Este problema permitirá discutir las propuestas sobre el papel de la frecuencia de ciertas palabras en la difusión de los cambios lingüísticos. El capítulo 5 aborda diferentes variables en el nivel fonológico, morfológico y léxico; en este se muestra que el proceso de variación en el nivel morfológico puede ser muy similar al del nivel fónico (de tipo regular) y que, en contraparte, la variación en el nivel léxico permite confirmar que la afirmación de Gillieron de que “cada palabra tiene su propia historia” es básicamente correcta. En el capítulo 6 se hace una recapitulación de los procesos. Uno de los hallazgos más interesantes es que la comunidad lingüística capitalina parece estar más nivelada; esto obligará a confrontar distintas hipótesis que expliquen estos procesos; finalmente, en el mismo capítulo se hace una propuesta sobre la distribución geolingüística de los informantes en la extensa Área Metropolitana de la Ciudad de México y algunos comentarios sobre posibles rutas de investigación. Además de la bibliografía, se incluye un apéndice al final, con las características sociales de los informantes de esta investigación.

CAPÍTULO 1

ASPECTOS TEÓRICOS Y ANTECEDENTES DESCRIPTIVOS

INTRODUCCIÓN

Esta tesis se enmarca en un contexto de investigación sociolingüística en el que el realismo es parte primordial del enfoque. Se trata, pues, de un acercamiento que busca librarse de apriorismos innecesarios (Haspelmath 2007) y, en cambio, trata de contribuir a la construcción de una teoría del lenguaje a partir de procedimientos principalmente inductivos. Como lo aclara Martín Butragueño (2014: 18), no debe confundirse la perspectiva *realista* con empirismo radical ni con descriptivismo sin falta de metas. De lo que se trata es de estudiar el lenguaje tomando en cuenta que nos enfrentamos a un fenómeno que se encuentra *heterogéneamente ordenado*, como lo plantearon en su momento Weinreich, Labov & Herzog (1968; en adelante “WL&H 1968”) al proponer un programa empírico de investigación del cambio lingüístico.

Las herramientas formales del variacionismo sociolingüístico fueron, desde sus inicios, derivaciones del generativismo chomskiano (cf. Martín Butragueño 2000). Por ejemplo, las *reglas opcionales* planteadas en Chomsky (1957) inspiraron la posibilidad de insertar *reglas variables* en la gramática comunitaria, unidad de análisis de la sociolingüística de la variación y el cambio (Labov 1969). En esta tesis se recuperan también formalismos generativistas, en este caso, adaptaciones de la Teoría de Optimidad (Prince & Smolensky 1993/2004) a los datos variables.

Por otra parte, es válido afirmar que el estudio científico del cambio lingüístico en marcha a través de la óptica variacionista es una corriente de investigación en consolidación en México (Martín Butragueño 2010; Serrano 2007), y por eso es importante ubicar esta tesis en su contexto de producción en un sentido histórico y teórico-descriptivo. La manera específica de discutir los alcances del programa variacionista, el problema de la formalización y la pertinencia de estos trabajos en el contexto mexicano de investigación lingüística se hace a través del estudio de algunos procesos variables en los niveles fónico, morfológico y léxico en español mexicano oral, en una muestra de 48 hablantes levantada en el *tiempo real* (Labov 1996: Cap. 4).

Este capítulo se propone, por lo tanto: 1) hacer una revisión de los avances en el programa empírico para el estudio del cambio lingüístico de WL&H (1968) a partir de dicho texto fundante y los resultados plasmados en los dos primeros volúmenes de la obra *Principios del cambio lingüístico* de W. Labov (1996, 2006); esta revisión implica relacionar dicho programa de investigación con los distintos problemas abordados en esta tesis; 2) discutir las posibilidades de formalización del cambio a

través de *reglas variables* (Labov 1969; Sankoff 1988; Guy 1998), de *jerarquías de restricciones* en términos de Teoría de Optimidad (Prince & Smolensky 1993/2004) y Teoría de Optimidad Estocástica (Díaz-Campos & Colina 2006), el modelo asumido en esta tesis; y 3) presentar los antecedentes de investigación (que incorporan también un estado de la cuestión sobre la investigación sociolingüística en México, particularmente del español) y una discusión sobre las posibilidades prácticas de hacer un estudio de cambio lingüístico a partir de los corpus con que se cuenta actualmente.

1.1. NECESIDAD HISTÓRICA DE UNA LINGÜÍSTICA EMPÍRICA Y SOCIALMENTE REALISTA: AVANCES DEL PROGRAMA DE INVESTIGACIÓN DE WEINREICH, LABOV & HERZOG (1968)

La creación de la lingüística como disciplina y la postulación de sus principios básicos, tales como la naturaleza binaria y arbitraria del signo lingüístico y las dicotomías sintagma/paradigma y *langue/parole* por parte de F. de Saussure, constituyó un hito científico (por cierto, con poco o nulo reconocimiento explícito en las escuelas americanas, llámese descriptivismo, generativismo o cognitivismo [cf. Lara 2000; Lazard 2012]). En el marco de una reflexión para defender la lingüística saussuriana o *Pure linguistics*, G. Lazard aclara que:

Its core is the idea of a radical separation between language (*la langue*) and speech (*la parole*), i.e. between the analysis of language as a system of signs and the observation of the conditions of its use in discourse. It is argued in the *Cours*

de linguistique générale [...] that these two aspects of language cannot be investigated together; there should therefore be two different disciplines, one applied to language as a system, the other to discourse, in other words two linguistic sciences (*deux linguistiques*) [Lazard 2012: 242].

Esta visión del estudio del lenguaje como una escisión en dos subdisciplinas es muy semejante a la adoptada por Chomsky en *Aspects of the Theory of Syntax* (1965), al postular una lingüística de la competencia (*competence*) por un lado y una lingüística de la actuación (*performance*), por otro. Sin embargo, el programa generativista chomskiano, sobre todo en sus primeras versiones (Chomsky 1957, 1965), era bastante enfático en resaltar el poco interés teórico del estudio de la actuación y, en un sentido metodológico, que la teoría lingüística no podía construirse con datos de ningún “corpus levantados sobre el terreno”¹. Aunque desde un punto de

¹ Por supuesto, el mismo Chomsky ha modificado su postura altamente idealista, de modo tal que ahora difícilmente defendería la idea de evitar el trabajo con corpus. El programa de *biolingüística*, que cobra fuerza a partir de 1974 y al que Chomsky se adhiere (cf. Chomsky 2007), incorpora el trabajo lingüístico con equipos multidisciplinarios provenientes de las ciencias del comportamiento, genética, neurociencias y lingüística computacional, entre otras disciplinas. Una discusión reciente sobre el papel de la variación en las teorías formales se encuentra en el interesante trabajo conjunto coordinado por E. Hernández y P. Martín Butragueño de 2013, “Qué es la variación lingüística y por qué puede interesar su estudio desde distintas perspectivas?”, en el que participan además R. Barriga Villanueva, V. Demonte, P. García Mouton, R. Gutiérrez Bravo, E. Herrera Zendejas, I. Pérez Jiménez y J. Pozas Loyo, especialistas en varias áreas de la lingüística (adquisición, sintaxis formal, dialectología, sociolingüística, sintaxis formal en lenguas indígenas, fonología y lingüística histórica), quienes aportan

vista historiográfico es entendible la necesidad teórica y metodológica de concebir el lenguaje humano como un sistema abstracto, ideal y homogéneo, con un objeto de estudio propio, lo cierto es que la comunidad científica de la primera mitad del siglo XX terminó confundiendo los sistemas lingüísticos idealizados obtenidos desde dicha aproximación con *la realidad* misma del lenguaje (Coseriu 1958) que es, en esencia, un hecho concreto, variable y condicionado lingüística y socialmente².

Puede afirmarse, por lo tanto, que el estructuralismo en lingüística, tanto en sus escuelas americana (v.g. Bloomfield 1933) como europea, asumía dos ideas básicas respecto a la variación y el cambio (una de impacto teórico y otra en el ámbito metodológico). La primera, que el cambio lingüístico sólo puede ser abrupto: respecto al cambio fónico, Hockett señala: “Sound change itself is constant and slow. A phonemic restructuring, on the other hand, must in a sense be absolutely sudden” (1958: 456-457). Esto es, sólo podemos ver las consecuencias del cambio; el proceso es imposible de describirse y explicarse desde una perspectiva que, de entrada,

su definición de “variación lingüística” desde sus propias perspectivas. En general, todos coinciden en reconocer la importancia teórica y descriptiva de la variación y la conveniencia de acercamientos empíricos al problema, aunque lo que se entiende por “variación” quizás no sea lo mismo para todos los participantes: por ejemplo, Gutiérrez Bravo señala que en las teorías formales la variación se entiende principalmente como una de tipo interlingüístico, “como resultado y no como un proceso”, mientras que para Martín Butragueño la variación (y su estudio) se inserta en la cuestión más general de una teoría del cambio lingüístico.

² Sobre el *idealismo* chomskiano opuesto al *realismo* variacionista, es muy pertinente la lectura de Figueroa (1994), en especial las págs. 79-84.

considera que el sistema de la lengua es homogéneo, un sistema que, además, se obtiene a partir de la competencia de un hablante-oyente ideal. Es importante mencionar aquí que el hecho de asumir que el cambio lingüístico se lleva a cabo *en el individuo* (idea de Hermann Paul [1880] recuperada --no explícitamente-- por Saussure y Chomsky) forzaba a ver el cambio como una simple congregación de cambios idiolectales generados durante la adquisición (cfr. WL&H 1968: 5-16). La segunda idea, que impacta directamente al método es que “el cambio lingüístico no es observable”. Sin embargo, los trabajos pioneros de W. Labov en la isla de Martha’s Vineyard (1963) y la ciudad de Nueva York (1966) (por mencionar los más emblemáticos) mostraron que era factible observar el proceso de cambio lingüístico en marcha a través de entidades teóricas como la *variable lingüística* o el recurso metodológico del *tiempo aparente*, esto es, el análisis de la distribución de la variación en diferentes grupos de edad (cf. Labov 1996: cap. 3). El punto crucial aquí, es que mientras que para el estructuralismo la lengua y su sistema se encuentra en el individuo (hablante-oyente ideal), el estudio empírico del cambio lingüístico y la variación, tal como lo proponen WL&H (1968) sólo puede llevarse a cabo si tomamos la *comunidad de habla* como la unidad de análisis.

Las mencionadas investigaciones de Labov, junto con el programa de WL&H (1968) se enmarca en un escenario de investigación que se contrapone fuertemente al idealismo radical del generativismo que dominaba en los años 60, sobre todo, en los Estados Unidos. Estos nuevos estudios serán agrupados bajo el término genérico de *sociolinguistics*, aunque desde sus inicios existieron aproximaciones muy distintas a la

compleja relación lenguaje-sociedad. Una de las primeras compilaciones norteamericanas es la de Bright (1966), *Sociolinguistics: Preceedings of the UCLA Sociolinguistics Conference*, donde Labov, Hymes, Gumperz, Fishman y otros presentan trabajos ahora clásicos sobre problemas clave de la relación lengua-sociedad, tales como el concepto de competencia comunicativa, el cambio lingüístico en marcha, el cambio de códigos y las actitudes ante la lengua³.

³ En el año 2005 tuve la oportunidad de conversar con John Gumperz, quien me comentó que precisamente él, Fishman, Hymes, Labov, Bright y otros comenzaron a estudiar la lengua en las calles y en el campo porque la mayoría de sus colegas en las distintas facultades, entusiasmados con el programa generativista, se concentraban en los estudios “de escritorio” que no reflejaban la lengua en su contexto natural de uso. Por supuesto, como ya se mencionó, debe tomarse muy en cuenta que por ‘sociolingüística’ pueden entenderse muchos enfoques particulares. Chambers (1995), por ejemplo, resalta una gran división entre a) las investigaciones que buscan comprender *la sociedad* a través del análisis que sobre ésta tienen el uso diferenciado de lenguas y dialectos particulares y b) las disciplinas que incorporan aspectos sociales en el análisis de objetos *lingüísticos* particulares. Ralph Fasold adopta esta macro-división de la sociolingüística al redactar dos volúmenes claramente diferenciados: *The Sociolinguistics of Society* (Fasold 1991 a) y *The Sociolinguistics of Language* (Fasold 1991 b). Finalmente, P. Trudgill (2000), además de reconocer esta división, sugiere tres grandes grupos de estudios: sociocientíficos (etnometodología), mitad sociológicos y mitad lingüísticos (etnografía de la comunicación, análisis conversacional, psicología social del lenguaje) y, finalmente, los propiamente lingüísticos (variacionismo sociolingüístico); a este último grupo de trabajos Trudgill los llama “lingüística secular” o “sociolingüística verdadera” (2000: 34). Por supuesto, no se pretende aquí hacer un recuento histórico de la sociolingüística. Sobre sus orígenes puede revisarse Shuy 1990 y Koerner 1991; discusiones más recientes son las de Moreno Fernández (2010) y Martín Butragueño (2010).

El seminal ensayo de WL&H (1968), “Empirical foundations for a Theory of linguistic change”, sienta las bases de una lingüística realista que buscará no sólo la descripción, sino las explicaciones últimas de los complejos procesos de cambio lingüístico atestiguados y en curso. Proponen dichos autores unos “Principios generales para el estudio del cambio lingüístico en marcha” (WL&H 1968: 49), que se presentan a continuación:

- (1) El cambio lingüístico no procede de la simple variación inherente a la comunidad: se inicia cuando una alternancia se generaliza en un subgrupo de la comunidad, asume una dirección y presenta diferenciación ordenada.
- (2) Es una ilusión pensar que existe una asociación entre estructura y homogeneidad.
- (3) No toda variabilidad y heterogeneidad termina en cambio, pero todo cambio implica heterogeneidad y variabilidad.
- (4) La generalización del cambio lingüístico en la estructura lingüística no es uniforme ni instantánea.
- (5) El cambio ocurre en la gramática comunitaria, no en la individual.
- (6) El cambio lingüístico se transmite en la comunidad como una totalidad. No se confina a pasos discretos en una familia.
- (7) Puede haber heterogeneidad y variabilidad sin cambio, pero todo cambio implica heterogeneidad y variabilidad.

(8) Factores lingüísticos y sociales están íntimamente relacionados en el cambio lingüístico.

Desde esta renovada perspectiva, WL&H (1968) plantean los cinco problemas básicos en el estudio del cambio lingüístico: *innovación*, *transmisión*, *inserción* (en la estructura lingüística, en la comunidad de habla y su difusión entre comunidades de habla), *evaluación* (psicológico-social) del cambio y *accionamiento* (el mecanismo general del cambio).

Como puede verse, es indispensable ver el lenguaje como una heterogeneidad ordenada. El *axioma de categoricidad* (“el sistema de la lengua es invariable”) que ha permeado el pensamiento lingüístico estructuralista (cf. Milroy & Gordon 2003: 1-8), al negar la posibilidad del análisis de la variación, vuelve necesario concebir el lenguaje como un sistema con reglas de distinta naturaleza: por un lado, un conjunto de *reglas categóricas* (o cuasi categóricas) que establecerían el núcleo (relativamente) estable del sistema, tal como es presentado en las gramáticas descriptivas, y por otro, subconjuntos de *reglas sociolingüísticas* como las que plantea Dittmar (1996) (discutidas con detalle en Martín Butragueño 2010: 61-66; los ejemplos se inspiran en los presentados ahí): 1) *reglas regulativas* (del tipo *si X, entonces Y*, ej.: ‘una vocal átona ante pausa se ensordece’), 2) *reglas constitutivas* (*en el contexto Z, x cuenta como Y*, ej.: ‘en una conversación entre hombres jóvenes, *guëy* cuenta como un vocativo neutro’) y 3) *instrucciones* (*encuentra una z minúscula, y ve que en el contexto z, x cuente como y*; ej.: las reglas para el saludo cotidiano, o el cambio de turno en una

conversación). Resume Martín Butragueño (2010) las diferencias en los siguientes términos:

Las reglas regulativas describen actividades y comportamientos independientes de las reglas; las reglas constitutivas se refieren a acciones que dependen de las reglas mismas; las instrucciones son formulaciones muy sensibles al contexto. En su manifestación lingüística, las expresiones regulativas son las propias de la variación fónica y gramatical, las constitutivas se asocian a las instituciones léxico-semánticas y las instrucciones describen el significado construido en la interacción sociolingüística. La diferencia principal entre los tres tipos de reglas es la cantidad de contexto que incorporan: poco y con aspiraciones de generalidad en las reglas regulativas [...], medianamente involucrado cuando se trata de la variación etnosemántica propia de las reglas constitutivas, y plenamente presente en el desarrollo de las instrucciones [...]. [Martín Butragueño 2010: 65-66].

Ejemplos de estos tipos de reglas serían las que se redactan para los problemas fónicos y morfológicos abordados en los Caps. 3, 4 y 5 de esta tesis (reglas regulativas), mientras que el seguimiento de ciertas palabras y su evolución (en el Cap. 5) puede verse como un tipo de regla constitutiva. Esto es, la propuesta de Dittmar (1996) permite un análisis homogéneo de la variabilidad en diferentes niveles de análisis lingüístico. Por supuesto, existe el problema de cómo formalizar estos tipos de reglas. En esta tesis se optará por aparatos formales como Teoría de

Optimidad Estocástica (en caso de los problemas fónicos) o desde la lexicología (en el seguimiento histórico de palabras individuales).

1.1.1. *Los problemas de la innovación (restricciones) y la transmisión*

WL&H (1968) señalan que la innovación es quizás uno de los problemas más importantes: ¿Cuáles son las restricciones del cambio? ¿Cuáles cambios son posibles? -sólo una teoría muy fuerte podría predecir esto-- y ¿cuáles cambios son imposibles? -una teoría un poco más débil, a la cual puede llegarse, ciertamente, sí podría responder a esta pregunta--⁴. Esta tesis abordará el léxico (en el cap. 5) como una zona privilegiada de innovación (y retracción).

⁴ Aunque cf. Croft (2000) y sus *desiderata* para una teoría del cambio lingüístico. Para dicho autor, no sólo habría que explicar el mecanismo del cambio, sino también por qué las lenguas preservan sus estructuras:

a theory of language change must explain why languages do NOT change in many ways, [...] Many theories of language change focus their attention on mechanisms to bring about change. But if those mechanisms were the only mechanisms around, then languages would be changing constantly in all of their respects. Yet they are not. A theory of language change must provide for mechanisms that act as forces for stability as well as for change [Croft 2000: 4, énfasis en el original].

Por supuesto, hay que entender también por qué se da la estabilidad del sistema, pero lo cierto es que a pesar de los innegables avances del programa laboviano, todavía no han quedado completamente claros los mecanismos del cambio lingüístico, de ahí el sentido de esta tesis.

Por otra parte, el problema de la transición o de la transferencia es también fundamental: ¿cómo van aprendiendo los hablantes las formas innovadoras? ¿Cuál es el papel del entorno familiar y social en la adquisición de la variación?, etc. Lo que WL&H (1968) encontraban es que dicha adquisición parece ocurrir en los años de la pre-adolescencia y conferían una gran importancia a los hablantes bidialectales en este proceso. Toda la evidencia disponible hasta ese momento apuntaba hacia el hecho de que los niños no mantienen el dialecto de los padres, sino el de sus amigos de la misma edad o de edades muy cercanas. Los posteriores estudios de Guy & Boyd (1990), Chambers (1992) y Roberts (1997) confirmaron esta idea⁵. Sobre todo, el proceso de *emparejamiento de la probabilidad* (Labov 1996: 882-887), compartido con otras especies animales, y que consiste en una imitación del comportamiento variable comunitario por parte de los aprendices, explica en parte el fenómeno de la adquisición de la variación y el mantenimiento de la variación estable. En los capítulos 3 (sobre asibilación de consonantes róticas) y 4 (sobre debilitamiento de vocales átonas) de esta tesis se confirman patrones de aparente no transmisión de las formas más extremas de divergencia con el estándar; por supuesto, se verá que pesan también los factores sociales en estos procesos de retracción del cambio.

⁵ En Serrano (2002) se documenta el caso de cinco inmigrantes dialectales que, al igual que ocurre en los estudios mencionados, abandonan las formas divergentes de sus padres –inmigrantes sonorenses en la ciudad de México.

1.1.2. *Los problemas del incrustamiento y la evaluación*

El tercer problema abordado en WL&H (1968) es el del incrustamiento (*embedding*). Este problema tiene 2 aspectos: a) el incrustamiento en la estructura lingüística: toda lengua tiene i) capas discretas, disponibles para toda la comunidad de habla; ii) variables intrínsecas, que covarían con elementos lingüísticos y extra-lingüísticos; y b) el incrustamiento en la estructura social; la cambiante estructura lingüística está inserta en la comunidad de habla de tal modo que las variaciones sociales y geográficas son elementos intrínsecos de la estructura. Normalmente se ha encontrado que una variable *no* tiene correlatos sociales de significación al inicio y al final del proceso de cambio. Muy asociado a este se encuentra el cuarto problema, el de la evaluación social del cambio, según el cual deben establecerse empíricamente los correlatos subjetivos de los diferentes estratos [sociales] y las variables lingüísticas en una estructura heterogénea; el nivel de consciencia social es una característica importante del cambio lingüístico que debe determinarse de manera directa. Respecto a estos problemas, el fenómeno de elisión de vocales átonas en español mexicano (cap. 4 de esta tesis) mostrará que los correlatos sociales son aún poco definatorios en la actual etapa de variación (no existe una fuerte evaluación social del proceso), mientras que lo que se podrá ver en el capítulo 5 es cómo el léxico sustantivo, con el paso del tiempo, se distribuye de manera cada vez más equilibrada entre grupos de hablantes de niveles socioeducativos altamente contrastantes.

1.1.3. *Problema del accionamiento (actuation)*

El quinto problema para el estudio del cambio, el *accionamiento*, es fundamental, y viene a englobar a los demás: ¿Por qué se origina, cómo se difunde y consolida un cambio lingüístico? ¿Cuál es el peso de los factores sociales e internos en el cambio lingüístico? ¿Es posible predecir cuándo surgirá un cambio, a través de qué procesos específicos y en qué parte de la estructura lingüística? WL&H (1968) sugieren el proceso de cambio lingüístico en los siguientes términos:

It is suggested that a linguistic change begins when one of the many features characteristic of speech variation spreads throughout a specific subgroup of the speech community. This linguistic feature then assumes a certain social significance — symbolizing the social values associated with that group (cf. Sturtevant 1947: 81 ff.). Because the linguistic change is embedded in the linguistic structure, it is gradually generalized to other elements of the system. Such generalization is far from instantaneous, and change in the social structure of the community normally intervenes before the process is completed. New groups enter the speech community and reinterpret the on-going linguistic change in such a way that one of the secondary changes becomes primary [...]. The advancement of the linguistic change to completion may be accompanied by a rise in the level of social awareness of the change and the establishment of a social stereotype. Eventually, the completion of the change and the shift of the variable to the status of a constant is accompanied by the loss of whatever social significance the feature possessed. The high degree of regularity which sound

change displays is the product of such loss of significance in the alternations involved, and the selection of one of the alternants as a constant [WL&H 1968: 48-49].

El análisis del proceso de asibilación de las róticas finales (Cap. 3 de esta tesis) permitirá discutir algunos de los pasos de este proceso: al parecer, cuando una variante cobra notoriedad en el imaginario lingüístico comunitario (volviéndola un estereotipo), puede ocurrir en realidad una de dos soluciones: se consolida el cambio, o se provoca todo lo contrario: la retracción del proceso. También el análisis lexicológico de ciertas piezas léxicas y su inserción en la comunidad de habla (en el cap. 5) permitirá ejemplificar el proceso de accionamiento de un cambio lingüístico en ese nivel particular.

Los avances respecto a los cinco problemas principales en el estudio del cambio lingüístico se discuten en la siguiente sección.

1.1.4. Tratando de encontrar respuestas a los problemas fundamentales del cambio

Las generalizaciones más importantes que WL&H encontraban estaban basadas en su mayoría en problemas de índole fónica (WL&H 1968: 47). Éstas fueron:

- a) cuando dos sistemas entran en contacto y uno tiene dos fonemas y otro uno, el resultado será tener un sistema con un solo fonema.

b) respecto a los cambios en cadena (*chain shifts*), las vocales periféricas⁶ (tensas) se elevan, pero no sucede a la inversa.

c) no se ha encontrado una variable que se origine como estereotipo y variación estilística y que después emerja como variable social sin variación estilística.

Tras poco más de 30 años de investigaciones, mucho se avanzó respecto a los cinco problemas del cambio, se han añadido nuevos principios y se han ido matizando las generalizaciones de (a-c)⁷. En el vol. 1 de los *Principios del cambio lingüístico*, Labov (1996 [1994]) propone varios principios; aunque la base empírica principal para estos principios descansa en macro proyectos como el “Estudio Cuantitativo del Cambio Fónico en Marcha” (LYS) (1968-1972) –con datos de dialectos americanos y británicos del inglés--; el de “Cambio y Variación Lingüística en Filadelfia” (LCV)

⁶ Definidas así: “las vocales periféricas están más cerca de la periferia del espacio vocálico que las vocales no periféricas, definición que se sostiene para la posición articulatoria además de para las propiedades acústicas. En las lenguas germánicas y bálticas, las vocales *tensas* son periféricas, y las vocales *laxas* son no periféricas. En el espacio fonológico de las lenguas románicas, las vocales anteriores no redondeadas y las posteriores redondeadas son periféricas; las vocales anteriores redondeadas o ‘mixtas’ son no periféricas” (Labov 1996: 913, cursivas en el original).

⁷ Aunque es importante aclarar que este no es lugar para describir en toda su complejidad estos principios: su planteamiento de origen, la evidencia empírica que los soporta y muchos matices descriptivos se encuentran claramente expuestos en los tres volúmenes de W. Labov (1996, 2006, 2010). La siguiente exposición sólo busca aclarar el tipo de aporte que puede hacer la presente investigación sobre el español mexicano respecto a la precisión y poder explicativo de estos principios generales.

(1973-1977); o los levantamientos dialectológicos que sirvieron para elaborar *The Atlas of North American English* (Labov, Ash & Boberg 2007); es importante señalar que todos los principios se ponen a prueba en varias lenguas con cambios atestiguados, lenguas pertenecientes, en la medida de lo posible, a familias no indoeuropeas⁸.

El primer grupo de principios tiene que ver con las mutaciones en cadena (*chain shifts*), analizadas en la Parte B del volumen 1. El *Principio de Mutación Vocálica* sintetiza otros tres (en Labov 1996: cap. 8): “En las mutaciones en cadena, las vocales periféricas se hacen más abiertas y las vocales no periféricas se hacen menos abiertas” (Labov 1996: 913). Como aclara Labov, este Principio, sin ninguna restricción, generaría sistemas vocálicos con acumulación de vocales altas en el subsistema tenso periférico y de vocales bajas en el subsistema laxo no periférico, y por lo tanto, propone otros tres principios de las mutaciones en cadena que impiden dicho resultado (pp. 913-914): el *Principio de la salida inferior* (“En las mutaciones en

⁸ Por ejemplo, respecto a los tres principios de mutación vocálica en cadena, Labov presenta las lenguas en las cuales se han evidenciado dichos principios (tabla 5.1. de la versión en inglés de 1994: 122) y son 22 en total; si bien la mayoría pertenecen al gran tronco indoeuropeo (de variadas ramas como la latina, eslava, germánica, celta, entre otras), también figuran lenguas semíticas (sirio), lolo-burmesas (akha) y coreano. Lenguas de la familia tupi-guaraní, por otra parte, permiten confirmar el “Principio VIII (Principio de desmarcación)” (pp. 288-291), mientras que lenguas como el eskikun (austronesia) y varias lenguas sino-tibetanas permiten discutir los principios de difusión léxica, las escisiones y las fusiones (pp. 444-456). Esto es, los principios sí son puestos a prueba con evidencia en una diversidad de familias lingüísticas.

cadena, las vocales bajas no periféricas se hacen periféricas”), el *Principio de la salida superior* (“En las mutaciones en cadena, la primera de dos moras altas puede cambiar la perifericidad, y la segunda puede hacerse no periférica”) y el *Principio de la salida media*: “En las mutaciones en cadena, las vocales periféricas que ascienden de posición media a alta desarrollan semivocales medias”.

Al *Principio de Mutación Vocálica* lo complementan otros dos que regulan la interacción entre los subsistemas vocálicos (p. 914): el *Principio de redefinición* (“La perifericidad se define con relación al sistema vocálico en conjunto”) y el *Principio de desmarcado* (“En las mutaciones en cadena, los elementos del sistema marcado se desmarcan”). Estos Principios constituyen un gran avance respecto al problema de las restricciones del cambio, particularmente respecto al cambio vocálico en cadena.

La Parte C de los *Principios del cambio lingüístico* trata sobre escisiones, fusiones y cuasi-fusiones. Los Principios propuestos son dos: el *Principio de Garde* (“Las fusiones son irreversibles por medios lingüísticos”) y el *Principio de Herzog* (“[e]l correlato espacial del Principio de Garde” [Labov 1996: 915]: “Las fusiones se extienden a expensas de las distinciones”). Éste último viene a reforzar la generalización mencionada en WL&H (1968), presentada *supra* en (a) y que podría parafrasearse de la siguiente manera: dados dos sistemas en contacto, prevalecerá aquél que tenga un sistema más simple⁹. Evidentemente, estos Principios aportan

⁹ Aunque es posible atestiguar cambios en la dirección opuesta, como el documentado por Villena Ponsoda (2008) sobre la reciente escisión /s:/θ/ en español de Andalucía.

para resolver el problema del incrustamiento en el sistema (que incluye el problema de la difusión del cambio en la estructura lingüística y social).

En la Parte D, Labov revisa el *Principio de regularidad* neogramático, reformulado de la siguiente manera: “El cambio fónico es un cambio en la realización fonética de un fonema, sin considerar la identidad léxica” (p. 916). Este principio se opone a la *Hipótesis de la difusión léxica* (Chambers & Trudgill 1998: 159-165), que propone que el cambio se difunde gradualmente, *palabra por palabra*; a este principio le renombrará Labov *Principio del cambio de categoría*: “Los cambios que afectan a varios rasgos de un sonido simultáneamente proceden alterando la filiación categorial de las palabras individuales” (p. 916). El problema del ensordecimiento y elisión de las vocales átonas en español mexicano (que se verá en el cap. 4 de esta tesis) permitirá hacer una revisión de este principio al analizar separadamente las palabras frecuentes (que presentan un mayor desgaste articulatorio) de las palabras no frecuentes.

Los aspectos relacionados con el cambio y la retención de significado (la cuestión ‘funcional’) se analizan en la Parte E. Labov concluye que existen pruebas suficientes a favor del *Principio mecánico* neogramático, formulado así: “El avance relativo del cambio fónico se determina por factores fonéticos sólo, sin considerar la preservación del significado” (p. 917) y comenta: “Buena parte de la discusión [en aquel volumen] tuvo que ver con variación sincrónica; se encontró que la elección de una variante lingüística u otra no la determina la necesidad de preservar información, sino que se ve influida por la tendencia a mantener estructura paralela y articulación

paralela” (p. 917). Un hallazgo muy interesante es que en los procesos de elisión de /s/ y /n/ finales en español y portugués opera un *Principio de compensación estructural*: “Cuando la tasa de elisión de un rasgo significativo de una lengua se incrementa, la frecuencia de los rasgos que de modo redundante portan este significado se incrementará” (p. 917). Por esto Labov afirma que los efectos funcionales son *consecuencia* del cambio –que se opone a la idea tradicional, funcionalista (Martinet 1974), de que es la necesidad de preservar el significado la que configura el rumbo del cambio lingüístico.

Una de las conclusiones más importantes de Labov (1996) es la siguiente: “En conjunto, los hallazgos de este volumen muestran que la caracterización neogramática de la estructura lingüística es esencialmente correcta. Esa estructura es un sistema en gran parte mecánico, fuera del alcance del reconocimiento o ajuste conscientes por parte de sus usuarios” (Labov 1996: 918). Sin embargo, no todos los cambios lingüísticos atestiguados recientemente muestran un desarrollo mecánico, y justo es aquí donde el peso de los factores sociales entra en la construcción de un modelo teórico del cambio lingüístico.

1.1.5. Los factores sociales y el problema del cambio

En las conclusiones del volumen 2 de los *Principios del cambio lingüístico* (2006), Labov enfatiza el hecho de que ciertos cambios no se apegan a los principios estructurales (internos y “mecanicistas”), por lo que hay que acudir a explicaciones desde el ámbito de lo social:

A pesar de los muchos triunfos de los principios estructurales y físicos al explicar el cambio, muchos cambios los desafían. El nuevo desarrollo en la fonología de Filadelfia [...] es el descenso de /e/. Ésta no es una mutación en cadena acompañada por movimientos correlacionados de /i/ y /æ/. La considerable superposición de /e/ y /æ/ corre en dirección contraria a las presiones conservadoras de la economía funcional, y al mecanismo de emparejamiento de la probabilidad propuesto en el capítulo 20 del volumen 1 [...]. Tenemos que inferir que el descenso de (e) ha recibido alguna otra fuente de apoyo más allá del principio relativamente débil de que las vocales breves descienden con más frecuencia de lo que ascienden. *En ausencia de una explicación estructural, este movimiento puede atribuirse sólo a una proyección de los valores sociales en el proceso fonético [...].* Una investigación en las causas de un cambio dado se llevará a cabo como una investigación de los líderes de ese cambio [Labov 2006: 746, subrayado mío].

Esto es, aunque algunos cambios lingüísticos pueden seguir principios mecánicos (internos), otros están fuertemente condicionados por la evaluación social de la variación. Por supuesto, el problema que se enfrenta cuando vemos los hechos sociales como factores explicativos, es que la evolución de las sociedades no tiene un hilo conductor, lo que obliga a postular un *Principio de contingencia*: “Los factores que determinan el curso del cambio lingüístico se derivan de un patrón de comportamiento social que no está vinculado de ninguna manera predecible con el

resultado lingüístico. Se establece una cierta vinculación entre las dimensiones sociales y los rasgos lingüísticos como resultado de una concatenación arbitraria y accidental en la historia” (Labov 2006: 751). Por lo tanto, los cambios en la composición demográfica de la comunidad estarían entre los primeros factores sociales que pueden explicar el desarrollo de un cambio lingüístico. Una versión actualizado de este principio lo encontramos en la perspectiva de Martín Butragueño (2010) en su análisis de la historia sociolingüística de México: *el catastrofismo* es la lente que mejor permitiría observar y entender la dinámica de las lenguas prehispánicas y los desarrollos sociolingüísticos en México; las catástrofes en este caso son la colonización española (siglo XVI), las revoluciones sociales de 1810 y 1910 y los fuertes procesos de migración y urbanización a partir sobre todo de la década de 1950 (Martín Butragueño 2010: 4).

Para Labov (2006), el estudio de la influencia de los factores sociales en el cambio lingüístico estará centrada en la búsqueda de los *líderes lingüísticos*. Los líderes lingüísticos no son necesariamente los innovadores, sino los personajes que tienen influencia suficiente en su entorno inmediato (redes sociales) para difundir los rasgos lingüísticos innovadores. El estudio del papel de los líderes inicia sobre todo con las investigaciones en redes sociales por parte de L. Milroy en la ciudad de Belfast (1980), quien encontraba que los miembros menos centrales en redes sociales menos densas y menos múltiples (esto es, redes donde no todos los miembros se conocen entre sí, y que no tienen relaciones múltiples: familia, trabajo, vecindario y tiempo libre) eran quienes difundían las innovaciones. Los hallazgos de Labov (2006: 292) le

permiten proponer un primer Principio de curvilinealidad: “el cambio lingüístico desde abajo se origina en un grupo social central, situado en el interior de la jerarquía socioeconómica”; más adelante añade:

No hay nada acerca de las ocupaciones de cuello blanco o azul que distinga a los líderes del cambio lingüístico desde abajo. Más bien, es su posición central en la comunidad: [...] no meramente centrales en términos de la jerarquía socioeconómica abstracta, sino centrales en términos de la actividad local, la interacción local y el prestigio local [Labov 2006: 293].

El perfil de los líderes que traza Labov es el de mujeres que en su juventud se inconformaban con las normas establecidas y que ascendieron en la escala social. De hecho, Labov propone una *Hipótesis de la inconformidad* que abarcaría los procesos variables que tradicionalmente se analizaban en un eje de formalidad/informalidad o a través de nociones como tabú lingüístico o diferencias entre grupos sociales, como la existente entre *jocks* y *burnouts* en las investigaciones de P. Eckert en escuelas preparatorias de la ciudad de Detroit (Eckert 1989). La Hipótesis de la inconformidad se plantea de la siguiente manera: “La primera estratificación social del lenguaje adquirida por los niños es la reinterpretación de la estratificación estilística

en la dimensión formal/informal como habla conformista/inconformista” (Labov 2006: 766)¹⁰.

Para el caso del español, Martín Butragueño (2006) extiende esta noción de líderes en el marco de una investigación sociolingüística de la ciudad de México, la

¹⁰ Croft (2000), desde su perspectiva “evolutiva”, también defiende una idea similar al principio de inconformidad:

[T]he Theory of Utterance Selection for language change puts linguistic convention at center stage [...]. Normal (i.e. identical) replication of linguemes in utterances is conforming to the linguistic conventions of the speech community. Altered replication of linguemes in utterances – the creation of variants – is a causal consequence of not conforming to the linguistic conventions of the speech community. The reasons for nonconformity are the causal mechanisms of altered replication [Croft 2000: 30].

El concepto *lingueme* (que traduzco *lingüema*) es central en la teoría de Croft: “[...] the paradigm replicator in language is the lingueme, parallel to the gene as the basic replicator in biology; an utterance is made up of linguemes, and linguemes possess linguistic structure. [...]” (Croft 2000: 28); cuando un lingüema replica uno de sus genes con una forma alternativa (un *alelo*) se produce una *variante de lingüema*:

[T]hat is, alternative structures used for a particular structural element, such an alternative phonetic realizations of a phoneme, alternative words for the same meaning, or alternative constructions used to express a complex semantic structure such as comparison. The locus for a set of variants is essentially the VARIABLE in the sociolinguistic sense of that term, that is, ‘two ways of saying the “same thing”’ (Weinreich, Labov & Herzog 1968: 162 [...]) [Croft 2000: 28].

comunidad objeto de análisis aquí. Martín Butragueño identifica al menos cinco tipos principales de líderes:

[...] *líderes del cambio* (las personas que van por delante en la difusión de los cambios lingüísticos); *líderes de la variación* (los hablantes adscritos con más firmeza a las variantes prestigiosas); *líderes en la interacción* (los individuos dominantes en las conversaciones y otros encuentros lingüísticos); *líderes en instituciones* (con más peso cuanto mayor es la participación de la institución en el mercado verbal); *líderes públicos* (constructores del discurso de liderazgo) [...] (Martín Butragueño 2006: 206-207, subrayado mío).

Esta tipología se discute a la luz de los procesos variables fónicos, léxicos y morfológicos de los Caps. 3-5 de esta tesis. El papel decisivo de las mujeres en el liderazgo de los procesos de cambio lingüístico en marcha se contrapone con su papel en los fenómenos de variación estable (los más comunes en las lenguas). Esta tendencia de las mujeres a favorecer las variantes prestigiosas en situación de variación estable se establece en el *Principio de conformidad lingüística de las mujeres*: “con las variables sociolingüísticas estables, las mujeres muestran una tasa más baja de variantes estigmatizadas y una tasa más alta de variantes prestigiosas que los hombres” (Labov 2006: 406). De hecho, el caso de las vibrantes asibiladas (Cap. 3) mostrará que se mantiene actualmente este principio tras perder su prestigio inicial la variante fricativa.

El problema de la evaluación también es abordado a través de pruebas de evaluación subjetiva del habla de 4 individuos de Filadelfia de clases socioeconómicas distintas. El resultado de estas pruebas muestra que existe en realidad mucha regularidad en todos los informantes en su manera de evaluar subjetivamente los distintos acentos de la prueba. De esta manera, se propone el *Principio de la evaluación uniforme*: “Una variable lingüística estratificada de modo regular es evaluada de manera uniforme por la comunidad de habla” (Labov 2006: 331-332). Dada la naturaleza de los corpus disponibles, la presente tesis no aborda directamente el problema de la evaluación, ya que los materiales de los años 1960-1970 no preveían el estudio de actitudes y creencias; sin embargo, el fenómeno de asibilación de vibrantes permite abordar un poco el problema, ya que se trata de un proceso variable que cada vez se encuentra más presente en el imaginario sociolingüístico de la comunidad estudiada, al grado de que existen estereotipos televisivos (de habla *gay*) que asibilan las vibrantes (Serrano 2008; Eller 2013).

La amplia evidencia empírica manejada en los dos volúmenes de *Principios del cambio lingüístico* permite a Labov proponer dos principios que explicarían la compleja interacción de factores como el sexo, la clase socioeconómica, la idea de prestigio y ascenso social. El primero de ellos sería el *Principio de la inconformidad*: “los cambios lingüísticos en curso son emblemáticos de la inconformidad a las normas sociales establecidas acerca del comportamiento apropiado, y se generan en el medio social que desafía de manera más consistente esas normas” (Labov 2006: 770). Ahora, dicho principio debe complementarse con uno más que lo vincule con el *Principio de*

curvilinealidad arriba mencionado (que son los estratos medios, de movilidad ascendente, los que difunden vigorosamente los cambios lingüísticos); este será el *Principio de la inconformidad constructiva*: “Los cambios lingüísticos los generalizan en la comunidad más amplia aquellos que muestran los símbolos de la inconformidad en un patrón mayor de movilidad social ascendente” (Labov 2006: 771).

Todos estos principios permiten a Labov proponer un guión del cambio lingüístico que incorpora precisión a la propuesta para resolver el problema del accionamiento de WL&H (1968), pero antes aclara que “las siguientes etapas representan la trayectoria social de un cambio lingüístico que asume evaluación social; no se aplica a la mayor parte de las fusiones y a otros cambios que recorren su curso sin aceptar valor simbólico de ninguna clase” (Labov 2006: 772). En esta tesis se revisan dos procesos variables fonéticos que contrastan en el sentido de la evaluación social: como ya se mencionó *supra*, la asibilación de vibrantes parece estar cobrando fuerza en el imaginario sociolingüístico, pero la elisión y ensordecimiento de vocales átonas no, así que podrán ser pruebas interesantes para este ‘guión’ del cambio:

- 1 Para un fonema dado con promedio P y vecinos asimétricos, las formas atípicas en la dirección del vecino más lejano V1 se oyen como casos válidos con más frecuencia que las formas atípicas en la dirección del vecino más cercano V2.
- 2 Los nuevos aprendices lingüísticos adquieren un valor promedio P' mudado en la dirección de V1.

3 Los pasos (1) y (2) continúan en las generaciones subsiguientes con independencia de la evaluación social, de modo que tanto P' como las formas atípicas que van hacia V1 se mueven constantemente en la dirección de V1.

4 Las formas atípicas en la dirección de V1 (formas atípicas tipo V1) se oyen como características de los hablantes jóvenes y del habla enfática y menos controlada, desviadas de la norma aceptada de los hablantes de más edad.

5 Los hablantes inconformistas más jóvenes usan por consiguiente una frecuencia más alta de formas atípicas V1.

6 Los hablantes incrementan su empleo de formas atípicas V1 más allá de los hombres, de modo que estas formas atípicas se oyen ahora como características del habla femenina.

7 Los hablantes varones inconformistas se retiran del uso de formas atípicas V1.

8 Los hablantes inconformistas con movilidad ascendente propagan el uso de formas atípicas V1 y mueven los promedios de P' a los límites de la comunidad de habla.

9 Se desarrolla un patrón de estratificación social de P, con los valores más altos de la variable en comunidades con la concentración más alta de hablantes inconformistas con movilidad ascendente.

10 El uso de formas atípicas V1 asciende a la consciencia pública y empieza una corrección social irregular [Labov 2006: 772-773].

El problema de las variantes asibiladas de las róticas (en Cap. 3), asociado sobre todo al habla femenina, se encuentra en el paso 10: la asibilación cobró

conciencia en la comunidad hablantes en los años sesenta; por lo tanto, será interesante observar el devenir de este proceso con la muestra en tiempo real propuesta, levantada alrededor del año 2000. En general, la enorme contribución de Labov al entendimiento de los mecanismos del cambio lingüístico permite una puesta a prueba de algunos de estos principios en un contexto distinto al de las comunidades angloparlantes sobre las que descansa buena parte de la base empírica laboviana ¹¹, tarea que se irá viendo en cada capítulo de esta tesis.

1.2. EL PROBLEMA DE LA FORMALIZACIÓN DEL CAMBIO LINGÜÍSTICO: REGLAS VARIABLES, TEORÍA DE OPTIMIDAD Y TEORÍA DE OPTIMIDAD ESTOCÁSTICA

Desde sus inicios, el estudio del cambio lingüístico en marcha se ha valido de las herramientas formales propias del generativismo para representar los procesos variables (WL&H 1968). Sin embargo, como ya se señaló en (1.1), la regla variable (Labov 1969) trataba de reflejar lo que ocurre en la comunidad de habla y no en la mente de los individuos, lo que llevó a una fuerte polémica sobre su estatus teórico dentro de la gramática generativa (Cedergren & Sankoff 1974; Fasold 1991 c). Las reglas variables se constituyen por una variable x que puede tomar una forma y en contextos determinados z, w ; hasta aquí terminan las similitudes con las reglas generativistas que se encuentran, por ejemplo, en *The Sound Pattern of English*

¹¹ Aunque el español ha jugado un papel importante en el variacionismo, particularmente cuando se discute la “sobreevaluación” del funcionalismo en el cambio con los procesos de elisión de $-s$ implosiva en dialectos caribeños (Labov 1996, Cap. 19).

(Chomsky & Halle 1968). Los factores contextuales anotados en la regla variable, a diferencia de la regla fonológica tradicional, están asociados a un coeficiente probabilístico y por lo tanto no constituyen contextos categóricos de aplicación/no aplicación: más bien, la regla variable debe interpretarse como “*x* se transforma en *y* con más probabilidad en el contexto *z, w*” lo que se indica con corchetes angulares:

$$(1) \quad (x) \rightarrow \langle y \rangle / \langle z_w \rangle$$

Actualmente los resultados del análisis probabilístico multivariable (realizados con *Varbrul*, programa computacional diseñado por D. Sankoff en 1978 y con versiones actualizadas para plataformas como Macintosh y Windows) se presentan en forma de tablas y la cuestión sobre el estatus de las reglas variables dentro de la gramática se ha dejado de lado (Fasold 1991 c), aunque hay quien afirma que la regla variable sigue siendo una forma adecuada de representar los datos sociolingüísticos (como en Guy 1997).

El marco general de análisis que se implementará en los capítulos sobre variación fónica será el denominado Teoría de la Optimidad (en adelante TO)¹². Esta perspectiva de origen generativista surge a principios de los años noventa y el primer planteamiento concreto del modelo lo encontramos en el trabajo de Prince & Smolensky (1993/2004). Como otros modelos generativistas, la TO propone una

¹² La base para este breve recuento se encuentra principalmente en los textos de Kager (1999) y Legendre (2001).

estructura profunda y una estructura de superficie. Sin embargo, aquí la lengua no es un sistema de reglas ordenadas que derivan una salida (*output*) a partir de una entrada (*input*) que sufre una serie de transformaciones; más bien, lo que ocurre entre una y otra representación es una interacción entre restricciones de distintos niveles del sistema (sintáctico, morfológico, prosódico o segmental), que dará como resultado la forma más “armónica”, la que mejor respete principios universales de buena formación estructural. Una de las grandes diferencias de este modelo de gramática respecto a los modelos generativistas como Principios y Parámetros (Chomsky 1981), donde se encuentra un conjunto de Principios propios de la Gramática Universal (y que por lo tanto funcionan para todas las lenguas) y Parámetros que se aceptan o cancelan y dan por resultado las diferencias entre las lenguas particulares; en TO las diferencias entre lenguas se producen por el distinto ordenamiento de *restricciones* universales en una jerarquía. Las *restricciones* pueden ser infringidas (quebrantables), con tal de que la forma lingüística seleccionada (la presente en el *output*) cumpla con restricciones de un nivel mayor en la jerarquía de restricciones de la lengua en turno. Los dos tipos principales de restricciones están en conflicto: las restricciones de fidelidad (que persiguen preservar los contrastes léxicos del *input*) por un lado, y por otro las de marcación (condiciones de buena formación estructural). De hecho se suele señalar constantemente que la TO es una perspectiva “orientada al *output*” (*output-oriented*).

Los componentes del modelo que propone la TO son: *Input*, que incluye únicamente elementos lingüísticamente bien formados (que respeten principios muy

generales de buena formación), un Generador (Gen) que sólo produce candidatos que respeten principios generales de la teoría X-Barra para el caso de la sintaxis, o geometrías de rasgos de tipo autosegmental en el caso de la fonología; un conjunto de Restricciones universales (Con), un componente Evaluador (Eval) y finalmente el Output. Como ya se mencionó, las restricciones de Con se encuentran ordenadas en una jerarquía que permite seleccionar la mejor salida de entre un conjunto de “candidatos” proporcionado por Gen. Existen dos grandes tipos de restricciones: las de MARCACIÓN, que prohíben formas lingüísticas marcadas (por ejemplo, una sílaba sin inicio es marcada frente a una sílaba que sí tenga inicio); las de FIDELIDAD, que exigen que se preserven en la salida las formas tal como son proporcionadas en la entrada. Cada candidato es evaluado por el componente Eval, que determina el output más armónico.

Ahora, esta tesis asume la “generalización estocástica” de Guy (2007)¹³, según la cual, los modelos fonológicos deben dar cuenta tanto de los fenómenos categóricos como de los fenómenos variables:

The central observation here is that variable processes display the same patterns of occurrence and non-occurrence that are found for categorical alternations, and hence are governed by the same principles and generated by

¹³ Esta generalización está en la base de la PREMISA I del “modelo multidimensional y metarregular del cambio” de Martín Butragueño (2014): “un modelo de cambio fonológico debe respetar los requisitos formales de la teoría fonológica general” (pág. 111).

the same processes of grammar. Since alternations are what phonological theories have classically been designed to account for, we can reasonably expect that extant theories incorporate explanatory principles and generalizations about linguistic structure that are relevant to variable alternations [Guy 2007: 6].

Como se ha señalado (Holt 2006), muy tempranamente se propusieron, en el marco de la TO, varios acercamientos para explicar los datos variables. Ejemplos de estos son las “restricciones no jerarquizadas” de Zubritskaya 1997; el modelo de “restricciones flotantes” de Nagy & Reynolds 1997 (y la adecuación de Guitart 1997, para el español); el modelo probabilístico del cambio de Anttila & Cho 1998; el *algoritmo de aprendizaje gradual* de Boersma & Hayes 2001; las discusiones de Anttila 2002 sobre las posibilidades formales de la TO para explicar a variación; la *gramática armónica ruidosa* de Coetzee 2008; 2009; Coetzee & Kawahara 2013; y la reciente propuesta *multicausal y multirrelacional* de Martín Butragueño 2014, inspirada en parte en Díaz-Campos & Colina (2006), quienes combinan las restricciones no jerarquizadas con pesos probabilísticos. En todos estos modelos se ha utilizado de una u otra manera la cuantificación como argumento para proponer jerarquías de restricciones.

En particular, a los modelos que parten del algoritmo de aprendizaje gradual de Boersma & Hayes (2001) se les conoce como de *Teoría de Optimidad Estocástica* (TOE) (Díaz-Campos & Colina 2006; Martín Butragueño 2014; Feldhausen & Vanrell 2014): “traditional OT assumes grammars of individual languages to be a strictly

ranked set of violable constraints C ($C_1 \gg C_2 \gg C_3$). SOT [Stochastic Optimality Theory], in contrast, ranks the constraints along a continuous scale. As such, the distance between the constraints can vary” (Feldhausen & Vanrell 2014).

Se ha señalado (Dittmar 1996; Guy 1997, los primeros; más recientemente Martín Butragueño 2000, 2002, 2014) que estos esfuerzos por adaptar la TO a los datos de variación no parecen solucionar el problema de fondo que es cómo tratar la variación estilística, sexual y social, tanto en la comunidad como en el individuo, de manera que esté realmente incorporada en una gramática *realista* de la competencia (o del uso, según se quiera ver). Por otra parte, un cambio en la jerarquía de restricciones implica desde la perspectiva de TO (e incluso TOE) un cambio de gramática y con ello surge, por ejemplo, el siguiente problema: ¿los individuos bidialectales poseen múltiples gramáticas? Valga un ejemplo: sea el caso de un hablante de español caribeño que tiene tres opciones de pronunciación dada una /s/ final de sílaba: [s], aspiración [h] o elisión [∅], y que en una misma oración produce un ejemplo de cada una:

(2) ayer fuimo[s] con lo[h] Rodrigue[∅]

¿Cambió el hablante de gramática tres veces en la misma enunciación? Finalmente, se ha criticado la visión abrupta del cambio lingüístico, que presupone la TOE, como los modelos generativistas clásicos¹⁴.

Pese a todo, implementar la TOE para esta investigación tiene sus ventajas evidentes sobre otras propuestas. En primer lugar, ofrece un modelo unificado de gramática, donde sintaxis, fonología segmental y fonología prosódica interactúan en forma paralela, a diferencia de modelos anteriores, en que tenía que utilizarse una forma de representación diferente para fonología y sintaxis. Por otra parte, la ingeniería optimal permite representar los fenómenos de cambio como una modificación en el orden de las restricciones; esto es, ofrece un mecanismo formal de representación del cambio lingüístico. Por último, la idea de variación de alguna manera ya está incluida en el modelo, algo que sólo las *reglas variables* (Labov 1969, 1996) habían hecho explícito. Por supuesto, la gran diferencia de la TOE respecto a la versión clásica de TO es que se incorpora la probabilidad (en este caso, se utilizará el algoritmo de regresión escalonada de VARBRUL (Sankoff, Tagliamonte & Smith 2005) para determinar las jerarquías de restricciones.

Los resultados cuantitativos se presentarán en cuadros y se formalizarán siguiendo las directrices generales de TOE, con las que se busca la jerarquía de restricciones que mejor explique las diversas variantes. Otra diferencia de los cuadros

¹⁴ Visión que se viene arrastrando desde las primeras versiones generativistas del cambio (WL&H 1968). Un análisis crítico sobre TO y variación se encuentra también en los trabajos de Martín Butragueño (2002; 2014).

optimales de TOE que se presentarán en esta tesis, respecto a los análisis clásicos en TO, es que todos los candidatos son reales, no hipotéticos. Cada restricción, por lo tanto, se ordena de acuerdo a la frecuencia (y probabilidad) con que impide la aparición de cierto candidato.

Como se ha visto en esta sección, la formalización del cambio lingüístico y su estatus dentro de la teoría lingüística no es un problema resuelto, sobre todo si se asume que el interés en todo caso debería estar *en los procesos*, no en las soluciones finales, como propone la PREMISA III del modelo de Martín Butragueño (2014: 111): “la explicación del cambio fonológico será, ante todo, la explicación de los procesos de cambio”. Por otra parte, y acompañando a los problemas teóricos y antecedentes descriptivos que conlleva toda investigación, se encuentra también la cuestión de las condiciones externas a la misma: las condiciones generales en que se lleva a cabo, incluida la historia de los agentes que investigan, de los avances y limitaciones en el campo de estudio, la perspectiva general sobre lo que debe ser central o prioritario en materia de investigación lingüística, etcétera. La siguiente sección trata de ubicar precisamente a esta tesis en su contexto científico, presentando los antecedentes de investigación a la vez que plantea la necesidad de trabajos como este en el marco de la lingüística mexicana contemporánea.

1.3. ANTECEDENTES DE INVESTIGACIÓN

Como lo he mencionado en otros trabajos, en México existe la necesidad urgente de estudiar no sólo el sistema de reglas categóricas, sino también las reglas variables de

la comunidad (Serrano 2007: 96-97), a través de muestreos socialmente realistas (Serrano 2011: 210-213; 2013). En lo general, es válido afirmar que la sociolingüística ha enfrentado cierta oposición en los círculos académicos, sobre todo en México¹⁵. Mi hipótesis es que, por una parte, existe una tradición de investigación en el país que promueve fuertemente las formas estructuralistas de ejercer la disciplina (ya sea descriptivismo, generativismo o funcionalismo), esto es, que implica el trabajo con sólo uno o dos informantes y existe una preponderancia del analista en la interpretación de los datos que está por encima de la realidad misma; por otro lado, en general existe el prejuicio de que los recursos humanos, tecnológicos y financieros que exige son altísimos para un solo investigador –ahora injustificado gracias al acceso generalizado a las tecnologías de grabación digital y a computadoras cada vez más poderosas en términos de procesamiento y almacenamiento de información--. Finalmente, en general se ignora que la sociolingüística enfocada en el lenguaje constituye un acercamiento estrictamente científico al estudio de la variación y el cambio lingüístico (Serrano 2007: 99-100).

¹⁵ En otras disciplinas sociales ocurre algo similar: “En economía, como en todas las ciencias, existe una inercia que hace muy difícil romper la ortodoxia. Esta inercia se explica por factores muy diversos: sistemas de educación, ideologías, falta de rigor analítico en teorías alternativas, incentivos de publicación en revistas científicas, requisitos de contratación para profesores jóvenes, y el tradicional resquemor del ser humano al cambio” (Castañeda 2010: 3). Incluso la resistencia al cambio de perspectiva de academias (y editores) se produce en la física actual: v. el interesante recuento de João Magueijo (2007) y sus dificultades para publicar y difundir la Teoría de la velocidad variable de la luz.

Se comparte aquí el señalamiento de Martín Butragueño (2014: 19-22) de que la lingüística “secular” en general se enfrenta con varias falacias que perviven en el ámbito académico: *cualitativa, cuantitativa, tipológica, funcional, formal, del muestreo, de los nuevos mundos y de la nuclearidad*. Desde mi experiencia de investigación, sobre todo en ciertas presentaciones en congresos de lingüística general, me he enfrentado con comentarios del público en los que se presenta alguna de estas falacias como “argumentos”: 1) la del muestreo, que siempre cuestiona un muestreo levantado en el campo, con varios informantes y con criterios claros de selección, pero que no cuestiona jamás el trabajo descriptivo a partir de las intuiciones de un solo informante (el procedimiento común en la descripción de lenguas indígenas), por ejemplo; 2) la falacia formal, que toma como “explicaciones” el hecho de poner orden en los datos, ya sea a través de reglas fonológicas de tipo autosegmental, cuadros de ingeniería optimal, estructuras arborescentes en sintaxis o dibujos en gramática cognitiva; 3) la falacia de la nuclearidad, que asume que existe un núcleo ‘duro’ de la lingüística (v.g. sintaxis) que debe tener ‘prioridad teórica’; 4) la falacia tipológica, que asume como indiscutibles las propuestas de ‘universales de base tipológica’ (jerarquías de sonoridad o de animacidad, por dar un par de ejemplos), cuando en realidad se trata de *hipótesis* más o menos fundamentadas¹⁶; finalmente, 5) la falacia de los nuevos mundos: “[i]mplica aceptar que un fenómeno en principio universal es inválido

¹⁶ Cf. “The myth of language universals” de Evans & Levinson (2009) donde hacen una fuerte crítica de los universales tipológicos *à la* Greenberg (1966); v. también las críticas de Lazard (2012) al *apriorismo* en el trabajo tipológico.

cuando las condiciones son completamente diferentes (por ejemplo, la estratificación lingüística en tiempos pretéritos y actuales, los principios geolingüísticos proyectados en áreas lingüísticas coloniales, etc.)” (Martín Butragueño 2014: 22); esta falacia pide considerar con cuidado la opción de proyectar los hallazgos de los procesos con datos modernos al pasado de las lenguas¹⁷. En la misma página, el autor concluye: “Todos los argumentos falaces se resumen en un principio metodológico general: las paráfrasis no deben confundirse con análisis y menos con explicaciones”.

En buena medida, se sigue concibiendo la sociolingüística como una ‘subdisciplina’ de la lingüística cuando en realidad la sociolingüística implica una **redefinición** de la lingüística. C. Lehmann (2007), en un ensayo sobre el concepto *competencia*, sintetiza la visión tradicional de lo que puede considerarse propiamente ‘lingüístico’ --visión que, ese es el punto, pervive en México también:

Here a remark on the term **linguistic** is in order. In the historical period when structuralism, including generative grammar, had its heyday in linguistics, the meaning of this term tended indeed to be restricted to ‘concerning the formal structure of language’. When it was therefore recognized, from the nineteen seventies on, what a restricted concept of language was behind that terminology, concepts like “communicative” and “pragmatic” started to be opposed to “linguistic”. Since then, many a relevant publication (e.g. Vollmer 1982: 50) takes

¹⁷ Al respecto también v. Aitchison (2001). Creo que puede afirmarse que M. Swadesh era un defensor a ultranza del principio de historicidad al proponer los principios metodológicos de la glotocronología (Swadesh 1955).

“linguistic” to denote some restricted set of structural phenomena corresponding more or less to “grammatical” and, even worse, to “concerning declarative knowledge of grammar” (cf. the discussion in Canale & Swain 1980: 5) [...] The reduction of linguistics to structural linguistics was an error in its history that has been corrected. The predicate “linguistic” comprises everything that has to do with language, including (among other things) its structural, communicative and pragmatic aspects [Lehmann 2007: 247, n. 31; énfasis en el original].

A pesar de que la lingüística mexicana (iniciada de manera profesional en la Escuela Nacional de Antropología e Historia en 1939) ponía énfasis en la relación lengua-cultura (en el marco de proyectos antropológicos de gran alcance, que incluían la etnografía, arqueología, descripción lingüística y educación de los pueblos indígenas, como el Proyecto Tarasco –cf. López Aguilar 2011), no puede hablarse de un desarrollo de la sociolingüística como tal. A pesar de ello, una rápida revisión de la investigación lingüística en México (Barriga 2013) permitiría suponer que el rubro de “sociolingüística” es el que aporta más trabajos en los años recientes:

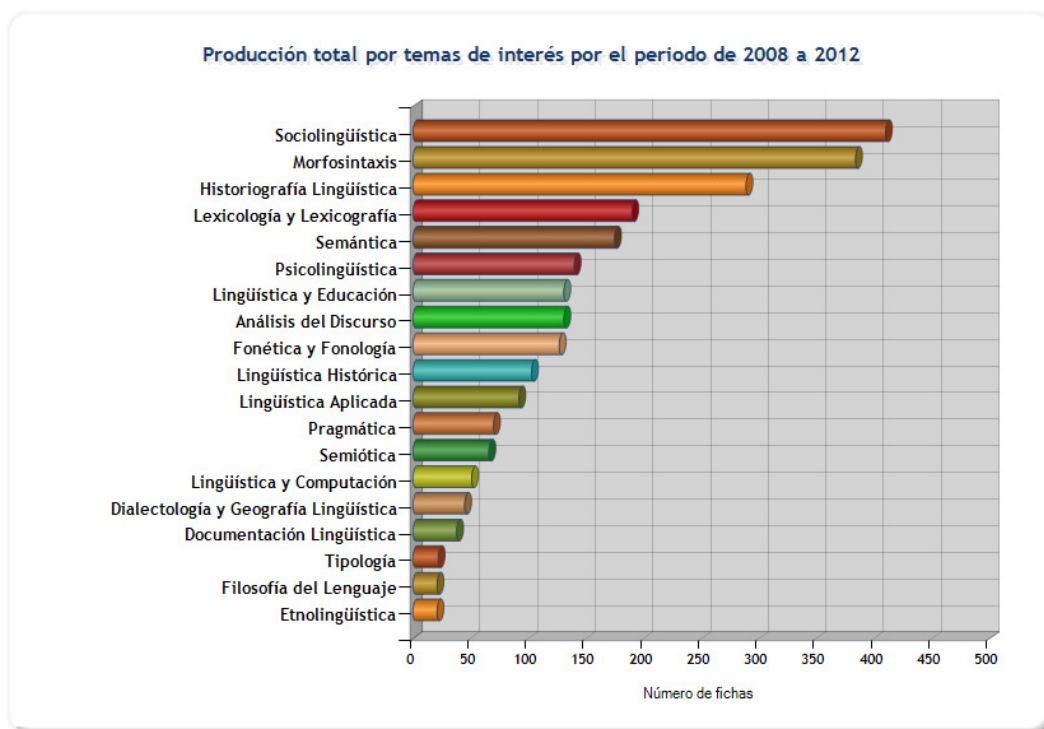


Figura 1.1. *Investigación lingüística en México por temas (2008-2012) en LingMex¹⁸, al 6 de mayo de 2013*

Sin embargo, la base de datos *LingMex* clasifica bajo el término “sociolingüística” trabajos de las siguientes sub-áreas: bilingüismo, actitudes, políticas lingüísticas y lectura y escritura (alfabetización: creación de alfabetos). Como puede verse, al menos los dos últimos dos subtipos se ubicarían más bien en el campo de la lingüística aplicada. Si se asume una idea de ‘sociolingüística’ muy amplia, que incluya los tipos propuestos por Trudgill (2000) (mencionado *supra*, n. 3), de cualquier modo,

¹⁸ V. <http://lingmex.colmex.mx/index.php/busquedas/graficas>.

relativamente pocos trabajos podrían ubicarse en esa categoría. Vale la pena mostrar algunos de los títulos marcados como ‘sociolingüísticos’ en *LingMex*¹⁹:

(3) *Independencia y educación. Cultura cívica, educación indígena y literatura infantil*

(4) *90 años de cultura. Centro de Enseñanza para Extranjeros*

(5) *Escollos y nuevos derroteros en el estudio de la literatura novohispana. De la paráfrasis a la imaginación crítica*

(6) “...Existe un universo inmenso de oportunidades...” *La educación para la interculturalidad en la enseñanza de L2 en la educación superior en México: ¿preparación para un mundo globalizado o un ejemplo de una nueva desigualdad educativa?*

(7) *Los libros de texto gratuitos en lengua indígena para la educación primaria en México*

Estos trabajos pertenecen a campos de investigación como la educación o la literatura, y sin embargo (normalmente como última opción, debe resaltarse), llevan la etiqueta de ‘sociolingüística’. Por otra parte, la calidad del trabajo sociolingüístico en México es --irónicamente— muy variable: no todos los trabajos, hay que reconocer, cuentan con el mismo nivel de rigor metodológico o analítico. La sociolingüística de la variación y el cambio en particular, como disciplina, tiene alrededor de 20 años (v. el recuento de Martín Butragueño, 2012). Por supuesto, los tres volúmenes publicados de la *Historia sociolingüística de México* (Barriga Villanueva & Martín Butragueño

¹⁹ Consultada el 7 de septiembre de 2013.

2010 a, 2010 b y 2014) permiten tener una idea más completa de la investigación actual no sólo en español, si no en varias lenguas indígenas.

Puede afirmarse que el español mexicano de enormes regiones del país es un completo desconocido, valga por ejemplo el norte: prácticamente no se conoce nada sobre las variedades de Coahuila, Chihuahua, Baja California, Durango, Zacatecas y Tamaulipas²⁰. En realidad es relativamente reciente el interés por acercamientos realistas al lenguaje. Variedades que cuentan con algunos estudios sociolingüísticos recientes son las de Sonora (Serrano 2002, 2009), Veracruz (Ceballos Domínguez 2006), Baja California (Gil Burgoin 2011) y El Ciruelo, Oaxaca (Espinosa Vázquez 2008); Existen proyectos sociolingüísticos de importancia en la ciudad de México, como el *Corpus Sociolingüístico de la Ciudad de México* (Lastra y Martín Butragueño 2011, 2012, en prensa), *El Habla de Monterrey* (Rodríguez Alfano 2005), *El Habla de Culiacán* (Mendoza 2011), levantamientos en Mérida y, muy recientemente, en Puebla, todos vinculados al proyecto panhispánico PRESEEA ('Proyecto para el Estudio Sociolingüístico del Estudio del Español de España y América') (Moreno Fernández 2003)²¹. El trabajo colegiado en sociolingüística empieza a surgir también: el Grupo de Investigación Sociolingüística, con base en el Distrito Federal, agrupa desde 2007 a investigadores de varias instituciones como El Colegio de México, Universidad Nacional Autónoma de México, Instituto Nacional de Antropología e Historia,

²⁰ Sobre la falta de estudios del español en el norte de México en general, v. Serrano (2010) y sobre el noreste en particular, Rodríguez Alfano (2014).

²¹ Próximamente se iniciarán los trabajos de PRESEEA para la ciudad de Querétaro.

Universidad Autónoma Metropolitana y Universidad Autónoma de la Ciudad de México. Dicho grupo de trabajo ha organizado la serie de coloquios sobre “Cambio y variación lingüística” (2008, 2010, 2012; actualmente prepara una cuarta edición, a celebrarse en octubre de 2014). En la primera compilación correspondiente a dichos congresos (Martín Butragueño [ed.] 2011), puede verse trabajos de variación en lenguas indígenas como náhuatl y zapoteco (M. San Giacomo, F. Arellanes & M. Chávez Peón), español en contextos de contacto dialectal (D. Pesqueira), variación y adquisición del lenguaje (C. Rojas Nieto), entre otros temas²². Los trabajos de variacionismo sociolingüístico en lenguas indígenas son todavía escasos (Guerrero, 2006; Cruz Gómez 2008, 2014; y San Giacomo 2009 son algunas excepciones) y muy recientemente la Lengua de Señas Mexicana ha sido objeto de estudio sociolingüístico en los trabajos de Cruz & Serrano (en prensa a y b).

Por otra parte, los estudios en el ámbito de la sociología del lenguaje tienen fuerte presencia en la Universidad Nacional Autónoma de México (Terborg & García Landa 2011). También es destacable que cada vez se presentan más trabajos sobre la relación lengua-sociedad en los congresos de lingüística más importantes en el país como el Encuentro Internacional de Lingüística en el Noroeste y el Congreso Nacional de Lingüística²³.

²² Se encuentra en prensa el correspondiente a la segunda edición del Coloquio (Martín Butragueño & Orozco [eds.], en prensa) y en preparación el del tercero (Guerrero & Orozco [en preparación]).

²³ V. Martín Butragueño (2012) para una revisión crítica de los estudios de variación y cambio en español mexicano. Para el autor, los clásicos problemas de la *transmisión* y la *difusión* del cambio

1.3.1. *Las posibilidades de estudiar la variación y el cambio lingüístico en español de la ciudad de México: corpus disponibles*

El español de la ciudad de México es el más y mejor estudiado de todas las variantes del español mexicano²⁴. Obviamente, los estudios sintácticos y léxicos sobre esta variante son los predominantes en la lingüística mexicana dedicada al español, pero la mayoría de estos trabajos no suele contar con una perspectiva sociolingüística o incorporar argumentos cuantitativos probabilísticos (cf. Barriga & Parodi 1996; Barriga 2013).

Respecto a su fonología, la primera descripción general es la de Charles C. Marden, *La fonología del español en la ciudad de México* (1896), traducido por el dominicano Pedro Henríquez Ureña en 1938, donde se hacen observaciones aisladas sobre algunos de los procesos que en esta tesis se abordarán, como la fricativización de (r) ante pausa (v. capítulo 3). P. Henríquez Ureña (1938), en el mismo volumen, presenta las características más importantes del dialecto del centro de México (Moreno de Alba 1994: 13-14): “timbre cerrado de las vocales, brevedad de las átonas, tensión y precisión en las consonantes [...]”. A pesar de estos trabajos, se trata aún de observaciones de carácter aislado sobre rasgos ‘tipificantes’ de este dialecto.

lingüístico son objetos de estudio prioritarios, sobre todo si tomamos en cuenta que la población mexicana ha pasado de ser mayoritariamente rural a predominantemente urbana en unas cuantas décadas (Martín Butragueño 2010).

²⁴ Para el siguiente recuento sobre el español capitalino sigo de cerca a Moreno de Alba (1994, 2013) y Martín Butragueño (2012).

Hay que esperar hasta la tesis de Joseph Matluck, *La pronunciación en el español del Valle de México* (1951), para tener una descripción detallada del componente fonológico de esta variante mexicana. En este y otros trabajos del mismo autor se plantean problemas específicos de la fonología capitalina, destacando entre otros el de la entonación circunfleja (Matluck, 1965). Juan M. Lope Blanch contribuye en los años sesenta con trabajos muy importantes, entre los que destaca “En torno a las vocales caedizas del español mexicano” (1963-1964), investigación minuciosa con datos de 100 hablantes en la que se resalta el carácter generalizado del fenómeno en todos los estratos sociales, educativos y entre hombres y mujeres; este problema se abordará en el capítulo 4 de esta tesis.

El primer trabajo propiamente sociolingüístico sobre la variación fonológica del español capitalino es la tesis doctoral de G. Perissinotto (1975). Entre los hallazgos más interesantes se encuentra el de la aparente expansión del fenómeno de asibilación de las vibrantes, que en el momento del trabajo se planteaba como un cambio lingüístico que se habría de consolidar en poco tiempo.

Después del trabajo de Perissinotto (1975) se produce un hiato de casi 20 años en el que no se producen trabajos de variación sociolingüística en el español de la ciudad de México. Las investigaciones de J. Moreno de Alba (1994) sobre la fonología del español en México basadas en el *Atlas Lingüístico de México* (Lope Blanch 1990-2000) pueden considerarse como las que marcan el regreso de los estudios interesados en la variabilidad del español mexicano. Si bien en estos trabajos trata el español de todo el país, Moreno de Alba hace observaciones muy detenidas sobre el

habla del valle de México. Por las mismas fechas, Martín Butragueño comienza a publicar varios trabajos sociolingüísticos de variación fónica, gramatical y entonativa en esta variedad (Martín Butragueño 1994, 2002, 2006, 2008, 2011, 2014). Serrano por su parte ha contribuido con algunos trabajos sociolingüísticos sobre fonética y léxico (Serrano 2006, 2008, 2010 b, 2011, en prensa).

Existen actualmente grandes corpus que permiten estudiar el español capitalino desde una perspectiva sociolingüística y que son aprovechados en esta tesis como fuentes de datos principales. El primer proyecto que documentó las hablas capitalinas fue el denominado “Norma culta de la ciudad de México”, levantado en la Universidad Nacional Autónoma de México (UNAM) entre los años 1967-1970 (cf. Lope Blanch 1971). Dicho proyecto grabó a 598 personas en el marco de 430 entrevistas, con una duración mínima de 30 min cada una. Inmediatamente después de este levantamiento (entre 1971 y 1975), le siguió la grabación de entrevistas para el proyecto “Habla popular de la ciudad de México” (Lope Blanch 1976), que buscaba contar con un corpus altamente contrastante con el de norma culta que permitiera tener acercamiento sociolingüísticos a los datos. El proyecto sobre habla popular grabó a 244 informantes en 149 entrevistas, también de un mínimo de 30 min de duración cada una. En total, ambos proyectos produjeron más de 550 horas de grabación con hablantes de 2 estratos sociales muy diferenciados en la ciudad (cf.

Serrano 2009)²⁵. Los trabajos de R. Ávila aprovecharon estos materiales para el análisis sociolingüístico del léxico (Ávila 1999 a y b).

Por otra parte, el *Corpus Sociolingüístico de la Ciudad de México* (Martín Butragueño & Lastra 2011, 2012, en prensa²⁶) cuenta con más de 300 entrevistas grabadas digitalmente y recogidas entre los años 1997-2007 que abarcan alrededor de 500 hrs. de grabación. El núcleo básico es de 108 entrevistas (36 por cada *modo de vida* principal [Højrup 1983; Lastra & Martín 2000]). Dicho proyecto forma parte del proyecto PRESEEA (mencionado *supra*), proyecto panhispánico que trata de documentar sociolingüísticamente las principales ciudades del mundo hispánico²⁷.

²⁵ Existe desde 2011 una página web con versiones de las 66 entrevistas originales en formato PDF y una muestra de audio de 5 min de cada una. En enero de 2014 se añadieron 24 nuevas transcripciones del proyecto “Norma culta” (algunas levantadas en 1964), con criterios de transcripción modernos y también con muestras de 5 min de audio: <http://www.iifilologicas.unam.mx/elhablamexico/>.

²⁶ Puede consultarse dicho corpus en: <http://lef.colmex.mx/index.php/investigaciones/corpus-sociolingueistico-de-la-ciudad-de-mexico-cscm>.

²⁷ V. <http://preseea.linguas.net/>. Un corpus más reciente, en construcción, es el coordinado por P. Martín Butragueño y L. Orozco: “El *Corpus oral del español de México* (COEM) pretende ofrecer una imagen actual, geográfica y sociolingüística, de las hablas mexicanas. Parte de la hipótesis de que los núcleos poblacionales se ordenan en diferentes subsistemas, relacionados de diversos modos, bajo la herencia de distintas circunstancias históricas y, sobre todo, en constante intercambio actual, social y lingüístico. De tal manera, la investigación descriptiva pretende convertirse en hipótesis sobre las formas de difusión (entre comunidades) y de transmisión (dentro de las comunidades)” (<http://lef.colmex.mx/index.php/investigaciones/corpus-oral-del-espanol-de-mexico>). Por otra parte, existe también el *Archivo de la palabra* del Instituto Nacional de Antropología e Historia (INAH), que

Si se suma todas las horas de grabación de los corpus de la UNAM y El Colegio de México, resulta que el español de la ciudad de México ha documentado más de 800 horas de grabación con alrededor de 1200 informantes. Si a este acervo se unen los del INAH, se contaría con un archivo masivo que permitiría el diseño de un hipotético “Corpus Histórico Oral del Español Mexicano” (CHOEM)²⁸. Dicho corpus incluiría también grabaciones de radio comercial, televisión, películas, discursos políticos, etc., para poder abarcar una serie de opciones estilísticas suficiente para la caracterización apropiada de esta variedad.

Como puede verse, se cuenta actualmente con recursos suficientes para realizar estudios de cambio lingüístico en tiempo real y a medida que pasen los años, los corpus mencionados irán adquiriendo cada vez mayor valor para la investigación, no sólo de la lengua, sino del cambio social, micro historia, antropología urbana, etc.

compila entrevistas levantadas principalmente en los años 1970 y 1980 con excombatientes y testigos presenciales de la Revolución Mexicana (1910-1928). Se ha anunciado la publicación de 328 entrevistas (<http://www.inah.gob.mx/boletines/1-acervo/4199-testimonios-de-la-revolucion>). La página web del Instituto Nacional de Antropología e Historia asegura que su Archivo de la Palabra cuenta “con casi un millón de entrevistas de personas que han formado parte de sucesos relevantes de la historia nacional”. Para acceder a este acervo hay que hacer una solicitud por escrito y realizar consultas en las instalaciones del Instituto: <http://www.inah.gob.mx/especiales/194-fonoteca-del-inah->.

²⁸ Existen esfuerzos interinstitucionales recientes por conformar una ‘Red de Archivos de Lenguas México’ (RALMEX), que permitiría enlazar los acervos sonoros digitalizados en varias lenguas mexicanas y, por supuesto, español. El *Primer Coloquio sobre Documentación y Archivos de Lenguas* llevado a cabo en la ENAH en noviembre de 2012 fue un primer producto de este trabajo conjunto.

CAPÍTULO 2

MÉTODOS¹

INTRODUCCIÓN: LA ESTRUCTURA LINGÜÍSTICA VISTA DESDE EL VARIACIONISMO

En el capítulo anterior se discutieron los presupuestos teóricos del variacionismo: la lengua tiene una estructura conformada por reglas de tipo categórico pero cuenta también con un componente de reglas sociolingüísticas (Dittmar 1996) que están guiadas por múltiples factores, tanto lingüísticos como extralingüísticos. Como lo señalan Milroy & Gordon (2003: 1-8), la sociolingüística, y en especial Labov (1963, 1966), vinieron a criticar el llamado *axioma de categoricidad* que está en la base de

¹ Este capítulo trata sobre el planteamiento metodológico de la tesis *en general*. Dado que los capítulos subsecuentes tratan sobre procesos muy particulares y en distintos niveles de lengua, no es posible mencionar todas las decisiones analíticas de detalle. Por supuesto, el lector encontrará una justificación de las mismas en cada capítulo.

toda la lingüística estructuralista (saussuriana, bloomfieldiana y chomskiana) de la primera mitad del siglo XX --y que pervive en la actualidad en muchos de los trabajos de “lingüística teórica”. Según el axioma de categoricidad, la lengua tiene una estructura invariante, compartida por toda la comunidad de habla; las variaciones individuales son, desde dicha perspectiva, periféricas para la teoría lingüística. La sociolingüística no solo cuestiona la naturaleza de la estructura lingüística, sino que también presenta un acercamiento distinto al objeto lingüístico, lo que tiene serias consecuencias metodológicas: se trabajará ahora con datos de producción de habla espontánea como dato primario, mientras que la elicitación y la introspección se convierten ahora en métodos secundarios.

En general, puede afirmarse que los métodos de investigación desarrollados para el estudio empírico de la variación y el cambio lingüístico en marcha han enriquecido los procesos de investigación sociolingüística en tres grandes aspectos: a) respecto a la obtención del habla más natural y diversificada posible en el seno de la comunidad lingüística de análisis --lo que unificaba a los pioneros del estudio del lenguaje “in its natural setting”, como W. Labov, o los acercamientos más etnometodológicos de J. Gumperz (v.g. Gumperz & Wilson 2000), entre otros--; b) el refinamiento del análisis cuantitativo de los datos de producción y c) recursos analíticos como el *tiempo aparente* y el *tiempo real* para el estudio de la difusión de los cambios lingüísticos (Labov 1996). Este capítulo por tanto abordará estos tres aspectos metodológicos: primero se discute el valor de la entrevista como fuente de datos principales y las características de los corpus base de esta tesis; en segundo

lugar se detallan los procedimientos de análisis cuantitativo, aspecto fundamental del acercamiento empírico a la variación y el cambio lingüístico; finalmente, se presentan las ventajas de un análisis combinado del tiempo aparente y el tiempo real para el estudio del cambio en español de la ciudad de México.

2.1. EN BUSCA DEL HABLA ESPONTÁNEA: LA ENTREVISTA COMO FUENTE DE DATOS SOCIOLINGÜÍSTICOS

Desde su concepción, esta tesis buscaba aprovechar los materiales grabados de conversaciones de los años sesenta y setenta para rastrear ciertos procesos variables en el tiempo real. Aunque existen otras fuentes de habla de la misma época que pueden ser útiles (entrevistas en radio, películas, telenovelas, programas de variedad, etc.), no dejan de ser formas de habla dramatizadas que pueden reflejar parte de una realidad sociolingüística, sí, pero que difícilmente escapan al riesgo de que el actor o locutor introduzca rasgos idiosincráticos en la interpretación del papel o, en todo caso, pueden reflejar tipos de habla muy formales en el caso de la radio².

Esta tesis se basa casi en su totalidad en habla grabada en entrevistas --el dato más natural que se pudo obtener de los años 1960-1970 en los archivos del Instituto

² Aunque la audición no sistemática de ciertas producciones televisivas o películas recientes mexicanas (como el programa *Big Brother* o la película *Amores perros*, de Alejandro González Iñárritu [2000]) parecen mostrar una creciente actitud de los productores por mostrar un lenguaje más “auténtico”. Por supuesto, esto tendría que estudiarse con detenimiento y al respecto se encuentran proyectos en marcha, como los de R. Ávila (2011).

de Investigaciones Filológicas. La muestra de esos años se replica con los materiales del *Corpus sociolingüístico de la ciudad de México* (Martín Butragueño & Lastra 2011, 2012, en prensa). Como se sabe, la intención del sociolingüista es estudiar el lenguaje “natural”, pero en la práctica esto es obtenible sólo hasta cierto punto, ya que es imposible salvar la *paradoja del observador* (Labov 1972), esto es, se quiere estudiar el lenguaje en su contexto natural, capturar el *vernacular*, que es la forma de habla que los informantes utilizan *cuando no están siendo observados* (entiéndase en este caso, *grabados*), sin embargo, no puede haber observación sin un observador quien, forzosamente, sesgará el dato.

La grabación secreta, por otra parte, es un recurso que ciertamente puede lograr la mayor espontaneidad por parte del informante, pero presenta al menos tres problemas de importancia: 1) no se salva por completo la paradoja del observador, ya que el entrevistador, al saberse grabado, quizás quiera que haya más intervenciones por parte del informante (lo que provocará una situación de diálogo menos natural), o puede medir sus propios comentarios al saber que terceras personas podrían escucharlo; 2) normalmente, al estar escondida la grabadora, se obtienen grabaciones de muy mala calidad; 3) no menos importante es el problema ético de grabar a una persona sin su conocimiento; por supuesto, el procedimiento normal tras grabar en secreto a un informante es comunicarle de inmediato que se realizó una grabación y preguntarle si quiere que sean utilizados estos datos; si el informante se niega, en ese momento debe borrarse la grabación (cf. Schilling-Estes 2007). En esta tesis sólo dos

de los 48 informantes fueron grabados en secreto (ambos en el corpus de los años sesenta-setenta).

La entrevista en presencia de una grabadora como medio de obtener datos en lingüística ha resultado un método adecuado para el tipo de variabilidad que se analiza en sociolingüística, que muchas de las veces se encuentra por debajo del nivel de conciencia del informante; nos referimos a los procesos fonéticos de detalle, sobre todo cuando no tienen una fuerte presencia en el imaginario lingüístico comunitario – como puede ser el caso de las vocales debilitadas estudiadas en el capítulo 4 de esta tesis.

El valor de la entrevista como recurso metodológico se discute a profundidad en un volumen monográfico de la revista *Language in Society* (2011, 1). En este número se compila varias investigaciones donde se estudian las *narraciones conversacionales* –aquellas construidas en el marco de la entrevista y distintas a las narraciones e historias tradicionales que suelen elicitar en otras investigaciones. De hecho, la introducción a dicho volumen se titula: “Interviews vs. ‘natural’ contexts: A false dilemma”. Las autoras (De Fina & Perrino) recuperan una conclusión interesante en el trabajo de De Fina, en el mismo volumen, con italo-australianos: “Like other authors in this special issue, De Fina notes that the effect of the Observer’s Paradox, that is, the highlighting of the presence of the researcher in the data collection event, is not always a problem, but may become a resource for the analyst” (2011: 8; subrayado mío).

Sólo por mencionar una de las investigaciones en dicho volumen, en el trabajo de M. Koven (2011) se compara la misma narración por parte de una joven, primero contada en el marco de una entrevista y después con amigos sin la presencia de la investigadora y, como es de esperarse, con un grado mayor de informalidad. Koven descubre que “[o]verall, these materials demonstrate that interview stories may be as interactionally complex and amenable to interactional analysis as conversational stories” (2011: 87). De hecho, la entrevista en sociedades industrializadas puede verse como un *acto de habla* en sus propios términos (Briggs 1986). Recientemente, en *Sociolingüística cognitiva*, Moreno Fernández (2012) hace una defensa muy explícita de la entrevista sociolingüística como la herramienta más adecuada para obtener datos de lengua en uso. En su Proposición 8.16.3 señala que “la entrevista es la técnica básica para la recolección de materiales de la lengua hablada” (2012: 174).

La audición directa de las entrevistas seleccionadas para esta tesis proporcionó abundantes ejemplos de los procesos fonéticos investigados -- fricativización de vibrantes (cap. 3) y debilitamiento de vocales átonas (cap. 4). La transcripción por escrito de las entrevistas, por su parte, proporcionó la base para el análisis de grupos consonánticos cultos, los morfemas *-mente* y *-dad*, para el estudio del léxico sustantivo global y de ciertas piezas léxicas emblemáticas (*onda*, *chido*, *padre*, entre otras), todos estos procesos son revisados en el cap. 5. El no contar con una variedad de *estilos contextuales* (Labov 1972) ciertamente es una limitante de la investigación; sin embargo, las condiciones de levantamiento de las encuestas fueron muy similares: estudiantes de licenciatura fueron los principales entrevistadores y

generalmente se partió de las redes sociales de ellos mismos para contactar a los informantes (cf. Aguirre y Chico 2011; Serrano 2009). Esto se verá con detalle en la siguiente sección.

2.2. LOS CORPUS UTILIZADOS

Las posibilidades de hacer un estudio en tiempo real sobre el español mexicano han aumentado con la distancia temporal que adquieren los corpus levantados en proyectos históricos y lingüísticos de los años 60, 70 y 80. Ejemplos de acervos importantes son el proyecto de “Historia Oral de la Revolución Mexicana” del Instituto Nacional de Antropología e Historia, en que se entrevistaron principalmente a participantes en las batallas revolucionarias de inicios del siglo xx. Dicho archivo cuenta al momento con más de 300 entrevistas digitalizadas³. El Archivo de la Palabra del Instituto de Investigaciones José Ma. Luis Mora cuenta también con entrevistas grabadas en la ciudad de México⁴.

Esta investigación aprovecha los materiales de tres importantes acervos sonoros: “Norma lingüística culta de la ciudad de México” (Lope Blanch 1971) y “Habla popular de la ciudad de México” (Lope Blanch 1976), del Instituto de

³ Puede consultarse el catálogo de publicaciones surgidas de estos materiales en: http://www.bibliotecas.tv/zapata/bibliografia/indices/catalogo_del_archivo_de_la_palabra.html.

⁴ El acceso a este archivo se encuentra mucho más restringido y forma parte del proyecto de investigación de “Historia oral” en el Instituto Mora (<http://www.mora.edu.mx/Investigacion/SitePages/Campos%20Investigaci%C3%B3n.aspx>).

Investigaciones Filológicas⁵ y las grabaciones del *Corpus Sociolingüístico de la Ciudad de México* de El Colegio de México (Lastra & Martín Butragueño 2000; Martín Butragueño & Lastra 2011, 2012, en prensa)⁶.

2.2.1. Breve historia de los corpus utilizados

En 1964, Juan M. Lope Blanch propuso durante el *Segundo Simposio del Programa Interamericano de Lingüística y Enseñanza de Idiomas (PILEI)* el *Proyecto de Estudio Coordinado del Habla Culta de las Principales Ciudades de Hispanoamérica*. Ese mismo año se inician las grabaciones de prueba de la sección mexicana y para 1970 se contaba con una muestra de más de 400 hablantes. En 1971 iniciaron las entrevistas para un segundo proyecto: *Habla Popular de la Ciudad de México*, que recogió poco más de 220 entrevistas⁷. Los proyectos “Norma culta” (NC) y “Habla popular de la ciudad de México” (HP), produjeron más de 635 entrevistas con poco más de 800 informantes pertenecientes a grupos socioeconómicos altamente diferenciados. El solo proyecto “Norma culta” está compuesto por 411 grabaciones totales realizadas en su mayoría entre los años 1967 y 1970, las cuales contienen 518 entrevistas en las que participaron un total de 592 informantes. Por otra parte, en el proyecto “Habla popular” se realizaron 149 grabaciones de una hora de duración que registran el habla de 244 informantes en 224 entrevistas. Todos los materiales de estos dos proyectos

⁵ Agradezco a Gloria Báez Pinal y a Cecilia Rojas Nieto su valioso apoyo para acceder a estos materiales.

⁶ Agradezco a Pedro Martín Butragueño el acceso a estos materiales y sus transcripciones.

⁷ Cf. Lope Blanch (1976).

fueron grabados en cinta magnetofónica y se utilizó el mejor equipo de grabación de audio en campo disponible en su momento (generalmente eran grabadoras y micrófonos de la marca alemana *Uher*). Los materiales obtenidos en dichos proyectos fueron aprovechados en los años 70 y parte de los 80 para producir un volumen importante de tesis, artículos e investigaciones y se han incorporaron a corpus más voluminosos, como el macro-corpus de ALFAL (Samper Padilla *et al.* 1998)⁸. Comparado este volumen de informantes y de horas de grabación con los producidos en proyectos panhispánicos como PRESEEA⁹ (con presencia en 39 ciudades), en los cuales se exige una cuota mínima de 108 informantes en las ciudades más grandes, el corpus del Centro de Lingüística Hispánica “Juan M. Lope Blanch” (CLH) puede considerarse todavía el más grande del mundo en cuanto al número de hablantes de español de una sola variedad¹⁰. El “Proyecto de recuperación de los archivos sonoros del CLH” (Serrano 2009), iniciado a mediados de 2007, tuvo como objetivo principal la digitalización de todos estos materiales (originalmente grabados en cintas de carrete abierto de 1/4 de pulgada), siguiendo parámetros estandarizados de documentación lingüística (cf. Austin 2007): muestreo de señal a 44.1 Khz, codificación PCM a 16 bits,

⁸ “Asociación de Lingüística y Filología de América Latina”. Página web: <http://www.mundoalfal.org/>.

⁹ Página web: www.linguas.net/portalpreseea.

¹⁰ El proyecto *El habla de Monterrey* (Rodríguez 2005) se acerca bastante en términos de volumen. Entre 1985 y 1986 se levantaron 600 encuestas para documentar el habla de dicha ciudad industrial del noreste mexicano. Además, añadieron en años recientes (2006-2010) un total de 177 nuevas entrevistas, 100 de ellas con los mismos informantes del primer muestreo, con lo que pueden hacerse estudios de panel en tiempo real (v.: <http://www.hablademonterrey.org/>).

en formato *wav*. Dicho proceso ya concluyó y produjo más de 550 horas de grabación digital y dos bases de datos de informantes que concentran 816 fichas (cf. Aguirre & Chico 2011). Las grabaciones presentan muy buena calidad para la transcripción fonética impresionística y muchas de ellas permiten análisis acústicos detallados. La digitalización de estos materiales en la UNAM permitió disponer de una muestra de 24 informantes: 12 de norma culta y 12 de habla popular. Al tratarse de dos grupos socioeducativos extremos, y que los grupos medios no estaban considerados en el levantamiento, se decidió que dos informantes en cada cruce de variables (sexo, grupo de edad y nivel de estudios) podía ser un número adecuado para dar cuenta de los procesos fónicos que originalmente inspiraron esta investigación. Por otra parte, trabajos previos con la mitad de estos informantes (Serrano 2008), o incluso menos (Serrano 2006), arrojaron resultados de probabilidad dentro del rango de 0.05 de significatividad, lo que sugería que la muestra, ya duplicada, podía ser suficiente para entender los fenómenos analizados.

El segundo corpus, que replica el de los años 60-70, lo constituyen 24 entrevistas del *Corpus Sociolingüístico de la Ciudad de México* (en adelante CSCM) recogidas en su mayoría entre 1997 y 2007 como parte del proyecto “Cambio y variación lingüística en la ciudad de México” dirigido por Pedro Martín y Yolanda Lastra en El Colegio de México. Este proyecto está vinculado a PRESEEA, *Proyecto para el Estudio Sociolingüístico del Español de España y América*, coordinado por Francisco Moreno Fernández desde 1996 (cf. Moreno Fernández 2003). Para esta investigación de tesis, las 24 entrevistas del CSCM se dividen en dos subgrupos de 12 informantes

que se denominarán “Alto” y “Bajo” en referencia al nivel de estudios de los informantes de cada grupo (bajo: estudios secundarios o menos; alto: estudios universitarios) y que tratan de equipararse con los correspondientes grupos “culto” y de “habla popular” del corpus de la UNAM. A diferencia de las entrevistas en los años 1970, el CSCM cuenta con entrevistas sociolingüísticas que duran alrededor de dos hrs., con un mínimo de 45 min. de conversación grabada cada una¹¹. Desde un principio dicho proyecto planteó la calidad acústica de las grabaciones como prioridad y por tanto se utilizó equipo digital que permitiera la investigación acústica a través de recursos computacionales como el *Speech Analyzer* de Kay Elemetrics. Por lo tanto, las entrevistas del CSCM fueron grabadas en su totalidad utilizando equipo de audio digital. La gran mayoría de las entrevistas se grabaron con equipo y micrófonos profesionales en formato DAT (*Digital Audio Tape*) y a partir del segundo año se empezaron a utilizar también grabadoras de formato *MiniDisc*. La calidad de estas grabaciones es excelente tanto para la transcripción impresionística como para el análisis acústico. Además de la página web oficial del proyecto¹², donde se han subido las transcripciones de las 108 entrevistas para el corpus PRESEEA (36 por cada nivel de instrucción: bajo, medio y alto), existen dos publicaciones en formato CD de las 36 entrevistas del grupo de alta instrucción y 36 del grupo de instrucción media;

¹¹ A lo largo de la tesis se hablará del primer muestreo como “datos de 1970” y los del muestreo moderno como datos “del año 2000” sólo como recurso expositivo.

¹² Dirección web: <http://lef.colmex.mx/index.php/investigaciones/corpus-sociolingueistico-de-la-ciudad-de-mexico-cscm>.

actualmente se encuentra en preparación la tercera publicación con las 36 entrevistas de los grupos de bajos estudios; todas las publicaciones incluyen versiones electrónicas en formato *.txt* y *.doc* de cada entrevista. Otras 200 entrevistas forman parte del CSCM, y todas juntas rebasan las 500 hrs. de grabación.

2.2.2. Características sociales de los informantes en 1970 y 2000

En cada muestra de 24 informantes se distribuyen los informantes según tres factores: *sexo, grupo de edad y nivel socioeducativo*, con 2 informantes en cada casilla. En estos cuadros se esquematiza la distribución de los informantes de acuerdo con estas variables. Se indica junto al nombre del informante su año de nacimiento.

Cuadros 2.1 y 2.2. Informantes de la muestra de los años 1960-1970

Cuadro 2.1. Nivel sociocultural bajo

	<i>Jóvenes (18-34)</i>	<i>Adultos (35-54)</i>	<i>Mayores (55+)</i>
<i>Hombres</i>	Encarnación, 1941 Roberto, 1948	Agustín, 1926 Isidro, 1936	Francisco, 1914 Salomón, 1900
<i>Mujeres</i>	Ana, 1951 Rosa, 1949	Guadalupe, 1935 Agripina, 1925	Genoveva, 1906 Guadalupe, 1915

Cuadro 2.2. Nivel sociocultural alto

	<i>Jóvenes (18-34)</i>	<i>Adultos (35-54)</i>	<i>Mayores (55+)</i>
<i>Hombres</i>	Marco, 1938 Raúl, 1942	Héctor, 1925 Raúl, 1928	Juan, 1912 Luis, 1899
<i>Mujeres</i>	Dolores, 1945 Sofía, 1939	Martha, 1934 Celia, 1935	Blanca, 1887 Ana, 1912

Los hablantes del grupo Norma Culta (NC) se distinguen por un alto nivel de estudios, pertenecen a una segunda generación de hablantes con nivel educativo y socioeconómico elevado, suelen hablar uno o dos idiomas extranjeros y tener acceso a experiencias culturales diversas en México y el mundo. Por otra parte, los 12 informantes del proyecto Habla Popular (HP), en su mayoría tienen estudios de primaria inconclusa, e incluso algunos informantes son analfabetos.

Es interesante notar que los grupos de edad de esta muestra coinciden con tres períodos importantes en la historia social y económica de México: los mayores nacieron entre 1887-1915, período que se inserta en la llamada “época porfirista” o “prerrevolucionaria”; las personas de mediana edad nacieron entre 1925-1936, precisamente en la época posrevolucionaria; mientras que los hablantes jóvenes, que nacieron entre 1939-1951, se ubican en el período de la modernidad, cuando México es gobernado por presidentes civiles y ya no existen conflictos armados de importancia. Asimismo, la profundidad temporal que tenemos es importante. Por ejemplo, Blanca, quien tenía 80 años en 1967, representa una muestra de habla que se consolidó hace más de cien años (si tomamos en cuenta –como ya se mencionó-- que tras la adolescencia se consolida el “vernacular” de los hablantes –cf. Bailey 2002: 324). Los siguientes cuadros muestran los nombres y año de nacimiento de los informantes de la muestra moderna, obtenida del CSCM.

Cuadros 2.3 y 2.4. Informantes de la muestra de los años 1997-2003

Cuadro 2.3. Nivel sociocultural bajo

	<i>Jóvenes (18-34)</i>	<i>Adultos (35-54)</i>	<i>Mayores (55+)</i>
<i>Hombres</i>	Guillermo, 1980 Pedro, 1981	Faustino, 1957 Rafael, 1956	Delfino, 1927 Felipe, 1946
<i>Mujeres</i>	Carmen, 1969 Sofía, 1979	Cristobalina, 1962 Irene, 1953	Berta, 1949 Ángela, 1910

Cuadro 2.4. Nivel sociocultural alto

	<i>Jóvenes (18-34)</i>	<i>Adultos (35-54)</i>	<i>Mayores (55+)</i>
<i>Hombres</i>	Antonio, 1972 Fernando, 1978	Eduardo, 1965 Javier, 1960	José Luis, 1940 Rodolfo, 1941
<i>Mujeres</i>	Eugenia, 1972 Maripaz, 1973	Edith, 1963 Ruth, 1967	Justine, 1936 Liliana, 1941

Es muy pertinente resaltar que, si bien los proyectos dirigidos por Lope Blanch en los años 60 y 70 tenían el propósito de documentar dos registros sociolingüísticos muy diferenciados (habla culta y popular), y que el CSCM buscaba un muestreo socialmente realista de la ciudad de México y su zona metropolitana (con grupos educativos altos, medios y bajos, lo que los podría volver estos corpus “poco comparables”), los materiales que se escogieron para esta tesis en realidad son una *submuestra* de estos corpus, submuestras que no son necesariamente “representativas” de dichos acervos: lo que se buscó en realidad fueron *muestreos pareados*. Esto es, se construyó un submuestreo de 48 informantes de características sociales lo más semejantes posibles para asegurar un grado de comparación

aceptable. Por otra parte, hay que considerar también que, si se trataba de un estudio de variación y cambio en tiempo real, los materiales del CLH eran (y siguen siendo) los únicos disponibles levantados con el suficiente rigor metodológico y con fines estrictamente lingüísticos que permitieran llevarlo a cabo.

En conclusión: sin bien el corpus de la UNAM de los años 1970 no es comparable con el CSCM del año 2000, las submuestras obtenidas para esta investigación sí lo son. El Apéndice al final de la tesis muestra las características sociales de estos 48 informantes con mayor detalle, lo que permite evaluar dicha comparabilidad. Se volverá a esta discusión sobre el muestreo en el capítulo 6.

2.3. EVALUACIÓN DE LA CALIDAD DE LOS DATOS Y PROCESOS DE CUANTIFICACIÓN EN ESTA TESIS

Silva-Corvalán (2001) y Schilling-Estes (2007) presentan reflexiones interesantes sobre la diversidad de problemas que el investigador se enfrenta en el campo, con los hablantes reales. Particularmente Schilling-Estes (2007: 165) destaca las siguientes preguntas:

(1) How do I decide who to get data from? (2) How do I get people to talk to me and let me record them, and can I get them to talk in a fairly relaxed way? (3) How do I explain to people why I'm recording them, and how much detail should I go into? (4) How do I make good-quality recordings, especially out in the "field" vs. in a quiet laboratory? and (5) How involved should I get with my research participants and in what ways?

Las preguntas 1, 2 y 4 sobre la muestra, la obtención de habla relajada y las características de una buena grabación pueden responderse con la información que tenemos sobre los corpus a los que se tuvo acceso. Las preguntas 3 y 5, sobre la cantidad de información que debe proporcionarse al entrevistado acerca de los intereses de la investigación y el compromiso ético con ellos son problemas que sólo parcialmente pueden abordarse aquí. En general, el muestreo para este estudio trata de tener un nivel aceptable de representación cualitativa de los extremos socioeconómicos en la ciudad de México, sin embargo y como ya se mencionó, la ausencia de hablantes de grupos medios constituye una limitante para explorar sobre todo a los líderes del cambio que, como se vio en el Cap. 1, suelen pertenecer a los grupos centrales de la sociedad (Labov 2006).

En general, las entrevistas utilizadas en esta tesis logran un alto grado de naturalidad, sobre todo en los grupos de más altos estudios. Un factor que influye en la espontaneidad (relativa) de los hablantes es que normalmente el entrevistador ya conocía al (o la) informante, por lo que se inicia la entrevista desde una base de confianza mutua. Además, varias entrevistas se realizaron con dos o más informantes, lo que añade aún más naturalidad a las conversaciones.

Otro aspecto importante es la calidad de las grabaciones. Como ya se mencionó, el equipo de grabación de los años 1960-1970 era de la más alta calidad (cintas de carrete abierto, marca Uher, micrófonos dinámicos de la misma marca); las grabaciones del CSCM se hicieron todas con tecnología digital *DAT* o *MiniDisc* y con micrófonos *lavalier* de condensador (*Electrect*).

Vale la pena mencionar que la utilización de los materiales de los proyectos NC y HP ha sido defendida anteriormente por Raúl Ávila, en el marco de sus investigaciones sobre “semántica social” y las diferencias de léxico entre los grupos cultos y de habla popular (Ávila 1988, 1999 a y b). Ávila apela al relativo nivel de inconsciencia lingüística de los hablantes entrevistados como un argumento a favor de la comparabilidad de este tipo de muestras:

A diferencia de otro tipo de investigaciones o encuestas, el análisis del léxico y de otros componentes del lenguaje mediante grabaciones tiene la ventaja de que el informante, aunque quisiera, difícilmente podría reaccionar y cambiar su conducta lingüística frente al investigador [...] No se da esa reacción precisamente por el nivel de inconsciencia que tienen los hablantes respecto al sistema de la lengua. Consecuentemente, este tipo de datos lingüísticos resulta altamente confiable para la caracterización de los sujetos investigados [Ávila 1998: 146-147].

En la nota 30, el mismo autor acude a Halliday (1978) para consolidar el razonamiento:

se podría argumentar que es posible que el entrevistado cambie de registro ante el entrevistador. Sin embargo [...] *las grabaciones se realizaron en una misma situación comunicativa. Esta situación podría, precisamente, condicionar el registro, y no al contrario.* De acuerdo con los planteamientos de Halliday

[1978: 31 y ss.], el registro es una forma de predicción. Si se conocen los factores que intervienen en la comunicación y el escenario en que éste ocurre, “we can predict a great deal about the language that will occur, with reasonably probability of being right” [Ávila 1998: 146-147; subrayado mío].

Pareciera razonable pensar que, en efecto, el relativo nivel de inconsciencia lingüística de los entrevistados, la gran similitud en las condiciones de entrevista y el hecho de que debe ser la situación la que condiciona el registro, y no al contrario, confieren un aceptable grado de confiabilidad a los datos aquí presentados. Por supuesto, existe la limitación general de los datos de no contar con una diversidad de registros sociolingüísticos o *estilos contextuales* (Labov 1972), sin embargo, creo que no es arriesgado afirmar que las evidentes similitudes en las condiciones de levantamiento otorgan un aceptable margen de confiabilidad en los resultados, al menos en un sentido estrictamente cualitativo¹³.

¹³ En forma de comunicaciones personales, tanto en foros académicos como en contextos más informales, se ha cuestionado la comparabilidad de las muestras en el tiempo real aquí utilizadas. El argumento base que se esgrime es que –parafraseando a una colega– “las entrevistas en 1970 implicaban un alto grado de formalidad; ahora, la gente está más acostumbrada a que la entrevisten y por ello está más relajada; esto explica por qué [por ejemplo] se encuentra un menor acervo léxico en 2000 que en 1970”. Obviamente esta idea --que recuerda la noción de *diseño de la audiencia* de A. Bell (1984), que se discutirá más adelante --, se trata de una especulación interesante que, sin embargo, es difícil de comprobar. En todo caso, no tengo conocimiento de un estudio histórico o de psicología social que permita demostrar que los habitantes de la ciudad de México están “más acostumbrados” a ser entrevistados. De hecho, la probabilidad de que una persona en México sea entrevistada en algún

Por último, valdría la pena mencionar que el volumen total del corpus (casi 250,000 palabras), no es tan pequeño si se considera que es la mitad de las utilizadas para el primer diccionario estadístico del español --el de Juilland & Chang Rodríguez (1964), que constaba de 500,000 *tokens* que produjeron una base de 5,000 vocablos para dicho diccionario.

2.3.1. *Criterios de transcripción fonética y léxica*

Se decidió realizar una transcripción auditiva de las variables fónicas (róticas /r/ y /r/ y vocales átonas). Aunque debe reconocerse las limitaciones de esta forma de transcribir, se buscó siempre que las variables transcritas presentaran parámetros acústicos y articulatorios que permitieran discernirlas impresionísticamente con un alto grado de precisión. Se transcribieron todas las ocurrencias de cada variable en 30 min. de entrevista.

momento de su vida para un estudio lingüístico o sociolingüístico sigue siendo ínfima: en los años 1970 se entrevistó a 800 personas en los proyectos norma culta y habla popular (Serrano 2009); en el muestreo sociolingüístico más reciente de Lastra & Martín Butragueño (2011, 2012, en prensa) se entrevistó poco más de 300 personas en una ciudad que aumentó su población de 6.9 millones de habitantes en 1970 a 8.6 en el año 2000 (INEGI 2012); basándonos en ambas experiencias de investigación, resulta que la probabilidad de ser entrevistado en el año 2000 es más baja que en 1970. Por otra parte, es imposible demostrar que, en efecto, la entrevista grabada con una persona en el año 2000 resulte ser una experiencia más “relajada” que en 1970 y que, además, propicie el uso de un lenguaje más informal por parte del entrevistado. Esto se discutirá a profundidad en el Cap. 6.

Para el estudio del léxico se utilizaron las transcripciones ya existentes para cada grupo de materiales. Los criterios de transcripción difieren según la época del muestreo. Los materiales de los años 1970 se transcribieron siguiendo normas ortográficas convencionales. Los materiales del CSCM (años 2000) se transcribieron siguiendo las normas generales del proyecto PRESEEA que tratan de ser más fieles a la realidad fónica y pragmática de la interacción entre informante y entrevistador. Las diferencias de transcripción se discutirán con mayor detalle en el capítulo 5.

2.3.2. Métodos cuantitativos y programas utilizados

La estadística inferencial es una herramienta esencial de la sociolingüística. Dado que la aproximación sociolingüística a la variación es de tipo “realista” (cf. Figueroa 1994), se requiere obtener un gran número de datos que idealmente puedan proyectarse a la comunidad de habla en su totalidad (cf. Moreno Fernández 1990). En este sentido, los análisis multivariantes de regresión logística son una pieza fundamental. Específicamente el algoritmo VARBRUL, diseñado por D. Sankoff en 1978 (Sankoff 1988), que se ha implementado en varias versiones en las plataformas Macintosh y Windows con el nombre de *GoldVarb* (Sankoff, Tagliamonte & Smith 2005)¹⁴, sigue siendo el más utilizado en los estudios de variación, además de ANOVA (análisis de varianza, cf. Woods, Fletcher & Hughes 1986). Aunque existen desarrollos recientes

¹⁴ Página web con la versión más reciente (*Goldvarb Lion*):
<http://individual.utoronto.ca/tagliamonte/goldvarb.htm>.

del algoritmo VARBRUL en la plataforma de código abierto *R*¹⁵, como *R-Brul*¹⁶ de D. E. Johnson (2009) o *R-Varb*¹⁷ de J. Paolillo (2004), la familiaridad de la comunidad académica con *GoldVarb* hace que siga siendo a la fecha el más utilizado y citado en revistas importantes como *Language Variation & Change*, o que existan monografías y manuales que le dedican muchísimo espacio al programa (v.g., Tagliamonte 2006, Paolillo 2004). *GoldVarb X*, además del análisis multivariable, también realiza la prueba de *hipótesis nula* χ^2 (ji-cuadrada) que permite medir el ajuste del modelo estadístico con los datos, y tabulaciones cruzadas. La decisión de utilizar VARBRUL es que el tipo de cálculo que realiza se basa en la hipótesis de que los cambios lingüísticos trazan una curva estadística en forma de “s” (Kroch 1989), patrón común a varios procesos de cambio social y sugerido desde el ensayo de WL&H (1968): esto es, el cambio lingüístico inicia de manera muy lenta, normalmente en los grupos más

¹⁵ Página web: <http://cran.r-project.org/>.

¹⁶ Página web: <http://www.danielezrajohanson.com/rbrul.html>. En particular el algoritmo de Johnson (2009) es interesante porque aprovecha las capacidades gráficas de *R*, permite la inclusión de variables continuas y ofrece además un ‘modelo mixto’ en el que los informantes pueden introducirse como variables independientes. Sin embargo, la diferencia sustancial entre *GoldVarb* y *R-Brul* es que el primero produce más errores del tipo I (falsos positivos); en cambio, *R-Brul* ofrece resultados más conservadores, esto es, produce más errores del tipo II (falsos negativos). Con ciertos criterios cuantitativos, sin embargo, es posible escoger modelos conservadores también en *GoldVarb*, por lo que se decidió mantener los análisis con este programa a lo largo de la tesis y generar las gráficas con apoyo de otros programas.

¹⁷ *R-Varb* al parecer tuvo muchos problemas de implementación desde el inicio (cf. Johnson 2009). No encontré una página actualizada sobre el programa.

bajos en la escala socioeconómica, después pasa a una etapa de aceleramiento pronunciado –patrocinado por los grupos medios (o insertos en un proceso de ascenso social)--, y finalmente la etapa de consolidación vuelve a darse de manera lenta¹⁸. Por otra parte, el análisis de la variación léxica --otro punto tratado en la tesis--, se realizó con ayuda del programa gratuito *AntConc*¹⁹, de L. Anthony (2014). Dicho programa permite la creación de listas, análisis de agrupamientos (*clusters*) y colocaciones, entre otras tareas.

2.4. EL ANÁLISIS EN TIEMPO APARENTE Y TIEMPO REAL

Antes que nada, vale la pena presentar unas definiciones operativas sobre estos importantes recursos metodológicos del variacionismo:

ESTUDIO DE VARIACIÓN Y CAMBIO EN *TIEMPO APARENTE*

Exploración de la distribución de la variación lingüística en hablantes pertenecientes a distintos grupos de edad (normalmente *jóvenes, adultos y mayores*), en la misma comunidad de habla, entrevistados en el marco de un

¹⁸ Existen otros algoritmos, como el de aprendizaje gradual de Boersma & Hayes (2001) o el de la *noisy Harmonic Grammar* de Coetzee & Pater (2008). VARBRUL es compatible con ambos. Por otra parte, y de manera por demás interesante, el modelo social de “modos de vida” (Højrup 1983) adaptado para la ciudad de México en el artículo de Lastra & Martín Butragueño (2000: 36-38) predice que los modos de vida medios son los que promueven más el cambio, a diferencia de los modos inferiores y superiores, produciendo también una curva en forma de “s”.

¹⁹ Página web: <http://www.antlab.sci.waseda.ac.jp/software.html>.

mismo levantamiento sincrónico. Es el método más común en sociolingüística para avanzar hipótesis sobre posibles cambios lingüísticos en marcha.

ESTUDIO DE VARIACIÓN Y CAMBIO EN *TIEMPO REAL*

Exploración de la evolución de la variación lingüística en muestreos de hablantes entrevistados (o documentados) en distintos levantamientos sincrónicos (separados estos levantamientos mínimamente por una generación, aproximadamente 30 años), en la misma comunidad de habla. Es el método de la lingüística histórica y permite confirmar si un cambio lingüístico se ha llevado a cabo.

Recuérdese que el estudio de la variación y el cambio lingüístico en marcha, tal como surgió en la década de 1960, planteaba dos grandes maneras de observar las variantes innovadoras que suponemos son cambios en curso: una de estas es hacer observaciones en *tiempo aparente* (Labov 1996: Cap. 3), es decir, se realiza una muestra de la comunidad de habla y se trabaja con grupos de edad bien diferenciados; si la variante innovadora es la preferida por las generaciones jóvenes y se encuentra menos favorecida por las generaciones mayores, entonces puede inferirse que el cambio se está consolidando. Sin embargo, esa es sólo una posible interpretación de la distribución de la variante innovadora en el tiempo aparente. También sucede que ciertas variantes “innovadoras” son preferidas durante la adolescencia y juventud, pero éstas se pierden cuando se alcanza la madurez; en este caso se habla de un

fenómeno de *age-grading* (o *variación estable*), y por lo tanto la variable simplemente se encuentra estratificada por edad en la comunidad; esta situación puede permanecer por un número indeterminado de años y de hecho puede no provocar un cambio lingüístico del todo.

Las observaciones en tiempo aparente, como puede inferirse, requieren necesariamente de las observaciones en el *tiempo real* (Labov 1996: Cap. 4) para confirmarse²⁰. Para realizar un análisis del cambio en tiempo real se puede proceder de dos maneras. La primera de estas es recuperar los escritos del pasado donde se menciona el cambio que se pretende estudiar. Los problemas de esta manera de proceder es que se depende demasiado de los detalles que hubieren llamado la atención del autor, casi siempre se trata de comentarios aislados y raramente se basan en análisis exhaustivos con un cierto número de informantes. Fuentes mucho más valiosas son los atlas dialectológicos. Labov (1996, 2006, 2010) se basó en las fuentes escritas de los grandes atlas dialectológicos norteamericanos como el de Nueva Inglaterra (Kurath *et. al* 1939) para rastrear los cambios en curso en los estados del noreste estadounidense. A pesar de las ventajas evidentes de trabajar con atlas dialectales, estos también presentan el problema de estar interesados en procesos no necesariamente coincidentes con los del trabajo del investigador moderno; además de que pueden existir restricciones metodológicas que pueden ser cruciales: recuérdese

²⁰ Es importante aclarar que la noción de *tiempo real* no está asociada con la noción epistemológica de *realismo*: tanto el estudio en tiempo aparente como el estudio en tiempo real implican aproximaciones realistas al problema del cambio (cf. Figueroa 1994).

que los atlas dialectales tradicionales se conformaban con datos de unos cuantos informantes (cuando no de uno solo) de cada población; además, generalmente se buscaba hablantes “NORM” (*non-migrant, old, rural, male*; cf. Chambers & Trudgill 1998: 29), esto es, un individuo que mantuviera el habla originaria de la manera más fiel posible, por lo que la estratificación social en la población es un aspecto que se pasaba por alto en dichos trabajos.

La segunda manera de observar el cambio en el tiempo real es levantando muestras de habla de la misma comunidad con una separación temporal de, al menos, una generación (normalmente 30 años) entre las mismas. Hay dos maneras de hacer este trabajo: entrevistando a los mismo informantes del estudio original (el estudio “de panel”), lo cual resulta sumamente complejo puesto que los informantes originales pudieron haber muerto o emigrado a otras ciudades; o levantando un nuevo muestreo que respete los criterios del estudio original, de manera que tengamos una sólida base de comparación entre las dos muestras.

Obviamente el tipo de preguntas que pueden responderse son distintas dependiendo del tipo de muestreos que se haga. Con el estudio de panel puede verse cuál es la evolución lingüística de los hablantes individuales: la ontogénesis de la variación (y el cambio) en cierto sentido; también puede determinarse si los fenómenos variables que se observan en el tiempo aparente son procesos de variación estratificada por edad y además se cuenta con referentes fiables sobre el cambio individual. Con los muestreos donde se reproduce el muestreo antiguo puede confirmarse el avance o retroceso del cambio comunitario, aunque se pierda

información más específica sobre el cambio en los individuos: se cuenta con la filogénesis de la variación y el cambio, en este caso. El muestreo aquí utilizado corresponde al segundo tipo: se tratará de rastrear los procesos en *muestreos pareados* separados por una distancia temporal de aproximadamente 30 años, muestreos ubicados alrededor de los años 1970 y 2000. Además, se explorarán estos muestreos de dos maneras: 1) a través del análisis de los muestreos de 1970 y 2000 “en bloque” y 2) tomando como referencia temporal el *año de nacimiento* de los informantes, independientemente de la época en que hubieran sido entrevistados.

Un clásico estudio en tiempo real (Cedergren 1984) permitió dejar claras las hipótesis que pueden trazarse dependiendo de las distintas distribuciones de una variable en el tiempo. Dicho trabajo trata sobre la fricativización de la africada (tʃ) en la ciudad de Panamá. En la Figura 2.1 se proyectan dos modelos de variación asociada al tiempo respecto a los niveles de producción de la variante africada. Por un lado está el modelo de cambio generacional: en 1969 (línea continua, con cuadros blancos) los dos grupos de jóvenes (15 a 26 y 27 a 32 años), a diferencia de todos los demás, favorecían la variante africada, por lo tanto, Cedergren proyectó (en línea punteada) un modelo donde los grupos más jóvenes y el primer grupo adulto siguen impulsando este proceso.

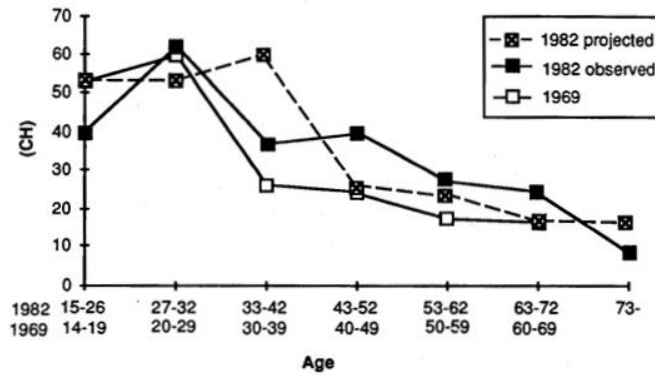


Figure 4.8b. Model of generational change of (ch) in Panama City: Projected and observed values for Cedergren's restudy

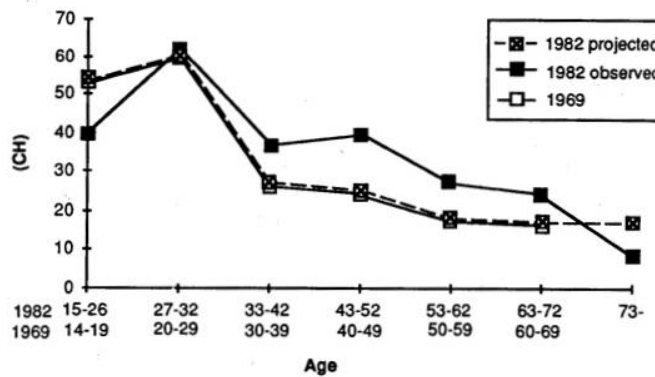


Figure 4.8c. Model of age-grading of (ch) in Panama City: Projected and observed values for Cedergren's restudy

Figura 2.1. El estudio de Cedergren sobre debilitamiento de (tʃ) en Panamá (Labov: 1994: 96, Figs. 4.8b y 4.8c)

Como puede constatarse en la parte superior de la Fig. 2.1, el levantamiento de datos de 1982 no confirmó este patrón. El segundo modelo, el de estratificación por edad (*age-grading*), es el que más se adecua a los valores encontrados en 1982 (en línea continua, cuadros negros): toda la comunidad presenta un patrón de aumento de las variantes fricativas, pero siguen siendo los dos grupos más jóvenes los que más la favorecen, tal como se encontró en 1969. Por lo tanto, puede concluirse que este

proceso variable es más bien un fenómeno de estratificación por edad: en su juventud, los hablantes panameños tienden a las variantes estándar para luego fricativizar más a medida que se vuelven adultos y pasan a la vejez.

Un ejercicio de este tipo puede realizarse con los procesos fonéticos de asibilación de róticas y debilitamiento vocálico que se analizarán en los Cap. 3 y 4 respectivamente. Las hipótesis que se plantean al respecto son que la asibilación de róticas se encuentra en proceso de cambio (cf. Lastra & Martín Butragueño 2003), mientras que el proceso de debilitamiento vocálico parece tratarse de un proceso de variación estable (cf. Serrano 2006). Por lo tanto, se espera que estas dos variables presenten un comportamiento estadístico distinto. Gráficamente, lo que se puede encontrar es lo siguiente:

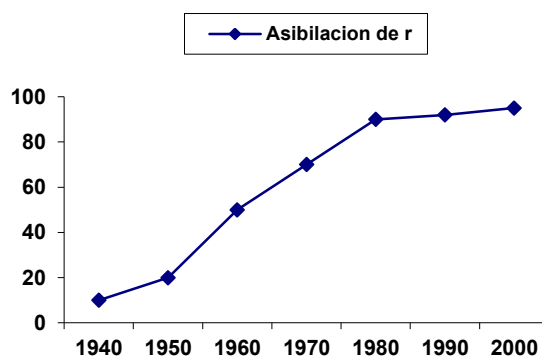


Figura 2.2. *Hipótesis sobre la asibilación de vibrantes como proceso de cambio lingüístico*

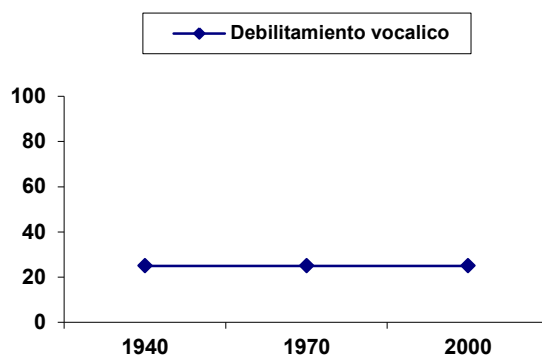


Figura 2.3. *Hipótesis sobre la estabilidad del debilitamiento vocálico*

Esto es, un cambio en curso típicamente se comporta como una curva en forma de “s” (Fig. 2.2), si se toma en cuenta las proyecciones de Perissinotto (1975), la variante asibilada de las róticas debería ir ganando terreno en la comunidad lingüística capitalina. Por otra parte, la hipótesis de la variación estable (Fig. 2.3) parece compatible con los estudios recientes sobre el debilitamiento vocálico en la ciudad de México (Serrano 2006).

Por último, los cambios léxicos que se analizarán en el capítulo 5 pueden presentar un comportamiento en forma de campana a lo largo del tiempo (Fig. 2.4): poca presencia estadística al inicio, una etapa de uso frecuente y finalmente una etapa de retracción:

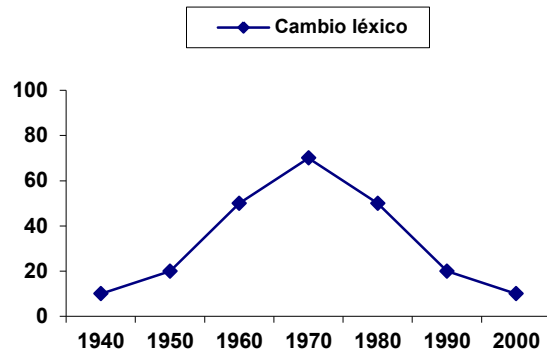


Figura 2.4. *Hipótesis sobre el cambio léxico*

El léxico típicamente presenta una curva en forma de campana: existe un rápido movimiento ascendente en su uso, se produce el período de auge de la forma y finalmente es desplazada a un ritmo también bastante rápido en términos relativos.

Los siguientes capítulos tratarán de comprobar si este es realmente el comportamiento de las variables que se analizarán, lo que puede tener importantes consecuencias teóricas sobre la difusión sociolingüística del cambio en distintos niveles estructurales.

CAPÍTULO 3

VARIABILIDAD DE LAS CONSONANTES RÓTICAS

3.1. REPRESENTACIÓN FONOLÓGICA DE LAS RÓTICAS DEL ESPAÑOL

Martín Butragueño (2014: 122-128) propone una geometría de rasgos fonológicos que permita trazar una fonología variable del español mexicano. Esto es, no se trata simplemente de la descripción de los principales alófonos de este dialecto del español, sino de construir una geometría que permita dar cuenta de la variabilidad inherente a muchos segmentos consonánticos y vocálicos: aspiración de [s], apertura y debilitamiento de [j], la propia asibilación de vibrantes, debilitamiento de vocales átonas, cierre de vocales medias, etc. A partir de la propuesta de rasgos segmentales de Hall (2007: 313), Martín Butragueño añade debajo del nodo DORSAL el rasgo [RLA] ('raíz lingual adelantada'), que aplica sólo para vocales y que, explica, sustituye al rasgo [tenso] de Chomsky & Halle (1968). El esquema queda de la siguiente manera:

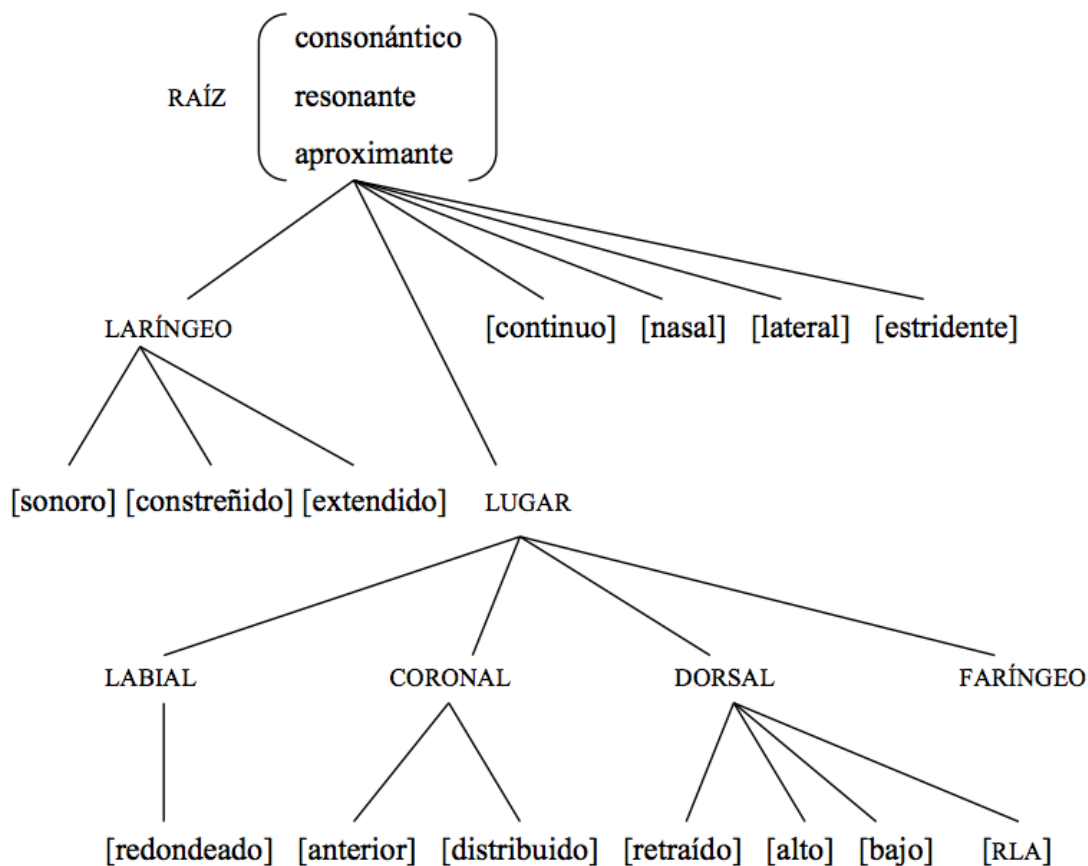


Figura 3.1. Geometría de rasgos segmentales de Martín Butragueño (2014: 123)

Las consonantes róticas compartirían los rasgos [+cons, +res, +aprox, +cor, +ant] y se distinguirían por el carácter [+cont] de [r] frente al de [-cont] de [r] (2014: 126-127)¹. La realidad fonética de las róticas fue abordada por Catford (1977: 130)

¹ Sin embargo, en el capítulo 5, Martín Butragueño (2014: 570-579) propondrá un análisis en términos de Teoría de Optimidad donde incorpora los rasgos fonéticos privativos [percusivo] para (r) y [vibrante] para (r). La geometría de rasgos, por lo tanto, incorpora dos ramificaciones del nodo “Modo”:

quien dejó claro que estos segmentos no comparten el mismo mecanismo articulatorio: mientras que la percusiva tiene un **gesto articulatorio** (un movimiento “balístico”) que puede analizarse en tres etapas (acercamiento, constricción, soltura) la articulación de la vibrante es en realidad una **postura articulatoria** en la que el cuerpo de la lengua se tensa, la punta de la lengua hace contacto con los alvéolos y por efecto Bernoulli la corriente de aire produce el característico movimiento vibratorio (pasivo) del ápice lingual (Bradley 2001: 122). Si se tratara de articular la vibrante a través de gestos como el necesario para producir /r/, sólo se obtendrían 5 ó 6 ciclos por segundo, en cambio, la vibrante en español presenta alrededor de 30 ciclos por segundo. Catford añade: “[T]here are languages in which flap and trill, for example, [ɾ] and [r], are opposed to each other as the 'short' and 'long' members, respectively of a phonological correlation of duration or 'quantity' as it is often called; but ... from a rigorous phonetic point of view a trill is not a lengthened flap” (1977: 196, *apud* Bradley 2001: 122).

Siguiendo este razonamiento, Bradley (2001) propuso una representación de las róticas que incluye una grada de duración (*timing*) por encima de la grada

la primera es “No continuo” (del que se desprenden los rasgos [africado] y [oclusivo]) y la segunda “Continuo”, del que se desprenden [fricativo], [aproximante], [lateral]; finalmente un nodo [rótico] de donde se desprenden los rasgos [percusivo] para la vibrante simple y [vibrante] para la múltiple. Los rasgos privativos [percusivo] y [vibrante] ofrecen la ventaja de evitar el uso del rasgo binario [+/-tenso] que, como había señalado Martín Butragueño en (2002) resulta bastante problemático caracterizar. Más adelante se discutirá esta propuesta.

[coronal] que permite explicar la función contrastiva de las róticas entre vocales². Además, bajo el nodo CORONAL, Bradley añade una grada “Apertura”, a partir de una versión ampliada de la “Teoría de la Apertura” de Steriade (1993, 1994), quien originalmente propuso tres grados de constricción: Apertura 0 (cierre oral), Apertura Máxima (soltura A_m) para las vocales y aproximantes y una Apertura f (A_f) que crea la turbulencia de las fricativas. Bradley añade (siguiendo a Inouye 1995) un grado de apertura intermedia entre la de las oclusivas y las fricativas, que será la propia para la vibrante (A_t); finalmente, añade un grado para la Apertura propia de las vocales: A_v , dejando A_m como un grado de apertura para las aproximantes. En resumen, los cinco grados de apertura que propone Bradley son: Apertura 0 (oclusivas), Apertura t (*trill*), Apertura f (fricativas), Apertura m (aproximantes) y Apertura v (vocales). Su propuesta es colocar el rasgo de apertura en el nodo CORONAL y no en el de MODO (como lo sugiere Inouye 1995: 91), esto por el comportamiento de las róticas en grupos consonánticos.

Desde la propuesta reseñada, percusiva y vibrante comparten la misma duración total, pero dicha duración se subdivide en tres momentos en el caso de la percusiva. Esquemáticamente, Bradley (2001: 131) los representa así (traducción mía):

² El problema del contraste fonológico entre /r/ y /r/, vale la pena aclararlo, no compete a esta tesis.

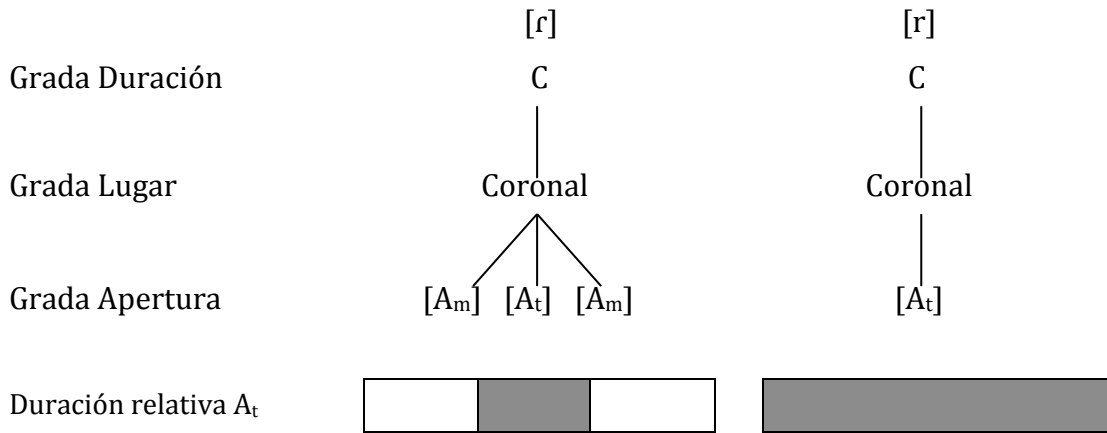


Figura 3.2. Una representación fonológica de las róticas del español (Bradley 2001: 131)

El grado de apertura “trill” (A_t) es muy breve en el caso de la percusiva y de hecho lo preceden y siguen movimientos de acercamiento y soltura con apertura aproximante [A_m], conformando un contorno de apertura que distingue a esta consonante de todas las demás en cuanto a sus movimientos articulatorios; en cambio, la vibrante sólo tiene el grado de apertura trill, pero sostenido durante más tiempo para obtener la duración equivalente a un segmento consonántico cualquiera³.

Dado lo anterior, en el presente análisis se asume una representación fonológica de la percusiva (r) como **contorno de apertura** ($A_m+A_t+A_m$) y de la vibrante (r) como un segmento producido por **un solo grado de apertura trill** (A_t) que se alarga. También se asume, en consonancia con Bradley (2001), que la Teoría de la Apertura por él modificada sólo aplica para los segmentos coronales. Sin embargo,

³ Intervendrían aquí criterios de tipo perceptual: según la Teoría de la Dispersión (Flemming 1995), debe asegurarse el contraste percusiva:vibrante en el output (Bradley 2001: 31-32).

vale la pena distinguir entre coronales anteriores (alveolares, dentales) y no anteriores (alveopalatales y palatales), aspecto necesario para describir la variación fonética en español (cf. Martín Butragueño 2014); en el caso específico de este estudio, para dar cuenta de la variante asibilada [ʒ] que es [+continua, +sonora, +coronal, -anterior] y que se opone a ciertas variantes fricativas que pueden caracterizarse fonéticamente como alveolares sonoras o sordas de constricción apical [ʒ̥] [+continua, +sonora, +coronal, -anterior]; esto es, ambas variantes son apicales y pueden estar ensordecidas, pero se distinguen por la posición más atrasada de la lengua en el caso de la asibilada. De esta forma, si se añade el Grado de Apertura en algún punto del nodo CORONAL (en este caso, debajo del rasgo [Anterior]) se obtiene una configuración en la que [Aproximante] desaparece del nodo RAÍZ ya que ahora se asume como un grado de apertura más. Así, puede proponerse una geometría de rasgos (del nodo coronal) para el análisis de las róticas y sus principales alófonos como la siguiente:

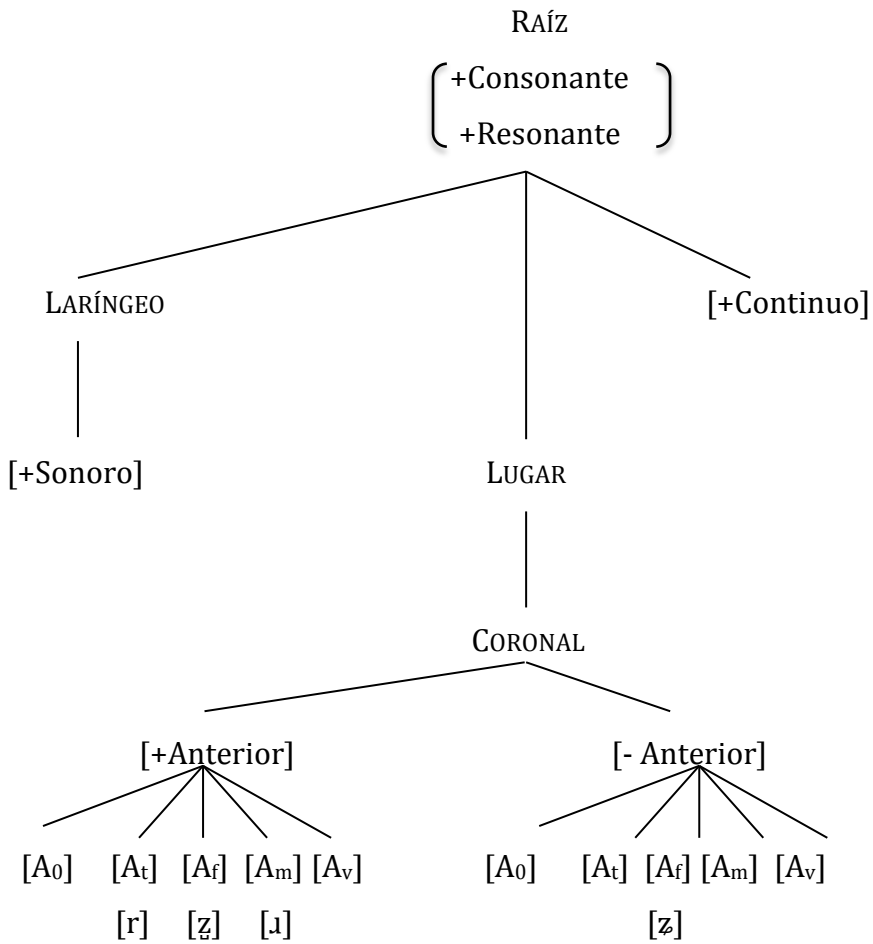


Figura 3.3. Geometría de rasgos para alófonos róticos del español mexicano

En la siguiente sección se presentan las cuatro variantes fonéticas principales que se distinguen para cada rótica a lo largo del capítulo y su correspondiente representación fonológica.

3.2. EL PROCESO FONÉTICO DE ASIBILACIÓN DE LAS CONSONANTES RÓTICAS EN ESPAÑOL

El proceso en que las consonantes róticas (r) y (r) se vuelven fricativas se conoce en la bibliografía lingüística tradicional como “asibilación” (v.g. Canellada & Zamora 1960;

Lope Blanch 1967; Penny 2004: 157-158), por el carácter estridente que suelen tener⁴. Lastra & Martín Butragueño (2003) (en adelante citado como “L&MB 2003”) hacen un completo estado de la cuestión sobre la asibilación en español americano en particular, donde señalan que es rasgo extendido en varios países del continente como Venezuela, Perú, Argentina, entre otros.

Enrique Obediente hace un análisis de la asibilación de (r) y (r) en la variedad de Mérida, Venezuela. Obediente define acústicamente la asibilada como: “caracterizada esencialmente por una difusión irregular de la energía en las altas frecuencias (de ahí su sibilancia), por una relación frecuencia/intensidad distinta a la de los sonidos propiamente vibrantes, y por un grado relativamente importante de tensión, solo superado por la *r* múltiple” (2008: 155). Estas anotaciones coinciden con lo mencionado en L&MB (2003), quienes hacen una revisión de la bibliografía sobre vibrantes asibiladas, las cuales:

dejan ver una estructura acústica de gran interés. Puede desaparecer de ellas la estructura formántica o no, pero lo que parece rasgo pertinente es la “fricación turbulenta que ocupa la mitad superior de su espectro” (Quilis 1981, p. 302), como ocurre con las fricativas de resonancias altas (Quilis 1993, pp. 262 y ss.; Martínez Celdrán 1998, pp. 69-70) [...]. Según el estudio acústico de Quilis y Carril de 1971 sobre asibilación de (rr), que parece ser

⁴ En ocasiones se les llama *rehiladas*, pero el llamado rehilamiento exige la sonoridad para llamarlo así (Torreblanca 1979: 461). Preferimos el término *asibilación* porque este incluye variantes sordas y sonoras del segmento en cuestión.

todavía el más detallado [...] la [rr asibilada] normalmente es sonora, y en pocos casos (12.5%) se ensordece, a veces en contacto con una consonante sorda. Esta es una observación muy interesante pues, en contraposición, muchas de las observaciones de los dialectólogos asocian la asibilación al ensordecimiento.

En Martín Butragueño (2008: 182, n. 14) se sintetiza esta descripción:

Se trata de fricativas alveopalatales cuyo timbre recuerda al de las sibilantes. Suelen presentar redondeamiento labial, flexión del ápice, que queda tras los incisivos inferiores, con constricción predorsoalveolar o mediodorsoprepalatal; se produce un sonido fricativo y tenso, sonoro aunque a veces se ensordezca o sea sordo. En el espectro, aparece una turbulencia en la parte alta; pueden poseer cierta estructura formántica. Si está presente el F2, en muchos casos es armónico; si no está, aumenta la percepción sibilante.

Aunque parece claro que por “asibilación” puede entenderse distintos grados de intensidad de fricación y de valor de la sonoridad, se ha decidido para esta investigación cerrar a cuatro el número de variantes fonéticas por considerar: vibrante canónica, aproximante, asibilada y alargada (o breve), mismas que caracterizamos a continuación.

3.3. LAS RÓTICAS EN ESPAÑOL DE LA CIUDAD DE MÉXICO: VARIANTES CONSIDERADAS

Como en todo trabajo de sociolingüística variacionista, la *variable lingüística* es un constructo descriptivo por parte del analista y por ello conlleva cierto grado de abstracción de las variantes que la componen⁵. En este caso, se decidió trabajar cuatro variantes fonéticas. Uno de los criterios para esta decisión es que se trata de las variantes tradicionalmente reconocidas en la bibliografía al respecto (Lope Blanch 1967; L&MB 2003; Obediente 2008), además de que dichas variantes presentan una relativa prominencia perceptual que facilita la tarea del investigador: esto es, se buscaron alternantes que fueran claramente discernibles unas de otras de manera impresionística y sin necesidad de realizar análisis acústico computarizado.

3.3.1. Variante 1: rótica canónica

La representación fonológica coincide con la realización fonética, por lo tanto: /r/ → [r]; y /r/ → [r]. Un ejemplo lo tenemos en la Fig. 3.4:

⁵ La variable lingüística para Labov (1966: 15) es “una clase de variantes ordenadas a lo largo de una dimensión continua y cuya posición está determinada por una variable independiente, lingüística o extralingüística” (en López Morales 1983: 18). Las características principales de la variable lingüística son: a) alta frecuencia en el habla cotidiana, b) se presentan todas las opciones en todos los individuos (aunque en distintas proporciones), c) es sensible a la variación estilística y d) que (idealmente) se asocia con variables sociales específicas (sexo, edad, grupo social) (cf. Tagliamonte 2006: Cap. 5).

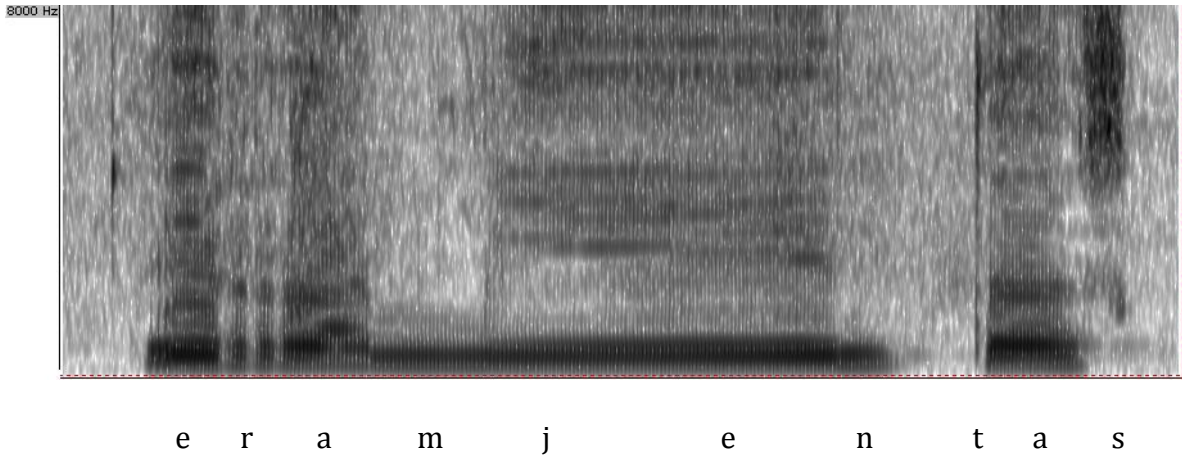


Figura 3.4. *La palabra herramientas mostrando una realización canónica de la rótica vibrante [r]*

En dicha figura se pueden observar los tres golpes de la vibrante (en forma de líneas en blanco, que indican la interrupción de la corriente de aire). Siguiendo la geometría de rasgos propuesta en la Fig. 3.3, la representación de las róticas canónicas sería de la siguiente forma (se trazan sólo los rasgos a partir del nodo “Lugar”):

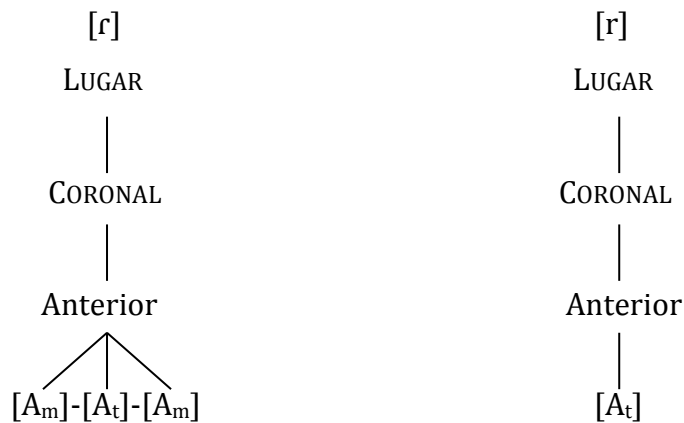


Figura 3.5. *Representación fonológica de las róticas canónicas del español (percusiva y vibrante)*

3.3.2. Variante 2: aproximante [ɹ]

Conocida en la tradición dialectológica como “fricativa”, la variante aproximante tiene como rasgos principales los siguientes: 1) se pierde el (los) momento(s) de constricción del ápice de la lengua en los alvéolos; 2) existe relativa distensión articulatoria; y 3) puede haber o no ensordecimiento. Un parámetro acústico que permite caracterizar esta variante como un tipo de aproximante es que cuando son plenamente sonoras presentar cierto grado de continuidad formántica con las vocales adyacentes. Para la variante aproximante tanto del fonema vibrante como percusivo se utiliza el mismo carácter fonético: [ɹ]. Aunque la realización como aproximante es la misma en ambos segmentos, las articulaciones fonéticas correspondientes a la vibrante múltiple (r) son más largas que las de su contraparte simple [ɹ:], como se puede apreciar en el siguiente espectrograma:

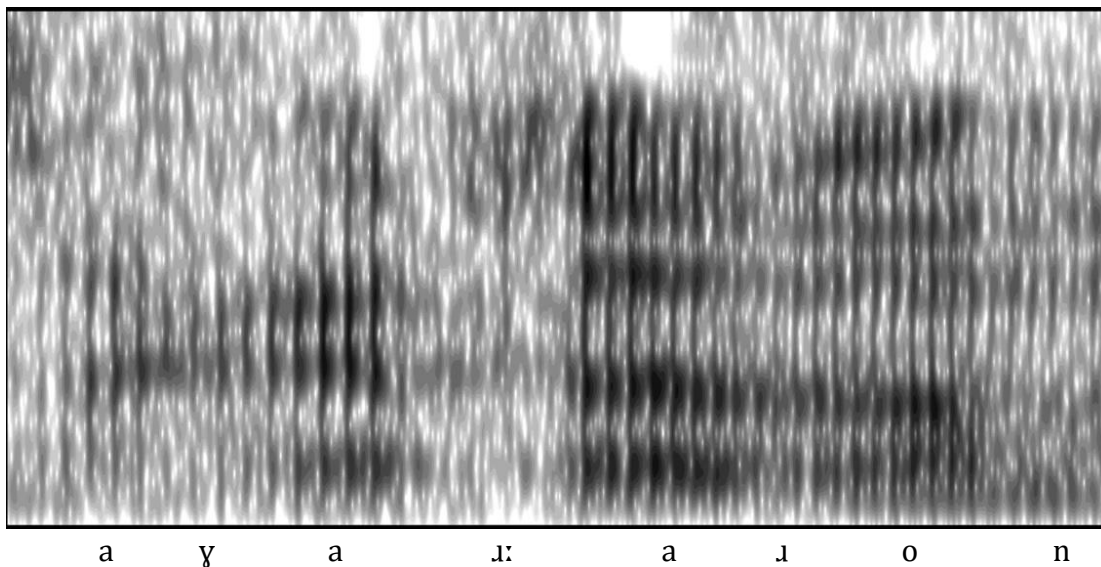


Figura 3.6. Realizaciones de róticas como aproximantes en la palabra *agarraron* (rango del espectrograma: 8000 Hz)

La duración de la primera aproximante (forma subyacente /r/) es de 0.069 s, mientras que la segunda aproximante (forma subyacente /r/) tiene una duración de 0.026 s⁶. En términos de Teoría de la Apertura, puede representarse la aproximante como una desasociación del momento de Apertura trill (A_t), quedando sólo el grado de apertura aproximante (A_m) en su articulación. La diferencia entre la aproximante proveniente de la percusiva y la derivada de la vibrante es sólo de duración:

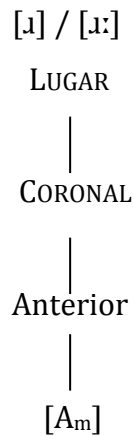


Figura 3.7. Representación fonológica de las róticas aproximantes

3.3.3. Variante 3: asibilada [ʒ]

Como ya se comentó, en realidad se reconoce en este trabajo como “asibiladas” a un conjunto de variantes que comparten las siguientes características acústico-

⁶ El hecho de que sea precisamente la duración del segmento lo que distinga (r) de (r) apoyaría el análisis fonológico de las vibrantes como simples realizaciones fonéticas de un solo segmento fonológico (esto es, podría tratarse la [r] como una geminada -cf. Martín Butragueño 2002: 93-94); sin embargo, v. *supra*.

articulatorias: tener modo fricativo, ser producidas en la zona alveopalatal y estar (o no) acompañadas de pulsos glóticos (sonoridad); por lo tanto, el mismo carácter fonético [z] se empleará para representar variantes sonoras y ensordecidas total o parcialmente. Debe anotarse que en términos perceptuales puede ser en ocasiones difícil distinguir claramente entre una aproximante ensordecida [ʝ] y una asibilada propiamente. Sin embargo, lo que distingue una asibilada sorda o ensordecida de una aproximante sorda, además del modo aproximante o fricativo, es el punto de articulación, que es más atrasado en el caso de la asibilada (normalmente posalveolar o prepalatal, según Zamora Munné & Guitart 1988 [*apud* Martín Butragueño 2014: 538]) y produce un timbre en la fricación más agudo (y muchas veces más intenso) que en su contraparte realizada en la zona alveolar o dentoalveolar. El siguiente espectrograma muestra una asibilada a final de emisión.

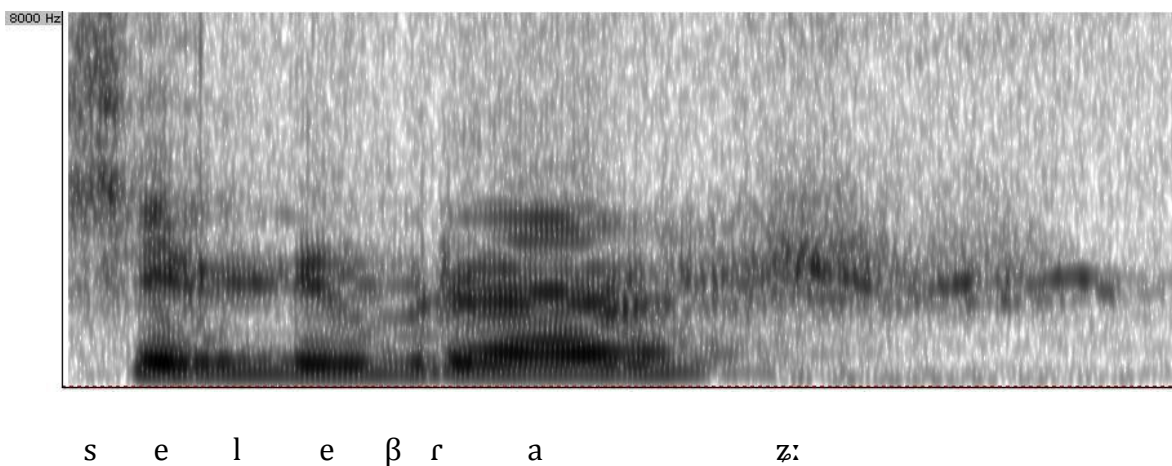


Figura 3.8. Asibilación de (r) final en la palabra celebrar

La asibilada por tanto implica dos procesos de desasociación de rasgos respecto a la rótica en el input: por un lado, se produce en la zona alveopalatal, por lo tanto es [- Anterior]; por otro, implica un grado de Apertura fricativa (A_f); su representación fonológica es la siguiente:

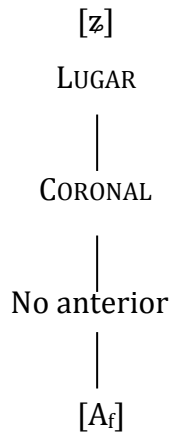


Figura 3.9. *Representación fonológica de las variantes asibiladas*

Es precisamente esta variante la de máximo interés para este trabajo, ya que los resultados de estudios recientes señalan la rótica asibilada como una variante fonética en proceso de retracción (L&MB 2003; Serrano 2008 a). Se habla de proceso de *retracción* porque Perissinotto (1975), quien trabajó con datos levantados entre los años 1963-1968 (algunos de sus informantes incluso forman parte de mi propio corpus) proyectaba la asibilada como la variante preferente en el futuro del español capitalino:

No ha de descartarse la posibilidad de que en un futuro más o menos cercano la vibrante pase a ser una articulación minoritaria destinada a desaparecer. Notamos, en efecto, que la [r asibilada] es común entre mujeres de cualquier grupo y entre algunos hombres. Cabe subrayar, además, que la asibilación es socialmente aceptada por la clase alta y, por consiguiente, común entre la media. Con estas consideraciones creo que se puede proyectar la asibilada hacia el futuro y afirmar que irá adquiriendo mayor popularidad debido al impulso recibido por las mujeres de todas las edades y grupos socioeconómicos, por la clase media de ambos sexos, y con el consentimiento tácito de los hombres de la clase alta [Perissinotto 1975: 115].

L&MB (2003) a partir de los datos de 54 informantes del *Corpus Sociolingüístico de la Ciudad de México* (CSCM) encontraron que, contra las expectativas de Perissinotto, la asibilada no ganó terreno sino todo lo contrario, disminuyó su presencia estadística en la comunidad. El presente estudio busca confirmar este hallazgo al confrontar datos levantados en la misma época que Perissinotto (1975) (datos extraídos, como ya se mencionó, de los proyectos “Norma culta” y “Habla popular de la ciudad de México” de la UNAM) con muestras actuales (las del CSCM).

Actualmente parece consolidarse la asibilada como un fuerte estereotipo del habla de varones homosexuales en México. Debe recordarse que muchos cambios lingüísticos suelen tener etapas de notoriedad sociolingüística y por lo mismo pueden devenir en estereotipos lingüísticos. En Serrano (2008 a) ya comentaba la cuestión de

las asibiladas en la televisión mexicana, particularmente en el caso de los personajes ‘La Jitomata’ y ‘La Perejila’, interpretados por los actores cómicos Alfredo y Germán Ortega, quienes presentan róticas en ocasiones exageradamente asibiladas.

Al respecto, Eller (2013) investigó la asociación sociolingüística del rasgo de asibilamiento de róticas en el habla de 21 hombres homosexuales del Distrito Federal. Doce de los 21 informantes de la muestra hablaban “marcadamente gay” (*gay-sounding*, cf. Smyth & Rogers 2002: 300). Como parte de su investigación, Eller transcribió el habla de algunos personajes televisivos “gay” que exageran la asibilación al grado de producir secuencias de percusiva más una fricativa, como [rs]:

- (1) “*por favors* [por.fa.βó.ɾs]”, y “*la verzdad* [la.βe.ɾz.ðáð]”.

(*La Jitomata y La Perejila*)

- (2) “*muerstos* [mwé.ɾs.tos]” y “*encuerars* [eɲ.kwe.ɾá.ɾs]”.

(*Cristian de El Mundo de Manigüis*)

[Eller 2013: 71]

Los análisis cuantitativos de Eller (2013) confirmaron la correlación del asibilamiento con los informantes homosexuales que marcaban su habla *gay*:

[...] la vibrante simple asibilada ha sido adoptada por los hombres gays con la voz marcada. El habla marcada como factor en la asibilación de esta vibrante presenta un peso de 0.576 frente al 0.329 del habla no marcada. Los hombres que marcan su habla asibilan en 23.5% de los casos

mientras los hombres gays que no suenan así, asibilan en sólo 4.9% de los casos [...]. [Eller 2013: 75].

En general, la red de hombres homosexuales entrevistados por Eller presentan una mayor variabilidad de las róticas y la rótica asibilada (sobre todo la prepausal) como uno de sus rasgos característicos: “Los hombres homosexuales entrevistados para esta obra, en particular los que marcan su habla como gay, asibilan más que la población general de hoy y más que los hombres gays que no marcan su habla como tal” (Eller 2013: 78)⁷.

Finalmente, otra característica sociolingüística de la variante asibilada es su asociación en México con el habla femenina, lo que se ha anotado en estudios pioneros como los de Cárdenas (1967: 66) sobre el español de Jalisco, o los de Canellada & Zamora (1960) y Lope Blanch (1967) sobre el español del centro del país. Dicha

⁷ Otro rasgo fonético segmental que parece ser estereotipo del habla gay es el de la /s/ interdental. En abril de 2009 yo mismo escuché /s/ interdentales [θ] en el doblaje mexicano de una serie cómica de televisión inglesa titulada *Little Britain*. El actor David Walliams interpreta a “Lou”, un homosexual que cuida de un personaje supuestamente inválido y en silla de ruedas (“Andy”) (<http://www.bbc.co.uk/comedy/littlebritain/>). Las /s/ de Lou son en su mayoría muy adelantadas e incluso algunas parecen interdentales. Eller (2013: 98) también estudió este estereotipo en su red de hombres homosexuales y confirmó una “/s/ modificada” (transcrita como interdental [s]) como característica de este grupo social. Por *estereotipo* se entiende aquí lo señalado por Hillary Putnam (revisada en Lara 2004): el estereotipo lingüístico es una forma que tiene una carga social y cultural especial para los hablantes. Su función simbólica rebasa el dominio propiamente lingüístico y pasa a la esfera de lo cultural.

asociación ha sido confirmada en los estudios de Perissinotto (1975) sobre el español de la ciudad de México (ECM) y más recientemente por Rissel (1986, 1989) para la variante de San Luis Potosí. Una investigación con la mitad de los informantes de esta tesis (24 personas) también confirmó esta tendencia: hacia el año 2001, las mujeres en la ciudad de México asibilan tres veces más que los hombres (21% en ellas frente a sólo 7% en ellos, cf. Serrano 2008 a: 205).

3.3.4. Variante 4: realización “corta” o “alargada” de la consonante rótica

Se añade al análisis de las róticas una variante más que implica un ‘trueque fonético’ de las entidades fonológicas involucradas: nos referimos a los casos en que, donde se tiene una vibrante en el input, se produce una percusiva en el output como en *carro*: /'ka.ro/ → ['ka.ro], y viceversa, donde de una percusiva en el input se obtiene una realización fonética vibrante en la realización de superficie: *comer*: /ko.'mer/ → [ko.'mer]⁸. Este último fenómeno ha sido trabajado por Lope Blanch (1967) y Moreno de Alba (1994); este último autor incluso dedica un apartado a la pronunciación de vibrantes simples como múltiples. Aunque el fenómeno se encuentra en muy diversos puntos del territorio mexicano, parece estar en relativa distribución complementaria respecto a las zonas del país donde la asibilación es más frecuente (cf. Moreno de Alba 1994: 134-138). Por su parte, Lope Blanch atribuye el fenómeno a simple énfasis que quiere dar el hablante a su expresión (Lope Blanch 1967: 13). Veamos un ejemplo:

⁸ En esta sección se llamará “variante corta (o breve)” a la solución como percusiva y “variante alargada” a la pronunciación como vibrante.

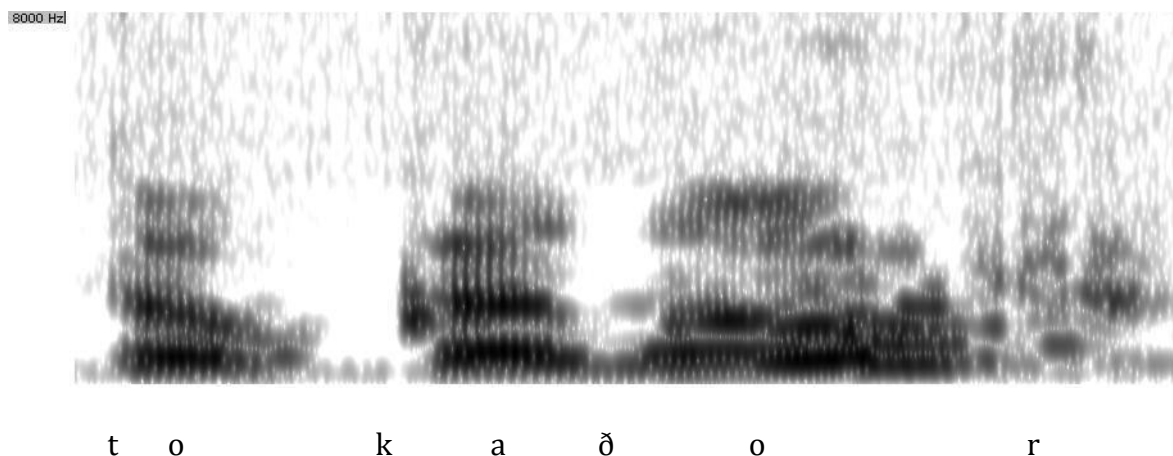


Figura 3.10. Rótica vibrante al final de la palabra *tocador*

En la figura 3.10 se aprecia los dos silencios en la vibrante, producto del contacto de la lengua en la zona alveolar⁹.

En la siguiente sección se presentan los resultados del análisis cuantitativo de las cuatro variantes consideradas.

3.4. ANÁLISIS ESTADÍSTICO DESCRIPTIVO DE LAS RÓTICAS EN ESPAÑOL DE CIUDAD DE MÉXICO

En este apartado se presenta la estadística descriptiva (sólo frecuencias relativas) para tener un panorama general de los procesos en ambas róticas y poder comparar en términos generales con los trabajos que incorporaban cuantificación también en forma de porcentajes, como los de Lope Blanch (1967) y Perissinotto (1975). Los

⁹ Aunque no he podido atestiguar ningún ejemplo, Bradley (2001: 150) asegura que también puede aparecer una percusiva realizada como vibrante en grupos consonánticos homorgánicos, como en las palabras *tren* o *creer*.

análisis en esta sección se realizaron a partir de los corpora totales de cada rótica, corpora que incorporan en una misma base de datos a las dos muestras en tiempo real (1970 y 2000; 1129 ejemplos de percusiva y 2274 de vibrante). También es importante aclarar que respecto a la percusiva sólo se analizan los casos *ante pausa* (como en *comer, traer, dormir, jalar* a final de emisión), ya que este es el contexto que favorece mayormente la variabilidad: en una primera exploración del proceso (Serrano 2005) se transcribieron las róticas percusivas en todos los contextos y sólo en el final de emisión se producía la pronunciación variable de manera significativa¹⁰.

No se presenta la interacción de factores lingüísticos y sociales respecto a cada variante en esta sección porque, como se verá más adelante, en realidad los resultados de 1970 son algo distintos a los del año 2000 y, por tanto, el peso de los factores internos y externos se vería sesgado si se combinaran los datos de ambos períodos. En las secciones correspondientes a cada rótica por separado, se presentará el análisis detallado de la interacción de factores distinguiendo precisamente según el año de muestreo (el tiempo real). La siguiente tabla muestra los resultados totales para la rótica percusiva ante pausa en el corpus analizado.

¹⁰ Este ejercicio de *circunscribir el contexto* es un paso necesario en el análisis cuantitativo de la variación (Tagliamonte 2006: 13). También en el análisis del debilitamiento vocálico (Cap. 4 de esta tesis) se circunscribieron los contextos fónicos.

Tabla 3.1. Variabilidad de la percusiva prepausal (r) en 48 hablantes de ECM

	F	F
[r]	433	38.4%
[ɹ]	241	21.3%
[ʒ]	260	23.0%
[r]	195	17.3%
Totales	1129	100%

El corpus total muestra que la pronunciación canónica de la percusiva es la de mayor peso estadístico (38.4%) seguida de la asibilada (23.0%) lo que no apoya la idea de una consolidación de esta última variante. Sin embargo, el análisis de las muestras de 1970 y 2000 por separado sugiere que ha habido cambios importantes en la distribución de las variantes.

Tabla 3.2. Evolución de la variabilidad de la rótica percusiva prepausal (r/) 1970-2000

	1970		2000	
[r]	145	(23.7%)	288	(55.7%)
[ɹ]	152	(24.8%)	89	(17.2%)
[ʒ]	192	(31.4%)	68	(13.2%)
[r]	123	(20.1%)	72	(13.9%)
Totales	612	(100%)	517	(100%)

Esta tabla revela claramente que el peso estadístico de la variante asibilada ha descendido drásticamente con el paso del tiempo: de un 31.4% en 1970 a sólo 13.2% en el año 2000 y que, además, la pronunciación canónica duplicó su presencia estadística, al pasar de 23.7% en 1970 al 55.7% en el año 2000 -- seguida de lejos por

la aproximante (17.2%) y la múltiple (13.9%). Aunque la asibilada es mayoritaria en los datos de 1970 (31.4%), llama la atención la disparidad de estos resultados con los reportados por Perissinotto (1975: 103) para el mismo período, quien registra un 68.1% de asibilación de (r) prepausal. Esta tabla indica un proceso de retracción del asibilamiento en el lapso de 30 años aproximadamente. En su estudio sobre el asibilación de las róticas en ECM, L&MB (2003) comparan sus resultados con los recogidos en el trabajo de Lope Blanch (1967), a quien citan:

De acuerdo con los materiales por mí reunidos [de 12 informantes nacidos o radicados en el Distrito Federal], la -r simple (vibrante o fricativa, sonora o ensordecida) aparece, en México, en un 75 por ciento de los casos aproximadamente; las variantes asibiladas, en un 13 por ciento; y las vibrantes múltiples, en un 12 por ciento tan sólo. Aunque estas proporciones pudieran alterarse un tanto si se analizase el habla de un mayor número de mexicanos, no creo que cambiaran hasta el grado de otorgar a las realizaciones múltiples de -r el primer puesto [Lope Blanch 1967: 5].

A partir de esta cita, L&MB discuten las similitudes de sus resultados con los de Lope y lo distintos que son respecto al trabajo de Perissinotto:

Obsérvese que las proporciones presentadas por Lope Blanch están mucho más cerca de las nuestras (51% de vibrantes y 21% de fricativas, lo que suma 72%; 10% de asibiladas y 16% de alargadas, contando todos los casos de _r//; si se

toma en cuenta sólo el estilo de conversación, tenemos un 24% de vibrantes y un 25% de fricativas, lo que da un 49%; un 27% de asibiladas y un 22% de alargadas) que las de Perissinotto [L&MB 2003: 22].

Los materiales de esta tesis, puede concluirse, coinciden con los de Lope Blanch (1967) y L&MB (2003) en términos generales. La única investigación que presenta cálculos distintos a los de todas las demás es la de Perissinotto (1975), lo que sugiere que dicho autor muy seguramente tenía criterios de transcripción distintos a los empleados en estos trabajos; es válido especular que quizás Perissinotto consideró como “asibiladas” muchos ejemplos que en realidad eran *aproximantes ensordecidas* – el parecido acústico entre la asibilada y la aproximante ensordecida es muy grande y yo mismo tuve dudas en algunos casos--, de aquí las diferencias en cuanto a la distribución estadística de las variantes. Esta disparidad de resultados se refleja también, aunque menos radicalmente, respecto a la rótica vibrante. Véase la siguiente tabla:

Tabla 3.3. Variabilidad de la rótica vibrante (r) en 48 hablantes de ECM

	<i>F</i>	<i>F</i>
[r]	1438	63.2%
[ɾ]	422	18.6%
[ʀ]	193	8.5%
[r̄]	221	9.7%
Totales	2274	100 %

En este caso, la asibilada es la variante menos presentada en el corpus (8.5%).

La siguiente tabla presenta esta variabilidad en el tiempo real:

Tabla 3.4. *Evolución de la variabilidad de la rótica vibrante (r) 1970-2000*

	1970		2000	
[r]	642	(62.1%)	796	(64.1%)
[ɾ]	220	(21.3%)	202	(16.3%)
[ʀ]	134	(13.0%)	59	(4.8%)
[r̄]	37	(3.6%)	184	(14.8%)
Totales	1033	(100%)	1241	(100%)

El 13% de asibiladas en 1970 contrasta notoriamente con los datos de Perissinotto (1975: 110), quien reporta un 31.5%. Nuevamente con datos del año 2000, L&MB (2003) obtienen resultados menos abultados (14% de vibrantes asibiladas) pero ciertamente muy distintos a los que aquí se recogen en la muestra del mismo año (4.8%). Llama la atención que, mientras la asibilada pierde terreno, aumentan los casos en que la vibrante se realiza como una percusiva [r̄], de modo que intercambiaron lugares con pesos estadísticos la asibilada y la variante breve [r̄]. Por otra parte, aumentó ligeramente la preferencia por la variante canónica (en dos puntos porcentuales) y disminuyó la proporción de variantes aproximantes. En resumen, puede concluirse con estos datos que se mantiene la variabilidad de la vibrante en español capitalino, sólo que los últimos candidatos seleccionados (asibilada y breve) intercambian posiciones estadísticas.

Conjuntando los datos de ambas róticas, puede afirmarse que en general la variabilidad es inherente a la producción de estas consonantes, y que el análisis en tiempo real permite confirmar que la asibilación de vibrantes se encuentra en franco retroceso en ECM. En las secciones siguientes se presenta la cuantificación en términos de probabilidad, combinando factores lingüísticos y sociales en el análisis.

3.5. *Análisis cuantitativo de la variabilidad de la rótica percusiva ante pausa*

Como se mencionó en el cap. 2, se utiliza el análisis VARBRUL (Sankoff, Tagliamonte & Smith 2005) para obtener el valor de probabilidad de los factores que intervienen en la selección de las distintas variantes de los segmentos róticos y un cálculo del error estadístico (la prueba ji-cuadrada o *chi-square* en inglés). En un trabajo anterior (Serrano 2005) se había explorado el papel de factores fónicos como el tipo de vocal precedente y el acento, pero no resultaron significativos. Además, es precisamente el contexto prepausal el que más favorece la asibilación de este segmento, como lo han mostrado otros estudios (Martín Butragueño 2008; Obediente 2008). Por tanto, no se consideraron factores del contexto fónico o gramatical para el análisis de probabilidad de la percusiva, sólo factores sociales: edad, sexo, año de muestreo, nivel educativo y año de nacimiento. La siguiente tabla presenta los resultados de los análisis de regresión logística de (*r*) respecto a la variable “año de muestreo” (1970 vs. 2000) *en el corpus total*, por esta razón la tabla debe leerse como “la probabilidad de que se presente la variante *X* en el año 1970 frente a la probabilidad de que se produzca en el año 2000”.

Tabla 3.5. Variabilidad de la rótica percusiva prepausal [ɾ/] en el tiempo real:
probabilidades (N=1129)¹¹

		1970			2000		
<i>Input</i>		<i>f</i>	<i>Pb 1</i>	<i>Pb U&D</i>	<i>f</i>	<i>Pb 1</i>	<i>Pb U&D</i>
[ɾ]	(.361)	23.7	.315	.334	55.7	.715	.694
[ɹ]	(.206)	24.8	.527	.553*	17.2	.468	.438*
[ʀ]	(.157)	31.4	.722	.722	13.2	.245	.245
[ɹ̥]	(.157)	20.1	.547	.543*	13.9	.445	.450*

La tabla muestra una clara tendencia a una pronunciación más conservadora de la percusiva en tiempos recientes; en especial es evidente la retracción del asibilamiento, que se redujo en 2000 a menos de la mitad del peso porcentual que tenía en los años 1970 (13.2 vs. 31.4%).

La siguiente gráfica, basada en la Tabla 3.5, compara la evolución de las cuatro variantes de la percusiva, donde es clara la tendencia al abandono de las formas innovadoras a favor de la variante conservadora (según análisis de regresión de subida y bajada).

¹¹ En esta tabla de análisis VARBRUL, *input* es la probabilidad general de que la regla variable se produzca. Abreviaturas: *f*= frecuencia relativa; *Pb 1*= Probabilidad análisis binomial de 1 nivel; *Pb U&D*= Probabilidad análisis binomial de subida y bajada (*Up & Down*). El asterisco significa que el análisis *Up & Down* seleccionó la variable en cuestión (en este caso “año de muestreo”) sólo en uno de los dos mejores recorridos; por lo mismo, dicho dato debe reconsiderarse. Finalmente, un guion “—” significa que el análisis de subida y bajada descarta dicho factor.

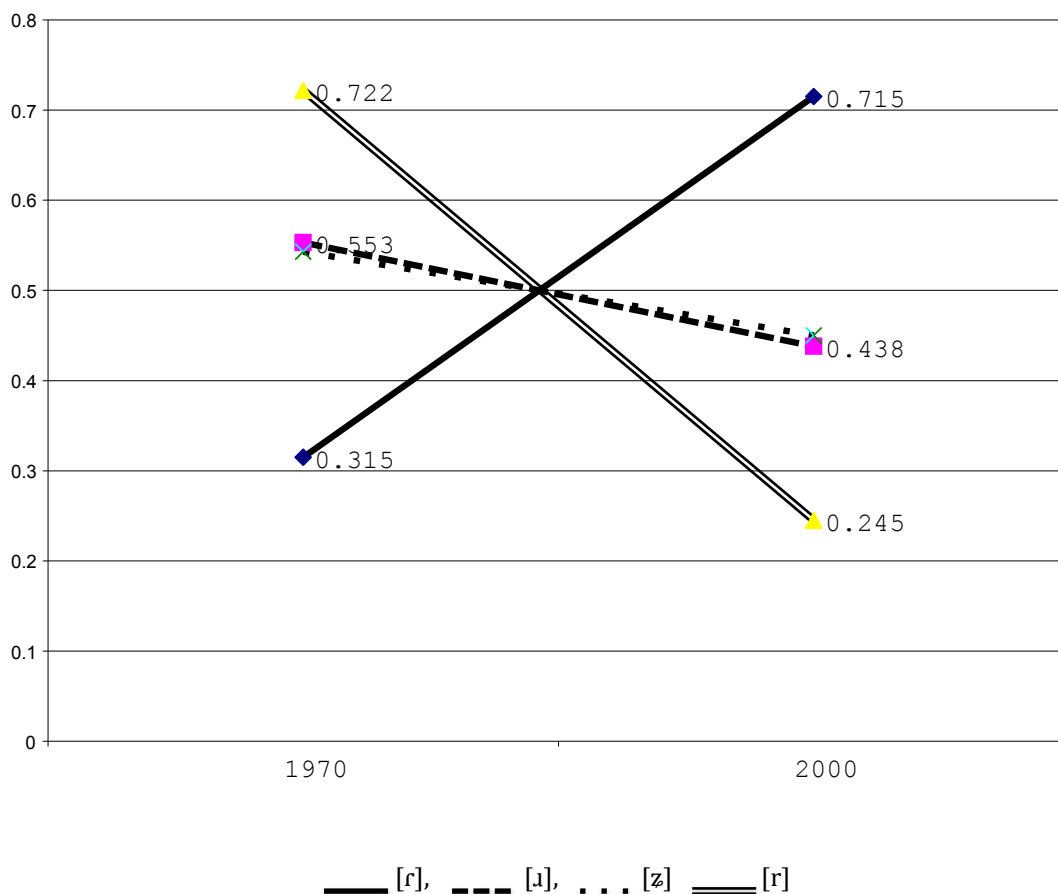


Figura 3.10. Evolución de la variabilidad de la rótica percusiva prepausal [-r/] 1970-2000 (Pb U&D, N=1129)

Se trata de dos procesos complementarios: abandono de las formas innovadoras (o no estándar) por un lado y regreso a la pronunciación normativa por otro. La asibilada es la que evoluciona de forma más acelerada, al igual que la variante canónica. La aproximante y la alargada, por su parte, también presentan un descenso, pero menos pronunciado. El análisis probabilístico de los factores sociales permite trazar el camino sociolingüístico que tomó este proceso de cambio. Como se mencionó arriba, se realizaron los cálculos VARBRUL con las muestras de cada año por separado.

La siguiente tabla resume el papel de los factores sociales respecto a la asibilación de la percusiva en 1970.

Tabla 3.6. *Peso de factores sociales en asibilación de percusiva [r/] 1970 (f=192/612)*¹²

	<i>Pb 1</i>	<i>Pb U&D</i>
SEXO		
<i>Hombres</i>	.269	.268
<i>Mujeres</i>	.745	.747
ESTUDIOS		
<i>Alto</i>	.482	!
<i>Bajo</i>	.523	!
EDAD		
<i>Mayores</i>	.400	.398
<i>Adultos</i>	.712	.712
<i>Jóvenes</i>	.439	.441

Atendiendo a los resultados del análisis de regresión *Up & Down*, sólo son dos los factores que están interactuando significativamente respecto a la asibilación de la percusiva prepausal: sexo y edad, específicamente las mujeres (.747) y las personas adultas nacidas alrededor de 1925 (.712) favorecen la asibilación; llama la atención que ni jóvenes ni mayores favorezcan esta pronunciación. Este resultado permite

¹² La variable “año de nacimiento” se ignora en estos recorridos porque no se cuenta con informantes de las cuatro generaciones en cada muestra, además de que quedan algo descompensadas estadísticamente (por ejemplo, la muestra del año 2000 tiene más de la mitad de datos del grupo nacido alrededor de 1970). Estas disparidades pueden provocar un sesgo estadístico en el valor de probabilidad de todas las demás variables, sesgo que debe evitarse.

sugerir que se trataba de un rasgo “de moda” en el grupo de edad adulto. Volveremos sobre este punto más adelante.

Complementario al análisis de probabilidad, puede realizarse uno de curvas de *frecuencia relativa acumulativa* (utilizado en Martín Butragueño 2006). Para construir esta curva, primero se ordena a los individuos según su índice de asibilación personal, que es igual al número de ejemplos de asibilada del informante entre el número total de casos en la muestra (esto es, el peso relativo con que contribuye el informante a la asibilación en la comunidad estudiada). Al siguiente informante en la lista se suma el valor obtenido para el informante 1, a su vez, al informante 3 se le asigna, además de su índice, la sumatoria de los dos informantes que le preceden y así sucesivamente. Martín Butragueño (2006) anota las ventajas de este tipo de gráfica: “las curvas de frecuencia relativa acumulativa son interesantes para describir procesos de cambio lingüístico, porque permiten considerar simultáneamente el papel de la muestra general y el de los casos individuales. Por otra parte, el uso de frecuencias y no de probabilidades permite un acercamiento menos abstracto a los rasgos sociales de los individuos” (2006: 194-195).

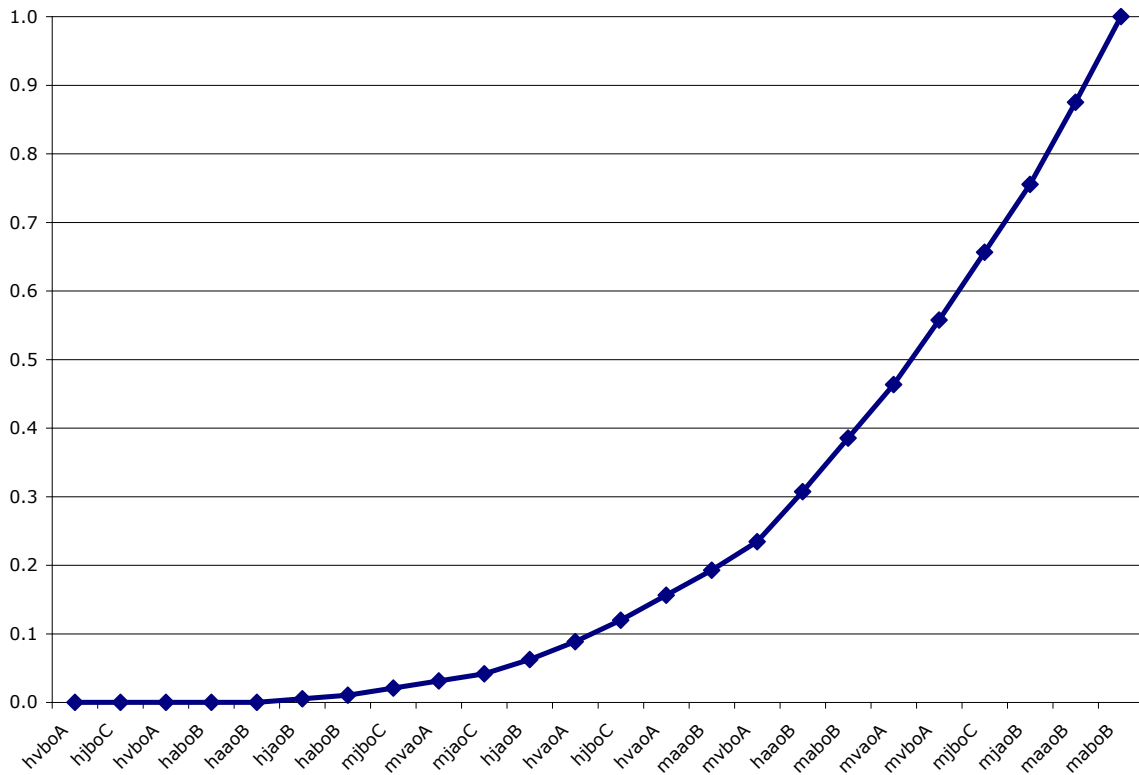


Figura 3.11. Frecuencia relativa acumulativa de asibilación de la percusiva en 1970

$$(f=192/612)^{13}$$

Para poder inferir las características de los líderes del cambio (Labov 2006) puede ponerse especial atención a la parte derecha de la gráfica. Los últimos cinco informantes contribuyen con cerca de la mitad de la frecuencia acumulada (44%). Todas son mujeres (de hecho, de los diez informantes que más asibilan, siete son mujeres). Entre estas siete mujeres, tres son adultas, dos son mayores y dos son

¹³ No se anota el nombre del informante sino una codificación de sus características sociales. Para esta y las demás gráficas, interprétese así: h/m=hombre/mujer; j/a/v=joven/adulto/mayor; a/b=nivel educativo alto/bajo; o/m=muestra 1970/2000; A/B/C/D= Año de nacimiento: 1905/1925/1945/1970.

jóvenes; cuatro nacieron en la generación de 1925, dos en la de 1905 y sólo una alrededor de 1945. Respecto al nivel de estudios, cuatro son de nivel bajo y tres de nivel alto. Concuerdan estos resultados con lo comentado por Perissinotto (1975: 115): la asibilación, rasgo ciertamente femenino, no parece estar claramente asociado a un grupo social o de edad particulares. En contraparte, es claro que los hombres de la muestra se agrupan con los índices más bajos de asibilación (los siete con los índices más bajos son hombres).

Es interesante mirar más de cerca el perfil sociolingüístico individual de las líderes de la asibilación de la percusiva en 1970. Por ejemplo, las dos mujeres que más asibilan en 1970 fueron del grupo de edad adulto: Agripina y Martha, que pertenecen a grupos sociales muy distintos. Por un lado, Agripina contaba 47 años al momento de la entrevista (agosto de 1972), estudió hasta segundo año de primaria, es casada, se dedica al hogar y atiende un pequeño puesto de dulces afuera de una escuela, por otro, Martha es maestra normalista, habla inglés y francés y se dedica a la traducción y la docencia. Curiosamente las siguientes dos mujeres que más asibilan pertenecen al mismo grupo de edad también (son jóvenes), pero a grupos sociales distintos: Sofía, de 30 años, tiene maestría en Geografía y habla inglés, mientras que Rosa, de 23 años, sólo cursó hasta 4to de primaria. Entonces, las líderes son mujeres de los grupos de edad joven y adulto y hasta ahí llegan las similitudes.

Por otra parte, la variabilidad inherente de la percusiva en el muestreo del año 2000 tiene también una distribución sociolingüística muy interesante.

Tabla 3.7. *Peso de factores sociales en la asibilación de la percusiva en 2000 (f=68/517)*

	<i>Pb 1</i>	<i>Pb U&D</i>
SEXO		
<i>Hombres</i>	.240	.240
<i>Mujeres</i>	.764	.764
ESTUDIOS		
<i>Alto</i>	.568	.583
<i>Bajo</i>	.409	.372
EDAD		
<i>Mayores</i>	.655	.655
<i>Adultos</i>	.455	.455
<i>Jóvenes</i>	.331	.331

Recordemos que se trata de sólo 68 ejemplos de la asibilada y por esta misma razón los análisis de regresión logística no tienen el mismo alcance que cuando contamos con un número importante de datos. El análisis *Up & Down* propone las tres variables estudiadas como pertinentes: sexo, grupo de edad y nivel educativo. En el año 2000 siguen siendo mujeres quienes más asibilan (.764), las personas de edad más avanzada (.655) y de más estudios (.583), lo que permite inferir que la generación adulta de los años 1970 (que se corresponde con la generación mayor del año 2000), mantuvo este rasgo, pero fue transmitido en proporciones mucho menores a las nuevas generaciones.

Este panorama sociolingüístico sobre la asibilación moderna de la percusiva puede confirmarse en términos generales a través del análisis de la frecuencia relativa acumulativa.

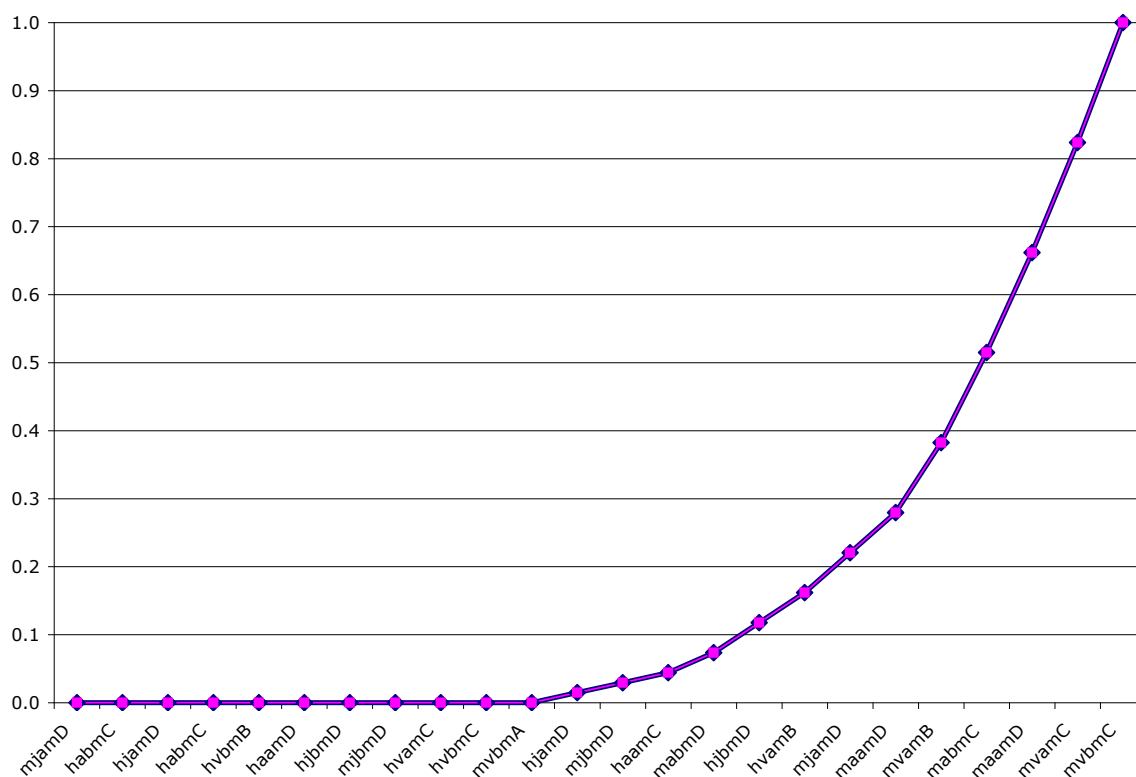


Figura 3.12. *Frecuencia relativa acumulativa de asibilación de la percusiva en el año 2000 (f=68/517)*

Nuevamente debe recordarse que son muy pocos datos de asibilación (13.2%) y que 11 de los 24 informantes no asibilan ni una sola ocasión, lo que ya es indicativo de la relativa poca pervivencia de la variante –al menos en esta muestra. Nuevamente son siete mujeres las líderes en la producción de la variante, de las cuales cuatro son del grupo de mayor edad, tres son adultas y sólo una es del grupo de jóvenes entrevistadas.

La mujer que más aportó asibiladas (12 en total) fue Berta, del grupo de edad mayor (57 años en 2007), quien dejó inconcluso el 1ro de primaria, y a pesar de tener

una procedencia socioeconómica humilde, ascendió su nivel de vida general, lo que le ha permitido pagar la escuela particular para dos de sus hijos y contar con casa propia, tres autos, salir de vacaciones al interior del país, entre otras actividades. La segunda informante con más asibiladas fue Liliana, de 62 años, quien cumple con el perfil señalado por Labov (2006) respecto a los líderes lingüísticos: mujeres exitosas ubicadas en el centro de la escala socioeconómica, que iban en su juventud a contracorriente de los patrones establecidos por su comunidad y con redes sociales extendidas. La historia personal de Liliana es muy interesante: hija de padres con estudios universitarios, fue de las primeras alumnas de arquitectura en la UNAM en los años 1960 (rama de estudio eminentemente masculina en la época), salió de su casa para estudiar en Inglaterra un año contra la voluntad de sus padres, rechazaba la idea del matrimonio, tuvo su primer hijo a los 36 años, etc. Lo que Liliana y Berta tienen en común, además de haber nacido en los años 1940, es el ascenso social (o como señala Martín Butragueño (2006), sentirse *insertas* en un proceso de ascenso social). Se volverá a discutir este punto en las conclusiones del capítulo.

Ahora, si se consideran los datos de frecuencia relativa de las variantes asibiladas de la rótica percusiva según las generaciones de informantes, el panorama es el siguiente.

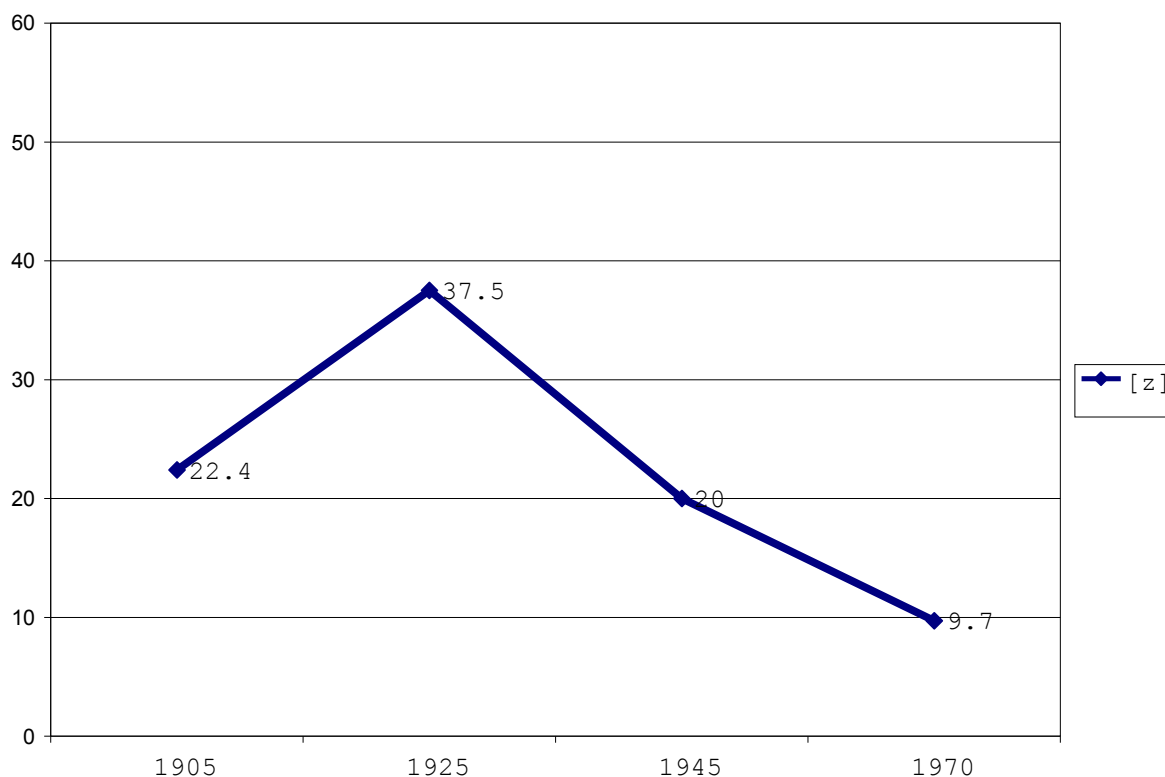


Figura 3.13. Frecuencia relativa de la variante asibilada de la percusiva prepausal por año de nacimiento (48 informantes, $n=1129$)¹⁴

Este recurso gráfico permite ver el desarrollo de la asibilada en el tiempo real desde otra perspectiva: en lugar de agrupar a los informantes por año de muestreo (1970 ó 2000), se agrupan según su año de nacimiento¹⁵. Por supuesto, esta figura implica la hipótesis de que los informantes mantienen en lo general su pronunciación

¹⁴ Se anotan los años 1905, 1925, 1945 y 1970 sólo como recurso expositivo. Debe recordarse que representan en realidad grupos de informantes cuyos nacimientos se produjeron *alrededor* de esos años (v. Cap. 2).

¹⁵ Agradezco a Natalie Schilling-Estes por esta propuesta de análisis alternativo del tiempo real.

a lo largo de toda la vida. Esta agrupación tiene la ventaja, además, de permitir emplear todos los datos a la vez. Lo que muestra la figura 3.13 es que, si bien en la generación más antigua existe la asibilación (22.4%), esta asciende de manera significativa en la siguiente (correspondiente a los “adultos” entrevistados de los años 1970) para convertirse, de hecho, en la variante más presentada en el habla de estos informantes (37.5%); las generaciones posteriores, sin embargo, presentan una curva de retracción de la asibilación, descendiendo esta variante a un 20% en la generación 1941-1960 y a un 9.7% en la generación más joven (nacidos en 1961-1981). El terreno que gana la variante canónica de la percusiva en el tiempo real es el espejo de aquella imagen:

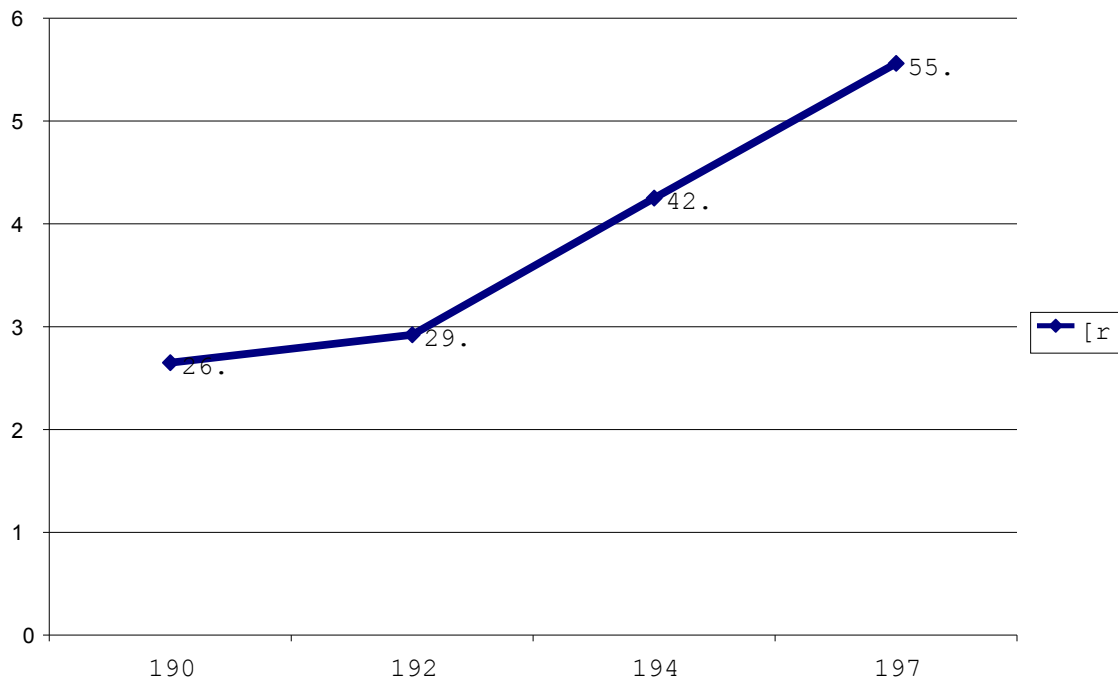


Figura 3.14. Frecuencia relativa de la variante canónica de la percusiva [r] por año de nacimiento (48 informantes, n=1129)

La gráfica es bastante clara: la variante canónica de la percusiva ha ido ganando terreno con los años, pasando de un 26.5% (sólo la cuarta parte de los ejemplos) en la generación más antigua (1887-1915) a un 55.6% moderno (los nacidos en 1961-1981). Esta recuperación de la pronunciación estándar de la percusiva se encuentra en distribución complementaria (estadística) con la retracción del proceso de asibilación.

En la siguiente sección se presenta el proceso respecto de la vibrante [r]. Esto permitirá verificar si son los mismos tipos de restricciones los que están en juego y, en caso de que coincidan factores internos y externos en ambas róticas, se tratará de encontrar una explicación unificada para el proceso en términos de Teoría de Optimidad.

3.6. Análisis cuantitativo de la variabilidad de la rótica vibrante (r)

La siguiente tabla presenta un resumen de los análisis de probabilidad de la vibrante en el total de la muestra.

Tabla 3.8. Variabilidad de la rótica vibrante (r) en el tiempo real: probabilidades

(n=2274)

		1970			2000		
<i>Input</i>		<i>F</i>	<i>Pb 1</i>	<i>Pb U&D</i>	<i>f</i>	<i>Pb 1</i>	<i>Pb U&D</i>
[r]	(.646)	62.1	.481	!	64.1	.515	!
[ɹ]	(.162)	21.3	.511	!	16.3	.491	!
[ʀ]	(.049)	13.0	.703	.700	4.8	.328	.331
[ɽ]	(.071)	3.6	.346	.346*	14.8	.629	.630*

Es importante recordar aquí que los cálculos en la tabla 3.8 surgen de los recorridos con todo el muestreo y todas las variables. La tabla pone en evidencia la importancia de la variante canónica sobre todas las demás en ambos muestreos, por lo que estamos ante un proceso que se comporta cuantitativamente de manera distinta a la rótica percusiva prepausal, donde la asibilada sí era la elección preferida en los años 1970. Incluso en los años 1970 el índice de asibilación de la vibrante es muy bajo (13%), sin embargo (como ya se había mencionado), contrasta notablemente con el peso de la variante en el año 2000 (4.8%). El hecho de que los recorridos VARBRUL no seleccionen el año de muestreo como factor de peso para la variante canónica y la aproximante (en la última columna: *Pb U&D*) es otro indicativo de que la variabilidad de la rótica vibrante es hasta cierto punto más estable en el tiempo respecto a la de la percusiva.

Para el análisis probabilístico de la vibrante se consideró un factor interno: el contexto fónico precedente. El contexto fónico subsecuente para la /r/ fonológica – que sólo se presenta a inicio de emisión o entre vocales— es siempre vocálico o aproximante, por lo que se decidió no tomarlo en cuenta. Se distinguieron cuatro contextos precedentes a la vibrante: vocálico (*la rosa, arroz*), consonántico (*un ratón*), pausa o “inicio absoluto” (*¡rómpele*) y segmento [s] (*las rosas, Israel*), que suele favorecer el asibilamiento de manera mucho más clara que otros elementos consonánticos¹⁶.

¹⁶ El segmento [s] precediendo a (r) también se exploró en Lastra & Martín (2003), y resultó significativo.

El análisis de subida y bajada en la muestra de 1970 (con 1033 datos) indica que el único contexto que favorece la vibrante canónica es el **vocálico (.586)**, los demás contextos desfavorecen el respeto a la variante estándar (en especial el segmento [s]): **consonante: 0.479 > inicio absoluto: 0.336 > [s]: 0.090**. El análisis de la variante asibilada en el mismo año complementa este panorama, ya que los contextos se ordenan de la siguiente manera: **[s]: 0.936 > inicio absoluto: 0.543 > consonante: 0.454 > vocal: 0.430**, esto es, sólo la s y el inicio absoluto favorecen la asibilación, creando una imagen estadística en espejo de la variante estándar. En términos absolutos, poco más de la mitad de casos de (r) después de [s] (como en *les robas*) fueron asibiladas (43/80) y por ello el gran peso en términos de probabilidad para este contexto específico.

En el año 2000, el panorama es muy similar. Se realizaron los análisis de regresión con sólo 59 ejemplos de asibiladas en 1241 tokens (sólo el 4.8%) frente a 796 ejemplos de vibrantes canónicas (64.1%). Nuevamente, el único contexto precedente que favorece la variante canónica es la vocal: **vocal: 0.570 > consonante: 0.387 > inicio absoluto: 0.288 > [s]: 0.199**. Los pesos varían, pero el ordenamiento es el mismo que en 1970. Respecto a la asibilada, el ordenamiento es muy parecido en los dos muestreos: **[s]: 0.883 > consonante: 0.652 > inicio absoluto: 0.587 > vocal: 0.421**; la diferencia es que en 2000 la consonante precedente sí favorece la asibilación, incluso más que el inicio absoluto; por otra parte, una vocal precediendo a la vibrante es ahora el único factor que desfavorece la asibilación (dato que coincide con lo reportado por L&MB 2003: Cuadro 19). Entonces, puede concluirse que la

asibilación de la rótica vibrante presenta un modelo multivariable más complejo pero, a pesar de ello, ha disminuido su presencia estadística en ECM. Por supuesto esta aseveración requiere de más datos para poder ser concluyente.

La vibrante realizada como aproximante [ɹ] es de sumo interés, ya que comparte algunos contextos con la asibilada. En 1970 el peso del contexto precedente para la aproximante es: **inicio absoluto: 0.682 > [s]: 0.596 > consonante: 0.538 > vocal: 0.460**. Sin embargo, en 2000 el ordenamiento selecciona el contexto vocálico como favorecedor de la aproximante: **[s]: 0.712; inicio absoluto: 0.659, vocal: 0.512, consonante: 0.254**. Este inusual papel de la vocal sugiere que la pronunciación de la vibrante en ECM actual tiende más al respeto de la forma canónica y que este proceso de regreso al estándar se produce con cierta independencia del tipo de elemento fónico que le preceda.

La realización breve [r] sólo se produjo en 37 ocasiones (3.6%) en 1970 (en 1033 tokens), de los cuales 21 se produjeron precedidos de vocal. En cambio, en el año 2000 aumentó considerablemente esta variante, ya que representa un 14.8% de los datos (184/1241). El contexto precedente se ordena de la siguiente manera: **consonante: 0.738 > inicio absoluto: 0.690, [s]: 0.512 > vocal 0.438**. Al contar con tan pocos casos en 1970, se vuelve muy difícil definir el modelo multivariable; el número de casos moderno, por otra parte, sí es más robusto y permite concluir que en general una vocal precedente [V_] es el único contexto fónico que desfavorece las variantes no estándar de la vibrante, ya que se encuentra al final de los recorridos logísticos para la asibilada, aproximante y variante breve.

Si se retoman las probabilidades del análisis binomial de 1 nivel para las cuatro variantes de la vibrante (Tabla 3.8), la gráfica queda de la siguiente manera:

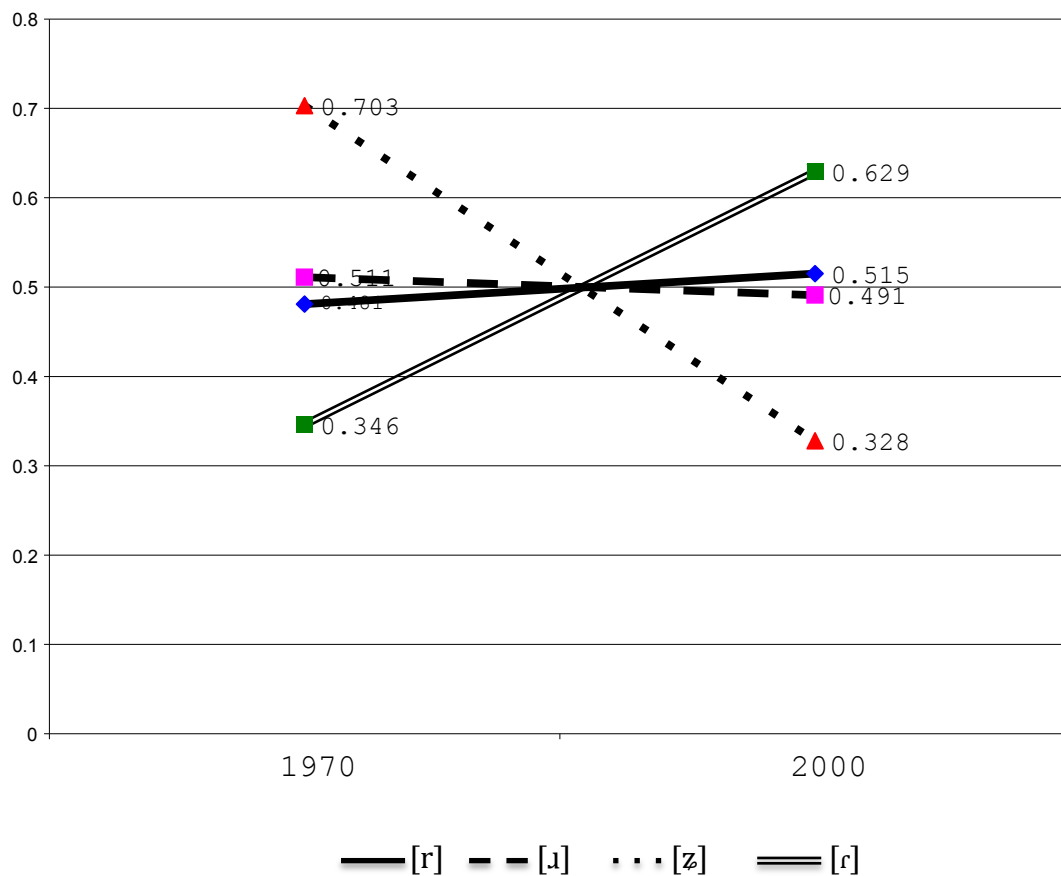


Figura 3.15. Evolución de la variabilidad de la rótica vibrante (r) 1970-2000 (Pb 1, n=2274)

Esta gráfica es muy similar a la presentada para la percusiva. Es claro el retroceso de la asibilada (línea punteada), mientras que las variantes aproximante (guiones) y canónica (sólida) se mantienen más o menos en los mismos niveles. El interés aquí es el comportamiento de la variante breve (línea doble) que aumenta a

.629 su probabilidad de aparición en el año 2000 (de hecho, en términos de frecuencia relativa cuadruplicó su presencia estadística: de 3.6% en 1970 a 14.8% en 2000). Este es un proceso interesante, porque en términos fonológicos implica la permanencia del rasgo [A_r], esto es, aumenta la preferencia por mantener el carácter rótico de la consonante, como sucede también con la percusiva, que gana mucho terreno en tiempos modernos (v. Fig. 3.10).

A continuación se explora el papel de los factores sociales en el proceso de retracción de la asibilación de la vibrante.

Tabla 3.9. *Peso de factores sociales en asibilación de la vibrante (r) 1970 (f=134/1033)*

	<i>Pb 1</i>	<i>Pb U&D</i>
SEXO		
<i>Hombres</i>	.326	.326
<i>Mujeres</i>	.646	.646
ESTUDIOS		
<i>Alto</i>	.447	.447
<i>Bajo</i>	.565	.565
EDAD		
<i>Mayores</i>	.413	.413
<i>Adultos</i>	.756	.756
<i>Jóvenes</i>	.408	.408

Los factores sociales se distribuyen exactamente igual tanto en la vibrante como en la percusiva: las mujeres (.646), personas de mediana edad (.756) y hablantes de pocos estudios (.565) favorecen la asibilada, por lo que puede concluirse que si un informante asibila la (r) prepausal, también lo hará con la vibrante –al

menos en 1970. El análisis de frecuencia relativa acumulativa debería confirmar estos resultados:

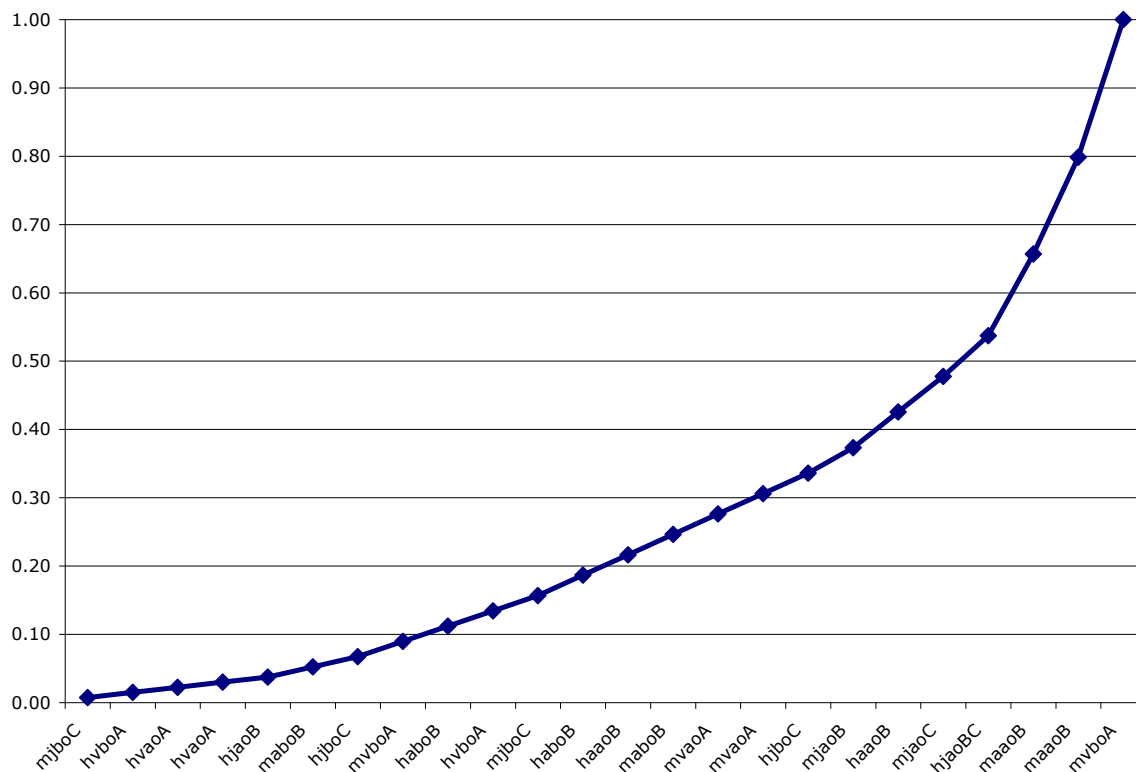


Figura 3.16. Frecuencia relativa acumulativa de la variante asibilada de (r) 1970
($f=134/1033$)

De entrada, todos los hablantes presentaron al menos un caso de asibilación de la vibrante. Entre los diez hablantes que concentran el mayor número de asibiladas, siete son mujeres, una repetición de lo que sucede con la percusiva. De hecho, tres mujeres (dos adultas y una mayor) y un hombre joven concentran el 46% de los datos totales, lo que confirma el carácter predominantemente femenino del fenómeno

aunque, como se ha visto a lo largo del capítulo, es un rasgo compartido por los hombres también, aunque en menores proporciones.

Ahora deben contrastarse estos resultados con lo obtenido en la muestra del año 2000.

Tabla 3.10. *Peso de factores sociales en asibilación de la vibrante en 2000 (f=59/1241)*

	<i>Pb 1</i>	<i>Pb U&D</i>
SEXO		
<i>Hombres</i>	.411	.411
<i>Mujeres</i>	.585	.585
ESTUDIOS		
<i>Alto</i>	.517	.517
<i>Bajo</i>	.418	.418
EDAD		
<i>Mayores</i>	.706	.706
<i>Adultos</i>	.513	.513
<i>Jóvenes</i>	.254	.254

La asociación de la asibilada con el habla femenina (con peso de .585) se confirma nuevamente en el análisis de subida y bajada; se asocia además con el nivel alto de estudios (.577) y con los grupos de hablantes mayores (.706) y adultos (.756). La siguiente gráfica permite identificar a los líderes de la asibilación.

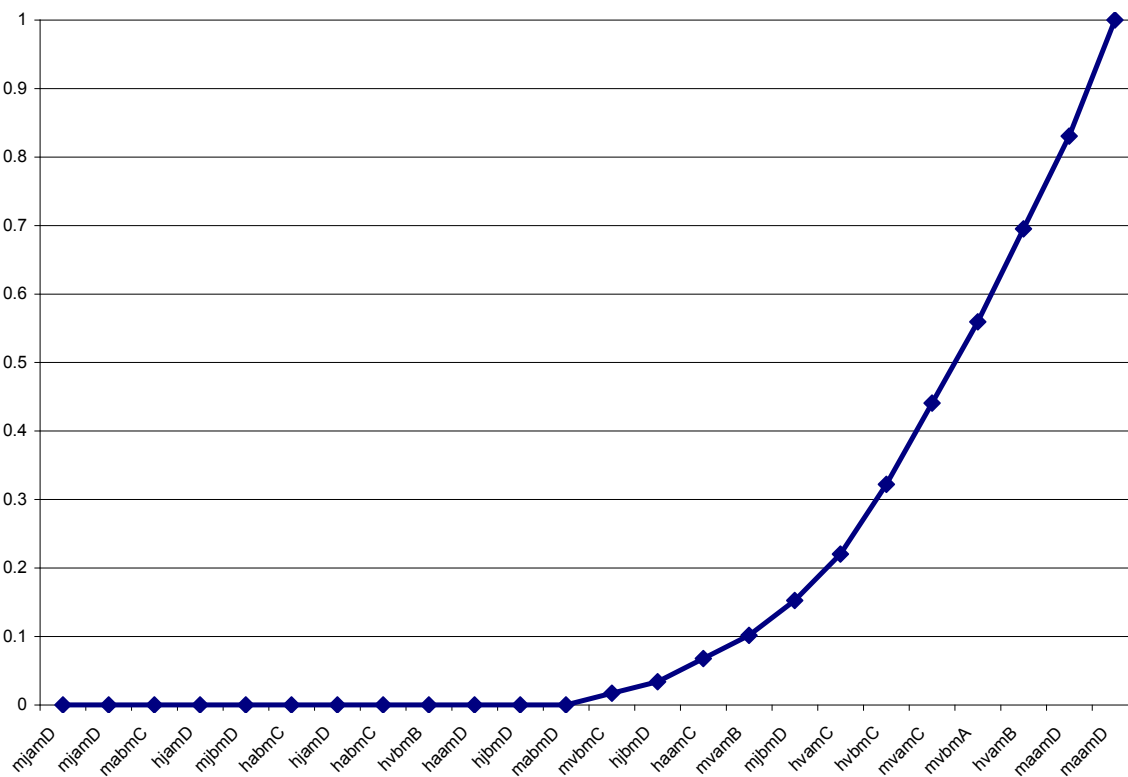


Figura 3.17. Frecuencia relativa acumulativa de asibilación de la vibrante en 2000
($f=59/1241$)

Los primeros 12 informantes en la gráfica no produjeron ni una sola asibilada, el 4.8% de presencia estadística de la variante en el año 2000 indica de entrada la poca productividad del proceso, y por otro confirma la idea de retracción de la asibilación de las róticas en ECM. Punto en común con los análisis anteriores es el predominio de las mujeres al frente de la asibilación: las primeras tres personas son mujeres, dos adultas y una mayor, seguidas de un hombre mayor. Las primeras dos líderes (Edith y Ruth) son de estudios universitarios, de mediana edad y nacidas en los

años 1960, época en que la asibilación tenía mayor presencia en ECM, por lo que puede sugerirse una adquisición temprana del rasgo.

Puede resumirse la asibilación de la vibrante como un proceso encabezado en los años setenta por mujeres de mediana edad y de pocos estudios; en tiempos modernos se desplazó a mujeres mayores y adultas pero de nivel educativo alto. Los resultados cuantitativos sobre variantes asibiladas de la rótica vibrante confirman una tendencia general a la disminución estadística de la variante en ECM. La siguiente gráfica presenta los pesos porcentuales en cada generación de informantes.

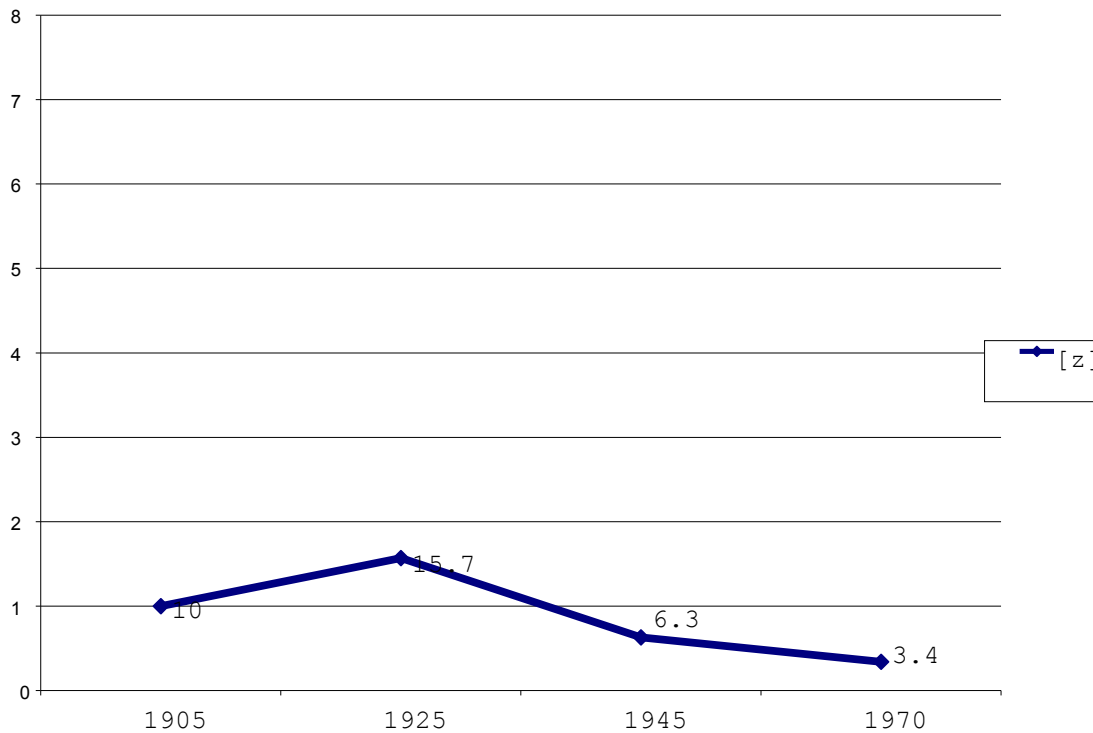


Figura 3.18. Frecuencia relativa de la asibilación de vibrante por año de nacimiento (48 informantes, N=2214)

La línea se ubica en la parte baja de la figura porque precisamente se trata de una variante poco productiva que tiene un comportamiento lineal muy parecido a la de las variantes asibiladas de la percusiva (Fig. 3.13): por un lado comienza en la primera generación con un nivel bajo, pero sube significativamente (en un 57%) a la siguiente generación (los hablantes adultos de los años 1970), para luego descender en las últimas dos generaciones: la generación nacida alrededor de 1970 tiene menos de la cuarta parte de asibiladas que tenía la generación de 1925. Ahora debe explorarse estas mismas generaciones respecto de la variante canónica.

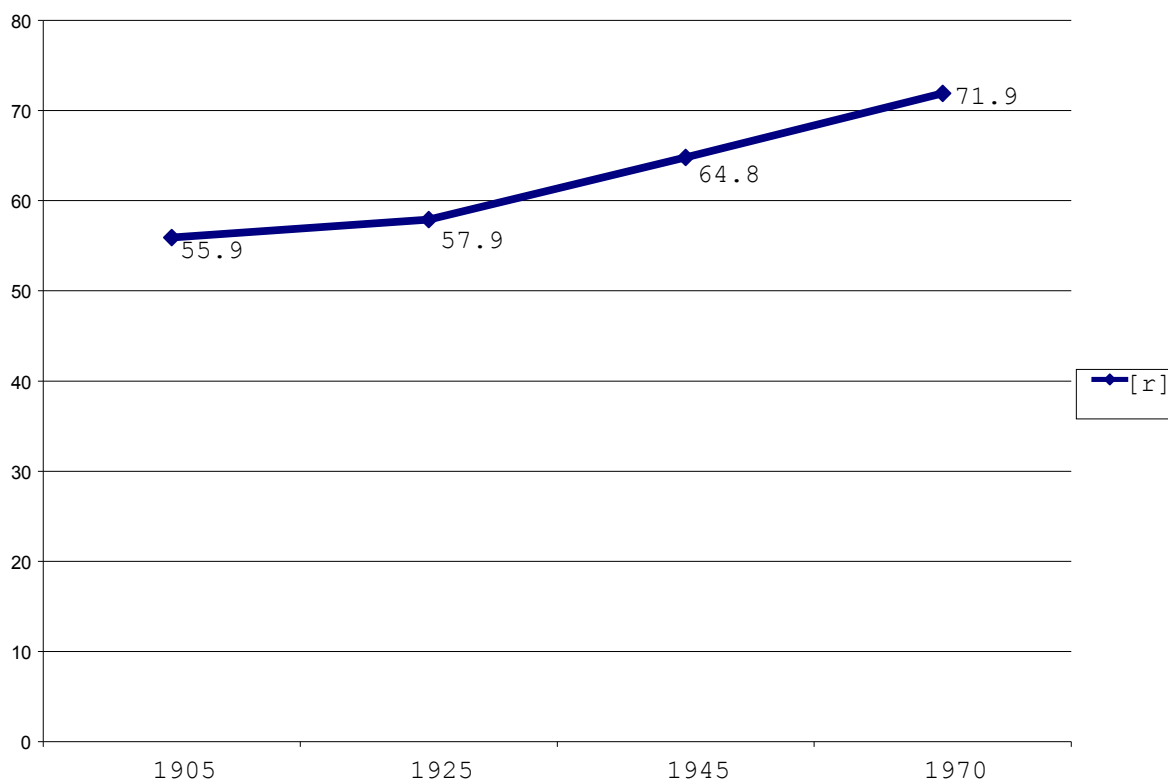


Figura 3.19. Frecuencia relativa de la vibrante canónica por año de nacimiento (48 informantes, N=2214)

La gráfica muestra claramente que en cada generación la vibrante canónica ha ganado terreno, aunque más tímidamente de la primera a la segunda generación (1905 a 1925) porque, recordemos, es esta última la que privilegia un poco más la forma asibilada, pero gana en total 14 puntos porcentuales en las dos generaciones siguientes (los nacidos entre 1945 y 1970). Al igual que sucede con la percusiva, la (r) se comporta con las variantes asibilada y canónica en espejo (en un sentido estadístico).

3.7. FORMALIZACIÓN DE LA VARIABILIDAD DE LAS RÓTICAS EN TEORÍA DE OPTIMIDAD

Para el análisis de la variabilidad en las róticas del español de la ciudad de México (ECM) se emplea aquí el modelo de Teoría de Optimidad Estocástica (TOE) (Boersma & Hayes 2001; Díaz-Campos & Colina 2006; Feldhausen & Vanrell 2014; Martín Butragueño 2014), una adaptación de la Teoría de Optimidad clásica (Prince & Smolensky 1993/2004) para formalizar los datos variables. Como se recordará, en TOE la jerarquía de restricciones está ordenada de forma tal que la variante óptima es la de mayor peso estadístico en el habla de la comunidad lingüística estudiada y la variante menos presente viola un mayor número de restricciones o viola restricciones de mayor peso en la jerarquía, de modo que por ello tiene una menor frecuencia estadística, pero esta nunca llega a cero. Es importante recordar también que una diferencia esencial del presente trabajo con los mencionados sobre TOE, es que en lugar de utilizar algoritmos como el de Boersma & Hayes (2001) para determinar la jerarquía (o *peso*, en *noisy HG*) de las restricciones, aquí se toma en cuenta el peso

probabilístico de las variantes obtenido con el algoritmo VARBRUL (Sankoff, 1988; Sankoff, Tagliamonte & Smith 2005), ya que la curva que presupone este último algoritmo es en forma de *s*, la configuración más documentada del cambio lingüístico (cf. Conde Silvestre 2007). De este modo, en los cuadros que se presentan aquí, *todos* los candidatos son seleccionados en algún momento, por tanto el cuadro debe interpretarse como uno en el que todos los candidatos son *relativamente* óptimos. Se trata de posibilidades reales y no sólo teóricas¹⁷.

Una ventaja del conjunto de restricciones aquí propuesto respecto a otros es que todas las restricciones funcionan en el nivel segmental y se evita el uso de restricciones como *COMPLEX (Martín Butragueño 2008: 181-190) que incorpora el nivel silábico. En un análisis más reciente, Martín Butragueño (2014: 572) propone la restricción *GEM en lugar de *COMPLEX, aunque básicamente se trata de la misma restricción: evítense sonidos geminados asociados a dos elementos en la grada prosódica. Otra ventaja formal del conjunto aquí propuesto es que el análisis unificado de (r) y (r) resulta más económico (sólo cuatro restricciones) frente al propuesto en Martín Butragueño (2014), donde se utilizan cinco restricciones para (r) y 4 para (r); es muy importante aclarar que el autor utiliza distintos conjuntos de restricciones porque, a diferencia del presente estudio, en su base de datos de 54 informantes se

¹⁷ “Es necesario subrayar que se trata de *preferencias*, y no de candidatos seleccionados o excluidos [...] Esta es una diferencia importante entre las representaciones fonológicas categóricas y las variables: los candidatos menos preferidos sí se documentan de todos modos y no quedan excluidos de la gramática” (Martín Butragueño 2008: 164, subrayado en el original).

presentaron muy pocos casos de (r) realizada como vibrante simple [r] (por lo tanto, su análisis da cuenta de tres variantes: canónica, aproximante y asibilada, y no cuatro, como se hace aquí); además, se analiza la percusiva en todos los contextos de aparición (a diferencia de aquí, donde sólo se incluyen los casos de la percusiva ante pausa), por lo que, en sentido estricto, no se está trabajando exactamente los mismos contextos (Martín Butragueño 2014: 571).

Las restricciones que se proponen para el conjunto de datos analizados en esta investigación son: una restricción general de IDENTIDAD I-O subdividida en dos restricciones que exigen mantener el valor de ciertos rasgos fonológicos del input en el output: apertura aproximante [A_m] y del rasgo [Anterior]; las restricciones de marcación que entran en juego son: *FAST (Steriade 1995, *apud* Bradley 2001: 137-138) que penaliza el rápido contorno de apertura que caracteriza a la percusiva (v. Fig. 3.5), y *HOLD (Bradley 2001: 143-147) que prohíbe un inusual alargamiento de las constricciones apicales (y con ello a la vibrante).

(3) UN CONJUNTO DE RESTRICCIONES PARA EL ANÁLISIS UNIFICADO DE (r) PREPAUSAL Y (r)

- IDENT_[Ant] El rasgo [Anterior] en el input permanece idéntico en el output.
- IDENT_[Am] Respete la asignación del rasgo [Am] dada en el input.
- *FAST Evítense transiciones articulatorias inusualmente rápidas en constricciones apicales.
- *HOLD Evítense constricciones apicales inusualmente largas.

Respecto a las restricciones de fidelidad, la restricción IDENT_[Ant] penaliza una modificación en el punto de constricción; en este caso, la única variante que lo infringe

es precisamente la asibilada [ʒ], que tiene un punto de constricción no anterior (cf. Martín Butragueño 2014: 572, 577; de hecho, suele ser alveopalatal). La segunda restricción IDENT_[Am] prohíbe la modificación del valor original asignado respecto al grado de apertura aproximante en el input; la (r) tiene asignada una *secuencia* de apertura [A_m-A_t-A_m], pero la variante aproximante [ɹ] viola esta restricción al realizarse sólo con la apertura [A_m]; las demás variantes no poseen este grado de apertura y por ello no violan esta restricción. En cuanto a las restricciones de marcación, si se asume la descripción fonológica de la percusiva como un contorno de apertura, ésta es la única variante que restringe la restricción *FAST; finalmente el alargamiento de la constricción del ápice de la lengua en la zona alveolar implica la infracción de *HOLD y con ello se descarta a la vibrante como candidata óptima.

3.7.1. Formalización en TOE de la variabilidad de la (r) ante pausa

A partir de los resultados presentados en la Tabla 3.5 (columna *f*), puede representarse la variabilidad de la rótica percusiva en 1970 de la siguiente manera:

Cuadro optimal 3.1. Jerarquía de restricciones para la variable percusiva (r) en 1970

/r /	*HOLD	*FAST	IDENT _[Am]	IDENT _[Ant]
☞ [ʒ]				*
[ɹ]			*!	
[r]		*!		
[r]	*!			

Como se desprende del cuadro, son las restricciones de marcación las que más fuertemente sancionan las distintas variables de la percusiva, de manera que las variantes propiamente róticas son las últimas seleccionadas. En general, podría interpretarse este resultado como la acción de una restricción general que penaliza los segmentos no continuos al final de emisión (de hecho ese es el análisis de Martín Butragueño 2008 y 2014: 572, quien propone la restricción *CODA[-Cont]) produciendo con ello una sílaba más natural para el español, lengua que suele evitar consonantes oclusivas en dicho contexto.

El cuadro optimal es muy distinto en el año 2000. Ahora la restricción Ident_[Ant] cobra peso y marca la asibilada como la última preferencia estadística. Nuevamente, partiendo de la frecuencia relativa de cada variante en dicho año de muestreo, se obtiene el siguiente cuadro.

Cuadro optimal 3.2. Jerarquía de restricciones para la variable percusiva (r) en 2000

/r /	IDENT _[Ant]	*HOLD	IDENT _[Am]	*FAST
[r]				*
[ɹ]			*!	
[r]		*!		
[ʒ]	*!			

La aproximante sigue ocupando el segundo lugar y gana un poco de terreno la variante alargada (vibrante). Sobre todo es claro que sube en la jerarquía la

restricción IDENT_[Ant] y con ello la asibilada deja de ser la variante más común, para ser ahora la de menor presencia estadística.

3.7.2. Formalización en TOE de la variabilidad de la rótica vibrante (r)

Partiendo solamente de la frecuencia relativa de las variantes, se perfila el siguiente orden de restricciones en los años 1970:

Cuadro optimal 3.3. Jerarquía de restricciones para la variable vibrante (r) en 1970

/r/	*FAST	IDENT _[Ant]	IDENT _[Am]	*HOLD
☞ [r]				*
[ɹ]			*!	
[ʒ]		*!		
[ɾ]	*!			

La restricción *FAST es la de mayor peso y marca a la variante percusiva como último candidato preferido, seguido por la asibilada, que viola la identidad de punto de constricción (IDENT_[Ant]). La aproximante, al igual que en los cuadros optimales 3.1 y 3.2, sigue estando poco penalizada, lo que asegura un segundo lugar en las preferencias estadísticas de los informantes. El cuadro optimal para la vibrante en el año 2000 no es muy distinto del cuadro para el año 1970:

Cuadro optimal 3.4. Jerarquía de restricciones para la variable vibrante (r) en 2000

/r/	IDENT _[Ant]	*FAST	IDENT _[Am]	*HOLD
[r]				*
[ɹ]			*!	
[r̥]		*!		
[ʒ]	*!			

En este caso la restricción IDENT_[Ant] es la de mayor peso, por lo que la asibilada es la última elección e intercambia posición con la variante breve. No puede concluirse que las restricciones de marcación e identidad se estén agrupando en uno u otro extremo de la jerarquía. La única variante que sistemáticamente es la segunda elección estadística, en ambas róticas, es la aproximante, por lo que puede afirmarse que en ECM la restricción IDENT_[Am] tiene un lugar relativamente fijo en la jerarquía.

3.8. DISCUSIÓN

3.8.1. Movimientos en la jerarquía de restricciones de las róticas en el tiempo real

En esta sección se explora la posibilidad de unificar el análisis fonológico de [r̥] y [r]. Primero se deben revisar las jerarquías de restricciones y el candidato óptimo seleccionado en cada período.

Cuadro 3.1. Jerarquías de restricciones para las róticas en tiempo real

	1970	2000
(r)	*HOLD >>*FAST >>IDENT _[Am] >>IDENT _[Ant]	IDENT _[Ant] >>*HOLD >>IDENT _[Am] >>*FAST
(r)	*FAST >>IDENT _[Ant] >>IDENT _[Am] >>*HOLD	IDENT _[Ant] >>*FAST >>IDENT _[Am] >>*HOLD

De las cuatro jerarquías, tres resuelven a favor de la variante canónica, lo que de entrada confirma que la asibilación, sobre todo en los años 1960, se producía básicamente en la rótica percusiva prepausal. Al ser la asibilada la variante más alejada del estándar de pronunciación (y sobre todo, las asibiladas sordas), requiere de una jerarquía que asigna un peso menor a las restricciones de identidad (respecto al rasgo [Apertura aprox] y el punto de constricción). Respecto a la vibrante, puede verse que el respeto a la postura articulatoria de esta rótica (representado por el débil peso de *HOLD) predomina en 1970 y 2000.

Una vista rápida a estas jerarquías sugeriría que no hay agrupación de las dos clases de restricciones (marcación e identidad). Sin embargo, puede experimentarse con ciertos artificios analíticos para encontrar una sistematicidad que en realidad sí existe. Por ejemplo, si se asigna un valor numérico que pondere el peso de cada restricción en la jerarquía pueden encontrarse resultados esclarecedores. Asígnese <1> a la restricción más baja en la jerarquía, <2> a la siguiente, y así sucesivamente. Los valores quedarían como en las siguientes tablas:

Cuadro 3.2. Peso relativo de restricciones para la rótica percusiva prepausal [r/]

	*HOLD	*FAST	IDENT _[Am]	IDENT _[Ant]
1970	4	3	2	1
2000	3	1	2	4

Cuadro 3.3. Peso relativo de restricciones para la rótica vibrante (r)

	*HOLD	*FAST	IDENT _[Am]	IDENT _[Ant]
1970	1	4	2	3
2000	1	3	2	4

Respecto a la percusiva prepausal (r), en 1970 las restricciones de MARCACIÓN pesan más (7 puntos) que las de IDENTIDAD (3 puntos), pero para el año 2000 se invierte la tendencia: ahora las restricciones de IDENTIDAD son las más pesadas (6 puntos), frente a 4 puntos de las restricciones de MARCACIÓN. Respecto a la vibrante (r) el panorama siempre ha mostrado una tendencia a preservar la pronunciación estándar: 7 puntos de las restricciones de IDENTIDAD frente a 5 de las de MARCACIÓN en 1970; en el año 2000 se mantiene la tendencia: 6 puntos en IDENTIDAD frente a 4 en MARCACIÓN.

Otra manera de analizar los cuadros 3.2 y 3.3 es haciendo una suma de los puntajes totales de cada restricción. Por ejemplo, la restricción *HOLD en la vibrante simple (Cuadro 3.2) suma 7 puntos totales, ya que estaba en el primer lugar en 1970 (por eso recibe 4 puntos) pero en el año 2000 cae al segundo puesto (y por ello recibe 3 puntos).

Si se suma el peso de las restricciones de MARCACIÓN por un lado y de IDENTIDAD por otro, el panorama general independientemente del tiempo real respecto a la percusiva es que pesan más las restricciones de MARCACIÓN (11 puntos en conjunto) que las de IDENTIDAD (9 puntos) y esta relación se invierte respecto a la vibrante: las restricciones de IDENTIDAD pesan más que las de MARCACIÓN, exactamente en las mismas proporciones: 11 puntos para IDENTIDAD, 9 para MARCACIÓN. Aunque es discutible, quizás estas tendencias reflejen el peso de un factor funcional: específicamente, que la vibrante debe ser más robusta para poderse diferenciar adecuadamente de la percusiva, sobre todo en el contexto intervocálico donde sí tienen oposición fonológica (Bradley, 2000: 9). Por otra parte, recuérdese que el contexto final de emisión (donde se produce la neutralización de la percusiva y la vibrante) no suele favorecer la percepción de contrastes fonológicos --cf. la noción de contrastes fonológicos “permitidos por pistas” (*licensed by-cue*) de Steriade (1997).

Conjuntando los pesos por año de muestreo, independientemente de la variable que se trate, resulta que 1970 es un año que privilegia estructuras menos marcadas (12 puntos frente a 8) y que en 2000 se revierte la tendencia y ahora son las restricciones de identidad las que más pesan en la jerarquía, con 12 puntos también. Estos pesos relativos, aunque obtenidos de manera artificial, confirman la idea del regreso a la pronunciación más conservadora de las róticas del ECM en un lapso de aproximadamente 30 años.

3.8.2. Sociolingüística de las róticas en el tiempo real

El papel de las mujeres, los grupos de edad, el nivel educativo y sobre todo la clara asociación de la asibilación de ambas róticas con los mismos grupos de hablantes permite también unificar el análisis sociolingüístico de estas consonantes. Lo interesante ahora es contrastar los perfiles sociolingüísticos de las variantes más contrastantes en términos tanto fonéticos como sociales: la variante asibilada y la variante canónica. Los siguientes cuadros presentan los resultados de regresión logística para las variables sociales en ambas variantes.

Cuadro 3.4. Perfil sociolingüístico de la asibilación de róticas en ECM (1970-2000)

	1970	2000
/ɾ/ → [ʒ]	Mujeres (.747)	Mujeres (.764)
	Adultos (.712)	Mayores (.583) Alto (.655)
/r/ → [ʒ]	Mujeres (.646)	Mujeres (.585)
	Adultos (.756)	Mayores (.706),
	Bajo (.565)	Adultos (.513) Alto (.577)

Cuadro 3.5. Perfil sociolingüístico de las róticas canónicas en ECM (1970-2000)

	1970	2000
/ɾ/ → [ɾ]	Hombres (.641)	Hombres (.630)
	Bajo (.567)	
/r/ → [r]	Jóvenes (.630)	Hombres (.545)
	Alto (.593)	Jóvenes (.695),
		Adultos (.533) Alto (.554)

Aunque no todas las variables entran en juego, sobre todo respecto a la vibrante, es claro el contraste en al menos dos variables: sexo y grupo de edad. Mientras que las mujeres adultas asibilaban más las róticas en 1970, los hombres jóvenes preferían la pronunciación estándar. La asibilada y la canónica son promovidas por personas de mayores estudios en el año 2000, pero pervive la asibilación como pronunciación femenina en general.

3.8.3. Hacia una explicación de la retracción de la asibilación

¿Por qué pierde terreno la asibilación? Existen algunas rutas de explicación. Por un lado, la asibilación, como toda variación lingüística en proceso de cambio, empezó a cobrar interés en la comunidad hablante desde la época del primer muestreo. Varios colegas que se formaron como lingüistas en la ciudad de México en los años 1960 me han comentado que era un rasgo fonético censurado por sus profesores, ya que se consideraba “habla afectada”, de “señoras riquillas” o simple rasgo de “afeminamiento”¹⁸.

Todavía no contamos con estudios sistemáticos sobre las actitudes subjetivas ante esta variante, en buena medida porque, como señalan L&MB (2003): “[l]a variación lingüística asociada a (r) y a (rr) no ha formado parte hasta ahora de la conciencia lingüística explícita de los hablantes. No es algo de lo que se hable, de lo que se escuchen opiniones o se dejen oír reconvenciones. La escuela no da ninguna recomendación abierta a los hablantes” (2003: 26). Sin embargo, la reciente

¹⁸ Mis colegas Ángeles Soler y Cecilia Rojas coinciden en esto.

estigmatización de la asibilación de róticas en los medios de comunicación masiva como la televisión es un indicio de que el fenómeno está pasando al plano consciente de la comunidad hablante (cf. Eller 2013).

Como ya se ha mencionado, recientemente la asibilada se utiliza como un estereotipo del habla de hombres homosexuales en programas televisivos de alta audiencia. Se trata, de hecho, de retratos cómicos y denigrantes de homosexuales (v. *supra*). Si es verdad que la asibilada se está asociando a este grupo altamente estigmatizado en la sociedad mexicana, el alejarse de dicho patrón de pronunciación sería la opción esperable en la población heterosexual que mantiene prejuicios negativos hacia los homosexuales. Siguiendo la argumentación basada en la discriminación de género, tampoco puede descartarse la posibilidad de que la variante pierda peso actualmente por considerarse un rasgo femenino, en una sociedad mexicana que discrimina altamente a las mujeres (cf. Pflieger Biering 2009)¹⁹. En este caso, se trataría de una retracción por estereotipo negativo, situación que se atestigua también en Chambers & Trudgill (1998):

¹⁹ Por supuesto, estas hipótesis se basan en factores que podrían denominarse “perceptuales”: la asociación que se hacía (y se hace) del rasgo en términos de percepción, en este caso, con mujeres de clase media alta o alta y con hombres homosexuales, no coincide totalmente con lo que se encuentra aquí: sí hay una preferencia de las mujeres por la asibilada, pero el grupo educativo bajo es el que lo promueve, no el alto. Por otra parte, no debe ignorarse el hecho de que las mujeres típicamente son *líderes* de los procesos de cambio, sobre todo aquellos vinculados con el prestigio (Labov 2006).

If very stigmatised forms become stereotyped in this way, it may only be a matter of time before they disappear altogether, as the New York [əɪ] pronunciation appears to be doing [...]. When changes of this kind take place, reversing as they often do the original direction of a linguistic change, they can be called CHANGES FROM ABOVE, i.e. from above the level of conscious awareness [Chambers & Trudgill 1998: 76; versalitas en el original].

Otra posibilidad que ya se había mencionado en un trabajo anterior (Serrano 2008 a) es que probablemente era rasgo compartido por la generación trabajadora y profesionalista en los años 1970 como marca de identidad, ya que la generación adulta suele estar más preocupada por la integración lingüística –esto si seguimos un modelo como el de *mercado lingüístico* (Bourdieu 1984, 1990; Milroy & Gordon 2003) en el que las interacciones comunicativas tienen un valor “de mercado” en la sociedad. El rasgo pervivió en dicha generación pero no se heredó completamente a la siguiente. L&MB (2003), al resumir sus hallazgos sobre la asibilación moderna de las róticas, plantean la siguiente hipótesis:

¿Qué tienen en común entonces las personas que asibilan? Nuestra hipótesis [...] es que es el *ascenso social*, en la medida en que se considere posible o deseable, con independencia de que se haya conseguido o no, el motor que mueve a muchos de los informantes que encabezan el uso de asibilaciones. Estos informantes son, de manera típica, mujeres, muchas de

ellas de origen humilde, trabajador o semirrural [...] involucradas en la consecución de objetivos muy específicos que las llevan a superarse por encima del entorno inmediato que las rodea. Cambio que puede retraerse, o desvanecerse incluso, una vez que las posiciones sociales se afianzan, como ocurre en varios casos en que las hijas no siguen en este aspecto a sus madres [subrayado mío].

Los resultados de esta investigación permiten apoyar esta hipótesis. Por supuesto, hacen falta entrevistas más detalladas con los informantes y empezar a rastrear, a través de estudios de actitudes o de percepción, el nivel de conciencia sociolingüística de este interesante fenómeno sociofonético.

CAPÍTULO 4

DEBILITAMIENTO DE VOCALES ÁTONAS

Entre los rasgos característicos del español de la ciudad de México se destaca el debilitamiento de sus vocales átonas, proceso conocido en la bibliografía tradicional como de “vocales caedizas”. Si bien la distensión vocálica se produce en otras variantes geográficas del español (Navarro Tomás 1963: 44; Moreno Fernández 2000) como Ecuador (Lipski 1990) y especialmente en el Perú (Delforge 2008), es llamativo el grado que alcanza en español de la ciudad de México (ECM), donde llega incluso a la elisión total del segmento (Henríquez Ureña 1921 [1976]; Lope Blanch 1963-1964; Moreno de Alba 1994; Serrano 2006; Lastra & Martín Butragueño 2003; Martín Butragueño 2014). Este capítulo abordará los antecedentes de investigación, la descripción fonética del proceso y propondrá un análisis sociolingüístico variacionista a partir del habla de los 48 informantes que conforman la base de datos principal de

esta tesis. Tras el análisis, se presentará una discusión sobre los patrones sociolingüísticos que sigue el proceso y una formalización en términos de Teoría de Optimidad Estocástica.

4.1. DEBILITAMIENTO VOCÁLICO EN ESPAÑOL DE LA CIUDAD DE MÉXICO: ANTECEDENTES DESCRIPTIVOS

Al menos desde el trabajo pionero de Charles C. Marden (1896 [1938]) se ha llamado la atención sobre la variabilidad en las vocales del español de la ciudad de México (ECM). Marden, en su *Fonología del español de la ciudad de Méjico* de 1896 --traducido por Pedro Henríquez Ureña y publicado en 1938 en la Biblioteca de Dialectología Hispanoamericana--, menciona la aféresis (elisión) vocálica y precisa que “el caso más común de aféresis es la caída de la *a* inacentuada inicial en verbos de más de dos sílabas: *atrasado* > **trasáu**, *apetecer* > **petece.i**, *arrastrar* > **řastra.i**, *arrancar* > **řaŋka.i** [...]” (Marden 1896 [1938]: 103)¹. Aunque Marden no reporta debilitamiento de las vocales átonas en forma de ensordecimiento o elisión, sí abunda en el fenómeno del cierre de /e/ y /o/ como un rasgo característico de esta variedad (1896 [1938]: 112-118). Sin embargo, Henríquez Ureña (quien había hecho algunas observaciones sobre debilitamiento de las vocales átonas en 1921) aclara en la nota 2 del capítulo III de Marden: “Observación general: las vocales de la altiplanicie central de Méjico se diferencian de las castellanas en que resultan comúnmente más breves; las

¹ Todos los ejemplos citados se transcriben respetando las convenciones de transcripción fonética del original (las de la *Revista de Filología Española* posteriores a 1915).

consonantes se destacan y adquieren predominio: *bloques para apuntes* > *blocs pr'apunts*; *pase usted* > *pas-sté*; *cómo está usted* > *cómo stá sté*. Sobre este asunto hace falta estudio especial” (Marden 1896 [1938]: 112, n. 2). Esto es, el fenómeno llamaba la atención al menos desde los inicios del siglo xx.

Otras observaciones del fenómeno de debilitamiento y elisión vocálica se encuentran en los trabajos de Boyd-Bowman (1952), y Canellada & Zamora Vicente (1960), quienes confirman el rasgo. Sin embargo, dichos trabajos estaban basados en impresiones aisladas de los autores. Para corroborar y corregir dichas afirmaciones, Juan M. Lope Blanch (1963-1964) recogió datos de 52 informantes grabados magnetofónicamente, de los que transcribió de 20 a 30 minutos; añadió anotaciones y observaciones espontáneas de otros 48 informantes, completando con ello los datos de 100 personas quienes --comenta Lope Blanch-- son “informantes pertenecientes a las más diversas clases sociales y culturales: personas de ambos sexos y de muy diferentes edades (entre los 15 y los 60 años), cuya habla representa la casi totalidad de las distintas modalidades lingüísticas que pueden escucharse en la ciudad de México” (Lope Blanch *ibid.*: 2). Los hallazgos de Lope Blanch, a partir de un total de 2,284 ejemplos, son múltiples y de gran interés, y de hecho constituyeron la base de una investigación previa a la aquí presentada (Serrano 2006), en la que, en lo esencial, se recuperaban las variantes principales que Lope Blanch mismo detectaba: relajación intensiva, vocal debilitada, vocal mínima y elisión. Dicho autor encuentra en general que “el debilitamiento o pérdida de las vocales no depende básicamente de la posición silábica que la vocal guarde con relación al acento principal de la palabra, sino del

entorno consonántico que la envuelva” (Lope Blanch *ibid.*: 7); también menciona la influencia decisiva de la consonante fricativa sorda coronal /s/ (pp. 8-9, 12), y de la posición prepausal, sobre todo a final de emisión (p. 18). Otras observaciones fonético-articulatorias de Lope Blanch se recuperarán en la descripción de estas variantes *infra*. También es muy importante destacar que el autor no hallaba patrones sociolingüísticos del debilitamiento:

Como resumen, puedo anticipar que en la debilitación y pérdida de las vocales no se aprecia disposición sistemática alguna por lo que al nivel sociocultural de los hablantes se refiere. Como bien observaron ya Ma. J. Canellada & A. Zamora [1960], el fenómeno *puede* aparecer “en todas las clases sociales, desde personas de cultura superior hasta los últimos estratos” (p. 225); y lo mismo en adolescentes y jóvenes que en personas maduras o ancianas. Junto a un estudiante que debilite las vocales hasta el máximo y a cada momento, otro de su misma edad y condición social las mantendrá todas con igual firmeza que un castellano. No hay, pues, sistematización social y cultural de ninguna clase. Si me viera obligado a declarar en qué ámbito *parece* producirse el fenómeno con alguna regularidad o intensidad, tal vez pudiera decir que los materiales por mí reunidos apuntan vagamente hacia personas jóvenes de cultura media o superior [Lope Blanch *ibid.*: 4; subrayado en el original].

Dado que en dicho trabajo se reportaba la total generalización del fenómeno y que no se encontraba ninguna asociación clara con alguna variable social, resulta de especial interés observar si la variabilidad del debilitamiento vocálico en ECM ha evolucionado en un período de más de 30 años y si existe ahora (o no) una asociación del fenómeno con factores sociales².

Otro trabajo que aborda el problema del debilitamiento vocálico pero en la variedad del Perú es el de Delforge (2008), quien pone especial atención al proceso fonético acústico y articulatorio del debilitamiento, sin atender a factores sociolingüísticos. En la introducción a su detallado análisis en términos de fonología articulatoria (Browman & Goldstein 1989, 1990), Delforge (2008: 107) comenta que: “Spanish [unstressed vowel reduction] appears to be a relatively homogeneous process, exhibiting the same basic characteristics in Mexico and the Andes”. Entre las características que son similares en ambos países, destaca que los efectos del debilitamiento son graduales y variables, que existe una tendencia a concentrarse en las vocales medias /e/ y /o/ y en contacto con /s/, y finalmente, que se presenta más en las sílabas a final de palabra (Delforge 2008: 107). La investigación de Delforge se basa en el habla de entrevista de 16 informantes: 14 hombres y dos mujeres,

² Aunque la reducción vocálica en ECM no parece estar muy presente en la conciencia sociolingüística de la comunidad, es posible encontrar ejemplos de reflexión sobre el proceso. En un artículo de opinión capitalino (*Metro*, 4/mayo/2010, p. 20) encontré la siguiente expresión: “¡Misión cumplida! A sus or’densss, jefe!”, donde se resalta la elisión de la segunda vocal /e/ en la palabra *órdenes* (quizás el autor quiso plasmar algo como [ór.de.ns:]).

hablantes que resultaron “frequent vowel reducers” (Delforge 2008: 109) dentro de una muestra de 180 personas de la ciudad andina de Cusco. Una diferencia sustancial respecto a la investigación de esta tesis (y con Lope Blanch 1963-1964) es que Delforge elimina del conteo las palabras *pues*, *entonces* y *digamos*, de alta presencia estadística y que típicamente presentan vocales debilitadas. Esta decisión analítica tendrá un impacto en la comparabilidad de los resultados, como se verá más adelante. Por el momento se pueden destacar algunos de los hallazgos más importantes de Delforge –basados en el análisis espectrográfico de 16,581 vocales átonas, entre las cuales detecta 1,648 debilitadas. Que por un lado, existe una tendencia al ensordecimiento en sílaba átona y en contexto de consonantes sordas; que las vocales /e/ y /o/ son las más susceptibles al ensordecimiento; que no se trata de un fenómeno de centralización (o “reducción”, como se llama tradicionalmente al fenómeno), sino principalmente de distintos grados de ensordecimiento, y que no existe un efecto significativo de la velocidad del habla³.

³ Para una discusión profusa de los hallazgos de Delforge (2008) véase Martín Butragueño (2014: 168-173). Respecto a la centralización de las vocales átonas en habla espontánea, Poch Olivé, Hermegnies & Martín Butragueño (2008), al comparar habla de laboratorio y de entrevista en dos informantes (de España uno, de México otro), encuentran que, en contraste con el habla de laboratorio, el habla de entrevistas promueve una menor distancia de las vocales en el espacio formántico. Esta tendencia a la menor cardinalidad y mayor traslape de los timbres vocálicos en habla de entrevistas sociolingüísticas se confirmó también en un estudio de Martín Butragueño (2014: 145-163) con ocho informantes de ECM.

En Serrano (2006) se hace el análisis sociolingüístico de este proceso en el habla de 12 informantes. A partir de una base de 847 ejemplos (obtenidos en 30 min de transcripción y la lectura de un texto de media cuartilla de extensión por informante), Serrano confirma básicamente los hallazgos de Lope Blanch sobre el contexto fónico del debilitamiento vocálico (consonantes sordas, especialmente /s/) y también la poca correlación con factores sociales, aunque ahora la variable *sexo* juega algún papel en las reglas variables trazadas en dicho trabajo: los hombres debilitan más que las mujeres; además, el debilitamiento es susceptible de cancelarse en el estilo de lectura en voz alta (Serrano 2006: 54).

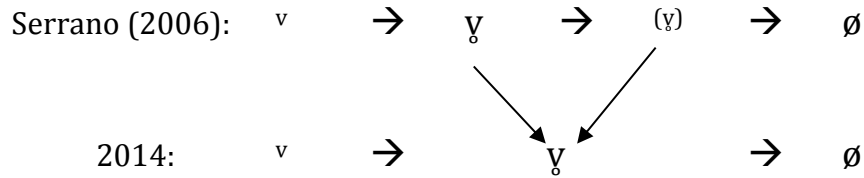
Por su parte, el reciente análisis sociolingüístico y dialectológico de Martín Butragueño (2014: Cap. 2) explora el problema a la luz de los datos del *Atlas Lingüístico de México* (Lope Blanch [dir.] 1990, 1992). Puede destacarse que dicho autor, además de confirmar lo reportado en la bibliografía aquí reseñada sobre el importante papel del fonema /s/ y el contexto final de emisión ([Vs#]) en el debilitamiento (Martín Butragueño 2014: 232-233), también encuentra que el proceso está significativamente asociado con hombres y con personas de nivel educativo bajo (2014: 239-240). Por otra parte, en un modelo de sólo cinco zonas geográficas, la región centro-este del país (Distrito Federal, Puebla, Tlaxcala, Oaxaca y Veracruz) es la que concentra mayores índices de debilitamiento (2014: 226, Cuadro 2.16) y que, en un modelo más fino, de nueve zonas, es precisamente la zona “México central” la que más debilita (2014: 228, Cuadro 2.17), lo que permite confirmar el rasgo como propio de la comunidad de habla estudiada en esta tesis.

Este breve recuento ha permitido contar con un panorama lingüístico y social general del proceso de debilitamiento vocálico, que se irá discutiendo con los datos de esta tesis. Como en el capítulo anterior, en las secciones siguientes se presentará una discusión sobre la fonética y fonología del proceso de debilitamiento vocálico y de las variantes consideradas, se discutirán los resultados del análisis cuantitativo y se pondrá especial atención en el papel de la frecuencia de uso y en el perfil de los líderes lingüísticos de la elisión vocálica.

4.2. FONÉTICA Y FONOLOGÍA DEL DEBILITAMIENTO VOCÁLICO EN ESPAÑOL DE LA CIUDAD DE MÉXICO

En trabajos anteriores se ha definido el proceso de debilitamiento de las vocales átonas como una afectación en el carácter sonoro o sordo de la vocal y/o en su duración (Lope Blanch *ibid.*; Serrano 2006; Delforge 2008). En un trabajo previo a la presente investigación (Serrano 2006), consideré cuatro variantes debilitadas: vocal relajada [v] de menor intensidad general, con o sin ensordecimiento), vocal completamente sorda [\bar{v}], vocal mínima [v^{m}] (poca duración y completamente ensordecida) y elisión [\emptyset], ello para poder realizar comparaciones con los trabajos de los años sesenta. Ahora puede resultar más interesante organizar las variantes en un continuum más simple, de mayor a menor sonoridad, que culmine en la elisión de la vocal. Por tanto, de las cuatro variantes originalmente transcritas en este corpus, se hizo una agrupación de las variantes sordas (las llamadas “sorda plena” y “vocal mínima” en Serrano 2006). Véase el siguiente esquema:

AGRUPAMIENTO DE VARIANTES DEL DEBILITAMIENTO VOCÁLICO



Las variantes ahora son sólo tres y se ordenan en un *continuum* que inicia en una vocal relajada pero no necesariamente sorda, le seguiría una variante sorda (con duración normal o reducida) y finalmente la elisión. A continuación se describen las características acústicas y articulatorias de cada una de estas variantes.

4.2.1. Variante relajada (o reducida) [v̥]

La realidad acústica de las variantes de vocales debilitadas se exploró con dos informantes jóvenes (hombre y mujer) de estudios universitarios. Se recogieron algunos ejemplos de las variantes que interesan aquí y se analizaron a través del *Computer Speech Lab* (Kay Elemetrics) con datos de entrevistas sociolingüísticas (para otros ejemplos, se utilizó *PRAAT*). Un ejemplo de variante relajada se presenta a continuación:

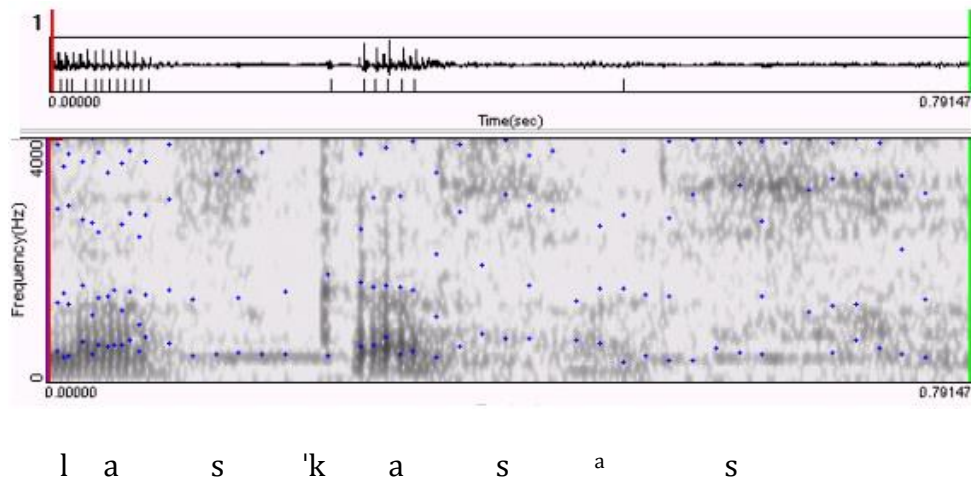


Figura 4.1. Oscilograma y espectrograma de la frase *las casas* que muestra una vocal relajada (ancho de banda: 4,000 Hz)

La última vocal *a* presenta sólo un pulso glótico (visible en el oscilograma superior) que contrasta con los cinco pulsos de la vocal acentuada en la sílaba ['ka], presenta una estructura formántica mucho menos definida que las dos precedentes y la duración es un poco más breve que en las demás. Coincidiría en lo general con la vocal que Delforge (2008: 111) llama *partially devoiced/shortened*. Martín Butragueño (2002: 88) propone una representación fonético-fonológica en la que se analiza la reducción de la duración de la vocal en el nivel moraico. La vocal breve llevaría asociado un valor moraico igual a *y* (un valor entre 0 y 1); la consonante siguiente absorbería la pérdida de duración aumentando su propio valor en moras mediante la operación *2-y*, lo que permite obtener siempre un valor de 2 moras para el núcleo y coda silábica de una palabra, como *pasos* pronunciada como [pá.s²s]. La representación de Martín Butragueño (2002: 88) es la siguiente:

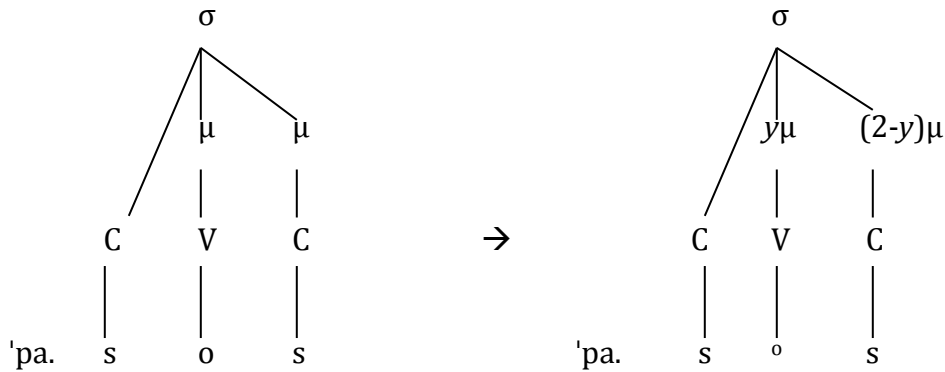


Figura 4.2. Representación fonológica de la pérdida de duración de vocales átonas

De esta manera, aunque se perdiera material fónico, la sílaba mantendría su peso moraico original. Dicho modelo se basa también en la observación de dicho autor (a partir del análisis del habla de un informante masculino, joven, originario de la ciudad de México) de que parece existir un efecto de alargamiento compensatorio de la segunda [s] en estructuras como *pesos*, *pasos* o *cosas*, en las que la vocal debilitada se encuentra entre dos fricativas sordas (2002: 72-81). El autor aclaraba, sin embargo, que hacían falta más datos para confirmar esta hipótesis. En su más reciente investigación, el autor dedica un importante apartado (Martín Butragueño 2014: 176-188) a la cuestión de la reducción de la duración de las vocales átonas y la interacción con el acento y la entonación como parte del complejo fenómeno del debilitamiento. Sus datos, obtenidos del *Corpus sociolingüístico de la Ciudad de México* (CSCM) confirman que la menor duración puede formar parte del esquema general de

debilitamiento vocálico, el cual implicaría también algún nivel de ensordecimiento y de centralización⁴.

4.2.2. Variantes sordas [ɥ]

Las variantes sordas, por otra parte, presentan total ausencia de pulsos glóticos, pero pueden tener o bien una duración estándar (lo que se llamaba “sorda plena” en Serrano 2006: 38) o reducirse al grado de resultar prácticamente imperceptibles (las vocales llamadas “mínima o reducida” de Lope Blanch [*ibid.*: 5] y “mínima” en Serrano 2006: 39). El ensordecimiento se representaría fonológicamente como la asociación del rasgo del nodo Laríngeo (en este caso [-sonoro]) del segmento sordo adyacente a la vocal afectada o, alternativamente, como la disociación del rasgo [+son] de la vocal. Se presenta un ejemplo en la siguiente figura.

⁴ “La complejidad y las diferencias entre cada uno de los ejemplos parecen argumentos para, en el futuro, considerar las correlaciones del debilitamiento vocálico a partir de varios parámetros acústicos independientes (además de por medio de la estructura prosódica), y no a partir de variantes “empaquetadas” y obtenidas por medio de métodos perceptuales. Otra de las preguntas fundamentales es si el fenómeno del debilitamiento es ante todo segmental o prosódico. Aunque normalmente se trata de sílabas átonas, y su colaboración en la portación del tono es modesta, también la hipoarticulación puede impedir que se cumplan con claridad todos los objetivos tonales” (Martín Butragueño 2014: 188).

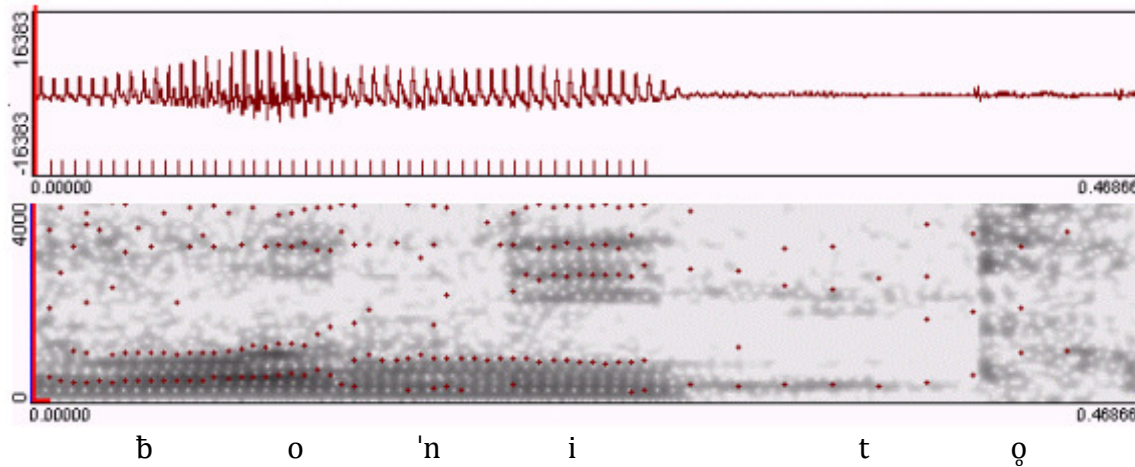


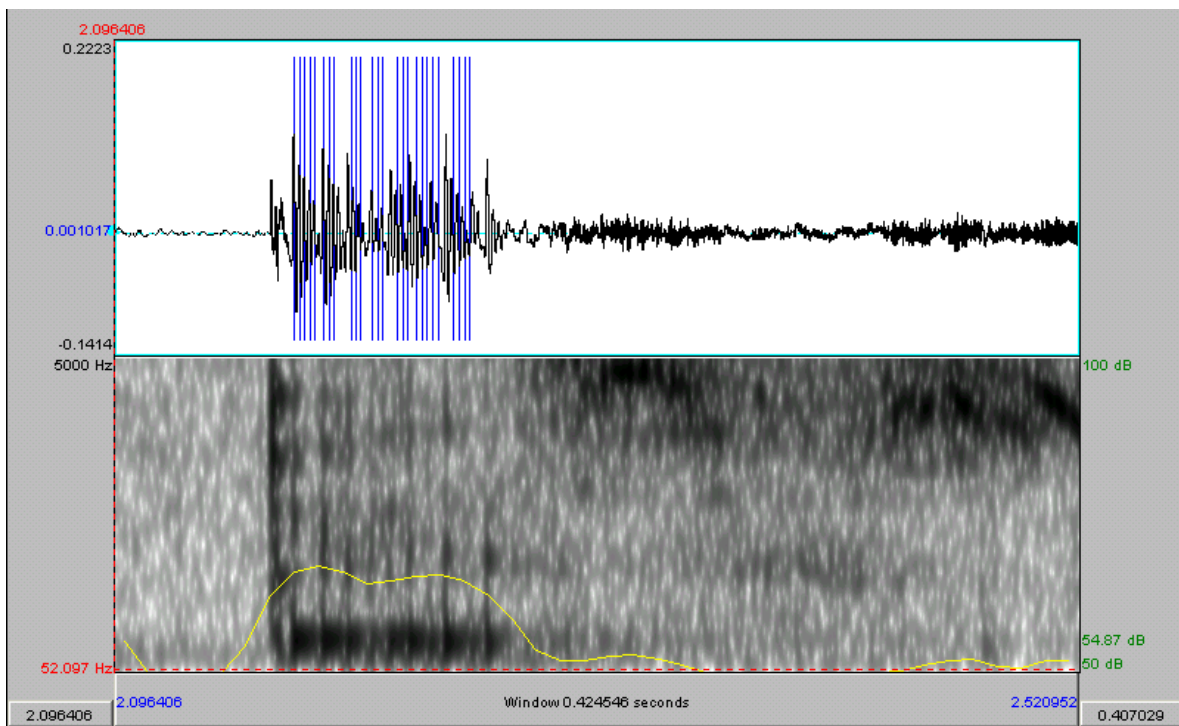
Figura 4.3. *Oscilograma y espectrograma de la palabra bonito con vocal final sorda*

Las manchas horizontales en la vocal sorda parecen vestigios de F_1 y F_2 . Al quedar rastros de estructura formántica, quizás esto permitiría a un oyente distinguir la cualidad de la vocal ensordecida --aspecto que se comenta un poco más adelante respecto a la elisión. Para Ohala (1983) el ensordecimiento de vocales puede deberse entre otras razones a: 1) la asimilación parcial con segmentos sordos adyacentes⁵ y la pausa (las cuerdas vocales se pueden separar un poco); 2) una presión subglotal reducida que se presenta en las vocales inacentuadas o en las que siguen a una consonante con una gran corriente de aire, como [p^h]; 3) la brevedad de la vocal, ya que entre más corta, es más susceptible al ensordecimiento (Ohala 1983: 203).

⁵ En el trabajo de Ohala (1990) sobre asimilación se sugiere que, en general, lo más probable es encontrar patrones en que se propague la sordéz antes que la sonoridad (1990: 266), justo lo que ocurre en ECM.

Coincidirían estas variantes con las que Delforge (2008: 111) denomina *completely devoiced*.

El siguiente ejemplo de vocal entre segmentos fricativos sordos [s] deja aún más claro el carácter sordo de esta variante y permitirá contrastarlo con la variante elidida en el mismo contexto (espectrogramas realizados a través de PRAAT⁶):



'p e s o s

Figura 4.4. Oscilograma y espectrograma de la palabra pesos con vocal /o/ sorda.

(Ancho de banda: 5,000 Hz)

⁶ Dado que estaba analizándose las vocales, el ancho de banda en estos espectrogramas se fijó a 5000 Hz, aunque debe reconocerse que las vocales sordas, al estar rodeadas de fricativas, se analizan mejor con un ancho de banda de 11,000 Hz.

En este ejemplo se aprecia el espacio sordo entre ambas fricativas y vestigios del F₂. Dentro de la variante sorda se contaron todos los ejemplos que se han descrito como “vocal mínima” porque son también sordas, aunque prácticamente no dejan huella en el espectrograma. Normalmente estas no permiten distinguir su timbre vocálico, algo que también señalaba Lope Blanch (*ibid.*), quien las describe como “vocal mínima *indiferenciada*, de timbre impreciso: [p^(e)stola], ‘pistola’, [ermán^{es}], ‘hermanos’, [est^(e)s personas], etc.” (Lope Blanch *ibid.*: 5, n. 14; énfasis y transcripción como en el original).

4.2.3. Variante elidida

El grado máximo de debilitamiento vocálico en ECM es el cero fonético. Compárese la Figura 4.4. con la siguiente, donde se muestra, en un ejemplo distinto de la misma palabra <pesos> y con el mismo informante, lo que se entiende en esta tesis como una ‘elisión total de la vocal con rastros articulatorios vocálicos en los segmentos consonánticos adyacentes’:

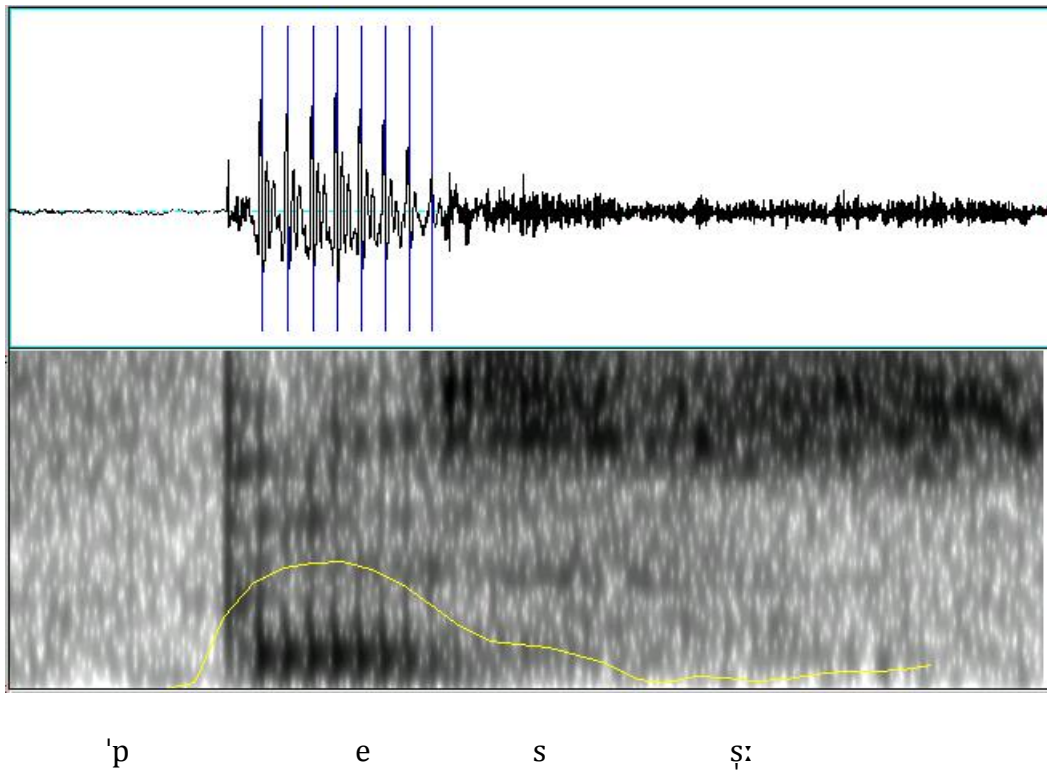


Figura 4.5. Oscilograma y espectrograma de la palabra pesos con elisión total de la vocal /o/. (Ancho de banda: 5,000 Hz)

Si se pone atención en la parte superior del segmento correspondiente a [s] en la figura, existen dos manchas de intensidad de la fricción, una al inicio seguida de una sección un poco más clara y otra al final (variaciones de intensidad marcadas también por la curva de intensidad, que sube un poco hacia el final del espectrograma), que sugieren una articulación de [s] en dos momentos, esto es, una especie de [s] rearticulada. Esta pronunciación había sido descrita casi en los mismos términos por Lope Blanch (*ibid.*) y Boyd-Bowman (1952), quienes la caracterizan como

una interrupción breve, pero evidente; su articulación no coincide con la de una *s* larga, sino que se acerca más a la de las consonantes geminadas [...]. Que el resultado de *svs* no es una sola *s* larga ya lo había advertido Boyd-Bowman: “Cuando hay encuentro de dos *s*, una es siempre silábica. En *es’s*, *gras’s*, *entons’s* no tenemos una *s* larga y uniforme (**entons:*), sino *dos*: la división silábica entre ellas va marcada por un ligero y brevísimo relajamiento de la sibilación, *sin que esta deje de ser continua y sorda*. Con una *s* fuerte y continua (**es*, **gras*;) las palabras quedarían extrañamente alteradas (n. 5) [Lope Blanch *ibid.*: 13].

Este aspecto fonético deberá analizarse en el futuro con detalle, pero independientemente del proceso exacto de su articulación, es evidente que la elisión de la vocal entre sibilantes produce secuencias fonológicamente inusitadas en otras variantes del español, como este ejemplo, donde la [s] tendría que constituir el núcleo silábico⁷. Mi hipótesis respecto a la conservación del sentido de la palabra con vocal elidida es que esta última deja un rastro articulatorio en la sibilante para que el hablante pueda reconstruir la vocal original; en algunos casos puede ser una

⁷ Para Francisco Arellanes (comunicación personal), más bien se trataría de una [s] “más robusta” que no constituye núcleo silábico pero que sí deja pistas al oyente para poder determinar la pieza léxica en cuestión. Para Perissinotto (1975: 26-33) --contrariamente a todo lo reportado en estudios previos-- resulta claro que **no** existen procesos de elisión total. Según reporta dicho autor, esto lo descubrió a partir de un breve ejercicio de fonética perceptual que realizó con 20 hablantes a quienes presentó frases con elisiones totales de vocal que se volvían ininteligibles (1975: 28).

pronunciación de [s] acompañada de redondeamiento de labios (v.g.: *pesos*) y en otros, de estiramiento de labios (como en *peces*).

Para poner a prueba esta hipótesis, en Serrano (2010 b) se analizaron 24 casos de elisión con datos obtenidos en el campo y otros ocho datos de habla de laboratorio. Se obtuvieron *spectral slices* de la parte central del segmento [s] dejado tras la elisión para ver si existía una correlación entre la frecuencia pico y el tipo de vocal elidida. Aunque los resultados son todavía muy preliminares, al menos las secuencias /sVs/, como en *cosas* o *peces*, muestran que la frecuencia pico de la [s] dejada tras una elisión de [o] es menor (5766 Hz, promedio de 5 casos) a la de las [s] dejadas tras elisión de [e] (6413 Hz, promedio en 3 casos), lo que puede sugerir que, en efecto, la vocal deja un rastro articulatorio que permite identificar correctamente la palabra (un par mínimo puede ser <besos> /'be.sos/ vs. <veces> /'be.ses/). Por supuesto, hacen falta muchos más datos y un estudio de percepción. Sin embargo, de confirmarse estas tendencias, permitirían entender también por qué se permite la elisión total de segmentos vocálicos en ECM.

Finalmente, la elisión vocálica puede producir también grupos consonánticos muy complejos en los cuales no parecen quedar rastros vocálicos. En el siguiente espectrograma se muestran dos elisiones de la vocal en la sílaba final: en la palabra <cuatrocientos> y en la palabra <pesos>; la que interesa es la primera de estas elisiones:

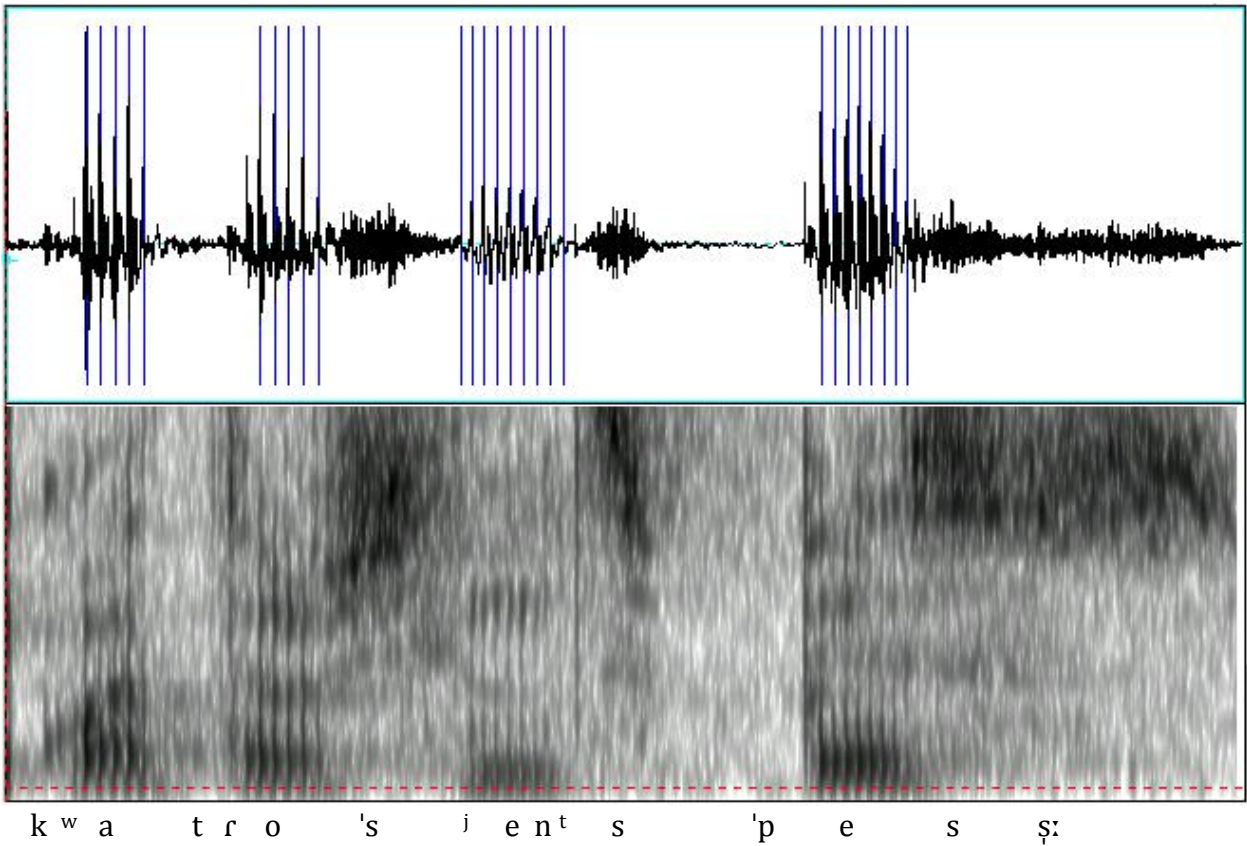


Figura 4.6. *Oscilograma y espectrograma de la frase cuatrocientos pesos. (Ancho de banda: 5,000 Hz)*

Llama la atención la brevísima duración de todos los elementos que se encuentran en coda en la secuencia [nts]: entre los tres suman una duración de 85 ms. Esto contrasta de forma muy interesante con lo reportado en Rodríguez & Mora (2001: Cuadro 4) en su trabajo sobre las consonantes en coda silábica en la misma variedad que estudiamos, quienes encontraron que un grupo de dos consonantes

sordas en coda tiene en promedio una duración de 166 ms⁸. En cambio, la secuencia [sʃ:] en *pesos* dura 238 ms, frente a los 44 ms que dura la primera [s] de la palabra *cuatrocientos*⁹.

Hasta aquí la descripción acústica y articulatoria de las variantes consideradas. En las siguientes secciones se detallan los factores internos tomados en cuenta para el análisis cuantitativo, los resultados y la discusión sobre la fonologización del proceso en términos de Teoría de Optimidad Estocástica.

4.3. VARIABLES INTERNAS EN EL ANÁLISIS DEL DEBILITAMIENTO VOCÁLICO

En un estudio previo de las vocales átonas en la ciudad de México (Serrano 2006) con una muestra de 12 informantes -distinta de la utilizada en esta tesis- se había considerado como variable independiente el acento de las sílabas adyacentes; los resultados mostraron que el contexto acentual que favorece el debilitamiento vocálico es [‘σ__σ] (Serrano 2006: 45, Tabla 5), de modo que palabras como *cómete* serían ideales para el debilitamiento en la sílaba intermedia [me]. Sin embargo, el nuevo grupo de datos (con 48 informantes) arrojó muy pocos casos de debilitamiento en

⁸ Por supuesto que aquí debe considerarse que los datos de Rodríguez & Mora (2001) son de habla de laboratorio y lectura de textos, estilos que producen un habla de tempo más lento que el habla espontánea de estas entrevistas.

⁹ Además de estos casos de rastros de material vocálico en [s], algunos ejemplos transcritos sugieren que la vocal puede dejar rastros también en la consonante previa (a través de sonorización) en secuencias como *pues sí* que se realizan como [p̥.sí] o [p̥.zí]. Nuevamente, estos casos deben analizarse acústicamente con sumo detalle.

palabras de tres o más sílabas. En un estudio con los primeros 24 informantes del presente corpus –658 casos totales (Serrano 2008 b)--, el 98% de casos de debilitamiento (648) se presentó ante [s] y (casi siempre) pausa, como en ['es.tos/], o ante pausa simple ['ka.sa/], por lo que se decidió trabajar solamente con ejemplos en contexto prepausal.

Dicha decisión analítica provocó que el factor tonicidad resultara superfluo, porque si se toma en consideración que la mayoría de las palabras en español son de acento paroxítono (graves), entonces el valor [+acento] en la penúltima sílaba habría resultado excesivamente predominante en la base de datos, lo que podría llevar a interpretaciones erróneas sobre el papel de dicho factor.

De este modo, el análisis fónico se simplificó para atender solamente los factores que: a) se presentaran de manera abundante en el corpus; y b) que hubieran mostrado un papel importante en estudios previos.

Las variables internas quedaron codificadas de la siguiente manera:

Cuadro 4.1. *Variables internas para el análisis del debilitamiento vocálico*

Variante	Contexto anterior	Contexto posterior	Factor léxico
1=relajada	k= cons. sorda	s= [s]	n= poco frecuente
2= sorda	g= cons. sonora	p= pausa	l= frecuente
4= elisión			

Dado que el papel de la frecuencia de uso es importantísima en los procesos de desgaste fónico como el que aquí se analiza (Bybee & Hopper 2001; Pierrehumbert

2001; Coetzee 2008), se elaboró una lista de palabras y morfemas considerados frecuentes o que ya se encuentran prácticamente lexicalizadas con una forma abreviada. En el trabajo de Lope Blanch (*ibid.*) se mencionan: “*antes, pesos, gracias, entonces, nosotros, tardes, precisamente, pues, casas, otros, morfema verbal -mos, sintagmas que está, que es*” (*ibid.*: 6, n. 18); a estas formas se añadieron los demostrativos *eso(s), esa(s)*; el morfema *-cientos*; palabras en diminutivos *-ito/a(s)* y aumentativos *-ote/a (s)*, y las palabras *noche(s), coche(s), bonito/a (s), veces, meses y parte (s)*.

4.4. CRITERIOS DE TRANSCRIPCIÓN AUDITIVA

En Serrano (2006) se realizó una transcripción auditiva que no distinguía entre debilitamiento debido al decaimiento de la energía articuladora a final de emisión y aquél que se producía en contextos de articulación plena. Por lo tanto, los índices de debilitamiento (en 12 informantes) en 2006 son llamativamente más altos que en el estudio con 24 hablantes reportado en Serrano (2008 b) y en esta tesis. En el primero (Serrano 2006: 44) el promedio fue de 63 vocales debilitadas por informante, mientras que en Serrano (2008 b) baja el promedio a 27 y para esta tesis el promedio es de 29. Debe aclararse que no se trata de los mismos informantes y que, según mi impresión, la espontaneidad del habla en el estudio de 2006 es mayor que en los otros dos, lo que también contribuye a estas diferencias tan marcadas en el índice de debilitamiento general¹⁰.

¹⁰ En la siguiente tabla se comparan los resultados de Serrano (2006) con los de esta tesis:

En esta ocasión se decidió poner atención en los debilitamientos que *no* fueran propiciados por el fenómeno de relajamiento articulatorio propio de los finales de emisión, para tener una visión más parecida a la que buscaba Lope Blanch en su ensayo de los años sesenta (*ibid.*).

Por otra parte es importante señalar que, contraviniendo el principio de exhaustividad o de “rendición de cuentas” (*principle of accountability*, Tagliamonte 2006: 12-13), no se transcribieron los casos de *no* debilitamiento. Como ya se comentó *supra*, en realidad fuera del contexto prepausal las vocales átonas no suelen debilitarse en índices importantes, por lo que se habría obtenido una cantidad

Serrano (2006)			2014	
[v]	83	(9.7%)	54	(3.8%)
[ʋ]	449	(53.0%)	760	(54.1%)
[ʷ]	124	(14.6%)	272	(19.3%)
[∅]	191	(22.5%)	317	(22.6%)
Totales	847	(100%)	1403	(100%)

A pesar de que los criterios de transcripción fueron un poco diferentes, llama la atención que la distribución de variantes es muy similar en ambos trabajos. Nótese que en Serrano (2006) se reporta un 53% de vocales sordas y ahora el porcentaje es de 54.1%; un resultado prácticamente idéntico respecto a la elisión: 22.5% en 2006, frente a un 22.6% en 2014. El único valor diferente (y esperable) es que en la transcripción de 2014 son menos los casos de la variante relajada (3.8% frente al 9.7% reportado en 2006). Por lo tanto, puede afirmarse que la tendencia estadística se ha repetido a pesar de las diferencias en los muestreos, lo que a su vez indica que podemos confiar en que el proceso de transcripción ha sido adecuado.

abrumadora de casos de no debilitamiento. Se prefirió en este estudio concentrar la atención en las restricciones fonéticas de detalle que intervienen, por ejemplo, en las elisiones, lo que permitirá abrir la discusión sobre los procesos de cambio fonético y la interacción de factores sociales en los mismos.

4.5. EL CORPUS TOTAL DE VOCALES DEBILITADAS: UN PRIMER ANÁLISIS

En total se transcribieron 1403 ejemplos de vocales debilitadas en el corpus de 24 horas de grabación transcribiendo todos los ejemplos surgidos en media hora por cada uno de los 48 informantes. Como ya se comentó, el promedio es de 29 casos por hablante, bastante menor al promedio de 44.3 ocurrencias en los 52 informantes transcritos por Lope Blanch (*ibid.*)¹¹. Estas diferencias, sin embargo, no son tan grandes si se compara sólo con los datos de 1970 en esta tesis, ya que estos informantes de dicho muestreo presentan 35.3 de vocales debilitadas (producto de dividir 847 casos entre 24 informantes). Por supuesto, siempre existirá la posibilidad de que los criterios impresionísticos para determinar una vocal “debilitada” hayan sido algo distintos (en este caso, más exigentes en esta tesis). Sin embargo, los criterios de transcripción aplicados son los mismos para ambos muestreos, lo que asegura la comparabilidad.

¹¹ Este promedio se obtiene de dividir los 2,284 datos totales entre 52 informantes. Hay que recordar que en dicho recuento Lope Blanch incluyó a tres informantes españoles que tenían más de 20 años en la ciudad de México, quienes presentaban cierto grado de adopción del fenómeno (Lope Blanch *ibid.*: 2-3, n. 5).

Realizar un análisis del corpus total permite tener un panorama muy general del fenómeno. A continuación se analiza la variabilidad vocálica en el corpus total para poder dar una idea de los pesos de cada factor y el efecto del tiempo en los mismos.

Las tres variantes organizadas en el *continuum* de sonoridad-sordez arriba propuesto se distribuyen de la siguiente manera:

Tabla 4.1. *Vocales debilitadas en 48 hablantes de ECM*

Variante	N	f
[^v]	54	(3.8%)
[V.]	1032	(73.6%)
[Ø]	317	(22.6%)
Totales	1403	(100%)

De entrada es claro que, de producirse una vocal debilitada, esta será predominantemente sorda (73.6%) o se elidirá (22.6%), mientras que en muy pocas ocasiones reducirá su duración (3.8%). Delforge (2008: 112) reporta 56% de variantes totalmente sordas (*completely devoiced*), 25% de vocales débilmente sonorizadas, 10% de vocales *partially devoiced/shortened* –que se corresponden aproximadamente con la variante “relajada” de esta tesis) y sólo 9% de vocales *apparently elided* (Delforge 2008: 111-112), lo que sugiere en realidad un gran parecido en cuanto a las proporciones de las variantes sordas, que en ambos estudios son las de mayor peso estadístico. Al contar con tan pocos casos de la variante relajada (sólo 54), en algunos de los análisis en *GoldVarb* se decidió sacar esta variante y

realizar los recorridos de probabilidad solamente con dos variantes: a) vocal con cierto nivel de ensordecimiento y b) cero fonético.

La siguiente tabla compara la distribución de las variantes debilitadas en el tiempo real:

Tabla 4.2. *Debilitamiento vocálico en ECM 1970-2000 (N=1403)*

	1970		2000	
[^v]	83	(6.8%)	10	(1.3%)
[v.]	449	(68.9%)	586	(77.5%)
[Ø]	191	(24.3%)	160	(21.2%)
Totales	847	(100%)	756	(100%)

El índice global de debilitamiento vocálico es menor en el año 2000 (pasó de 847 ejemplos en 1970 a 756 en 2000, una disminución del 10.7%), lo que ya de entrada sugiere una tendencia general a la preferencia por las vocales plenas en tiempos modernos. Dentro de las variantes debilitadas, es llamativo el avance de las variantes sordas en un plazo de 30 años, de más de ocho puntos porcentuales. En sentido contrario, las variantes relajadas descienden muchísimo (de 6.8 a 1.3%). Finalmente, el porcentaje general de elisión baja 3 puntos (de 24 a 21%), pero sigue siendo una característica importante del ECM. Gráficamente las tendencias quedan de esta manera:

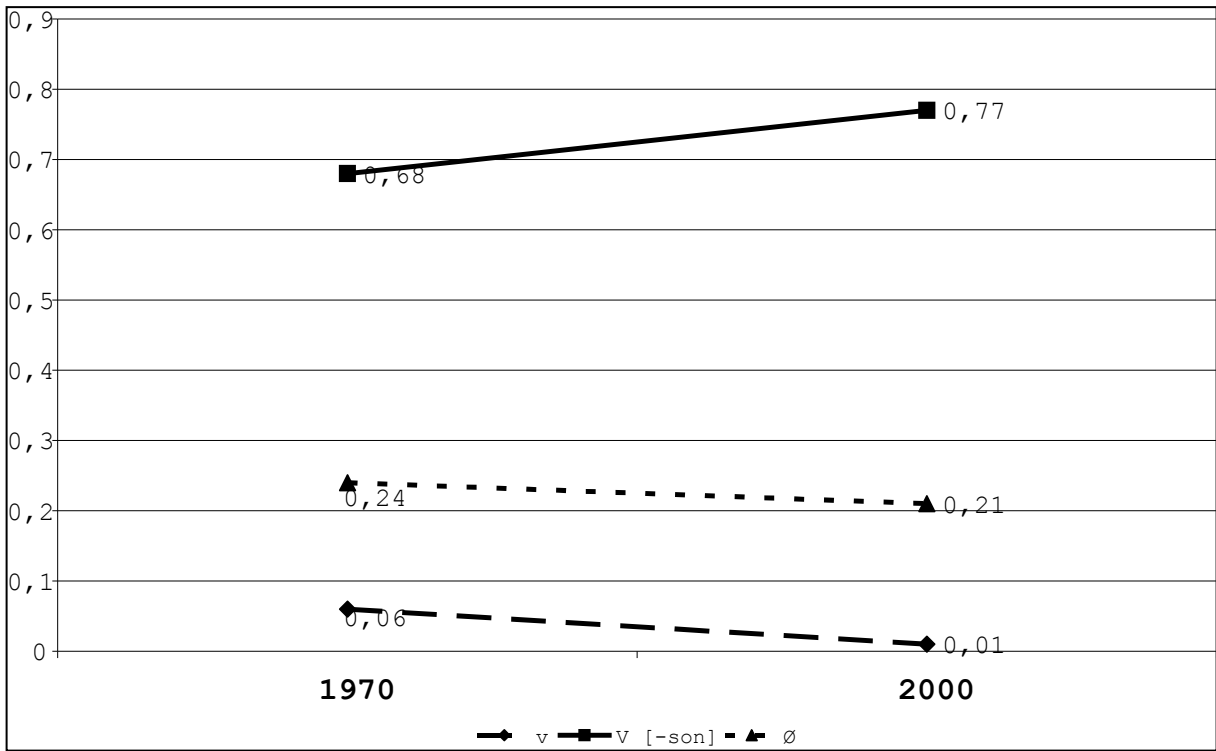


Figura 4.7. Evolución del debilitamiento vocálico 1970-2000 (N=1403)

La siguiente tabla añade los resultados de probabilidad general de cada variante según el año de muestreo en un análisis sobre el corpus total, esto es, se trata de responder a la pregunta: “¿qué peso tiene el año de muestreo en la selección de cada variante?”.

Tabla 4.3. *Peso del año de muestreo en el debilitamiento vocálico (N=1403)*¹²

		1970			2000		
<i>Input</i>		<i>f</i>	<i>Pb 1</i>	<i>Pb U&D</i>	<i>f</i>	<i>Pb 1</i>	<i>Pb U&D</i>
[<i>v̥</i>]	(.021)	6.8	.722	.722	1.3	.307	.306
[<i>v̥</i>]	(.853)	68.9	.442	.443	77.5	.550	.549
[<i>∅</i>]	(.054)	24.3	.492	-	21.2	.507	-

Son muy pocos los datos de la variante relajada (54), pero aun así los análisis binomial de un nivel y de subida y bajada (en adelante *Pb1* y *PbU&D*, respectivamente) permiten confirmar la tendencia a la baja en los datos modernos (pasó de .722 en 1970 a .306 en 2000). Un poco más tímidamente (siempre cerca de .500), pero el análisis de probabilidad señala a los informantes del año 2000 como más favorecedores de las variantes sordas. Finalmente, el año de muestreo no es significativo para la elisión, como lo muestra su no selección en el análisis de regresión escalonada (*Bin U&D*) y el valor de .507 para el año 2000 en el análisis binomial de un nivel.

¹² Hay que aclarar que todos los cálculos de probabilidad en este capítulo (y el anterior) son *multivariantes*: a menos que se indique lo contrario, se incluyen todas las variables en los recorridos, tanto internas como externas. Para los cálculos de esta tabla se ignoraron las variables 'Año de nacimiento' y 'Edad' para evitar el sesgo estadístico que produciría la inclusión de tres variables vinculadas al tiempo.

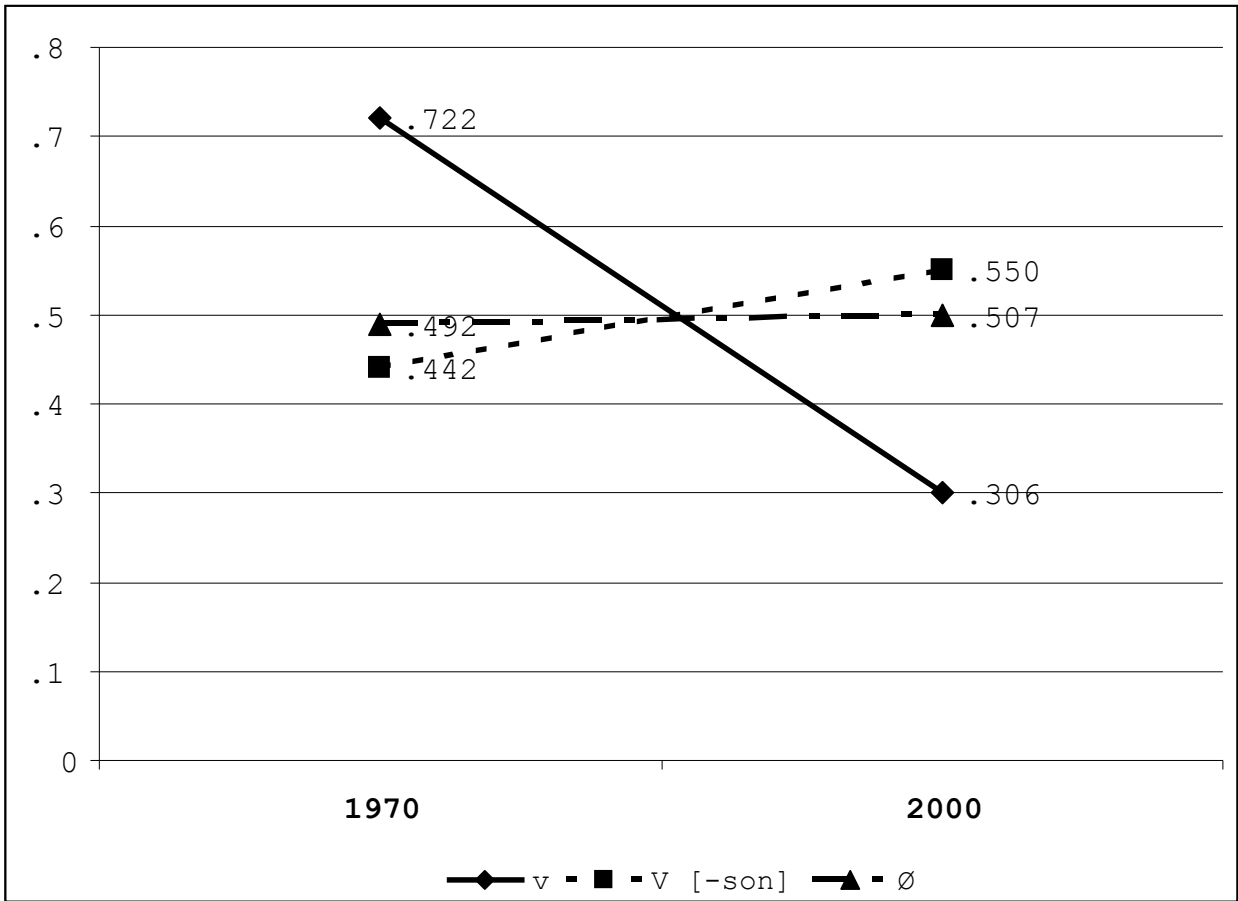


Figura 4.8. Evolución probabilística del debilitamiento vocálico 1970-2000 (N= 1403)

El 33.5% de los ejemplos recogidos en la presente investigación (471/1403) formaban parte de un elemento frecuente. La siguiente tabla compara la distribución de las variantes según se trate de elementos frecuentes o no frecuentes.

Tabla 4.5. Debilitamiento vocálico y peso del factor léxico (N=1403)

	Frecuentes		No frecuentes	
[^v]	14	(3%)	40	(4.3%)
[V.]	201	(42.7%)	831	(89.2%)
[Ø]	256	(54.4%)	61	(6.5%)
Totales	471	(100%)	932	(100%)

La distribución diferenciada de las variantes permite asegurar que existe una correlación con la frecuencia de uso: mientras que la mayoría de las palabras no frecuentes terminó en una vocal sorda (89.2% de las ocasiones) y con muy pocas elisiones (6.5%), las palabras frecuentes presentan como variante mayoritaria la elisión (54.4%). Dicho de otra manera, de todas las elisiones vocálicas en el corpus (317), el 80% (256) se produjo en palabras frecuentes, correlación que se confirmó claramente con el análisis de probabilidad (v. *infra*). También los factores internos están vinculados a la frecuencia de uso:

Tabla 4.6. Contexto fonético y factor léxico en el debilitamiento vocálico

	Frecuentes		No frecuentes	
C[-son]__	460	(97.7%)	644	(69.1%)
C[+son]__	11	(2.3%)	288	(30.9%)
Totales	471	(100%)	932	(100%)
__//	88	(18.7%)	583	(62.6%)
__[s]	383	(81.3%)	349	(37.4%)
Totales	471	(100%)	932	(100%)

La tabla muestra una fuerte correlación estadística entre el carácter más frecuente de la palabra y un contexto fonético específico: consonante sorda precedente y una [s] después. Por supuesto, la lista de elementos que sirvió para clasificar los ejemplos contiene una mayoría de palabras con esos contextos (*cosas, muchos, -cientos, pues*, morfema plural, etc.) y a ello se debe también dicha tendencia. La elevada proporción de estas palabras con una [s] final (81.3%), sin embargo, no está tan claramente determinada por la lista; más bien, puede responder a una tendencia general en español a que una gran proporción de finales de emisión sean en [s], segmento muy productivo a final de palabra en esta lengua. Se confirma en todo caso el patrón reportado por Lope Blanch (*ibid.*: 8-9): “La consonante que más favorece el debilitamiento y pérdida de cualquier vocal es, sin duda alguna, la /s/, como ya había advertido muy atinadamente Boyd-Bowman [...]. De los 2,284 casos registrados, 2,036 presentan contacto anterior o posterior de la vocal con /s/, lo cual equivale a casi el 90% del total absoluto”.

4.6. PESO DE LOS FACTORES LINGÜÍSTICOS Y SOCIALES SOBRE EL DEBILITAMIENTO VOCÁLICO

En este apartado se comentan los resultados del análisis VARBRUL de regresión logística en cada variante, atendiendo sobre todo a su evolución en el tiempo real.

4.6.1. Variante relajada

De la variante relajada sólo surgieron 54 casos en la muestra total. Sin embargo, se realizaron recorridos de regresión logística para tener una idea de las tendencias

generales respecto a esta variante. El único factor interno que motiva la relajación se encuentra en distribución complementaria con aquellos que favorecen el ensordecimiento y la elisión: el carácter sonoro de la consonante precedente (.753), por lo que palabras como *sigo* o *nada* pueden ser candidatas a presentar su segunda vocal de forma relajada.

Por otra parte, de los factores externos, el análisis de subida y bajada sólo selecciona dos: la edad (jóvenes: .614; adultos: .532) y el sexo, en este caso, masculino (.637). En la muestra del año 2000 sólo se registraron 10 casos, por lo que es inútil hacer análisis más refinados. Lo único que podemos mencionar es que ocho ejemplos se dieron ante pausa y siete en palabras no frecuentes.

4.6.2. Variante sorda

Al contar con un mayor volumen de datos, la matriz sociolingüística respecto a esta variante es mucho más clara, aunque diferenciada según se exploren los años 1970 o 2000. De acuerdo con el análisis de subida y bajada ($p=0.000$), en el habla de 1970 inciden dos factores internos: contexto subsecuente (la pausa: .731) y que no se trate de una palabra frecuente (.606). El único factor social seleccionado es el nivel socioeconómico: los hablantes de mayores estudios favorecen claramente el ensordecimiento (.765).

Para el año 2000 el perfil sociolingüístico se vuelve más complejo, porque ahora inciden varios factores:

Tabla 4.4. Modelo sociolingüístico de las *variantes sordas* [ɣ] en el año 2000 (p=0.000)

VARIABLE	Pb U&D
CONTEXTO POSTERIOR	
Pausa	.790
[s]#	.216
FACTOR LÉXICO	
Frecuente	.212
Poco frecuente	.654
SEXO	
Mujeres	.619
Hombres	.431
EDAD	
Mayores	.544
Adultos	.396
Jóvenes	.553
ESTUDIOS	
Alto	.565
Bajo	.433

Se repiten los factores de 1970 (pausa, palabras no frecuentes y hablantes de mayores estudios) y se añade el sexo (mujeres: .619) y la edad de los informantes (jóvenes [.553] y mayores [.544] favorecen ahora las variantes sordas).

Como se vio en el capítulo anterior, el análisis en *tiempo aparente* (Labov 1996: Cap. 3) a través de los grupos de edad nos permite sugerir hipótesis interesantes sobre el posible devenir de un proceso de cambio lingüístico en marcha. Sin embargo, este análisis debe complementarse con los de tiempo *real* (Labov 1996: Cap. 4). El agrupamiento de los informantes por su año de nacimiento es una forma alternativa de analizar el paso del tiempo. Desde esta perspectiva se asume que, en lo general, los hablantes mantienen a lo largo de toda su vida los patrones fonéticos del *vernacular* adquirido. Véase la siguiente gráfica:

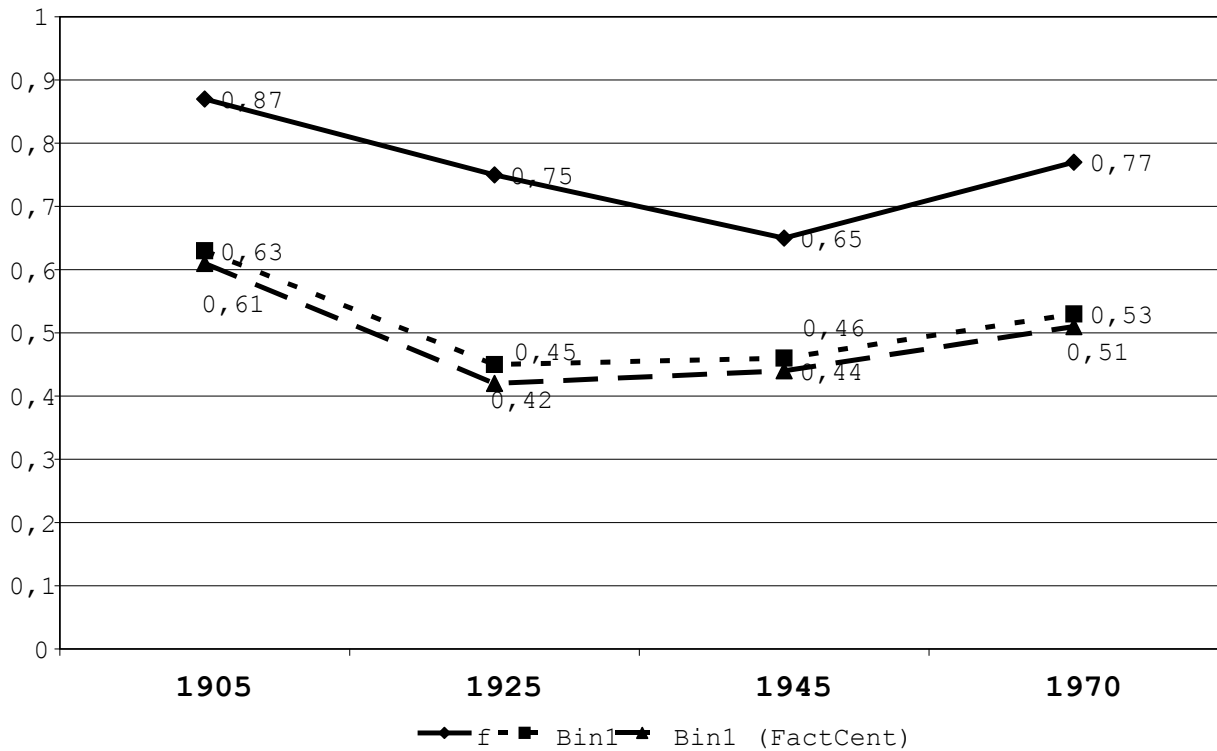


Figura 4.9. Evolución cuantitativa de las variantes sordas en el tiempo real según tres tipos de análisis (N=1403)

Las distintas líneas representan análisis estadísticos alternativos para verificar si las tendencias se repiten, independientemente del cálculo efectuado. En este caso, la línea continua (con nodos romboides) indica los valores de frecuencia relativa (*f*), la línea punteada la tendencia según el análisis binomial de 1 nivel (*Bin1*) y la línea continua de nodos triangulares el análisis binomial de 1 nivel, pero con los factores compensados (*center factors* en el programa *GoldVarb*), esto es, el programa no toma en cuenta el hecho de que haya casillas que acumulen muchos datos y otras que acumulen muy pocos. Si se toma en cuenta que los grupos de año de nacimiento en

este estudio no tienen el mismo número de informantes y por ende, de datos, es importante analizar qué pasaría si se le otorgara el mismo peso virtual a cada factor en cada grupo, independientemente del número de ejemplos con los que hubieran contribuido al total de casos. Los resultados en los tres tipos de análisis indican básicamente lo mismo: que existe una reducción gradual del ensordecimiento en las primeras dos o tres generaciones (según se vea frecuencias o probabilidad), pero termina la línea en un ascenso decidido, sobre todo en la última generación. Esta gráfica se complementa con la de la Fig. 4.8, que también muestra la tendencia general al ascenso de las variantes sordas con el tiempo.

4.6.3. *Elisión vocálica*

Para el año 1970 el análisis *Bin U&D* ($p=0.009$) selecciona tres factores internos y sólo uno social: intervienen el contexto precedente (consonante sorda: .596), contexto subsecuente (consonante [s]: .838) y que se trate de una palabra frecuente (.709); esta variante es favorecida claramente por hablantes de pocos estudios (.795). De alguna manera, este perfil es el espejo del correspondiente a las variantes sordas. Dada una elisión, es muy probable que se dé en palabras muy frecuentes con /s/ final como *pesos*, *pues* o *muchos* y en el habla de un informante de pocos estudios. El modelo estadístico sobre la muestra del año 2000 (análisis *Bin U&D*, $p=0.037$), además de seleccionar los mismos factores que en 1970 (consonante sorda precedente: .562; segmento [s] subsecuente: .875; palabra frecuente: .772; nivel bajo de estudios: .580), añade el del sexo del informante, pues en este caso son los hombres (.579) quienes

más favorecen la elisión. La distribución de la elisión según el año de nacimiento de los informantes traza también un patrón interesante:

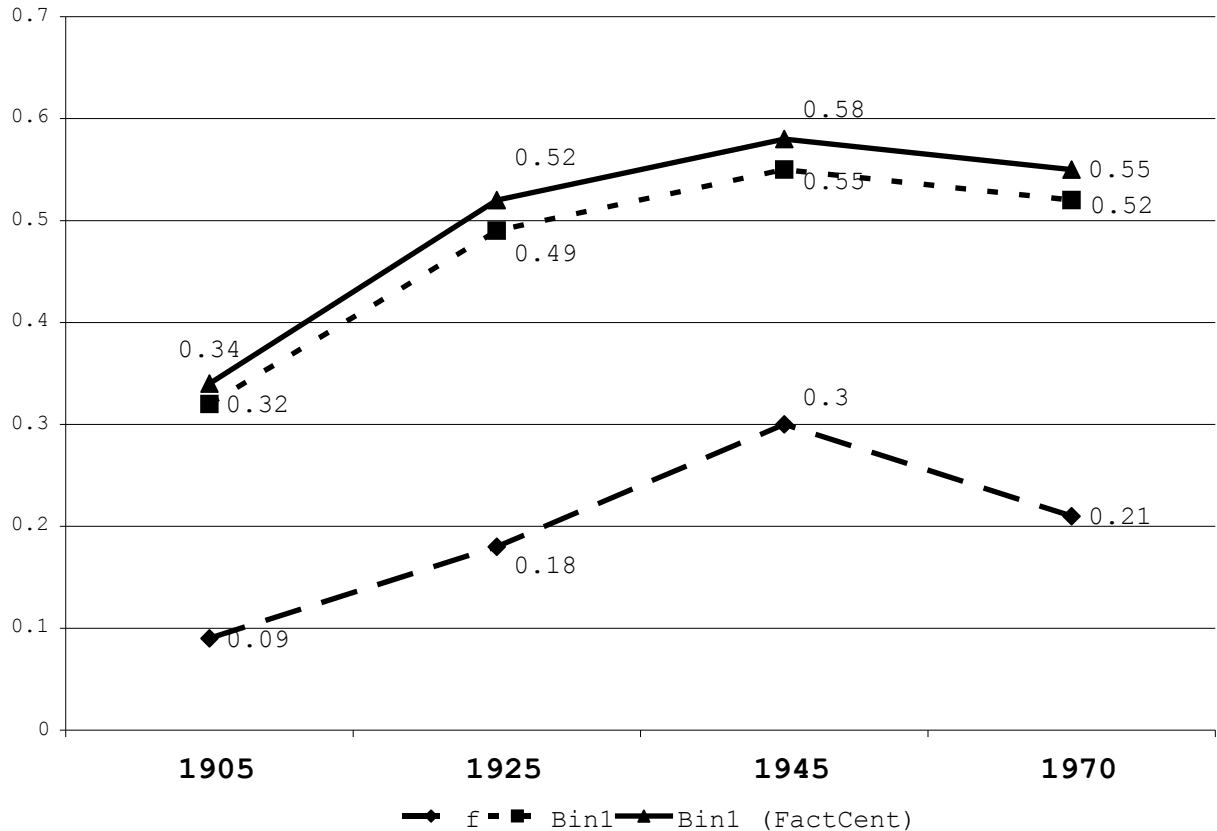


Figura 4.10. Evolución cuantitativa de la elisión vocálica en el tiempo real según tres tipos de análisis (N=1403)

Al igual que sucede con la asibilación de vibrantes, el patrón indica un ascenso de la variante más extrema en la tercera generación (los nacidos alrededor de 1945), pero en la última generación hay una retracción hacia pesos más conservadores que se complementa con el ascenso de las variantes sordas (Fig. 4.9). Con esto se confirma

nuevamente que en este dialecto se prefiere conservar material fónico a pesar de que se aumente el número de variantes no necesariamente fieles al input. En el siguiente apartado se discute con más detenimiento el perfil social de los informantes y su contribución al debilitamiento vocálico en ECM.

4.7. EN BUSCA DE LOS LÍDERES DEL DEBILITAMIENTO VOCÁLICO

Para complementar el perfil sociolingüístico de las vocales sordas en ECM se recurre nuevamente a la *Frecuencia relativa acumulativa* empleada para el análisis de las róticas en el cap. 3.

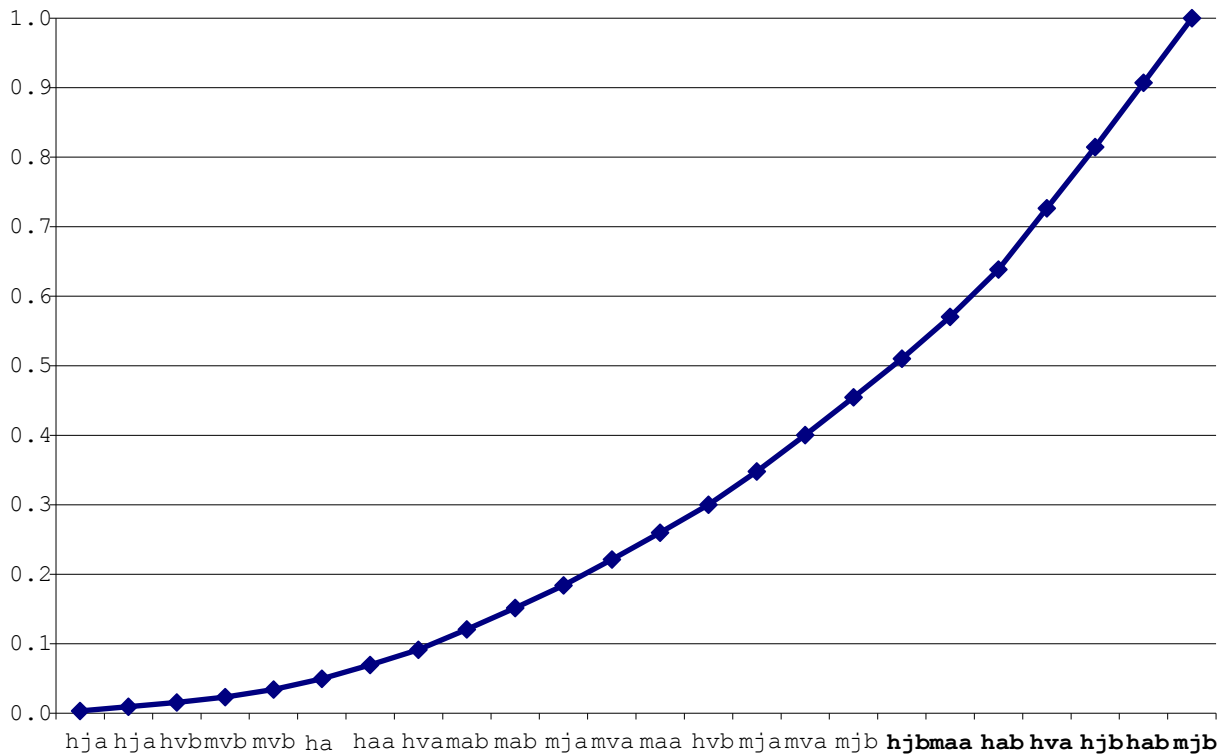


Figura 4.11. *Frecuencia relativa acumulativa del debilitamiento vocálico en 1970*

(N=847)

Son siete los hablantes que concentran alrededor del 50% de todos los casos de debilitamiento vocálico (codificados según sus características sociales y resaltados en negritas). De este grupo, cinco son personas de escasos estudios, cinco son hombres, y sólo uno de ellos es del grupo de mayor edad. Los líderes son tres hablantes con estudios de primaria: Rosa, de 23 años, ama de casa; Agustín, de 46 años, carpintero y Roberto, obrero de 24 años. Esta gráfica complementa el perfil del debilitamiento como un proceso que se concentra en los estratos de baja escolaridad de la ciudad. Ahora compárese con los líderes en el año 2000.

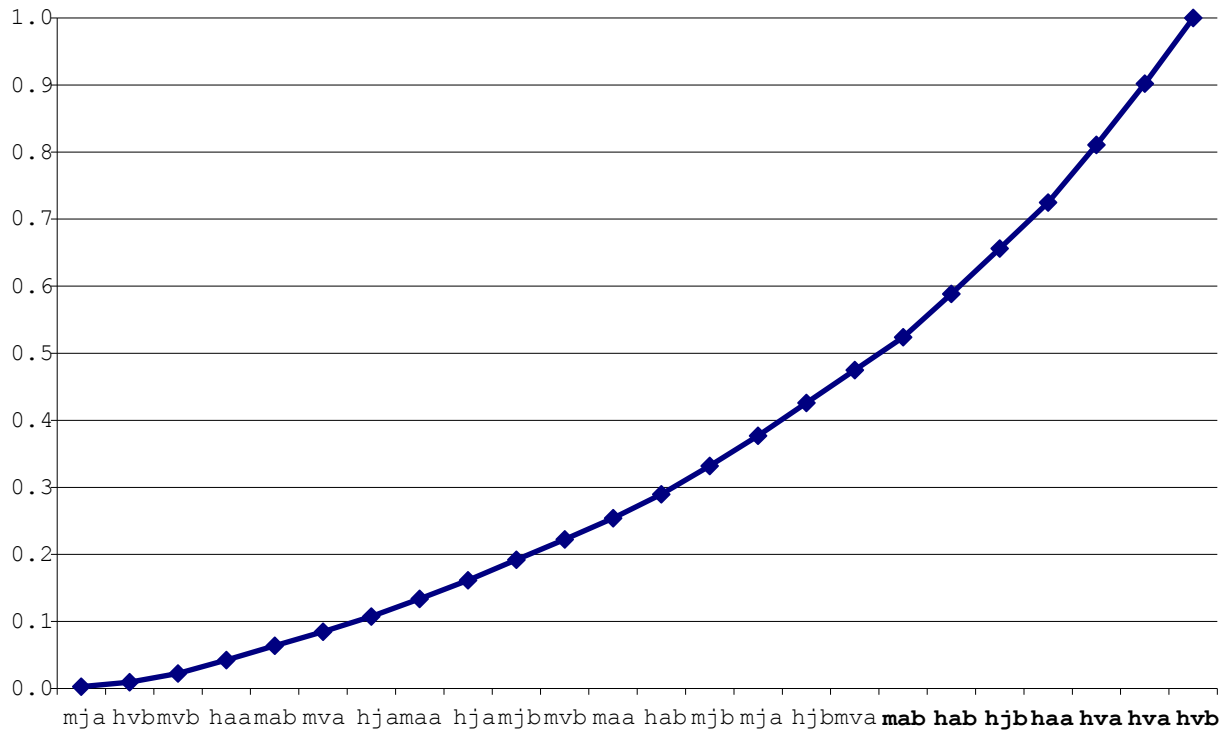


Figura 4.12. Frecuencia relativa acumulativa del debilitamiento vocálico en 2000

(N=756)

En el año 2000 los líderes siguen siendo predominantemente hombres (los primeros seis lo son), pero los primeros tres pertenecen al grupo de mayor edad, lo que sugiere que forman parte de una generación que presentaba un índice de debilitamiento mayor y este patrón lo conservaron, ellos son: Felipe, de 55 años, estudios de primaria; José Luis, de 64 años, con licenciatura y Rodolfo, de 56 años y estudios universitarios. A diferencia de 1970, ahora algunos de estos líderes no pertenecen a los estratos de baja escolaridad, pero sí constituyen la primera generación de universitarios en sus familias.

La naturaleza de los líderes ha cambiado respecto al tipo de variante con la que contribuyen al proceso de debilitamiento vocálico. Los líderes del debilitamiento en 1970 son sólo cinco informantes, pero acumulan 103 elisiones de las 157 totales, esto es, el 65%; el mismo grupo contribuye con la tercera parte de las vocales sordas (32%, 105/328). Es muy llamativo que uno de estos líderes, Luis –con estudios universitarios y entrevistado en el proyecto Norma culta-- sólo haya aportado dos elisiones, a diferencia de Agustín, de oficio carpintero y con estudios de primaria, quien aportó 33; Luis en realidad está en el grupo de líderes porque fue el máximo articulador de vocales sordas de toda esa muestra (48 ejemplos). En el año 2000 el panorama de los líderes es distinto: ahora el 43% de las elisiones totales (70/160) se concentra en los primeros cinco hablantes (esto es, ya no contribuyen con la mayoría de los ejemplos de elisión, pero siguen teniendo un peso importante) y en cambio sí concentran ahora una proporción mayor de vocales sordas que en 1970: 42.5% (184/432). Entonces, el líder de los años 1970 elidía más y ahora (año 2000) el líder

ensordece más, lo que nuevamente confirma el patrón general a evitar formas extremas de debilitamiento en tiempos recientes.

4.8. ANÁLISIS EN TEORÍA DE OPTIMIDAD ESTOCÁSTICA DEL DEBILITAMIENTO VOCÁLICO

La siguiente propuesta parte de un contexto fónico que, de inicio, favorece el debilitamiento (ante pausa o en sílaba cerrada por segmento sordo), por lo que no se presenta un candidato que cumpla completamente con los rasgos asignados en la entrada. Se decidió así debido a la escasísima presencia de debilitamiento en otros contextos. Por lo tanto, si consideramos solamente tres candidatos (vocal relajada, vocal sorda y cero fonético) se pueden proponer las siguientes restricciones:

UN CONJUNTO DE RESTRICCIONES PARA EL ANÁLISIS DEL DEBILITAMIENTO VOCÁLICO

IDENT_[μ]: El valor moraico atribuido en la entrada debe permanecer idéntico en la salida.

*PCO_[L]#: Prohíbese la aplicación del Principio de Contorno Obligatorio para rasgos del nodo laríngeo en final de emisión.

MAX-IO: Todo elemento en el input debe tener uno correspondiente en el output.

La primera restricción (IDENT_[μ]) penaliza la reducción moraica de la vocal y con ello a las variantes que llamamos 'relajadas'. La segunda restricción (*PCO_[L]#: penaliza sílabas como *-sas#* o *-sa#* que preserven la sonoridad de la vocal en un contexto de consonantes sordas. Finalmente, MAX-IO penaliza el cero fonético. En un

sentido fonológico, las restricciones en juego se jerarquizan de la misma manera para 1970 y 2000. La jerarquía de restricciones –atendiendo al peso estadístico de cada una de ellas- sería la siguiente:

Cuadro optimal 4.1. *Jerarquía de restricciones para el debilitamiento vocálico en ECM*

/(C)V(C)#/	IDENT _[μ]	MAX-IO	*PCO _[L] #
☞ [V]			*
[Ø]		*!	
[v]	*!		

En este caso las restricciones de fidelidad se encuentran arriba en la jerarquía. La restricción anti-PCO (de marcación) estaría al final, favoreciendo así la selección de variantes sordas. Esta jerarquía se modifica cuando interviene el factor léxico. Al respecto Bybee & Hopper (2001) y Pierrehumbert (2001) han señalado que un modelo que incorpore la frecuencia de uso en los procesos de variación y lenición fonética puede dar cuenta de los procesos fonéticos de detalle --que suelen ocupar, entre otras disciplinas, a la sociolingüística-- de una manera mucho más adecuada que los modelos categóricos y jerárquicos en fonología.

Siguiendo dichos razonamientos, A. Coetzee (2009) propone que las restricciones fonológicas pueden ser las mismas para todo el lexicón, pero cada palabra tendría una función asignada que daría indicaciones al hablante sobre su frecuencia de uso y cuestiones de detalle sobre su realización fonética (como sugiere Pierrehumbert, 2001). De esta manera, el hablante asigna la palabra a una clase léxica

que, como en el caso de palabras como *pues* [pus] o *entonces* [tons], privilegia formas reducidas siguiendo un principio de menor esfuerzo, esto a pesar de que violen restricciones de marcación muy altas en español, como las que penalizan secuencias de consonantes heterorgánicas a inicio de sílaba como [ps] como en *pues sí* [#'ps:i#] o incluso segmentos consonánticos como núcleo de sílaba en frases como *pues no* [pɤ.'no]).

Atendiendo el orden de las variantes en las palabras frecuentes de nuestro corpus, la jerarquía de restricciones queda de la siguiente manera:

Cuadro optimal 4.2. *Jerarquía de restricciones para el debilitamiento vocálico en palabras frecuentes en ECM*

/(C)V(C)#/	IDENT _[μ]	*PCO _[L] #	Max-IO
☞ [Ø]			*
[V̥]		*!	
[v]	*!		

Las restricciones *PCO_[L]# y Max-IO intercambian posiciones en la jerarquía, dejando la elisión como primera opción, mientras que la restricción sobre la reducción moraicica permanece al inicio, penalizando las variantes relajadas. La interesante propuesta de Coetzee (2008, 2009), al incorporar de manera explícita el peso de la frecuencia de uso de palabras o grupos de palabras específicos en modelos derivados

de la Teoría de la Optimidad, permite explicar de manera congruente la movilidad en la jerarquía de restricciones para el debilitamiento vocálico en ECM¹³.

La tendencia a que las variantes sordas ganen terreno en el año 2000, junto con el descenso del nivel de elisión y relajación, son indicios de una mayor preferencia por las formas más fieles en ECM. El patrón en este sentido es similar a lo que ocurre con las róticas (Cap. 3), en las que la forma menos fiel (la asibilada) está dejando su lugar a las variantes de relajación intermedia (las aproximantes). Estas similitudes se discutirán en el Cap. 6.

4.9. CONCLUSIONES

El análisis del debilitamiento vocálico en ECM confirma el patrón que sigue la variabilidad en las róticas de retracción de las variantes más extremas (en este caso la elisión) a favor de variantes intermedias (el ensordecimiento). La jerarquía de restricciones --que en este caso favorece el ensordecimiento-- es la misma desde los años 1970, lo que difiere respecto a las claras modificaciones de la jerarquía que

¹³ El trabajo de Jurafsky *et al.* (2001) también muestra que los ítems léxicos con mayor frecuencia de uso tienden a favorecer el desgaste fonológico, que se manifiesta --vale la pena subrayarlo-- como *ensordecimiento de vocales, reducción de la duración o elisión de elementos consonánticos finales*. A esta hipótesis la llaman de Reducción Probabilística. Estaría asociada esta reducción, de alguna manera, con procesos de coarticulación a nivel segmental que buscan la concatenación de elementos homorgánicos --esto es, el uso preferente de los mismos articuladores para segmentos distintos-- siguiendo un principio general de economía articulatoria (cf. Farnetari & Recasens 1999).

sufren las consonantes róticas en el tiempo. Una consonante sorda previa, sumada a una /s/ final [Ç_s/] (*cosas, pesos, meses, muchos, (en)tonces*, etc.) y a la posición prepausal constituyen el contexto ideal para que se active este proceso.

En general, el perfil sociolingüístico del hablante que elide es el de un hombre de bajos estudios, mientras que hablantes de estudios altos y las mujeres prefieren el ensordecimiento. Por lo tanto, este proceso en ECM confirma la muy documentada tendencia de las mujeres a evitar las formas estigmatizadas o más alejadas del estándar de pronunciación (cf. Labov 2006; Martín Butragueño 2006; Lastra & Martín Butragueño 2003).

Los datos de esta investigación en ECM confirman los hallazgos sobre el importante peso de la frecuencia de uso de las palabras y la mecanización de los gestos articulatorios en los procesos de desgaste fónico. Será muy interesante en el futuro estudiar más a fondo la correlación entre estudios bajos y un mayor índice de elisión que seguramente está emparejada con un acervo léxico menos voluminoso. En otras palabras, es esperable que hablantes con menos escolaridad presenten mayores índices de elisión debido a un mayor uso de palabras como *entonces (tons)*, *pues (pus)* –palabras con función de ‘comentador’ o para mantener el turno de habla (Soler & Serrano 2010)--, *pesos*, etc. El hecho de que haya menos elisión en el año 2000, cuando la mayoría de la población ha aumentado su nivel educativo respecto a 1970¹⁴, puede ser evidencia a favor de esta hipótesis.

¹⁴ En 1970 el promedio de años de escolaridad en la población mayor a 15 años en el D. F. era de 5.8 (menos de la primaria), mientras que para el año 2000 es de 9.7 (secundaria concluida) (cf. Aguilar

Por otra parte, algunos de los líderes del debilitamiento vocálico en el año 2000 son hombres que en su juventud se asentaron en la periferia del Área metropolitana de la ciudad de México: Felipe, residente en Ciudad Netzahualcóyotl; Guillermo, quien viven en Tlalnepantla; y Rodolfo, en la región boscosa-rural de la Delegación Álvaro Obregón. Esto sugiere otra hipótesis que podría explorarse en el futuro: que quizá las hablas caracterizadas como “populares” en el español capitalino de los años sesenta y setenta (y sus rasgos característicos, como la elisión vocálica) pervivieron entre los grupos sociales con menos recursos económicos que tuvieron que alejarse de la zona principal del D. F. para vivir en colonias relativamente alejadas del núcleo urbano o en los municipios conurbados del Estado de México, muchas de las veces, en asentamientos irregulares o complejos habitacionales de interés social donde era común la falta de oportunidades de desarrollo económico, social, político y cultural (cf. Villaseñor 2002). No deja de ser llamativo que estos líderes del debilitamiento compartan el rasgo de residir en estas zonas que –al menos hace 30 años-- se encontraban en la periferia de la Zona Metropolitana de la ciudad de México. Un seguimiento a través de historias de vida de estos hablantes puede arrojar luz respecto a esta hipótesis que podríamos llamar de “desplazamiento de registros sociolingüísticos populares del centro a la periferia de la metrópoli”. Esta hipótesis se discutirá nuevamente en el cap. 6.

Morales 2010). Por supuesto, debe tomarse en cuenta que el CSCM no tiene una muestra proporcional, sino en forma de cuotas.

CAPÍTULO 5

LOS LÍMITES DE LOS CORPUS ORALES: UNA REVISIÓN DE LA HIPÓTESIS DE LA DIFUSIÓN LÉXICA

En este capítulo se aborda la complejidad de tratar de buscar procesos de cambio lingüístico en corpus orales que vayan más allá de los aspectos puramente fonético-fonológicos que, como se vio en los capítulos anteriores, se analizan con relativa suficiencia en corpus orales bien estructurados --aunque sean relativamente pequeños, como el utilizado en esta tesis.

La pregunta de este capítulo, por lo tanto, es si existen procesos variables en otros niveles estructurales del español de la ciudad de México (ECM) que puedan ser analizados con relativa exhaustividad a partir de un corpus oral como el que aquí se utiliza. Se explorarán diversos problemas en tres niveles: fonológico, morfológico y léxico. Primero se abordará la distribución de dos marcas morfológicas vinculadas con los grupos educativos altos: los adverbios terminados en *-mente* y los sustantivos terminados en *-dad*, asociados a entidades abstractas. Después se analiza el caso de los grupos consonánticos cultos, un fenómeno que puede ubicarse en un nivel “fonológico”. Por último, se explora la medida en que el nivel léxico puede ser un índice sociolingüístico de diferenciación social. Para ello, se investiga las diferencias lexicoestadísticas del acervo de sustantivos entre los grupos sociales y temporales del corpus a través de una diversidad de métodos analíticos para finalmente estudiar la aparición o desaparición de ítems léxicos con carga sociolingüística específica (palabras como *onda*, *chido* o *súper*). Tales problemas ilustrarán la forma específica en que ciertos procesos variables (y cambios) se difunden a través del léxico. Por supuesto, el capítulo no pretende agotar la diversidad de problemas susceptibles de estudiarse, sino que busca testigos de diferenciación sociolingüística que puedan estudiarse con un nivel de detalle razonable.

Por otra parte, uno de los propósitos específicos del capítulo es mostrar que en estos niveles de análisis es necesario valerse de corpora más voluminosos y diversificados, algo que yo había señalado en Serrano (2011). El hilo conductor del capítulo es la cuestión de la hipótesis de la difusión léxica. Se discute su pertinencia

respecto a los procesos fonéticos revisados en los capítulos 3 y 4 y, tras esta revisión, se analizan los problemas fonológicos, morfológicos y léxicos arriba enumerados. El capítulo cierra con una recapitulación de todos los procesos.

5.1. EL CAMBIO NEOGRAMÁTICO Y LA HIPÓTESIS DE LA DIFUSIÓN LÉXICA

Los atlas dialectológicos de G. Wenker (*Sprachatlas des Deutschen Reichs*, de 1881) y de Jules Gilliéron (*Atlas Linguistique de la France*, publicado entre 1902 y 1910) surgen en buena medida para poner a prueba las leyes neogramáticas sobre la regularidad del cambio fónico: “The Neogrammarian principle stated: *Ausnahmslosigkeit der Lautgesetze*, or ‘sound changes are exceptionless’. With such a hypothesis, so bold and admirably refutable, dialect evidence would be obviously relevant” (Chambers & Trudgill 1998: 14). La ley de Verner, propuesta en 1877, y que corregía las excepciones de la Ley de Grimm sobre el cambio fonético del proto-indo-europeo a los distintos dialectos germánicos, permitió postular –en definitiva-- que “el cambio fonético se produce sin excepciones” (cf. Moreno de Alba 2013: 7; Chambers & Trudgill 1998: 21) y que, además, el cambio actuaba “ciegamente”, sin importar los efectos en la estructura de las palabras o en nuestra habilidad para comunicarnos (Hock & Joseph 1996: 148). Desde esta perspectiva, el cambio lingüístico se produce

de manera regular en todos los contextos que lo permitan (como las correspondencias entre las /t/ iniciales en inglés con la /ts/ en alemán)¹.

Sin embargo, no todos los hallazgos de la geografía lingüística permitían confirmar la regularidad del cambio, como el problema de la evolución de /u:/ del alemán medieval a /au/ en alemán moderno, explorado en el atlas de Wenker² (cfr. Chambers & Trudgill 1998: 133). Otro hallazgo de la dialectología fue que los cambios fonéticos no afectan a todas las palabras que presenten el mismo contexto fónico, lo que sugería la hipótesis de que “cada palabra tiene su propia historia” (Warburg 1951). De hecho, los hallazgos principales de Gilliéron, que están detrás de algunos de los procesos revisados en esta tesis, pueden resumirse de la siguiente manera:

1. Los dialectos no tienen a menudo homogeneidad porque reciben influjos de distintas procedencias, a consecuencia de los cuales se originan movimientos de avance y retroceso.

¹ Aunque básicamente correcta, desde la perspectiva de Roger Lass (1997), esta historia sobre la hipótesis neogramática y las aportaciones de Verner está muy simplificada. Lass (1997: 132-135) discute desde los alcances y términos de la hipótesis, hasta el hecho mismo de que Verner no perteneciera a ningún círculo de neogramáticos y que tampoco veía sus propias correcciones como “leyes”, sino como una simple heurística para la reconstrucción.

² El atlas está disponible digitalmente en: <http://www.diwa.info/>.

2. Un proceso ya producido puede desandarse (idea de regresión); esta regresión puede llevar caminos distintos de los que recorrió el movimiento de avance (falsa regresión).

3. La evolución de una palabra depende también de que corresponda a una idea abstracta o a un objeto importado de fuera, por ejemplo, de la capital del país.

[Warburg 1951, *apud* Arellano 1977: 278].

La dialectología, por lo tanto, sugería la llamada hipótesis de la *difusión léxica*: en lugar de un mecanismo “ciego”, el cambio lingüístico va afectando a palabras específicas (o grupo de palabras) a la vez y de manera acumulativa a través del lexicón, hasta que se generaliza la forma innovadora: “Changes filter or diffuse through available environments rather than hitting all at once. In sound change these environments are typically item-specific (either lexical or morphological), so that at a given stage only certain lexical items or morph(eme)s will be affected” (Lass 1997: 140).

Labov (1996, 2006) discute a profundidad los alcances de ambas teorías del cambio. En general, los hallazgos de Labov y de otros investigadores en todo el mundo permiten afirmar que los principios neogramáticos del cambio fónico se sostienen, sobre todo cuando se oponen los argumentos articulatorios a los argumentos de tipo funcional (Labov 1996: 918). Sin embargo, los procesos mecánicos no parecen funcionar claramente en ciertos procesos de difusión de variantes innovadoras. Como

ya se mencionó en el Cap. 1 de esta tesis, Labov encontró que el proceso de descenso de /e/ en Filadelfia sólo podría explicarse atendiendo al efecto de las presiones sociales (Labov 2006: 746-751). Esto lleva a Labov a postular que existen cambios que siguen un patrón mecanicista (sensibles, sobre todo, a los factores internos) y otros que siguen un patrón de difusión léxica (más sensibles a la frecuencia y los factores sociales)³.

Además del programa de investigación variacionista de Weinreich, Labov & Herzog (1968) y los avances de Labov (1996, 2006, 2010) –reseñados en el Cap. 1–, existe una fuerte línea de investigación concentrada en el papel de la frecuencia en los procesos de cambio lingüístico. Investigaciones que analizan con detalle los mecanismos específicos de los patrones de difusión léxica y la frecuencia de las palabras se reúnen en el influyente volumen *Frequency and the Emergence of Linguistic Structure* (Bybee & Hopper 2001). Algunos trabajos de dicho volumen que son altamente pertinentes para esta tesis son, por un lado, el de B. Phillips, quien propone la sugerente “Hipótesis de implementación de la frecuencia” que dicta que “Changes which require analysis –whether syntactic, morphological, or phonological—during their implementation affect the least frequent words first;

³ Lass (recuperando las hipótesis de McMahon 1994 y Ogura 1995) sugiere que en realidad los cambios fonológicos surgen con un patrón de difusión léxica pero que se consolidan siguiendo principios neogramáticos, “given enough time” (Lass 1997: 141, n. 41).

others affect the most frequent words first” (2001: 134)⁴, por otra parte, en el mismo volumen Pierrehumbert proporciona evidencia a favor del *modelo de ejemplares* (originalmente un modelo de percepción) en la producción lingüística; dicha autora encuentra que el hecho de que las categorías fonológicas sean aprendidas a través de la exposición a un gran número de *tokens* permite entender por qué es posible que los hablantes aprendan también la variación fonética de detalle propia de los procesos de lenición (Pierrehumbert 2001: 154)⁵.

Como pudo verse en el cap. 3, el proceso de asibilación de las róticas parece comportarse de manera ciega, al estilo neogramático: el contexto prepausal afecta por igual a toda la clase de infinitivos (que constituyen la gran mayoría de palabras en dicho contexto). Por otra parte, el proceso de debilitamiento vocálico (cap. 4) sí está vinculado a ciertas palabras más frecuentes (como *pues* y *entonces*), por lo que puede verse más bien como un proceso de difusión léxica y en el que la alta frecuencia de las palabras afectadas desempeña un papel crucial.

Si en los procesos fonéticos analizados hasta el momento parece haber comportamientos neogramáticos y de difusión léxica, ¿qué ocurrirá con la productividad de fenómenos del nivel morfológico? Esto es precisamente lo que se

⁴ Una sugerencia más reciente de Phillips (2006) es que algunos patrones de difusión en los individuos reflejan el patrón comunitario, lo que permitiría contar con un modelo unificado de difusión y justificar a la comunidad de habla como unidad de análisis (según la reseña de Tipton 2009: 280).

⁵ Otro trabajo importante en dicho volumen es el de Jurafsky *et al.* (2001), ya mencionado en el Cap. 4.

explora en las siguientes subsecciones sobre la distribución social y etaria de los adverbios terminados en *-mente* y de los sustantivos abstractos terminados en *-dad*.

5.2. MORFOLOGÍA CULTA I: EL CASO DEL MORFEMA *-DAD*

En esta búsqueda de índices de diferenciación sociolingüística y la ruta de difusión que siguen, se eligió el morfema derivativo *-dad* por ser el que suele expresar la cualidad abstracta del sustantivo, y por lo mismo, estar (esperablemente) más vinculado con el lenguaje culto en general: *humanidad, nacionalidad, bondad*, etc. Se realizaron búsquedas de concordancias de todas las palabras con esta terminación. En la cuenta se descartó la palabra *verdad* cuando funcionaba como marcador discursivo (de funcionamiento fático, o *tag question*)⁶, pero se mantuvieron los ejemplos en los que el contenido léxico cabal de la palabra se mantenía.

Se usaron archivos de texto de los 48 informantes totales editados por Jerid Francom (Wake Forest University)⁷. En estos archivos sí se incluye el habla del entrevistador, pero el texto está muy depurado en cuanto a palabras fuera de la entrevista y marcas especiales (*I, E*, transcripciones fonetizadas, palabras incompletas, etc.).

⁶ Sobre *verdad* como marcador discursivo en español mexicano, véase Orozco (en prensa).

⁷ Aprovecho para agradecer a Jerid su valioso conocimiento y tiempo invertidos en esta tarea.

Para este análisis se buscaron las siguientes secuencias de caracteres con el programa *AntConc* 3.2.4 (Anthony 2014): *-edad*, *-idad*, *-rdad* y *-rtad* (las únicas variantes que aparecieron en el corpus). En el siguiente cuadro se indica el número de ejemplo totales de cada variante en los dos grupos educativos explorados.

Cuadro 5.1. *Palabras terminadas con variantes del morfema -dad en 1970*

	<i>-EDAD</i>	<i>-IDAD</i>	<i>-RDAD</i>	<i>-RTAD</i>
CULTA (230)	36 (<i>edad</i> , <i>enfermedad</i> , <i>sociedad</i>)	179 (<i>universidad</i> , <i>cantidad</i> , <i>calidad</i>)	7 (<i>verdad</i>)	8 (<i>libertad</i>)
POPULAR (69)	25 (<i>edad</i>)	38 (<i>facilidad</i>)	0	6

En las búsquedas en los cuatro grupos de hablantes principales se decidió eliminar varios casos de palabras como *verdad* y *realidad* cuando funcionaban como marcadores discursivos o que formaban parte de colocaciones, ya que implican algún grado de desemantización de la forma. Por ejemplo, en los resultados de los hablantes cultos se eliminaron los casos de *verdad* como pregunta retórica (1) o marcador discursivo (2):

- (1) Pues, ahí es lógico que se requiera una ingeniera química ¿*verdad*? para desarrollar ese trabajo [Hombre, joven, culto, 1970].

(2) pues eh no sé *la verdad* yo creo que es una cantidad que ¡nunca se llega a saber!

[Hombre, mayor, culto, 2000]

Entre estos hablantes se documentaron 250 datos de la forma *verdad*, de las cuales sólo 7 fueron casos del sustantivo propio⁸. Del sufijo *-idad* se descartaron 10 casos de la frase interjectiva *¡qué barbaridad!* y 13 del marcador discursivo *en realidad*; el vocablo mayoritario para este sufijo fue *universidad* (32 ocasiones). Finalmente, el vocablo *libertad* apareció en 8 ocasiones. El gran total, tras la eliminación de las formas como marcadores discursivos o en estructuras fraseológicas es de 230 casos del morfema *-dad* entre los hablantes cultos de 1970. Ese morfema al parecer sí permite diferenciar a los grupos sociales: nótese que los hablantes de pocos estudios presentan sólo 69 ejemplos. En el siguiente cuadro se presentan los resultados en el año 2000, obtenidos por el mismo procedimiento.

Cuadro 5.2. *Palabras terminadas con variantes del morfema -dad en 2000*

	<i>-EDAD</i>	<i>-IDAD</i>	<i>-RDAD</i>	<i>-RTAD</i>
CULTA (215)	60 (<i>edad</i> , <i>enfermedad</i> , <i>sociedad</i>)	149 (<i>universidad</i> , <i>especialidad</i>)	0	6 (<i>libertad</i>)
POPULAR (52)	23 (<i>edad</i>)	28 (<i>responsabilid</i> <i>ad</i>)	0	1 (<i>libertad</i>)

⁸ Respecto al sufijo *-edad*, se descartaron 3 casos de *quedado*.

Hay una ligera disminución comunitaria en los usos del morfema (299 casos totales en 1970 frente a 267 en 2000), pero las proporciones son similares en ambos muestreos. Existen, sin embargo, algunas anotaciones interesantes sobre algunas de las formas desemantizadas (y descartadas). La expresión *¡qué barbaridad!*, que era común en las entrevistas de los años 1970 (aparece 25 veces), al parecer va desapareciendo en las hablas actuales, en las que sólo parece en 4 ocasiones (de hecho, he atestiguado que ciertas personas la asocian con hablas rurales o de gente poco educada). Por otra parte, la palabra *verdad* ganó terreno en las hablas populares, ya que apareció 103 veces en total, frente a los 38 casos de 1970, lo que indica una clara tendencia ascendente de esta palabra como marcador discursivo, sobre todo como *en verdad* y *la verdad*. Los resultados globales para el morfema *-dad* se muestran en la siguiente gráfica:

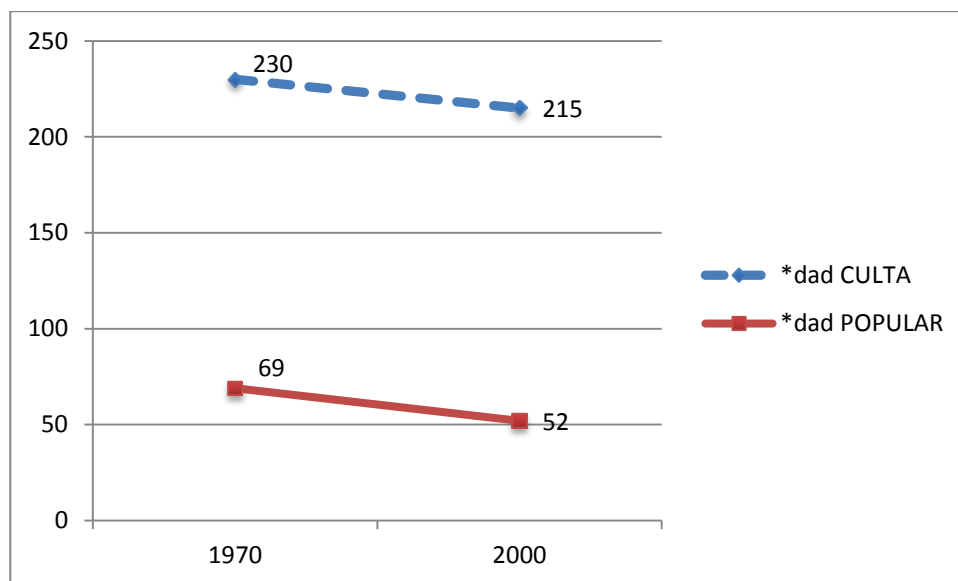


Figura 5.1. Palabras con sufijo *-dad* en hablantes cultos y populares 1970-2000

El descenso es comunitario y los grupos sociales parecen haber bajado en las mismas proporciones. Para confirmar que la pendiente de descenso es similar, se realizó una gráfica de tendencia lineal que se presenta a continuación:

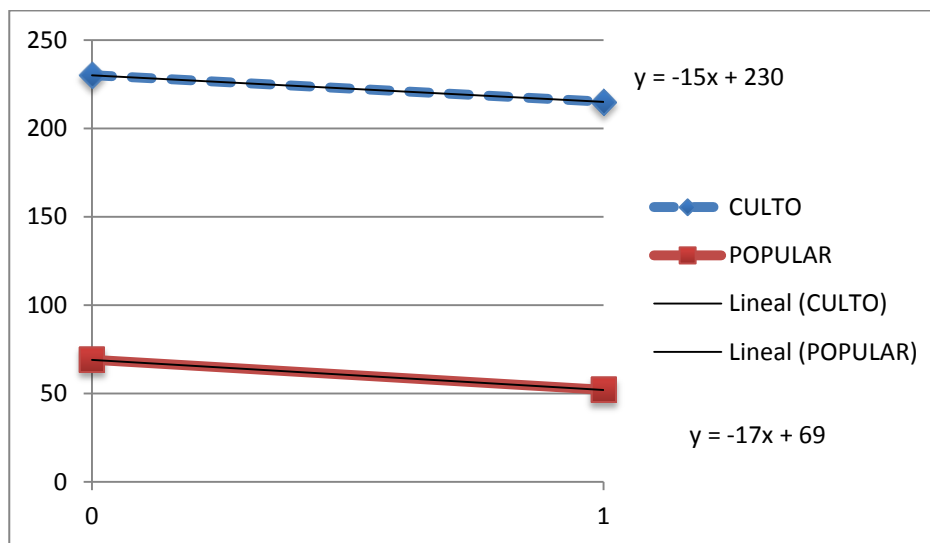


Figura 5.2. Gráfica de tendencia lineal del morfema -dad en ECM 1970-2000 (0=1970, 1=2000)

Sigue habiendo gran distancia entre los grupos sociales, pero el aspecto general de la reducción es en paralelo. Obsérvese que las pendientes son muy semejantes en ambos casos, -15 para habla culta, y -17 para habla popular. Por lo tanto, se mantienen las diferencias sociolingüísticas, con un ligero descenso global, lo que podría indicar cierta estabilidad o quizás una menor formalidad en el habla comunitaria en estilo de entrevista. Lo que puede interpretarse sociolingüísticamente a partir de esta gráfica

es muy interesante: por un lado, que los grupos sociales siguen estando altamente diferenciados en cuanto a la productividad estadística del sufijo *-dad*; sin embargo, la proporción en que ambos disminuyen es la misma: esto es, siguen siendo hablantes distintos, pero no dejan de pertenecer a la misma comunidad de habla. El sufijo adverbializador *-mente*, tradicionalmente vinculado a las palabras cultas, permitirá confirmar (o descartar) esta tendencia del nivel morfológico.

5.3. MORFOLOGÍA CULTA II: EL CASO DE LOS ADVERBIOS TERMINADOS EN EL MORFEMA *-MENTE*

El uso del sufijo adverbializador *-mente*, de función semántica básicamente modificadora de verbos, adjetivos y otros adverbios, es de origen decididamente culto (Company 2012 a) y está siendo sustituido actualmente, sobre todo en lengua oral, por *adverbios adjetivales* (Kaul de Marlangeon 2002: 73-74) que utilizan sólo la base adjetiva, como en *Juliana lo escribió rápido* (en vez de *rápidamente*). De hecho, su *productividad*, entendida como la capacidad del morfema para sufijarse a bases adjetivas, actualmente sigue siendo alta (84 bases --Company 2012 b: 130), pero la *productividad* en términos de *frecuencia* de cada uno de estos adverbios es mucho menor, ya que se limita a unos cuantos tipos (p. ej., sólo cuatro tipos conforman el 34% de ocurrencias en un corpus reciente de estos adverbios en posición inicial -- levantado en 2011, con 417 casos totales reportados en Arias Leal 2013); estos son (en orden decreciente): *obviamente*, *finalmente*, *simplemente* y *afortunadamente* (cf. Company 2012 b: 132, Cuadro 4). Dado el origen culto que tuvo este tipo de adverbio

(Company 2012 a), puede servir como otro índice sociolingüístico que permita diferenciar grupos de altos y bajos estudios, como los analizados en esta tesis.

En la exploración realizada en el corpus, la tendencia es a una baja generalizada en el tiempo real, como puede verse en la siguiente gráfica:

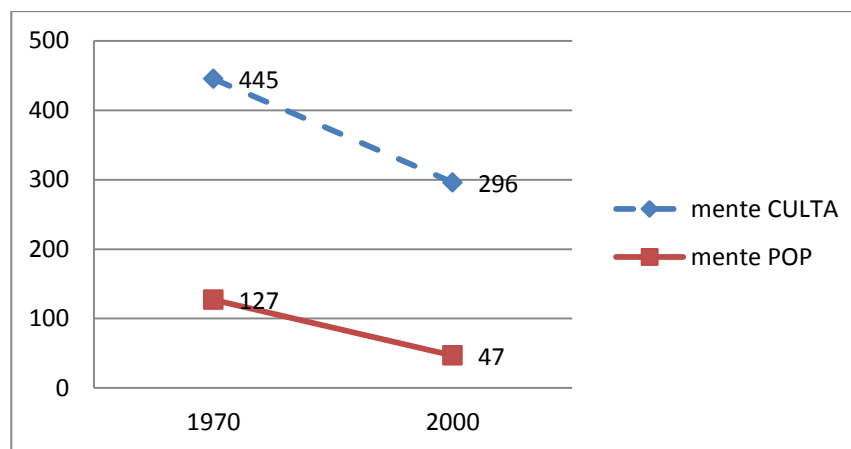


Figura 5.3. Palabras con sufijo -mente en hablantes cultos y populares 1970-2000

El paralelismo de las líneas de los grupos sociales es muy llamativo y replica el patrón del sufijo *-dad* revisado en el apartado anterior, incluso la gráfica de tendencia lineal es relativamente similar:

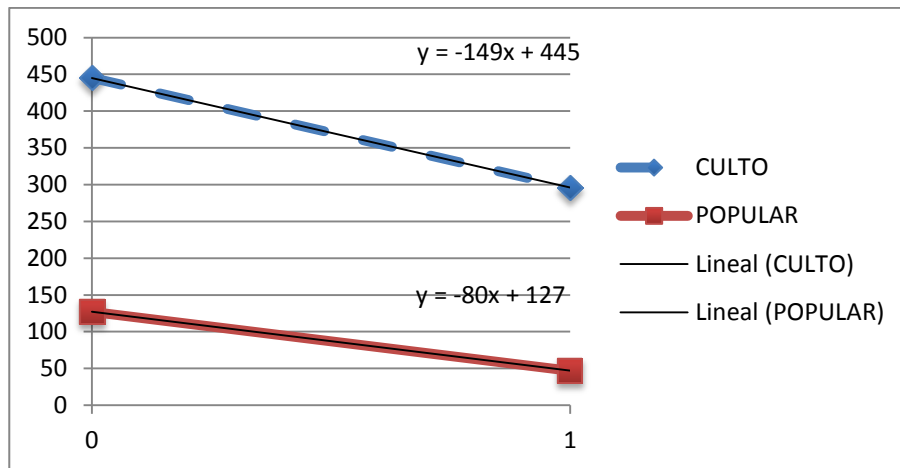


Figura 5.4. Gráfica de tendencia lineal del morfema *-mente* en ECM 1970-2000 (0=1970, 1=2000)

El grado de descenso de *-mente* es bastante mayor respecto al del sufijo *-dad*, sin embargo, lo que tiene en común es que se mantienen las diferencias entre hablantes cultos y populares, si acaso, con un nivel algo menor en el año 2000. Por supuesto, muy probablemente no se trata de que, por ejemplo, haya menos expresiones adverbiales en el año 2000, sino que están siendo expresadas, como se había dicho, a través de adverbios adjetivales (sólo el adjetivo base, sin el morfema *-mente*). Si bien la tendencia de ambos grupos sociales es a la baja, lo cierto es que, aunque el grupo culto sigue arriba, su pendiente de descenso es proporcionalmente más prominente que la de los hablantes de pocos estudios, quienes parecen estar abandonando el morfema de manera decidida, algo que ya se había indicado en las investigaciones recientes de Company, quien encuentra que este adverbio se encuentra muy

vinculado, diacrónica y sincrónicamente, con la *lengua escrita* (Company 2012 a: 21).

Una conclusión importante en el mismo texto es la siguiente:

Con base en los datos arrojados por los corpus base y adicional, hemos intentado probar que la formación de estos adverbios es un cultismo peculiar, que por ello se retrajeron las variantes morfofonémicas patrimoniales o vernáculos, que por ello se manifiestan diversificados léxicamente y muy bien generalizados en la lengua escrita literaria, *pero que fue una gramaticalización en cierto sentido débil en la lengua oral*, debilidad no estructural sino léxica, dada la constante repetición de ciertos adverbios, y frecuencial, dada su baja aparición en la oralidad [Company 2012 a: 37; énfasis mío].

Esto es, no se trata, al menos en lengua oral, de un proceso de estricta gramaticalización, sino de inserción léxica. Dado que los morfemas *-dad* y *-mente* presentan un patrón de disminución general en la comunidad, no es fácil rastrear el patrón de abandono de estas formas; en todo caso, si existe un cambio lingüístico en marcha (p. ej., el del descenso generalizado en el uso de los adverbios en *-mente*, defendido por Company [2012; en prensa]), sería un patrón de cambio de tipo mecánico, relativamente ciego ante la naturaleza formal o funcional de las palabras involucradas.

En el siguiente apartado se analiza otro proceso variable, y que puede denominarse *fonoléxico*: el de los grupos consonánticos cultos, que permitirá observar la distribución de un índice sociolingüístico de naturaleza distinta a la de los revisados hasta aquí en el tiempo real.

5.4. EL CASO DE LOS GRUPOS CONSONÁNTICOS CULTOS

En los apartados anteriores se exploró la pertinencia de sufijos relacionados con las hablas educadas (*-dad* y *-mente*) como factores de diferenciación sociolingüística. La pregunta de esta subsección es: ¿también los grupos consonánticos cultos (GCC) pueden servir como índices de diferenciación sociolingüística? Si es así: ¿qué patrón evolutivo han tenido estos grupos en los períodos analizados? Los datos del corpus oral de esta investigación sugieren que la mayor escolaridad e inmersión en los modos de vida urbanos en tiempos recientes funcionan como factores de movilidad sociolingüística, movilidad reflejada en las hablas recogidas en entrevistas. Los GCC, como se sabe, son *bs*, *cc*, *cn*, *cs*, *ct*, *ds*, *gn*, *gd*, *mn*, *ns*, *ps* y *pt* (Díaz Rojo 2002). Aunque antes se ha afirmado que esos grupos se introdujeron a la norma escrita del español por presión de la Real Academia de la Lengua en el siglo XVIII, el estudio de Ramírez Luengo (2011) sobre estos grupos en documentos de Bilbao de los siglos XVI y XVII sugiere que, en realidad

el proceso de recuperación de esta coda silábica no tiene lugar en el siglo XVIII, como tradicionalmente se ha mantenido [...], sino que comienza mucho antes de esta centuria, por más que sea en ella cuando –al menos, en el caso de la documentación bilbaína aquí analizada– termina por generalizarse [Ramírez Luengo 2011: 463].

Existen pocos trabajos sobre estos grupos consonánticos en español hablado contemporáneo. Uno de estos es el estudio sociolingüístico de Donni (1974) en español de Rosario, Argentina, que muestra evidencia a favor de los grupos consonánticos en general como índices sociolingüísticos. Aunque la autora menciona que se basa en grabaciones con alrededor de 180 personas, no ofrece datos cuantitativos, sino una síntesis cualitativa del fenómeno. Donni encuentra que

Los límites entre los niveles sociolingüísticos son fluidos y en cada nivel se conserva la estratificación estilística. Desde otro punto de vista, el del prestigio de las formas lingüísticas, sólo se consideran incorrectas las realizaciones con cero fónico del primer elemento y que son características del nivel popular/vulgar en contextos informales. Sin embargo, estas formas a veces penetran también en los grupos de hablantes cultos en estilo informal y ritmo rápido [Donni 1974: 537].

Un trabajo sobre los GCC en español mexicano que combina el estudio de documentos coloniales y habla popular actual es el de Arias (2008). En dicho ensayo, Arias trata de demostrar, a través de la revisión de 24 documentos coloniales de los siglos XVI y XVII, cómo ciertos procesos como la vocalización de consonantes implosivas ([dok.'tor → doj.'tor]) o el cambio de punto de articulación del primer miembro del par (como en [dok.'tor → dot.'tor]), presentes en las hablas populares contemporáneas, ya se veían registrados en la época colonial (Arias 2008: 22). Al igual que el trabajo de Donni (1974), se trata de una investigación de índole cualitativa y no incorpora proporciones cuantitativas de cada fenómeno.

En la presente revisión del fenómeno se estudian los GCC por medio de las transcripciones ortográficas de las 48 entrevistas que conforman la base de datos principal de esta tesis. Al no hacer una transcripción auditiva de cada ejemplo, los siguientes resultados deben verse más como una investigación léxica (o fonoléxica). Aunque seguramente una transcripción auditiva confirmaría que se producen reducciones, mutaciones consonánticas o elisiones en la realización de estos grupos consonánticos⁹, esta primera aproximación tratará de interpretar los volúmenes que cada grupo de hablantes (por edad, sexo o nivel de estudios) aporta al corpus total, para así tratar de encontrar patrones de diferenciación sociolingüística que puedan complementar el análisis de la variación fonética revisada en los capítulos anteriores.

⁹ Por supuesto, este análisis sí es posible si se atiende a las precisiones fonéticas que se anotan en ambos corpus, análisis que merece una revisión detenida que por el momento no se hace por tratarse en este capítulo de fenómenos de índole léxica y morfológica, ante todo.

Obsérvese que se estudia en realidad un grupo de palabras vinculadas por una propiedad fónica, la presencia de grupos cultos. En consecuencia, se trata de una variable ante todo *léxica*.

Es necesario hacer algunas precisiones metodológicas. Los grupos consonánticos se buscaron con el programa de concordancias *AntConc* 3.2.4 (Anthony 2014). Lo que se buscó fue el grupo consonántico en sí, y no palabras particulares, lo que implicó un proceso de revisión y descarte de las concordancias que no tenían que ver con el grupo; p. ej., para el grupo /ks/ es necesario buscar la grafía <x>, sin embargo, una gran proporción de concordancias es de topónimos como *México*, *Xochimilco*, *Uxmal*, entre otros que, por supuesto, se descuentan del análisis. Debe aclararse que se mantuvieron los datos de los entrevistadores. Hay tres razones para tomar esta decisión. Por un lado, normalmente el sesgo estadístico introducido por los datos del entrevistador es muy pequeño: la mayor parte del texto de entrevista es aportado por el entrevistado. Por otro, puede especularse que si el entrevistador emite una palabra particular es porque de alguna forma asume que esta forma parte del repertorio lingüístico del entrevistado. Finalmente, con esta decisión se trataba de no perder la información pragmática y el hilo conversacional del texto.

5.4.1. Los grupos consonánticos cultos en los hablantes de altos estudios

La siguiente gráfica muestra las diferencias cuantitativas absolutas en cuanto a los GCC de los grupos de mayores estudios en el tiempo real:

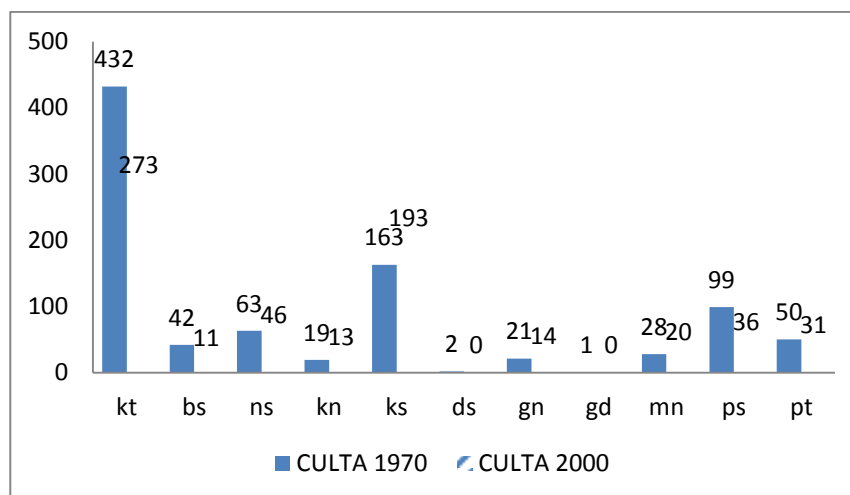


Figura 5.5. Grupos consonánticos cultos en hablantes de estudios altos 1970-2000

En general, los hablantes cultos de 1970 presentan más GCC totales que sus pares del año 2000 y lideran en todos estos, excepto en el grupo /ks/. Los vocablos que lideran el grupo /kt/ (ortográfico <ct>) entre los hablantes cultos de 1970 son *aspecto* (80 veces, pero debe tomarse en cuenta que 70 ejemplos son de un solo informante, varón y joven), el vocablo *doctor* (80 veces) y *exactamente* (24), mientras que en el año 2000 las principales palabras con este grupo son *arquitectura* (29 veces) y *actor* (24), pero ambas son dichas por un solo informante; otras fueron *proyecto* (23, con 16 ejemplos de un solo informante) y *efectivamente* (12)¹⁰. La poca dispersión de las palabras frecuentes en el año 1970 y 2000 hace que estos datos deban tomarse con

¹⁰ Precisamente debido a la distribución muy concentrada en algunos pocos informantes de grupos específicos es que la variable es *GCC de cualquier tipo*; con esto se evita el sesgo que puede producir la alta frecuencia de palabras específicas vinculadas al tema de conversación abordado en la entrevista.

ciertas previsiones. Respecto al grupo /ks/ (<cc, cs, x>), los vocablos *exactamente* (24), *producción* (11 veces, en un solo informante) y *examen* (10) son los mayoritarios para los hablantes cultos de 1970, mientras que los informantes del año 2000 presentan *experiencia* (20) y *explicar* (10) como los vocablos mayoritarios. El tercer grupo consonántico que aporta más datos es el de /ps/¹¹; las palabras más comunes en 1970 con este grupo son *psiquiatría* (25 veces) y *psicología* (16); el total es de 99 datos; en el año 2000 hay sólo 36 ejemplos de este grupo, siendo *psicología* la palabra más común (16 veces). Como puede concluirse, al menos en los grupos de mayores estudios, no puede hablarse de que sean algunos marcadores discursivos (muy frecuentes generalmente) los que lideran la inserción de los GCC, sino que se trata de palabras de claro contenido léxico y que por lo mismo pueden estar condicionadas por la temática de la entrevista. Llama la atención la disminución general de estos grupos consonánticos en el tiempo real. Podría tratarse de un proceso de nivelación de estas hablas cultas hacia un estándar relativamente más informal.

5.4.2. Los grupos consonánticos cultos en los hablantes de bajos estudios

Una hipótesis por revisar aquí sería: los hablantes de pocos estudios, que presentan un menor volumen de GCC respecto a los informantes de altos estudios, pero que aumentaron el volumen en el tiempo real, ¿han introducido estos GCC recientemente

¹¹ Aunque no se trata de un grupo consonántico que se pronuncie en español mexicano, se decidió analizar porque, a final de cuentas, pertenece al léxico “cultivado”.

gracias a piezas léxicas particulares (y frecuentes)? Los resultados con los hablantes de pocos estudios muestran que ellos también tienen un comportamiento diferenciado en el tiempo real. A continuación la gráfica que presenta estas diferencias:

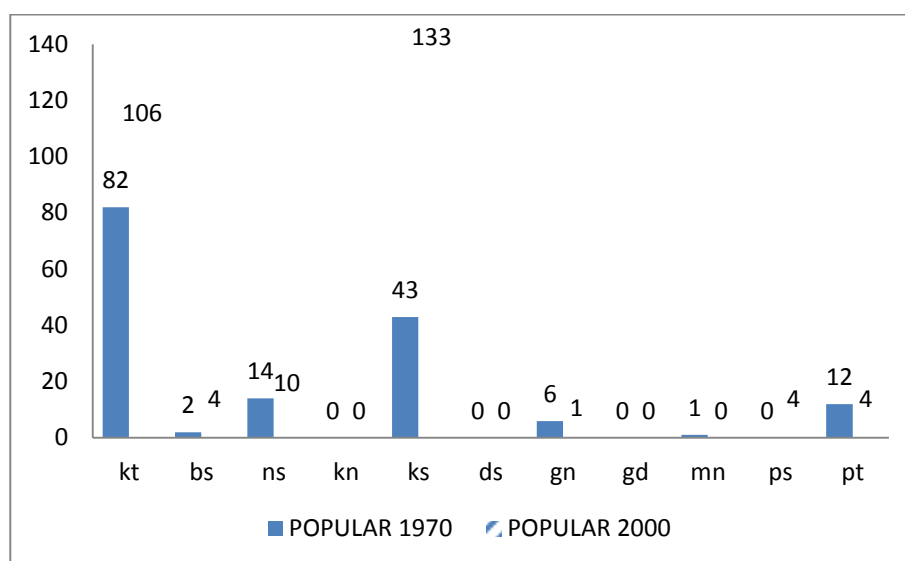


Figura 5.6. Grupos consonánticos cultos en hablantes de bajos estudios 1970-2000

Al igual que ocurre con los hablantes de altos estudios, el grupo /ks/ tiene mayor presencia en el año 2000. Sin embargo, hay que tomar en cuenta que este mayor número de ejemplos del grupo /ks/ en el *habla popular* del 2000 está asociado con dos entrevistas principalmente, donde se hablaba de *bicitaxis*, *taxis*, *taxistas* y *taxímetros*: las palabras más frecuentes con <x> son, en orden descendente: *exacto/exactamente* (22, además, hubo 2 casos del uso de *exacto* como adjetivo),

taxi(s) (21 casos, en 4 entrevistas), *taxímetro* (20, en 2 entrevistas) y *bicitaxi(s)* (15, en una sola entrevista). Ha ganado terreno *exacto/exactamente*, que en 1970 había aparecido sólo en 6 ocasiones en los hablantes de bajos estudios; los 22 ejemplos del año 2000 representarían cerca de la tercera parte de todos los casos de este grupo si sacáramos las palabras con el lexema *taxi* (61 en total; esto dejaría un total de 72 palabras portando el grupo, lo que sigue siendo un aumento significativo respecto a 1970, donde sólo hubo 43 casos totales de /ks/). Aunque son pocos datos en general, podría especularse que quizás se trate de una forma prestigiosa tomada de los hablantes cultos, quienes presentan muchos menos casos de estas dos palabras en el tiempo real: 24 ejemplos en 1970 frente a 7 en 2000; esto es, los grupos sociales intercambiaron los patrones de uso. Aparte de la posibilidad del uso adverbial de *exacto/exactamente* como portadora del grupo consonántico, no existen otras que pudieran vincularse en el proceso de inserción de este grupo consonántico.

Por supuesto, las palabras *exacto/exactamente* también contribuyen al aumento del grupo /kt/, que en el año 2000 va liderado por palabras derivadas a partir de la raíz de *doctor* (30 en total); en el año 1970 las palabras derivadas de *doctor* también eran las más frecuentes (20 en total).

En cuanto a las diferencias entre grupos educativos en el tiempo real, es llamativo que los GCC /kn/ (*técnica*) /ds/ (*adscrita*) no tengan ejemplos en los grupos de pocos estudios. De hecho, son raros algunos grupos incluso entre hablantes

cultivados: p. ej., sólo hay dos casos de /ds/ en 1970 (*adscrita*, en hablantes cultos) y uno sólo de /gd/ (*amígdalas*) (contra cero casos en toda la muestra de 2000).

Finalmente, el aumento de los GCC en las hablas de pocos estudios del año 2000 no parece presentar una asociación con el género de los hablantes: baste como ejemplo que respecto a /kt/, 50 datos son de mujeres y 56 de hombres; mientras que respecto a /ks/, 12 datos son de mujeres y 12 de hombres.

En general, el aporte del uso del marcador discursivo *exacto/exactamente* en el aumento de dos GCC en los grupos populares (/kt/ y /ks/) podría ser el único caso compatible con la hipótesis de difusión léxica que, en este caso, permitiría confirmar un acercamiento sociolingüístico de los hablantes estudiados, lo que se ve en el siguiente subapartado.

5.4.3. Panorama general de los grupos consonánticos cultos en ECM

Una revisión en tiempo real de la comunidad de ECM muestra que los grupos sociales del año 2000 parecen haber acortado la brecha sociolingüística que los separaba ampliamente en los años 1970:

Tabla 5.1. Grupos consonánticos cultos en ECM (1970-2000) (48 informantes)

	1970	2000
CULTA	920	637
POPULAR	160	262
TOTALES	1080	899

Gráficamente, estos resultados vuelven más clara la disminución global de los GCC y el acercamiento de hablantes cultos y populares:

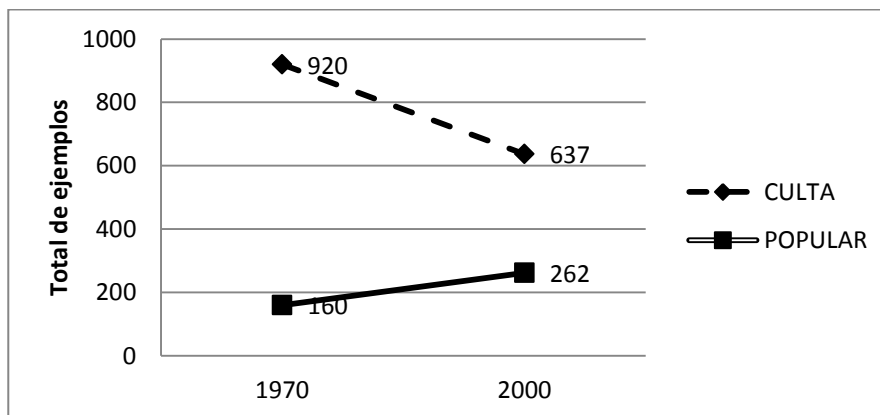


Figura 5.7. Grupos consonánticos cultos en hablantes cultos y populares en ECM (1970-2000)

Esto es, hay menos GCC en el año 2000 (899 totales) respecto a 1970 (1080, una disminución global del 16.8%). Por otra parte, los hablantes cultos de los años 2000 presentan un menor número de GCC que los informantes cultos de los años 1970 (de hecho, el descenso es de 30.8%); entre los grupos de pocos estudios, se da el proceso contrario: un mayor crecimiento de los GCC (en el orden del 63.7%). La comunidad capitalina, por lo tanto, parece más nivelada en el tiempo real.

Los datos sobre GCC sugieren que –con excepción quizás de los grupos /kt/ y /ks/ vinculados a *exacto/exactamente*--, estos no se han insertado recientemente a través de palabras particulares (como defendería la hipótesis de la difusión léxica); la duda es si la disminución generalizada de estos GCC se debe a las limitaciones del

propio corpus y al fuerte condicionamiento del tema de conversación, o si se trata de un fenómeno real de disminución en la comunidad. Por supuesto, aquí se defiende la idea de que el corpus es lo suficientemente fidedigno como para apoyar la hipótesis del cambio comunitario, aspecto que se discutirá en el cap. 6.

Hasta aquí, la revisión de los procesos fonéticos, morfológicos y fonológicos parece confirmar que el cambio lingüístico en este corpus oral de ECM puede seguir distintas rutas: una básicamente neogramática y mecanicista, como sucede con la variación fonética de las róticas, revisada en el cap. 3 y los morfemas *-dad* y *-mente*, y otra compatible con la hipótesis de la difusión léxica –la elisión vocálica analizada en el cap. 4 y, quizás, el crecimiento de algunos GCC en hablantes de bajos estudios. En la siguiente subsección se pasa al nivel léxico para tratar de confirmar las tendencias que los distintos rasgos fonéticos y morfológicos han documentado, en el sentido de que los grupos sociales parecen estar menos diferenciados en 2000, lo que podría sugerir un cambio lingüístico comunitario tendiente a la nivelación dialectal.

5.5. ACERCAMIENTOS LEXICOESTADÍSTICOS A LA DIFERENCIACIÓN SOCIOLINGÜÍSTICA

En esta sección se aborda principalmente la cuestión del papel que juega el léxico en la diferenciación sociolingüística basada en la escolaridad. Se analizarán los repertorios léxicos sustantivos y se utilizará una medida específica de densidad léxica global por informante y por grupos.

Es muy importante aclarar que, en buena medida, esta sección se inspira en los estudios de *sociosemántica* de R. Ávila (1999 a y b), quien defiende la idea de que el léxico es un índice importante de diferenciación sociolingüística; de hecho, Ávila se basó en los materiales de “Norma culta” y “Habla popular” de los años 60-70 para discutir la distribución social del léxico en varios trabajos; de esta manera, resultaba muy interesante confrontar sus hallazgos en el tiempo real¹². Sobre el papel del léxico como factor de diferenciación sociolingüística, Moreno Fernández (2012) ha propuesto varios principios desde una perspectiva sociolingüístico-cognitivista. Cito algunos de ellos, que considero los más pertinentes para el presente trabajo:

Principio 5.1.5

Las comunidades lingüísticas no son homogéneas en cuanto al conocimiento léxico-semántico, dado que está desigualmente distribuido entre sus miembros.

Principio 5.19

La capacidad léxica se muestra como un síntoma o como una variable predictora de la posición social de los hablantes y está fuertemente ligada a su educación.

Principio 5.21

¹² Algo que puede verse también a través de metodologías como las de *disponibilidad léxica* (Ávila Muñoz & Villena Ponsoda 2010).

El sexo y la edad tienen poca influencia en la adquisición, el conocimiento y la distribución de las unidades léxicas.

[Moreno Fernández 2012: 117-118]

De este modo, una pregunta base para esta tesis es si se mantienen las diferencias sociales que Ávila (1999 a y b) reportaba y para ello se compara con el *Corpus Sociolingüístico de la Ciudad de México* (CSCM), un corpus que añade alrededor de 30 años de brecha temporal. Por otra parte, la presente exploración no sólo se vale de técnicas similares a las de Ávila, sino que explora otras alternativas de analizar las diferencias en el léxico de los informantes y evalúa cada una de estas. Al tiempo, servirán estos acercamientos para evaluar la pertinencia de corpus orales basados en entrevistas para el estudio sociolingüístico del léxico.

A lo largo de las siguientes subsecciones se utilizará el concepto de *vocablo* y, alternativamente, *palabra*. Aunque ambas son diferenciadas por Lara (2006) como términos técnicos para la descripción lingüística, el mismo autor reconoce que ambas comparten el hecho de referir en esencia a una *unidad de cita*, que en sus términos se define de la siguiente manera:

La unidad de cita es una forma representante de la palabra, en la que se abstrae la variedad morfológica con que aparece usada en el discurso, para poder

mencionarla con facilidad y para poderla objetivar como unidad de reflexión. En cuanto tal, estrictamente hablando, es la representación del paradigma que forman la raíz o el núcleo morfemático de la palabra y los morfemas ligados en cohesión estrecha con ellos, sobre la base de la unidad de denominación [Lara 2006: 137].

Por tanto, aquí se asume el sentido de *vocablo* (o *palabra*) como “una unidad de cita que representa todo el paradigma de flexiones, derivaciones o conjugaciones que se forman a partir de una raíz o de un núcleo morfemático y que se reconocen como palabras” (Lara 2006: 138).

5.5.1. *Volumen y características textuales de los corpus utilizados*

Los corpus utilizados emulan los que se emplearon en los caps. 3 y 4 sobre cambio fonético. Los informantes de la muestra del año 2000, con algunas excepciones, son los mismos de los capítulos sobre variación fonética. Sin embargo, respecto a los años 1970, no se utilizan los datos de los mismos informantes de los caps. 3 y 4. En su lugar, se emplearon las transcripciones que en su momento estaban disponibles en los libros editados por J. M. Lope Blanch, *El habla de la ciudad de México* (1971) y *El habla popular de la ciudad de México* (1976). Se utilizó una versión digitalizada de dichos materiales, llevada a cabo por investigadores del Centro de Lingüística Hispánica

“Juan M. Lope Blanch” del Instituto de Investigaciones Filológicas de la UNAM¹³. Nuevamente, se utilizan los datos de 30 min. de grabación por persona. Los cuadros 5.3 a 5.6 presentan a los informantes, con los datos de nombre del archivo, edad (en los de 1970) o año de nacimiento (sólo de los informantes del año 2000) y el código utilizado en el proyecto del que se seleccionaron.

Cuadro 5.3. *Habla popular 1970*

	<i>Jóvenes (18-34)</i>	<i>Adultos (35-54)</i>	<i>Mayores (55+)</i>
<i>Hombres</i>	MUESTRA I - JH-POP-01 (19 años)	MUESTRA VI - AH-POP-01 (42 años)	MUESTRA X - VH-POP-01 (59 años)
	MUESTRA II - JH-POP-02 (22 años)	MUESTRA VII - AH-POP-02 (48 años)	MUESTRA XII - VH-POP-02 (72 años)
<i>Mujeres</i>	MUESTRA III - JM-POP-01 (19 años)	MUESTRA VIII - AM-POP-01 (36 años)	MUESTRA XIV - VM-POP-01 (60 años)
	MUESTRA IV - JM-POP-02 (30 años)	MUESTRA IX - AM-POP-02 (45 años)	MUESTRA XV - VM-POP-02 (71 años)

Cuadro 5.4. *Habla culta 1970*¹⁴

	<i>Jóvenes (18-34)</i>	<i>Adultos (35-54)</i>	<i>Mayores (55+)</i>
<i>Hombres</i>	MUESTRA I - JH-CULT-01 (26 años)	MUESTRA IV - AH-CULT-01 (49 años)	MUESTRA VIII - VH-CULT-01 (62 años)
	MUESTRA II - JH-CULT-02 (25 años)	MUESTRA XXXI - AH-CULT-02 (52 años)	MUESTRA XV - VH-CULT-02 (60 años)
<i>Mujeres</i>	MUESTRA III - JM-CULT-01 (25 años)	MUESTRA V - AM-CULT-01 (42 años)	MUESTRA X - VM-CULT-01 (80 años)
	MUESTRA XXVIII - JM-CULT-02 (26 años)	MUESTRA VI - AM-CULT-02 (49 años)	MUESTRA XI - VM-CULT-02 (71 años)

¹³ Agradezco a Cecilia Rojas y Alejandra Viguera el acceso a dichas versiones electrónicas.

¹⁴ Las muestras de 1970 marcadas con negritas indican informantes de quienes se aprovecharon también sus datos en los capítulos sobre fonética. La entrevista XXVIII (mujer joven, culta) fue secreta.

Cuadro 5.5. *Habla popular 2000*¹⁵

	Jóvenes (18-34)	Adultos (35-54)	Mayores (55+)
Hombres	Guillermo (1980) Entrevista 74 JH-POP-01 Pedro (1981) (Sustituido por hombre de 26 años, Entrevista 76) JH-POP-02	Faustino (1957) Entrevista 86 AH-POP-01 Rafael, (1956) Entrevista 88 AH-POP-02	Delfino (1927) Entrevista 97 VH-POP-01 Felipe (1946) (Sustituido hombre de 73 años, Entrevista 102) VH-POP-02
Mujeres	Sofía (1979) Entrevista 79 JM-POP-01 Carmen (1969) (Sustituida por mujer de 32 años, Entrevista 83 JM-POP-02)	Cristobalina (1962) Entrevista 91 AM-POP-01 Irene (1953) ME-286-12M-07 Entrevista 92 AM-POP-02	Ángela (1910) Entrevista 104 VM-POP-01 Berta (1949) Entrevista 106 VM-POP-02

Cuadro 5.6. *Habla culta 2000*

	Jóvenes (18-34)	Adultos (35-54)	Mayores (55+)
Hombres	Fernando (1978) ME-190-31H-05 Entrevista 5 JH-CULT-01 Antonio (1972) ME-197-31H-01 Entrevista 6 JH-CULT-02	Eduardo (1965) ME-254-32H-05 Entrevista 17 AH-CULT-01 Javier (1960) ME-257-32H-05 Entrevista 18 AH-CULT-02	Rodolfo (1941) ME-009-33H-97 Entrevista 25 VH-CULT-01 José Luis (1940) ME-249-33H-05 Entrevista 29 VH-CULT-02
Mujeres	Eugenia (1972) (Sustituida por mujer de 25 años, Entrevista 7) JM-CULT-01) Maripaz (1973) (Sustituida por mujer de 21 años, Entrevista 12) JM-CULT-02)	Ruth (1967) Entrevista 22 AM-CULT-01 Edith (1963) Entrevista 23 AM-CULT-02	Justine (1936) Entrevista 31 VM-CULT-01 Liliana (1941) Entrevista 35 VM-CULT-02

¹⁵ En los cuadros del año 2000 se indica en negritas los informantes sustituidos en el estudio léxico.

5.5.2. Regularización del texto original de las muestras

Los textos de cada submuestra fueron transcritos siguiendo criterios algo distintos en 1970 respecto al año 2000. Aunque en general el sistema en ambos es básicamente ortográfico, en los textos de 1970 (Lope Blanch 1971, 1976) se respetan las convenciones de puntuación generales del español normativo (uso de coma, punto y coma, sangría, etc.). El CSCM, por otra parte, sigue los lineamientos generales de transcripción apegados a las normas propuestas para PRESEEA (Proyecto para el Estudio Sociolingüístico del Español de España y América; Moreno Fernández 2004) que tiene como marco general una perspectiva pragmático-discursiva en la que se pone especial atención a los inicios y toma de turno, traslapes, distintos tipos de pausa, recursos paralingüísticos como la risa, chasquidos, clics o alargamientos vocálicos, etc.; estos últimos se anotan entre paréntesis '()' y las pronunciaciones particulares entre corchetes angulares '< >’.

Los textos originales, además, tuvieron que depurarse para evitar ciertas palabras ajenas al texto propio de la entrevista, como los datos del informante o los encabezados de turno. Jerid Francom (Wake Forest University) realizó la depuración de los textos a base de *scripts* creados por él mismo¹⁶. Los siguientes son fragmentos de una entrevista del corpus Norma culta antes de la depuración y después de esta:

¹⁶ Dichos *scripts* no están compilados en un software autónomo, por ello no puedo citarlo.

EJEMPLO DE TEXTO ORIGINAL DE NORMA CULTA DE 1970

ME-1. Hombre de 26 años. Contador público

Inf.- Bueno, de la natación te puedo platicar muchas cosas. Yo, el deporte de la natación lo seguí, porque mi mamá durante mucho tiempo fue clavadista. Entonces, cuando yo nací, ya mi mamá ya era clavadista... Durante el lapso entre el nacimiento de mi hermano mayor y yo, mi mamá compitió, y posteriormente a mi nacimiento, siguió compitiendo. Yo nací en el año de mil novecientos cuarenta y dos, y la última competencia de ella fue en el año de mil novecientos... cuarenta y ocho, cuando fue la Olimpiada... de Londres.

Enc.- Perdona que te interrumpa un momento. ¿Sabes qué, Laura? ¿Me traes una mesita?

Otra persona.- Una silla o algo, ¿verdad?

Enc.- No... Me estabas hablando de la última competencia de tu mamá; fue en el año ¿qué?

Inf.- Entonces, la última competencia en la que ella participó fue en la Olimpiada de Londres, en mil novecientos cuarenta y ocho. Posteriormente a esa competencia, creo que tuvo una competencia más, que fue en Barranquilla. Entonces ella, dentro del ambiente de la natación en el cual ya tenía mucho tiempo... nos inició a nosotros. Y, entonces, yo empecé a aprender a nadar a la edad de dos años.

EJEMPLO DE TEXTO DEPURADO DE NORMA CULTA DE 1970

Bueno, de la natación te puedo platicar muchas cosas. Yo, el deporte de la natación lo seguí, porque mi mamá durante mucho tiempo fue clavadista. Entonces, cuando yo nací, ya mi mamá ya era clavadista... Durante el lapso entre el nacimiento de mi hermano mayor y yo, mi mamá compitió, y posteriormente a mi nacimiento, siguió compitiendo. Yo nací en el año de mil novecientos cuarenta y dos, y la última competencia de ella fue en el año de

mil novecientos... cuarenta y ocho, cuando fue la Olimpiada... de Londres. Perdona que te interrumpa un momento. ¿Sabes qué, Laura? ¿Me traes una mesita? Una silla o algo, ¿verdad? No... Me estabas hablando de la última competencia de tu mamá; fue en el año ¿qué? Entonces, la última competencia en la que ella participó fue en la Olimpiada de Londres, en mil novecientos cuarenta y ocho. Posteriormente a esa competencia, creo que tuvo una competencia más, que fue en Barranquilla. Entonces ella, dentro del ambiente de la natación en el cual ya tenía mucho tiempo... nos inició a nosotros. Y, entonces, yo empecé a aprender a nadar a la edad de dos años.

En estos textos se eliminan los encabezados de turno de informantes y entrevistadores (*Inf.* y *Enc.*). Es importante aclarar que para esta sección y con excepción de las muletillas fáticas como *mh* o *mm*, no se elimina el habla del encuestador, que nunca es demasiado voluminosa respecto a la del informante; además, se consideró interesante mantener en lo posible el *input* que recibe el informante durante la entrevista.

En el caso del corpus CSCM, se buscaron y remplazaron con espacio en blanco las secuencias que encabezaban los turnos, como "I: (tab)" o "E: (tab)" y se eliminaron del recuento todas las formas entre corchetes angulares (que indican pronunciaciones particulares). Un ejemplo de transcripción CSCM ya depurada se presenta a continuación:

EJEMPLO DE TEXTO DEPURADO DE HABLA CULTA 2000 (CORPUS CSCM)

a ver I okey dime dime ¿cómo te la pasaste en la universidad? yo me la pasé muy digo yo me la estoy pasando muy bien este aprendo y pues diría que bien bien lo que pasa es de que como yo fui en vocacional por ahí había problemillas en cuestión de porros y todo eso mh y luego el ambiente estaba un pesado y era un poquito difícil y vivir con la tensión y no sé no podías llevar walkman o cosas así porque te robaban a mí ya me robaron una vez y no fue una muy buena experiencia y como que lo que yo quería era salir de la escuela mh y el cambio fue algo muy bueno porque cuando yo entré a la superior ajá este ya en la superior nadie se mete contigo yo estuve ahí en Zacatenco en Esime y este y pues ya ahí pues tú a tu rollo ¿no? ya nadie se metía contigo y de hecho hasta los maestros en la vocacional los maestros por la forma por tu aspecto físico meterse contigo

Las diferencias entre los tipos de transcripción en 1970 y 2000 forzaron a implementar criterios de análisis lexicográfico y lexicoestadístico algo distintos para cada año. Por ejemplo, el programa utilizado --*AntConc*-- permite que se excluya de las listas de formas gráficas (*types*) toda la información entre corchetes angulares '< >', cuadrados '[']' o paréntesis '()', pero no en dos o tres tipos de paréntesis a la vez. Para el caso del CSCM, que utiliza dos tipos (paréntesis redondos y corchetes angulares), se decidió excluir la información entre corchetes angulares y por lo tanto hubo que hacer una *stop list* (lista de exclusión) que ignorara las formas que se presentaron entre paréntesis redondos y los falsos comienzos que también se indican en dicho corpus (transcribiendo la letra seguida de un guión breve, ejs.: *f- falso* o *tamb- también*). Por supuesto, los textos de 1970 también indican algunos falsos comienzos y se utilizaron

los corchetes cuadrados para las expresiones paralingüísticas, como *risas*. La mencionada limitación del programa *AntConc* de sólo permitir excluir un tipo de información parentética a la vez, exigió la revisión exhaustiva de las más de 23,000 formas gráficas (*types*) obtenidas y verificar las cadenas de caracteres inusuales en todas sus concordancias.

Un problema adicional en el corpus de 1970 es que presenta transcripciones alternativas (fuera de la norma) para la misma palabra, sobre todo en habla popular, como <vinu> por *vino* en el habla de informantes de zonas rurales. Dada la multitud de estas transcripciones alternativas, no pudo evitarse que aparecieran en las listas finales estas formas pero, por supuesto, no se contabilizaron como tipos aparte, sino como simples alternativas de sus formas canónicas y se asignaron en el análisis lexicológico al mismo vocablo.

5.5.3. Palabras y tipos totales en el corpus

Como se mencionó anteriormente, para el análisis de los acervos léxicos se utilizó el programa gratuito *AntConc 3.2.4w* de Laurence Anthony (2014). El programa permite crear listas de palabras, encontrar concordancias, hace análisis de *clusters* y colocaciones, entre otras funciones. Previo al recuento de palabras, se aplicaron *stop lists* (listas de exclusión) para cada submuestra. Las listas de exclusión indican al programa que ignore en los recuentos las cadenas de caracteres que el analista indique. El número de palabras gráficas por submuestra se presenta a continuación.

Cuadro 5.7. *Palabras gráficas totales (tokens) y tipos (types) en cada submuestra*

	<i>1970</i>		<i>2000</i>	
	<i>types</i>	<i>tokens</i>	<i>types</i>	<i>tokens</i>
<i>Culta</i>	6,549	56,663	6,225	61,880
<i>Popular</i>	5,362	57,786	4,703	55,584

La diferencia entre grupos cultos y de habla popular se mantiene en el tiempo: los hablantes de menos estudios presentan menos *types* que los hablantes cultos.

Esta es la base de datos para el análisis del vocabulario. En los siguientes apartados se presentan distintos acercamientos al problema del léxico, al tiempo que se discuten sus ventajas e inconvenientes.

5.5.4. *Estudio I: listas de vocablos sustantivos más frecuentes*

Como se mencionó en la introducción de este capítulo, el análisis de la variación social asociada al léxico es un tema tratado por Raúl Ávila en varios de sus trabajos sobre el español mexicano (Ávila 1988, 1999 a y 1999 b). En dichas investigaciones, Ávila ha propuesto ciertos acercamientos metodológicos para tratar de determinar qué es lo que distingue a los grupos sociales en cuanto a su acervo léxico. Por ejemplo, en el trabajo “Sobre semántica social: conceptos y estratos en el español de México” (Ávila

1999 a: 59-104)¹⁷, encontró que los primeros 20 referentes sustantivos en los grupos socioeconómicos alto y bajo de un corpus levantado a nivel nacional (en el marco del ALM), eran los siguientes.

Tabla 5.2. *Vocablos sustantivos propios de habla culta y popular (1970) (basada en Ávila 1999 a: 62)*

	<i>Culta</i>	<i>Popular</i>
1	cultura	chile
2	aspecto	tortilla
3	serie	caña
4	valor	kilo
5	profesión	caldo
6	país	litro
7	curso	cebolla
8	obra	gallina
9	educación	baile
10	sala	sal
11	sistema	pescado
12	estrella	sopa
13	historia	pollo
14	idea	semilla
15	interés	milpa
16	médico	molino
17	dato	manteca
18	arte	patrón
19	examen	verdura
20	tema	tamal

Para Ávila estas listas evidencian una fuerte tendencia a los referentes “no perceptibles” en el habla de los grupos más educados (17 vocablos están en dicha

¹⁷ Otro trabajo que se comentará es el de “Sociosemántica: referentes sustantivos y verbales en el habla culta y popular de la ciudad de México” (Ávila 1999 b: 153-219), donde se hace una propuesta de clasificación de referentes sustantivos que se utilizará en este capítulo.

clasificación), que contrasta con los referentes concretos (“perceptibles por alguno de los cinco sentidos”) del grupo bajo. El procedimiento que siguió dicho autor para obtener las listas se detalla a continuación:

Los porcentajes se obtuvieron de acuerdo con la ecuación $\text{Frecuencia Mayor} - \text{Frecuencia Menor} \text{ entre } \text{Frecuencia Menor} \text{ por } \text{Cien}$. Esto quiere decir que si en un estrato la frecuencia fue de 10, en el otro debió ser por lo menos de 20 (tener un crecimiento porcentual del 100.00%) para que fuera incluido el vocablo en la lista [Ávila 1999 a: 63, n. 5].

Los resultados de la comparación del léxico entre grupos sociales (o en el tiempo) pueden ser algo distintos según sea el procedimiento analítico seguido. En el caso de la presente investigación, y en contraste con las investigaciones de Ávila, se decidió explorar otras posibilidades, como lo es la identificación del léxico *no* compartido entre los distintos grupos sociales. La idea es que las palabras que aparezcan en los datos de un grupo social y nunca en el otro, pero que tengan cierta dispersión y frecuencia mínima (esto es, que sean frecuentes en ese grupo), podrían ser más claramente “propias” del grupo que las palabras compartidas con diferente distribución en términos de frecuencia. Los procedimientos específicos se detallan en el siguiente apartado.

5.5.4.1. Un método alternativo para comparar el acervo léxico

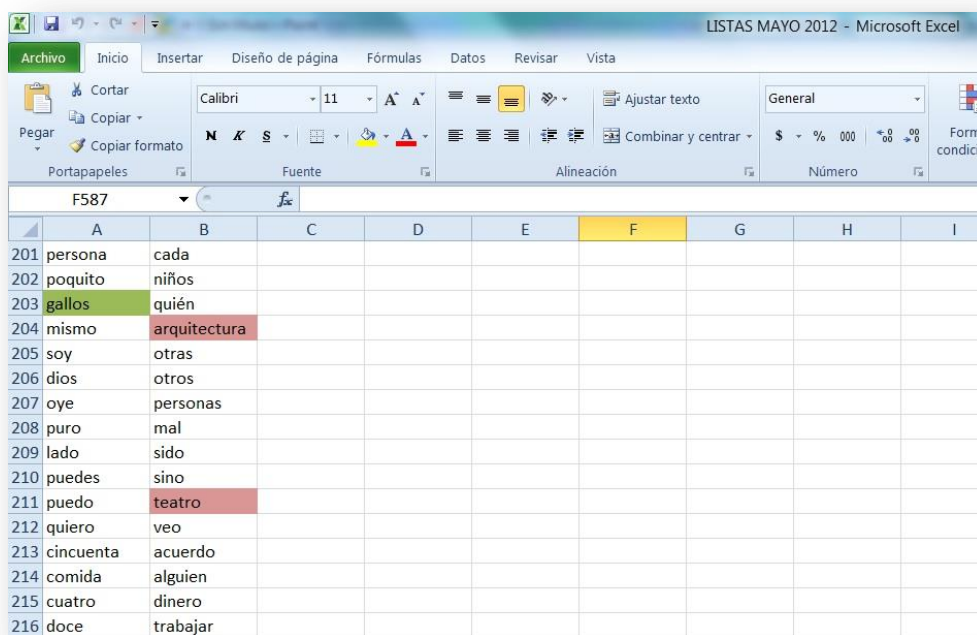
El procedimiento para extraer el léxico no compartido fue el siguiente: se obtuvo una lista de palabras con el programa *AntConc* en cada una de las siguientes 4 submuestras: hablantes cultos de 1970, de habla popular 1970, hablantes cultos del año 2000 y de habla popular del mismo año. Las listas son como la mostrada a continuación.

Corpus Files		Concordance		Concordance Plot	File View	Clusters	Collocates	Word List	Keyword List
		Hits		Total No. of Word Types: 6356		Total No. of Word Tokens: 64813			
Rank	Freq	Word	Lemma Word Form(s)						
1	2834	que							
2	2574	y							
3	2563	no							
4	2078	de							
5	1612	a							
6	1573	la							
7	1226	sí							
8	1170	en							
9	1026	es							
10	965	el							
11	891	pues							
12	885	me							
13	786	lo							
14	763	o							
15	733	yo							
16	698	pero							
17	687	se							
18	656	ya							
19	617	un							
20	567	los							
21	525	por							
22	520	como							
23	516	este							
24	489	con							

Figura 5.8. Ejemplo de lista de palabras obtenida con el programa AntConc

El programa construye una lista de palabras tipo (*types*); cada palabra tipo (en adelante, ‘tipo’) va acompañada de su frecuencia total en el corpus (*Freq*) y su valor

ordinal (*Rank*). El siguiente paso es asignar cada tipo a un solo *vocablo*, lo que se anotó en una nueva columna de vocablos. Al terminar la lista de vocablos finales, se realizó la comparación de las listas de cada grupo social y temporal para obtener las palabras compartidas y las palabras exclusivas en cada uno de ellos¹⁸. Tras ajustar los parámetros respecto al número de casillas por comparar, las casillas que incluyen palabras exclusivas de una lista se marcan con el color que el usuario indique, como en la figura siguiente.



	A	B	C	D	E	F	G	H	I
201	persona	cada							
202	poquito	niños							
203	gallos	quién							
204	mismo	arquitectura							
205	soy	otras							
206	dios	otros							
207	oye	personas							
208	puro	mal							
209	lado	sido							
210	puedes	sino							
211	puedo	teatro							
212	quiero	veo							
213	cincuenta	acuerdo							
214	comida	alguien							
215	cuatro	dinero							
216	doce	trabajar							

Figura 5.9. Ejemplo de listas comparadas en el programa Excel

¹⁸Se utilizó para este proceso una fórmula en el programa *Microsoft Excel 2010*: $=CONTAR.SI(\$B\$2:\$B\$10,A2)=0$.

La última etapa consistió en hacer búsquedas de concordancias en *AntConc* y verificar que los tipos de cada vocablo fueran estrictamente sustantivos y que tuvieran el mismo valor referencial. La figura siguiente muestra el uso de los programas durante este proceso. En el ejemplo, se está verificando la palabra *proyecto*.

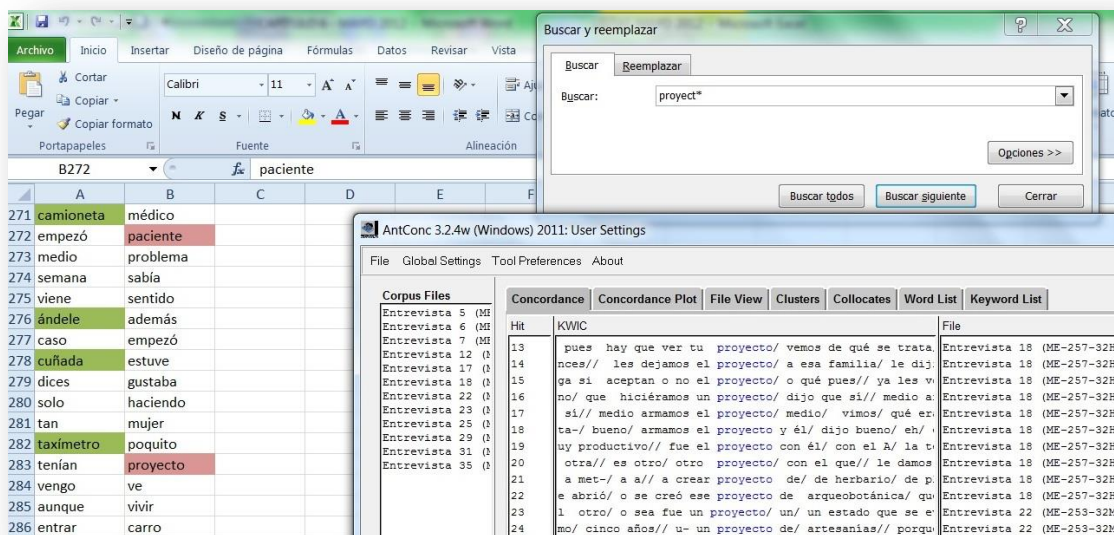


Figura 5.10. Uso de AntConc y Excel en la verificación de las palabras exclusivas de cada grupo social

Tras la extracción de los vocablos exclusivos de cada grupo, se obtuvieron las siguientes listas de los primeros 20 sustantivos con mayor presencia estadística (que llamaremos listas “revisadas”). Los criterios para que los vocablos entraran en la lista (Lara 1990: 100) fueron: a) que tuvieran al menos 3 apariciones; b) que hubieran

aparecido al menos en el habla de 2 informantes distintos (dispersión = 2); c) que no fueran nombres propios o topónimos¹⁹.

Tabla 5.3. Los 20 vocablos sustantivos más frecuentes en habla popular 1970 (revisada)

	Vocablo	Frecuencia	Dispersión
1	zapato	120	6
2	centavo	21	7
3	terreno	21	4
4	máquina	20	3
5	maíz	19	4
6	fútbol	16	3
7	taller	16	3
8	pueblo	15	5
9	tortilla	14	5
10	llave	13	3
11	rancho	10	4
12	frijol	10	3
13	banca	10	2
14	rancho	10	2
15	grasa	9	3
16	kínder	9	2
17	basura	7	2
18	cajón	7	2
19	cera	7	2
20	trabajador ²⁰	4	4

¹⁹ Los criterios a y b (frecuencia mínima de 3 y dispersión=2) son los sugeridos por Lara (1990) respecto al *Diccionario del español de México*: “En realidad, ya tres ocurrencias de un vocablo, repartidas en más de un texto, son suficientes para considerarlo ‘existente’ en una lengua (es decir, que no se trate de un vocablo de uso estrictamente individual y que tampoco sea un *hápax* [cadena de caracteres que aparece una sola vez en todo el corpus]), por lo que, quizás con poco valor estructural, pero con validez documental, se puede confiar uno en los resultados que ofrece el CEMC [Corpus de Español Mexicano Contemporáneo] con tan baja frecuencia” (1990: 100). Si se toma en cuenta que el CEMC estaba conformado por cerca de 2, 000,000 de palabras y el presente corpus oral es de 250,000 (la octava parte de aquel), los criterios son aún más pertinentes.

En esta lista de vocablos en hablantes de habla popular todos los referentes son concretos –algo coincidente con Ávila (1999 a), pero sólo se comparte una palabra con aquella lista: *tortilla*. Otra diferencia con la lista revisada es que la mayoría de vocablos en la lista de Ávila (15 de 20) pertenecen al campo de la alimentación —son comestibles o pertenecen al campo de la cocina y la preparación de alimentos, como *chile, frijol, sal, tortilla, manteca, litro, caldo, kilo*, etc.— mientras que la nueva lista presenta una mayor diversidad de campos referenciales: ahora sólo tres vocablos son comestibles (*tortilla, frijol, maíz*), los demás se concentran en lugares (*rancho, kínder*), objetos de la vida cotidiana (*basura, llave, zapato, centavo*) o vocablos asociados con el trabajo u oficios (*grasa, cera, taller, máquina*). El promedio de dispersión es de 3.1 informantes por vocablo.

Tabla 5.4. *Los 20 vocablos sustantivos más frecuentes en habla culta 1970 (revisada)*

	Vocablo	Frecuencia	Dispersión
1	aspecto	80	6
2	materia	39	4
3	química	38	2
4	médico	35	7
5	caso	33	7
6	elemento	28	2
7	curso	26	5
8	competencia	23	2
9	barco	22	3
10	vocación	20	3
11	tesis	18	6

²⁰ Obviamente se trata de los ejemplos sustantivos. Hubo 3 casos de la palabra *trabajador* con valor adjetivo.

12	país	18	4
13	investigación	18	2
14	siglo	17	4
15	objeto	15	3
16	templo	15	3
17	psicoanálisis (y sicoanálisis)	15	2
18	serie	13	5
19	sistema	13	4
20	psicología	13	3

La lista revisada de vocablos más frecuentes en grupos cultos comparte 5 vocablos con la lista de Ávila (1999 a): *aspecto, curso, médico, país, serie y sistema*. En la lista revisada, aparece casi el mismo número de vocablos perceptibles que en la lista original: 6 en la lista revisada, 5 en la de Ávila. El promedio de dispersión es un poco más alto que en los grupos de pocos estudios: 3.85 informantes en promedio por cada vocablo enlistado. El hecho de que coincidan 5 vocablos en las listas revisadas y las originales de Ávila (1999 a) indican que muy probablemente dicho autor también contó palabras exclusivas de hablantes cultos (y por lo tanto inexistentes en el acervo de habla popular).

Esta breve comparación pone en evidencia que lo que es “propio” léxicamente de un grupo u otro puede variar fuertemente según el método aplicado. En este caso, al focalizar la atención sobre los vocablos no compartidos, se destaca que las listas de los 20 vocablos más recuentes de 1970 y 2000 comparten muy pocos ítems y que la lista en los grupos de menos estudios presenta en la revisión –de manera más realista, desde mi perspectiva-- una mayor diversidad de campos referenciales, lo que sugiere que estos hablantes no hablan de comida o su preparación tan marcadamente, como

lo indicaría la lista con datos nacionales de Ávila (1999 a)²¹ --que después se corregiría con datos de la ciudad de México en Ávila (1999 b).

Tras comparar los datos de esta tesis con los de Ávila --ambos estudios concentrados en muestras de los años 1970--, se decidió crear las mismas listas comparativas en la muestra del año 2000 siguiendo, por supuesto, los mismos criterios de inclusión utilizados en las tablas anteriores.

Tabla 5.5. *Los 13 vocablos sustantivos más frecuentes en habla popular 2000*

	<i>Vocablo</i>	<i>Frecuencia</i>	<i>Dispersión</i>
1	gallo	44	2
2	fruta	25	3
3	camioneta	24	5
4	taxímetro	24	2
5	tianguis	18	6
6	bicitaxi	16	2
7	Revolución (época histórica)	15	2
8	sobrino	11	5
9	pollo	11	2
10	centavo	8	3
11	barrio	7	2
12	obrero	7	2
13	comadre	6	2

De entrada, no existen siquiera 20 vocablos en la lista que no hayan aparecido en habla culta y que tengan al menos una dispersión de 2, lo que puede considerarse un indicio de un repertorio “exclusivo” disminuido en las entrevistas de habla popular respecto a las de los años 1970, o también como un cierre de la brecha entre grupos

²¹ Hay que reconocer la posibilidad de que en el levantamiento popular haya muchos informantes de zonas rurales.

cultos y de pocos estudios. Sólo un vocablo coincide en el tiempo real: *centavo*. Doce de los trece vocablos son perceptibles.

Tabla 5.6. *Vocablos sustantivos más frecuentes en habla culta 2000*

	<i>Vocablo</i>	<i>Frecuencia</i>	<i>Dispersión</i>
1	paciente	44	2
2	arquitectura	33	2
3	teatro (como arte)	31	3
4	proyecto	26	5
5	alumno	18	4
6	boda	17	2
7	playa	16	6
8	inglés	16	5
9	psicología	16	4
10	personaje	16	2
11	edificio	15	4
12	país	14	4
13	etapa	13	7
14	medicina (como disciplina)	13	4
15	oficina	12	4
16	biología	12	2
17	especialidad	12	2
18	sociedad	10	3
19	conciencia	9	5
20	asunto	9	4

Al igual que en los años 1970, se mantiene la tendencia a que prevalezcan vocablos de objetos o entidades no perceptibles (14 de 20). Se repiten sólo dos vocablos en el tiempo real: *país* y *psicología*. A continuación se evalúan estos resultados.

5.5.4.2. Evaluación del método alternativo de comparación de vocabularios

Este método mostró que los resultados pueden ser muy distintos a los seguidos en Ávila (1999 a); las diferencias pueden deberse a que las listas revisadas presentan vocablos con una dispersión mínima de 2 y además reflejan una mayor diversidad de referentes, lo que sugiere que quizás sea un método más adecuado. Por otra parte, para mejorar el procedimiento, pueden establecerse criterios más exigentes de inclusión para evitar vocablos prácticamente idiosincráticos que entraron en las listas; varios vocablos en las 4 listas obtenidas concentran demasiados ejemplos en un solo informante; estos son: *química*, *elemento* (en culta 1970), *gallo*, *taxímetro*, *bicitaxi*, *Revolución* (popular 2000), *zapato*, *fútbol*, *rancho*, *grasa* (popular 1970), *arquitectura*, *teatro* y *boda* (culto 2000).

Algo que sí puede confirmarse con cualquiera de los dos métodos, es que los referentes perceptibles son predominantes en las hablas populares, de manera opuesta a las hablas cultas, donde los primeros vocablos tienen referentes más abstractos. Las ocupaciones de los informantes, por supuesto, condicionan muchísimo las elecciones léxicas: los informantes de habla popular trabajan en oficios como albañilería, carpintería y peluquería, en el servicio doméstico o el comercio, actividades que obligan a tener referentes concretos; mientras que los hablantes cultos son en su mayoría profesionistas, investigadores, profesores o estudiantes universitarios y principalmente los tópicos en la entrevista giran en torno a su profesión, o intereses que implican hablar de referentes más abstractos. Por supuesto,

no creo en las posibles consecuencias y limitaciones cognitivas o psico-sociales de estas diferencias como en la propuesta de códigos amplios y restringidos de Bernstein (1984). Se asume, más bien, que los seres humanos hablan del entorno social y económico que les toca vivir y adaptan su discurso a dichas circunstancias, algo discutido ampliamente por Ávila (1999 b: 176-178).

En el siguiente apartado se revisará un análisis global de todos los referentes sustantivos en estos grupos sociales en el tiempo real, para tener otro parámetro de comparación sobre la evolución del léxico en español capitalino.

5.5.5. Estudio II: Un análisis alternativo de la variación léxica. El caso de los sustantivos

Otra manera de investigar las diferencias sociales o temporales reflejadas en el léxico es a través del análisis del volumen de vocablos totales²², en lo que podría verse como

²² La base de esta investigación se realizó en el segundo semestre de 2011 y se presentaron los resultados preliminares en el *XI Congreso Nacional de Lingüística* (Chetumal, Quintana Roo) (Serrano 2011 b). La base de datos ahí agregó los datos de dos informantes al grupo culto de 1970 (haciendo un total de 14 informantes) para acercar el número de tokens al de los demás subgrupos, que en ese momento era de alrededor de 60,000 por submuestra. En sus comentarios en el mencionado congreso, Luis Fernando Lara rechazaba la idea de observar diferencias sociolingüísticas en corpus basados en entrevistas (aunque sí reconocía esta posibilidad en materia de variación fónica); parafraseándolo, su argumento era el siguiente: “todos los hablantes tienen los mismos niveles de riqueza léxica, siempre y cuando la temática de la conversación gire en torno ‘de lo que realmente *saben*’, de sus especialidades y oficios”. En el corpus analizado existen informantes de pocos estudios que hablan de sus oficios y, sin

un ejercicio de exploración de la *riqueza léxica* que, por lo general, suele abordarse con variados recursos de elicitación, como en los atlas dialectológicos o los estudios de *disponibilidad léxica* (v.g., Ávila Muñoz & Villena Ponsoda 2010). Esta investigación exploró las diferencias sociales y en el tiempo real en los vocablos sustantivos. Dado que se partió de una base de cerca de 240,000 palabras y más de 20,000 sustantivos, y que no se contaba en el momento del análisis con un programa automatizado que hiciera el reconocimiento de vocablos sustantivos, la solución para este análisis fue contar los sustantivos “cien por ciento seguros”, esto es, aquellas formas que no requerían una revisión de cada concordancia para estar seguros de que pertenecían a dicha clase léxica: de esta manera, tipos como *sueldo*, que podía tomarse como forma verbal conjugada (1ra sg, pres, ind del verbo *soldar*) o un sustantivo propiamente ('salario'), se marcaron en la base de datos con un signo de interrogación para un análisis ulterior. El cuadro siguiente muestra el total de tokens, types y sustantivos

embargo, no presentan niveles de riqueza léxica especialmente elevados. Una hipótesis que podría explorarse es que la riqueza en las hablas “populares” se encuentra en la morfología derivativa o en la gran cantidad de recursos lúdicos como fraseologismos, juegos de palabras, doble sentido, etc., como se puede observar en tesis recientes sobre el tema (Sánchez de la Barquera 2014). En todo caso, ya se ha señalado que las clases medias no son necesariamente más eficaces para comunicarse: “Labov (1972: 179) ha señalado la grandilocuencia de la clase media, su tendencia a matizar en exceso y su frecuente actitud de buscar *status* mediante recursos lingüísticos, lo que a veces produce discursos vacíos o difíciles de comprender” (Ávila 1999 b: 177). En realidad, la eficacia comunicativa no está relacionada con un gran acervo léxico, sino con la manera de explotar las posibilidades expresivas de la lengua en todos sus niveles estructurales.

donde se incluyen tanto los sustantivos seguros como aquellos marcados con signo de interrogación.

Cuadro 5.8. Total de tokens, types y sustantivos en 1970 y 2000

	1970			2000		
	<i>Tokens</i>	<i>Types</i>	<i>Sust.</i>	<i>Tokens</i>	<i>Types</i>	<i>Sust.</i>
<i>Culta</i>	58,290	6,755	2,013	66,253	6,505	1,842
<i>Popular</i>	63,205	5,209	1,269	59,677	4,947	1,365

Se podría argumentar que las desproporciones en el volumen de datos totales pueden explicar las diferencias en el número de vocablos encontrados, pero un análisis detenido de cada grupo de hablantes demuestra que no existe una correlación constante entre el número de palabras totales y el número de tipos. La muestra con más palabras (habla culta 2000 –en adelante HC 2000) tiene 13.6% más palabras que la que menos palabras presenta (habla culta 1970 –HC 1970) y sin embargo presenta un menor acervo de sustantivos. Por otra parte, el grupo de habla popular 1970 (en adelante HP 1970) tiene 8.4% más datos que HC 1970 y 5.9% más que el grupo de habla popular del año 2000 (HP 2000); aun así, su acervo de sustantivos es menor en ambos casos. Puede suponerse que de tener el grupo culto moderno el mismo número de datos que habla popular 1970, se mantendría la tendencia de este último grupo a presentar menos vocablos.

Cada forma (*type*) en las listas obtenidas con el programa *AntConc* se asignó a un vocablo en una quinta columna²³. En la siguiente figura se muestra como las formas *amiga*, *amigas*, *amigo*, *amigos*, *amigueta*, *amiguitas* se asignan al mismo vocablo *amigo*.

	A	B	C	D	E	F	G	H	I
259		259	1 amas	?					
260		260	5 amater						
261		261	1 amateur						
262		262	1 ambientaban						
263		263	6 ambiente	S	ambiente				
264		264	1 amenazaba						
265		265	1 amenazaron						
266		266	1 americana						
267		267	3 amiga	S	amigo				
268		268	8 amigas	S	amigo				
269		269	9 amigo	S	amigo				
270		270	18 amigos	S	amigo				
271		271	2 amiguera						
272		272	1 amigueta	S	amigo				
273		273	1 amiguitas	S	amigo				
274		274	1 amistad	S	amistad				
275		275	3 amistades	S	amistad				
276		276	1 amistoso						
277		277	1 amistá	S	amistad				
278		278	2 amolar						
279		279	1 amonestación	S	amonestación				
280		280	3 amor	S	amor				

Figura 5.11. Ejemplo de identificación de vocablos sustantivos en la listas de palabras del subcorpus habla popular 1970

Ya agrupadas las distintas realizaciones de cada sustantivo en un solo vocablo (por ej.: *abuelo*, *abuelitos*, *abuelitas*, se agrupan en el vocablo *abuelo*), se depuró la lista

²³ Agradezco a Yulma Mejía su ayuda para identificar vocablos sustantivos en una sección del grupo culto del año 2000.

para incluir sólo los sustantivos ‘seguros’. A pesar de no considerar los datos de los sustantivos dudosos, el corpus sigue siendo suficientemente voluminoso como para poder realizar exploraciones confiables. Los acervos de cada submuestra quedan distribuidos de la siguiente manera:

Cuadro 5.9. Número de vocablos sustantivos seguros por submuestra

	1970	2000
<i>Culta</i>	1,571	1,453
<i>Popular</i>	949	1,036

La proporción de vocablos sustantivos ‘seguros’ ronda del 25 al 30% en todos los grupos; dichas proporciones se ordenan de la siguiente manera: 29.8% de sustantivos en HC 1970 > 28.3% en HC 2000 > 27.5% en HP 2000 > y finalmente 24.3% en HP 1970.

Una manera de ver la evolución respecto al volumen de sustantivos en el tiempo real puede ser la siguiente:

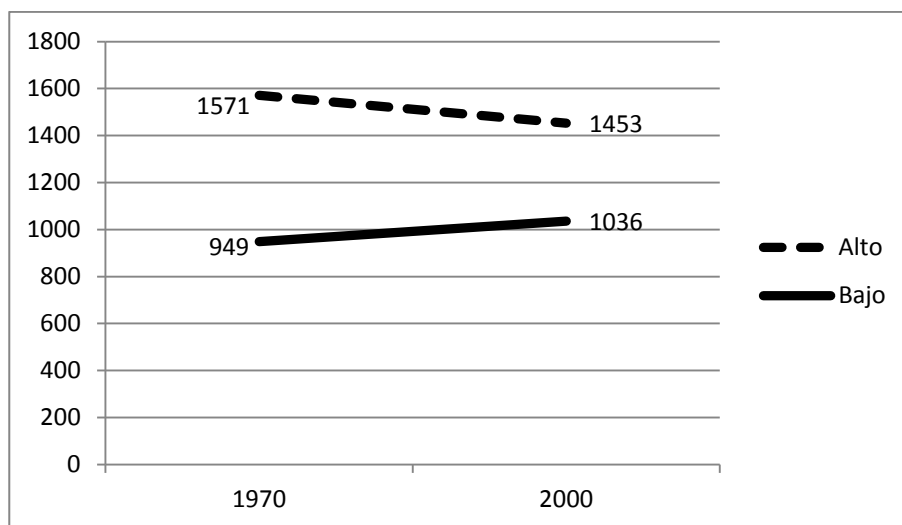


Figura 5.12. *Volumen de vocablos sustantivos en hablantes de estudios Alto y Bajo (1970-2000)*

Los hablantes cultos presentan una disminución del 7.6% en 30 años. Por el contrario, los grupos de habla popular han aumentado en el mismo período su acervo (un 9.1% para ser exactos). Esta tendencia de los grupos a reducir la brecha sociolingüística se puede confirmar de otras maneras. Por ejemplo, puede analizarse el volumen de vocablos comunes en ambos grupos sociales en cada período o, por el contrario, el número de vocablos que no comparten estos mismos grupos --las 'peculiaridades léxicas' de cada submuestra, por decirlo de otra manera.

De esta forma se realizaron comparaciones entre grupo educativos sincrónica y diacrónicamente. En las siguientes gráficas se muestra la variación en el volumen de vocablos compartidos y no compartidos entre los grupos educativos en el tiempo.

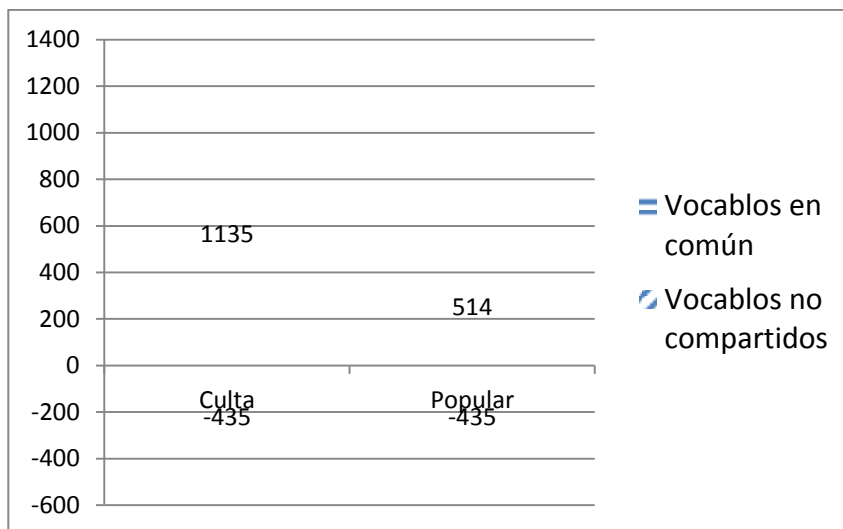


Figura 5.13. Acervo de vocablos compartidos y no compartidos en habla culta y popular (1970)

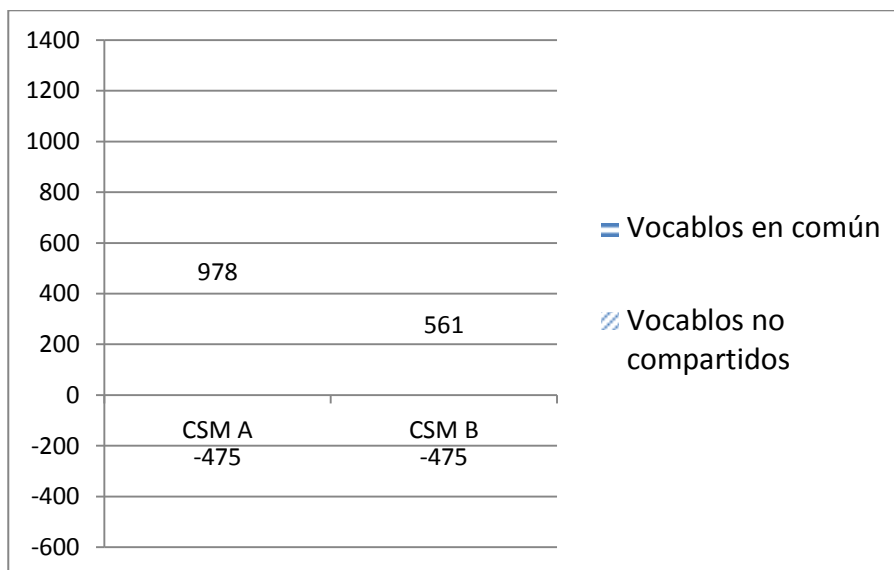


Figura 5.14. Acervo de vocablos compartidos y no compartidos en habla culta y popular (2000)

El léxico compartido entre HC y HP en 1970 es de sólo 435 palabras. El grupo HP tiene 514 palabras propias al comparar su acervo con el de los hablantes HC,

quienes presentan por su parte un total de 1135 palabras que nunca aparecen en los hablantes de HP. Esto es, nos encontramos ante grupos de hablantes fuertemente diferenciados en el nivel léxico nominal.

Las diferencias entre grupos sociales, sin embargo, parecen disminuir con el paso del tiempo. El léxico que comparten ahora estos grupos educativos es un poco mayor que en los años 1970: un total de 475 lexemas. Esto equivale a un incremento del léxico sustantivo compartido a una tasa de 9.1% en 30 años. Para el año 2000, los hablantes cultos tienen 978 palabras propias, frente a los hablantes de habla popular, quienes presentan 561 palabras propias. Esto es, los grupos populares incrementaron su índice de vocablos originales alrededor de un 9.1% (47 ítems más) y el grupo culto bajó en este rubro en un 13.8% (157 palabras menos). Aunque sigue existiendo una brecha importante entre los grupos educativos alto y bajo, es llamativo su relativo acercamiento en cuanto al número de referentes sustantivos característicos.

Finalmente, vale la pena comparar el acervo global en los años 1970 y 2000, independientemente de la diferenciación social.

Cuadro 5.10. Sustantivos totales en 1970 y 2000

	1970	2000
<i>Sustantivos</i>	2084	2014

La diferencia global es mínima: sólo 70 vocablos distintos tiene de más la muestra de 1970 frente a la de 2000, esto es, una diferencia en volumen de 3.35%, lo que significa que la comunidad de habla conserva prácticamente un mismo número de vocablos, pero *la distribución social* de estos no es exactamente la misma, como lo muestran las gráficas siguientes.

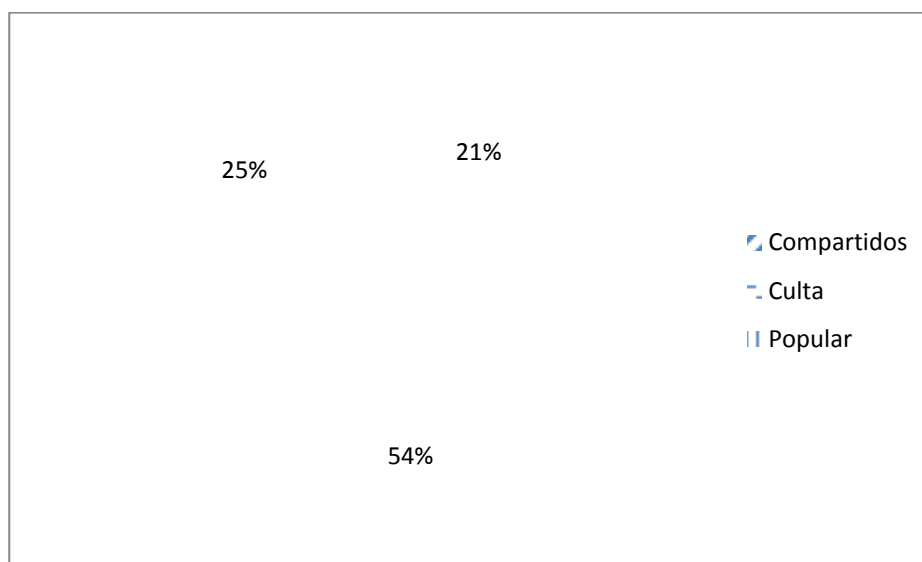


Figura 5.15. *Distribución de vocablos sustantivos en 1970*



Figura 5.16. *Distribución de vocablos sustantivos en el año 2000*

Evaluando estos resultados, puede concluirse que los grupos sociales popular y culto del año 2000 comparten más vocablos y tienden a diferenciarse menos en su repertorio. Todo esto refuerza la idea de que las diferencias sociolingüísticas entre grupos sociales a nivel léxico sí deben ser significativas porque se trata de datos levantados en condiciones muy similares. Los grupos educativos, altamente diferenciados en los años 1970, parecen acercarse estadísticamente en el año 2000. Esta disminución de la brecha entre grupos de alta y baja instrucción en el tiempo real es paralela a la reducción de la brecha sociolingüística que se produce también en el nivel fonético y de cierta manera se complementa con la observación del movimiento paralelo de disminución de los morfemas *-dad* y *-mente* revisados en los subapartados 5.2 y 5.3. Todo esto se discutirá a profundidad en el cap. 6.

5.5.6. Estudio III: Campos referenciales en sustantivos

Tras el análisis en términos de volumen de sustantivos, se decidió analizar los campos referenciales en que se distribuyen los mismos. Para ello se retomaron los campos propuestos en Ávila (1999 b: 162-163). Debido a la enormidad de la tarea de asignar todos los vocablos a un campo referencial, se decidió asumir un criterio numérico para determinar los vocablos 'más productivos' y más distribuidos en cada muestra. Se decidió que para entrar en esta lista, el vocablo debía cumplir con un criterio de descripción elevado: que se presentara al menos 12 veces (esto es, que tuviera la posibilidad de haber aparecido al menos una vez en cada uno de los informantes)²⁴. Dicho criterio resultó útil, ya que los sustantivos que entran en la lista concentran más del 50% de los sustantivos totales en los cuatro grupos analizados. Los campos referenciales implementados en el análisis de Ávila (1999 b) se definen de la siguiente manera (los ejemplos fueron extraídos del corpus de esta tesis):

CAMPOS REFERENCIALES DE SUSTANTIVOS

ELAN: entorno, lugares naturales y artificiales (*vida, lugar, pueblo, país*)

HAPV: ser humano: atributos psicológicos y valores (*miedo, persona, punto, motivo*)

²⁴ Nota importante: los datos de esta sección son los mismos presentados en Serrano (2011 b), por ello para el grupo culto 1970 se exigió una dispersión de 14 porque este era el número de informantes en dicho subgrupo.

HCRP: ser humano: cuerpo (*mano, cerebro, ojo*)

HMNO: ser humano (*gente, niño, muchacho, hombre*)

LMNA: elementos naturales y artificiales (*día, tamal, zapato, aerosol*)

SCED: sociedad: cultura y educación (*cosa, escuela, verdad, dios*)

SESP: sociedad: esparcimiento (*cine, teatro, canción, actor*)

SOSR: sociedad: ocupaciones y servicios (*dinero, arqueólogo, camión, doctor*)

SRIN: sociedad: relaciones interpersonales (*mamá, hijo, hermano, padre*)

TRTM: tiempo: referencias personales (*tiempo, año, vez, momento*)

En la siguiente figura se muestra cómo se fue asignando el campo referencial a cada vocablo.

	A	B	C	D	E	F	G	H	I	J	K	L
1	POPULAR 2000		CULTA 2000				HABLA POPULAR 1970			NORMA CULTA 1970		
2	vida	57 ELAN	vida	92 ELAN	campo	28 ELAN	mundo	65 ELAN				
3	lado	48 ELAN	lugar	50 ELAN	lugar	25 ELAN	vida	61 ELAN				
4	lugar	43 ELAN	lado	42 ELAN	pueblo	24 ELAN	lugar	52 ELAN				
5	parte	25 ELAN	mundo	39 ELAN	lado	22 ELAN	país	43 ELAN				
6	centro	22 ELAN	casa	34 ELAN	colonia	21 ELAN	campo	38 ELAN				
7	cuarto	22 ELAN	punto	22 ELAN	vida	18 ELAN	estado	29 ELAN				
8	pueblo	19 ELAN	ciudad	20 ELAN	miedo	13 HAPV	punto	28 ELAN				
9	tianguis	18 ELAN	espacio	17 ELAN	ojo	23 HCRP	lado	21 ELAN				
10	terreno	14 ELAN	centro	14 ELAN	cabeza	14 HCRP	medio	19 ELAN				
11	vecindad	14 ELAN	colonia	14 ELAN	mano	12 HCRP	naturaleza	19 ELAN				
12	miedo	17 HAPV	país	14 ELAN	señor	100 HMNO	tierra	17 ELAN				
13	mano	14 HCRP	campo	13 ELAN	niño	91 HMNO	iglesia	15 ELAN				
14	gente	89 HMNO	hospital	12 ELAN	muchacho	56 HMNO	hospital	14 ELAN				
15	persona	41 HMNO	zona	12 ELAN	gente	51 HMNO	motivo	19 HAPV				
16	muchacho	31 HMNO	miedo	17 HAPV	persona	51 HMNO	necesidad	18 HAPV				
17	mujer	26 HMNO	cerebro	16 HCRP	señorita	50 HMNO	estilo	14 HAPV				
18	niño	26 HMNO	muerte	16 HCRP	hombre	28 HMNO	niño	132 HMNO				

Figura 5.17. Proceso de asignación de cada vocablo a un campo referencial

Los vocablos que aparecen al menos 12 ocasiones rondan el número 100:

Cuadro 5.11. *Vocablos sustantivos de mayor concentración en las cuatro submuestras*

	1970		2000	
	<i>Culta</i>	<i>Popular</i>	<i>Culta</i>	<i>Popular</i>
N	120	99	101	76

Al marcar cada uno de los vocablos en el cuadro 5.11, se obtuvieron los siguientes totales por cada campo referencial. Se organizan en la siguiente tabla de mayor a menor:

Tabla 5.7. *Campos referenciales en 4 submuestras de ECM*

1970		2000	
<i>Culta</i>	<i>Popular</i>	<i>Culta</i>	<i>Popular</i>
SCED (46)	SCED (22)	SCED (34)	SRIN (15)
TRTM (14)	SOSR (15)	ELAN (14)	SCED (14)
ELAN (13)	SRIN (13)	SRIN (12)	ELAN (10)
SOSR (13)	HMNO (10)	SOSR (10)	TRTM (10)
SRIN (11)	SESP (10)	TRTM (9)	HMNO (9)
HMNO (9)	TRTM (10)	HMNO (8)	SOSR (8)
SESP (7)	LMNA (9)	SESP (7)	LMNA (6)
LMNA (4)	ELAN (6)	HCRP (3)	SESP (2)
HAPV (3)	HCRP (3)	LMNA (3)	HAPV (1)
HCRP (0)	HAPV (1)	HAPVC (1)	HCRP (1)

A diferencia de lo reportado por Ávila (1999 a y b), los campos referenciales que más sustantivos concentran tienen una distribución relativamente similar en los cuatro grupos: para 3 de las 4 submuestras, es el campo SCED (Sociedad: educación y cultura) el que más vocablos aporta; sólo en la submuestra popular 2000 cede ante el campo SRIN (sociedad: relaciones interpersonales: *hijo, cuñado, papá*, etc.). También es llamativo que los campos referenciales con elementos más concretos como LMNA o los referidos al cuerpo humano (HAPV, HCRP) ocupen los últimos 2 lugares en 3 de las 4 listas, con muy pocos vocablos.

A partir de estos resultados puede concluirse que, respecto a los campos referenciales, existe una tendencia generalizada a concentrar vocablos en los campos de lo social (con referentes más abstractos) en las submuestras, si bien, el grupo Culto 1970 es el que presenta más datos en este campo –lo que es congruente con los resultados sobre vocablos con referentes más abstractos en el Estudio I sobre los vocablos más frecuentes. Por otra parte, los campos con referentes más concretos no son privilegiados en ningún grupo social particular. De aquí se sugiere una revisión crítica de las hipótesis sobre diferencias sociolingüísticas basadas en la oposición *concreto/abstracto* y sus correlatos psicosociales y cognitivos (nuevamente, como las de Bernstein 1984).

5.5.7. Estudio IV: el uso de la Relación Type/Token para el análisis del léxico

Para evitar los sesgos que pueden producir la observación de frecuencias absolutas y relativas, se utiliza a lo largo de este apartado la *Relación Type/Token* (RTT) como parámetro de comparación global entre los grupos. La RTT puede verse como una medida muy simple de riqueza léxica y morfológica, ya que resalta el nivel de diversidad de palabras tipo en un corpus. Matemáticamente, es el producto de dividir el número de tipos encontrados entre el número de palabras gráficas totales (Baayen 2008: 244). La siguiente gráfica muestra el valor de RTT de las cuatro grandes submuestras.

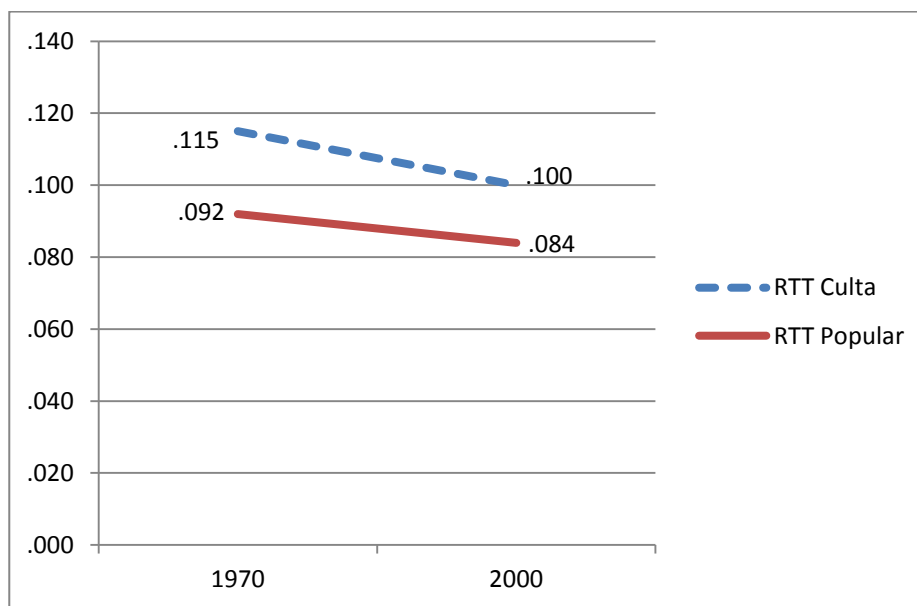


Figura 5.18. Relación type/token en grupos cultos y populares en el tiempo real

Hay una baja generalizada en la RTT en el tiempo real, aunque ciertamente los grupos del año 2000 parecen guardar una menor diferencia entre ellos (de .016 puntos frente a los .023 entre los grupos del año 1970), lo que indicaría, al igual que en el nivel fónico, cierto nivel de acercamiento de los grupos sociales que puede interpretarse también como un proceso global de nivelación de las hablas urbanas en el que los grupos con más estudios presentan una menor diversidad léxica interna, acercándose con ello a los niveles de los grupos de habla popular²⁵. En las siguientes secciones se analizan las diferencias entre grupos de edad, sexo y nivel de escolaridad.

5.5.7.1. RTT en los distintos grupos de edad

Tomando como índice sólo los vocablos sustantivos, la distribución de estos en los distintos grupos de edad (tanto en los muestreos de 1970 como los de 2000) confirma lo hallado en otros trabajos sobre el hecho de que el acervo léxico suele crecer con la edad.

²⁵ Nuevamente, esto se discutirá con detenimiento en el capítulo 6.

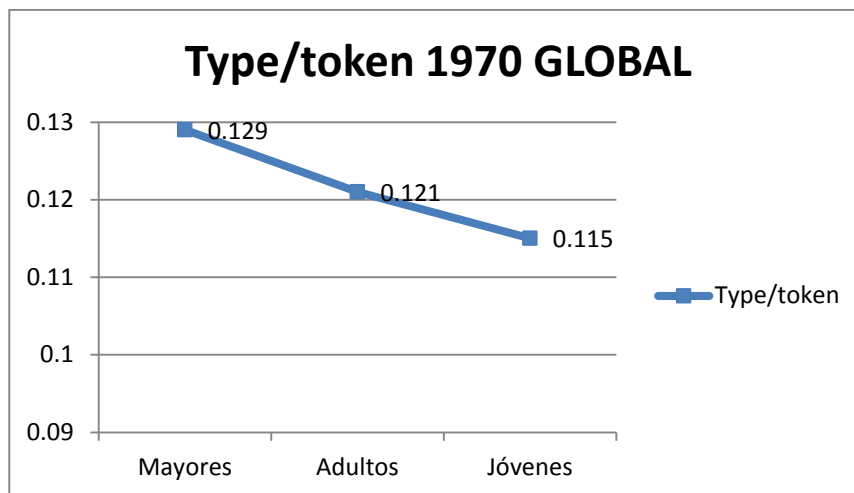


Figura 5.19. Relación *type/token* en 1970 según los grupos de edad

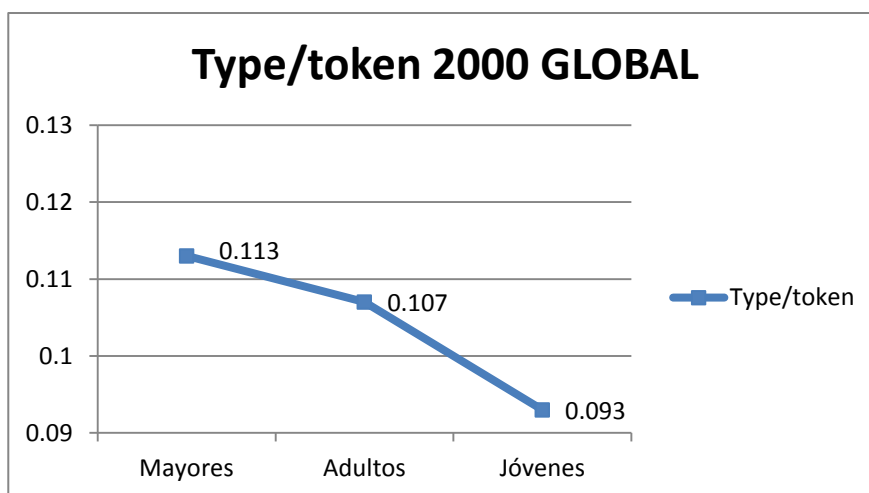


Figura 5.20. Relación *type/token* en el año 2000 según los grupos de edad

El hecho de que el patrón de ascenso del léxico conforme a la edad no sea distinto entre grupos educativos puede ser evidencia a favor del *Principio 5.21* sobre la poca influencia de la edad (y el sexo) en la adquisición, conocimiento y distribución

del léxico (Moreno Fernández 2012: 118). En los siguientes apartados se analiza con más detalle cada muestra e individuo según su valor RTT.

5.5.7.2. Valores de RTT individuales

Como ocurre en muestreos sociolingüísticos limitados en cuanto al número de hablantes entrevistados, la variación intragrupal puede ser alta. Esta muestra no es la excepción respecto al número de palabras. En la siguiente tabla se muestran los datos del muestreo de 1970 por cada informante, ordenados de mayor a menor según su relación type/token individual.

Tabla 5.8. 24 hablantes de 1970 ordenados según RTT²⁶

	Archivo	Types	Tokens	RTT
1	VH-CULT-01	1107	4018	0.27551020
2	VM-CULT-01	947	3633	0.26066612
3	JM-CULT-01	1027	4192	0.24499046
4	AM-CULT-01	1008	4137	0.24365482
5	VH-CULT-02	1311	5488	0.23888484
6	AH-CULT-01	988	4229	0.23362497
7	JH-CULT-01	1017	4376	0.23240402
8	AM-CULT-02	836	3615	0.23125864

²⁶ Para entender la clave de cada informante: el primer carácter indica el grupo de edad: J=joven, A=adulto, V=Mayor; el segundo carácter indica el sexo: H=hombre, M=mujer. El tercer dato indica el grupo educativo: POP=Popular, CULT=culto y los números finales (01 y 02) permiten identificar al informante en esa casilla.

Capítulo 5. Los límites de los corpus orales y la hipótesis de la difusión léxica

9	AM-POP-01	1021	4668	0.21872322
10	AM-POP-02	881	4046	0.21774592
11	VM-CULT-02	1162	5406	0.21494636
12	VH-POP-02	937	4363	0.21476049
13	JM-CULT-02	1073	5079	0.21126206
14	VM-POP-01	966	4579	0.21096309
15	VH-POP-01	854	4126	0.20698013
16	AH-POP-02	976	4768	0.20469799
17	JH-POP-01	1029	5302	0.19407771
18	AH-CULT-02	853	4411	0.19338019
19	AH-POP-01	937	4853	0.19307645
20	JH-CULT-02	1548	8049	0.19232203
21	JM-POP-01	738	3869	0.19074696
22	VM-POP-02	1046	5871	0.17816386
23	JM-POP-02	1061	6137	0.17288577
24	JH-POP-02	838	5204	0.16102998
	<i>Total</i>	9,776	114419	
	<i>Promedio</i>		4767.45833	0.21403151
	<i>Desv std</i>		973.417689	0.02772792

La tabla muestra grandes diferencias entre individuos en 1970: el informante con la RTT más baja es un hombre joven de bajos estudios, con 0.161, mientras que el hablante con la mejor RTT es un hablante mayor de edad, hombre, de altos estudios, con 0.275, esto es, 0.114 puntos de diferencia. Lo cierto es que la tabla confirma la tendencia general a que los hablantes cultos presenten en general una RTT más alta que los hablantes de pocos estudios: los ocho punteros son de aquel grupo educativo; los cuatro informantes con menor RTT, por el contrario, son de habla popular. Otra manera de ver la estratificación es la siguiente: de los 12 primeros, 9 son del grupo culto y sólo tres son informantes de habla popular. La siguiente tabla presenta los valores de RTT en los informantes del año 2000.

Tabla 5.9. 24 hablantes del año 2000 ordenados según RTT

	<i>Archivo</i>	<i>Types</i>	<i>Tokens</i>	<i>TTR</i>
1	AH-CULT-01	1443	5919	0.24379118
2	AM-CULT-02	1016	4195	0.24219309
3	VH-CULT-01	968	4186	0.23124701
4	VH-POP-01	788	3447	0.22860458
5	VH-CULT-02	1025	4499	0.22782841
6	VM-CULT-01	1087	4887	0.22242685
7	VM-CULT-02	906	4143	0.21868211
8	AH-CULT-02	1188	5685	0.20897098
9	VH-POP-02	794	3866	0.20538024
10	AM-POP-01	789	3970	0.19874055
11	JH-CULT-01	984	5018	0.19609406
12	JM-CULT-01	819	4177	0.19607374
13	JM-POP-01	863	4410	0.19569161
14	AH-POP-02	893	4603	0.19400391
15	AM-CULT-01	1275	6595	0.19332828
16	JH-POP-02	891	4635	0.19223301
17	VM-POP-01	1084	5668	0.19124912
18	JH-POP-01	908	4991	0.18192747
19	JH-CULT-02	1293	7288	0.17741493
20	JM-CULT-02	939	5419	0.1732792
21	AM-POP-02	980	5710	0.17162872
22	AH-POP-01	847	4954	0.17097295
23	JM-POP-02	819	4803	0.17051843
24	VM-POP-02	794	4829	0.16442328
	<i>Total</i>		117897	
	<i>Promedio</i>		4912.375	0.19986265
	<i>Desv std</i>		896.580548	0.02361757

Siete de los ocho informantes con RTT más alta en el año 2000 son del grupo culto. El cuarto lugar lo ocupa un hombre mayor de pocos estudios. Examinando los datos de este informante resaltan dos aspectos importantes. Delfino (el nombre del

informante) tiene la entrevista más breve de todas y en general cuando existen menos tokens, la RTT es más alta (cf. Baayen 2008: 244); por otra parte, Delfino puede verse más como un “informante experto” que como un informante promedio. El señor Delfino ha sido colaborador de antropólogos e historiadores de su lugar de origen (Milpa Alta) desde los años 1970, habla español, náhuatl y algo de inglés, es una persona destacada en su comunidad y fue entrevistada porque los habitantes del lugar lo sugirieron por ser un gran narrador. Evidentemente, el señor Delfino tiene recursos léxico-discursivos mayores a los del hablante promedio del lugar y por ello pertenece a una categoría aparte.

Al igual que en la tabla para 1970, la tabla del año 2000 confirma la tendencia a concentrar la RTT más alta en los informantes cultos: nuevamente, 9 de los primeros 12 informantes son de este último grupo.

Para evitar el sesgo estadístico de la variación intragrupal, se decidió obtener la desviación estándar (s^2) del número de tokens y sumarla y restarla al promedio para poder descartar entrevistas que presenten demasiados o muy pocos datos, esto porque, como evidencia la tabla, a pesar de que regularmente se cuenta con la transcripción de 30 mins por informante, varias entrevistas son inusualmente largas (como la núm. 19 de 2000 y la 20 de 1970) o inusualmente breves (como la entrevista 3 de 1970 y la 4 de 2000). En el siguiente cuadro se presentan los límites inferior y superior que permiten decidir qué informantes entran o no al análisis grupal.

Cuadro 5.12. Límites superior e inferior de tokens para muestreo final 1970 y 2000

Año	Tokens promedio (\bar{x})	Desv. Std. (s^2)	Límite inferior ($\bar{x} - s^2$)	Límite superior ($\bar{x} + s^2$)
1970	4767	973	3794	5740
2000	4912	896	4015	5808

Con estos criterios se sacrifican los datos de cinco informantes en 1970: tres que sobrepasan el límite superior (un hombre joven culto [JH-CULT-02] y dos mujeres de habla popular, una joven [JM-POP-02] y otra mayor [VM-POP-02]) y dos que quedan debajo del límite inferior (dos mujeres cultas, una mayor [VM-CULT-01] y una adulta [AM-CULT-02]); los datos de cuatro mujeres se van y se descartan los datos de sólo un hombre, por lo que queda algo descompensada la muestra en cuestión de género. En total quedan 19 informantes en la muestra final: 10 de habla popular y 9 de habla culta.

Tabla 5.10. RTT en 19 informantes en 1970

	Archivo	Types	Tokens	RTT
1	VH-CULT-01	1107	4018	0.27551020
2	JM-CULT-01	1027	4192	0.24499046
3	AM-CULT-01	1008	4137	0.24365482
4	VH-CULT-02	1311	5488	0.23888484
5	AH-CULT-01	988	4229	0.23362497
6	JH-CULT-01	1017	4376	0.23240402
7	AM-POP-01	1021	4668	0.21872322
8	AM-POP-02	881	4046	0.21774592
9	VM-CULT-02	1162	5406	0.21494636

10	VH-POP-02	937	4363	0.21476049
11	JM-CULT-02	1073	5079	0.21126206
12	VM-POP-01	966	4579	0.21096309
13	VH-POP-01	854	4126	0.20698013
14	AH-POP-02	976	4768	0.20469799
15	JH-POP-01	1029	5302	0.19407771
16	AH-CULT-02	853	4411	0.19338019
17	AH-POP-01	937	4853	0.19307645
18	JM-POP-01	738	3869	0.19074696
19	JH-POP-02	838	5204	0.16102998
	<i>Total</i>		87114	
	<i>Promedio</i>		4584.94737	0.21586631
	<i>Desv Std</i>		510.089913	0.02549181

Respecto a la muestra del año 2000, se descartaron los datos de seis informantes: tres que sobrepasan el límite superior de tokens (los tres son cultos: un hombre joven [JH-CULT-02] y un hombre y una mujer adultos [AH-CULT-01, AM-CULT-01]) y tres el inferior (los tres son de habla popular: dos hombres mayores [VH-POP-01, VH-POP-02] y una mujer adulta [AM-POP-01]); se descartan los datos de cuatro hombres y de sólo dos mujeres. La muestra queda constituida por 18 informantes con 9 hablantes en cada grupo social.

Tabla 5.11. *RTT en 18 informantes en el año 2000*

	Archivo	Types	Tokens	RTT
1	AM-CULT-02	1016	4195	0.242193
2	VH-CULT-01	968	4186	0.231247
3	VH-CULT-02	1025	4499	0.227828
4	VM-CULT-01	1087	4887	0.222427

5	VM-CULT-02	906	4143	0.218682
6	AH-CULT-02	1188	5685	0.208971
7	JH-CULT-01	984	5018	0.196094
8	JM-CULT-01	819	4177	0.196074
9	JM-POP-01	863	4410	0.195692
10	AH-POP-02	893	4603	0.194004
11	JH-POP-02	891	4635	0.192233
12	VM-POP-01	1084	5668	0.191249
13	JH-POP-01	908	4991	0.181927
14	JM-CULT-02	939	5419	0.173279
15	AM-POP-02	980	5710	0.171629
16	AH-POP-01	847	4954	0.170973
17	JM-POP-02	819	4803	0.170518
18	VM-POP-02	794	4829	0.164423
	Total		86812	
	Promedio		4822.88889	0.19719133
	Desv Std		527.265631	0.02347288

A pesar de tener descartados a los informantes en los extremos de cada muestra, las diferencias básicas se mantienen: la RTT es mayor en 1970 que en 2000, pero en general permanece la tendencia de los grupos cultos a tener mayor número de formas –algo muy documentado en los trabajos de Ávila (1999 a y b).

5.5.7.2.1. Informantes con mayor y menor densidad léxica en 1970

El informante con mayor densidad léxica en 1970 es un médico otorrinolaringólogo de 62 años. Tiene bastantes recursos léxicos, incluso en una narración de temática alejada de su profesión:

Ya con una idea fija, con una idea fija de... de adentrarme un poco más en un campo desconocido totalmente en México. Lo único que podía hacerse por aquella época, si acaso, pues... era una vieja escuela de sordomudos que fundó don Benito Juárez y que, a la usanza de lo que en aquella época se... se hacía, pues... sigue haciéndose, es decir, con procedimientos muy... muy peculiares, que... todavía perduran y que... y que a mí en lo personal no me satisfacen, ni me satisfacían en aquella época. Entonces... empezamos... pues, una nueva etapa, en el año cincuenta y uno, y transcurrieron cinco años, para mí, de una posición profesional un tanto ambigua, porque, por una parte, la larga prole con que ya me había obsequiado esta señora... pues... me obligaba... a muchas... necesidades que cubrir, a muchas cosas que hacer y, por otro lado, el lanzarme de lleno a una disciplina que aún empezaba a esbozarse en mi mente y que todavía no teníamos ni siquiera oportunidad de echar en marcha de una manera integral... pues... resultaba una aventura. La aventura estaba hecha, desde el momento mismo en que quemamos nuestras naves y dimos fin con todo: hipotecamos lo que teníamos y montamos el Centro Audiológico y Foniátrico de México.

El informante con menor densidad léxica es un joven de 22 años, de oficio zapatero. En este fragmento, el informante hace comentarios sobre su oficio:

Para yo aconsejarle... ee... que el zapato dure... le voy a decir una cosa: El zapato tiene que... este... cuidarse; no pisar cosa lodosa y no andar... este... en...- digamos- en... andar en... caminos... este... muy *escarbosos* -digamos- como tierra, como piedras; allí que se... que sube y baja, y toda la cosa. Como muchas personas: agarran y se van a sus pueblos, y se agarran los zapatos y se... se meten al cerro, se meten a muchas partes, ¿entiende? Entonces es donde el zapato empieza a desgastarse bastante, y el corte se empieza a... quebrar de la piel. Eso es lo que pasa. Pero si, en cambio, a él le agarra... dice:

"Mi zapato me costó tanto; yo agarro y... *pos*, en lugar... Yo veo un lodo allí o yo veo un charco... *ps* agarro y me voy ... doy la vuelta, y ya". Entonces, el zapato lo bolea, lo cuida, ¿entiende? No meterse en... en digamos, en esas... en esos trances así. *Usté* agarra y... se va por otro lado, y lo cuida; lo bolea bastante. Porque ése es... la vida de... de un zapato. En la piel... es la boleada, porque lo lustra, lo cuida, ¿entiende? Al lustrar, al lustrarse -le voy a decir una cosa- al lustrarse, el zapato se entiende que no lo va a ensuciar en cualquier momento. Que el polvo le llegue, es diferente; pero que lo... lo trate de meter en algún lodo, no. ¿Por qué? Porque le costó un trabajo de *boliarlo*, o pagó cierta cantidad porque se lo *bolién*.

En este fragmento el informante repite verbos que podrían admitir sinónimos, como *agarrar* en el fragmento siguiente: "Como muchas personas: *agarran* y se van a sus pueblos, y se *agarran* los zapatos y se... se meten al cerro"; utiliza demasiados marcadores para mantener turno como *este*, *eh* o *digamos*; por tratarse de su oficio, aparece muchas veces la palabra *zapato*; finalmente, el tipo de habla es muy dubitativo –y por ende repetitivo. Todos estos factores indudablemente contribuyen a que baje el nivel de densidad léxica de su discurso.

5.5.7.2.2. Informantes con mayor y menor densidad léxica en 2000

La informante con mayor densidad léxica es una mujer médico de 41 años. En realidad, en esta entrevista particular no existen fragmentos muy largos sin intervención del entrevistador, pero aquí se ofrece una porción del texto:

I: fijate que acabo de tener un problemilla ahí ahorita con la maestra de ballet// me habló a la casa el jueves// por horarios/ a veces la dejamos súper temprano/ porque no la podemos dejar exactamente/ al cinco para la hora que va a empezar su clase/ tampoco la podemos recoger al cinco después de la hora que terminó/ o sea por los horarios de trabajo// y hay veces que sí se queda tarde pero hay ve-/ o sea normalmente los martes la recogemos/ a buena hora/ los jueves son los días que sí/ que llega/ llegamos a recogerla tarde// entonces/ me habla la maestra/ que es la directora/ que en su vida me había hablado/ y yo estaba espantada/ dije/ “algo le pasó a M”

E: pues <~pus> sí

I: y me dijo// este/ “buenas...”/ por su forma de ser que ya es grande/ ¿no?/ “buenas tardes señora/ habla miss M/ este”

E: la mismísima en persona/ ¡huy! [(risa)]

I: [yo dije/ “santa María”]/ dije este/ “sí dígame miss/ ¿en qué le puedo ayudar?”/ “no pues <~pus> este/ M/ nos damos cuenta que está/ desde las cuatro hasta las nueve de la noche// eh <~eh:>// cuando ya empieza la clase todavía no se ha cambiado/ eh <~eh:>/ de repente trabaja muy bien/ de repente trabaja muy mal/ eh <~eh:>/ me dijo que iba a hablar con usted/ le dije a M que iba a hablar/ que quería hablar con usted/ y me dijo M que no/ podía”/ que yo tenía mu-/ “que usted tenía muchos pacientes y que no iba a poder ir/ y que además este/ que iba a ver”// algo así como/ “no vas a lograr que ella me saque de la escuela”/ algo así

La densidad léxica de esta informante en términos de RTT es de .242, algo menor al .275 del informante con mayor densidad en 1970. Sus recursos léxicos son claramente menores que su par de 1970 y además parecen reflejar un estilo más relajado. Este relajamiento en el estilo de habla podría deberse a un cambio en las circunstancias particulares de la entrevista en 30 años, como primera hipótesis, pero como se mencionó al inicio del capítulo, es la situación entrevista la que debe

condicionar el registro y no al contrario; la segunda hipótesis sería que dicho relajamiento estilístico no se debe a la situación de entrevista, sino a que se trata de una propiedad generalizada en los hablantes de altos estudios en general. Esta última hipótesis es la que se defiende aquí. El hecho de que todas las entrevistas del año 2000 presentan un menor valor de densidad léxica que en 1970, a pesar de levantarse en condiciones similares, permite plantear la hipótesis de un cambio generalizado en la comunidad de habla, lo que se discutirá en el Cap. 6.

El hablante con menor densidad léxica en el año 2000 es una mujer, ama de casa, de 57 años.

I: en esa escuela me/ me inscribieron/ pero yo creo que quedaron en que iban a lle-/ a/ a llevar mi acta de nacimiento// como no la <~la:> tenía mi mamá/ y tenía que ir m-/ hasta Celaya/ entonces// pues <~pus> me sacaron// ya no// entonces ya pasó tiempo/ y hasta/ cuando tuve once años/ ya fui a otra escuela/ en la tarde/ pe-/ ya tenía yo el acta de nacimiento/ pero fue acá por este mm/ de donde está la iglesia de San Miguel// por ahí no me acuerdo cómo se llama// es hasta abajo/ José Morán/ todo José Morán [hasta abajo]

E: [mh]

I: está la iglesia de San Miguel/ ahí a una/ dos cuadras hacia abajo/ cerca de <~jeláti>/ está este/ estaba una escuela chiquita

E: y ahí

I: y/ y ahí me apuntaron// y ahí/ un año nada más fui// porque/ yo ya me sentía mal/ en primer año/ pues <~pus> ya me sentía yo muy grande/ aparte de eso/ mi mamá estaba por aliviarse (silencio)

E: [¿y usted qué?]

I: [y al tener un bebé]

E: ¿cuidando chamacos?

I: pues <~pus> yo tenía que cuidar a mis hermanos/// sí// y por eso ya no estudié/ me apuntó en la nocturna/ después de que se alivió/ no terminé el primer año/ me apuntó en la nocturna/ pero pues/ se pelearon <~pelearon>/ se fue la luz/ y no/ me dio miedo/ ya no fui/ ya no quise ir/ se pelearon <~pelearon> dos señores/ o sea muchachos ya grandes/ pues <~pus> yo era la única más chiquita de todo el grupo/ pura gente grande

Al igual que el informante de menor densidad léxica en 1970, la informante repite algunos marcadores discursivos como *pues*, o utiliza formas pronominales que podrían considerarse redundantes, como en el siguiente fragmento: “porque/ yo ya me sentía mal/ en primer año/ *pues* <~pus> *ya me sentía yo muy grande...*”; rasgos de este tipo disminuyen su riqueza léxica en términos de RTT –aunque no impactan en la expresividad o la eficacia de su discurso.

Como puede observarse en los fragmentos de los cuatro informantes con mayor y menor densidad léxica, todos tratan tópicos relacionados con sus historias de vida, ocupación u oficio; sin embargo, los repertorios son muy distintos, sobre todo cuando entra en juego la variable escolaridad.

5.5.8. Conclusiones sobre el nivel lexicoestadístico

El acercamiento general en términos de riqueza léxica medida a través de vocablos sustantivos, los campos referenciales concentrados mayoritariamente en referentes

de lo social/abstracto y una tendencia comunitaria a la baja en términos de RTT pero con grupos sociales un poco menos diferenciados en el año 2000, vuelven evidente cierta tendencia a la informalidad lingüística en el habla capitalina reciente. Por otra parte, los grupos de menores estudios han enriquecido su acervo sustantivo y hacen un mayor uso de algunos rasgos que los asemejan a los hablantes de mayor escolaridad, como ocurría en el reflejo léxico de los grupos consonánticos /ks/ y /kt/. Como se pudo observar con estos resultados, existe una coincidencia con los capítulos sobre variación fónica (3 y 4) en dos patrones sociolingüísticos: uno que indica cierto alejamiento de los hablantes cultos del año 2000 de las hablas ‘cultivadas’ de los años 1970 y por otro, un ligero acercamiento a lo que podría llamarse una “nivelación léxica” por parte de los hablantes de nivel bajo de estudios. Las evidencias serían el descenso comunitario en la densidad léxica (en términos de RTT) en el año 2000, descenso más marcado en los grupos altos y, por otra parte, un aumento en el repertorio léxico en los hablantes de baja escolaridad del año 2000.

Si es correcta la hipótesis de que se están produciendo procesos de convergencia o nivelación dialectal en ECM, ésta bien podría ponerse a prueba a través de una revisión detenida de algunos ítems relativamente emblemáticos de las hablas “cultas” (o asociadas a grupos prestigiosos) y “populares”. Si los procesos son de nivelación dialectal, sería esperable que los grupos cultos, en procesos de préstamo “desde arriba”, difundieran algunas de sus palabras y fraseologismos más característicos (o estereotípicos); por supuesto, también deberían poderse observar procesos de difusión léxica “desde abajo”, en los que los grupos socialmente

privilegiados “tomen prestadas” ciertas frases e ítems léxicos normalmente considerados “populares”. Esta exploración servirá también para documentar las rutas de difusión que siguen las palabras sociolingüísticamente marcadas.

5.6. “CADA PALABRA TIENE SU PROPIA HISTORIA”. PROCESOS DE DIFUSIÓN, INNOVACIÓN Y RETRACCIÓN LÉXICA²⁷

Esta sección trata sobre la manera en que surgen, se difunden o desaparecen algunos vocablos seleccionados como representativos del habla cotidiana de la ciudad de México de alrededor de los años setenta y cuál fue su evolución en un período aproximado de 30 años. Como se ha hecho a lo largo de la investigación, el corpus base es de lengua oral en tiempo real. En buena medida, esta exploración pondrá en evidencia que los corpus orales como el aquí utilizado pueden ser limitados para rastrear ítems léxicos específicos.

En esta sección se entenderá por “cambio en el nivel léxico” tres tipos de procesos. Por un lado, el de *cambio léxico* propiamente, entendido como la inserción (o desaparición) de un vocablo en el lexicón (Crowley 1997); en particular, se verá la inserción por acuñación (*coinage*) o préstamo de otras lenguas debidas a un cambio

²⁷ Una versión previa de esta sección se publicó en forma de artículo con el título “Retracción e innovación léxica en español de la ciudad de México: 1970-2000” (Serrano 2011 a). El presente texto incorpora más ejemplos, extiende los comentarios ahí plasmados y precisa la redacción en varios puntos.

social o un cambio en la cultura material de los hablantes (Hock & Joseph 1996: § 8-9). El segundo proceso es el de *cambio semántico*, entendido como un proceso de modificación del significado de un vocablo a través de mecanismos como *extensión metafórica*, *extensión metonímica*, *extensión de significado*, *reducción*, etc. (Fortson IV 2003: 648-650)²⁸. Finalmente, el tercer tipo de proceso investigado será el de *aumento en la dispersión* de algunos vocablos en la comunidad, aumento que puede estar conducido por rutas sociales o cognitivas “desde arriba” o “desde abajo” (Labov 1972)²⁹.

Antes de abordar los resultados de esta investigación, se presenta un breve recorrido sobre los estudios de vitalidad y mortandad léxica en español, lo que permitirá consolidar los objetivos de la sección.

5.6.1. *Índices de mortandad y vitalidad léxica y el estudio del léxico de la ciudad de México*

J. López Chávez aseguró recientemente (2011) que el *léxico básico* del español de México en el período 1964-2004 no ha cambiado sustancialmente: “Se han dado cambios en el inventario de lexemas, algunos han salido, otros han entrado, otros más

²⁸ Existen varios tipos de cambio léxico ligados a procesos de *cambio analógico* (fónico, morfológico, rimas --cf. Hock & Joseph 1996: 292-293) que no se tratarán en esta sección.

²⁹ Chambers & Trudgill (1998: 157-159) presentan el ejemplo del desplazamiento de la palabra *chesterfield* (‘sofá’) por las opciones *couch* y *sofa* en Canadá; dicho desplazamiento léxico, que sirve a los autores para ejemplificar el cambio estratificado por edad (*age-grading*), es llamado por los autores como *innovación léxica*.

han modificado algunos de sus aspectos, pero al final de la jornada todo sigue ahí” (2011: 415). Por “léxico básico” López Chávez (2011: 399-400) entiende el conjunto limitado de vocablos basado en frecuencias que refleja el 98% del vocabulario usual³⁰. Por supuesto, en períodos tan breves como el aquí revisado es imposible que se produzcan cambios sustanciales en el acervo de una lengua, sin embargo, no deja de ser interesante investigar, por ejemplo, lo que puede ocurrir con ciertos vocablos que funcionan como *marcadores* sociolingüísticos, que son aquellos que pueden devenir en *estereotipos* sociolingüísticos (cf. Chambers & Trudgill 1998: 70-76).

Existen ciertos trabajos que hablan sobre índices de vitalidad o mortandad léxica en español mexicano, principalmente centrados en la permanencia de indigenismos, como el de Lope Blanch (1969), *Léxico indígena del español de México*. En dicha investigación, basada en un enorme corpus de lengua hablada (343 entrevistas con 490 informantes de la ciudad de México, entrevistados entre 1964 y 1965 --Lope Blanch 1969: 22-23) y complementada con un volumen similar de material escrito (2,393,750 palabras totales), se encuentra que la proporción de los indigenismos no es estadísticamente representativa de las hablas mexicanas (10,478 en lengua oral, que sólo llegan al 0.478% del total de dicha sección del corpus, calculado en 2,211,000 palabras totales y de los cuales el 84% corresponde a topónimos y gentilicios); sin embargo, además del indudable valor descriptivo de dicha investigación, ésta tiene la ventaja adicional de detallar ciertos procedimientos

³⁰ López Chávez toma como punto de partida para su comparación diacrónica el *Frequency Dictionary of Spanish Words* de Juilland & Chang-Rodríguez (1964) y compara los primeros 5,000 vocablos.

para calcular la vitalidad de dichos indigenismos. Como se documenta en el mismo libro, se realizó un estudio vinculado al de corpus, que consistió en levantar 100 encuestas para verificar el “léxico pasivo”. Se creó un “índice de vitalidad del léxico indígena *pasivo*” con cinco apartados:

I, voces de conocimiento absolutamente general; II, voces también generalmente conocidas, pero sin la firmeza y seguridad de las anteriores [es decir, casos en los que algunos informantes titubeaban, o en que era preciso que hicieran un pequeño esfuerzo para reconocer la palabra o su significado...]; III, voces de reconocimiento *medio* [...]; IV, voces poco conocidas o de significado impreciso para la mayoría de los informantes; V, voces prácticamente desconocidas; VI; voces enteramente desconocidas [...] [Lope Blanch 1969: 33-34].

Los resultados indicaron que las voces de conocimiento general llegan a 95 (p. ej., *apapachar, jícama, pepenar*), de conocimiento casi general son 60 (*achichinle, chilango, tlaconete*), las de conocimiento medio son 62 (*cacahuacincle, memela, tlacoyo*), las poco conocidas son 27 (*acocil, chiltepín, temascal*), las muy poco conocidas 38 (*achiote, cuescomate, uchepo*) y 30 voces prácticamente desconocidas (*tequescamote, paxclal, yagual*) (cf. Lope Blanch 1969: 35-37)³¹.

³¹ Trabajos recientes sobre vitalidad de indigenismos en español mexicano son los de Rosado (2013) con hablantes yucatecos y el de Lozanova (2000) sobre indigenismos en la ciudad de México.

Por supuesto, existen trabajos sobre vitalidad y mortandad en otras variedades hispánicas. En su trabajo “A propósito del *Diccionario usual del español de Puerto Rico. Índices de mortandad léxica*”, H. López Morales (2006) presenta los resultados de un estudio de mortandad de vocablos patrimoniales “anotados como arcaísmos españoles vivos en Puerto Rico”, de indigenismos y afronegrismos. López Morales describe el proceso de mortandad léxica de la siguiente manera:

El proceso por el que pasan estas palabras antes de llegar a la tumba nos es bien conocido: su frecuencia va decreciendo cada vez más, perdiendo presencia en ciertos estilos lingüísticos, pasa de nómina activa a la pasiva –las que se comprenden pero no se usan—y, por último, desconocimiento total: ni se usan ni se reconocen como unidades del inventario léxico propio. El proceso ha llegado a su fin [López Morales 2006: 277].

Por supuesto, López Morales aclara la diferencia entre las palabras o vocablos *obsolescentes* (que sólo son producidos por la generación de mayor edad) de las palabras poco usadas o *poco frecuentes*, pero que sí forman parte del *léxico disponible*, aquel que “formando parte de la competencia léxica de los hablantes, de todos los hablantes o de parcelas de ellos, no se actualiza en el discurso hablado o escrito a menos que el tema lo permita” (2006: 278, n. 2). Para determinar el índice de mortandad léxica en el español usual de Puerto Rico, López Morales recupera los resultados de tres cuestionarios: uno de 160 términos patrimoniales que se pone a prueba con 210 informantes en 25 municipios; uno más de Vaquero (1983, 1986) de

183 indigenismos auscultados a 100 informantes y finalmente el estudio de López Morales (1988) sobre 131 afronegrismos puestos a prueba con 256 encuestas levantadas a lo largo de toda la isla. A partir de ciertos criterios estadísticos (v.g., una palabra con índice menor a 10% se considera en mortandad, y aquellas entre 10 y 30% se consideran “de poco uso” –López Morales 2006: 280) se determinan los grados de mortandad de los vocablos propios de esta variedad hispánica.

En el presente estudio, a diferencia de las investigaciones mencionadas, no se pone a prueba en un cuestionario la vitalidad de los vocablos que aquí interesan, sino que se explora solamente el uso particular de los mismos en el habla de entrevistas –lo que ciertamente representa una desventaja, porque no es posible acceder al conocimiento del léxico pasivo de los informantes. La hipótesis es que ciertos vocablos, sobre todo aquellos ligados al habla juvenil de los años 1960 y 1970 (y reflejados en la literatura de la época, como *onda*, *chido* o *súper*) tienen una fuerte presencia en la actualidad y que, además, presentan un uso más generalizado (más difundido) en términos de la edad y nivel educativo de los hablantes. Además, se exploran vocablos ligados a la tecnología recientemente insertados en ECM --como (*teléfono*) *celular*-- y cambio de significado en vocablos como *género* (en lugar de *sexo*).

Una pregunta importante de este subapartado también es en qué medida los dos corpus de español hablado en la ciudad de México que se han utilizado para el nivel fónico, morfológico y lexicoestadístico podrían ser útiles para rastrear la aparición, desaparición o difusión de un grupo de ítems léxicos particulares. Como se

verá a lo largo de la sección, las primeras exploraciones con el corpus original de 48 informantes muestran que es insuficiente conformarse con estos datos de lengua oral, y que es imperativa la necesidad de rastrear la evolución de las palabras elegidas en otros corpus, como el *Corpus Diacrónico del Español* (CORDE) y el *Corpus de Referencia del Español Actual* (CREA), ambos de la Real Academia Española (www.rae.es), como fuentes principales de lengua española escrita³².

5.6.2. Búsqueda de palabras en los corpus originales

Utilizando el programa *Simple Concordance Program* 4.0.9 (SCP)³³ se exploraron varios grupos de palabras en distintos campos referenciales en los dos corpus. La decisión sobre qué palabras estudiar se basó en las intuiciones sociolingüísticas del autor sobre lo que es común escuchar cotidianamente en español de la ciudad de México, y no en el trabajo *a priori* sobre elementos léxicos más frecuentes. Los vocablos pertenecen a distintos campos referenciales: formas de tratamiento y estereotipos sociales (*v.g. güey, naco*), tecnología (*casette*), expresiones superlativas “juveniles” (*chido, súper*) y términos provenientes del discurso cultivado (*género vs.*

³² No se utilizan los datos del *Corpus del español del siglo XXI* (CORPES XXI) (<http://web.frl.es/CORPES/view/inicioExterno.view>) por contener datos, en su mayoría, fuera del período de estudio (1970-2000).

³³ Diseñado por Alan Reed, el programa se descarga gratuitamente en: <http://www.textworld.com/scp/> para plataformas Macintosh y Windows.

sexo). En la siguiente tabla se presentan los resultados de la exploración en los corpus orales.

Tabla 5.12. *Ítems léxicos en dos corpus de español de la ciudad de México*³⁴

	1970 popular	1970 culto	2000 bajo	2000 alto
Total de palabras	65,290	57,772	129,080	136,871
Núm. de types	5,313	6,742	7,605	9,818
Relación type/token	0.08137540	0.11670013	0.05891695	0.07173178
Habla juvenil				
Súper	0	0	3	31
Onda	1	0	4	32
Padre	0	0	0	32
Chido	0	0	3	17
Formas de tratamiento, estereotipos sociales				
Güey	0	0	16	19
Naco (a)	0	0	0	7
Fresa	0	0	0	2

³⁴ En esta sección se utilizaron *todos* los datos de las 24 entrevistas del CSCM, lo que duplica los datos totales del año 2000 respecto a 1970; la razón es que se requería compilar la mayor cantidad de texto posible. Por *token* se entiende cualquier grupo de caracteres entre dos espacios en blanco y por *type* una o más formas gráficas que pueden pertenecer a un solo “vocablo” o palabra de contenido léxico o gramatical individual (v.g., *fresa*, *fresita*, *fresón* son *types* del vocablo adjetivo ‘fresa’).

Capítulo 5. Los límites de los corpus orales y la hipótesis de la difusión léxica

Ñoño	0	0	0	1
Tecnología				
Disco	8	1	2	7
Computadora	0	0	6	6
Celular	*	*	1	3
Grabadora	3	1	3	1
Diskette	*	*	0	0
Cassette	0	0	2	0
Cultismos				
Género (fem/masc)	0	0	0	0
Sexo (fem/masc)	0	0	0	0

Lo que la tabla indica de entrada es que los corpus documentan de manera pobre muchos de los vocablos estudiados (*ñoño, fresa, naco*), pero sí permite rastrear la evolución y vitalidad de algunos vocablos que resultaron relativamente productivos (*súper, onda, padre, chido*). Otro dato llamativo es la relación *type/token* (tipo/ocurrencia, RTT) entre los diferentes grupos de datos. Por ejemplo, nótese que la RTT es llamativamente mayor en la columna “1970 culto” que en “2000 alto”: 0.1167 en el primero, frente a un 0.0717 en el segundo, lo que sugeriría una constricción del acervo léxico en el lapso de 30 años. Por supuesto, deben tomarse con tiento estos resultados ya que, como se mencionó *supra*, se cuenta con el doble de datos en el año 2000 (136,871 palabras de los hablantes cultos del año 2000 frente a 57,772 de los hablantes cultos de 1970).

Las palabras asociadas a la tecnología son poco productivas y hubo algunas que, por supuesto, no se buscaron en los datos de 1970 (marcadas con asterisco) por tratarse de objetos de muy reciente aparición. Otra aclaración importante que debe hacerse es el término *disco*, cuyo referente ha cambiado de 1970 a 2000: antes refería a los discos de acetato (de 45, 75 y 33 1/3 rpm) y ahora se utiliza la palabra para designar generalmente a los “discos compactos” (*CD*); aunque conviven los términos *CD*, *compacto* y *disco*, al parecer este último presenta gran vitalidad. Es el mismo caso de *carro*, que refería en el siglo XVIII a un vehículo de transporte tirado por animales y que ahora refiere principalmente a un vehículo automotor: la palabra sobrevive, aunque el referente se haya transformado sustancialmente.

A partir de estos primeros resultados se decidió trabajar solamente con las palabras de la tabla que fueran más productivas, con excepción de la palabra *género* y su significado relativamente innovador para designar al sexo femenino/masculino, significado de reciente introducción (al menos en el discurso político mexicano) y del que puede rastrearse su aparición con relativa facilidad. La lista final por tanto es la siguiente: *onda*, *género* (y su sinónimo *sexo*), *celular*, *súper*, *chido*, *padre* y *güey*³⁵.

Además de las palabras en la tabla 5.12 se realizaron búsquedas de ciertos ítems léxicos que consideró se encontraban en retracción, como *valedor*, *carnal* o *mano* (apócope de *hermano*) y, por otra parte, también palabras consideradas groseras (*pendejo*) para verificar si se distribuían de manera interesante en los muestreos, pero ninguna muestra arrojó ejemplos de dichos vocablos, por lo que

³⁵ Se excluyeron los vocablos *naco*, *fresa*, *ñoño*, *computadora*, *grabadora*, *diskette* y *cassette*.

resultó imposible sugerir cualquier tendencia. Estos resultados evidencian nuevamente que los corpus utilizados presentan limitaciones cuantitativas y que, al tratarse de materiales de entrevistas, no son representativos de la diversidad de registros de habla en español capitalino –sobre todo de los registros más informales que favorecen la interdicción³⁶ o el uso de términos de relación afectiva como los explorados.

5.6.3. En busca de otras fuentes

La primera exploración con los corpus orales evidenció que se requería de un mayor volumen de datos y diversidad de registros sociolingüísticos que permitieran trazar las tendencias de innovación o retracción léxica. De esta manera, se decidió explorar en el CREA, en una primera etapa en la sección de *lengua oral, México*; si no arrojaba datos, entonces se buscaban concordancias en *México, todos los registros*. Por lo tanto, se sacrificó la naturalidad de los datos de lengua oral por una mayor representatividad de estilos, aunque estuvieran de alguna manera “disimulados” en la ficción escrita.

Se realizaron búsquedas en el CORDE para determinar las primeras documentaciones de los vocablos estudiados, pero no hubo resultados para algunos de estos (al menos no con el sentido particular que se buscaba). En los siguientes

³⁶ Por *interdicción lingüística* se entiende “la prohibición social por la cual ciertas palabras no pueden ser usadas en el discurso. Los motivos de esta prohibición pueden ser religiosos, mágicos, políticos, morales, etc.” (Palacios 2002: 41).

apartados se presenta con detalle los resultados sobre cada una de las palabras investigadas³⁷.

5.6.4. "¡Qué buena onda!"

El *Diccionario del español de México* (en adelante, DEM) (<http://cemc.colmex.mx/>)

recoge los diversos sentidos de la palabra *onda* que inspiraron esta búsqueda:

II (*Coloq*)

- 1 Comportamiento o forma de actuar, actitud o actividad que se adopta en un momento dado: "Andaba en otra *onda*", "Está en la *onda* de la música", *una onda gruesa, una onda padre, ondas diferentísimas*, "Esa maestra es muy buena *onda*", "¡Qué mala *onda*, no devolverle el libro que le prestó", *respetar la onda de cada quien*
- 2 *Agarrar, coger o captar la onda* Adquirir una habilidad o entender algo: "Ya le *agarré la onda* a las computadoras"
- 3 *Írsele la onda a uno* Perder el hilo del discurso, la corriente del pensamiento o el curso de las ideas: "*Se me fue la onda* a la hora del examen y no supe qué contestar"
- 4 *Sacar de onda* Confundir, ser difícil de interpretar: "Me *sacan de onda* sus explosiones de violencia"
- 5 *Entrar en onda* Entrar en un determinado ambiente y adaptarse o asimilarse a él: "Son bien fresas, no *entran en onda*"

³⁷ Por supuesto, el Corpus del Español Mexicano Contemporáneo (CEMC) de El Colegio de México, base del *Diccionario del Español de México* (Lara [dir.] 2011) es, con sus 2 millones de palabras, otra fuente importantísima de datos.

6 ¡Qué onda! interj Saludo informal

III *La onda* Corriente que a mediados de los años 60 tomó como tema a los jóvenes, con un énfasis especial en el rock y en la experiencia de las drogas, como la mariguana, el peyote y el LSD, y dio tratamiento literario a su lenguaje coloquial, lenguaje juvenil de aquella época que hoy se ha extendido ampliamente. Entre sus principales exponentes están José Agustín y Parménides García Saldaña.

En general estos usos de *onda* (o al menos muy similares) están documentados para muchas variedades del español (Colombia, España, Perú, Argentina, etc.) y en la publicidad reciente puede encontrarse la expresión *¡buena onda!* como estereotipo del habla de los años 1960³⁸. Se buscó en el CORDE para rastrear el surgimiento de la palabra. La primera mención se da en *Susana y los jóvenes* (1954) de Jorge Ibargüengoitia, quien trataba de reflejar el habla juvenil en sus obras de teatro. Véase el ejemplo:

(1) **Tacubaya:** No se puede platicar contigo.

Susana: Al contrario, es lo único que se puede hacer.

Tacubaya: Nunca encontramos **la onda**.

Susana: Bueno, Tacubaya, amigo, me suicidaré.

³⁸ En las salas de cine mexicanas se veía hace unos años (2008-2009) un comercial del refresco *Pepsi retro* donde un personaje *hippie* exclama al probarlo: “Pepsi Retro: “¡buena oonda!”.

En el corpus del CREA aparece desde 1975, en la obra de Rosario Castellanos, *El eterno femenino*:

(2) **Prostituta:** Chócala. Oye ¿y cómo viniste a parar en esto?

Lupita:(Despreocupada.) **Pura onda.** Desde chiquita me gustaba darle vuelo a la hilacha, y una vez que ya no tuve respeto de padre agarré y dije: ya vas.

Tabla 5.13. Resultados para onda por año en el CREA (México)

1975-79	1980-84	1985-89	1990-94	1995-99	2000-04
65	29	19	11	17	6

Si se realiza una revisión detenida de los resultados para cada período en el CREA salta a la luz que en realidad los tipos de texto y estilos literarios específicos son los que fuerzan la aparición u omisión de la palabra. Por ejemplo, en 1976 hay 62 casos (de 65) que provienen de *Las rojas son las carreteras*, de David Martín del Campo, una obra de temática juvenil. De igual forma, 22 de los 25 ejemplos de 1980 provienen de una misma novela: *Fábrica de conciencias descompuestas*, de Gerardo María. Nótese también que entre 1975 y 1989 se concentran 103 ejemplos y a partir de 1990 baja considerablemente su documentación -a sólo 34 ejemplos. Parece claro que los datos del CREA sugieren una disminución en el uso del vocablo que no

coincide con los datos de lengua oral de nuestro corpus original, donde se ve que, en realidad, la forma era inexistente en el habla culta de los años 1970 y aparece sólo una vez en el habla popular del mismo período, pero sí se documenta para el año 2000 en 36 ocasiones, especialmente en el grupo “alto” (32 ejemplos). Si nos guiáramos por los datos de lengua escrita podríamos concluir que el uso de *onda* está en franco retroceso en español mexicano; pero si se siguieran los datos de lengua oral se tendría que reportar lo contrario, que la expresión ha cobrado fuerza en el español capitalino. Esta contradicción según el tipo de lengua que observemos obliga a cuestionar qué dato puede ser más revelador, el oral o el escrito; también debería cuestionarse si la no aparición de la expresión en los últimos tiempos se debe a la aleatoriedad de la elección de textos en el CREA, y si el uso de la expresión se debió a una simple moda literaria que no reflejaba necesariamente el habla de las calles de la ciudad de México. Por lo pronto, a partir de los resultados en el corpus oral debe concluirse que los usos de *onda*, en los años setenta vinculados al habla juvenil e informal, se han generalizado a toda la comunidad de ECM.

5.6.5. La palabra género

El Diccionario de la Real Academia (DRAE) (<http://www.rae.es/recursos/diccionarios/drae>) no cuenta todavía con una definición de *género* que incorpore la noción de estudios de equidad entre hombres y mujeres, algo que sí registra el DEM:

III

Conjunto de rasgos culturales, antropológicos, sociales, económicos, etc. asociados con el sexo o la sexualidad de una persona o un grupo de personas, especialmente las mujeres: *estudios de género*³⁹

Este sentido *académico* del vocablo *género* –utilizado en lugar de *sexo*– es de reciente introducción en el repertorio léxico hispánico general, y está fuertemente asociado a la lengua escrita –recuérdese que no aparece ni un solo ejemplo en el corpus original de lengua oral. Véase la siguiente tabla de distribución de ese sentido de la palabra *género* en el tiempo.

Tabla 5.14. Resultados para género en el CREA (México)

1975-79	1980-84	1985-89	1990-94	1995-99	2000-04
0	0	9	9	17	71

La primera aparición en el CREA (México) es en 1987; el mismo texto da un total de nueve ejemplos:

³⁹ En una versión previa de este trabajo (Serrano 2011 a), basada en el *Diccionario del español usual de México* (DEUM, Lara 1996), se reconocía un sentido de *género* algo distinta: “5 *Estudios de género*: Aquéllos dedicados a investigar el lugar que corresponde a las mujeres en la sociedad”.

- (3) la ciencia predominante en una sociedad burguesa está sellada por los intereses de clase; de la misma manera que está en función de los intereses de **género**, puesto que de manera obviamente mayoritaria son los hombres quienes detentan el “saber”.

Este ejemplo aparece en la obra de Eli Bartra (1987), *Frida Kahlo: Mujer, Ideología, Arte*. (Icaria Barcelona, 1994). En la siguiente entrevista periodística se observa una alternancia entre “sexo” y “género” muy interesante:

- (4) [Entrevistador] Mucha gente dice que usted tiene más valor que los hombres, ¿es cuestión de **sexo** el saber torear?

No. Cada persona es capaz de hacerlo de una manera distinta, sin importar si se es mujer o hombre, creo que el arte del toreo no depende del **sexo**, ni el valor va con el **género**. Hay hombres muy miedosos y son hombres, así como mujeres muy valientes. Que yo me ponga delante de un toro no significa que tenga un valor mayor que el hombre.

Prensa, “El mundo del toreo, machista, pero no especulo con mi condición femenina...”, *Proceso* 22/12/1996.

La revisión detallada de los ejemplos en CREA deja claro que este sentido de la palabra se introdujo casi exclusivamente en textos que tratan sobre equidad de género. Al aumentar el interés académico en estos temas y verse obligados por las tendencias mundiales al respecto, el gobierno de México (sobre todo a partir del año 2000) pone en marcha programas institucionales sobre equidad de género, lo que

vino acompañado de la engorrosa tarea de dirigirse siempre “a los dos géneros”: “maestros y maestras, chiquillos y chiquillas, ciudadanos y ciudadanas”, etc. en los textos políticos (y oficiales) mexicanos⁴⁰. Se encuentra un total de 106 ejemplos de *género* en este corpus, de los cuales 71 (prácticamente el 70%) se producen después de 1999. Por tanto, la palabra parece ir cobrando fuerza con el paso del tiempo. Los ejemplos “sin fecha” son 34, y muchos de ellos provienen de los discursos en la Cámara de Diputados o del Senado de la República en los años 2000-2004.

Esta palabra debe contrastarse con el uso de *sexo femenino/masculino*. En el CREA se buscaron las concordancias de la frase *sexo femenino/masculino*, la cual arrojó 37 casos en total.

Tabla 5.15. *La secuencia sexo femenino/masculino en CREA (México)*

1975-79	1980-84	1985-89	1990-94	1995-99	2000-04
11	2	9	1	6	6

Como puede observarse, los años setenta y ochenta concentran 22 de los 37 casos totales, lo que sugiere que la forma *género* empezó a sustituir a *sexo* en los años

⁴⁰ Reconozco que tuve mis dudas respecto a los textos médicos, donde se habla de “género femenino” ya que no pude determinar si el término fue tomado directamente de las ciencias médicas y biológicas, o si se trata de una adaptación al lenguaje “políticamente correcto” en boga.

1990. Podría suponerse que se sustituyó la palabra 'sexo' en el discurso oficial para evitar el *tabú* de dicha palabra (Hock & Joseph 1996: 231-234). La conclusión provisional es que la palabra *género* está desplazando a la palabra *sexo*, al menos en los formatos discursivos más formales.

5.6.6. *Celular (teléfono, telefonía)*

La introducción de la palabra *celular* puede clasificarse como cambio léxico onomasiológico desde dos direcciones (Zgusta 1990): por un lado, la aparición de un nuevo referente en la vida cotidiana de los hablantes y, por otro, la expansión semántica del término *celular* a la esfera referencial de la tecnología, algo registrado en el DEM: "que tiene estructura u organización de célula o como una célula: *tejido celular, teléfono celular*". *Teléfono móvil* compite con *teléfono celular* en distintas partes del mundo hispanohablante. Al parecer, en México se prefirió *celular* por la influencia tecnológica de EE. UU., donde se les conoce como *cell phones*. Su primera aparición en CREA (México) es en 1991:

- (5) Se estima que las comunicaciones móviles de larga distancia en tierra firme se efectuarán principalmente por medio de **redes de radio móvil celular** [...]

Rodolfo Neri Vela, *Satélites de comunicaciones*.

El segundo ejemplo es: “sistema europeo digital celular” y el tercero “sistema celular terrestre” en la misma obra. En todos los casos la forma cumple una función adjetiva.

En 1995, en *La casta divina*, Felipe Victoria Zepeda utiliza “aparato celular”. En la misma obra se documenta el primer uso de *celular* como sustantivo:

(6) pero no sabía de plano cómo se manejaba un **celular**

Tabla 5.16. *La palabra celular en el CREA (México)*

1975-79	1980-84	1985-89	1990-94	1995-99	2000-04
0	0	0	4	52	48

De los 29 casos de 1996, 28 son formas adjetivas de *celular*, principalmente “telefonía celular” y “teléfono celular”. En 1997 son 8 de 10 los usos adjetivos, en 2001 los 18 ejemplos totales son usos adjetivos. En 2002, 7 de los 15 ejemplos ya son usos de la palabra como sustantivo. Actualmente, en la publicidad y el habla cotidiana, se utiliza *celular* casi exclusivamente como sustantivo y los datos del CREA parecen confirmar este desplazamiento de categoría léxica de la palabra. Resulta interesante cómo podemos ver la evolución del término, de cumplir función adjetiva a nominal en alrededor de 20 años. Sería muy interesante explorar si este es el patrón normal que

siguen las innovaciones léxicas vinculadas a la ciencia y la tecnología, como (en tiempos más recientes) *chat*, *escanear* o *memoria USB*.

5.6.7. La palabra *súper*

Se buscó la palabra *súper* con valor adjetivo o adverbial. En CREA aparece por primera vez con dicha función en 1995 en una novela de Felipe Victoria Zepeda, *La casta divina. Historia de una narcodedocracia*:

(7) yo mandé a dos hombres de **súper** confianza a traernos el botín

En total sólo aparece en siete ocasiones en todo el CREA. No se tomaron en la cuenta los ejemplos que son calcos del inglés, como *peso súper pesado* (hablando de boxeo), *Súper Tazón*, y en nombres compuestos como *pistola .38 Súper* -donde forma parte indisociable del nombre del arma.

Tabla 5.17. La palabra *súper* en CREA (México)

1975-79	1980-84	1985-89	1990-94	1995-99	2000-04
0	0	0	0	4	3

Recuérdese que la búsqueda en el corpus de 1970 arroja *ceros* casos de la palabra. En cambio, el CSCM presenta un total de 34 ejemplos, que se distribuyen de manera diferenciada según el estrato sociocultural: tres casos en habla de personas de bajos estudios frente a 31 ocurrencias en habla de informantes de alto nivel de instrucción.

(8) la **súper** fiesta [mujer, adulta, culta]

(9) a veces la dejamos **súper** temprano/ porque no la podemos dejar exactamente/ al cinco para la hora que va a empezar su clase [mujer, adulta, culta]

De esta manera, podría concluirse que el uso de *súper* como modificador superlativo adjetival o adverbial (o incluso como interjección), es una innovación en ECM, que ha sido introducida en los últimos 15 años (aproximadamente, si se toma en cuenta las fechas del levantamiento de encuestas para el CSCM), y que es utilizada principalmente en los grupos de mayor educación.

También se realizó la búsqueda del prefijo intensificador *re-* que, en mi experiencia en el D. F., suele asociarse con el habla rural: de hecho, pueden atestigüarse en ocasiones expresiones como *¡están reteharto pobres!*⁴¹, *¡está rebonito!*, entre otras, como maneras de burlarse de los acentos estereotípicamente rurales.

⁴¹ Usualmente pronunciada [re'tjarto#'pobres].

Aparece la forma tres veces en el muestreo de 1970, en habla popular: “bailaba *rebien* los ojos”, “yo lloraba *reteharto*”, y “¡se bañaban *rechistosos!*”. En el año 2000 aparecen solo dos ejemplos en el grupo de altos estudios. En total se trata de sólo cinco casos del prefijo *re-* frente a los 34 de *súper*. ¿Podrían estar compitiendo estas formas de intensificación? Si es así, los datos confirmarían una creciente predilección por la forma innovadora *súper*.

5.6.8. La palabra chido

El surgimiento o expansión en el uso de este adjetivo es difícil de determinar con los datos orales de esta tesis. Sin embargo, coincide la primera documentación en CREA (con fecha 1984) con el uso de esta expresión por parte de personajes cómicos en la televisión mexicana de mitad de los ochenta⁴². En los datos de 1970 no aparece la cadena *chid** en ninguna ocasión (lo que es esperable al tratarse de una expresión informal y, en todo caso, juvenil). En el CREA aparece sólo en siete ocasiones, y el primer ejemplo aparece en 1984 en *El Rayo Macoy y otros cuentos* de Rafael Ramírez Heredia:

⁴² El mejor ejemplo es el Luis de Alba, actor que personificaba a “Juan Camaney” en el programa de alto rating televisivo *La carabina de Ambrosio* en 1985. El grito de guerra de este personaje era la expresión *¡chido, chido, chido!* El personaje tenía todas las pautas de comportamiento del *naco*, que refiere a una persona sin estudios, grosera e indolente.

(10) si Acapulquito es el más **chido** lugar del Filaneta y la Huerta es lo más granado del puerto de palos, pues no hay duda que están ustedes en la mera papa de la galaxia [...]

Como puede observarse, dicho texto trata de reflejar un tipo muy rebuscado de habla popular, de barrio. El CREA la documenta débilmente:

Tabla 5.18. La palabra chido(a) en CREA (México)

1975-79	1980-84	1985-89	1990-94	1995-99	2000-04
0	1	2	1	0	3

Las búsquedas de concordancias en el CSCM, por otra parte, arrojan resultados un poco más interesantes. Al igual que con las palabras *súper* y *onda*, el grupo de mayor escolaridad presenta más ocurrencias de la palabra (17 vs. 3 del nivel bajo):

(11) pues fotografiar y eso sería algo muy *chido* pero bueno a ver [hombre, joven. culto]

Los resultados combinados permiten sugerir la pervivencia del vocablo en el registro oral del español capitalino, pero ya no asociado con el habla popular como

ocurría en los años 1980. Sería, por lo tanto, una innovación léxica *desde abajo*, que se ha ido adoptando en los demás estratos socioeducativos.

5.6.9. La expresión adjetiva con padre/padrísimo

Se buscaron expresiones como *muy/bien padre* y *padrísimo*. El corpus de 1970 no documenta ni un solo caso, en habla culta o popular. En los datos del año 2000 sólo aparecen estas expresiones en el habla de informantes de estudios universitarios (32 ocasiones), lo que permite suponer que sí está asociado a este nivel educativo⁴³:

(12) ay no sé porque tú ves los libros de anatomía y dices “¡ay! pues qué bonito libro y qué *padre* y todo” pero ya cuando lo ves dices “guau” y entonces te interesa más [mujer, joven, culta]

(13) y después nos fuimos a Boca del Río a un hotel *padre* y está muy *padre* Boca del Río pues es la zona más bonita ¿no? [mujer, joven, culta]

⁴³ De hecho, mi intuición de hablante con poco más de 20 años radicando en el Distrito Federal, es que *padre* es una versión aceptada o neutra para calificar algo que en un registro más popular (y grosero) sería *chingón*. Se exploró esta última palabra, pero arrojó sólo 6 datos en el corpus del año 2000. Nuevamente, la explicación debe encontrarse en el hecho de que la situación de entrevista propicia que se eviten las palabras consideradas groseras.

La primera documentación de *padrísimo* (en CORDE) la encontramos en *La región más transparente*, de Carlos Fuentes (1958):

(14) Doscientos más que aquí, y en un barrio *padrísimo*: por el rumbo de Nuevo León

Tabla 5.19. *Expresiones adjetivas con padre en el CREA (México)*

1975-79	1980-84	1985-89	1990-94	1995-99	2000-04
4	3	1	0	2	4

Un total de 14 ejemplos se encuentran en el CREA. Curiosamente la curva en el tiempo sería inversa respecto a una curva de distribución normal, pero se trata de muy pocos datos y no es prudente hacer interpretaciones por el momento. En todo caso, puede en el futuro cercano rastrearse su evolución en corpus orales más extensos, que son los que favorecen su aparición.

5.6.10. *La palabra güey*

El trabajo de Palacios (2002) es quizás el único que documenta con cierta exhaustividad los diferentes usos de la expresión *güey* en México, palabra que es analizada en dicho trabajo como un marcador discursivo, ya que independientemente de su función gramatical, cumple funciones de tipo semántico-pragmático como

retención del turno, introducción de información nueva, conservar la atención del interlocutor, reforzar sentimientos de identidad de grupo, etc. Se basa en datos obtenidos a través de grabaciones secretas entre adolescentes de la ciudad de Puebla, sin la intervención del investigador, por lo que se trata de datos muy naturales y espontáneos⁴⁴. La autora destaca que las principales funciones de *güey* son: 1) establecer la identificación de los participantes al evidenciar una relación de confianza y solidaridad, al propiciar retroalimentación afectiva, reforzar identidad, autoestima y dignidad de los hablantes; y 2) estructurar la información: permite ganar la atención de los interlocutores, introduce información nueva y marca una solicitud de refuerzo, aprobación o seguimiento por parte de los interlocutores (Palacios 2002: 110-111, Cuadros 2 y 3).

En los datos orales de esta tesis la palabra *güey* nunca aparece en los años 1970, pero los dos subgrupos del año 2000 presentan un número parecido de ejemplos: 16 los del grupo educativo bajo y 19 los del grupo alto. Esta no diferenciación entre niveles también se documenta en el mencionado trabajo de Palacios (2002), por lo que puede suponerse que se trata de una expresión “general” en español mexicano, al menos entre jóvenes:

⁴⁴ La autora incluso desechó las grabaciones en las que se notaba que quien portaba la grabadora estaba forzando la aparición de la interdicción (Palacios 2002: 46).

(15) y ayer le decía a a J le decía “*güey* es que esta impotencia no la había sentido” dice “sí *güey* acuérdate Michoacán” y yo “ah neta ya nos había pasado” [hombre, joven, culto]

Nuevamente es en *La región más transparente* de Carlos Fuentes donde se tiene la primera documentación de esta expresión en datos del CREA:

(16) ¡Si no venimos a tomar tecito! -- ¡Ordéñalas, *güey*!

Si se observa la distribución temporal de la palabra en CREA, parecería que la expresión *güey* se encuentra en franco retroceso:

Tabla 5.20. La palabra *güey* en CREA (México)

1975-79	1980-84	1985-89	1990-94	1995-99	2000-04
67	33	20	16	13	1

Sin embargo, volviendo a los datos del CSCM, puede verse que en realidad se trata de una expresión productiva aún⁴⁵. Se encuentra, entonces, otra contradicción entre los datos de lengua hablada y lengua escrita⁴⁶.

5.6.11. Síntesis de resultados del análisis de palabras aisladas

La siguiente tabla y su gráfica permiten entender mejor el patrón que ha seguido cada palabra en el tiempo real, en los datos de CREA.

Tabla 5.21. *Evolución reciente de algunas palabras en español mexicano (CREA)*

	1975-79	1980-84	1985-89	1990-94	1995-99	2000-04
<i>onda</i>	65	29	19	11	17	6
<i>género</i>	0	0	9	9	17	71
<i>sexo</i>	11	2	9	1	6	6
<i>celular</i>	0	0	0	4	52	48
<i>súper</i>	0	0	0	0	4	3

⁴⁵ Un trabajo sociolingüístico de Wendianne Eller (2008) localizó sólo 33 casos de la palabra *güey* en 24 entrevistas del CSCM; cuando realizó grabaciones secretas, recogió 434 ejemplos en menos de dos horas de grabación, lo que evidencia la alta productividad de la palabra en el discurso oral espontáneo de la ciudad de México.

⁴⁶ El panorama cambia muchísimo al revisar el CORPES XXI, que proporciona 388 datos de *güey* (al 20 de junio de 2014). Esto se debe al peso estadístico de los materiales mexicanos, que es mucho mayor que en CREA.

<i>chido</i>	0	1	2	1	0	3
<i>padre</i>	4	3	1	0	2	4
<i>güey</i>	67	33	20	16	13	1

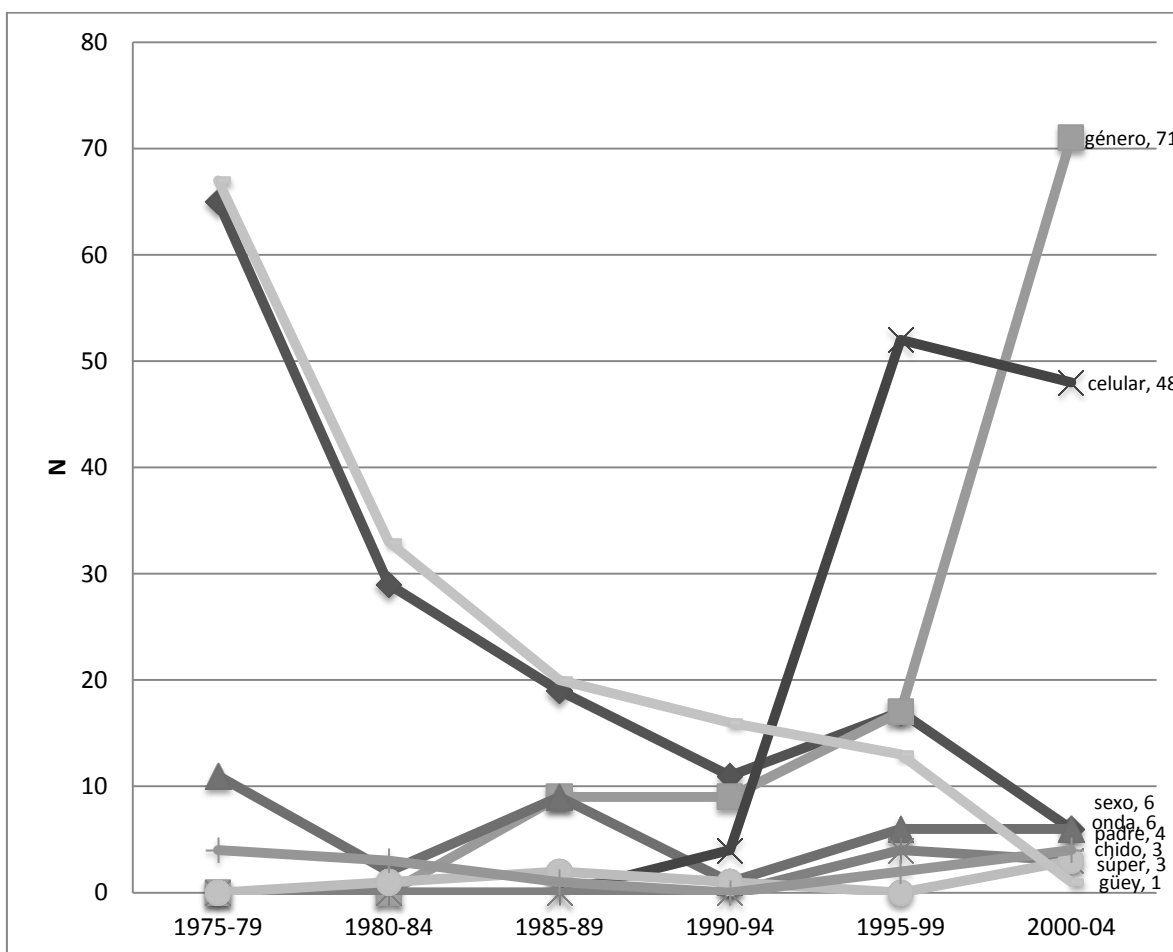


Figura 5.21. Evolución reciente de algunas palabras en español mexicano (CREA)

La gráfica indica que en realidad cada palabra sigue su propio patrón evolutivo. Mientras que *género* y *sexo* parecen seguir direcciones totalmente opuestas (descenso

total en el uso de *sexo* frente al aumento decidido de *género*), las palabras *onda* y *güey* van en retroceso; palabras como *padre*, *súper* y *chido*, por su parte, parecen tener una presencia mínima pero relativamente estable. Por último, la palabra *celular* presenta un patrón ascendente muy pronunciado a partir de los años 1990 que se ha estabilizado durante los últimos diez años. Esta gráfica permitiría apoyar la hipótesis de Guilliéron de que “cada palabra tiene su propia historia”.

Por otro lado, el panorama es muy distinto si sólo se atiende los datos de los corpus originales de lengua oral de esta tesis. Véase la siguiente tabla.

Tabla 5.22. *Evolución reciente de algunas palabras de español mexicano oral*

	1970	2000
<i>onda</i>	1	36
<i>género</i>	0	0
<i>celular</i>	0	4
<i>súper</i>	0	34
<i>chido</i>	0	20
<i>padre</i>	0	32
<i>güey</i>	0	35

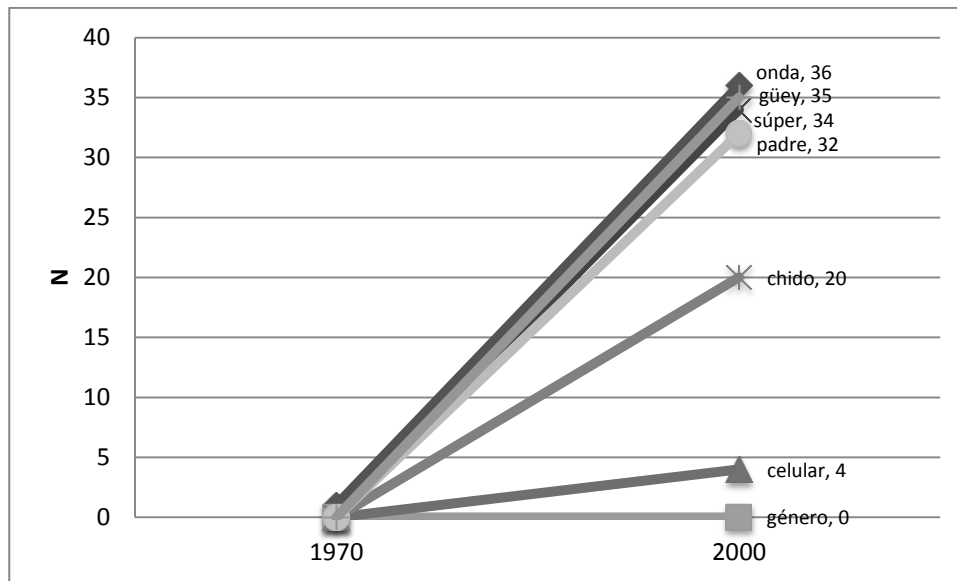


Figura 5.22. Evolución reciente de algunas palabras en español mexicano oral

Esta gráfica, a diferencia de la anterior, en que varios de los ítems léxicos analizados presentan “claros” procesos de retracción, mostraría que todas las palabras (con excepción de *género*) son innovaciones léxicas en el español de la ciudad de México. Observando las gráficas 1 y 2, puede notarse que las palabras *celular* y *género* presentan un patrón estadístico un poco más claro que los demás: podemos determinar prácticamente cuándo se documentan por primera vez y también confirmar que la tendencia es a una mayor frecuencia en su uso. La explicación de este ascenso es diferente en cada caso. El uso del sustantivo *celular* es producto del avance tecnológico y de telecomunicaciones que ha permitido la generalización del uso del teléfono celular en prácticamente todos los estratos socioeconómicos de la ciudad de México, por lo que se vuelve un referente sumamente

cotidiano para toda la comunidad de habla. La palabra *género*, como construcción socio-cultural, por otra parte, ha sido adoptada en el lenguaje político y académico principalmente, siguiendo las normas internacionales propuestas por la UNESCO para diferenciarlo de *sexo* como realidad biológica. Podríamos decir –extendiendo un poco la tipología de cambios lingüísticos de Labov (2006)- que se trata de una innovación léxica *desde arriba*. Por otra parte, el corpus oral base de la tesis muestra que las palabras *súper*, *padre*, y expresiones que incorporan la palabra *onda*, parecen ser marcas sociolingüísticas de los grupos de mayores estudios, quienes en el año 2000 incorporan también las palabras *chido* y *güey*, que se habrían adoptado “desde abajo”. Todos estos vocablos gozan de relativa vitalidad, según la Fig. 5.22.

¿Cuál de los dos patrones es el más fidedigno? Por un lado, las muestras aquí analizadas constituyen apenas un volumen de 389,013 palabras, pero tienen la ventaja de tratarse de datos de lengua oral. El corpus de CREA, por su parte, cuenta con cerca de 17 millones de palabras para la llamada “zona mexicana” (México, El Salvador, Honduras, Guatemala y Suroeste de EE. UU.), pero el 90% de los datos es de lengua escrita. ¿Son realmente equiparables? ¿Cuál refleja mejor la “realidad” de la lengua? Se tratará de buscar respuestas en el siguiente apartado.

5.7. DISCUSIÓN: LA VIABILIDAD DE LOS CORPUS ORALES BASADOS EN ENTREVISTAS PARA EL ANÁLISIS SOCIOLINGÜÍSTICO DEL LÉXICO

El debate lingüístico sobre la mejor forma de construir un corpus lingüístico es todavía muy reciente y no ha llegado a conclusiones satisfactorias. Para Tognini-Bonelli (2001: 55-62) todo trabajo de corpus se enfrenta a tres grandes cuestiones: la autenticidad de los textos, su representatividad y el tipo de muestreo. Si se aplican estos tres criterios para los corpus de español con que se cuenta actualmente, es evidente que todos tienen puntos débiles. En general, lo que caracteriza tanto al CORDE, como al CREA, es que la lengua escrita tiene una preeminencia importantísima en estos y que, por el contrario, la oralidad queda relegada a un incómodo segundo lugar (esto si un 10% de datos orales en el CREA puede ser suficiente para concederle un “segundo lugar”). De hecho, la sociolingüística enfrenta este problema con la mayoría de los corpus actuales, que están basados mayormente en lengua escrita: “Perhaps the main [problem] is the emphasis within corpus linguistics on standard written forms of language. Texts found within most corpora do not contain the kind of material of greatest interest to most sociolinguists, namely, casual everyday speech, often from non-standard language varieties” (Romaine 2008: 97).

Porque una pregunta pertinente para este estudio particular es la siguiente: ¿es suficiente el registro escrito (y el grado de formalidad, estandarización y fosilización que conlleva) para dar cuenta de procesos de cambio lingüístico en *lapsos tan breves*

como el que se estudia aquí? La respuesta provisional es “no”. En esta exploración pude reconocer que mi experiencia cotidiana como hablante (y lingüista) está lejos de ser confirmada por los hallazgos en muestreos de habla tan estructurados como los empleados aquí (NC, HP y CSCM) que, sin embargo, reflejan una realidad lingüística del español que puede estar muy alejada del *input* cotidiano con el que se interactúa, en el que la lengua oral tiene un papel primordial. La relativa escasez de ejemplos en estos conjuntos de datos orales obliga a concluir que, al menos en materia de léxico, la diferencia entre *lo posible* desde la competencia lingüística y *lo manifestado* en la actuación puede ser enorme; por lo tanto, hay que atender siempre al otro extremo del problema: aquí se muestra el producto --en un sentido *humboldtiano*, el *ergon*-- pero hay que abordar la *energeia* (el mecanismo de producción) para entender mejor la evolución del léxico, y de cualquier nivel estructural para el caso (sobre todo si se aborda el lenguaje desde una perspectiva no modular, como la cognitivista --cf. Morenos Fernández 2012: 38-39).

Por supuesto, una solución sería construir un corpus lo suficientemente representativo de los usos cotidianos de la lengua para poder obtener conclusiones más sólidas. Ha habido propuestas muy interesantes de corpus basados en el uso, como el de Quirk para el inglés (1974), quien consideraba dos grandes tipos de lengua: hablada y escrita. Su esquema, previo a la era de los grandes corpus electrónicos, era bastante detallado:

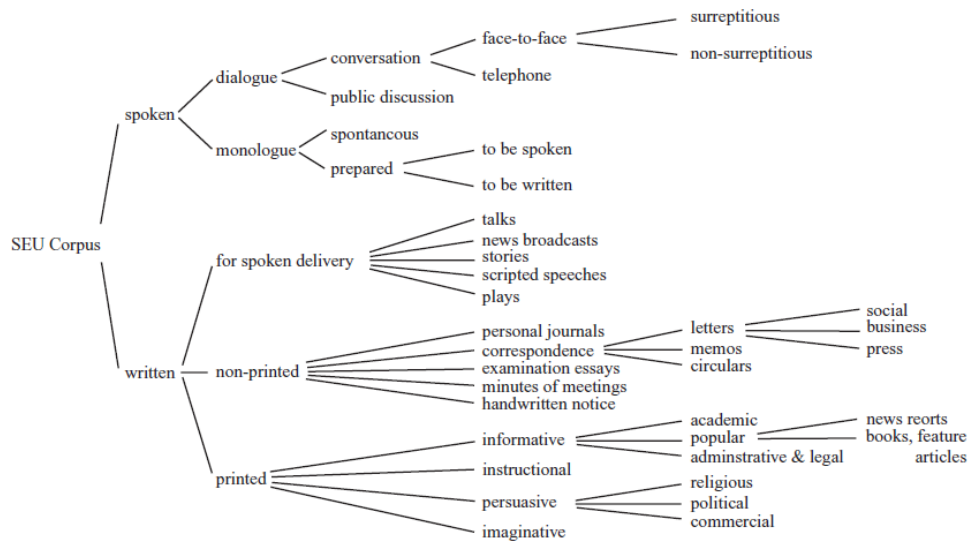


Figura 5.23. *Survey of English Usage* (Quirk 1974)

Por supuesto, en estos tiempos habría que añadir a tal esquema la lengua de internet (o “tecleada”, como la llama González Cáceres 2012) como otra gran variedad en el continuum diamésico (Berruto 2010). Esto implicaría incluir en un corpus así géneros televisivos relativamente recientes, como los *reality shows* y *talk shows*, o los propios de internet, como *videoblogs*, textos en redes sociales como *Facebook* y de comunicación por teléfonos celulares e inteligentes (mensajes SMS).

Volviendo a la cuestión oralidad-escritura, piénsese un experimento en el que se realiza un conteo, en una muestra cualitativa y cuantitativamente representativa (esto es, que fuera una muestra ponderada) de hablantes del español

de la ciudad de México, del número de oraciones –cualquiera que sea la noción de “oración” adoptada- que escuchan y emiten por un lado, y también de aquellas que leen o escriben por otro. ¿Cuál sería la proporción entre un registro y otro? Seguramente, para la mayoría de participantes en este hipotético muestreo, las oraciones construidas e interpretadas desde la oralidad tendrían un peso estadístico mayor que las provenientes de la lengua escrita. Por tanto, si este es el input al que están expuestos y, como se sabe, los cambios lingüísticos se propagan en su gran mayoría por la interacción oral cara a cara, sobre todo cuando el informante presta la menor atención a la “corrección” en su forma de hablar (Chambers & Trudgill 1998: 48), la utilización de un corpus basado en su mayoría en registros escritos resulta una ruta ciertamente inadecuada. Las preguntas que vale la pena formular son: ¿cuál es el sentido de marginar los textos orales de los grandes corpus? ¿Qué tipo de español se está documentando: el de la población hispanohablante general, o sólo el de algunos sectores sociolingüísticos? ¿Qué proporción de la población hispanohablante interactúa con los textos escritos? Y en última instancia: ¿cuál es el peso de la lectoescritura en los procesos de variación y cambio?

Incluso para pensar un corpus de lengua escrita deberían considerarse los tipos de texto escrito que la población muestreada realmente consume. La *Encuesta Nacional de Lectura* realizada en México (CONACULTA 2006) señala que el perfil lector mayoritario -en una muestra de 4,080 personas- es el de “lector esporádico” (19%), aquél que presenta hábitos de lectura “utilitarios”, esto es, que lee libros de texto escolares, manuales de usuario, periódicos, etc., sólo para obtener información

de utilización inmediata (63.8%). El segundo tipo de lector es el que lee “publicaciones diversas” (periódicos, historietas, revistas, libros, etc.) y su tipo textual favorito (para un 27.6% de ellos) es precisamente la historieta. ¿Por qué no construir un corpus en el que estos tipos textuales más populares -libros de texto, historietas y periódicos- fueran mayoritarios? Otro tipo de texto que no suele incluirse en los macro corpus es el publicitario: carteles, anuncios espectaculares, etiquetas, comerciales televisivos, slogans, entre otros. También deberían incorporarse los libros más vendidos, los *best-sellers*. Esto es, si uno de los objetivos es contar con una muestra del input escrito al que se expone la comunidad de habla, debería muestrearse precisamente lo que “las mayorías” leen, lo que desde mi punto de vista no está suficientemente bien representado en los corpus mencionados, algo que he discutido recientemente (Serrano 2013).

Evidentemente, estas exigencias a los macro corpus están pensadas desde la perspectiva del lingüista, quien prefiere una idea más realista de la lengua para poder llegar a generalizaciones de mayor alcance descriptivo y teórico, y no desde el punto de vista del hablante o del purista que prefiere recopilar el “tesoro” de la lengua. Una cuestión es tratar de documentar la riqueza léxica de una lengua -que suele ser mejor cultivada en los textos literarios y académicos- y otra muy distinta es tratar de documentar el acervo propio de la actuación lingüística, los procesos de cambio lingüístico en marcha o simplemente obtener un panorama de tipo realista al acervo léxico de una lengua. En este último caso, el uso de un corpus representativo que

tomara en cuenta los hábitos de lectura de la comunidad de habla sería el más apropiado.

Regresando al CREA, los 160 millones de palabras que lo constituían en mayo de 2008 estaban distribuidos de la siguiente manera: 90% lengua escrita; 10% lengua oral; 50% español de España y 50% español de América⁴⁷. Afortunadamente, la propuesta del CORPES XXI (RAE 2014) propone una distribución demolingüística más realista, de manera que ahora las proporciones son de 30% de español de España y 70% de español de América. Por supuesto, lo ideal sería que este corpus tomara en cuenta otros factores –mucho más difíciles de definir– como el área de influencia de dialectos particulares en otros países, su prestigio, tradición oral y literaria, etc. De seguir estos criterios, el habla mexicana debería ser la mayoritaria en un corpus del español, simplemente por ser la variedad que cuenta con mayor número de hablantes y con gran influencia mediática en Latinoamérica (Ávila 2011).

Resultados como los de esta exploración obligan a plantear una serie de preguntas fundamentales para el quehacer lingüístico: ¿cuál es el valor de nuestros datos? ¿Es suficiente con observar lo manifestado en un corpus o se deben explorar otros métodos de recolección de datos? ¿Cómo trascender los límites de los corpus lingüísticos de habla levantada en el campo? De entrada puede pensarse en un enriquecimiento de tipo cualitativo; pueden incorporarse materiales de distintos estilos

⁴⁷ En esta época se recogieron los datos aquí discutidos. Se mantiene los datos de esta versión por tratarse de una comparación con datos orales levantados entre 1970 y 2000.

de habla en un eje de formalidad/informalidad; pueden incorporarse textos orales provenientes de conversaciones telefónicas, de locución de radio y televisión, *blogs* de internet, etc. Por otra parte, con las facilidades de almacenamiento y procesamiento de las computadoras actuales, un solo investigador puede manejarse corpus masivos, de cientos de millones de palabras, de manera automatizada y rápida. Finalmente, este trabajo ha evidenciado que la introspección del investigador, la obtención de datos a través de la elicitación con informantes o el levantamiento de cuestionarios detallados son herramientas que deben complementar el estudio lexicoestadístico y de los procesos de variación y cambio léxico de palabras particulares que, como se ha visto en este capítulo, son altamente dependientes de la constitución particular del corpus que se utilice.

5.8. CONCLUSIONES

El estudio de niveles estructurales distintos del fonético abre la discusión sobre los rumbos distintos que pueden tener los procesos de variación y cambio según el lugar donde se ponga el foco de investigación. En general, los hallazgos de este capítulo, en particular los fenómenos de los grupos consonánticos cultos y las distintas aproximaciones lexicoestadísticas (en particular los índices de RTT y el volumen general del vocabulario sustantivo), permiten reforzar la idea sugerida en los Caps. 3 y 4 sobre un relativo acercamiento sociolingüístico de los grupos educativos estudiados en el tiempo real. Por otra parte, la última sección de este capítulo, sobre palabras

aisladas (5.6), hizo evidentemente algunas limitaciones de los corpus orales basados sólo en entrevistas para el análisis de palabras “emblemáticas” de ciertos grupos sociales y que siempre es necesario acudir a corpus de lengua escrita para poder trazar las historias individuales.

La hipótesis de que ciertos cambios lingüísticos se llevan a cabo “palabra por palabra” se refuerza con los hallazgos sobre los grupos consonánticos cultos y los adverbios terminados en *-mente*, que complementan los resultados del Cap. 4 sobre el debilitamiento vocálico, también vinculados a la frecuencia de ciertas palabras con el contexto más favorecedor.

Por supuesto, la retracción registrada en el Cap. 3 sobre la asibilación de róticas, de comportamiento más al estilo neogramático, provoca la pregunta de por qué pueden ocurrir los dos tipos de cambio. Una respuesta puede estar en los distintos grados de procesamiento cognitivo implicados: el comportamiento fonético variable es, al menos en sus primeras etapas, menos prominente que, por ejemplo, la incorporación de un grupo de palabras (p. ej., las vinculadas con la tecnología de cómputo en tiempos recientes) al acervo léxico comunitario, o la elección de ciertas formas morfológicas. Quizás se trata de procesos que involucran *indicadores* (no presentes en la conciencia sociolingüística de los hablantes pero caracterizadores de dialectos y sociolectos) por un lado, y *marcadores* (procesos variables que sí son sensibles a la evaluación social y que pueden devenir en *estereotipos*), por otro (cf. Chambers & Trudgill 1998: 71-76).

La discusión general se retoma en el siguiente capítulo, donde se detallarán los hallazgos de la tesis y sus posibles explicaciones.

CAPÍTULO 6

RECAPITULACIÓN. PROCESOS SOCIOLINGÜÍSTICOS DE CONVERGENCIA Y NIVELACIÓN EN ESPAÑOL DE LA CIUDAD DE MÉXICO

Los diferentes acercamientos a la variabilidad del español de la ciudad de México (ECM) en esta tesis permiten trazar ciertas historias sociolingüísticas de interés. Algunas siguen un camino de retracción (asibilación de róticas), otras, de creciente marcación en la conciencia sociolingüística comunitaria (debilitamiento vocálico). Finalmente, se producen procesos en el nivel morfológico y léxico que parecen reflejar con sorprendente fidelidad el desarrollo de la comunidad lingüística estudiada como una entidad que, antes que nada, es una comunidad histórica y social.

Este capítulo, que funcionará como un espacio para la discusión y la presentación de las conclusiones generales de la tesis, se organiza en cuatro apartados principales. En el primero se analizan en conjunto los procesos variables del nivel fonético, morfológico y léxico que indican cierta nivelación sociolingüística de las hablas capitalinas. En el segundo se discuten distintas explicaciones posibles de estos procesos de convergencia y nivelación. En el tercero se presenta un ensayo de geolingüística urbana en el que se trata de ubicar a los líderes de algunos de los procesos variables estudiados en la geografía del Área Metropolitana de la Ciudad de México. Finalmente, se evalúa la importancia de los hallazgos, las ventajas y limitaciones de los procedimientos analíticos abordados en la tesis y se señalan algunos otros fenómenos dignos de estudiar en el tiempo aparente y el tiempo real en ECM.

6.1. PROCESOS DE NIVELACIÓN SOCIOLINGÜÍSTICA EN EL NIVEL FONÉTICO, MORFOLÓGICO Y LÉXICO

En esta subsección se muestra cómo en el nivel fonético se presenta una tendencia comunitaria en el año 2000 a un estándar de pronunciación más conservador respecto a los años 1970. Estas tendencias a la convergencia sociolingüística parecen reflejar un proceso más generalizado en el país, debido en buena medida a los grandes procesos de urbanización y flujos migratorios a partir de los años 1940 (Martín Butragueño 2010).

6.1.1. Róticas asibiladas

Frente a la hipótesis de Perissinotto (1975), de que las variantes asibiladas de (r) y (r) serían mayoritarias en el futuro, los resultados en el tiempo real mostraron que, en realidad, el proceso tomó una dirección opuesta: la asibilación sigue siendo una variante presente en ECM, pero su probabilidad de aparición, con el paso del tiempo, es mucho menor, tal como se muestra en la siguiente figura:

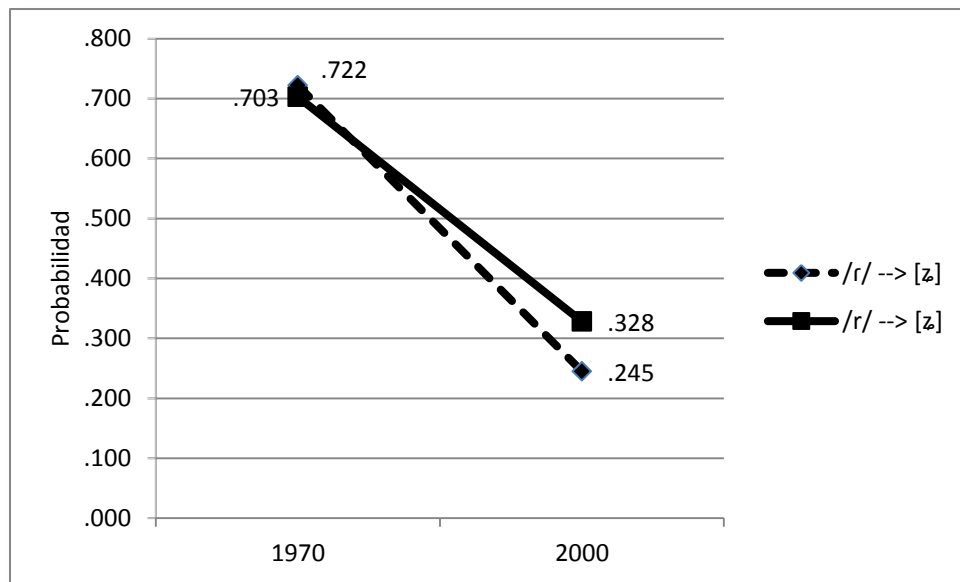
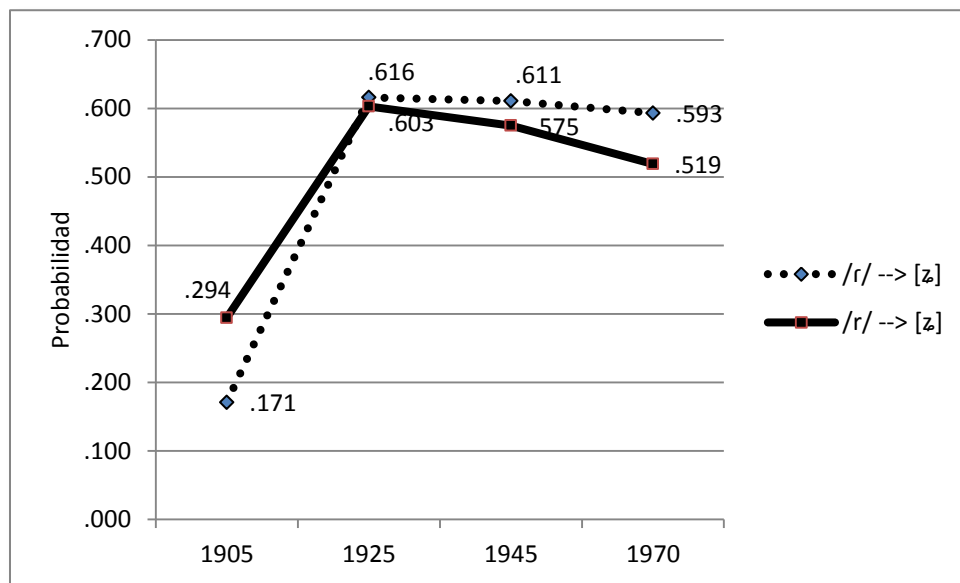


Figura 6.1. Retracción de la asibilación de róticas (1970-2000)

La gráfica anterior muestra cómo desciende fuertemente la probabilidad de selección de las asibiladas en el tiempo real –de hecho, pasan de ser favorecidas (esto es, de tener una probabilidad de aparición con valor arriba de .500) a no serlo del todo

en el año 2000. Al analizar la muestra total de 48 informantes, agrupados por año de nacimiento, se confirma esta tendencia a la retracción del rasgo de asibilación:



Gráfica 6.2. Asibilación de (r) y (r) según el año de nacimiento de 48 informantes (datos de regresión escalonada, (r): $f=260/1129$, sig.=.000; datos de (r): $f=193/2274$, sig.=.000)

Esta gráfica permite observar el efecto del tiempo en la comunidad, a través de los años de nacimiento de cada informante; la ventaja evidente es que aquí no importa el año en que se realizó la entrevista. En ambas róticas el proceso es muy similar: hay un punto máximo de asibilación de los hablantes nacidos en la segunda generación (1925) --el grupo adulto de los años 1970--, para luego descender en las siguientes dos generaciones. Como se vio en el Cap. 3, las variantes más fidedignas (en el sentido

optimal) ganan terreno en el tiempo real, presentando gráficamente un comportamiento en espejo a la figura 6.2.

6.1.2. Elisión de vocales átonas

En este caso, el análisis de regresión escalonada no seleccionó como significativa la variable año de muestreo, lo que indicaría que no habría grandes diferencias entre 1970 y 2000 en cuanto a los índices de elisión. Sin embargo, la variable año de nacimiento (que permite combinar los dos tiempos aparentes con el tiempo real), sí es significativa en los recorridos logísticos:

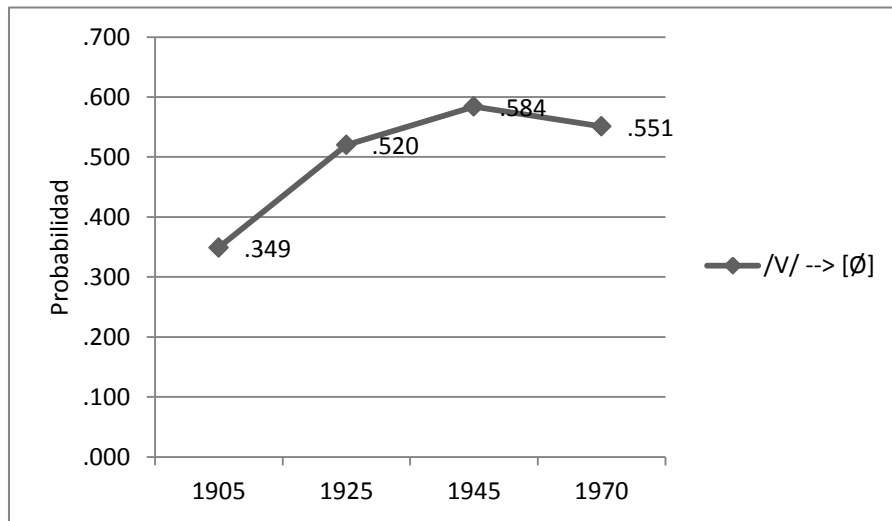


Figura 6.3. Elisión vocálica según el año de nacimiento de 48 informantes (datos de regresión escalonada, $f=317/1403$, sig.=.025)

Nuevamente, se traza una curva de ascenso que se retrae en la última generación (“1970”, los nacidos entre 1961 y 1981). La recuperación del material fónico, atestiguada en un mayor porcentaje de vocales sordas plenas (que pueden verse como “sustitutas” de las elisiones, v. Cap. 4), es evidencia de que se trata del mismo proceso de alejamiento de las formas innovadoras --o un acercamiento a las formas estándar, como quiera verse-- tanto en vocales como en vibrantes.

En un sentido sociolingüístico, los grupos más extremos parecen acercarse entre sí en el tiempo real. Véanse las siguientes figuras, donde se muestran los casos totales de variantes asibiladas de la percusiva y de elisión vocálica.

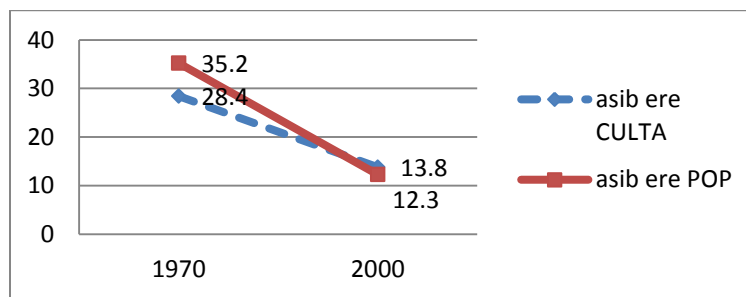


Figura 6.4. Frecuencia relativa de asibilación de (r) en ECM (1970-2000) (N=1129)

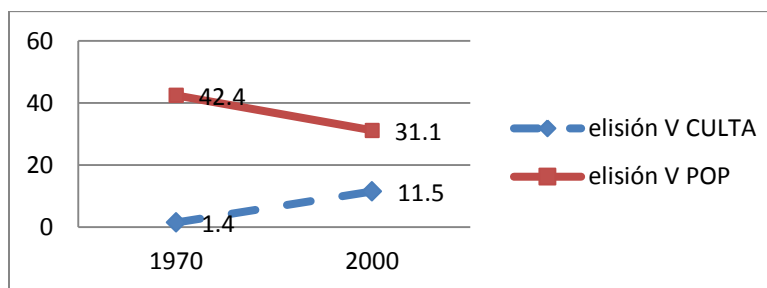


Figura 6.5. Frecuencia relativa de elisión vocálica en ECM (1970-2000)

Puede verse en la Fig. 6.5 cómo el grupo culto presenta en 2000 algunas elisiones (prácticamente ausentes en 1970), pero el grupo de pocos estudios baja en el mismo rubro, acortando con esto las diferencias. Por eso se interpreta éste como un proceso de nivelación sociolingüística. La Fig. 6.6. muestra un descenso comunitario respecto a la asibilación de (r), aunque los grupos parecen un poco más diferenciados en el año 2000.

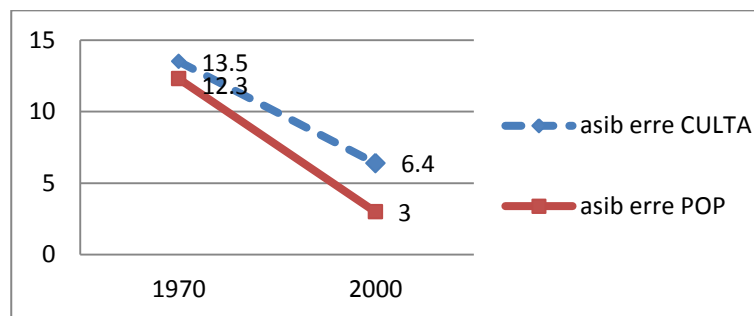


Figura 6.6. Frecuencia relativa de asibilación de (r) en ECM (1970-2000) (N=2274)

Por lo tanto, existe una baja global de los tres procesos, pero en especial las gráficas 6.4 y 6.5 (que muestran el abandono de la variante asibilada de /r/ y de las elisiones vocálicas), son evidencia de un acortamiento de la brecha sociolingüística entre los grupos educativos estudiados.

6.1.3. Los morfemas -dad y -mente como testigos de nivelación sociolingüística

El nivel morfológico muestra un patrón general de descenso comunitario en el uso de sufijos asociados con palabras “cultivadas”. Véanse las siguientes figuras.

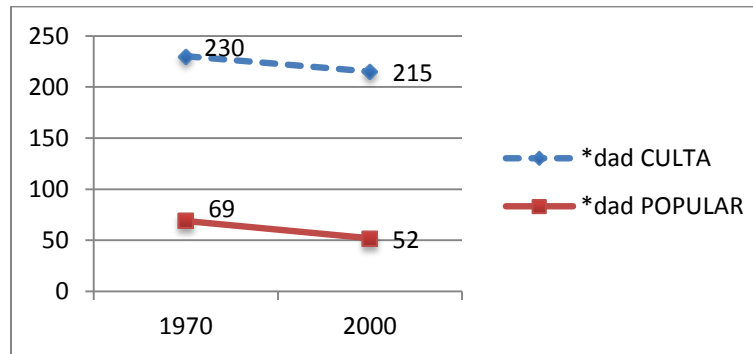


Figura 6.7. Palabras con sufijo -dad en ECM (1970-2000) (N=566)

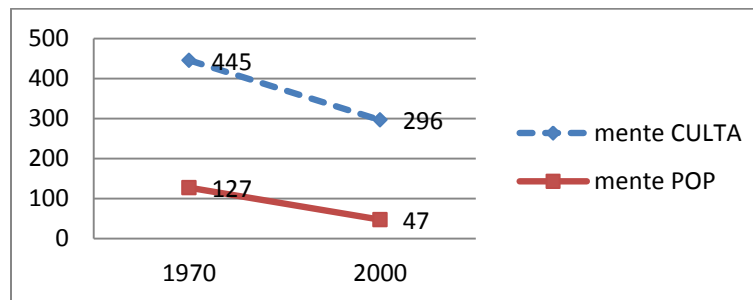


Figura 6.8. Adverbios terminados en -mente en ECM (1970-2000) (N=915)

También en cuanto a la morfología “cultivada” los hablantes de ECM de esta investigación comparten el comportamiento hacia la baja. Las líneas de los grupos cultos y populares en esta ocasión corren de manera casi paralela, aunque puede

notarse un ligero acortamiento de la brecha en el caso de los adverbios terminados en *-mente*.

Por lo tanto, si se toma en cuenta los niveles estructurales analizados hasta aquí, resulta que se cuenta en el año 2000 con una comunidad que está dejando rasgos morfológicos que indicarían un alto nivel de formalidad en el estilo de entrevista; a la vez, recorren estos hablantes, en términos de pronunciación, un camino que los acerca a un estándar de pronunciación un tanto más conservador. Podría decirse entonces que están menos diferenciados entre ellos, pero que se comportan como una misma comunidad en cuanto al abandono de la formalidad en el lenguaje.

El nivel léxico confirma este patrón general, como se ve en el siguiente subapartado.

6.1.4. Sociolingüística de los grupos consonánticos cultos y el léxico sustantivo

El patrón de convergencia sociolingüística se confirmó con los grupos consonánticos cultos (GCC), a pesar de tratarse de un problema más bien vinculado al léxico, ya que en realidad los GCC no fueron el objeto en sí de análisis, sino que sirvieron para *localizar* palabras cultas. El hecho de que se trate de un problema léxico podría confirmarse con la coincidencia de este patrón de nivelación sociolingüística con el del análisis del léxico sustantivo y la relación type/token (RTT), tal como se muestra en las siguientes figuras.

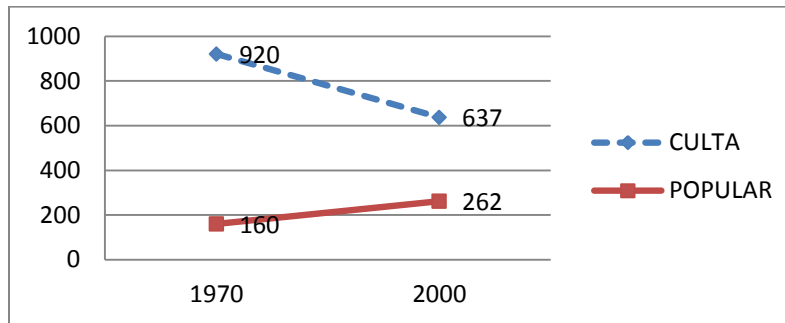


Figura 6.9. Grupos consonánticos cultos en ECM (1970-2000) (N=1979)

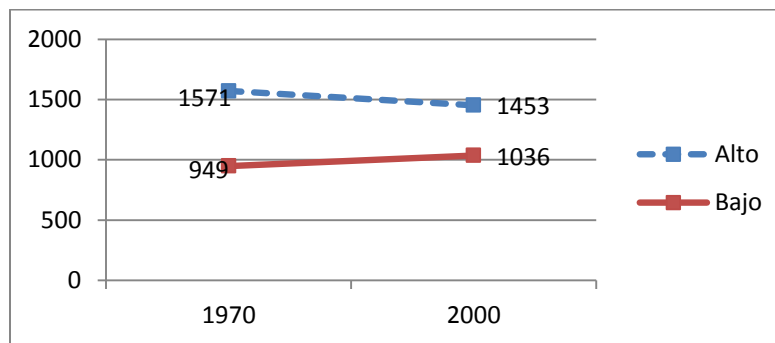


Figura 6.10. Vocablos sustantivos en ECM (1970-2000) (N=5009)

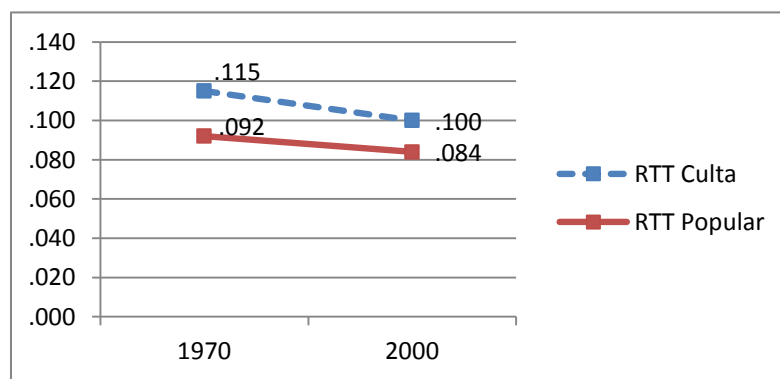


Figura 6.11. Relación type/token (RTT) global de 48 informantes (1970-2000)

La RTT comunitaria muestra una tendencia hacia la baja, pero los grupos sociales están un poco menos diferenciados; este acercamiento sociolingüístico es más patente en el léxico sustantivo global y los GCC (Figs. 6.9 y 6.10), procesos en los que los grupos de bajos estudios incluso aumentaron su contribución estadística. ¿Qué puede explicar estos procesos? Esto es lo que se analiza en la siguiente sección.

6.2. PROCESOS DE CONVERGENCIA SOCIOLINGÜÍSTICA EN ECM: EN BUSCA DE EXPLICACIONES

En general, los patrones de variación en el tiempo real sugieren una comunidad lingüística más homogénea y que tiende a la informalidad, con grupos sociales extremos más cercanos entre sí en el tiempo real, lo que puede observarse en los tres niveles estructurales revisados. En los aspectos fonéticos, lo que se documenta es una tendencia comunitaria a un estándar de pronunciación normativo, en el que se penalizan las variantes menos fieles al input rótico (esto es, /r/ y /r/ asibiladas) y se presenta una baja en los índices de elisión vocálica. Morfológicamente y fonoléxicamente, se produce un descenso generalizado de las formas cultivadas, acompañada de una diferenciación sociolingüística cada vez menor. Finalmente, el nivel léxico en general muestra una comunidad con un acervo sustantivo de iguales dimensiones en el tiempo real, pero con una distribución sociolingüística que también apunta hacia la equidad entre grupos sociales.

Con excepción en la asibilación de (r) y el número de adverbios terminados en *mente*, el resto de gráficas indican una brecha sociolingüística reducida en el tiempo

real. Este panorama puede tener al menos tres posibles explicaciones: a) las entrevistas levantadas en épocas distintas siguieron metodologías tan diferentes que se trataría más bien de diferencias de material debidas al método de recolección utilizado, antes que un cambio lingüístico comunitario; b) el *diseño de la audiencia* (Bell 1984) ha cambiado con el paso de los años, de modo que en 1970 la entrevista como acto comunicativo exigía una mayor formalidad y en el año 2000 se experimenta la entrevista como un acto menos formal que se reflejaría necesariamente en el comportamiento lingüístico de los informantes; finalmente, c) en efecto, sí hay un cambio lingüístico comunitario que refleja lo que ha ocurrido en la sociedad capitalina en los últimos años, que pasó de una de tipo conservador en los años 1970 a una más “de avanzada” en algunos aspectos como mayor movilización política, ciertos procesos de equidad económica y el aumento generalizado en el nivel educativo, factores que explicarían el acercamiento de los grupos sociales extremos¹. A continuación se discuten estas tres posibilidades.

¹ Para Bell (1984: 156), el caso del persa de Teherán (Jahangiri 1980), en el que la generación más joven reducía las diferencias sociolingüísticas, no se debía al efecto de la equidad social porque las muestras se habían levantado *antes* de la Revolución iraní de 1979. En el caso que nos ocupa –y como se discutirá más adelante--, sí existen movimientos sociales, políticos y económicos que indican una significativa transformación de la sociedad capitalina a partir de los años 1960.

6.2.1. Hipótesis sobre las diferencias en el levantamiento de los datos

El propósito principal de esta revisión es mostrar que, a pesar de que los muestreos se realizaron con una distancia temporal mayor a los 30 años, en realidad se trató de procesos de recolección de datos muy similares.

6.2.1.1. La situación de entrevista en 1967-1975

En general, fueron estudiantes de la Licenciatura en Letras Hispánicas de la Facultad de Filosofía y Letras de la UNAM quienes realizaron la gran mayoría de las entrevistas para el proyecto Norma culta. Estas se realizaban partiendo de las redes sociales de los mismos estudiantes. Estas redes sociales se caracterizan por un nivel de estudios superior (más de 12 años de escolaridad) en la mayoría de los informantes y un nivel de ingresos que les permitía viajes al extranjero, aprender otros idiomas (especialmente inglés y francés) y ciertas experiencias en el ámbito de “lo cultural”, como acceso al teatro, danza, cine, etc. --para todo esto, cf. Aguirre & Chico (2011).

Respecto a las entrevistas para el proyecto Habla popular, la mayoría de los entrevistadores son los mismos estudiantes de la Facultad de Filosofía y Letras de la UNAM; los informantes solían ser trabajadores de servicio doméstico en sus casas o en las de sus familiares. Por supuesto, muchas otras entrevistas se realizaron en otros contextos, como en los alrededores de Ciudad Universitaria, peluquerías, hospitales o en la zona semi-rural del Ajusco, al sur de la ciudad.

Los temas abordados en las entrevistas suelen ser: ocupación del informante, historia de vida, la familia, viajes y tiempo libre (Serrano 2009).

6.2.1.2. La situación de entrevista en 1997-2007

El CSCM, al ser un levantamiento estrictamente sociolingüístico (vinculado, como ya se ha mencionado, al macro-proyecto PRESEEA), está muy diversificado, por lo que se entrevistó a personas de todas las clases sociales, grupos de edad y niveles de estudio que fuera posible. Lo que se hizo para esta tesis es una selección que pudiera emular en lo posible las características sociales de los informantes de los años sesenta-setenta, que como ya se comentó arriba, estaban concentrados en los extremos de la estratificación socioeconómica capitalina. En las entrevistas seleccionadas, en general, fueron alumnos de la Licenciatura en Letras Hispánicas de la Facultad de Filosofía y Letras de la UNAM y de la Licenciatura en Lingüística de la Escuela Nacional de Antropología e Historia quienes realizaron las entrevistas. La metodología para levantar el *Corpus Sociolingüístico de la Ciudad de México* (en adelante CSM) (Martín Butragueño y Lastra 2011) es de *redes sociales* (Milroy 1980; Lastra y Martín Butragueño 2000). La mayoría de las redes tienen como punto de partida al entrevistador. Estas redes sociales, al igual que en las entrevistas de los años 1970, se caracterizan por un nivel de estudios superior en la mayoría de los informantes, aunque el nivel de ingresos del año 2000 no permite los privilegios de los años 1970. En general, la preparación en términos de idiomas y el acceso a la cultura son

similares a sus pares de 1970, pero ahora existen más restricciones económicas respecto a la posibilidad de hacer viajes al extranjero.

Las entrevistas de los informantes de baja escolaridad también las realizaron en su mayoría los mismos estudiantes que entrevistaron a los de niveles altos de estudio. Los contextos son variados, generalmente se trata de personas en el área vecinal, en su contexto de trabajo o familiares, y los informantes habitan en diferentes zonas de la ciudad, incluso en delegaciones con población semi-rural como Tláhuac, al sureste de la ciudad.

Los temas tratados durante las entrevistas son en su mayoría: ocupación del informante, historias de vida, la familia y tiempo libre (cf. Martín Butragueño y Lastra 2011).

6.2.1.3. Evaluación de la comparabilidad de los muestreos de 1970 y 2000

El recuento de la situación de entrevista en una época y otra en realidad denota que se llevaron a cabo en situaciones sociolingüísticas y de interacción social muy similares: ambos corpora partieron de proyectos de investigación con objetivos propiamente lingüísticos, casi todos los entrevistados formaban parte de redes sociales que partían de los entrevistadores, todas las entrevistas se grabaron magnetofónicamente con grabadoras en presencia del entrevistado (excepto dos), casi todos los entrevistadores fueron estudiantes de licenciatura, etc. En conclusión, no existen diferencias

metodológicas significativas entre ambos muestreos, por lo que no puede estar aquí la explicación de los resultados².

6.2.2. Hipótesis sobre el cambio en el diseño de la audiencia

Esta hipótesis plantearía que la manera en que los informantes responden a una entrevista socio-lingüística ha cambiado entre 1970 y el año 2000, en el sentido de un cambio en el *diseño de la audiencia* (Bell 1984) y que ello explicaría las diferencias en su comportamiento lingüístico. Por *audiencia*, Bell entiende lo siguiente:

The audience is, at one level, simply the people who hear the speaker's utterances. Yet their role is by no means passive. As in a theatre, the audience is the responsive, critical forum before whom the utterances are performed. Under an older meaning, speakers "have an audience with" their hearers. They are in a real sense subject to their audience, dependent on its goodwill, responsive to audience response. It is that responsiveness which informs a

² Por supuesto, se reconoce aquí que las posibles diferencias en cuanto a la percepción y dinámica de la situación de entrevista entre hablantes de los distintos estratos, sexo y edad juega un papel, "siquiera parcialmente" (cf. Moreno Fernández 2012: 2010-211) sin embargo, también es cierto que las diferencias de percepción, de haberlas, se dieron en condiciones socio-situacionales equivalentes en el tiempo real, lo que otorgaría comparabilidad a los datos.

speaker's style design. [Bell 1984: 161]

La siguiente figura muestra el modelo básico de Bell:

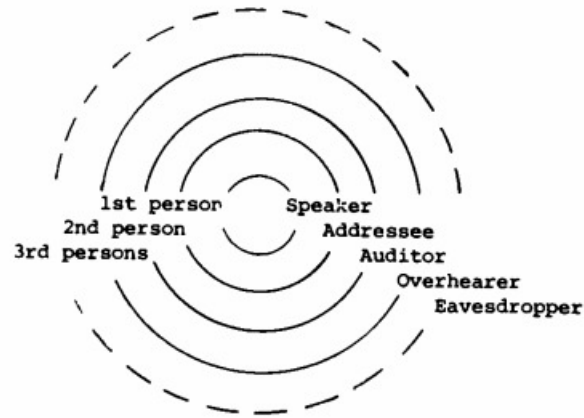


FIGURE 5: Persons and roles in the speech situation.

Fig. 6.12 Modelo básico de diseño de la audiencia de Bell (1984: 159)

El importante trabajo experimental de Douglas-Cowie (1978) sobre cambio de estilo en un poblado rural irlandés apoya la idea de Bell de que este cambio por parte del hablante está más orientado hacia el destinatario del mensaje (el *addressee* o “segunda persona” en la figura anterior) que hacia el oyente-audiencia (*auditor*, tercera persona presente en la entrevista –Bell 1984: 172). La naturaleza del destinatario está construida culturalmente y puede ser más imaginada que concreta (como sucede en los medios masivos como TV o radio); el diseño de la audiencia por parte del hablante es tan importante que lo obliga a mantener su estilo a pesar de que cambie el entrevistador (un destinatario en sí), como sucede en el clásico estudio de

Labov (1966) sobre la variable (r) en tiendas departamentales de Nueva York, en el que, a pesar de tener al mismo entrevistador anónimo, los hablantes de las tres tiendas se comportaron como lo hacen con su clientela habitual (Bell 1984: 170).

En todo caso, debe recordarse que Bell reconoce dos tipos de cambio de estilo: *responsive* (como reacción ante la situación) y *por iniciativa* del hablante, pero el más importante de los dos sigue siendo el primero: “It is clear that the use of style to redefine situations draws its force from the regular association of certain styles with certain situations. Without the basis that a given style regularly *results* from a given situation type, the style could not be used as a means to *create* that type of situation” (Bell 1984: 183-184, énfasis en el original). Esta idea coincide con lo propuesto por Halliday (1978: 31 y ss.) –ya citado en el Cap. 2–, respecto a que el registro es una forma de predicción.

La propuesta de Bell (1984) permite suponer dos aspectos fundamentales: 1) que el cambio de estilo del hablante sí está condicionado por el diseño de audiencia, y 2), que dicho cambio es ante todo una respuesta a la situación comunicativa particular –de esta segunda suposición puede concluirse, por ejemplo, que todos los informantes de este estudio, al encontrarse en la misma situación de entrevista, presentan un estilo de habla perfectamente comparable–. Un punto que resulta crucial, sin embargo, es que el modelo de Bell (1984) estaba pensado desde una perspectiva

básicamente sincrónica³. ¿Cómo demostrar, entonces, que el diseño de la audiencia en los entrevistados ha cambiado con los años (específicamente, que una entrevista no es un evento comunicativo altamente formal en el año 2000, como sí lo era en 1970) y que los cambios en el comportamiento lingüístico comunitario son el resultado de dicho efecto⁴? Creo que es lícito suponer que el cambio en el diseño de la audiencia puede ser el resultado de un cambio de perspectiva más general. La pregunta que se hace aquí es: ¿acaso este cambio de perspectiva puede estar reflejando un cambio social más general? Esta es la ruta que se explora a continuación.

³ La propuesta de Bell (1984), que discutía problemas esenciales en sociolingüística, como la noción de estilos contextuales –basados en la atención prestada al habla, según la propuesta original de Labov (1966, 1972)—y su distinto estatus respecto a los factores sociales interindividuales como sexo, edad, etc., o los efectos de la acomodación lingüística (Douglas-Cowie 1978, Trudgill 1981, Coupland 1984), ha tenido un gran impacto en la investigación del estilo en sociolingüística y se ha explorado a lo largo de los años en ámbitos tan diversos como el lenguaje infantil (Youssef 1993), la entrevista judicial (Haworth 2013) o los monólogos realizados por *video bloggers* (o *vloggers*) (Frobenius 2014).

⁴ Aunque debe tomarse en cuenta que “las percepciones que intervienen en la dinámica de la entrevista pueden tener proyecciones de naturaleza lingüística; esto significa que existen rasgos del discurso hablado capaces de reflejar tales percepciones y que, en consecuencia, es posible categorizarlos y analizarlos” (Moreno Fernández 2012: 204, Escolio 10-B). No puedo ofrecer en este momento un análisis del discurso de las entrevistas de esta naturaleza, sin embargo, mi percepción general sobre la interacción en las entrevistas es que ésta no ha cambiado sustancialmente entre 1970 y 2000, sobre todo en aquellas entrevistas donde entrevistador y entrevistado no se conocían.

6.2.3. *La hipótesis histórica: los cambios lingüísticos documentados reflejan un cambio social más general*

La idea que se defendería desde esta perspectiva es que la comunidad lingüística de ECM, más nivelada, tiende a la *informalidad lingüística* como parte de un proceso social más general. Existen varios indicadores sociales, políticos y económicos que permiten confirmar un cambio social en la comunidad estudiada (el Área metropolitana de la ciudad de México --AMCM), que sería compatible con una comunidad lingüística relativamente más igualitaria (o *nivelada*).

El interesante estudio de Vázquez Laslop (2014), “El discurso político en México (1968-1994): la emergencia del diálogo”, en palabras de la autora, “busca demostrar [...] la participación cada vez más diversificada de la deliberación de los ciudadanos en las instituciones del estado mexicano” (2014: 1783). En su extenso ensayo se documenta y explica, a través del análisis del discurso político de tres grandes períodos (los sucesos de 1968, la reforma política de 1977 y el levantamiento del Ejército Zapatista de Liberación Nacional en 1994) cómo el estado mexicano tuvo que abrir las puertas al diálogo con fuerzas políticas y movimientos ciudadanos que tradicionalmente no tenían “ni voz ni voto” en el devenir político del país. En sus conclusiones, Vázquez Laslop destaca que “la emergencia del diálogo político deliberativo se observa más claramente en los momentos de las últimas cuatro décadas del siglo xx en las que se construyeron mecanismos de interlocución

innovadores, con la participación de los actores en conflicto y, en ocasiones, de otros grupos o individuos de diversos perfiles sociales” (2014: 1889).

En efecto, la evolución política del AMCM hacia formas más democráticas de acción inicia claramente con los movimientos sociales de 1968, según comenta Álvarez Enríquez (2005):

En muchos sentidos el movimiento estudiantil de 1968 representó un parteaguas tanto en la historia nacional como en la historia local recientes. Significó, por una parte, el inicio del debilitamiento de las bases de consenso del Estado fuerte emergido de la Revolución Mexicana [...]. Por otra parte, significó también el límite de la tolerancia ciudadana ante la política autoritaria del grupo en el poder, el desbordamiento de las vías tradicionales para la canalización de la demanda social y el inicio de un reclamo sistemático de democratización del sistema político mexicano [Álvarez Enríquez 2005: 140]

Producto directo de los conflictos de 1968 fue la reforma política del estado mexicano en forma de la *Ley federal de organizaciones políticas y procesos electorales (LFOPPE)*, publicada el 30 de diciembre de 1977 y que permitía a distintas agrupaciones políticas constituirse como partidos políticos y salir así de la clandestinidad (cf. Vázquez Laslop 2014: 1814). Esta tendencia a prácticas más

democráticas se consolida, “desde abajo”, con la fuerte politización y cambio en el tipo de participación ciudadana tras los terremotos de septiembre de 1985 (Álvarez Enríquez 2005: 205).

Incluso las *formas* de manifestación política se han emulado entre los distintos grupos sociales en la metrópoli. En el trabajo de S. Loaeza (1995), sobre una historia política de la ciudad de México en el siglo XX, se señala que:

Unos de los aspectos más importantes del desarrollo político del Distrito Federal, durante la década de los ochenta, se refiere a la *generalización de los comportamientos políticos, como si los capitalinos se imitaran unos a otros y recurrieran a formas de participación política que durante mucho tiempo les habían sido ajenas, porque estaban supuestamente identificadas con un determinado estatus social. En la última década [1985-1995], la población del Distrito Federal ha demostrado una increíble flexibilidad y capacidad de adaptación a las demandas del medio político [...]. Si las formas colectivas y no convencionales de participación estuvieron identificadas durante años con la acción comunitaria de los barrios pobres de la ciudad, su eficacia fue una invitación para que las colonias de clases media y alta, también recurrieran a plantones, bloqueos y marchas para hacerse escuchar por las autoridades [...]* [Loaeza 1995: 148, énfasis mío]

Esto es, si existe una tendencia a una mayor participación ciudadana en la vida política del país, a una mayor interacción entre todos los grupos sociales y a una asimilación de las formas de expresión ciudadana, ¿puede este comportamiento político compartido venir aparejado con un acercamiento sociolingüístico? Es posible.

Por otra parte, la economía mexicana, que se ha caracterizado tradicionalmente como generadora de desigualdad en el siglo xx (Tello 2010), en realidad es más equitativa en cuanto a la distribución del ingreso, importante índice económico. Esta es la tesis de Cortés (2013), quien hace una revisión crítica de varias fuentes para medir el ingreso en México en los últimos 50 años (como *la Encuesta Nacional de Ingresos y Gastos de los Hogares* [INEGI 1977, 1984]) y encuentra que México presenta fluctuaciones en la desigualdad que pueden verse en tres fases (o tal vez cuatro):

La primera se extendería desde los años sesenta hasta mediados de la década de los ochenta y se caracterizaría por una reducción [de la desigualdad en la distribución del ingreso] a ritmo lento. La primera medición de la segunda fase, que se extendería hasta el año 2000, se realizó en 1989 y estaría signada por un aumento marcado en los niveles de desigualdad [...]. La tercera iniciaría después del año 2000, con una reducción de la desigualdad en 2002, localizándose en un escalón inferior que en la segunda fase; no es muy claro si el dato de 2010 pertenece a la serie que inicia en 2002 o es el comienzo de una nueva etapa [...] [Cortés 2013: 25].

Por supuesto, la primera fase vino acompañada de una baja generalizada en los ingresos de toda la población (Cortés 2013: 17)⁵. En general, si la brecha económica entre ricos y pobres en realidad ha disminuido en los primeros 20 años del período analizado, ¿podría ser un factor que propicie la nivelación *social* en la capital del país? ¿Esta nivelación social podría asociarse también a una nivelación *lingüística*? Esta es otra posibilidad.

Otro índice puede ser la reivindicación de las manifestaciones culturales “populares” rescatadas –“desde arriba”, podría decirse-- en el movimiento literario de “la onda”, surgido precisamente en los años 1960. En obras como *De perfil* de José Agustín (1966), o *Pasto verde* (1968) y *El rey criollo* (1972) de Parménides García Saldaña, donde existe una reivindicación del habla juvenil, relativamente inédita respecto a la generación “de medio siglo” (la de Carlos Fuentes y José Emilio Pacheco, entre otros), se recuperaba las hablas cotidianas, sobre todo de los jóvenes: “[...] la

⁵ “La disminución de la desigualdad en 1984 con respecto a 1977 [...], desató una ríspida controversia [...] y fue un acicate para emprender investigaciones sobre el tema. Considerando que el país había atravesado por una fuerte recesión económica iniciada en el año 1982 se esperaba que la desigualdad hubiese aumentado entre esos dos años. Los estudios realizados mostraron, por primera vez, que la crisis fue seguida por fuertes reducciones en la participación relativa del décimo decil [...], que se complementó con los aumentos en las participaciones de los deciles inferiores (del primero al séptimo).

¶ Por otra parte, la disminución de la desigualdad fue acompañada por una fuerte retracción en los ingresos de todos los hogares del país (lo que llevó a acuñar el concepto “equidad por empobrecimiento” (Cortés y Rubalcava 1991), pero que fue mucho más marcada en los hogares de los deciles superiores” [Cortés 2013: 17].

obra de Parménides [García Saldaña] muestra una irreverencia desenfadada que criticó sin temor su entorno social. Así lo reflejaba en el lenguaje que utilizaba. Él mismo decía: ‘El lenguaje de los barrios bajos es escudo y puñal; afrenta, reto, desafío a las buenas costumbres y defensa de costumbres prohibidas. Diferencia de un mundo que vive en la aventura, y de otro que niega toda posibilidad de vivirla’” (Aguilar García (2013)⁶. Otro movimiento cultural que indicaba una visión fresca de lo popular es el de “Tepito Arte Acá”; también surgieron grupos musicales como *Botellita de Jerez*, quienes defienden que “lo naco es chido”⁷, frase de gran fuerza en los años 1980.

Los medios masivos, como la televisión abierta y el radio, también han recuperado, sobre todo a partir de la modificación en 2002 de la *Ley de Radio y Televisión*, las formas de hablar cotidianas, al grado de que los niveles de interdicción lingüística, inserción de coloquialismos y fraseologismos “populares”, con poca presencia en la programación anterior a 2002, reflejan (al menos al oído de este investigador), los niveles que uno encuentra en la vida cotidiana capitalina. La impresión general de personajes importantes en los medios masivos de México a finales del siglo xx, como Jacobo Zabłudowsky o Sergio Sarmiento, es de que el

⁶ Juan Carlos Aguilar García, “A 25 años de su muerte, Crónica evoca a Parménides García Saldaña, notable escritor, cronista y crítico de rock, a quien ‘se lo tragara la locura’”, la *Crónica de Hoy*, 10 de febrero de 2013. Consultado en: <http://www.cronica.com.mx/notas/2007/324134.html>.

⁷ En la serie *Tragicomedia Mexicana* de José Agustín (1990, 1992, 1998, México, Ed. Planeta), de hecho, se analizan estos movimientos culturales y sociales como síntomas de una relativa nivelación cultural de la sociedad mexicana.

lenguaje en los medios se ha “empobrecido” en los años recientes (v. el recuento de Ávila 2006: 118-121), aunque, como lo señaló Sarmiento en el *Primer Congreso Internacional de la Lengua Española en Zacatecas* (1997), este lenguaje “no es más pobre de lo que usualmente encontramos en las calles de nuestras ciudades [...] El bajo nivel educativo de nuestra sociedad se manifiesta en los medios” (*apud* Ávila 2006: 121). Al parecer, la locución como estilo de habla altamente regulado y estandarizado ha pasado a segundo término en la mayoría de las emisiones, al menos en las de alta audiencia, con excepción eso sí, de los noticieros, que conservan la tendencia a la unidad de la lengua (Ávila 2006), aunque no faltan, sobre todo en los últimos años, alternativas como *El Mañanero*, noticiero de radio y TV conducido por el comediante Víctor Trujillo, quien representa a un payaso (*Brozo, El Payaso Tenebroso*), que posee un estilo de habla desenfadado, o *El Almohadazo*, noticiero nocturno conducido por Fernanda Tapia, también con un fuerte componente cómico. Por supuesto, se trata de desarrollos mediáticos posteriores a los años investigados, pero que sin duda pueden verse como consecuencias de un proceso de informalidad que se inició al menos a finales del período analizado⁸. Aquí sí puede hablarse de un

⁸ Para una visión de los medios como impulsores y a la vez obstaculizadores de la información y la cultura democrática en México, v. Sánchez Ruiz (2005). Otro índice no menos importante que ilustra esta tendencia a la informalidad se encuentra también en la comunicación de, por ejemplo, importantes líderes políticos: simplemente compárese la oratoria tradicional de Luis Echeverría (1970-1976) con la de Vicente Fox en su campaña política que lo llevó a la presidencia de la república en el año 2000, de carácter muy informal.

cambio en el diseño de la audiencia, algo que Bell ha investigado en los medios masivos (Bell 1994), pero el punto es: ¿qué está detrás de este cambio?

En suma, en realidad existen evidencias históricas, políticas, económicas y sociológicas sobre una transformación social en la ciudad de México que puede estar aparejada con un cambio lingüístico comunitario. Esta hipótesis es compatible con los recientes estudios del cambio y su relación con factores sociales y cognitivos. Los estudios sociolingüísticos de cambio en tiempo real, que tienen su inicio más genuino en la importante investigación de Labov en Martha's Vineyard (1963), suelen destacar el peso de los factores sociales --en ese caso, los efectos de la migración y el cambio en la actitud de los jóvenes por preservar rasgos estereotípicos del dialecto originario (la centralización de [ay] y [aw]). Por mencionar sólo algunos ejemplos recientes, en el trabajo de Poplack & St. Amand (2007) en francés de Quebec se documenta la importancia de las mujeres en el cambio socioestilístico del morfema de negación *ne* (extremadamente raro ya en el siglo XIX):

Our comparison of *ne* usage in 19th- and 20th-century Québec French leads us to conclude that there has in fact been a recent change in *ne* usage, but, contrary to claims based on written representations of speech, it does not involve a decrease in rate. Rather, the change is FUNCTIONAL: *ne* has become a sociostylistic marker of formality. Interestingly, this change seems to have been driven by women, traditionally the exponents of standard speech [Poplack & St. Amand 2007: 726; énfasis en el original].

En Nevalainen, Raumolin-Brunberg & Mannila (2011) se documentan seis cambios lingüísticos en inglés (entre los siglos xv y xvii) y se discute el papel de hablantes conservadores e innovadores; las redes sociales a las que pertenecen los hablantes juegan un papel importante en los patrones de difusión. Incluso en estudios con poblaciones muy homogéneas, como en el de Wagner (2012) sobre la variable (-*ing*) en un panel de adolescentes de Filadelfia, se encuentra que:

But even in this relatively homogenous sample, the retreat from [ɪn] over time was not evident for all speakers. A minority exhibited little to no change in the direction of the standard norm, and this was especially the case for speakers who were affiliated with Second Street, an iconic Irish neighborhood [...]. A second minority—those speakers who enrolled in nationally oriented colleges—exhibited a significant decline in [ɪn] use after high school, when compared with peers at regionally oriented colleges [Wagner 2012: 197].

La importancia del contacto dialectal y el *catastrofismo* (fenómenos sociales como migraciones masivas o guerras de conquista, como ocurrió en el México del siglo xv y colonial) como motores del cambio han sido señalados por Martín Butragueño (2004 y 2010, respectivamente). Por su parte Kerswill habla sobre el contacto como la principal fuente de cambios (Kerswill 1996); finalmente, los mencionados estudios de cambio en tiempo real en diferentes lenguas confirman el peso de factores extralingüísticos en los cambios estudiados. Como parte de dicha argumentación

basada en el peso de factores externos, en la siguiente sección se revisa la posibilidad de que los fenómenos de nivelación documentados en ECM estén vinculados con movimientos migratorios dentro de la gran AMCM.

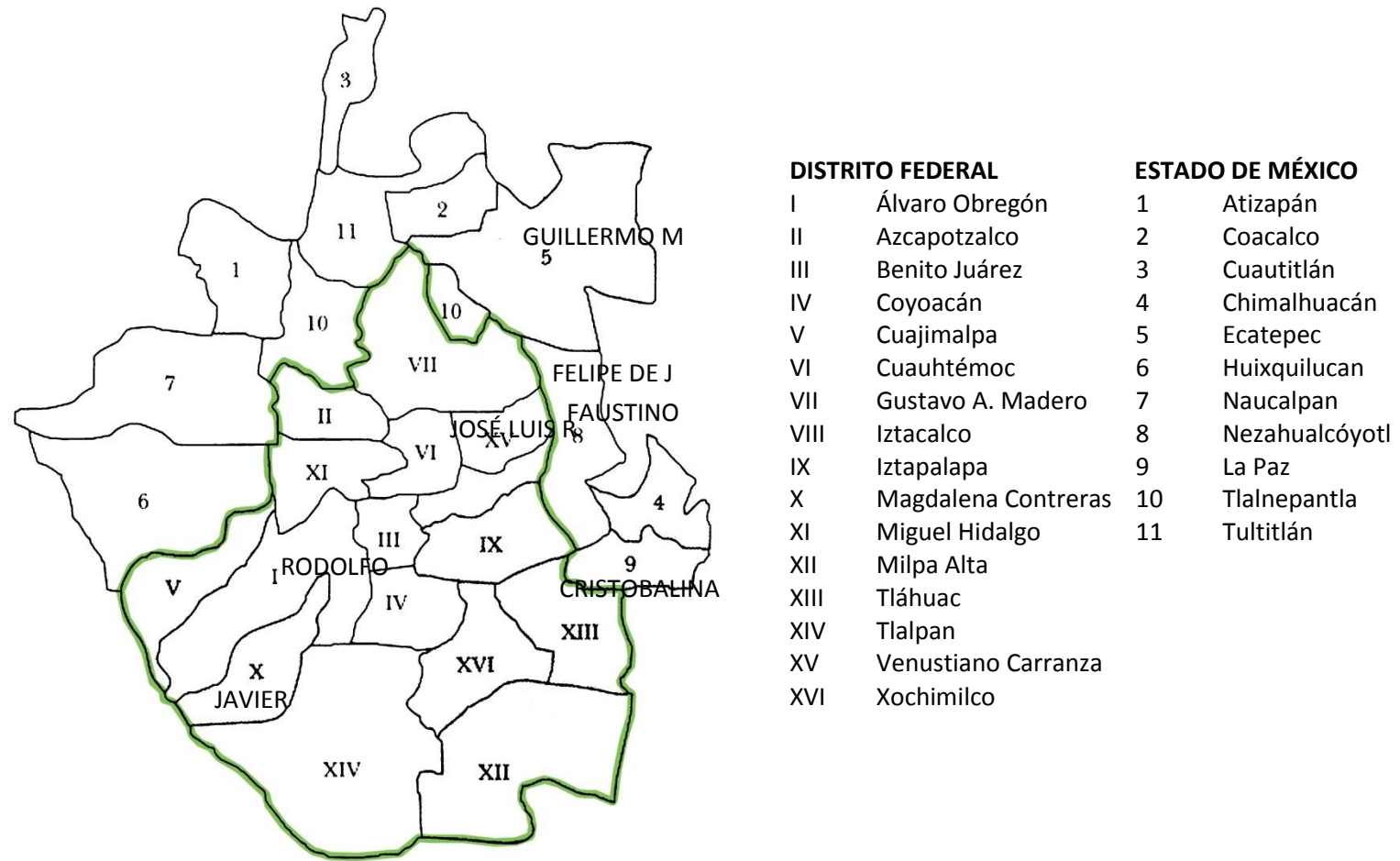
6.3. ENSAYO DE GEOLINGÜÍSTICA URBANA: UBICACIÓN DE LOS LÍDERES DE LA VARIACIÓN

Al investigar a los líderes de los procesos variables investigados a nivel fonético y léxico, un fenómeno que no se había pronosticado es que la distribución espacial de los rasgos lingüísticos mostraría un patrón relativamente regular: en este caso, una interesante oposición entre el centro y la periferia del AMCM, particularmente en los casos de elisión vocálica y riqueza léxica (RTT). Aunque estos resultados requieren una revisión a fondo y son todavía muy preliminares, vale la pena mostrar lo que parecen ser indicios de una distribución geolingüística particular de los procesos lingüísticos involucrados.

El patrón geolingüístico muestra que los hablantes que patrocina la elisión vocálica y un menor índice de RTT (ambos rasgos asociados a los grupos con menos estudios) se localizan en la periferia de la metrópoli, mientras que sus contrapartes (las personas bajos índices de elisión vocálica y con mayor índice de RTT) se encuentran ubicadas en las zonas centrales.

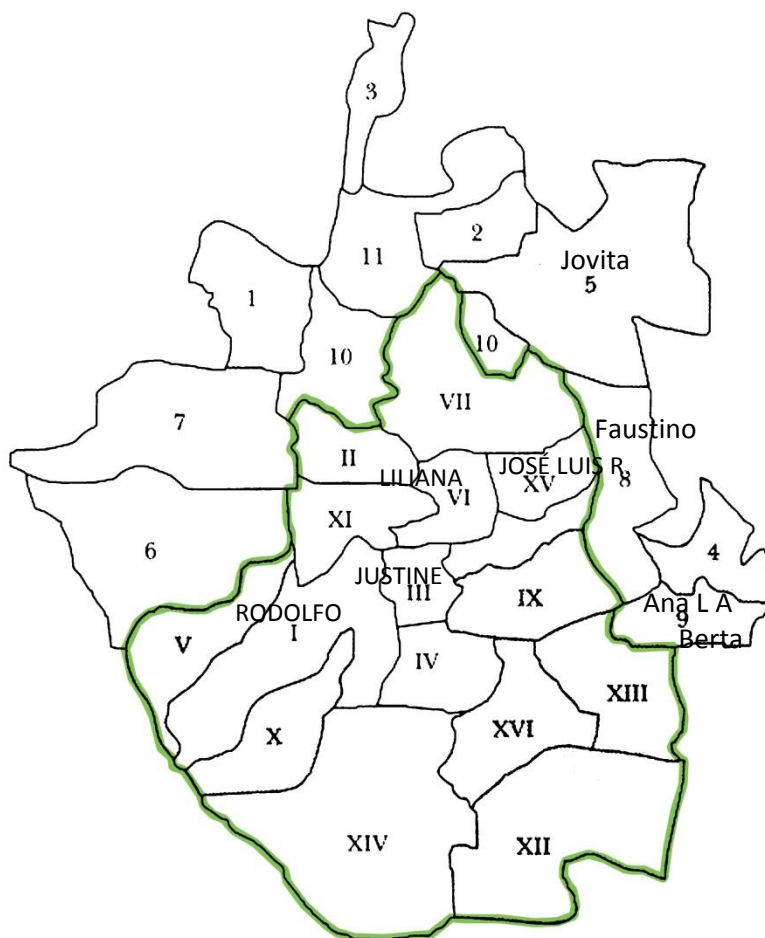
En el siguiente mapa se ubica geográficamente a los hablantes que concentran más del 50% de los casos totales de elisión en el AMCM⁹. Estos mismos hablantes son los punteros en la gráfica de Frecuencia relativa acumulativa de la elisión (cap. 4, Fig. 4.12):

⁹ Ese mapa no muestra toda el AMCM, sino la llamada “zona pertinente” para levantar el CSCM (Lastra & Martín Butragueño 2000). Agradezco a Itzel Nájera su apoyo para la elaboración de estos mapas.



Mapa 1. Distribución espacial de los líderes de la elisión vocálica en el AMCM, año 2000

El mapa muestra que los hablantes que más eliden vocales en el año 2000 tienen su domicilio en zonas periféricas o relativamente alejadas del centro de la ciudad: Ecatepec, Nezahualcóyotl, La Paz o Magdalena Contreras. Por otra parte, el léxico ofrece un panorama similar y complementario. El siguiente mapa muestra la distribución de los hablantes con mayor y menor índices de riqueza léxico-morfológica (RTT):



Mapa 2. Distribución espacial de los hablantes con MAYOR (en mayúsculas) y menor (en minúsculas) riqueza léxico-morfológica en el AMCM, año 2000

La distribución de la riqueza léxico-morfológica confirma el patrón de la elisión vocálica. Es muy importante recordar que estos resultados se basan en los recorridos de regresión escalonada de las vocales (con buena significatividad), en la gráfica de frecuencias acumulativas y en los índices de Relación Type Token; esto es, no se trata sólo de frecuencias absolutas o relativas individuales: estos hablantes líderes emergen del análisis de la muestra total de 48 informantes.

Los resultados permiten sugerir que algunos rasgos de las hablas “populares” de los años 1960-1970 se han desplazado a las regiones periféricas de la metrópoli. Por otra parte, hay que considerar que dichas áreas se habitaron con una población mixta, de habitantes de los estados circunvecinos de pocos recursos (Martín Butragueño & Lastra 2011) y población originaria de la zona central del AMCM, desplazada por razones económicas y de vivienda (como los terremotos de 1985); esto vuelve plausible que dichos inmigrantes hubieran formado una koiné dialectal (Berruto 2010)¹⁰, en este caso particular, de base sub-estándar (Carriscondo 2001). En parte, este proceso permitiría entender el porqué de esta tendencia a la nivelación dialectal en la metrópoli.

6.4. PROCESOS SOCIOLINGÜÍSTICOS EN LA CIUDAD DE MÉXICO

Los resultados del presente estudio en tiempo aparente y tiempo real permiten afirmar que nos encontramos con una comunidad lingüística de la ciudad de México y

¹⁰ “A *koiné*’, from Greek *κοινή* ‘common (language)’, is “the stabilized result of mixing of linguistic subsystems such as regional or literary dialects” (Siegel 1985: 363, *apud* Berruto 2010: 232).

su área metropolitana que tiende a la recuperación del estándar de pronunciación y que, a la vez, reduce la brecha sociolingüística entre los grupos extremos de esta sociedad capitalina en términos fónicos, morfológicos y léxicos. Evidencia de esto es la gran regularidad que se presenta en casi todos los fenómenos analizados, en particular, el retroceso estadístico de las formas fonéticas muy desviadas del estándar de pronunciación (elisión vocálica y róticas asibiladas), la disminución con el tiempo de la riqueza léxica individual (*relación type/token*), y el descenso comunitario en el uso de grupos consonánticos cultos, de adverbios terminados en *-mente* y de sufijos abstractos como *-dad* (cap. 5), todos rasgos asociados a las variedades cultas o estándar.

En general, los resultados del análisis de las variables fonéticas confirman los hallazgos en otros estudios sobre la preferencia de las mujeres por las variantes prestigiosas --la asibilación de róticas en los años sesenta-setenta-- y el evitar las variantes innovadoras no prestigiosas como la elisión vocálica¹¹.

En general, los resultados de la presente tesis pueden verse como argumentos a favor de una perspectiva que, como señalara Meillet (1921), debía estudiar qué cambios sociales están vinculados con qué cambios lingüísticos. Quedan varios procesos por analizar. Particularmente fenómenos fonéticos, léxicos y sintácticos de los últimos 10 años, muchos de los cuales podrían verse como productos del contacto

¹¹ Queda pendiente el análisis detenido de la riqueza léxica de mujeres y hombres. Los primeros resultados indican un patrón en el que las mujeres presentan una menor riqueza léxico-morfológica que los hombres. Debe profundizarse más en el análisis para poder asegurar que existe una diferencia sexual correlacionada significativamente con el volumen léxico.

con el inglés: pronunciaciones de (r) como aproximante con cierto grado de retroflexión, una incorporación ingente de anglicismos vinculados a la tecnología de las telecomunicaciones e internet (*accesar*, *u-ese-bé* (USB), *no-break*); innovaciones fraseológicas y léxicas, a partir de anglicismos, que pueden atestiguar en habla de adolescentes, como *LOL* (de *lot of laugh*, ‘muchacha risa’); *YOLO* (acrónimo de *you only live once*, ‘sólo se vive una vez’), *parijardear* (de *party hard*, ‘estar de fiesta’); etc.; falsos amigos en ciertos estilos periodísticos (*facilidades* [< ingl. *facilities*] por ‘instalaciones’); incorporación del sentido de agentividad en verbos como *crecer* –y su conjugación correspondiente--: *crece tu empresa* por ‘haz crecer tu empresa’, entre otros procesos. La futura exploración de fenómenos de este tipo permitirá confirmar, matizar o poner en tela de juicio los hallazgos, métodos y perspectivas analíticas de este trabajo de investigación.

BIBLIOGRAFÍA

Aguirre, Nazyheli & Gustavo Chico (2011). "Perfil socioeducativo de los hablantes entrevistados en los proyectos Norma culta del español de la ciudad de México (1967-1971) y El habla popular de la ciudad de México (1972-1974)", en: <http://www.iifilologicas.unam.mx/elhablamexico/uploads/Aguirre%20y%20Chico%20-%20Perfil%20socioeducativo%20NC%20y%20HP.pdf>, consultado en marzo de 2014.

Aitchison, Jean (2001). *Language Change: Progress or Decay?* Cambridge: Cambridge University Press.

Álvarez Enríquez, Lucía (2005). *Distrito Federal: sociedad, economía, política y cultura*. 2da. ed. México: Centro de Investigaciones Interdisciplinarias en Ciencias y Humanidades, Universidad Nacional Autónoma de México. [1ra ed.: 1998].

- Anttila, Arto (2002). "Variation and Phonological Theory", en *The Handbook of Language Variation and Change*. Jack Chambers, Peter Trudgill & Natalie Schilling-Estes (eds.). Oxford: Blackwell, pp. 206-243.
- Anttila, Arto & Young-Mee Yu-Cho (1998). "Variation and change in Optimality Theory", *Lingua* 194, 31-56.
- Arellano, Fernando S.J. (1977). *Historia de la lingüística. Tomo I. Desde sus orígenes hasta el siglo XIX inclusive*. Caracas, Venezuela: Universidad Católica Andrés Bello.
- Arias, Beatriz (2008). "Grupos consonánticos cultos en el siglo XVI y XVII: ¿representación fonémica o norma gráfica?", *Lexis* XXXII (1), 5-26.
- Arias Leal, Ileana (2013). *Adverbios terminados en -mente en posición inicial o parentética en el español actual*. Tesis de Maestría. UNAM.
- Austin, Peter K. (2007). "Los datos y la documentación lingüística", en *Bases de la documentación lingüística*. John Haviland & R. Flores Farfán (coords.). México: Instituto Nacional de Lenguas Indígenas, pp. 111-140.
- Ávila, Raúl (1988). "Lengua hablada y estrato social: un acercamiento lexicoestadístico", *Nueva Revista de Filología Hispánica* XXXVI (1), 131-148.
- Ávila, Raúl (1999 a). "Sobre semántica social: conceptos y estratos en el español de México", en *Estudios de semántica social*. México: El Colegio de México, pp. 59-104.

- Ávila, Raúl (1999 b). "Sociosemántica: referentes sustantivos y verbales en el habla culta y popular de la ciudad de México", en *Estudios de semántica social*. México: El Colegio de México, pp. 153-219.
- Ávila, Raúl (ed.) (2011). *Variación del español en los medios*. México: El Colegio de México.
- Ávila Muñoz, A. M. & J. A. Villena Ponsoda (eds.) (2010). *Variación social del léxico disponible en la ciudad de Málaga. Diccionario y análisis*. Málaga: Sarriá.
- Baayen, R. H. (2008). *Analysing Linguistic Data. A Practical Introduction to Statistics*. Cambridge: Cambridge University Press.
- Bailey, Guy (2002). "Real and apparent time", en *The Handbook of Language Variation and Change*. J. K. Chambers, P. Trudgill & N. Schilling-Estes (eds.). Oxford: Blackwell, pp. 312-332.
- Barriga Villanueva, Rebeca (coord.) (2013). *Lingmex: Bibliografía Lingüística de México desde 1970*. 15a. ed. Hermelinda Mendoza Filio (colab.). México: El Colegio de México. Disponible en: <http://lingmex.colmex.mx/>.
- Barriga Villanueva, Rebeca & Claudia Parodi (1996). *La lingüística en México, 1980-1996*. México: El Colegio de México, Universidad de California, Los Ángeles [614 pp. + 1 disco compacto].
- Barriga Villanueva, Rebeca & P. Martín Butragueño (dirs.) (2010 a). *Historia Sociolingüística de México. Vol. 1: México prehispánico y colonial*. México: El Colegio de México.

- Barriga Villanueva, Rebeca & P. Martín Butragueño (dirs.) (2010 b). *Historia Sociolingüística de México. Vol. 2: México contemporáneo*. México: El Colegio de México.
- Barriga Villanueva, Rebeca & P. Martín Butragueño (dirs.) (2014). *Historia Sociolingüística de México. Vol. 3: Espacio, contacto y discurso político*. México: El Colegio de México.
- Bell, Allan (1984). "Language style as audience design", *Language in Society* 13, 145-204.
- Bell, Allan (1994). *The Language of News Media*. Cambridge Mass.: Blackwell.
- Bernstein, Basil (1984). "Códigos amplios y restringidos: sus orígenes sociales y algunas consecuencias", en *Antología de estudios de etnolingüística y sociolingüística*. Paul Garvin & Y. Lastra de Suárez (eds.). México: UNAM, pp. 357-374.
- Berruto, Gaetano (2010). "Identifying dimensions of linguistic variation in a language space" en *Language and Space. An International Handbook of Linguistic Variation. Theories and Methods*. Peter Auer & Jürgen Erich Schmidt (eds.). Germany: Walter de Gruyter Mouton, pp. 226-241.
- Bloomfield, Leonard (1933). *Language*. Chicago: Chicago University Press.
- Boersma, Paul & Bruce Hayes (2001). "Empirical test of the Gradual Learning Algorithm", *Linguistic Inquiry* 32, 45-86.
- Boersma, Paul & David Weenink (2014). *PRAAT. Doing Phonetics by Computer*. Versión 5.3.78. Universidad de Ámsterdam. <http://www.fon.hum.uva.nl/praat/>.

- Bourdieu, Pierre (1984). "Capital et marché linguistique", *Linguisticshe Berichte* 90, 3-24.
- Bourdieu, Pierre (1990). "El mercado lingüístico", en *Sociología y cultura*. M. Pou (trad.). México: Grijalbo, pp. 143-158.
- Boyd-Bowman, Peter (1952). "La pérdida de las vocales átonas en la altiplanicie mexicana", *Nueva Revista de Filología Hispánica* VI, 138-140.
- Bradley, Travis G. (2000). "A typology of rothic duration and neutralization [Version 1]". Versión preliminar de la publicada en *Proceedings of the North East Linguistic Society* 31, en: <http://roa.rutgers.edu/files/436-0601/436-0601-BRADLEY-0-0.PDF>, 08 de octubre de 2012.
- Bradley, Travis G. (2001). *The Phonetics and Phonology of Rhotic Duration Contrast and Neutralization*. Tesis Doctoral. The Pennsylvania State University.
- Briggs, Charles (1986). *Learning How to Ask. A Sociolinguistic Appraisal of the Role of the Interview in Social Science Research*. Cambridge: Cambridge University Press.
- Bright, William (ed.) (1966). *Sociolinguistics: Proceedings of the UCLA Sociolinguistics Conference 1964*. La Haya, París: Mouton & Co.
- Browman, Catherine & Louis Goldstein (1989). "Articulatory gestures as phonological units", *Phonology* 6, 201-252.
- Browman, Catherine & Louis Goldstein (1990). "Gestural specification using dynamically-defined articulatory structures", *Journal of Phonetics* 18, 299-320.

- Bybee, Joan (2000). "La difusión léxica y el modelo de ejemplares". Conferencia magistral presentada en el *VI Encuentro Internacional de Lingüística en el Noroeste*, Hermosillo, Sonora (ejemplario).
- Bybee, Joan & Paul Hopper (eds.) (2001). *Frequency and the Emergence of Linguistic Structure*. Ámsterdam: John Benjamins.
- Canale, Michael & Merrill Swain (1980). "Theoretical bases of communicative approaches to second language teaching and testing", *Applied Linguistics* 1, 1-47.
- Canellada, María Josefa & Alonso Zamora Vicente (1960). "Vocales caducas en el español mexicano", *Nueva Revista de Filología Hispánica* XIV, 221-241.
- Cárdenas, Daniel N. (1967). *El español de Jalisco*. Madrid: CSIC.
- Carriscondo Esquivel, Francisco M. (2001). *La lexicografía en las variedades no-estándar*. Jaén: Publicaciones de la Universidad de Jaén.
- Castañeda, Gonzalo (2010). *Introducción a la sociomática. El estudio de los sistemas adaptables complejos en el entorno socioeconómico*. Ms. El Colegio de México.
- Catford, J. C. (1977). *Fundamental Problems in Phonetics*. Bloomington: Indiana University Press.
- Ceballos Domínguez, Rubí (2006). "Hacia un mayor consonantismo en la zona conurbada", en *Líderes lingüísticos. Estudios de variación y cambio*. Pedro Martín B. (ed.). México: El Colegio de México, pp. 13-36.

- Cedergren, Henrietta J. (1984). "Panama revisited: Sound change in real time". Ponencia presentada en *New Ways of Analysing Language Variation (NWAV)* 13. Filadelfia.
- Cedergren, Henrietta J. & David Sankoff (1974). "Variable rules: performance as a statistical reflection of competence", *Language* 50, 333-355.
- Chambers, Jack (1992). "Dialect acquisition", *Language* 68 (4), 673-705.
- Chambers, Jack (1995). *Sociolinguistic Theory*. Oxford: Wiley-Blackwell.
- Chambers, Jack & Peter Trudgill (1998). *Dialectology. Second Edition*. Cambridge: Cambridge University Press. [Versión española de 1994: *La dialectología*. Carmen Morán González (trad.), adaptación y anotación de Eugenio Bustos Gisbert. Getafe, Madrid: Visor Libros].
- Chomsky, Noam (1957). *Syntactic Structures*. La Haya: Mouton.
- Chomsky, Noam (1965). *Aspects of the Theory of Syntax*. Cambridge, Mass.: MIT Press.
- Chomsky, Noam (1966). *Cartesian Linguistics. A Chapter in the History of Rationalist Thought*. Nueva York: Harper & Row.
- Chomsky, Noam (1975 [1955]). *The Logical Structure of Linguistic Theory*. Nueva York: Plenum Press.
- Chomsky, Noam (1981). *Lectures on Government and Binding*. Dordrecht: Foris.
- Chomsky, Noam (2007). "Of Minds and Language", *Biolinguistics* 1, 9-27.
- Chomsky, Noam & Morris Halle (1968). *The Sound Pattern of English*. Nueva York: Harper & Row.

Coetzee, Andries W. (2008). "Phonological Variation and Lexical Frequency", en: <http://roa.rutgers.edu/article/view/982>. [Consultado el 3 de octubre de 2012].

Coetzee, Andries W. (2009). "Learning lexical indexation", *Phonology* 26 (1), 109-145.

Coetzee, Andries W. & Joe Pater (2008). "Weighted constraints and gradient restrictions on place co-occurrence in Muna and Arabic", *Natural Language and Linguistic Theory* 26, 289-337.

Coetzee, Andries W. & Shigeto Kawahara (2013). "Frequency biases in phonological variation", *Natural Language and Linguistic Theory* 31, 47-89.

Company Company, Concepción (2012 a). "Condicionamientos textuales en la evolución de los adverbios en *-mente*", *Revista de Filología Española* XCII (1), 9-42.

Company Company, Concepción (2012 b). "La <<engañosa>> productividad de los adverbios en *-mente* de la lengua española", en *En pos de la palabra viva: huellas de la oralidad en textos antiguos. Estudios en honor al profesor Rolf Eberenz*. V. Béguelin-Argimón, G. Cordone & M. de la Torre (eds). Berlín, Nueva York: Peter Lang, pp. 119-136.

Company Company, Concepción (en prensa). "Los adverbios en *-mente*", en *Sintaxis histórica de la lengua española. Tercera parte: Adverbios, preposiciones y conjunciones. Relaciones interoracionales*. México: Fondo de Cultura Económica, UNAM.

CONACULTA (2006). *Encuesta nacional de lectura*. México: Consejo Nacional para la Cultura y las Artes, en: http://sic.conaculta.gob.mx/publicaciones_sic.php?&l=&estado_id. [Octubre de 2008].

Conde Silvestre, Juan Camilo (2007). *Sociolingüística histórica*. Madrid: Gredos.

Cortés, Fernando & Rosa María Rubalcava (1991). *Autoexplotación forzada y equidad por empobrecimiento: la distribución del ingreso familiar en México, 1977-1984*. México: El Colegio de México.

Cortés, Fernando (2013). "Medio siglo de desigualdad en el ingreso en México", *Economía UNAM* 10 (29), 12-34.

Coseriu, Eugenio (1958). *Sincronía, diacronía e historia. El problema del cambio lingüístico*. Montevideo: Universidad de la República.

Coupland, Nikolas (1984). "Accommodation at work: Some phonological data and their implications", *International Journal of the Sociology of Language* 46, 49-70.

Croft, William (2000). *Explaining Language Change. An Evolutionary Approach*. Harlow: Longman.

Crowley, Terry (1997). *An Introduction to Historical Linguistics*. Oxford: Oxford University Press.

Cruz, Miroslava & Julio Serrano (en prensa a). "Elementos alfabéticos en la Lengua de Señas Mexicana: un acercamiento sociolingüístico", en *II Coloquio de Cambio y*

- Variación Lingüística*. Pedro Martín Butragueño & Leonor Orozco Vaca (eds.). México: El Colegio de México.
- Cruz, Miroslava & Julio Serrano (en prensa b). “El peso de la historia en la conformación de la estructura sociolingüística: la comunidad sorda en México”, en *Estudios de variación geolingüística*. Leonor Orozco & Alonso Guerrero (eds.). México: Instituto Nacional de Antropología e Historia.
- Cruz Gómez, Sandra Rocío (2008). “Variación estilística del tseltal de Oxchuc. El caso de la circunclisis determinante”. Ponencia presentada en el *X Encuentro Internacional de Lingüística en el Noroeste*. Hermosillo, Sonora, 12-15 de noviembre.
- Cruz Gómez, Sandra Rocío (2014). “Panorama sociolingüístico de las lenguas indígenas del Chiapas actual”, en *Historia Sociolingüística de México. Vol. 3: Espacio, contacto y discurso político*. Rebeca Barriga Villanueva & Pedro Martín Butragueño (dirs.). México: El Colegio de México, pp. 1721-1755.
- Cukor-Avila, Patricia (2000). *The Stability of Individual Vernaculars*. MS. University of North Texas.
- De Fina, Anna & Sabina Perrino (2011). “Introduction: Interviews vs. ‘natural’ contexts: A false dilemma”, *Language in Society* 40 (1), 1-11.
- Delforge, Anne Marie (2008). “Unstressed vowel reduction in Andean Spanish”, en *Selected Proceedings of the 3rd Conference on Laboratory Approaches to Spanish Phonology*. Laura Colantoni & Jeffrey Steele (eds.). Somerville, MA: Cascadilla Proceedings Project, pp. 107-124.

- Díaz-Campos, Manuel & Sonia Colina (2006). "The interaction between faithfulness constraints and sociolinguistic variation", en *Optimality-Theoretic Studies in Spanish Phonology*. Fernando Martínez-Gil & Sonia Colina (eds.). Amsterdam, Filadelfia: John Benjamins, pp. 424-446.
- Díaz Rojo, José Antonio (2002). "Elementos de origen grecolatino en la terminología médica (VI): Grupos consonánticos cultos", *El Trujamán. Revista Diaria de Traducción*. 19 de febrero de 2002, consultada en: http://cvc.cervantes.es/trujaman/anteriores/febrero_02/19022002.htm, marzo de 2014.
- Dittmar, Norbert (1996). "Descriptive and explanatory power of rules in sociolinguistics", en *Towards a Critical Sociolinguistics*. Rajendra Singh (ed.). Amsterdam: John Benjamins, pp. 115-149.
- Donni de Mirande, Nélica Esther (1974). "Grupos consonánticos en el español de Rosario", *Thesaurus* XXXIX (3), 526-538.
- Douglas-Cowie, E. (1978). "Linguistic code-switching in a Northern Irish village: Social interaction and social ambition", en *Sociolinguistic Patterns in British English*. Peter Trudgill (ed.). Londres: Edward Arnold, pp. 37-51.
- Eckert, Penelope (1989). *Jocks and Burnouts: Social Categories and Identities in the High School*. Nueva York: Teachers College Press.
- Eller, Wendianne Alice (2008). "Sociolingüística de los distintos usos del vocablo 'güey' en la ciudad de México". Ponencia presentada en el X Encuentro

- Internacional de Lingüística en el Noroeste*, Hermosillo, Sonora, 12-15 de noviembre.
- Eller, Wendianne Alice (2013). *Sociolingüística del español gay mexicano. Variación fónica, estereotipos, creencias y actitudes en una red social de hombres homosexuales*. Colección Posgrado. México: UNAM.
- Espinosa Vázquez, Alejandra Aide (2008). "Variación del segmento /-s/ en El Ciruelo, Oaxaca", en *Fonología instrumental. Patrones fónicos y variación*. Esther Herrera Zendejas & Pedro Martín Butragueño (eds.). México: El Colegio de México, pp. 107-128.
- Evans, Nicholas & Stephen C. Levinson (2009). "The myth of language universals: Linguistic diversity and its importance for cognitive science", *Behavioral and Brain Sciences* 32, 429-492.
- Farnetari, Edda & Daniel Recasens (1999). "Coarticulation models in recent speech production theories", en *Coarticulation*. William J. Hardcastle & Nigel Hewlett (eds.). Cambridge: Cambridge University, pp. 31-65.
- Fasold, Ralph (1991 a). *The Sociolinguistics of Language: Introduction to Sociolinguistics*. Oxford: Wiley-Blackwell.
- Fasold, Ralph (1991 b). *The Sociolinguistics of Society*. Oxford: Wiley-Blackwell.
- Fasold, Ralph (1991 c). "The quiet demise of variable rules", *American Speech* 66, 3-21.
- Feldhausen, Ingo & María del Mar Vanrell (2014). "Prosody, focus and Word order in Catalan and Spanish: An Optimality Theoretic Approach", en *Proceedings of the 10th International Seminar on Speech Production (ISSP)*. Colonia, Alemania, 5-8

de mayo de 2014, consultado en: <http://www.academia.edu/6626178/>, consultado el 9 de septiembre de 2014.

Figuerola, Esther (1994). *Sociolinguistic Metatheory*. Oxford: Pergamon.

Flemming, Edward (1995). *Auditory Representations in Phonology*. Tesis doctoral. University of California, Los Angeles.

Fortson IV, Benjamin W. (2003). "An approach to semantic change", en *The Handbook of Historical Linguistics*. Brian D. Joseph & Richard D. Janda (eds.). Oxford: Blackwell, pp. 648-666.

Fowler, Joy (1986). *The Social Stratification of (r) in New York City Department Stores, 24 Years after Labov*. MS. New York University.

Frobenius, Maximiliane (2014). "Audience design in monologues: How vloggers involve their viewers", *Journal of Pragmatics* (en prensa).

Gil Burgoin, Carlos Ivanhoe (2011). *El español del municipio de La Paz, Baja California Sur: Variación fónica y entonativa*. Tesis de Licenciatura. México: Escuela Nacional de Antropología e Historia.

González Cáceres, Susana (2012). *La "lengua tecleada": ¿Se observan marcas de su uso en escritos escolares de estudiantes de los niveles de bachillerato y licenciatura? Una aproximación desde la sociolingüística variacionista*. Tesis de maestría. Universidad Nacional Autónoma de México.

Greenberg, Joseph H. (1966). *Language Universals*. La Haya: Mouton de Gruyter.

- Guerrero, Alonso (2006). "Hablamos(s) así todo(s) iguale(s): concordancia plural en un contexto bilingüe", en *Líderes lingüísticos. Estudios de variación y cambio*. Pedro Martín B. (ed.). México: El Colegio de México, pp. 89-110.
- Guerrero, Alonso & Leonor Orozco (eds.; en preparación). *Estudios de variación geolingüística*. México: Instituto Nacional de Antropología e Historia, Escuela Nacional de Antropología e Historia.
- Guitart, Jorge M. (1997). "Variability, multilectalism, and the organization of phonology in Caribbean Spanish dialects", en *Issues in the Phonology and the Morphology of the Major Iberian Languages*. Fernando Martínez Gil & Alfonso Morales Front (eds.). Washington: Georgetown University Press, pp. 515-536.
- Gumperz, John J. & Robert Wilson (2000). "Convergencia y criollización. Un caso de la frontera entre lenguas indoarias y dravídicas en India", en *Estudios de sociolingüística*. Yolanda Lastra (comp.). México: UNAM, pp. 217-238.
- Guy, Gregory R. (1997). "Violable is variable: Optimality Theory and linguistic variation", *Language Variation and Change* 9, 333-347.
- Guy, Gregory R. (2007). "Variation and phonological theory", en *Sociolinguistic Variation. Theories, Methods, and Applications*. Robert Bailey & Ceil Lucas (eds.). Cambridge: Cambridge University Press; pp. 5-23.
- Guy, Gregory R. & Sally Boyd (1990). "The development of a morphological class", *Language Variation and Change* 2, 1-18.
- Hall, T. A. (2007). "Segmental features", en *The Cambridge Handbook of Phonology*. Paul de Lacy (ed.). Cambridge: Cambridge University Press, pp. 311-334

- Halliday, M. A. K. (1978). *Language as Social Semiotics*. Londres: Edward Arnold.
- Haspelmath, Martin (2007). "Pre-established categories don't exist: Consequences for language description and typology", *Linguistic Typology* 11, 119-132.
- Haworth, Kate (2013). "Audience design in the police interview: The interactional and judicial consequences of audience orientation", *Language in Society* 42, 45-69.
- Henríquez Ureña, Pedro (1921 [1976]). "Observaciones sobre el español de América (I)", en *Observaciones sobre el español en América y otros estudios filológicos*. Juan Carlos Ghiano (ed.). Buenos Aires: Academia Argentina de Letras, pp. 1-44.
- Henríquez Ureña, Pedro (1938). "Mutaciones articulatorias en el habla popular", en *El español en México, los Estados Unidos y la América Central*. Biblioteca de Dialectología Hispanoamericana 4. Buenos Aires: Universidad de Buenos Aires; pp. 329-279.
- Hernández, Esther & Pedro Martín Butragueño (coords.) (2013). "Qué es la variación lingüística y por qué puede interesar su estudio desde distintas perspectivas", con: Rebeca Barriga Villanueva; Violeta Demonte; Pilar García Mouton; Rodrigo Gutiérrez-Bravo; Esther Herrera Zendejas; Isabel Pérez Jiménez & Julia Pozas Loyo (preprint), en: <https://www.academia.edu/4024366>, abril de 2014.
- Hock, Hans Henrich & Brian D. Joseph (1996). *Language History, Language Change, and Language Relationship: An Introduction to Historical and Comparative Linguistics*. Berlin: Mouton de Gruyter.
- Hockett, Charles F. (1958). *A Course in Modern Linguistics*. Nueva York: The Macmillan Company.

- Højrup, Thomas (1983). "The concept of life-mode. A form-specified mode of analysis applied to contemporary western Europe", *Ethnologia Scandinavica. A Journal for Nordic Ethnology*, 15-50.
- Holt, D. Eric (2006). "Optimality Theory and language change in Spanish", en *Optimality-Theoretic Studies in Spanish Phonology*. Fernando Martínez-Gil & Socia Colina (eds.). Amsterdam, Filadelfia: John Benjamins, pp. 378-398.
- Hymes, Dell (1974). *Foundations in Sociolinguistics: An Ethnographic Approach*. Filadelfia: University of Pennsylvania Press.
- INEGI (1977, 1984). *Encuesta Nacional de Ingresos y Gastos de los Hogares: ENIGH*. México: Instituto Nacional de Estadística, Geografía e Informática.
- INEGI (2000). *XII Censo general de población y vivienda, 2000*. Base de datos de la muestra censal. México: Instituto Nacional de Estadística, Geografía e Informática <http://www.inegi.gob.mx/est/contenidos/espanol/tematicos/mediano/ent.asp?t=medu14&c=3282>, 15 de mayo de 2014.
- INEGI (2012). *Cuéntame... Información por entidad. Distrito Federal. Dinámica de Crecimiento. Gráfica: "Población total del Distrito Federal (1970-2010)"*, en <http://cuentame.inegi.org.mx/monografias/informacion/df/poblacion/dinamica.aspx?tema=me&e=09> [consultada el 3 de junio de 2012].
- Inouye, Susan (1995). *Trills, Taps and Stops in Contrast and Variation*. Tesis doctoral. University of California, Los Angeles.

- Jacobs, Haike (1995). "Optimality Theory and sound change". Ponencia presentada en *NELS* 25. [Consultado el 11 de junio de 2004 en <http://roa.rutgers.edu/files/129-0496-JACOBS-0-0.PDF>].
- Jahangiri, N. (1980). *A Sociolinguistic Study of Tehrani Persian*. Tesis doctoral, University of London.
- Johnson, Daniel Ezra (2009). "Getting off the GoldVarb standard: Introducing Rbrul for mixed-effects variable rules analysis", *Language and Linguistics Compass* 3 (1), 359-383.
- Juilland, Alphonse & E Chang-Rodríguez (1964). *Frequency Dictionary of Spanish Words*. La Haya: Mouton and Company.
- Jurafsky, Daniel; Alan Bell; Michelle Gregory & William D. Raymond (2001). "Probabilistic relations between words: Evidence from reduction in lexical production", en *Frequency and the Emergence of Linguistic Structure*. Joan Bybee & Paul Hopper (eds.). Amsterdam: John Benjamins, pp. 229-254.
- Kager, René (1999). *Optimality Theory*. Cambridge: Cambridge University Press.
- Kaul de Marlangeon, Silvia Beatriz (2002). *Los adverbios en -mente del español de hoy y su función semántica de cuantificación*. Madrid: Iberoamericana Vervuert.
- Kerswill, Paul (1996). "Children, adolescents, and language change", *Language Variation and Change* 8, 177-202.
- Kiparsky, Paul (1993). "Variable rules". Ponencia presentada en *Rutgers Optimality Workshop*, Universidad de Rutgers.

- Koerner, Konrad (1991). "Toward a history of modern sociolinguistics", *American Speech* 66, 57-70.
- Koven, Michele (2011). "Comparing stories told in sociolinguistic interviews and spontaneous conversation", *Language in Society* 40 (1), 75-89.
- Kroch, Anthony (1989). "Reflexes of grammar in patterns of language change", *Language Variation and Change* 1, 199 - 244.
- Kurath, Hans; Marcus L. Hansen; Julia Bloch & Bernard Bloch (1939). *Handbook of the Linguistic Geography of New England*. Providence: American Council of Learned Societies.
- Labov, William (1963). "The social motivation of a sound change", *Word* 19, 273-303.
- Labov, William (1966). *The Social Stratification of English in New York City*. Washington, D. C.: Center for Applied Linguistics.
- Labov, William (1969). "Contraction, deletion and the inherent variability of the English copula", *Language* 45, 715-762.
- Labov, William (1972). *Sociolinguistic Patterns*. Filadelfia: University of Pennsylvania Press. [En español: *Modelos sociolingüísticos*, Madrid, Cátedra, 1983].
- Labov, William (1996). *Principios del cambio lingüístico*. Vol. I: *Factores internos*. Pedro Martín Butragueño (Trad.). Madrid: Gredos.
- Labov, William (2006). *Principios del cambio lingüístico*. Vol. II: *Factores sociales*. Pedro Martín Butragueño (Trad.). Madrid: Gredos.
- Labov, William (2010). *Principles of Linguistic Change*. Vol. III: *Cognitive Factors*. Oxford: Wiley-Blackwell.

- Labov, William; Sharon Ash & Charles Boberg (2007). *The Atlas of North American English. Phonetics, Phonology, and Sound Change. A Multimedia Reference Tool*. Berlín, Nueva York: Mouton de Gruyter.
- Lara Ramos, Luis Fernando (1990). *Dimensiones de la lexicografía. A propósito del Diccionario del español de México*. México: El Colegio de México.
- Lara Luis Fernando (2000). "Estructuras sintácticas en perspectiva crítica", en *Estructuras sintácticas 40 años después*. Luis Fernando Lara (ed.). México: El Colegio de México, pp. 61-78.
- Lara Ramos, Luis Fernando (2004). *Lengua histórica y normatividad*. México: El Colegio de México.
- Lara Ramos, Luis Fernando (2006). *Curso de lexicología*. México: El Colegio de México.
- Lara Ramos, Luis Fernando (dir.) (2011). *Diccionario del español de México*. México: El Colegio de México. Consultado en: <http://dem.colmex.mx/>, junio de 2014.
- Lass, Roger (1997). *Historical Linguistics and Language Change*. Gran Bretaña: Cambridge University Press.
- Lastra, Yolanda & Pedro Martín Butragueño (2000). "El modo de vida como factor sociolingüístico en la ciudad de México", en *Estructuras en contexto. Estudios de variación lingüística*. Pedro Martín Butragueño (ed.). México: El Colegio de México, pp. 13-43.
- Lastra, Yolanda & Pedro Martín Butragueño (2003). "Un posible cambio en curso: el caso de las vibrantes en la ciudad de México", en *Jornadas de sociolingüística*.

- Francisco Moreno Fernández *et al.* (eds.). Alcalá de Henares: Universidad de Alcalá.
- Lazard, Gilbert (2012). "The case for pure linguistics", *Studies in Language* 36 (2), 241-259.
- Legendre, Géraldine (2001). "An introduction to Optimality Theory in syntax", en *Optimality-Theoretic Syntax*. G. Legendre, J. Grimshaw & S. Vikner (eds.). Cambridge, Mass.: MIT Press, pp. 1-27.
- Lehmann, Christian (2007). "Linguistic competence: Theory and empiry", *Folia Linguistica* 41 (3-4), 223-278.
- Lipski, John M. (1990). "Aspects of Ecuadorian vowel reduction", *Hispanic Linguistics* 4, 1-19.
- Loeza, Soledad (1995). "Perspectivas para una historia política del Distrito Federal en el siglo XX", *Historia Mexicana* XLV (1), 99-158.
- Lope Blanch, Juan M. (1963-1964). "En torno a las vocales caedizas del español mexicano", *Nueva Revista de Filología Hispánica* XVII, 1-19.
- Lope Blanch, Juan M. (1967). "La -r final del español mexicano y el sustrato nahua", *Thesaurus* XXII (1), 1-20.
- Lope Blanch, Juan M. (1969). *El léxico indígena en el español de México*. México: El Colegio de México.
- Lope Blanch, Juan M. (dir.) (1971). *El habla de la ciudad de México: materiales para su estudio*. México: Universidad Nacional Autónoma de México.

- Lope Blanch, Juan M. (dir.) (1976). *El habla popular de la ciudad de México: materiales para su estudio*. México: Universidad Nacional Autónoma de México.
- Lope Blanch, Juan M. (dir.) (1990). *Atlas lingüístico de México*. 1, 1: *Fonética*. El Colegio de México - Fondo de Cultura Económica, México.
- Lope Blanch, Juan M. (dir.) (1992). *Atlas lingüístico de México*. 1, 2: *Fonética*. El Colegio de México - Fondo de Cultura Económica - UNAM, México.
- Lope Blanch, Juan M. (1995). *El habla popular de la República Mexicana: materiales para su estudio*. México: Universidad Nacional Autónoma de México.
- López Aguilar, Fernando (2011). "Los inicios de la Escuela Nacional de Antropología e Historia: 1937-1942", en *Escuela Nacional de Antropología e Historia. 70 años*. Alejandro Villalobos Pérez (coord.). México: Instituto Nacional de Antropología e Historia, Escuela Nacional de Antropología e Historia, pp. 21-41.
- López Chávez, Juan (2011). "Un acercamiento a la estructura del mundo narrativo del léxico básico del español de México", en *De la lengua sólo por la extrañeza. Estudios de lexicología, norma lingüística, historia y literatura en homenaje a Luis Fernando Lara*. Vol 1. México: El Colegio de México, pp. 397-416.
- López Morales, Humberto (1983). *Estratificación social del español de San Juan de Puerto Rico*. México: Universidad Nacional Autónoma de México.
- López Morales, Humberto (1988). "Índices de mortandad léxica en Puerto Rico: afronegrismos", *Nueva Revista de Filología Hispánica* XXXVI, 733-751.
- López Morales, Humberto (2006). "A propósito del *Diccionario usual del español de Puerto Rico*. Índices de mortandad léxica", en *El español en América. Diatopía*,

- diacronía e historiografía. Homenaje a José G. Moreno de Alba en su 65 aniversario*. Concepción Company Company (ed.). México: Universidad Nacional Autónoma de México, pp. 277-288.
- Lozanova, Elena (2000). "Notas sobre la vitalidad del léxico indígena en el español contemporáneo de la ciudad de México", en *Estructuras en contexto. Estudios de variación lingüística*. Pedro Martín Butragueño (ed.). México: El Colegio de México, pp. 61-79.
- Magueijo, João (2007). *Más allá de la velocidad de la luz. Historia de una especulación científica*. México: Fondo de Cultura Económica.
- Marden, Charles C. (1896 [1938]). *La fonología del español de la ciudad de Méjico*. Pedro Henríquez Ureña (trad.). Biblioteca de Dialectología Hispanoamericana IV. Buenos Aires: Universidad de Buenos Aires.
- Martín Butragueño, Pedro (1994). "Hacia una tipología de la variación gramatical en sociolingüística del español", *Nueva Revista de Filología Hispánica* XLII (1), 29-75.
- Martín Butragueño, Pedro (2000). "Las repercusiones del primer generativismo en el estudio de la variación y el cambio lingüístico", en *Estructuras sintácticas 40 años después*. Luis Fernando Lara (ed.). México: El Colegio de México, pp. 181-204.
- Martín Butragueño, Pedro (2002). *Variación lingüística y teoría fonológica*. México: El Colegio de México.

- Martín Butragueño, Pedro (2004). "El contacto de dialectos como motor del cambio lingüístico", en *El cambio lingüístico. Métodos y problemas*. Pedro Martín B. (ed.). México: El Colegio de México, pp. 81-144
- Martín Butragueño, Pedro (2006). "Líderes lingüísticos en la ciudad de México", en *Líderes lingüísticos. Estudios de variación y cambio*. Pedro Martín B. (ed.). México: El Colegio de México, pp. 185-208.
- Martín Butragueño, Pedro (2008). "Retracción lingüística", en *Teoría de la optimidad: estudios de sintaxis y fonología*. Rodrigo Gutiérrez Bravo & Esther Herrera Zendejas (eds.). México: El Colegio de México, pp. 159-196.
- Martín Butragueño, Pedro (2010). "Perspectiva sociolingüística de la historia", en *Historia Sociolingüística de México. Vol. 1: México prehispánico y colonial*. R. Barriga & P. Martín Butragueño (eds.). México: El Colegio de México.
- Martín Butragueño, Pedro (ed.) (2011). *Realismo en lingüística. Primer coloquio de cambio y variación lingüística*. México: El Colegio de México.
- Martín Butragueño, Pedro (2012). "Variación y cambio lingüístico en el español mexicano", *Español Actual. Revista de Español Vivo* 98, 11-38.
- Martín Butragueño, Pedro (2014). *Fonología variable del español de México*. México: El Colegio de México.
- Martín Butragueño, Pedro & Leonor Orozco (eds., en prensa). *II Coloquio de Cambio y Variación Lingüística*. México: El Colegio de México.

- Martín Butragueño, Pedro & Yolanda Lastra (2011). *Corpus Sociolingüístico de la Ciudad de México. Materiales de PRESEEA. Vol. 1: Hablantes de instrucción superior*. México: El Colegio de México. [Estudio introductorio y CD].
- Martín Butragueño, Pedro & Yolanda Lastra (2012). *Corpus Sociolingüístico de la Ciudad de México. Materiales de PRESEEA. Vol. 2: Hablantes de instrucción media*. México: El Colegio de México. [Estudio introductorio y CD].
- Martín Butragueño, Pedro & Yolanda Lastra (en prensa). *Corpus Sociolingüístico de la Ciudad de México. Materiales de PRESEEA. Vol. 3: Hablantes de instrucción baja*. México: El Colegio de México. [Estudio introductorio y CD].
- Martinet, André (1974). *Economía de los cambios fonéticos*. Madrid: Gredos.
- Matluck, Joseph H. (1951). *La pronunciación en el español del Valle de México*. México: Edición de Autor.
- Matluck, Joseph H. (1965). "Entonación hispánica", *Anuario de Letras* V, 5-32.
- McMahon, April (1994). *Understanding Language Change*. Cambridge: Cambridge University Press.
- Meillet, Antoine (1921). *Linguistique historique et linguistique generale*. 2 vols. Paris: Société de Linguistique.
- Mendoza Guerrero, Everardo (2011). *El habla de Culiacán*. Culiacán: Universidad Autónoma de Sinaloa, El Colegio de Sinaloa, Instituto Municipal de Cultura Culiacán.
- Milroy, Lesley (1980). *Language and Social Networks*. Londres: Basil Blackwell.

- Milroy, Lesley & Matthew Gordon (2003). *Sociolinguistics. Methods and Interpretation*. Oxford: Blackwell Publishing.
- Moreno de Alba, José G. (1994). *La pronunciación del español en México*. México: El Colegio de México.
- Moreno de Alba, José G. (2013). *Notas de gramática dialectal. El Atlas Lingüístico de México*. México: Universidad Nacional Autónoma de México.
- Moreno Fernández, Francisco (1990). *Metodología sociolingüística*. Madrid: Gredos.
- Moreno Fernández, Francisco (1994). "Sociolingüística, estadística e informática", ms.
- Moreno Fernández, Francisco (2000). *¿Qué español enseñar?* Madrid: Arco/Libros.
- Moreno Fernández, Francisco (2003). *Metodología del "Proyecto para el Estudio Sociolingüístico del Español de España y de América" (PRESEEA)*, en: <http://preseea.linguas.net/Portals/0/Metodologia/METODOLOG%C3%8DA%20PRESEEA.pdf>, (10 de junio de 2014).
- Moreno Fernández, Francisco (2010). "Historia, sociedad y lengua", en *Historia Sociolingüística de México. Vol. 1: México prehispánico y colonial*. R. Barriga Villanueva & P. Martín Butragueño (eds.). México: El Colegio de México.
- Moreno Fernández, Francisco (2012). *Sociolingüística cognitiva. Propositiones, escolios y debates*. Madrid: Iberoamericana/Vervuert.
- Nagy, Naomi & Bill Reynolds (1997). "Optimality Theory and variable word-final deletion in Faetar", *Language Variation and Change* 9, 37-55.
- Navarro Tomás, Tomás (1963). *Manual de pronunciación española*. 11ª edición. Madrid: CSIC.

- Nevalainen, Terttu; Helena Raumolin-Brunberg & Heikki Mannila (2011). "The diffusion of language change in real time: Progressive and conservative individuals and the time depth of change", *Language Variation and Change* 23, 1-43.
- Núñez Cedeño, Rafael & Alfonso Morales-Front (1999). *Fonología generativa contemporánea de la lengua española*. Colaboran Pilar Prieto i Vives y José Ignacio Hualde. Washington D. C.: Georgetown University Press.
- Obediente Sosa, Enrique (2008). "Más sobre variación y cambio fonológico: el caso de los fonemas vibrantes en el español de Mérida (Venezuela)", en *Fonología instrumental: patrones fónicos y variación*. Esther Herrera Z. & Pedro Martín B. (eds.). México: El Colegio de México, pp. 151-170.
- Ogura, M. (1995). "The development of Middle English *ī* and *ū*: a reply to Labov (1992, 1994)", *Diachronica* 12(1), 31-54.
- Ohala, John J. (1983). "The origin of sound patterns in vocal tract constraints", en *The Production of Speech*. P. F. MacNeilage (ed.). Nueva York: Springer-Verlag, pp. 189-216.
- Ohala, J. J. (1990). "The phonetics and phonology of aspects of assimilation", en *Papers in Laboratory Phonology I: Between the Grammar and the Physics of Speech*. J. Kingston & M. Beckman (eds.). Cambridge: Cambridge University Press, pp. 258-279.

- Palacios Cuahtecontzi, Niktelol (2002). *La interdicción lingüística en el habla de los adolescentes mexicanos*. Tesis de licenciatura. Benemérita Universidad Autónoma de Puebla.
- Paul, Hermann (1880). *Prinzipien der Sprachgeschichte*. Halle: Niemeyer. [4ta ed. de 1909, 5ta. de 1920].
- Penny, Ralph (2004). *Variation and Change in Spanish*. Cambridge: Cambridge University Press. [Versión electrónica. Original de 2000].
- 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.
- Pfleger Biering, Sabine (2009). *Frontera, mujeres y hombres oscuros. La conceptualización narrativa del mundo: un estudio de caso de la construcción mediática del feminicidio en Ciudad Juárez*. Tesis doctoral. UNAM.
- Phillips, Betty S. (2001). "Lexical diffusion, lexical frequency, and lexical analysis", en *Frequency and the Emergence of Linguistic Structure*. Joan Bybee & Paul Hopper (eds.). Amsterdam: John Benjamins, 123-136.
- Phillips, Betty S. (2006). *Word Frequency and Lexical Diffusion*. Basingstoke: Palgrave Macmillan.
- Pierrehumbert, Janet B. (2001). "Exemplar dynamics: Word frequency, lenition, and contrast", en *Frequency and the Emergence of Linguistic Structure*. Joan Bybee & Paul Hopper (eds.). Amsterdam: John Benjamins, pp. 137-157.

- Poch Olivé, Dolors; Bernard Harmegnies & Pedro Martín Butragueño (2008). "Influencia del estilo de habla sobre las características de las realizaciones vocálicas en el español de la ciudad de México", en *Actas del XV Congreso Internacional ALFAL*. Montevideo, Uruguay, 18-21 de agosto. [CD-ROM].
- Poplack, Shana & A. St. Amand (2007). "A real-time window on 19th-Century vernacular French: The "Recits du français québécois d'autrefois", *Language in Society* 36 (5), 707-734.
- Quilis, Antonio (1981). *Fonética acústica de la lengua española*. Madrid: Gredos.
- Quilis, Antonio (1993). *Tratado de fonología y fonética españolas*. Madrid: Gredos.
- Quilis, Antonio & R. B. Carril (1971). "Análisis acústico de [ř] en algunas zonas de Hispanoamérica", *Revista de Filología Española* 54, 271-316.
- Ramírez Luengo, José Luis (2011). "Sobre la historia de los grupos consonánticos cultos: el caso del español de Bilbao en los siglos XVI y XVII", *Oihenart* 26, 455-472.
- Real Academia Española: Banco de datos (CORDE) [en línea] (2008). *Corpus diacrónico del español*. <<http://www.rae.es>>. [Octubre de 2008].
- Real Academia Española: Banco de datos (CREA) [en línea] (2008). *Corpus de referencia del español actual*. <<http://www.rae.es>>. [Octubre de 2008].
- Real Academia Española: *Diccionario de la lengua española* (2008). 22^a. Ed. <<http://buscon.rae.es/drae/>>. [Junio de 2014].

- Real Academia Española: Banco de datos (CORPES XXI) [en línea] (2014). *Corpus del español del siglo XXI*. <<http://web.frl.es/CORPES/view/inicioExterno.view>>. [Versión beta 0.6, junio de 2014].
- Rissel, Dorothy (1986). "La dinámica social de la asibilación de vibrantes en San Luis Potosí, México", en *Actas del II Congreso Internacional del Español de América*. José G. Moreno de Alba (ed.). México: UNAM, pp. 357-361.
- Rissel, Dorothy (1989). "Sex, attitudes, and the assibilation of /r/ among young people in San Luis Potosí, Mexico", *Language Variation and Change* 1 (3), 269-283.
- Roberts, Julie (1997). "Hitting a moving target: Acquisition of sound change in progress by Philadelphia children", *Language Variation and Change* 9, 249-266.
- Robinson, John; Helen Lawrence & Sali Tagliamonte (2001). *GoldVarb 2001. A multivariate analysis application for Windows [User's Manual]*. University of York.
- Rodríguez Alfano, Lidia (coord.) (2005). *Investigación sociolingüística. El habla de Monterrey. Su trayectoria en una página electrónica*. Monterrey: Trillas, Universidad Autónoma de Nuevo León.
- Rodríguez Alfano, Lidia (2014). "El noreste de México: panorama sociolingüístico en diacronía", en *Historia Sociolingüística de México. Vol. 3: Espacio, contacto y discurso político*. Rebeca Barriga Villanueva & Pedro Martín Butragueño (dirs.). México: El Colegio de México, pp. 1569-1615.
- Rodríguez, Yolanda & Armando Mora (2001). "Las consonantes oclusivas en posición de coda silábica: estudio sobre el español de la ciudad de México", en *Temas de*

- fonética instrumental*. Esther Herrera Z (ed.). México: El Colegio de México, pp. 147-167.
- Romaine, Suzanne (2008). "Corpus Linguistics and Sociolinguistics", en *Corpus Linguistics. An International Handbook. Vol. 1*. Anke Lüdeling & Merja Kytö (eds). Nueva York: Walter de Gruyter, pp. 96-111.
- Rosado, Leonor (2013). *Estudio sociolingüístico de la ciudad de Mérida, Yucatán*. Tesis de Maestría. Universidad Nacional Autónoma de México.
- Samper, J. A., C. E. Hernández Cabrera & M. Troya (eds.). (1998). *Macrocorpus de la norma lingüística culta de las principales ciudades del mundo hispánico (MC-NLCH)*. Edición en CD-ROM. Las Palmas de Gran Canaria: Servicio de Publicaciones de la Universidad de las Palmas de Gran Canaria.
- San Giacomo, Marcela (2009). *La présence de l'espagnol en nahuatl: une étude sociolinguistique des adaptations et non-adaptations des emprunts*. Tesis de doctorado. Universidad de París VIII.
- Sánchez de la Barquera, Ana Maurel (2014). *Variación semántico-léxica en español mexicano: estudio sociolingüístico de dos redes sociales*. Tesis de licenciatura. Benemérita Universidad Autónoma de Puebla.
- Sánchez Ruiz, Ernesto E. (2005). "Los medios de comunicación masiva en México, 1968-2000", en *Una historia contemporánea de México: actores*. I. Bizberg & L. Meyer (coords.). México: Océano, pp. 403-447.

- Sankoff, David (1988). "Variable rules", en *Sociolinguistics. An International Handbook of the Science of Language and Society*. Ulrich Ammon, Norbert Dittmar & Klaus J. Mattheier (eds.). Nueva York: Walter de Gruyter, pp. 984-997.
- Sankoff, David; Sali A. Tagliamonte & Eric Smith (2005). *GoldVarb X. A Variable Rule Application for Macintosh and Windows*. Department of Linguistics, University of Toronto: <http://individual.utoronto.ca/tagliamonte/goldvarb.htm>, 3 de octubre de 2012.
- Saussure, Ferdinand de (1916 [1988]). *Curso de lingüística general*. 3a. edición. Traducción de Mauro Armiño. México: Fontamara.
- Schilling-Estes, Natalie (2007). "Sociolinguistic fieldwork", en *Sociolinguistic Variation: Theories, Methods, and Applications*. R. Bayley & C. Lucas (eds.). Cambridge: Cambridge University Press, pp. 165-189.
- Serrano, Julio (2002). *Dialectos en contacto. Variación y cambio lingüístico en migrantes sonorenses*. Tesis de Licenciatura. México: Escuela Nacional de Antropología e Historia.
- Serrano, Julio (2005). "Las vibrantes asibiladas del español de la ciudad de México: un estudio en tiempo aparente y tiempo real", ponencia presentada en el *1er. Congreso de Sociolingüística y Sociología del Lenguaje*, UNAM, Ciudad de México.
- Serrano, Julio (2006). "En torno a las vocales caedizas del español mexicano: una aproximación sociolingüística", en *Líderes lingüísticos. Estudios de variación lingüística*. Pedro Martín Butragueño (ed.). México: El Colegio de México, pp. 37-59.

- Serrano, Julio (2007). “¿Quién impone los límites de la lingüística? Una reflexión sobre la lingüística en México”, *Lingüística Mexicana* 4 (1), 93-108. [Fecha de publicación: 2010].
- Serrano, Julio (2008 a). “Vibrantes asibiladas en español de la ciudad de México (1964-1972)”, en *Fonología instrumental: patrones fónicos y variación*. Esther Herrera Z. & Pedro Martín B. (eds.). México: El Colegio de México, pp. 191-210.
- Serrano, Julio (2008 b). “Debilitamiento vocálico en español de la ciudad de México (1970-2000)”. Ponencia presentada en el *VII Coloquio de Lingüística en la ENAH*. México, D. F., 16-18 de abril.
- Serrano, Julio (2009). “Rescate de los archivos sonoros del Centro de Lingüística Hispánica”, en *Jornadas Filológicas 2007*. México: Instituto de Investigaciones Filológicas, UNAM, pp. 297-306.
- Serrano, Julio (2010 a). “La variación en el español del Norte de México: investigaciones pendientes”. Ponencia presentada en el *Coloquio Estado Actual y Perspectivas del Estudio de la Variación Lingüística en México y Latinoamérica*. El Colegio de México, 12 de marzo.
- Serrano, Julio (2010 b). “¿Existen esas silábicas en español? Un análisis acústico del debilitamiento vocálico en la variedad de la ciudad de México”. Ponencia presentada en *VIII Coloquio de Lingüística en la ENAH*, México, 21-23 de abril.
- Serrano, Julio (2011 a). “Retracción e innovación léxica en español de la ciudad de México: 1970-2000”, en *Realismo en lingüística. Primer coloquio de cambio y*

- variación lingüística*. Pedro Martín Butragueño (ed.). México: El Colegio de México, pp. 189-213.
- Serrano, Julio (2011 b). "Cambio léxico en español mexicano: estudio sociolingüístico en tiempo real". Ponencia presentada en el *XI Congreso Nacional de Lingüística (AMLA)*. Chetumal, Quintana Roo, 8-11 de noviembre.
- Serrano, Julio (2013). "Una propuesta de corpus para el estudio de la estructura argumental del español mexicano". Ponencia presentada en el *XII Congreso Nacional de Lingüística (AMLA)*. Santiago de Querétaro, 9-12 de octubre.
- Serrano, Julio (en prensa). "Detalle fonético y variación sociolingüística en Teoría de la Optimidad: dos procesos en español mexicano", en *Nuevos estudios de Teoría de Optimidad*. Francisco Arellanes Arellanes; Rodrigo Gutiérrez Bravo & Mario Chávez Peón (eds). México: El Colegio de México.
- Shuy, Roger W. (1990). "A brief history of American sociolinguistics, 1949-1989", *Historiographia Linguistica* 17 (1-2), 183-209.
- Siegel, Jeff (1985). "Koinés and koineization", *Language in Society* 14, 357-378.
- Silva, David James (1997). "The variable deletion of unstressed vowels in Faialense Portuguese", *Language Variation and Change* 9, 295-308.
- Silva-Corvalán, Carmen (2001). *Sociolingüística y pragmática del español*. Washington D. C.: Georgetown University Press.
- Smyth, Ron & Henry Rogers (2002). "Phonetics, gender and sexual orientation", en *Proceedings of the Annual Meeting of the Canadian Linguistics Association*.

- Montreal: University of Toronto, l'University du Quebec au Montréal, pp. 299–301.
- Soler, María Ángeles & Julio Serrano (2010). “El marcador discursivo *este*. Algunos aspectos prosódicos y estadísticos”, en *Memorias del X Congreso Nacional de Lingüística*. Toluca de Lerdo: Asociación Mexicana de Lingüística Aplicada. [CD-ROM].
- Steriade, Donca (1993). “Closure, release, and nasal contours”, en *Nasals, Nasalization, and the Velum*. M. Huffman & R. Krakow (eds.). Nueva York: Academic Press, 401–470.
- Steriade, Donca (1994). “Complex onsets as single segments: The Mazateco pattern”, en *Perspectives in Phonology*. J. Cole & C. Kisseberth (eds.). Stanford: CSLI, pp. 203–291.
- Steriade, Donca (1995). “Licensing retroflexion”. Ms. University of California at Los Angeles.
- Steriade, Donca (1997). “Phonetics in Phonology: The case of laryngeal neutralization”. Ms. University of California at Los Angeles.
- Stockwell, Robert (1972). “Problems in the interpretations of the Great Vowel Shift”, en *Proceedings of the Ninth International Congress of Linguists*. H. Lundt (ed.). La Haya: Mouton, pp. 663-671.
- Sturtevant, E. H. (1947). *An Introduction to Linguistics*. New Haven: Yale University Press.

- Swadesh, Morris (1955). "Towards greater accuracy in lexicostatistic dating", *International Journal of American Linguistics* 21, 121-137.
- Tagliamonte, Sali A. (2006). *Analyzing Sociolinguistic Variation*. Cambridge: Cambridge University Press.
- Tello, Carlos (2010). "Estancamiento económico, desigualdad y pobreza: 1982-2009", *Economía UNAM* 7 (19), 5-44.
- Terborg, Roland & Laura García Landa (coords.) (2011). *Muerte y vitalidad de las lenguas indígenas y las presiones sobre sus hablantes*. México: Centro de Enseñanza de Lenguas Extranjeras, UNAM.
- Tipton, Phillip (2009). "Betty S. Phillips, *Word Frequency and Lexical Diffusion*. Basingstoke: Palgrave Macmillan, 2006", *Language in Society* 38, 279-280.
- Tognini-Bonelli, Elena (2001). *Corpus Linguistics at Work*. Amsterdam, Filadelfia: John Benjamins.
- Torreblanca, Máximo (1979). "Un rasgo fonológico de la lengua española", *Hispanic Review* 47 (4), 455-468.
- Trudgill, Peter (1981). "Linguistic accommodation: Sociolinguistic observations on a socio-psychological theory", en *Papers from the Parasession on Language and Behavior*. C. S. Masek; R. A. Hendrick & M. F. Miller (eds.). Chicago: Chicago Linguistic Society, pp. 218-237.
- Trudgill, Peter (2000). "Sociolingüística y sociolingüística", en *Estudios de sociolingüística*. Yolanda Lastra (comp.). México: Universidad Nacional Autónoma de México, pp. 21-37.

- Vaquero, María (1983). "El léxico indígena en el español hablado en Puerto Rico", en *Philologica Hispaniensia in honorem Manuel Alvar*. Madrid: Gredos, pp. 621-640.
- Vaquero, María (1986). "Índices sociolingüísticos de los indigenismos en Puerto Rico", en *Léxico marinero en Puerto Rico y otros estudios*. San Juan: Editorial Playor, pp. 149-193.
- Vázquez Laslop, María Eugenia (2014). "El discurso político en México (1968-1994): la emergencia del diálogo", en *Historia Sociolingüística de México. Vol. 3: Espacio, contacto y discurso político*. R. Barriga Villanueva & P. Martín Butragueño (dirs.). México: El Colegio de México, pp. 1783-1896.
- Villena Ponsoda, Juan Andrés (2008). "La formación del español común en Andalucía: un caso de escisión prestigiosa", en *Fonología instrumental: patrones fónicos y variación*. Pedro Martín B. y Esther Herrera Z. (coords.). México: El Colegio de México, pp. 211-256.
- Vollmer, Helmut (1982). *Spracherwerb und Sprachbeherrschung*. Tubinga: G. Narr.
- Wagner, Suzanne Evans (2012). "Real-time evidence for age grad(ing) in late adolescence", *Language Variation and Change* 24, 179-202.
- Warburg, W. von (1951). *Problemas y métodos de la lingüística*. Madrid: Gredos.
- Weinreich, Uriel; William Labov & Marvin I. Herzog (1968). "Empirical foundations for a theory of language change", en *Directions for Historical Linguistics. A Symposium*. W. P. Lehmann & Y. Malkiel (eds.). Austin: University of Texas Press, pp. 95-195.

- Woods, Anthony; Paul Fletcher & Arthur Hughes (1986). *Statistics in Language Studies*. Cambridge: Cambridge University Press.
- Youssef, Valerie (1993). "Children's linguistic choices: Audience design and societal norms", *Language in Society* 22 (2), 257-274.
- Zamora Munné, Juan C. & Jorge M. Guitart (1988). *Dialectología hispanoamericana. Teoría, descripción, historia*. 2da. Ed. Salamanca: Almar.
- Zgusta, Ladislav (1990). "Onomasiological change: Sachen-change reflected by Wörter", en *Research Guide on Language Change*. Edgar C. Polomé (ed.). *Trends in Linguistics. Studies and Monographs* 48. Werner Winter (ed.). Nueva York: Mouton de Gruyter, pp. 389-398.
- Zubritskaya, Katia (1997). "Mechanism of sound change in Optimality Theory", *Language Variation and Change* 9, 121-148.

APÉNDICE

CARACTERÍSTICAS SOCIALES DE LOS 48 INFORMANTES BASE

En los siguientes cuadros, los datos sobre profesión u ocupación se anotan tal como se registraron en las bases de datos originales. El orden de los informantes en cada cuadro es por grupos de edad: de jóvenes a mayores. Los 48 informantes de las siguientes tablas constituyeron la base de datos para los capítulos 3 y 4 sobre variación fónica. Tal como se indicó en el cap. 5, no siempre se contaba con la transcripción completa de la entrevista de todos estos informantes, por lo que algunos de ellos fueron sustituidos por informantes con características sociales equivalentes (mismo grupo de edad, sexo y tipo de ocupación) en los análisis fonológicos, morfológicos y lexicológicos abordados en dicho capítulo.

La disposición de los cuadros permite hacer la comparación directa entre los informantes según el año de la entrevista (1970 y 2000) para facilitar al lector la evaluación de la comparabilidad de los informantes de una y otra época.

INFORMANTES BAJA INSTRUCCIÓN 1970
INFORMANTES BAJA INSTRUCCIÓN 2000

Sexo	Edad	Año Nac.	Estudios	Ocupación	Sexo	Edad	Año Nac.	Estudios	Ocupación
Masc	31	1941	Primaria	Peluquero	Masc	26	1980	Primaria concluida	Mensajero
Masc	24	1948	5to año primaria ("semianalfabeta")	Obrero	Masc	22	1981	1er año bachillerato	Policía
Fem	21	1951	6to año primaria ("semianalfabeta")	Obrera (desempleada)	Fem	30	1969	Primaria	Cocinera
Fem	23	1949	4to. año primaria ("semianalfabeta")	Hogar	Fem	27	1979	1ro de primaria	Comerciante
Masc	46	1926	Primaria	Carpintero	Masc	44	1957	2do de primaria	Albañil
Masc	36	1936	No se indica ("semianalfabeta")	Conserje	Masc	51	1955	No especificado	Chofer de taxi
Fem	37	1935	3er año primaria	Trabajadora doméstica	Fem	49	1962	Primaria	Limpieza, hogar
Fem	47	1925	2do año primaria	Hogar, comerciante	Fem	54	1952	No especificado	Hogar
Masc	58	1914	Ninguno ("analfabeta")	Jardinero	Masc	69	1927	No especificado	Comerciante, campesino
Masc	72	1900	Semianalfabeta	Pensionado	Masc	55	1946	Primaria inconclusa	Pintor
Fem	66	1906	Sabe leer y escribir ("semianalfabeta")	Lavandera	Fem	57	1949	1ro primaria inconcluso	Ama de casa
Fem	57	1915	2do año primaria ("semianalfabeta")	Cocinera	Fem	91	1910	2do primaria inconcluso	Vendedora de flores

INFORMANTES ALTA INSTRUCCIÓN 1970
INFORMANTES ALTA INSTRUCCIÓN 2000

Sexo	Edad	Año Nac.	Estudios	Ocupación	Sexo	Edad	Año Nac.	Estudios	Ocupación
Masc	26	1938	3er año Derecho, 1 año Ciencias Pol.	Estudiante	Masc	29	1972	Licenciatura	Arquitecto
Masc	26	1942	Contador Público ("culto")	Funcionario bancario	Masc	26	1978	Licenciatura	Ingeniero
Fem	22	1945	Educadora. Lee inglés.	Educadora	Fem	25	1972	Licenciatura	Estudiante
Fem	30	1939	Maestra en Geografía. Inglés.	Maestra	Fem	23	1973	Licenciatura	Estudiante
Masc	39	1925	Abogado	Investigador UNAM	Masc	39	1965	Maestría	Musicólogo, escritor
Masc	36	1928	Químico bacteriólogo.	Investigador IPN	Masc	45	1960	Licenciatura	Biólogo
Fem	35	1934	Normal Superior. Inglés y francés.	Traductora, maestra inglés	Fem	41	1963	Licenciatura	Cirujano dentista
Fem	35	1935	Normal Superior. 3 años Letras Hisp.	Directora de escuela	Fem	37	1967	Pasante de licenciatura	Hogar
Masc	56	1911	Medicina	Médico	Masc	64	1940	Licenciatura	Médico, investigador
Masc	68	1899	Primaria y comercio ("culto")	Jubilado, exmilitar	Masc	56	1941	3er semestre licenciatura	Contratista
Fem	80	1887	Obstetricia en escuela libre	Hogar	Fem	65	1936	Licenciatura	Arquitecta, profesora
Fem	55	1912	Maestra normalista. Inglés.	Maestra de secundaria	Fem	62	1941	Licenciatura	Maestra jubilada