



UNIVERSIDAD NACIONAL  
AUTÓNOMA DE  
MÉXICO

UNIVERSIDAD NACIONAL AUTÓNOMA DE MÉXICO  
FACULTAD DE FILOSOFÍA Y LETRAS  
COLEGIO DE LETRAS HISPÁNICAS

---

---

---

REGLAS DE CORRESPONDENCIA ENTRE  
SONIDO Y GRAFÍA EN EL ESPAÑOL  
HABLADO EN MÉXICO EN EL SIGLO XVI  
PARA LA CREACIÓN DE UN TRANSCRIPTOR  
AUTOMÁTICO.

UNA APORTACIÓN AL *CORPUS HISTÓRICO  
DEL ESPAÑOL EN MÉXICO* (CHEM)

T E S I S

QUE, PARA OBTENER EL TÍTULO DE  
LICENCIADA EN LENGUA Y LITERATURAS HISPÁNICAS,  
PRESENTA

TERESITA ADRIANA REYES CAREAGA

ASESOR: MTRO. JAVIER OCTAVIO CUÉTARA PRIEDE  
CO-ASESOR: DR. ALFONSO MEDINA URREA



CIUDAD UNIVERSITARIA, 2008



Universidad Nacional  
Autónoma de México



**UNAM – Dirección General de Bibliotecas**  
**Tesis Digitales**  
**Restricciones de uso**

**DERECHOS RESERVADOS ©**  
**PROHIBIDA SU REPRODUCCIÓN TOTAL O PARCIAL**

Todo el material contenido en esta tesis esta protegido por la Ley Federal del Derecho de Autor (LFDA) de los Estados Unidos Mexicanos (México).

El uso de imágenes, fragmentos de videos, y demás material que sea objeto de protección de los derechos de autor, será exclusivamente para fines educativos e informativos y deberá citar la fuente donde la obtuvo mencionando el autor o autores. Cualquier uso distinto como el lucro, reproducción, edición o modificación, será perseguido y sancionado por el respectivo titular de los Derechos de Autor.

A mi abuelo, José de Jesús Careaga Manrique† (1914-2004),  
porque él me enseñó la perfección implícita en hacer una cosa a la vez.

*-Paréceme, Sancho, que tienes mucho miedo.  
-Si tengo-, respondió Sancho; -mas ¿en qué lo echa de ver  
vuestra merced ahora más que nunca?  
-En que ahora más que nunca hueles, y no a ámbar-, respondió don Quijote.  
Cervantes*

Agradezco a todos aquellos que estuvieron conmigo durante la realización de esta tesis, en especial:

A mi mamá y a Maru, porque siempre creyeron en mí y me apoyaron en cada cosa que emprendí.

A mis hermanos, Cata, Pepe, Oly, Pili y Migue, porque sé que estuvieron conmigo aunque no me diera cuenta; porque ustedes me vieron dando tumbos por la vida sin intervenir para criticar, sino para orientarme. A Cuquis, porque aparte de que eres mi amiga, formas parte de mi familia, y a José Xavier por darme la energía de cada día.

A mis amigos, Po, Octa, Andrea, George y Ross, por los quince años juntos. A Rui y al Pollo, porque me enseñaron este camino. A Gabino y a Roberto, porque me enseñaron que, no importa donde se encuentren, hay que salir a perseguir nuestros sueños. A Jéssica por los proyectos juntas y a Bertha porque más que mi amiga, eres mi colega y una crítica más de este trabajo.

A Barrón y Fonseca, por hacer el ambiente de trabajo menos denso y mucho más divertido. A Laura y Ari, por su apoyo, y a los demás miembros del GIL por ser parte de esto.

A Bulmaro y a Fulvia porque, al convertirse en mis amigos, me enseñaron más en la vida que en el salón de clases. A Sergio porque sin tu ayuda no hubiera llegado al término de esta tesis, de nuevo, muchas gracias.

A Javier y Amú por su amistad y su guía en este trabajo, por creer en mí y hacerme crecer, y por los proyectos futuros. A Concepción, Sergio y Graciela por enseñarme tantas cosas y por leer este trabajo.

A Jesús Careaga†, Catalina Hernández† y Lulú† porque siempre van conmigo y no los olvido.

A la UNAM, porque siempre me dio, gratuitamente, una educación de alto nivel.

A DGAPA porque gracias a la beca que me otorgó pude dedicarme a concluir esta tesis sin la necesidad de salir a buscar trabajo.

Esta tesis fue realizada dentro del proyecto DGAPA-UNAM, PAPIIT IN400905 llamado  
“Constitución del *Corpus Histórico del Español de México* (CHEM)”.

# ÍNDICE

|          |                                                                                    |    |
|----------|------------------------------------------------------------------------------------|----|
| 1.       | INTRODUCCIÓN                                                                       | 1  |
| 2.       | APORTACIONES DE LOS CORPUS ELECTRÓNICOS A LAS INVESTIGACIONES LINGÜÍSTICAS         | 8  |
| 2.1.     | <b>Corpus electrónicos</b>                                                         | 8  |
| 2.2.     | <b>Sistemas de codificación de corpus en internet</b>                              | 15 |
| 2.3.     | <b>Caracteres: ASCII, Unicode</b>                                                  | 17 |
| 2.4.     | <b>Caracteres de representación fonética: RFE, IPA, Mexbet</b>                     | 18 |
| 2.5.     | <i>TranscribeMex: antecedente de la aplicación</i>                                 | 19 |
| 2.6.     | <i>El Corpus Histórico del Español en México: necesidades</i>                      | 22 |
| 3.       | FENÓMENOS FONÉTICOS GENERALES EN LA HISTORIA DEL ESPAÑOL DE AMÉRICA                | 27 |
| 3.1.     | <b>Cambios fonéticos del español medieval al moderno</b>                           | 27 |
| 3.1.1.   | <i>Reajuste de sibilantes</i>                                                      | 31 |
| 3.1.2.   | <i>Confusión gráfica y fonética de b y v</i>                                       | 38 |
| 3.1.3.   | <i>Aspiración de f- inicial</i>                                                    | 41 |
| 3.1.4.   | <i>Yeísmo</i>                                                                      | 46 |
| 3.2.     | <b>Trasplante del español a América</b>                                            | 49 |
| 3.3.     | <b>Andalucismo del español americano</b>                                           | 53 |
| 4.       | REGLAS DE CORRESPONDENCIA ENTRE SONIDO Y GRAFÍA                                    | 63 |
| 4.1.     | <b>El corpus: características</b>                                                  | 63 |
| 4.1.1.   | <i>Características editoriales de los textos que integran el corpus de trabajo</i> | 67 |
| 4.2.     | <b>Sonidos vocálicos</b>                                                           | 70 |
| 4.2.1.   | <i>Vocales altas</i>                                                               | 70 |
| 4.2.1.1. | Vocal alta anterior [i]                                                            | 71 |
| 4.2.1.1. | Vocal alta posterior [u]                                                           | 74 |
| 4.2.2.   | <i>Vocales medias</i>                                                              | 76 |
| 4.2.3.   | <i>Vocal baja</i>                                                                  | 77 |
| 4.2.4.   | <i>Variación de vocales átonas</i>                                                 | 77 |
| 4.3.     | <b>Sonidos consonánticos</b>                                                       | 78 |
| 4.3.1.   | <i>Consonantes oclusivas sordas</i>                                                | 79 |

|                                                                            |     |
|----------------------------------------------------------------------------|-----|
| 4.3.2. <i>Consonantes oclusivas sonoras</i>                                | 84  |
| 4.3.3. <i>Consonantes africadas sordas</i>                                 | 88  |
| 4.3.4. <i>Consonantes fricativas sordas</i>                                | 91  |
| 4.3.5. <i>Consonante fricativa sonora</i>                                  | 103 |
| 4.3.6. <i>Consonantes nasales</i>                                          | 105 |
| 4.3.7. <i>Consonantes vibrantes</i>                                        | 107 |
| 4.3.8. <i>Consonante lateral</i>                                           | 109 |
| 4.3.9. <i>Grupo [ks]</i>                                                   | 110 |
| <b>4.4. Problemas específicos</b>                                          | 111 |
| 4.4.1. <i>Sonidos que permanecieron temporalmente por léxico indígena</i>  | 111 |
| 4.4.2. <i>Grupos de consonantes</i>                                        | 113 |
| 5. CONCLUSIONES                                                            | 117 |
| 6. BIBLIOGRAFÍA                                                            | 123 |
| 7. APÉNDICE 1. <b>Alfabetos fonéticos (tabla de sonidos del siglo XVI)</b> | 129 |
| 8. APÉNDICE 2. <b>Condensación de reglas</b>                               | 130 |
| 9. APÉNDICE 3. <b>Lista de excepciones</b>                                 | 134 |

## ÍNDICE DE FIGURAS

|            |                                                                                             |    |
|------------|---------------------------------------------------------------------------------------------|----|
| Figura 1.  | Búsqueda de la palabra <i>cibdad</i> en el CORDE de la RAE (siglo XVI en México)            | 10 |
| Figura 2.  | Resultados de la palabra <i>cibdad</i> en el CORDE                                          | 11 |
| Figura 3.  | Resultados de la palabra <i>cibdad</i> en el Corpus del Español                             | 11 |
| Figura 4.  | Resultados de la palabra <i>cibdad</i> en el CHEM                                           | 12 |
| Figura 5.  | Resultados de la palabra <i>çibdad</i> en el CHEM                                           | 13 |
| Figura 6.  | Resultados de la palabra <i>çiudad</i> en el CHEM                                           | 13 |
| Figura 7.  | Resultados de la palabra <i>ciudad</i> en el CHEM                                           | 14 |
| Figura 8.  | Transcripción de <i>Sacra çesárea catholica majestad</i> en <i>TranscribEMex</i>            | 20 |
| Figura 9.  | Transcripción de <i>no puedo dexar de hazer lo que soy oblygado</i> en <i>TranscribEMex</i> | 21 |
| Figura 10. | Sistema fonológico latino                                                                   | 29 |
| Figura 11. | Sistema fonológico del castellano medieval                                                  | 29 |
| Figura 12. | Sistema fonológico del español moderno (modalidad atlántica)                                | 30 |
| Figura 13. | Variación gráfica de <i>licenciado</i> en los <i>DLNE</i>                                   | 65 |

# 1. INTRODUCCIÓN

Empezaré con una idea de Lope Blanch (1969a), con relación a las tareas que él consideraba más urgentes dentro del campo filológico mexicano, cuando opinaba que el estudio y la descripción de la lengua española que —desde España y desde las Antillas— llegó a México en el siglo XVI era la principal de estas tareas. Él trató, con la publicación de sus *Estudios sobre el español de México* (1972), de estimular a los estudiantes de filología para “emprender investigaciones personales en torno a este tema” (1972). Quizá la esperanza que tenía (en sus propias palabras) “se transformaría en viva satisfacción” cuando este trabajo de tesis haya culminado.

En esta tesis se pretende combinar la investigación lingüística con el quehacer computacional. Es muy importante para mí realizar un estudio sobre el español que llegó a México en el siglo XVI y utilizar este análisis para contribuir al conocimiento de la historia del español hablado en México en dicho siglo mediante las reglas de correspondencia entre sonido y grafía que se implementarán en un transcriptor automático. Sin embargo, no es fácil la tarea de combinar el trabajo lingüístico —que, por ser sincrónico en el siglo XVI y referente a la pronunciación de los hablantes, es en gran medida un ejercicio de imaginación—, con el trabajo computacional —que requiere aterrizar las ideas y transformarlas en reglas de correspondencia que puedan ser utilizadas por un programa computacional.

El objetivo central de esta tesis será establecer el sistema fonológico del español hablado en México en el siglo XVI para crear las reglas de correspondencia entre sonido y grafía que se incorporarán a un transcriptor automático (Fonseca en proceso), a partir

del análisis en documentos novohispanos del siglo XVI. Cabe aclarar que nunca tendremos un acercamiento real al español hablado en México en dicho periodo, ya que la escritura no necesariamente refleja el habla a causa de ciertos hábitos escriturarios heredados de los amanuenses. El término *hablado* que utilizo en el título de esta tesis hace referencia al acercamiento que intenté rastrear entre la escritura de mi corpus de trabajo y la pronunciación de los hablantes del siglo XVI. Esto en la medida en que los textos escritos puedan reflejar el registro hablado de un sector de la población establecida en México en dicho siglo.<sup>1</sup>

Lo anterior requiere establecer antes algunos aspectos, que conforman los objetivos específicos de este trabajo. En el área de la lingüística son: observar el proceso de transición del sistema fonológico del castellano medieval al español moderno; observar cuáles fonemas representaban problemas y vacilaciones gráficas por su evolución natural y establecer en qué etapa de esta evolución se encontraban; observar bajo qué condiciones se dio el trasplante del español al continente americano y, con relación a este punto, observar si la mayor parte de la población española que llegó a América era de origen andaluz, como se ha discutido ampliamente en la tradición hispanista. Si así fuera, otro objetivo será observar qué fenómenos de ese dialecto permearon el español americano y cómo se encuentran reflejados en los documentos del corpus de trabajo. Si bien lo anterior ya se ha hecho, en esta tesis se realizará de manera especial, para reunir los estudios al respecto y prestando atención particular para el español mexicano.

Los temas anteriores son relevantes ya que proporcionan información valiosa acerca de las condiciones en las cuales se estableció el español en América en el periodo de orígenes. Durante los primeros cien años después del descubrimiento de América, no

---

<sup>1</sup> La escritura puede ser un disfraz de la producción oral, a esto se refiere Wright (1982/1989:41) como *fonoestética* o *simbolismo fónico*; es decir, “una razón no fonética que podría haber hecho que una palabra optara por no cambiar y quedarse rezagada”.

existió un español que pudiera llamarse americano, por lo que el español que se hablaba en México después del descubrimiento y hasta la primera mitad del siglo XVI dependía de los hablantes españoles y de sus propios dialectos. Finalmente, la mezcla que se dio entre estos dialectos y las lenguas nativas americanas proporcionó la base del español americano; por todo lo anterior también es importante determinar quiénes eran los hablantes que sentaron esa base para la conformación del español mexicano.

Para establecer el sistema fonológico del español hablado en México en el siglo XVI, se tendrán en cuenta factores como el grado cultural de los amanuenses bajo la premisa de que cuanto más bajo fuera su nivel cultural, se reflejarán en su escritura más fenómenos del habla real y popular de la población, mientras que, si su origen y cultura eran mayores, se observarán factores con tendencia cultista y/o conservadora, los cuales, como se sabe, frenan los cambios en la lengua. Asimismo, se tendrá en cuenta la cronología del documento con atención a la premisa de que, cuanto más temprano en el siglo XVI se haya escrito un texto reflejará menor avance de algún fenómeno en específico; por lo contrario, si un texto se ha escrito en el ocaso de dicho siglo, presentará los mismos fenómenos con un grado mayor de avance o totalmente cumplidos.

Otro factor que será observado es el origen geográfico del amanuense con relación al andalucismo mencionado arriba pues los fenómenos que permearon el español americano serán diferentes si la influencia provenía de un español del norte o del sur de la Península.

Todos estos aspectos se reflejarán de alguna manera en el corpus de trabajo y este brindará los fenómenos posibles que se estuvieran dando en dicho periodo en el español hablado en México. Como menciona Menéndez Pidal (1904/1940:31), “la fonética histórica, que estudia las transformaciones de la pronunciación desde la época

latina a la actual se funda casi únicamente en el estudio de los sonidos tal como han sido escritos”, de manera que sólo el análisis del corpus, confrontado con lo que dicen los teóricos, el conocimiento de la fonética actual y los diversos cambios de grafía para representar un sonido pueden dar pistas sobre la fonética histórica del español mexicano.

Dentro de la parte computacional, y como mencioné arriba, para los fines de este trabajo se deberán establecer correspondencias lo más abarcadoras y generales entre una grafía (en su debido contexto) y su equivalencia al sonido.

Los residuos de la investigación, bajo la premisa de que toda regla tiene excepciones, se ubicarán, para hacer más sencillo el trabajo del transcriptor automático, en una lista de excepciones, llamada en computación lista de filtrado. Las transcripciones de las palabras que se encuentren en esta lista se harán primero, antes de ejecutar las reglas arrojadas por el análisis de manera automática.

La hipótesis central de este trabajo es que el sistema fonológico del español hablado en el siglo XVI era ya semejante al actual, con pequeñas excepciones como la permanencia temporal de los sonidos [š] y [š̃] por influencia de las lenguas indígenas americanas (los cuales después se ajustaron a la pauta de evolución del español); la presencia del sonido [h] por influjo del dialecto andaluz y la convivencia del sonido [š] con el velar [x] porque el proceso de velarización no se había cumplido totalmente (Parodi 1995:43).

Mi investigación se enfocará en el siglo XVI, pues es el periodo en el que el español se asienta en América y en el cual, incluso en España, no había todavía una norma ortográfica bien establecida. Además porque esto es conveniente para propósitos cronológicos del *Corpus Histórico del Español en México* (CHEM),<sup>2</sup> en el cual se

---

<sup>2</sup> En §2.6. se hablará del corpus de trabajo y del proyecto del CHEM.

inserta esta investigación, ya que posteriormente se harán estudios sobre otros periodos que abarca este corpus (contiene textos del siglo XVI al siglo XIX, como se explicará más adelante) y conviene tratarlos en estricto orden cronológico para que los estudios subsecuentes puedan apoyarse en las investigaciones previas.

La investigación es importante, primero, por rastrear el origen posible del español mexicano; es decir, establecer si hay una influencia real del dialecto andaluz que se mencionó antes. Segundo, porque no hay muchos estudios específicos para el español hablado en México, sino descripciones sobre el español en España de ese siglo y sobre generalidades americanas.

En el capítulo 2 de esta tesis, el lector encontrará la descripción de las aportaciones de los corpus lingüísticos electrónicos o informatizados a las investigaciones lingüísticas; algunos ejemplos de corpus electrónicos disponibles en internet, los sistemas de codificación y etiquetado que requieren estos corpus; los caracteres de representación fonética que se utilizarán en el transcriptor automático; el antecedente de este transcriptor y la descripción general del *Corpus Histórico del Español en México* (CHEM), así como sus trabajos y necesidades.

En el capítulo 3 describiré algunos de los fenómenos fonéticos generales en la historia del español que se mencionaron arriba, estos son: el cambio fonético del castellano medieval al español moderno, el trasplante del español a América y el andalucismo en el español americano.

En el capítulo 4 se encontrará el análisis realizado para establecer las reglas de correspondencia entre sonido y grafía para el español hablado en el siglo XVI — abordando cada sonido según su modo de articulación y su rasgo de sonoridad— y se tratarán los problemas específicos que plantean algunos de estos sonidos con relación a su graficación. Para esta parte del análisis hice una selección de los documentos del

siglo XVI en los que se pudieran observar vacilaciones grafía-sonido. Asimismo, se verificó, en una muestra lo necesariamente amplia, que se tomaran todos los contextos posibles para cada grafía; posteriormente, se hizo el análisis, a mano, de cada contexto para cada grafía, después, con la confrontación del análisis arrojado y la opinión de algunos teóricos, se sugirió a qué sonido correspondía cada grafía en ese periodo y se establecieron las reglas de correspondencia.

Por supuesto que no deberá esperarse una correlación de uno a uno entre grafía y sonido, pues los sonidos correspondientes dependen del contexto fónico y, por ello, las grafías que representen dicho sonido también estarán sujetas a contextos gráficos específicos.

Por último, al final de esta tesis, se podrán encontrar tres apéndices: el primero es una tabla de fonemas para el español hablado en México en el siglo XVI y su correspondencia entre los alfabetos fonéticos de AFI, RFE y Mexbet. El lector también obtendrá, en el apéndice 2, la condensación de las reglas que se establecerán a lo largo del capítulo 4. Y en el apéndice 3 se encontrará la lista de excepciones (o lista de filtrado) que se mencionó anteriormente, en la cual el lector podrá observar la totalidad de las palabras que no se ajustaron a las reglas obtenidas por el análisis con su debida transcripción fonológica.

Así, el presente trabajo arroja luz sobre el español hablado en México en el siglo XVI y lo enlaza, a través del transcriptor automático, con las nuevas tecnologías, en una labor que resulta novedosa, por un parte —ya que mejorará en gran medida el trabajo que se hace actualmente en la interfaz del corpus—, e importante por la otra —por el acercamiento al conocimiento del español asentado en México después de la conquista. Sin duda, esta combinación le hubiera agradado al profesor Lope Blanch, pues a partir

del estudio filológico del español hablado en México en el periodo abordado habrá una aportación a las tecnologías y directamente al CHEM.

Es importante recalcar, una vez más, que esta tesis es una aproximación al español hablado en México en el siglo XVI tomando como punto de partida la escritura de dicho periodo. La escritura no reflejará de manera exacta la forma de hablar de los pobladores americanos, ya que los hábitos de escritura siempre van un paso atrás de las producciones orales de los hablantes; sin embargo, la escritura es la única pista que tenemos para darnos una idea de cómo hablaba la gente que habitaba el México del siglo XVI.

## 2. APORTACIONES DE LOS CORPUS ELECTRÓNICOS A LAS INVESTIGACIONES LINGÜÍSTICAS

En esta sección hablaré de manera general del problema que plantea la investigación lingüística en los corpus en papel, las ventajas que ofrecen los corpus lingüísticos electrónicos o informatizados frente a aquellas investigaciones y mencionaré los sistemas de codificación y caracteres que se utilizan para manejar los corpus electrónicos.

También expondré la aplicación final que tendrá la investigación de esta tesis que será un transcriptor automático. Mostraré su antecedente, llamado *TranscribEMex* (Cuétara 2004) y los caracteres de representación fonética con los que se trabajará en el programa.

Finalmente, presentaré el *Corpus Histórico del Español en México* (CHEM) y su interfaz que ya se encuentra en línea en <http://www.iling.unam.mx/chem>, las opciones de búsqueda que ofrece actualmente, sus necesidades, trabajos en curso e investigación a futuro.

### 2.1. Corpus electrónicos

Actualmente, es un tema conocido la desventaja que significa trabajar con corpus en papel, ya que representa utilizar mayor tiempo de análisis y resultados no tan precisos como se espera. Es por eso que recientemente han surgido herramientas automatizadas para manejar grandes cantidades de información y manipularla para beneficio de investigaciones científicas.

Corpora which are machine-readable possess several advantages over the original written or spoken format. The first and most important advantage of machine-readable corpora is that they may be searched and manipulated in ways which are simply not possible with the other formats (McEnery y Wilson 1996).

Hoy en día cualquier persona puede diseñar un corpus afín a su propia investigación y hacerlo electrónico para manipular los datos obtenidos. Algunos de estos corpus están disponibles para que el público en general pueda hacer uso de ellos en sus investigaciones. Entre los más conocidos, para el español, se encuentran los de la Real Academia Española: el *Corpus de Referencia del Español Actual* (CREA) <http://corpus.rae.es/creanet.html>, con 160 millones de palabras gráficas y el *Corpus Diacrónico del Español* (CORDE) <http://corpus.rae.es/cordenet.html>, con 250 millones de palabras gráficas; también es conocido el Corpus del Español de Mark Davies <http://www.corpusdelespanol.org/> que cuenta con cien millones de palabras gráficas. Más adelante veremos las interfaces de internet de estos corpus y su funcionamiento.

Las ventajas que representa trabajar con corpus electrónicos son muchas; por ejemplo, manipular información en varios niveles de lengua, como léxico, sintáctico, fonológico, morfológico, etc. También, típicamente, se pueden hacer búsquedas especializadas por palabra, por lema o por clase de palabra. Asimismo, en los corpus abiertos (es decir, en crecimiento) se puede incrementar fácilmente la cantidad de textos utilizados sin representar ningún problema, enriquecer la información mediante nuevos análisis lingüísticos, así como ampliar su temática o su representatividad (sea diatópica, diastrática o diacrónica).

Nos dice Torruella (1999:4) que “las ventajas de trabajar con corpus informatizados, sobre todo con aquellos que están anotados, es tan grande, que está obligando a los lingüistas *tradicionales* a trabajar conjuntamente con lingüistas computacionales”.

Un corpus anotado es una de las herramientas más útiles para el investigador, ya que estas marcas o etiquetas,<sup>1</sup> si están bien asignadas, ofrecen información léxica, ortográfica, sintáctica, fonológica, etc. del documento original.

A corpus, when annotated, may be considered to be a repository of linguistic information, because the information which was implicit in the plain text has been made explicit through concrete annotation (McEnery y Wilson 1996).

Una de las finalidades que pretende el CHEM es ofrecer su interfaz para que el público general pueda acceder a ella y hacer uso de sus recursos para manipular material lingüístico de la Nueva España y del México del siglo XIX, y con ello hacer investigaciones lingüísticas empíricas.

Veamos ejemplos de los corpus electrónicos mencionados anteriormente; la siguiente figura (1) muestra la interfaz de búsqueda del CORDE de la Real Academia Española:

| Real Academia Española - Corpus Diacrónico del Español (CORDE)                  |                                                                                                                    |        |                                                                     |
|---------------------------------------------------------------------------------|--------------------------------------------------------------------------------------------------------------------|--------|---------------------------------------------------------------------|
| Consulta:                                                                       | cibdad                                                                                                             |        |                                                                     |
| Criterios de selección:                                                         |                                                                                                                    |        |                                                                     |
| Autor:                                                                          |                                                                                                                    | Obra:  |                                                                     |
| Cronológico:                                                                    | 1501 1599                                                                                                          | Medio: | (Todos)<br>Libros<br>Periódicos<br>Revistas<br>Miscelánea<br>Orales |
| Tema:                                                                           | (Todos)<br>11.-Lírica<br>12.-Narrativa<br>121.-Breve<br>1211.-Relato breve tradicional<br>1212.-Relato breve culto |        |                                                                     |
| Geográfico: España<br>Filipinas<br>Guatemala<br>Honduras<br>México<br>Nicaragua |                                                                                                                    |        |                                                                     |
| [Buscar] [Limpiar]                                                              |                                                                                                                    |        |                                                                     |
| Consulta CREA    Nómina de autores y obras    Cómo citar el CORPUS    Ayuda.    |                                                                                                                    |        |                                                                     |

Figura 1. Búsqueda de la palabra *cibdad* en el CORDE de la RAE (siglo XVI en México)

<sup>1</sup> Una etiqueta, en el ámbito de la ingeniería lingüística, es una marca de XML con información lingüística.

Observamos en la figura 1 que el CORDE de la RAE permite hacer búsquedas por autor, obra, año, medio, región y tema. En el ejemplo seleccioné el periodo que corresponde al siglo XVI, marcando la cronología entre 1501 y 1599, y la ubicación geográfica correspondiente a México para hacer la búsqueda por palabra, como una muestra de léxico del siglo XVI, de *ciudad*.

A continuación (figura 2) presento los contextos en concordancias arrojados para esta búsqueda específica:



Figura 2. Resultados de la palabra *ciudad* en el CORDE

La siguiente figura (3) muestra también la búsqueda de la palabra *ciudad* en la interfaz que presenta el Corpus del Español de Mark Davies:



Figura 3. Resultados de la palabra *ciudad* en el Corpus del Español

En la figura 3 se presentan los resultados que arrojó el Corpus del Español para la palabra *ciudad*. Esta interfaz permite seleccionar el periodo que interesa al usuario (en este caso se eligió el periodo para los años 1500), pero no permite seleccionar zona geográfica y por ello mostró muchos más ejemplos para la palabra buscada (en comparación con los resultados del CORDE), pues arrojó ejemplos de todas las áreas de habla hispana incluidas en ese corpus.

En el CHEM, al ser un corpus exclusivo del español de México, sólo se muestran textos que hayan sido escritos en la Nueva España (siglos XVI, XVII y XVIII) y en el México del siglo XIX. Esta interfaz permite hacer búsquedas para los cuatro siglos de los que se compone el corpus, así como ampliar el contexto de las concordancias arrojadas; y, por el momento, sólo permite búsquedas por palabras completas.

Se presentan a continuación (figura 4) los resultados obtenidos por la búsqueda de la palabra *ciudad* para el español hablado en México en el siglo XVI.

The screenshot shows the CHEM search interface. At the top, it says 'CHEM Corpus Histórico del Español de México'. Below that, there are navigation links: 'Presentación', 'Notas', 'Participantes', 'Publicaciones', and 'Buscar en el CHEM'. The search form is titled 'Consulta de concordancias en el CHEM' and includes a search box with 'Palabra: ciudad' and a note '(Actualmente sólo es posible la búsqueda de palabras completas)'. There are radio buttons for 'Siglo: XVI', 'XVII', 'XVIII', and 'XIX', with 'XVI' selected. A 'Vencidas: 100 caracteres' indicator and a 'Buscar!' button are also present. Below the search form, the results are shown in a table with columns for 'Siglo', '#', 'Pal.', and 'Referencia'.

| Resultados de la búsqueda |   |                                                                                                                    |                                       |
|---------------------------|---|--------------------------------------------------------------------------------------------------------------------|---------------------------------------|
| Siglo                     | # | Pal.                                                                                                               | Referencia                            |
| XVI                       | 1 | ortante, a otro que es en la calçada que desta <b>ciudad</b> de Mexico sale a Tacuba, por ser al mejor y má        | Company C., DLNE, México, UNAM, 1994. |
| XVI                       | 2 | no está en más de tenerla esta <b>ciudad</b> , y esta <b>ciudad</b> no la puede tener sy esto no se haze, y muy de | Company C., DLNE, México, UNAM, 1994. |
| XVI                       | 3 | le llevado la principal joya de su casa. Desta <b>ciudad</b> de Mexico desta nuestra Nueva España, 29 de ha        | Company C., DLNE, México, UNAM, 1994. |

Figura 4. Resultados de la palabra *ciudad* en el CHEM

Los resultados sólo arrojan tres apariciones de la palabra *ciudad* en los documentos que integran el corpus, ya que esta forma, en el siglo XVI, alternaba gráficamente con *çibdad*, *ciudad* y con *çitudad* (véase §4.2.1.2) como veremos en las siguientes figuras (5, 6 y 7); sin embargo, sugiero que aunque alternara la forma gráfica, la pronunciación obedecía a una misma realización [sjuˈdád], aunque no debemos

descartar que en algunas situaciones comunicativas el hablante pronunciara, de hecho, [sib·dád] o [ʃib·dád].

The screenshot shows the search interface of the CHEM Corpus. The search term is 'ciudad', and the results table lists 24 entries. Each entry includes a row number, century (XVI), a snippet of text, the word 'ciudad', and a reference to a document.

| #  | Siglo | Pal.   | Referencia           |
|----|-------|--------|----------------------|
| 8  | XVI   | ciudad | Company C., DLNE, Mé |
| 9  | XVI   | ciudad | Company C., DLNE, Mé |
| 10 | XVI   | ciudad | Company C., DLNE, Mé |
| 11 | XVI   | ciudad | Company C., DLNE, Mé |
| 12 | XVI   | ciudad | Company C., DLNE, Mé |
| 13 | XVI   | ciudad | Company C., DLNE, Mé |
| 14 | XVI   | ciudad | Company C., DLNE, Mé |
| 15 | XVI   | ciudad | Company C., DLNE, Mé |
| 16 | XVI   | ciudad | Company C., DLNE, Mé |
| 17 | XVI   | ciudad | Company C., DLNE, Mé |
| 18 | XVI   | ciudad | Company C., DLNE, Mé |
| 19 | XVI   | ciudad | Company C., DLNE, Mé |
| 20 | XVI   | ciudad | Company C., DLNE, Mé |
| 21 | XVI   | ciudad | Company C., DLNE, Mé |
| 22 | XVI   | ciudad | Company C., DLNE, Mé |
| 23 | XVI   | ciudad | Company C., DLNE, Mé |
| 24 | XVI   | ciudad | Company C., DLNE, Mé |

Figura 5. Resultados de la palabra *ciudad* en el CHEM

The screenshot shows the search interface of the CHEM Corpus. The search term is 'ciudad', and the results table lists 16 entries. Each entry includes a row number, century (XVI), a snippet of text, the word 'ciudad', and a reference to a document.

| #  | Siglo | Pal.   | Referencia           |
|----|-------|--------|----------------------|
| 1  | XVI   | ciudad | Company C., DLNE, Mé |
| 2  | XVI   | ciudad | Company C., DLNE, Mé |
| 3  | XVI   | ciudad | Company C., DLNE, Mé |
| 4  | XVI   | ciudad | Company C., DLNE, Mé |
| 5  | XVI   | ciudad | Company C., DLNE, Mé |
| 6  | XVI   | ciudad | Company C., DLNE, Mé |
| 7  | XVI   | ciudad | Company C., DLNE, Mé |
| 8  | XVI   | ciudad | Company C., DLNE, Mé |
| 9  | XVI   | ciudad | Company C., DLNE, Mé |
| 10 | XVI   | ciudad | Company C., DLNE, Mé |
| 11 | XVI   | ciudad | Company C., DLNE, Mé |
| 12 | XVI   | ciudad | Company C., DLNE, Mé |
| 13 | XVI   | ciudad | Company C., DLNE, Mé |
| 14 | XVI   | ciudad | Company C., DLNE, Mé |
| 15 | XVI   | ciudad | Company C., DLNE, Mé |
| 16 | XVI   | ciudad | Company C., DLNE, Mé |

Figura 6. Resultados de la palabra *ciudad* en el CHEM

The screenshot shows the CHEM search interface. At the top, it says 'CHEM Corpus Histórico del Español de México'. Below that are navigation tabs: 'Presentación', 'Notas', 'Escribano', 'Publicaciones', and 'Escribanos en CHEM'. A search bar contains the word 'ciudad' and a 'Buscar' button. Below the search bar, there are radio buttons for centuries: XVI (selected), XVII, XVIII, and XIX. A character count shows 'Ventana: 100 caracteres'. The search results table has columns for '#', 'Siglo', 'Cub.', and 'Referencia'. The results list 13 entries, each with a century, a snippet of text, the word 'ciudad', and a reference number.

| #  | Siglo | Cub.   | Referencia                                                                                  |
|----|-------|--------|---------------------------------------------------------------------------------------------|
| 27 | XVI   | ciudad | de México, A.G.I.: Audiencia de México 280. Carta autógrafa de fray Miguel Navarro p...     |
| 28 | XVI   | ciudad | de México, A.G.I.: Indiferente General 2051. Carta autógrafa de Gonzalo Román pidi...       |
| 29 | XVI   | ciudad | de México, A.G.I.: Indiferente General 2051. Dos cartas autógrafas de Juan López de...      |
| 30 | XVI   | ciudad | de México, A.G.I.: Audiencia de México 280. Carta autógrafa de fray Juan de Córdoba...      |
| 31 | XVI   | ciudad | de México, A.G.I.: Indiferente General 2056. Carta autógrafa de Juan López Texera a...      |
| 32 | XVI   | ciudad | de México, A.G.I.: Indiferente General 2056. Carta autógrafa de Alonso López de Ald...      |
| 33 | XVI   | ciudad | de Sevilla. Muy magnífico sañor. La carta de v. m. recibí con Martín Hernández; y...        |
| 34 | XVI   | ciudad | de México, A.G.I.: Indiferente General 2056. Carta autógrafa de Alonso Ortiz pidiend...     |
| 35 | XVI   | ciudad | de México, A.G.I.: Indiferente General 2056. Segunda carta autógrafa de Alonso Ortiz...     |
| 36 | XVI   | ciudad | de México, A.G.I.: Indiferente General 2056. Carta de Rodrigo Muñiz a la esposa de...       |
| 37 | XVI   | ciudad | de México, A.G.I.: Indiferente General 2057. Carta autógrafa de Alonso Montaña, uno...      |
| 38 | XVI   | ciudad | de México, A.G.I.: Indiferente General 2057. Carta autógrafa de Francisco de Sosa in...     |
| 39 | XVI   | ciudad | de los Anxeles, porque yo acudiré luego con lo que fuere menester para pagar el Res...      |
| 40 | XVI   | ciudad | de México, A.G.I.: Indiferente General 2059. Carta autógrafa de Alonso González a s...      |
| 41 | XVI   | ciudad | de México, A.G.N.: Inquisición 127-1, ff. 78r-79r. Denuncia contra Diego Moreno por at...   |
| 42 | XVI   | ciudad | de México, A.G.N.: Inquisición 127-2, ff. 1r-2r. Denuncia de una mujer, natural de Méjic... |
| 43 | XVI   | ciudad | de México, veinte y siete días del mes de noviembre de mill y quinientos y setenta y...     |

Figura 7. Resultados de la palabra *ciudad* en el CHEM

En las figuras 5, 6 y 7 se muestra que la palabra *ciudad* alternaba gráficamente con *çibdad*, *çitudad* y *ciudad*, en el español hablado en México en el siglo XVI obedecía quizá a una misma pronunciación, como mencioné anteriormente, dependiendo de la situación comunicativa; estas tres últimas formas gráficas se presentan en el corpus, en 72, 59 y 65 ocasiones respectivamente.

La aportación final de esta investigación permitirá, en la interfaz del CHEM, relacionar estas tres formas gráficas según su similitud fonética a una sola búsqueda en cualquiera de las tres variantes. Es decir, que al hacer la búsqueda por palabra de *ciudad*, *çibdad*, *çitudad* o *ciudad*, se mostrarán resultados de todas las variantes gráficas encontradas en el corpus que correspondan a la transcripción fonética relacionada [sju·dád]. Este punto se explicará más ampliamente en §2.6.

Un dato importante que mencionar es que los corpus electrónicos están conformados por textos que han sido publicados por escritores, recopiladores, editores, etc., por lo que es importante el trabajo del autor de cada texto para la integración de un corpus electrónico; se debe atender, principalmente, al manejo cuidadoso de cada texto

por respeto a los derechos de autor. Más adelante (en §2.6) se presentará cómo está integrado el CHEM y mencionaré a los autores de los textos utilizados.

## **2.2. Sistemas de codificación de corpus en internet**

El CHEM, así como los corpus mencionados anteriormente son visibles a los usuarios de internet gracias a sus interfaces electrónicas. Una interfaz de internet requiere algunas etiquetas de codificación de texto para hacer que la información sea visible a los usuarios tal y como acceden a ella a través de las páginas electrónicas. Las etiquetas sirven, también, para marcar información lingüística (categorías gramaticales, lemas, transcripciones fonéticas o fonológicas, etc.) y, como ilustraré abajo, para desambiguación de dicha información.

La presentación de los corpus arriba mencionados se realiza mediante etiquetas a modo de marcas textuales con lenguajes de codificación como HTML (*Hyper Text Markup Language*), SGML (*Standard General Markup Language*) y XML (*eXtensible Markup Language*).<sup>2</sup>

El CHEM se mantiene en una base de datos XML lo que permite, en dado caso, transformarlo fácilmente a otros formatos (Medina y Méndez 2006). Esta marcación incluso acepta, dentro del marco de esta investigación, el almacenamiento de información, en el documento mismo, para desambiguar pares de palabras semejantes en su graficación. La desambiguación, en este caso fonológica, será de pares de palabras

---

<sup>2</sup> Estos lenguajes de codificación se usan para clasificar y manipular grandes cantidades de información. El origen de GML (*Generalized Markup Language*) fue gracias a la necesidad de la empresa IBM de almacenar su información en los años 60. A partir de este lenguaje, y gracias a su eficacia, se creó SGML en 1986, que se adaptó a muchos problemas. Desde la creación de la *web* en 1989 por Tim Berners y por la necesidad de una atractiva presentación de las páginas *web*, se hizo otro lenguaje de marcado llamado HTML que es el más conocido. HTML es un lenguaje de marcado muy popular, que predefine etiquetas. Algunos estudios indican que hay cerca de 800 millones de páginas *web* basadas en este lenguaje de marcado, como exploradores, correos electrónicos, administradores, entre otras. También existe XML que se creó para definir objetos en los documentos. Este lenguaje supera las limitaciones que tiene HTML al predefinir etiquetas, pues XML se adapta a las necesidades del usuario por medio de un lenguaje extensible (Marchal 2001).

semejantes gráficamente pero con diferente sonido, que el hablante puede desambiguar por contexto. En el corpus de trabajo, por ejemplo, se encontró el par de palabras *hoja* y *oja*, que corresponden a *hoja* y *oía*, respectivamente, según se pudo atestiguar por su contexto:

- a. la gente desta çiudad /<sup>20</sup> está sin sosiego, como la **hoja** en el arbol, a cabsa de lo que a *buestra majestad* he /<sup>21</sup> dicho (DLNE, 16)
- b. él la importunava a que tuviera /<sup>29</sup> açeso no save en qué forma, más de que este *testigo* /<sup>30</sup> **oja** cómo ella le dezia: "¡puto, dexame! (DLNE, 55)

Este par de palabras gráficas se codifican en XML de la siguiente manera (lo que constituye la desambiguación en sí):

- (1) a. <tok c="NCCS" f="ó·xa" l="hoja">hoja</tok>
- b. <tok c="VMII3S0" f="o·í·a" l="oír" n="oía">oja</tok>

En donde *tok* representa el nombre de la etiqueta XML utilizada para codificar palabras gráficas. La etiqueta tiene los siguientes atributos opcionales: **c**, **f**, **l** y **n** en los que se codifican, respectivamente, la categoría gramatical, la transcripción fonológica, el lema y, opcionalmente, una normalización (que está muy relacionada con la lista de excepciones o lista de filtrado que se mencionó antes; las excepciones para cada regla aparecerán, dado el caso, en su sonido correspondiente del capítulo 4).

En (1a) arriba, *hoja* tiene la categoría gramatical **c**, "NCCS" que corresponde, en el estándar EAGLES<sup>3</sup>, a nombre común, género común, singular. Mientras que en (1b), *oja* tiene la categoría "VMII3S0" que corresponde a verbo principal, imperfecto de indicativo, 3ª persona singular, sin marca de género. El atributo **f** representa, en ambas instancias, la transcripción fonológica asociada y el atributo **l** corresponde al lema (la forma canónica). Todo esto corresponde a la desambiguación, con relación a su categoría léxica; es decir, cuando se trate de verbo lo transcribirá como /o·í·a/, y

---

<sup>3</sup> *Expert Advisory Group on Language Engineering Standards*, cuya página electrónica se encuentra disponible en <http://www.ilc.cnr.it/EAGLES/home.html>.

exhibirá el lema *oír*, y cuando se trate de sustantivo como /ó·xa/, que tendrá como lema *hoja*. Finalmente, el atributo **n** consigna una normalización que permite descifrar la forma original atestiguada en el contexto original y que también podrá contribuir a la desambiguación de dicha forma.

Como se vio, estas etiquetas contienen información lingüística que se añadirá a cada una de las palabras del corpus, de tal suerte que el CHEM tendrá etiquetadas sus palabras con datos lingüísticos. Entre estos se contará con la transcripción fonética asociada a cada palabra y, mediante este dato, se podrán ver las variantes gráficas de cada una de ellas, como se mencionó en §2.1., en las búsquedas que se hagan en la interfaz de internet.

### **2.3. Caracteres de codificación: ASCII y Unicode**

Todos los caracteres que se utilizan en el trabajo de la computadora tienen códigos específicos asignados para que esta pueda procesar la información. Me refiero a los sistemas estándares de codificación como ASCII y Unicode.<sup>4</sup>

Esto es importante porque nos indica que, así como a cada letra que escribimos en un texto se le asigna una codificación para que la computadora pueda procesarla y hacerla presentable a nuestros ojos; de la misma manera, el transcriptor automático requerirá de la codificación de diferentes símbolos para cada fonema transcrito. Ya que el teclado de una computadora no contiene símbolos fonéticos, (por ser estos especiales y porque sólo un sector de la población los utiliza, ya que son de difícil manejo), se

---

<sup>4</sup> El código ASCII (*American Standard Code for Information Interchange*) se creó en los años 60 como una evolución de los códigos usados en telegrafía. Así, se implementó su utilización en el mundo computacional y se asignó una codificación que especifica una cadena de *bits* y símbolos alfanuméricos para procesar y almacenar información, así como para la comunicación entre medios digitales. Más adelante se creó Unicode (que contiene en sí mismo a ASCII) como un estándar industrial que asigna un único punto de código para cada carácter. Esto es, que para un mismo carácter se asignan diferentes códigos según ASCII o Unicode. Estos códigos no añaden información extra al texto, pues esto lo hacen las etiquetas de marcado de lenguajes como HTML y XML de los que se habló en la nota 4 (Wikipedia. 2007. *Código ACSII y Unicode*).

utilizan estos sistemas de codificación para hacer transcripciones fonético-fonológicas. Por ejemplo, a un símbolo fonético como [s], se le asignará un código especial y a un símbolo como [ʃ] se le asignará otro.

Unicode y ASCII por sí mismos contienen caracteres de representación fonológica, pero no permiten realizar transcripciones finas, ya que carecen de muchos caracteres necesarios para este tipo de investigaciones. Para ello se crearon, con base en estos códigos, alfabetos fonéticos que permiten hacer análisis finos en el área de fonética y fonología, estoy hablando de alfabetos como AFI, RFE y Mexbet que explicaré a continuación.

#### **2.4. Caracteres de representación fonética: RFE, AFI, Mexbet**

La *Revista de Filología Española* (RFE) creó su propio alfabeto fonético por las insuficiencias que presentaba el Alfabeto Fonético Internacional (AFI) de la *International Phonetic Association* para representar los múltiples alófonos del español (Navarro 1966-1967:7).

El alfabeto de la RFE se creó con símbolos base y con diacríticos para representar las diferentes producciones de los hablantes y fue de gran utilidad al hacer análisis fonéticos finos, como en el área de dialectología.

En la presente investigación utilizo el alfabeto de la RFE, ya que es el alfabeto utilizado por la comunidad de la tradición hispanista y empleado en la gran mayoría de los textos consultados para la realización de esta tesis.

En el Apéndice 1 hago una correlación entre el RFE y el AFI. Asimismo, se encontrará, en ese apéndice, una correlación entre los alfabetos mencionados y Mexbet, desarrollado por Cuétara como parte de su tesis de maestría en 2004, ya que el transductor automático que se elaborará arrojará las transcripciones en los tres alfabetos.

Mexbet tiene como antecedente varios alfabetos que contenían algunos errores por influencia de lenguas como el inglés (por ejemplo, clasificaban como semivocales a las líquidas) y de dialectos como el castellano (incluían la pronunciación de /j/). Cuétara mejoró los alfabetos previos y propuso un sistema de diacríticos compatibles con los sistemas de codificación computacionales. Por ejemplo, marca como a<sub>2</sub> un alófono de /a/ velarizado. Sin embargo, también mantuvo algunos signos de tradición, como B, D, G, para los alófonos de /b/, /d/, /g/ fricativizados (Cuétara 2004).

La finalidad de Cuétara al desarrollar este alfabeto fue “contar con un alfabeto específico para el español de la ciudad de México que [pudiera] aplicarse a las tecnologías del habla, buscando conciliar en la medida de lo posible los principios de los alfabetos de la AFI y de la RFE” (Cuétara 2004:70).

La aplicación final relacionada con esta investigación (Fonseca, en proceso) proporcionará una correlación entre AFI, RFE, Mexbet y los caracteres de codificación explicados arriba (ASCII y Unicode) que utiliza la computadora para manejar y procesar la información.

## **2.5. *TranscribEMex*: antecedente de la aplicación**

*TranscribEMex* (Transcriptor fonético automático para el español de México) es una herramienta computacional que realiza transcripciones automáticas a nivel fonético y fonológico. Fue desarrollada por Cuétara y Villaseñor (2004) para un proyecto del IIMAS de la UNAM en el lenguaje computacional conocido como *Perl*.

Esta herramienta presenta, además de las transcripciones en los niveles fonético y fonológico, una segmentación silábica de cada palabra. Puede transcribir desde palabras y oraciones hasta archivos de texto completos.

*TranscribEMex* es el antecedente de la aplicación final que desarrollará Fonseca porque este programa no puede, por sí solo, adaptarse a las necesidades que requiere el CHEM. Las reglas funcionales de transcripción para *TranscribEMex* fueron desarrolladas para el español de la ciudad de México del siglo XX, y, como veremos más adelante (en §2.6.), el CHEM se compone por textos del español de México (de varios estados de la República) que abarcan los siglos XVI al XIX, por lo que las reglas de transcripción del *TranscribEMex* no son apropiadas para estos periodos previos de la lengua española.

En la figura 8 podemos ver cómo funciona *TranscribEMex* para un texto del siglo XVI a partir del sintagma *Sacra çesárea catholica majestad*:

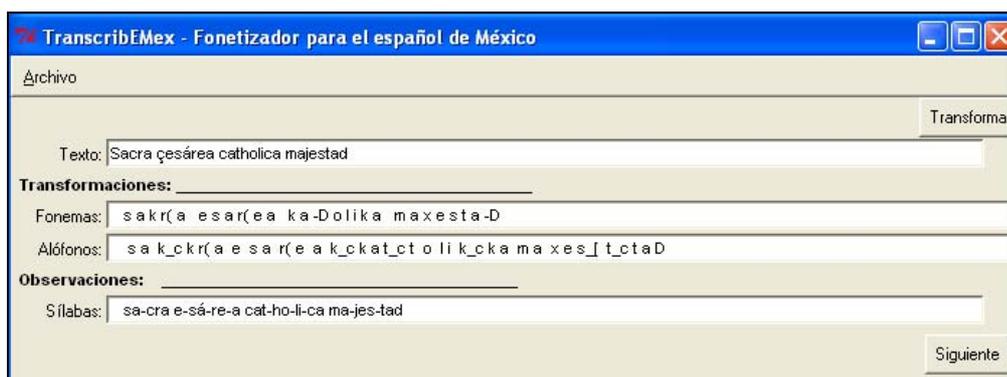


Figura 8. Transcripción de *Sacra çesárea catholica majestad* en *TranscribEMex*

Es evidente que la interfaz de *TranscribEMex* no funciona para textos del siglo XVI. En la figura 8 tenemos un ejemplo con el sintagma *Sacra çesárea catholica majestad*, fórmula que se encuentra abundantemente en los documentos del CHEM, y que contiene una serie de fenómenos frecuentes en el español hablado en México en el siglo XVI, los problemas que muestra *TranscribEMex* al transcribir esta frase son:

- El programa no lee la grafía ç, así que simplemente la ignora y no la transcribe.
- El dígrafo *th*, no lo transcribe como /t/, ya que lo segmenta erróneamente (*cat-hó-li-ca*) y presenta *t* como archifonema /D/.

En la figura 9 vemos otro fragmento de texto del español del siglo XVI que corresponde a la frase *no puedo dexar de hazer lo que soy oblygado* y así se ve en *TranscribEMex*.

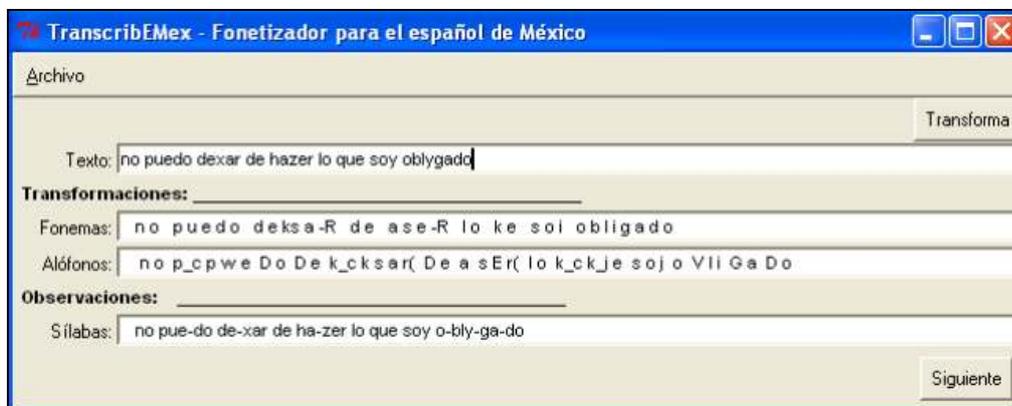


Figura 9. Transcripción de *no puedo dexar de hazer lo que soy oblygado* en *TranscribEMex*

En este ejemplo (figura 9) vemos que este programa transcribe la grafía *x* de *dexar* como /ks/ de *examen*. Esta grafía presenta complicaciones tanto para el español actual cuanto más para el español hablado en el siglo XVI, ya que en ese entonces correspondía al sonido prepalatal<sup>5</sup> fricativo sordo [ʃ] que se encontraba en una etapa avanzada de velarización; también correspondía, en algunos casos ya, al velar, o se usaba, como actualmente, para reflejar la secuencia [ks].

En la actualidad, la letra *x* representa un problema en sí, ya que para el español puede corresponder a varios sonidos, es decir, puede simbolizar [x], [s], [ks] y [ʃ]. *TranscribEMex* no resolvió este problema, pues todas las grafías *x* las transcribe como /ks/. Para hacer las transcripciones correctas, (al ser esta una grafía que no depende del contexto para tener una correspondencia fónica), se debió haber establecido una lista de excepciones, o de filtrado, para asignar a cada palabra su correcta transcripción fonológica.

<sup>5</sup> Algunos investigadores utilizan el término postalveolar o prepalatal para marcar la zona de articulación de este sonido [ʃ], al igual que la del sonido africado [ʃ]. Algunos otros prefieren utilizar el término dorsopalatal para resaltar el modo de articulación de la lengua en la zona palatal. A lo largo de esta investigación me referiré a ambos sonidos como prepalatales.

Los anteriores son algunos de los problemas que plantea este transcriptor al trabajar fragmentos de textos del siglo XVI, lo cual es lógico ya que su finalidad es transcribir el habla de la ciudad de México del siglo XX.

Con base en las reglas que arrojará el presente análisis para cada grafía observada en los documentos del español hablado en México en el siglo XVI, se hará un nuevo transcriptor fonológico en el lenguaje de programación *Java*. Este transcriptor primero transcribirá las palabras contenidas en la lista de filtrado de la que he hablado antes (Apéndice 3) y después ejecutará de manera automática las reglas obtenidas en este análisis.

Posteriormente, partiendo de esta investigación y con el transcriptor funcionando, se podrán modificar las reglas y los sonidos para cada periodo de los que está conformado el CHEM y, así, transcribir fragmentos de cualquier periodo del corpus.

## **2.6. El Corpus Histórico del Español en México: necesidades**

El *Corpus Histórico del Español en México*<sup>6</sup> (CHEM) es un proyecto del Instituto de Ingeniería (II), dirigido por el Dr. Alfonso Medina Urrea como parte de los trabajos del Grupo de Ingeniería Lingüística (GIL).

El objetivo central del proyecto es “constituir un corpus diacrónico del español de México que pueda ser utilizado por investigadores del lenguaje y cultura de la Nueva España y el México del siglo XIX, mediante una interfaz de internet” (Medina y Méndez 2006:248).

---

<sup>6</sup> Proyecto DGAPA-UNAM, PAPIIT IN400905, “Constitución del *Corpus Histórico del Español en México* (CHEM)”.

El corpus está conformado por textos de los siglos XVI, XVII, XVIII y XIX. Son de diferentes temáticas y de diversos orígenes diatópicos. Para el periodo del siglo XVI el corpus está conformado por tres textos: *Cartas de Diego de Ordaz* (Lope Blanch 1998), *Documentos lingüísticos de la Nueva España. Altiplano Central* (Company 1994) y *Los procesos inquisitoriales contra indígenas que realizó Fray Juan de Zumárraga en Nueva España (1536-1543)* (Buelna 2006)<sup>7</sup>.

El siglo XVII lo conforman: los *Documentos Lingüísticos de la Nueva España. Altiplano Central (DLNE)*. Para el siglo XVIII se cuenta con: los *DLNE*, el *Mercurio volante* (Bartolache 1979), *Obras I-Periódicos* (Alzate 1980) y el *Theatro americano* (Villa-Señor 1746). Y el siglo XIX está conformado por: los *DLNE*, un *Recetario* (1831), y varias constituciones federales y de los estados.

Es importante mencionar que el CHEM es un corpus abierto, por lo que en el transcurso del proyecto se le incorporarán nuevos documentos, se ampliará su información lingüística y se mejorará su interfaz. Cabe enfatizar que la futura inclusión de nuevos documentos, seleccionados metódicamente para llenar vacíos de áreas geográficas y temáticas (principalmente), garantizará el acercamiento del repositorio con el que se cuenta a los ideales de representatividad y equilibrio que todo corpus debe buscar satisfacer.

Comparado con los corpus mencionados en §2.1. el CHEM se diferencia de los corpus de la RAE porque es un corpus exclusivo para México y para los siglos en los que empezó a existir la lengua española en América (específicamente en México) hasta el siglo XIX.

Comparado con el Corpus del Español de Mark Davies, el CHEM se propone llevar a cabo la consideración de similitudes fonológicas entre palabras gráficas

---

<sup>7</sup> Véase el §4.1. y la bibliografía de esta tesis.

diferentes y etiquetar el corpus. Con respecto a las dos cuestiones anteriores, para el Corpus del Español, *vivir* y *bibir*, por ejemplo, son palabras diferentes, mientras que el CHEM hará la relación entre estas dos formas con base en su similitud fonológica, lo que constituye la aportación principal de esta tesis. Por otra parte, el Corpus del Español de Mark Davies, etiqueta n-gramas<sup>8</sup> en una base de datos y no el corpus en sí.

Para el estudio del español mexicano se encuentran problemas, “algunos de ellos son la escasa documentación sobre el dialecto mexicano en siglos previos, la falta de normas ortográficas que impiden identificar rasgos lingüísticos consistentes y el manejo de caracteres que hoy en día ya no se utilizan” (Medina y Méndez 2006:248).

Gracias a este corpus se podrán realizar estos estudios con base en la información lingüística que se ha estado añadiendo al mismo. Medina y Méndez afirman que:

Muchas investigaciones lingüísticas descansan bajo la premisa de que los datos empíricos que proporcionan los corpus son la mejor manera de conocer diversos fenómenos del lenguaje. Este es el caso, en especial, de los estudios diacrónicos, para los que no se puede contar con los conocimientos y la intuición de hablantes de los estados de lengua que se comparan (2004:5).

Para el caso específico de la investigación que me atañe, no se cuenta con datos reales de hablantes (grabaciones) por razones obvias, así que sólo disponemos de las vacilaciones gráficas de las palabras para la representación de los sonidos.

La investigación realizada en esta tesis constituye una de las aportaciones que se implementarán a la interfaz del CHEM para su mejoramiento. Algunos otros trabajos dentro de este proyecto son: un etiquetador de categorías gramaticales (o partes de la oración, calco del término POS, por sus siglas en inglés *parts of speech*) en realización por Carlos Méndez, de la Maestría en Lingüística Hispánica; el transductor ortográfico-fonológico que está en proceso por Carlos Fonseca; un sintetizador de voz hecho por el

---

<sup>8</sup> Dentro del ámbito de la ingeniería lingüística y en un nivel sintáctico, un n-grama corresponde a una palabra gráfica.

equipo de Abel Herrera, en la Facultad de Ingeniería, entre otros. Habrá que mencionar también la implementación de estadísticas en el CHEM para hacer búsquedas por palabras asociadas (por ejemplo, los constructos verbo-nominales, como  *echar ojo*  y  *echar aguas* , o formas asociadas con otras como  *sal*  y  *pimienta*  o  *sal*  y  *azúcar* ) y las frecuencias de aparición.

Son muchas las aportaciones que se pueden hacer a un corpus abierto; entre las necesidades del CHEM se encuentran: hacer las reglas de correlación entre sonido y grafía para los demás siglos y ver la cronología de los cambios fonéticos que ocurren en el español de México, hacer la lematización de las palabras para asignar el atributo lema de la etiqueta  *tok*  que se mencionó en §2.2. Etiquetar manualmente categorías gramaticales para entrenar el programa que lo hará automáticamente. Se necesitará aumentar el inventario de documentos y adecuarlos al formato del CHEM.

Además serán útiles otros estudios lingüísticos; por mencionar algunos, se planea hacer la clasificación del  *se*  y la ampliación de su funcionalidad en el español de México, la clasificación de la morfología del español de México a través de los documentos, el ordenamiento de afijos o afitáctica, estudios sobre léxico indígena, el uso del léxico dentro del tema jurídico, o culinario, etcétera.

Como vimos en este capítulo, el CHEM, al ser un corpus específico de México, necesita estudios especiales para el español mexicano. A través de las reglas que arroje esta investigación y que finalmente se implementarán a un transcriptor (aparte de la transcripción fonética correcta a las palabras de la lista de filtrado), será posible que en la interfaz de internet del corpus se asocien por su sonido palabras gráficas diferentes y, de esta manera, aparecerán todas las formas asociadas en las búsquedas en el CHEM como se ilustró en §2.2. Para hacer las correspondencias entre grafía y sonido no basta saber en qué contextos aparece determinada grafía, sino asociar a dicha grafía su sonido

correspondiente; por ello se abordarán los temas del español asentado en México en el siglo XVI en el siguiente capítulo.

### **3. FENÓMENOS FONÉTICOS GENERALES EN LA HISTORIA DEL ESPAÑOL DE AMÉRICA**

*Todo fenómeno pretérito de evolución fonética se presta siempre a diversas interpretaciones históricas. Casi todas parecen ser posibles, pero todas resultan, en último caso, indemostrables. La historia es en un elevado porcentaje recreación, suposición.*  
Salvador, 1982.

En este apartado veremos tres temas fundamentales para la configuración del español de México del siglo XVI. En primer lugar, la transición del castellano medieval al español moderno dentro de España, pues esta se da de diferentes maneras según la región de la que se trate, y como veremos en §4.1., este hecho determina la vacilación gráfica presentada en los textos según la región originaria de cada amanuense o testigo. Después, abordaré el tema del español trasplantado a América, cómo se da y de qué manera influye en la pronunciación americana. Finalmente, veremos cómo se presenta el andalucismo en América y cómo este permea los sonidos con los que se hablaba en México hacia el siglo XVI.

#### **3.1. Cambios fónicos del español medieval al moderno**

Los cambios más complejos en el sistema fónico del español se dan en el sistema consonántico. Podemos apreciar que en el cambio del latín al castellano antiguo y de este al español moderno, las consonantes son las que sufren mayores transformaciones para llegar a su estatus actual.

Como veremos a continuación, en el sistema fonológico latino (figura 10) no hay fonemas en la zona palatal del espacio articulatorio, en cambio, en el sistema fonológico del castellano medieval (figura 11) proliferan los fonemas en esta zona de articulación

(cf. §4.3.3. al §4.3.5.). Para el paso del sistema fonológico del castellano medieval (figura 11) al español moderno (figura 12) podemos ver un ajuste en esta zona de articulación y, consecuentemente, la pérdida de algunos fonemas palatales y la conservación de otros. Veamos las tres figuras siguientes en las que se muestran los sistemas fonológicos del latín (figura 10), del castellano medieval (figura 11) y del español moderno (figura 12).

En la figura 10 (Lloyd 1987, Penny 1993/2001) se aprecia que el sistema fonológico consonántico latino era muy sencillo, observamos que no existen fonemas africados y que tampoco hay fonemas en la zona palatal del espacio articulatorio, (faltan en la figura los correspondientes geminados; que son fonemas de cantidad larga, por ejemplo /p:/ corresponde a una doble *p* gráfica: SUPPLICARE > *suplicar*).

En cambio, en el sistema fonológico del castellano medieval (Lloyd 1987, Penny 1993/2001), figura 11, hay un incremento en los fonemas, pues, según la zona de articulación, surgen los que se articulan en el paladar y, según su modo de articulación, los africados.

En la modalidad atlántica del español moderno (Quilis 1993/1999), figura 12, vemos que algunos de los fonemas creados en romance se mantienen, pero otros desaparecen. Justamente, es el reajuste de los fonemas de la zona palatal, lo que veremos en la presente tesis, pues es en el siglo XVI cuando se consolidan los cambios que venían gestándose durante los siglos XIV y XV (algunos aun antes), y se nivela el sistema fonológico actual. Para García de Diego “la ortografía española no ha llegado a su admirable equilibrio de fonetismo y tradición y a su admirable sencillez y fijeza sino después de vicisitudes seculares” (1970:52), lo cual se refleja en la diferencia de la cantidad de fonemas que presentan los tres siguientes sistemas fonológicos consonánticos.

|                    | labiales | labiodentales | dentales | alveolares | prepalatales | palatales | velares | glotales |
|--------------------|----------|---------------|----------|------------|--------------|-----------|---------|----------|
| Oclusivas sordas   | p        |               | t        |            |              |           | k       |          |
| Oclusivas sonoras  | b        |               | d        |            |              |           | g       |          |
| Africadas sordas   |          |               |          |            |              |           |         |          |
| Africadas sonoras  |          |               |          |            |              |           |         |          |
| Fricativas sordas  |          | f             |          | s          |              |           |         | h        |
| Fricativas sonoras |          |               |          |            |              |           |         |          |
| Nasales            | m        |               |          | n          |              |           |         |          |
| Vibrantes          |          |               |          | r          |              |           |         |          |
| Laterales          |          |               |          | l          |              |           |         |          |

Figura 10. Sistema fonológico latino<sup>1</sup>

|                    | labiales         | labiodentales | dentales | alveolares | prepalatales | palatales | velares | glotales |
|--------------------|------------------|---------------|----------|------------|--------------|-----------|---------|----------|
| Oclusivas sordas   | p                |               | t        |            |              |           | k       |          |
| Oclusivas sonoras  | b                |               | d        |            |              |           | g       |          |
| Africadas sordas   |                  |               |          |            | ʃ̂           | ç̂        |         |          |
| Africadas sonoras  |                  |               |          |            | ʒ̂           |           |         |          |
| Fricativas sordas  |                  | f             |          | s          | ʃ̃           |           |         | h        |
| Fricativas sonoras | [b] <sup>2</sup> |               |          | z          | ʒ̃           | y         |         |          |
| Nasales            | m                |               |          | n          |              | ɲ̃        |         |          |
| Vibrantes          |                  |               |          | r / r̄     |              |           |         |          |
| Laterales          |                  |               |          | l          |              | ɭ̃        |         |          |

Figura 11. Sistema fonológico del castellano medieval

<sup>1</sup> Nótese que se han anexado filas para fonemas africados y columnas para realizaciones en la zona palatal. Aunque el sistema consonántico fonológico latino no poseía fonemas con estas características y estas posiciones se encuentran vacías, es pertinente para que el lector pueda observar que la zona palatal fue la que tuvo mayor movimiento en el cambio del sistema fonológico del castellano medieval al español moderno, como se observará en las dos figuras siguientes.

<sup>2</sup> Este sonido se vio envuelto en una polémica acerca de su estatus fonológico. Algunos estudiosos, como veremos en §3.1.2., aseguran que nunca hubo diferenciación de sonido entre las grafías *b* y *v*, otros opinan que la única diferencia era el modo de articulación [b] oclusivo y [β] fricativo, ambos labiales, y que la diferenciación se perdió hasta el siglo XVI.

|                    | labiales | labiodentales | dentales | alveolares | prepalatales     | palatales        | velares | glotales |
|--------------------|----------|---------------|----------|------------|------------------|------------------|---------|----------|
| Oclusivas sordas   | p        |               | t        |            |                  |                  | k       |          |
| Oclusivas sonoras  | b        |               | d        |            |                  |                  | g       |          |
| Africadas sordas   |          |               |          |            |                  | ç                |         |          |
| Africadas sonoras  |          |               |          |            |                  |                  |         |          |
| Fricativas sordas  |          | f             |          | s          | (š) <sup>3</sup> |                  | x       |          |
| Fricativas sonoras |          |               |          |            |                  | y                |         |          |
| Nasales            | m        |               |          | n          |                  | ɲ                |         |          |
| Vibrantes          |          |               |          | r / ̄r     |                  |                  |         |          |
| Laterales          |          |               |          | l          |                  | (ʎ) <sup>4</sup> |         |          |

Figura 12. Sistema fonológico del español moderno (modalidad atlántica)

Ahora bien, después de confrontar los sistemas fonológicos de los tres estados de lengua (latín, castellano medieval y español moderno) y ver las diferencias que se perciben de manera evidente, explicaré cómo fue el proceso de cambio entre el castellano medieval y el español moderno<sup>5</sup> en España y de qué manera se verá reflejado en el español americano y en los textos del corpus de trabajo.

Los principales cambios que ocurren son:

- El ensordecimiento de los fonemas prepalatal africado sonoro /ʒ̄/, prepalatal fricativo sonoro /ž/ y alveolar fricativo sonoro /z/ en /š/, /š̄/ y /s/, respectivamente (Lapesa 1981, Lloyd 1987).
- La confusión del fonema prepalatal africado sordo /ʒ̄/ con el alveolar fricativo sordo /s/, en la modalidad atlántica del español (Alvar 1982).

<sup>3</sup> Veremos más adelante §4.4.1. que este fonema se comporta como alófono de /s/.

<sup>4</sup> Se conserva como dialectalismo en muy pocas zonas americanas.

<sup>5</sup> La transición del sistema fonológico latino al castellano medieval es muy interesante pero no atañe su explicación en esta tesis.

- La velarización del fonema prepalatal fricativo sordo /ʃ/<sup>6</sup> → /x/ (Menéndez Pidal 1982).
- La total confusión del fonema bilabial oclusivo sonoro /b/ y el bilabial fricativo sonoro /β/ (Alatorre 1979/2003, Lathrop 1984/1992).
- La pérdida del fonema labiodental fricativo sordo /f/ → Ø en algunas voces de origen latino (Pensado 1984, Menéndez Pidal 1986).
- La confusión del fonema palatal lateral /ʎ/<sup>7</sup> con el palatal central /y/ (Parodi 1995).

### 3.1.1. Reajuste de sibilantes

Sobre el proceso de ensordecimiento de las prepalatales africada y fricativa /ʧ/ y /ʝ/ y la alveolar fricativa /z/; la confusión del fonema prepalatal africado sordo /ʃ/ con el alveolar fricativo sordo /s/ y la velarización del fonema prepalatal fricativo sordo /ʃ/ → /x/, hay diferentes posturas, como veremos enseguida.

Algunos estudiosos, como veremos en seguida, proponen que el ensordecimiento de los fonemas sonoros ocurrió primero y después modificaron su zona de articulación, en el caso de /ʃ/ → /x/, o su zona y modo de articulación, en el caso de la confusión de /ʃ/ → /s/. Otros lo proponen exactamente al revés, es decir, que primero modificaron su zona y/o modo de articulación y después se ensordecieron. Veamos.

André Martinet en su *Economía de los cambios fonéticos* (1955/1974:421) se sorprende por la sencillez del sistema fonológico del español actual, pero dice que para

<sup>6</sup> Este fonema se conserva, en el español de México, en algunas voces de origen náhuatl, aunque casi no tiene rendimiento fonológico sino que se comporta como variante alofónica de /s/ (cf. Lope Blanch 1972:94).

<sup>7</sup> Se conserva muy esporádicamente, en algunos dialectos americanos.

ello el español tuvo que pasar por varios cambios fonológicos como los de las sibilantes y sitúa en la segunda mitad del siglo XVI y primeras décadas del XVII la confusión de los tres fonemas sonoros /ʒ/, /ʒ̃/ y /z/ con sus correspondientes sordos. Este autor diferencia el ensordecimiento y el desplazamiento de las articulaciones pero no por ello deja de clasificarlos como fenómenos concomitantes.

También Antonio Alatorre (1979/2003:305, 306) acepta que el ensordecimiento ocurrió antes que el cambio de punto y modo de articulación para estos fonemas. Ubica el fenómeno hacia 1540 como un cambio que ocurrió de norte a sur, de manera que cuando en Burgos ya habían adoptado la forma innovadora *açer* /atsér/, en Toledo y en las provincias del sur aún pronunciaban *hazer* /xadsér/. La igualación de sordas y sonoras iniciada en el norte se generalizó en la segunda mitad del siglo XVI por el resto de España. Hacia 1540, la pronunciación innovadora de Burgos comenzó a extenderse también por Andalucía y de esta provincia se propagó hacia el Nuevo Mundo.

Por esta misma línea de investigadores se encuentra Rafael Lapesa (1981:373), quien también asegura que este fenómeno va de norte a sur. Así, indica que fue irradiado desde Aragón y Castilla la Vieja y que después se expandió por Toledo, Extremadura, Murcia, Andalucía y América, sin dejar más que algunos reductos dialectales.

Señala los cambios en el modo y zona de articulación como posteriores: “el aflojamiento de las africadas en fricativas, atestiguado en la Andalucía occidental desde principios del siglo XV, se produjo también en el Norte y meseta septentrional con independencia respecto al fenómeno andaluz y probablemente con posterioridad a él” (Lapesa 1981:373). Atestigua el fenómeno desde principios del XV pero este no se generaliza hasta el siglo XVII.

Catalán (1982:103) propone que es una práctica castellano-vieja la de ensordecer las sibilantes sonoras y que finalmente se adoptó en el habla descuidada de Madrid y Toledo en la segunda mitad del siglo XVI. Se impuso, pues, en el habla cortesana de Felipe II, y de allí adquiere estatus social, por ello se extiende hacia otras provincias, entre ellas Valencia y Sevilla.

Durante el siglo XVI y a principios del XVII se trata, por parte de algunos puristas, de hacer las distinciones pertinentes entre sonidos, es decir, respetar la ortografía, pero sólo es un afán normativo de utilizar *bien* la lengua. Esto es sólo un intento pues la confusión de grafías estaba totalmente establecida desde finales del siglo XVI como señal de que la gente común y corriente confundía sordas con sonoras, Catalán menciona que esta práctica “fue imponiéndose en el habla descuidada de la mayoría a lo largo de esta segunda mitad de siglo, hasta tal punto que, en el último cuarto del siglo XVI, había triunfado incluso entre los más letrados, con la consiguiente alarma de gramáticos y preceptistas (1982:103).

Por su parte, Lloyd (1987:330) también propone que el ensordecimiento ocurrió primero: “as late as the last decade of the sixteenth century, there is evidence of a maintenance of the phonemic distinction between /s/ and /z/ in parts of the south [de España], although it was undoubtedly archaic then” y después el cambio de modo de articulación: “by the end of sixteenth century even many conservative speakers had adopted the fricative pronunciation” (Lloyd 1987:334). Sin embargo, no descarta la posibilidad de que estos cambios ocurrieran al mismo tiempo, menciona que empezaron en el siglo XIII y finalmente triunfaron en el XVI.

Para Cano (1988:239) empiezan primero las confusiones de sonoridad y después el reajuste de punto y modo de articulación en las sibilantes. También menciona que el fenómeno empieza por el norte y el centro de la Península y que después se propaga por

el sur. Explica que “tras el ensordecimiento de sibilantes y palatales, otros dos fenómenos vinieron a ordenar las distinciones entre los fonemas que quedaron, con el objeto de evitar las confusiones entre unos y otros” (Cano 1988:239); es decir, ocurrió una refonologización<sup>8</sup> consistente en la velarización de /ʃ/ → /x/ y, en Castilla, la refonologización de /ʂ/ → /θ/. Respecto a esta refonologización, recordemos que en el español atlántico resultó una confusión (o desfonologización)<sup>9</sup> con el fonema alveolar fricativo sordo /s/ que ya existía en latín.

Otra voz que se suma a señalar antes el ensordecimiento que el cambio de articulación de estos fonemas es la de Cabrera (1992:4, 5) y ubica su origen incluso en la segunda mitad del siglo XIV. Sin embargo, la definitiva consolidación del cambio la reporta a finales del siglo XVI. Se agrega a la teoría que afirma que el fenómeno se propagó desde el norte de la Península, pues la idea ya había sido sugerida por lingüistas como Cuervo y respaldada por autores como Alonso y Menéndez Pidal.

Lope Blanch (2000:204-205) apunta que “el ensordecimiento de la apicoalveolar fricativa sonora /z/ (escrita como -s- sencilla en posición intervocálica) y su consiguiente confusión con la sorda correspondiente /s/ (ortográficamente -ss- o s- inicial) se documenta desde los comienzos mismos del siglo XV, y era ya relativamente común a finales de esa centuria”. Y esto lo confirma en su análisis de las *Cartas de Diego de Ordaz*, pues el conquistador sólo utiliza la grafía *s*. Las *Cartas*, que se utilizaron en esta investigación, reflejan en gran medida el habla del español usado en México en el siglo XVI y, por ende, aportarán importantes datos para este análisis.

---

<sup>8</sup> Jakobson (1971:115) define la refonologización como “la transformación de una diferencia fonológica en una diferencia fonológica heterogénea que se encuentra frente al sistema fonológico en una relación distinta de la primera”.

<sup>9</sup> Para Jakobson (1971:108), una desfonologización es “la supresión de una diferencia fonológica”.

Blanco (2006:85, 86) indica que el reajuste de sibilantes es un fenómeno tardío dentro del siglo XVI, ubica el proceso de pérdida de sonoridad en el norte de la Península, con un avance significativo (es decir, era frecuente) hacia mediados del siglo y generalizado hacia el XVII. La pérdida de oclusión la documenta, en las sonoras, hacia mediados del siglo XVI y en la sorda hasta el XVII. Por lo que podemos intuir que eran fenómenos que convivían. Documenta la velarización del fonema prepalatal fricativo sordo /ʃ/ hacia finales del siglo XVI y cumplido hasta bien entrado el siglo XVII.

Hasta ahora hemos visto las investigaciones que sustentan que el ensordecimiento de los fonemas sonoros ocurrió antes que el cambio en el punto y modo de articulación de los mismos. A continuación presento a los autores que defienden la postura contraria, es decir, que plantean que primero hubo un cambio de correlación entre los fonemas y que después se ensordecieron.

Amado Alonso (1955/1976a:314) propone, con base en el análisis de documentos, que el cambio de africadas a fricativas sucedió alrededor de la segunda mitad del siglo XVI y que hasta el último tercio de ese mismo siglo sufrieron el proceso de ensordecimiento. Sin embargo, menciona que *ç*, /ʃ/, permaneció africada hasta el siglo XVII. Podemos intuir que la conservación de este fonema se puede deber a su revitalización en léxico americano de lenguas nativas o a fenómenos de hipercorrección.

Menéndez Pidal (1904/1940:113) también comenta que a partir del siglo XVI estos fonemas habían cambiado de africados a fricativos, y menciona la aparición del fonema interdental /θ/ para el español castellano y la velarización de /ʃ/ → /x/. Sitúa el ensordecimiento en los comienzos del siglo XVII.

De igual manera, Alvar (1982:132) propone primero la pérdida de oclusión y después el ensordecimiento. Así, nos explica el cambio de correlaciones entre estos fonemas y también la forma en la que se produjo el seseo andaluz:

El sistema medieval castellano con sus dos pares de sibilantes (*s* [z] fricativa sonora –*ss* [s] sorda, y *z* [ʒ] africada sonora *ç* [ʃ] sorda) estaba caracterizado por el carácter apical de las primeras y el predorsal de las segundas. Al perderse la oclusión de [ʒ] y [ʃ] surgió una oposición mínima entre articulaciones apicales y predorsales, que era de difícil sostenimiento por la proximidad tanto articulatoria como de timbre. El castellano adelantó hasta  $\theta$  las *z* y *s* predorsales –con lo que vinieron a distinguirse de las *z*, *s* apicales; mientras que el andaluz las atrajo al punto de articulación de las predorsales, neutralizándolas. Como, por otra parte, se había anulado la oposición de sonoridad, el castellano creó una oposición  $\theta$ -*s* mientras que, en andaluz, todo quedó en una neutralización, /s/ (Alvar 1982:132).

Abad (1982:89) propone que desde el siglo XV se hallaba generalizada en el habla común del Reino de Sevilla la pérdida del carácter africado de ambas consonantes (*z* y *ç*); sin embargo, menciona que la ausencia de oclusión alcanzó primero a la *z* y después a la *ç*.

También Diego Catalán (1989:70) habla sobre la fricativización de estas consonantes desde el siglo XV, sobre todo en el habla popular. Menciona que por esta falta de africación de las consonantes se produjo el seseo en el habla de Sevilla:

Podemos afirmar que en el siglo XV se hallaba tan generalizada en el habla común del reino de Sevilla la pérdida del carácter africado de /ç/ y /z/, que la /ç/ se asemejaba peligrosamente a la /ss/ y la /z/ a la /s/, dando lugar a una creciente tendencia a identificar estos fonemas en una pareja única de dorso-dentales fricativas, sorda y sonora (Catalán 1989:70).

Como se aprecia en la cita, Catalán aún menciona la diferencia entre sordas y sonoras, entonces podemos inferir que el ensordecimiento se produjo después de la pérdida del carácter africado de estas consonantes.

De igual manera Penny (1993:121) resalta la diferencia de sonoridad después de que ya había ocurrido la fricativización de estas consonantes: “Las sibilantes africadas (sorda y sonora) del español medieval se convirtieron en fricativas al inicio del siglo

XVI. Las demás sibilantes se mantuvieron en fricativas alveolares y fricativas prepalatales (sordas y sonoras)”.

Los testimonios anteriores no significan, sin embargo, que la pérdida de sonoridad y la pérdida de oclusión fueran fenómenos separados o que se hayan diferenciado tajantemente entre sí, eran fenómenos en convivencia que podían estar sucediendo al mismo tiempo. Así lo menciona Lathrop (1984:223) pues no separa los dos cambios, solamente apunta su inicio en fechas diferentes, después menciona que un proceso se suma al otro, ya que “tenemos ciertas noticias sobre confusiones y trueques entre sordas y sonoras desde finales del xv. Además, en el xvi empieza a perderse la pronunciación africada de la pareja [ʃ]-[ʒ]”.

Con base en todo lo anterior y en el análisis del corpus (el cual se mostrará en el Capítulo 4), sugiero que el ensordecimiento ocurrió primero y después el cambio en la correlación de los fonemas indicados. Esto lo propongo porque en el corpus de trabajo para esta tesis la *s* y la *z* alternan asistemáticamente en los mismos contextos y no se encontraron grafías que señalen sonoridad exclusiva, como *i*, *gi*, *ggi*, *gh* para el sonido [ʒ], *s* para [z] y *z* para [ʒ]. En cambio, sí se encontraron muestras del sonido prepalatal africado sordo [ʃ], es decir, graficaciones con *ç* y muestras del sonido prepalatal fricativo sordo [ʃ], graficado con *x*, lo cual nos indica que la confusión de [ʃ] con [s] y la velarización de [ʃ] → [x] eran fenómenos en proceso, quizá en etapa de iniciación, pero de ninguna manera generalizados ni concluidos, mientras que la pérdida de sonoridad ya estaba en una etapa más avanzada o totalmente cumplida.

Con esto vimos cómo fue la transición del español medieval al moderno en lo que respecta al ensordecimiento de sibilantes y al cambio en el punto y modo de articulación de las mismas: la confusión de /ʃ/ con /s/ y la velarización de /ʃ/ → /x/.

### 3.1.2. Confusión gráfica y fonética de *b* y *v*

Ahora veremos las discusiones teóricas en lo que se refiere al estatus fonológico del fonema labiodental fricativo sonoro /v/ y su total confusión con el labial oclusivo sonoro /b/.

En primer lugar todos los estudiosos coinciden en que nunca existió el fonema labiodental fricativo sonoro /v/, sino que, si alguna vez hubo diferenciación entre estos dos fonemas, fue en que *b* y *v* eran: bilabial oclusivo sonoro /b/ y bilabial fricativo sonoro /ɸ/, respectivamente, es decir, la única diferencia radicaba en el modo de articulación, uno oclusivo y el otro fricativo.

Amado Alonso (1955/1976a:21) sostiene que en el español medieval los dos fonemas se distinguían, y que la *v* era bilabial /ɸ/ porque lo atribuía a un sustrato ibérico, ya que lo relacionaba con el cambio *f* > *h*-, apoyándose en la falta de labiodentales en las lenguas ibéricas. En español, clasifica la confusión como fenómeno iniciado en el norte, en Burgos y las dos Castillas y con cierta resistencia en el sur, es decir, en Andalucía los distinguían todavía hacia el siglo XVII.

Sin embargo, Alonso cae en una contradicción al decir que era un fenómeno reciente en Zamora hacia 1560, pues esta provincia está en el norte de la península, muy cerca de Burgos, en el Reino de León, y por ello lo más lógico sería pensar que la confusión fuera temprana también allí. Veremos en §4.1. que Diego de Ordaz, autor de las *Cartas* utilizadas en esta tesis era originario de Zamora y que en las cartas se ve claramente la distinción de *v* y *b* en la escritura, sin embargo, esto no necesariamente indica que distinguiera los dos sonidos, sino que era un hombre culto que tenía una ortografía rigurosa.

Dice Lope Blanch en su estudio sobre “La labiodental sonora en el español de México” (1988) que los escribanos coetáneos de Diego de Ordaz confundían totalmente los dos fonemas. Y en *El habla de Diego de Ordaz* menciona acerca de este problema que:

En las cartas escritas por Ordaz mismo parece haberse conservado rigurosamente la distinción entre *b* bilabial oclusiva /b/, y la fricativa labiodental (o también bilabial) /v/, en posición intervocálica. (...) Pero en esta ocasión, las dos cartas no autógrafas se apartan radicalmente de la fonética ordaciana. En ambas la confusión de las labiales es total, en cualquier posición o secuencia. (...) En resumen, confusión total y absoluta, donde Ordaz mantenía clara distinción fonética... o milagrosa diferenciación ortográfica (Lope Blanch 1998:208).

Martinet (1955/1974:444) acepta que “hasta finales del siglo XVI, los casos de confusión de *b* y *v* se limitan, en líneas generales, a Castilla la Vieja y a las provincias septentrionales que le son vecinas”, por lo que concluyo que, como los casos de confusión eran limitados, supone que había distinción en el resto de las regiones españolas.

Alarcos (1981:133) también dice que las grafías *b* y *v* remitían a fonemas diferentes en castellano y catalán medievales y explica que “cuando /v/ se realizó bilabial [b̥], se confundió con la realización intervocálica del fonema /b/, que también era [b̥], resultando que la relación primitiva /b/: /v/ se hizo [b]: [b̥], diferencia no distintiva entre variantes combinatorias”.

Por otro lado Menéndez Pidal (1904/1940:114) acepta que “la lengua antigua distinguía una *b* oclusiva sonora y una *v* fricativa sonora” pero que “hacia el siglo XVI se confundieron ambos sonidos”. Indica que “la diferencia que hace la ortografía moderna entre *b* y *v* quiere ser etimológica, pero no responde a la pronunciación” (Menéndez Pidal 1904/1940:97).

Penny (1993/2001:119) por su parte, documenta la diferencia de estos dos fonemas y propone la poesía como prueba de ello, pues estos fonemas oponían

diferentes rimas. Pero dice que para finales del siglo XIV, /b/ y /β/ iniciales “se neutralizarían en todos los contextos y la oposición fonológica sobreviviría tan sólo en posición intervocálica”. Y para el siglo XV señala la fusión completamente. Para el siglo XVI, menciona que la confusión era total en todas las variedades del español.

Blanco (2006:44) apunta una clara distinción entre las grafías *b* y *v* correspondientes, respectivamente, a los fonemas bilabial oclusivo sonoro /b/ y bilabial fricativo sonoro /β/ en el siglo XV. Hacia finales de esa centuria atestigua la incipiente confusión de grafías, que se extiende a lo largo del siglo XVI y se generaliza en el XVII.

Hasta ahora todos estos estudiosos proponen que existía una diferencia entre estos dos fonemas, por lo menos en lo que respecta a su modo de articulación, pero vemos que documentan las confusiones desde temprano. Enseguida veremos lo que dicen los autores que sostienen que nunca hubo tal diferenciación entre *b* y *v*.

Por ejemplo, Alatorre (1979/2003:314) dice que “esta igualación parece haber existido en Burgos y La Rioja, que fue donde tuvo origen el castellano y como nuestra lengua es el castellano, convendrá que ciertos pedantes o despistados sepan que en ella nunca se ha distinguido entre *v* y *b*”. Es curioso ver cómo trata de “pedantes” y “despistados” a los que actualmente, mediante una hipercorrección, tratan de distinguir /b/ y /v/ en un afán artificial de hablar *bien* la lengua y sólo ponen en evidencia su total ignorancia del tema en cuestión.

De la misma manera Lathrop (1984/1992:30, 107) documenta esta confusión “desde la época del latín vulgar, tanto *b-* como *v-* se pronunciaban como [b], y este rasgo ha llegado hasta el español moderno”.

Cabe mencionar que, en sendos trabajos para el español de México del siglo XVII y del siglo XVIII, Company (1993) y Reyna (2005), respectivamente, atestiguan la

confusión de estos sonidos. Company porque no ve en los documentos ninguna graficación con *v* y Reyna porque en sus documentos se utilizan ambas grafías en las mismas palabras.

En los documentos del corpus utilizado en esta tesis las grafías *b* y *v* alternan en los mismos contextos (con excepción de las cartas escritas por Diego de Ordaz), por lo que concluyo que la diferenciación de los sonidos respectivos o se había perdido o los amanuenses no la podían distinguir porque nunca existió una diferencia entre estos sonidos.<sup>10</sup>

### 3.1.3. Aspiración de *f*- inicial

Las consonantes iniciales latinas, o por lo menos la mayor parte de estas, se conservaron en su paso al castellano. Algunas sí sufrieron cambios como la *f* que cambió a *h* y después tendió hacia su desaparición *f*- > *h*- >  $\emptyset$ , aunque en algunos casos se reforzó y originó un fonema velar *f*- > *h*- > *x*-.

Adelante explicaré las diferentes teorías que tratan la causa del cambio de *f*- > *h*- en español. Existen dos principales corrientes que explican dicho cambio; una de ellas, a la que se ciñe la mayor parte de los estudiosos, con algunas diferencias en cuanto a su postura, propone el cambio como influencia de las lenguas prerromanas como el vasco. Por otro lado Salvador (1982) propone una teoría más innovadora en la que plantea una causa geológica para la pérdida de *f*- inicial en español. A continuación apuntaré detalles sobre una y otra teorías.

Menéndez Pidal (1926/1950:201) propone que “el paso de *f* > *h* es primitivo y consustancial en el habla castellana del Norte y debido a la ausencia de *f* en los idiomas

---

<sup>10</sup> Sigo hablando de dos sonidos diferentes por la discusión teórica presentada en este apartado en la cual Amado Alonso (1955/1976a), Martinet (1955/1974), Alarcos (1981), Menéndez Pidal (1982), Penny (1993/2001) y Blanco (2006) sostienen la diferencia entre /b/ y /b̄/.

ibéricos”. Es decir, como una influencia de sustrato debida al vasco, lengua de origen desconocido hasta el momento, con peculiaridades que no comparte con el español y con ninguna de las lenguas latinas y la cual no posee el fonema /f/. Así, para Menéndez Pidal es un fenómeno que tiene origen en el norte de la Península Ibérica y que de ahí se propaga hacia el sur, cambiando las *f-* iniciales del latín a *h-* en el castellano antiguo. Incluso en la toponimia de la zona hay pruebas del cambio de *f-* a *h-*.

Deberá observar el lector que esta *h* que sustituyó a *f* no se comportaba como hoy, es decir, si actualmente una *h* gráfica representa un cero fonético Ø, en ese entonces sí representaba un fonema, aspirado y débil, pero fonema al fin /h/.

Sigue diciendo Menéndez Pidal, que el cambio de *f* > *h-* en castellano antiguo ocurrió primero en la pronunciación y después en la ortografía; es decir, podría encontrarse una palabra escrita con *f-* en la que el hablante ya pronunciaba una *h-*. Para este autor todo el latín vulgar pronunciaba /h/ la *f* y aunque se conserva la grafía *f* en todos los escritos de los siglos X al XII, sí aparecen ejemplos de pérdida o de escritura con *h* (Menéndez Pidal 1926/1950:208).

Por supuesto que esta *f-* inicial no cambió a *h-* en todos los contextos, en algunos se conservó como ante el diptongo *ue* en *fuego*, ante la vibrante *r* como en *frente* y en palabras cultas *fe*, *fiel*. En el resto de los casos, alternaban las palabras con *f-* y *h-*, pero los más cultos acabaron seleccionando la *f-* como muestra de conservación del fonema original (Lathrop 1984/1992:110), ya que la *h* en un principio fue considerado un barbarismo de la gente inculta del Norte de Castilla y sus alrededores (Menéndez Pidal 1926/1950:220).

Posteriormente, durante la reconquista de Alfonso VII del reino de Castilla, la *h-* penetra hacia el sur, como una cuña, y así, se desaloja la *f-* del territorio mozárabe. Con

esto la pérdida del fonema *f*- adquiere estatus social alto, pues obedece a una influencia cortesana.

Sólo hasta los siglos XV y XVI esta *h*- inicial en la lengua literaria se vuelve tolerada e incluso es considerada expresión más destacada y llana. Este hecho, sin embargo, es solamente la generalización de una práctica antigua de pronunciación que antes se había considerado como dialectal y vulgar.

Menéndez Pidal en sus *Orígenes del español* (1926/1950:199) también menciona la opinión de otros estudiosos sobre el mismo tema, por ejemplo, Delius, Luchaire, Meyer-Lübke y Wartburg opinan, a grandes rasgos, que la sustitución de *f*- por *h*- se originó porque en el norte de la Península Ibérica, gracias al influjo vasco, no se pronunciaba esta *f* labiodental sino bilabial, lo que la llevó a su aspiración y posteriormente a su pérdida.

Martinet (1955/1974:434), sin embargo, no se explica por qué un hablante cambió una firme labiodental fricativa sorda por una débil aspiración glotal, es decir, cómo una pronunciación fuerte se hacía débil en los hablantes originarios del norte de España. Veremos más adelante cómo Salvador (1982) liga estas dos ideas y propone una solución singular a este fenómeno.

Lloyd (1987:213) comparte la opinión del sustrato vasco para la pérdida de *f*- a favor de *h*-, pero propone, asimismo, la pronunciación de la *f* como una bilabial [ɸ] y no como labiodental [f].

Pensado (1993:155) presenta un panorama amplio en su estudio sobre el cambio de *f*- a *h*-: muestra el cambio en todos los territorios romances y los atribuye a desarrollos independientes aunque muy parecidos entre sí. El único que sí acepta que se puede ligar con el español es el caso del gascón y añade que es posible que se trate de una evolución compartida entre ambas lenguas.

Pensado hace una crítica a la teoría de Menéndez Pidal, la cual consiste en que, mientras Menéndez Pidal presenta primero la aspiración de *f*- inicial latina y después la sonorización de *-f*- intervocálica (*-f* > *-v*-), Pensado lo sitúa exactamente al revés, es decir, para ella primero ocurrió la sonorización de *-f*- intervocálica (*-f* > *-v*-) y después la aspiración de la *f*- inicial (Pensado 1993:152). Por ello, dice que si la aspiración ocurrió primero, la *f* debió aspirarse en todos los contextos, incluso en posición intervocálica. Menciona que:

Claramente cualquier argumentación que no reconozca que la doble evolución de la *-f*- intervocálica del castellano no se explica simplemente porque el resultado *h* es posterior a la sonorización romance, está repleta de dificultades. La cronología impide considerar *f* > *h* como un fenómeno de sustrato (Pensado 1993:169).

Pensado concluye que:

Todo indica que la *f* latina se mantuvo en castellano sin problemas hasta el momento de la sonorización. En fecha posterior, la *f* pasó a *h* en posición intervocálica [en las escasas palabras en las que aún se conservaba] y en posición inicial. El proceso no llegó a alcanzar a la *f* postconsonántica, que se ha conservado inalterada desde el latín hasta nuestros días (Pensado 1993:170).

También habla de las hipercorrecciones y dice que la aparición de *f* gráficas en la época medieval no debe descartar que la pronunciación, de hecho, fuera [h].

Ahora veremos que para Gregorio Salvador (1982) la causa del cambio de *f*- a *h*- en español no corresponde a un sustrato vasco o de lenguas prerromanas, sino que va más atrás en el tiempo pues, para él, responde a un origen geológico y es cuestión bien sabida que las condiciones geológicas no cambian sino en millones de años.

En las zonas en las que Gregorio Salvador estudió el fenómeno (norte de la Península Ibérica) el agua carece de flúor; según su estudio, la gente que bebía de esta agua perdió tempranamente la dentadura y pronunció una *f* no labiodental [f] sino bilabial [ɸ] lo que pudo ser la causa de la aspiración de la *f* > *h* y su consecuente desaparición. Salvador dice que incluso los niños y jóvenes que aún tenían dentadura pronunciaban la *f* bilabial pues esa era la que habían aprendido de sus mayores. Es

decir, si los hablantes no tenían dientes, no se podía esperar que pronunciaran un sonido labiodental, y fue por ello que el sonido que lograban realizar cuando escuchaban y querían reproducir una [f], era un sonido cien por ciento bilabial, una [ɸ].

Como se ha mencionado arriba, a causa del origen geológico de esta teoría, se puede inferir que la situación, antes, durante y después de que el Imperio Romano se estableciera en la Península Ibérica y de que la lengua romance se hablara en la zona, era la misma, por lo que el origen de la aspiración y pérdida de *f* que propone Salvador no es de sustrato sino que tiene origen geológico consustancial a la zona en cuestión.

Moreno de Alba (1972:15) también propone que el fenómeno de pérdida de *f* inicial es de sustrato pre-romano ya que se presenta en una sola lengua romance, el español y que se comparte con el vasco, lengua viva de origen no romano.

Lapesa (1981:368) atestigua todavía en la primera mitad del siglo XVI la conservación de *f* y, en el lenguaje arcaizante de notarios y abogados, se conserva incluso hasta el XVII. La *f* fue sustituida por *h* excepto en cultismos y *casos especiales*<sup>11</sup> (como frente a consonante líquida o diptongo *ue*), por influencia de Castilla la Vieja que desde mucho antes ya no aspiraba esta *h*.

Para Alarcos (1981:256) esta aspiración de *f* > *h*- se da en ámbitos lingüísticos diferentes: en el norte de España se aspira primero y se considera habla rústica y vulgar, en el sur, la conservación de *f* persiste hasta el siglo XV pues su habla es más conservadora y mejor aprendida. Sin embargo, vimos que a medida que avanza el tiempo esta aspiración y pérdida del fonema adquiere estatus social y se acepta inclusive en lengua literaria. Hoy en día la pronunciación con *f*- inicial latina, en los contextos

---

<sup>11</sup> El subrayado es mío, los *casos especiales* completos se mencionan en la página 42 en la cita de Lathrop.

arriba mencionados, se siente arcaizante y sólo se conserva en palabras muy específicas (*fondo, profundo*) y en el lenguaje artificial de los escritos legales (*foja, folio, fecho*).

En los documentos del corpus se presentan *f*- iniciales de palabra como arcaísmos en los documentos legales, como era de esperarse, y en los verbos conjugados de *hacer*.

#### **3.1.4. Yeísmo**

El fenómeno conocido como yeísmo consiste en producir el sonido /ʎ/ graficado // como /y/. Actualmente en casi toda América existe yeísmo, con excepción de algunas regiones de Colombia (aunque no se descartan focos en otras zonas americanas) en donde persiste la distinción entre /ʎ/ y /y/. En la Península Ibérica se vive una condición similar: casi toda España es yeísta con excepción de algunos focos: la mitad norte, con excepción de sus ciudades y las zonas de influencia de estas (Alarcos 1981:133).

El yeísmo fue considerado por mucho tiempo, junto con el seseo, un rasgo característico de las hablas andaluzas y el español americano fue caracterizado como andaluz por ser yeísta (entre otras cosas). Sin embargo, Amado Alonso cree que el yeísmo es fenómeno consustancial al español americano, pues piensa que la confusión es fenómeno hispánico y no influencia del dialecto andaluz. Veamos.

Para Amado Alonso (1953/1976a:159) el cambio /ʎ/ > /y/ es parte de la revolución fonética que se dio en el paso del español medieval al moderno; lo sitúa entre 1560 y 1630 y lo clasifica como el más importante de los cambios. Según él, el yeísmo era práctica común en Andalucía desde antes que en Madrid y Toledo, únicos focos de yeísmo en el centro-norte de la Península. Como Alonso es antiandalucista (como veremos en §3.3.), no ve ni el seseo ni el yeísmo en América como fenómenos de

origen andaluz. Incluso lo documenta antes en América que en España: “el cambio es en todas partes moderno. La más antigua documentación es de hacia 1680 para Lima. Sólo de un siglo más tarde son nuestras primeras noticias del yeísmo andaluz” (Alonso 1953/1976a:206).

La cronología para Alonso es, por lo tanto, que el yeísmo apareció primero en América, después en Andalucía y más tarde en Madrid y Toledo. Asimismo, propone que no hay solamente un foco de propagación del fenómeno.

Alonso plantea (aunque con ciertas reservas, pues dice que le faltan datos para atribuir una generalidad) que el yeísmo es un fenómeno urbano, cuyo foco de propagación son las ciudades y para ello muestra el mencionado caso de Toledo y Madrid, en el centro de España, y Sevilla como irradiación en el área de Andalucía. En América también lo atestigua en las grandes ciudades y sobre todo en las capitales. Este autor, junto con Menéndez Pidal, Américo Castro y Tomás Navarro, recomienda la práctica castellana de distinguir la *ll* /*ʎ*/ de la /*y*/ por ser esta la pronunciación tradicional. Incluso dice que esta *ll*, de palabras derivadas del latín, se pronunció /*ʎ*/ en España hasta el siglo XVIII.

Con lo anterior encuentro ciertas contradicciones aparentes entre lo que propone Alonso. Primero menciona que *ll* > *y* es un cambio que se da entre 1560 y 1630, y más adelante dice que las palabras con *ll* se pronunciaron /*ʎ*/ hasta el siglo XVIII. Después nos dice que el foco de irradiación de la práctica confundidora fue Madrid y Toledo (zona castellana) y más adelante recomienda la práctica castellana de distinguir por ser esta la pronunciación tradicional. Es evidente que ninguna de las cosas que dice Amado Alonso puede ocurrir al mismo tiempo.

Según Parodi (1977:241), Cuervo encontró ejemplos de yeísmo en textos de principios del siglo XIX, pero en las investigaciones hechas después de Cuervo se ha ido puntualizando la fecha de aparición de este fenómeno hasta la segunda mitad del siglo XVI. Parodi, de hecho, lo adelanta hasta la primera mitad del siglo XVI —lo documenta en 1527 igual que Lapesa (1981:564)— como un fenómeno bastante extendido en América, aunque su generalización ocurrió después. Presenta su análisis basado en textos de principios del siglo XVI, utiliza datos del documento, como el lugar y la fecha en la que fue redactado, así como datos del amanuense, tales como el nombre, el origen geográfico y si presenta casos de seseo. Todo esto porque se considera al seseo y al yeísmo como rasgos caracterizadores del dialecto andaluz.

Parodi concluye que el yeísmo en América era proveniente del español, por su cronología temprana y porque los individuos que confundían provenían de zonas yeístas españolas. Duda que el fenómeno sea urbano y que se haya dado primero en las zonas bajas.

Para Cano (1988:241, 242) hay casos de yeísmo en la Edad Media, pero son esporádicos. Considera al yeísmo americano desde principios del siglo XVI y sin relación con el andaluz, pues a este lo sitúa en la segunda mitad del XVI.

Para Lloyd (1987:344-346), el yeísmo se presenta primero en las clases bajas de la zona de Andalucía. En Aragón y partes del este de Castilla se documenta pero es frenado por una fuerte reacción en contra de la indistinción, hasta que en el siglo XVI emerge otra vez sin control. En América lo documenta desde tiempos de la conquista aunque el fenómeno no tuvo la aceptación que tuvo el seseo.

Señala Lathrop (1984/1992:220) que no hay nada especial que advertir en el fenómeno del yeísmo con excepción de que se atestigua desde el siglo XIII en algunos puntos de la Península. Con lo anterior sitúa al fenómeno como previo para España. Y

para Moreno de Alba (1988/2001:199) el yeísmo es un fenómeno urbano que se propagó en Madrid procedente de Andalucía, y que de Madrid se extendió hacia América a través de las grandes metrópolis como México y Lima.

Boyd-Bowman (1988:86) no encuentra casos de yeísmo americano hasta el siglo XVII periodo en el que, para él, el cambio era aún esporádico.

La temática de los documentos del corpus de trabajo no permite que aflore yeísmo en los amanuenses, ya que se encontró que las grafías *ll* y *y* no alternan en los mismos contextos.

### **3.2. Trasplante del español a América**

El estatus que el español tenía en México en el siglo XVI depende de los hablantes venidos de España al Nuevo Mundo. En el apartado siguiente §3.3. veremos que, si bien la mayor parte de los españoles que se establecieron en América provenía de la zona de Andalucía, llegaron a las nuevas tierras españoles procedentes de todas las regiones de aquel país. Se dice que el español de América es un español andaluzado pues sus rasgos fonológicos y su léxico se asemejan a los que poseen los habitantes de Andalucía, sin embargo, esto no es del todo así. Veamos.

Durante los primeros años de la conquista, en el periodo de orígenes, el español andaluz convivió con los demás dialectos españoles y también con las lenguas nativas americanas y se creó una koiné<sup>12</sup>. Las lenguas nativas de las Antillas desaparecieron muy pronto pero dejaron rasgos en el léxico del naciente español americano. Cuando los españoles se establecen en tierra firme, la koiné de los diferentes dialectos españoles y

---

<sup>12</sup> Cualquier lengua común que proceda de una reducción a unidad, más o menos artificial, de una variedad idiomática (Lázaro Carreter 1953/1977).

las lenguas caribeñas entra en contacto con las lenguas dominantes de la zona en la que hoy se encuentra México: el maya y el náhuatl, entre otras muchas.

Se ha dicho que estas lenguas no permearon el sistema fonológico del español, y que su influencia se limitó al léxico, pero me parece, basándome en el análisis realizado para esta tesis (véase el §4.4.1.), que sí hubo cierta influencia del náhuatl en el sistema fonológico del español debido a la permanencia temporal (en lo que se adaptaron a la pauta del español) de dos sonidos que se mantuvieron en el siglo XVI y que aún se mantienen en el español americano actual (como alófonos, pero aún se escuchan) mientras que en España ya estaban en proceso de desaparición (véase §3.1.): el sonido africado [ʃ] y el fricativo [ʃ̄], ambos prepalatales sordos. A continuación expondré el porqué de mi propuesta y explicaré cómo se da la nivelación en el trasplante del español hacia América.

Primero hablaré sobre la población que llegó a América y la nivelación que tuvo que haber entre los diversos dialectos del español que llegaron a las nuevas tierras. Pues a ese respecto algunos estudiosos afirman que la población que llegó al continente americano provenía de todas las provincias españolas, otros afirman (y lo veremos en el §3.3.) que la mayoría de la población era de origen andaluz y más específicamente de Sevilla.

El otro punto en discusión es el estrato social de dichos pobladores: se creyó durante mucho tiempo que las personas que se embarcaban hacia América eran de estrato social bajo, presidiarios, bandidos, etc. Sin embargo, Alonso (1953/1976:15) apunta que aquellos pobladores eran de todos los estratos y que el pueblo que habitó en América los primeros años después del descubrimiento representaba en el mismo porcentaje al pueblo español peninsular.

Lope Blanch (2000:177-184) plantea el caso de nivelación que se dio en la Nueva España y considera que a través de esta nivelación fue que el castellano se empezó a hacer español. Los conquistadores que llegaron a las Antillas conservaban su dialecto originario y, consecuentemente, había una mezcla de estos. Existe una fuerte influencia de andalucismo; sin embargo, no todos los fenómenos fonológicos o fonéticos andaluces llegaron al altiplano central mexicano, sólo el seseo y el yeísmo fueron generalizados en América.

Después de varias décadas, dice Lope Blanch (1969a:13), tal vez un siglo completo, se llega a un estado de español nivelado que ya puede ser llamado americano, en su correspondiente zona dialectal, para esta tesis, mexicana. Sigo a Lope Blanch cuando dice que “tendremos que precisar hasta qué punto la pronunciación de los conquistadores era de tipo andaluzado o en qué medida y proporción conservaban éstos el habla de corte *castizamente* castellano”.

Respecto de la koiné que se formó en la nivelación del castellano que llegó a América, Lope Blanch dice que la rapidez con la que fueron aceptados fenómenos como el seseo y la confusión de *x* con *j* prueba que estos cambios ya se estaban produciendo en España. Lo que hizo esta koiné fue acelerar estos procesos.

Parodi (1995) también nos habla de esta nivelación en los primeros años del descubrimiento de las tierras americanas y lo prueba con el análisis de sus documentos:

Pocos son los casos en que un autor de los textos aquí analizados refleja consistentemente la pronunciación de un solo dialecto en las grafías de sus manuscritos. Ello sugiere que la mayoría de los primeros habitantes hispanos de América no hablaban un solo dialecto peninsular, sino un español nivelado, el cual contenía rasgos de todos los dialectos, aunque fuera predominantemente andaluzado (1995:28).

Con el análisis de sus documentos, Parodi presenta evidencia de que el español americano, desde sus orígenes fue un español nivelado.<sup>13</sup>

El contacto que hubo entre la Nueva España y la Península hizo que la mayor parte de los cambios que la lengua experimentaba en España ocurriera también en las principales metrópolis americanas. Sin embargo, el español americano es más cercano a la modalidad andaluza por sus rasgos fonológicos y léxicos (Lapesa 1981:565).

En lo que respecta al nivel social y cultural de la población que llegó a América, Granda (1994:89) menciona que hubo una heterogeneidad lingüística tanto en lo que respecta al índice diatópico como al diastrático. A la colonización americana contribuyeron pobladores de varias áreas culturales y lingüísticas y hablantes de los diferentes dialectos castellanos.

Sin embargo, para Frago (1999) sí hubo diferenciación diastrática:

Puesto que en la emigración a Indias predominó con mucho el elemento popular, para más señas salido de todos los rincones de España, será necesario preguntarse si al Nuevo Mundo se trasplantó un español básicamente nivelado o más bien un español con específicos rasgos regionales, sujetos a una ulterior y original nivelación americana (Frago 1999:12).

Es decir, Frago sugiere que el español que se formó en América fue una nivelación proveniente de gente de estrato popular con claros rasgos dialectales.

Como mencioné antes, sugiero que las lenguas indígenas permearon al español en la permanencia temporal de dos sonidos [ʃ] y [ʂ]. Estos fonemas estaban en proceso de desaparición en España (recordemos que el africado se confundió, en el español atlántico, con el alveolar fricativo [s] y que el fricativo se velarizó [x]), este proceso fue refrenado en México gracias a la influencia del náhuatl que poseía estos sonidos o unos muy próximos. De manera que la influencia de las lenguas indígenas fue frenar el

---

<sup>13</sup> También comparten esta opinión del español americano nivelado autores como Alvar (1982), Lüdtkke (1994) y Alarcos (1981).

proceso de pérdida que ya estaba iniciado en España y mantener temporalmente estos sonidos en lo que el léxico indígena que se incluyó en el inventario del español general se adaptó a la pauta de la lengua española.

### **3.3. Andalicismo del español americano**

Finalmente habrá que mencionar que la mayoría de los españoles que llegaron a la Nueva España era de origen andaluz (cf. Boyd-Bowman 1985:v) y por ello se dice que el español de América tiene rasgos articulatorios de dicha región española. Este es el llamado andalicismo del español americano, y en los apartados del Capítulo 4 correspondientes a cada fonema se hablará de algunos detalles en específico.

Dentro del andalicismo, clasificaré sus estudios dentro de tres grandes bloques: en el primero, que cronológicamente es de los inicios del siglo XX, hablaré sobre los estudios que suscitaron una polémica entre Henríquez Ureña y Wagner; ya que este acepta el andalicismo en América y aquél lo rechaza. Esta polémica fue originada a partir de una mala interpretación por parte de Henríquez Ureña de las palabras de Rufino José Cuervo.

Ahora bien, argumenta Cuervo, la época relativamente reciente en que comenzaron a producirse el yeísmo y la pérdida de -d- explica que aún no se hayan generalizado en América (ni en España), por contraste con rasgos como el seseo y la aspiración de la *f*, que dominan uniformemente en toda el habla americana (cit. por Guitarte 1991:34).

Henríquez Ureña malinterpretó la teoría de Rufino José Cuervo, cuando menciona que Cuervo tiene razón al afirmar que el seseo (confusión de *z* y *c* con la *s*) es de origen americano y que no fue trasplantado de España. Pero Cuervo nunca dice esto en el texto, precisamente sugiere que el seseo, al ser más antiguo y generalizado en América, fue trasplantado de España por la población que mayoritariamente era de Andalucía.

El segundo bloque de estudios que propongo, lo hago a partir de la investigación de Peter Boyd-Bowman iniciada en 1950. Esta investigación tiene un enfoque más científico debido a que Boyd-Bowman la hace con datos provenientes del Archivo de Indias y de algunas otras fuentes que proporcionan el origen de cada persona que se embarcaba hacia América. Con esta investigación se termina de manera tajante la polémica sobre el andalucismo, pues en ella Boyd-Bowman concluye contundentemente que la mayor parte de la población española que llegó a América era de origen andaluz (sobre todo de Sevilla) y por ello la pronunciación del español en América es andalucista.

El tercer bloque lo asigno a estudios más recientes que toman como base la investigación de Boyd-Bowman, sin dejar de lado el tema de la polémica entre Henríquez Ureña y Wagner.

Rufino José Cuervo estudió el andalucismo en América, y si bien dice que la población que llegó al nuevo continente proveniente de España era de todas las regiones, nunca rechaza el andalucismo. A continuación expongo una idea de Cuervo en la que sugiere que su postura es andalucista y, aunque nunca lo expresa claramente, sí se puede ver que de ninguna manera podría haber rechazado la idea de la influencia andaluza en las nuevas tierras:

El movimiento de Andalucía fue el que prevaleció en América, sin que llegara a ésta la escisión o la reacción que se verificó en algunos lugares de aquélla, que sólo conocen la z (Cuervo 1901/1954:534).

En este pasaje, Cuervo está hablando del ceceo y seseo<sup>14</sup> andaluces. El seseo como “el movimiento de Andalucía [que] fue el que prevaleció en América” y el ceceo

---

<sup>14</sup> Parodi (1995:79) dice que “tras haberse generalizado el fonema /θ/ en el castellano viejo, durante la segunda mitad del siglo XVI, se ha entendido por seseo como la pérdida de la oposición /s/-/θ/ a favor de /s/ o de /ʝ/, y el ceceo como la pérdida de la misma oposición a favor de /θ/”. De manera que el seseo que llegó a América se entiende como la confusión de /ʝ/ con /s/ ya que a este continente nunca llegó /θ/.

como “la escisión o la reacción que se verificó en algunos lugares de” Andalucía. De manera que observa cierta influencia andaluza en el español americano.

Fontanella (1976) confirma esta idea de Cuervo, pues dice que al atestiguar el seseo americano como fenómeno temprano, Cuervo afirma que el seseo americano procede del andaluz. Sin embargo, no admite como andaluz el yeísmo, pues este fenómeno no era exclusivo de Andalucía, sino que se extendía a Castilla, y en América no estaba generalizado.

Cuervo, como dije arriba, mencionó muchas veces que la población que llegó a América provenía de todas las regiones españolas, a partir de esta idea Pedro Henríquez Ureña rechaza rotundamente el andalucismo y no admite ninguna relación entre ambos dialectos (andaluz y americano) sino que atribuye su similitud a desarrollos paralelos.

Henríquez Ureña y Wagner se enfrascan en una disputa sobre el andalucismo, en la que cada uno de ellos publica su propia teoría. Ureña rechaza el andalucismo con ayuda de los estudios de Cuervo que él malinterpretó y Wagner reduce el andalucismo a las tierras bajas americanas.

A continuación una y otra posturas explicadas por Moreno de Alba (1988/2001), Primero, la postura de Henríquez Ureña:

[Según Pedro Henríquez Ureña] si había cuatro sibilantes en el siglo XV, el español de América las redujo a una *s*, frente a Andalucía, que hizo o bien *s* o bien una interdental; esta última realización no se da en América; el yeísmo en la Península Ibérica, no es exclusivo de Andalucía, pues se extiende por Castilla la Nueva y, en América, hay zonas (Colombia, Ecuador, área andina, partes de Chile y Argentina) que conservan la *ll* (Moreno de Alba 1988/2001:31).

Con esto vemos que Henríquez Ureña rechaza el andalucismo con los argumentos del seseo y el yeísmo, ya que para él el seseo americano era diferente del andaluz, pues en Andalucía se había optado por dos soluciones para la evolución del fonema /ʃ/ (/s/ seseo, /θ/ ceceo) mientras que en América sólo por una (/s/ seseo). En el

caso del yeísmo argumentaba que también este fenómeno estaba extendido en Castilla y que no era general en América.

La postura de Wagner era diferente:

[Según Wagner] en Andalucía es más fuerte la tendencia a relajar las consonantes finales, particularmente la *s*, y es la misma articulación que puede escucharse en las tierras bajas de América; cuando, en América, la *s* no se relaja (por ejemplo, en posición explosiva), su timbre, de carácter predorsal, no se parece al que priva en el centro y norte de España, que suele ser más bien apical (Moreno de Alba 1988/2001:31).

Wagner compara las realizaciones finales de las *s* de las costas americanas con las andaluzas cuando son relajadas y las halla semejantes; cuando la posición de alguna *s* no permite su relajamiento dice que las americanas son diferentes de las del norte y centro de España, por lo que se infiere que liga el español de las costas americanas con el español de la zona de Andalucía al sur de la Península Ibérica.

Es por todos conocida la polémica sobre el andalucismo que protagonizaron Wagner y Henríquez Ureña. Con sendos trabajos de estos dos autores se inició la polémica sobre el andalucismo americano. Wagner llegó a formular una teoría climatológica: nos dice Fontanella (1976:26) que ese autor amplía la tesis andalucista pues afirma que los dialectos del sur de España influyeron en el español americano, sin embargo limita esta influencia a las zonas costeras americanas y con esto formula su teoría climatológica sobre el andalucismo americano.

A este respecto Penny (1993/2001:40) nos explica que esta teoría climatológica se basa en que se ha creído que los pobladores originarios de la meseta castellana preferían instalarse en las tierras altas de América y que los originarios de las tierras bajas y costeras españolas —especialmente de Andalucía— preferían las tierras bajas americanas.

Amado Alonso (1952, 1953/1976a:12) es uno de los estudiosos que siguen la línea de investigación de Henríquez Ureña, pues también rechaza el andalucismo del

español americano con los mismos argumentos de Ureña. Ubica al seseo americano y andaluz como desarrollos paralelos, y sitúa al yeísmo americano antes que el español y al yeísmo español con focos diversos en Castilla y Andalucía.

Para terminar con esta polémica, llegó la investigación de Peter Boyd-Bowman, iniciada en 1950, en la que afirma con datos suficientes y probatorios, como se verá adelante, que el español de América sí tiene como base el español andaluz (y mejor dicho, sevillano) ya que la mayor parte de los pobladores que llegaron al Nuevo Mundo eran de esa zona.

Boyd-Bowman (1985:v) proporciona datos de más de 56 mil pobladores de América durante los primeros cien años “en los que se echaban los cimientos de la cultura y la lengua españolas”, con esto afirma que “no tardamos en darnos cuenta del papel decisivo que desempeñaron, en cada una [de las salidas de España], los andaluces, y sobre todo los sevillanos”.

Con datos contundentes Boyd-Bowman concluye que la mayor parte de la población que se embarcaba a América era de origen andaluz y más específicamente de Huelva y Sevilla: “en la época primitiva o antillana el grupo más numeroso, en cada año, y en todas las expediciones, fueron con mucho los andaluces, de los cuales más del 78% procedían de las dos provincias de Sevilla y de Huelva” (Boyd-Bowman 1985:XII).

Sigue diciendo Boyd-Bowman (1985: XIII), con base en dos cortes cronológicos que: “si dividimos la época antillana en dos partes casi iguales, desde 1493 a 1508 y desde 1509 a 1519, resulta que en la primera mitad la aportación andaluza fue un 60%”.

Acerca del español que se asentó en la altiplanicie mexicana, que son los datos que interesan para esta investigación, confirma lo mismo: que la mayor parte de los que vinieron a tierra firme eran andaluces, con lo cual se echa por tierra la teoría climatológica que proponía Wagner. “Los primeros conquistadores de México fueron

reclutados casi todos en Cuba y de estos compañeros de Cortés y Narváez he identificado la procedencia de 743, es decir, como la tercera parte. [...] Una vez más ocupa el primer lugar Andalucía con 227 (30%)” (Boyd-Bowman 1985:xv).

Y sobre la teoría de Henríquez Ureña y Alonso en la que propugnaban que la población española asentada en América procedía de todas las provincias españolas, con mayoría de castellanos, nos dice Boyd-Bowman (1985:xv) lo contrario: “las dos Castillas aportaron sólo una cuarta parte de los primeros conquistadores de México”.

En otro artículo de Boyd-Bowman (1988:75, 76) el autor dice que para el siglo XVI la mayor parte de la población blanca de América (80%) era oriunda del sur de la Península Ibérica: andaluces, extremeños y castellanos nuevos.

Acerca del estudio de Boyd-Bowman, Penny (1993/2001:41) explica que el español de América es de origen andaluz porque además de que la mayor parte de los viajeros eran de la zona de Andalucía, también la mayor parte de las mujeres que viajaban a América eran de esta zona y como son las madres las que transmiten la lengua a sus hijos, transmitían el dialecto andaluz a los muy jóvenes o nacidos en América. Por otra parte, la zona de Andalucía era también en la que salían las embarcaciones hacia América y los aspirantes a la travesía debían permanecer allí por varios meses, por lo que podían adoptar las características fonológicas propias de la zona aunque procedieran de otras regiones del imperio.

Lope Blanch (1969a:32) resalta también la importancia del papel de las mujeres como transmisoras de la lengua a sus hijos: “durante la primera mitad del periodo inicial —o antillano— de la conquista, el número de andaluces llegados a América representaba el 78% del total de los colonizadores; y el 67% de las mujeres eran también andaluzas” con lo que confirma, también, la idea del andalucismo americano.

Otra investigación emprendida por Cock (1969) y dirigida por Guitarte arroja los mismos resultados y en ella se afirma el andalucismo del español americano con base en el análisis del seseo. En esta investigación se contradice la teoría de Alonso (1953/1976) en la que el seseo americano se habría cumplido en tres etapas, pues los documentos analizados por Cock muestran al seseo general y absoluto desde los primeros testimonios americanos.

Afirma Cock (1969:16) que “el seseo, que es rasgo peculiar del español de América, muy posiblemente fue traído a este continente desde Andalucía donde, como hemos visto, era fenómeno muy arraigado en el último cuarto del siglo XV”.

Alvar (1982:139) menciona lo mismo y dice que la doctrina de Amado Alonso no puede sostenerse pues en el siglo XVI había confusión total de las sibilantes; es decir, el seseo se había generalizado.

Por supuesto Guitarte se muestra partidario del andalucismo americano, critica la postura de Henríquez Ureña y Alonso, y aclara que Cuervo nunca se ocupó de la después llamada teoría andalucista (y, por lo tanto, nunca la rechazó) con los siguientes argumentos:

En este punto he llegado al final de mi razonamiento y creo haber puesto en claro que debe retirarse de la circulación la imagen de un Cuervo que hubiera refutado la teoría andalucista, porque: 1) no se ocupó de ella; 2) si lo hubiera hecho, habría rechazado una identificación lingüística vulgar entre Andalucía y América, pero si hubiera aceptado la perduración de rasgos andaluces en el español de América, es decir, si se me permite la expresión, hubiera sido un andalucista “crítico” (Guitarte 1991:37).

La mayor parte de los investigadores, con base en el análisis de Boyd-Bowman, se ciñe a la postura andalucista, y acepta, en mayor o menor medida, que el español americano tiene rasgos andaluces.

Lapesa (1964) señala que “de todo lo expuesto se deduce que hoy no cabe ya duda posible respecto al origen andaluz de algunos de los rasgos más peculiares de la

pronunciación americana: el más general, el seseo; muy probablemente, el yeísmo; seguros, aunque no generales en América, la confusión de *r* y *l* finales, la aspiración de *-s* final y la sustitución de *j* por *h* aspirada” (cit. por Fontanella 1976:37).

También Fontanella ofrece la postura de Menéndez Pidal (1962). Este autor opina que “en la base de la lengua colonial no sólo está la norma general de la lengua común, sino también un dialecto particular de ésta destacado sobre los otros desde comienzos del siglo XVI; así el español ultramarino recibió un marcado tinte andaluz al aceptar la simplificación fonológica del çeçeo-zezeo surgida en el Reino de Sevilla” (cit. Por Fontanella 1976:38).

Un dato significativo nos muestra Moreno de Alba (1988/2001:187) cuando afirma que el primer momento de importancia diacrónica con respecto al seseo fue cuando la modalidad andaluza se impuso sobre la castellana. En este momento se debe situar a Sevilla como foco y se ubica cronológicamente en los siglos XV y XVI, esto repercutió en la pronunciación americana y tiene como rasgo predominante la pérdida de la *s* apicoalveolar. “El cambio de *s* apicoalveolar a *s* dorsal o interdental, queda documentado en Andalucía desde 1500 y, en México, desde 1525” dice Moreno (1988/2001:187).

Alatorre (1979/2003:309) asegura que “la variedad del español implantada en las Canarias y en el Nuevo Mundo fue la andaluza. En esos años iniciales del siglo XVI se desata, incontenible, el torrente de *faltas de ortografía*”.

Las voces de otros investigadores se suman a la idea del andalucismo americano, entre otros: Martinet (1955/1974), Zamora y Guitart (1982), Abad (1982), Lloyd (1987), Cano (1988), Catalán (1989), Granda (1994), Lüdtke (1994) y Frago (1994, 1996, 1999). Así como García Carrillo (1988) y Company (1993) en análisis propios para el siglo XVI y XVII, respectivamente.

Para concluir, confrontando las opiniones de los estudiosos arriba citados, y con base en el análisis de los datos del corpus, el lector deberá tener en cuenta que los temas anteriores no son sencillos. El español que llegó a América no es el preclásico, no es el de los marineros andaluces, no es el de los sevillanos. El español que llegó a América se gesta durante mucho tiempo, primero en las costas caribeñas y, después, de la zona caribeña pasa al altiplano central, con la convivencia de variantes dialectales españolas de cada hablante, e incluso con las lenguas indígenas. Es un proceso lo suficientemente complejo que no permite ser tratado a la ligera, no podemos señalar un andalucismo o antiandalucismo sin caer en contradicciones y en errores o en inexactitudes.

El conocimiento de los ajustes del sistema fonológico del español del siglo XVI, las condiciones en las que se produjo el trasplante del español hacia América y la procedencia de los habitantes españoles que se establecieron en México, temas que se vieron en este capítulo, son de gran utilidad al tratar de estudiar el español de México del siglo XVI. Estos datos pueden otorgar pistas importantes para intuir la correspondencia entre grafía y sonido, y con ello establecer las reglas que se pretenden implementar al transcriptor automático.

Como mencioné anteriormente, el español mexicano no era tal en el primer siglo posterior a la conquista, por ello, debemos establecer el español que se hablaba en México a través de los hablantes españoles que se establecieron en nuestro país. El conocimiento de estos temas fundamentales para establecer quiénes hablaban el español en México en el siglo XVI, será de invaluable utilidad para establecer los correlatos entre las grafías y sus fonemas correspondientes que se harán en el siguiente capítulo.

Habrá que mencionar, de nuevo, que el presente análisis sólo es un afán por intentar acercarnos al español hablado en México en el siglo XVI, en la medida en la que

los documentos del corpus de trabajo puedan ser reflejo de los hábitos orales de los pobladores de la Nueva España.

## 4. REGLAS DE CORRESPONDENCIA ENTRE SONIDO Y GRAFÍA EN LOS DOCUMENTOS DEL SIGLO XVI

*Spelling can never be a certain guide to pronunciation  
except when a new spelling appears.*  
Lloyd, 1987.

En el presente capítulo describiré las características de los textos que conforman el corpus utilizado en esta investigación, los criterios de edición de cada uno de ellos, sus diferencias y semejanzas. Posteriormente, y ya adentrándome en el tema central de esta tesis, propondré las reglas de correspondencia al sonido para cada grafía dependiendo del contexto en el que se encuentre, para que dichas reglas se implementen en el transcriptor automático.

En los casos que considere pertinentes, abordaré el tema desde una perspectiva teórica para los sonidos que representan mayores cambios y remitiré al lector al párrafo correspondiente del capítulo anterior. Por ejemplo, se hablará del seseo, de la aspiración y del yeísmo, siempre aludiendo al andalucismo como tópico central dentro de estos temas.

### 4.1. El corpus: características

La muestra textual que utilicé para esta investigación forma parte del *Corpus Histórico del Español en México* (CHEM), consiste sólo en documentos del siglo XVI, que pertenecen a los siguientes tres conjuntos de textos:

- 1) *Documentos Lingüísticos de la Nueva España. Altiplano Central (DLNE)*, edición de Concepción Company (1994).
- 2) *Cartas de Diego de Ordaz (DO)* edición de Lope Blanch (1998, pp. 193-228).

3) *Procesos inquisitoriales contra indígenas que realizó Fray Juan de Zumárraga en Nueva España (1536-1543) (PI)* edición de María Buelna (2006).<sup>1</sup>

En cuanto a su extensión, el corpus empleado tiene alrededor de trescientas mil palabras gráficas: ochenta y cinco mil conformadas por los *DLNE*, quince mil por las *Cartas de Diego de Ordaz* y doscientas mil que integran los *Procesos inquisitoriales*.

Los documentos del corpus están repartidos en el periodo comprendido entre 1525 y 1585. Las *Cartas de Diego de Ordaz* se escribieron entre 1529 y 1530. Los *Procesos inquisitoriales* entre 1536 y 1543 y los *DLNE* se agrupan en dos cortes cronológicos: el primero de 1525 a 1540 y el segundo de 1570 a 1585. Como veremos abajo, en la figura 13, mostraré, a manera de ejemplo, las diferentes graficaciones que presentaba la palabra *licenciado* en el corpus; con este ejemplo sugiero que la fecha en que se escribe cada documento determina la pronunciación del hablante y muestra mayor o menor grado de vacilación gráfica por la vigencia de los sonidos. Otros ejemplos de vocablos que presentan también esta variación cronológica son *cibdad* > *cuidad*, *çesárea* > *cesárea*, entre otras palabras.

Se observa en la figura 13 que la variación gráfica de la palabra *licenciado* en el primer periodo de los *DLNE* oscila entre *liçenciado*(21) y *licenciado*(17); para el segundo periodo podemos apreciar que se dirige hacia la forma moderna *licenciado* y no alterna, de manera importante, con otra forma gráfica. Esto porque, para el segundo periodo, se refleja en la graficación de la palabra la pérdida del sonido prepalatal africado sordo [ʃ̄]. Por ello sugiero que la fecha en la que fue escrito un documento determina la vacilación gráfica de las palabras, pues el siglo XVI es un periodo de

---

<sup>1</sup> Para ver las referencias completas de los textos que conforman el corpus, consulte el apartado de la bibliografía al final de esta tesis.

reajuste del sistema fonológico y de los usos gráficos recurrentes en la escritura (que era la ortografía anterior al siglo XVIII)<sup>2</sup>.

|           | Forma gráfica      | 1525 a 1540 | 1570 a 1585 |
|-----------|--------------------|-------------|-------------|
| con dos ç | <i>liçençiado</i>  | 2           | 2           |
|           | <i>lliçençiado</i> | 0           | 1           |
| con una ç | <i>liçenciado</i>  | 21          | 1           |
|           | <i>licençiado</i>  | 0           | 3           |
| sin ç     | <i>ljenciado</i>   | 3           | 1           |
|           | <i>lljenciado</i>  | 0           | 1           |
|           | <i>licenciado</i>  | 17          | 7           |
|           | <i>llicenciado</i> | 2           | 0           |

Figura 13. Variación gráfica de *licenciado* en los *DLNE*

Otro factor que determina el cambio de pronunciación en los hablantes es su ubicación geográfica y origen. Diego de Ordaz, aunque escribió las cartas en la Nueva España, era originario de Castroverde de Campos, Zamora. Esta provincia se ubica en León al norte de España. En el tema del reajuste de sibilantes (§3.1.1.) vimos en qué fenómenos repercutió este origen.

Fray Juan de Zumárraga, autor de los *Procesos Inquisitoriales*, era originario de Durango, provincia de Vizcaya ubicado en el País Vasco, también al norte de España; sin embargo, él no firma los procesos inquisitoriales; siempre hay un notario encargado de asentar las declaraciones. Podemos intuir que en los documentos se reflejará la pronunciación del amanuense. Sin embargo, debemos recordar que hay hábitos en la escritura de los amanuenses que se conservan como fórmulas y que han sido heredados por generaciones, de manera que no reflejan fielmente la pronunciación real, sino sólo una aproximación a esta.

Los *DLNE* están escritos por diversas personas, algunos son españoles de variadas regiones y otros mexicanos. Al respecto nos dice Company (1994:6) que “en todos los casos, con excepción claro está de los primeros cincuenta o sesenta años de

<sup>2</sup> Esta normatividad de la ortografía a partir del siglo XVIII fue gracias a la creación de la Real Academia Española.

existencia de la Nueva España, he procurado que el autor del documento, el testigo o denunciante fuera mexicano de nacimiento”.

Como nos dice Mondéjar (1991) en su análisis para determinar la dialectología andaluza:

Los documentos que analizamos desde el punto de vista grafemático para sacar conclusiones de fonética histórica y de organización del sistema fonológico, pueden engañarnos si no los encuadramos en los parámetros de cultura y de procedencia social y geográfica de los que los escriben o copian (320).

Igualmente acerca del grado de educación existente en el siglo XVI, Mondéjar (1991:323) nos dice que “no hay que olvidar que en cualquier caso se trata de hábito escriturario de especialistas, de escribanos, no del hombre o de la mujer, las menos, que en tiempos del altísimo grado de analfabetismo se veían constreñidos a escribir una carta”. Es decir, la gente analfabeta que se veía en la necesidad de escribir una carta, acudía con un amanuense y por ello, en la carta se ven reflejados los hábitos de escritura del amanuense y no su pronunciación en sí.

Otro factor de importancia lingüística es la coloquialidad. Los documentos, cuanto más coloquiales, mejor reflejan el habla real, mientras que cuando son oficiales reflejan un habla menos real, más cuidada y artificial. El carácter de las *Cartas de Diego de Ordaz* es más coloquial, ya que, aunque Diego de Ordaz era un hombre instruido y cuidadoso de su escritura, le escribió las cartas a su sobrino, así que se encuentra cierto aire de familiaridad en ellas.

Por su parte, los documentos de los *Procesos Inquisitoriales* son oficiales, su lenguaje es más cuidado y artificioso, son juicios y hay cartas dirigidas a las autoridades españolas. Sin embargo, hay residuos de habla coloquial ya que los procesos y juicios son contra indígenas por casos como poligamia, brujería, estupro, etc. y en ellos se manifiesta el carácter coloquial de cierto registro de población en la Nueva España por medio de los enjuiciados y testigos.

Los *DLNE* están integrados “fundamentalmente por materiales de carácter coloquial, que se aproximan algo más —en la medida en que la lengua escrita sea reflejo de la lengua hablada— al habla del periodo colonial mexicano” (Company 1994:5). Son de cuatro tipos: cartas, denuncias y testimonios en juicios, inventarios y testamentos, y peticiones e informes.

#### **4.1.1. Características editoriales de los textos que integran el corpus de trabajo**

En la presentación de los *DLNE*, Lope Blanch menciona que “todas las transcripciones se harán con el máximo rigor paleográfico, atendiendo puntualmente a las grafías, sobre todo en el caso de fonemas que estuvieran en crisis durante la época a que pertenezca cada documento” (Company 1994:XVI).

Ciertamente, hay más semejanzas que diferencias entre los tres conjuntos de textos que integran el corpus. En los tres se puede ver la preocupación por parte de los editores de que se haga un trabajo paleográfico riguroso, por lo que respetan, en la medida de lo posible, el documento original. Digo en la medida de lo posible, porque sin convenciones editoriales se dificulta (para un lector del siglo XXI no instruido en estos temas) la lectura del juicio, carta o texto en cuestión.

De manera evidente, y siendo tres grupos de textos editados por diferentes editores y para distintos propósitos, hay variaciones en cuanto a las convenciones editoriales y se presentan como a continuación se describe:

La única grafía que cambia la editora en los *DLNE* es la *v* y la *u*; a cada una se le asigna su debido valor vocálico o consonántico. Por ejemplo, podemos ver que una *u* gráfica en el documento original que tenga valor consonántico /b/ se transcribe, a juicio de la editora, con una *v* gráfica. Encontraremos, por ejemplo, *avedes* por *auedes*. En las *Cartas de Diego de Ordaz* y en los *Procesos inquisitoriales* no tenemos esta convención

editorial, por lo que veremos gráficamente *auedes* y habrá que asignarle a esta *u* gráfica su debido valor consonántico mediante una regla. De la misma manera no encontramos en los *DLNE* la grafía *v* con valor vocálico /u/, es decir, se cambió la grafía a *u*, se encontrará, por ejemplo, *ciudad* por *civdad*.

Igualmente, en el caso de las abreviaturas o siglas, en los tres conjuntos de textos se desatan haciendo que se note la intervención del editor, por medio de cursivas, en los tres textos. Vemos a continuación los ejemplos:

- (2)
- a. Muj *jllustre señor*. \\\ Aunque por quedar en la cama con muj grandes dolores de la gota dexo de escryvjr <sup>2</sup> a todos esos señores, no qujse a njnguna *quenta* se fuese este navjo sin carta /<sup>3</sup> mja para v. m. (*DLNE*, 76, 1583)
  - b. Dos vezes os *e* escrito *desque* vine a España, la vna de Madrid i la otra de aquí, de Toledo, por la vía de Santo Domingo (*DO*, Carta 1)
  - c. *Secretario Xristoval Larios* (*PI*, Caso contra Diego Díaz, f. 2)

Las cursivas presentes en estos ejemplos son reposiciones del editor, pues en estos casos se encontraban expresadas mediante abreviaturas, vemos en los ejemplos (2a) y (2c) que se dan frecuentemente en los encabezados o en las firmas de las cartas, pues en ellos se encuentran frases hechas o convencionalismos de respeto hacia las autoridades. En el ejemplo (2b) observamos que Diego de Ordaz usa las abreviaturas en habla coloquial hacia su sobrino, como lo haríamos hoy en día.

Otra convención editorial consiste en la presencia de diagonales (2a) para indicar el cambio de línea en el documento original, en los tres textos se presenta de la misma manera. También, en los tres textos se moderniza la puntuación para facilitar la lectura de los mismos.

Por último, otra de las convenciones que cabe mencionar es la de omisiones del texto original. Se marca con puntos suspensivos dentro de corchetes y es un fenómeno que se da por dos diferentes razones: a) es obvio que los documentos del siglo XVI han sido rescatados y conservados por diferentes medios. Estos medios no son siempre los

ideales, por lo que el documento en cuestión tiende a dañarse. En este caso puede ser que un fragmento del documento no esté presente o esté demasiado dañado como para ser legible; b) cuando el documento no es comprensible por sí mismo, es decir, cuando las grafías son ilegibles a causa del amanuense o cuando hay correcciones (tachones o sobreescritura) hechas por el que redactó como vemos en el ejemplo 3 y marcadas entre corchetes.

- (3) a. Resulta culpa contra él de *haver* cometido muchos/ adulterios, incestos y stupros y *haver* tenido *açeso*/ con su propia hija diciéndole que no era peccado des[...]/ Consta por el dicho de la *propia* hija (*PI*, Caso contra Diego Díaz, f. 2)

En ocasiones se pueden inferir las palabras que faltan y estas aparecen con cursivas para señalar que son reposiciones hechas por el editor. Por supuesto, estas palabras no son objeto de estudio lingüístico. De las convenciones editoriales anteriores, sólo las correcciones del editor (los cambios de grafías) influyen en el análisis lingüístico y hay que tomarlas en cuenta para no hacer juicios erróneos. Debe tomarse en consideración, asimismo, que desatar las abreviaturas y poner fragmentos faltantes en el documento original es responsabilidad del editor y que en algunos casos no es válido en el análisis lingüístico; por ejemplo, en los casos en los que las frases se encuentran en proceso de lexicalización (*vuestra merced* > *usted*). Aunque en otros casos sí pueden proporcionar evidencia fonética, como en el caso de la palabra *veedor*, cuya segunda *e* está repuesta por el editor, lo cual nos sugiere que el hablante sólo pronunciaba una sola *e* (véase §4.2.2).

Todas estas convenciones editoriales deben tenerse en cuenta para que pueda realizarse una investigación seria. Estos tres conjuntos de textos se unen ahora para llevar a cabo un estudio de correlaciones entre las grafías y el sonido del español hablado en México en el siglo XVI y es invaluable el trabajo de los editores que respetan los textos originales para que con ellos se puedan hacer estudios lingüísticos, pues en

este caso, la variación gráfica proporciona elementos para determinar la vigencia de algunos sonidos.

Como mencioné anteriormente, existen ocasiones en las que la escritura disfraza la pronunciación real, pues el carácter de los documentos de trabajo, por estar escritos por amanuenses o notarios, reflejan, no la pronunciación del que escribe, sino los hábitos de escritura que este poseía.

## **4.2. Sonidos vocálicos**

En latín, las vocales se distinguían entre breves y largas (ōs, vocal larga, *boca*; ős, vocal breve, *hueso*). Al paso del tiempo perdieron este rasgo para diferenciarse según su timbre; así, una vocal breve se escuchaba abierta y una vocal larga, cerrada (Lathrop 1984/1992:Cap.2). En español actual ya no se distingue este rasgo de apertura; son cinco únicas vocales que, abiertas o cerradas, largas o breves, ya no nos dan información fonológica para distinguir pares de palabras como sucedía en latín. El sistema vocálico del español es bastante estable, pero sí podemos encontrar en el español hablado en México en el siglo XVI diferentes formas para graficar las vocales. A continuación veremos cómo se presentan en el corpus.

### **4.2.1. Vocales altas**

Es un hecho conocido que las vocales altas son las que presentan mayores variaciones en cuanto a su representación gráfica. Esto se manifiesta de manera general en los textos del español en el siglo XVI y también lo comprobamos en los textos del corpus, como veremos a continuación.

#### 4.2.1.1. Vocal alta anterior [i]

En el corpus de trabajo, se puede ver que la vocal alta anterior [i] está graficada de tres maneras diferentes: una convencional o prototípica *i*, la *i* griega *y* y, finalmente, la *i* larga o *j*, que todavía se encontraba desempleada. Me refiero a *desempleada* en el sentido de que en ese momento no representaba, como hoy, al fonema velar fricativo sordo /x/, pues aún estaba en proceso la velarización del fonema prepalatal fricativo sordo /š/, como veremos en §4.3.4. En (4) se presentan muestras con cada tipo de grafía.

- (4)
- a. Por interpretación de los padres Frai Joan López, Vicario en el monasterio del Chimalhuacán, y/ de Juan González, clérigo... (*PI*, Caso contra Diego Díaz, f.4)
  - b. Porque luego en salyendo de la canal de Bahama, a la primera tormenta se ahogaron las gallynas, i después quedamos a toçino i queso (*DO*, Carta 1)
  - c. Siervos humjldes i continuos capellanes (*DLNE*, 36, 1569)<sup>3</sup>

En (4a) vemos una *i* prototípica en *frai* (que actualmente se graficaría con *y*); en (4b), la palabra *gallynas* incluye, como se ve, la *i* griega *y* cuyo valor es vocálico [i] y en (4c) apreciamos tanto el uso de la *j* con valor de [i] en *humjldes*, como una *i* que actualmente se graficaría con *y* pues está funcionando como conjunción.

Los puntos de encuentro entre grafías ponen en evidencia el hecho de la complicación de las representaciones ortográficas, como lo menciona Lathrop:

En la serie palatal la vocal *i* presentaba los mismos inconvenientes gráficos [que la graficación de la /u/]. Para la vocal se utilizaban los signos *i*, *j*, *y*. Con valor consonántico se utilizaban *y* por un lado y *j*, por otro. La alternancia en las vocales dependía un poco de escribas y tipos de letras (1984/1992:217).

García de Diego también habla sobre el uso gráfico de la *y* griega “y era corriente, formara o no diptongo, después de otra vocal, *cuydado*, *treynta*: era frecuente en posición inicial, *ynfante*, *yqual*, y como primera de un diptongo, *syempre*, pero en los textos preclásicos abunda en cualquier otro caso, *rryco*, *myo*, etc.” (1970:54). Estos datos se reflejan en las reglas obtenidas en el presente trabajo de investigación.

---

<sup>3</sup> Las negritas en los ejemplos son mías para resaltar las grafías bajo análisis.

Las reglas introducidas a continuación tienen las siguientes equivalencias según los símbolos utilizados: **V** vocal, **C** consonante, **#** margen de palabra (principio o final), **Ø** vacío gráfico o fónico, **\_** cualquier segmento.

Con el análisis detallado de cada grafía, determiné, según los contextos, las siguientes reglas de correspondencia grafía-sonido para identificar sonidos con valor vocálico [i] conforme a las grafías *i*, *y* y *j*.

### **Grafía *i* = [i]**

La *i* gráfica no representa mayores problemas (no puede representar otro sonido que no sea el de la vocal alta anterior) y, por ello, siempre tendrá una correspondencia con el sonido [i]:

| <b>Contexto</b>      | <b>Regla</b> |
|----------------------|--------------|
| Una <i>i</i> gráfica | $i = [i]$    |

### **Grafía *y* con valor vocálico [i]<sup>4</sup>**

Una *y* gráfica tendrá su correspondencia con el sonido vocálico [i] en los siguientes casos:

| <b>Contexto</b>                                                                | <b>Regla</b>                  |
|--------------------------------------------------------------------------------|-------------------------------|
| Si es una conjunción; es decir, si está entre dos espacios vacíos              | $\emptyset y \emptyset = [i]$ |
| Si está entre una vocal y una consonante                                       | $VyC = [i]$                   |
| Si es inicio de palabra y le sigue una consonante                              | $\#yC = [i]$                  |
| Si es final de palabra, no importando si está precedida por vocal o consonante | $y\# = [i]$                   |
| Si está antecedida por una consonante, sin importar qué grafía siga            | $Cy = [i]$                    |

### **Grafía *j* con valor vocálico [i]**

Una *j* gráfica sonará con valor vocálico [i] en los siguientes casos:

<sup>4</sup> Para las grafías *y* y *j* con valor consonántico, véase §4.3.4. y 4.3.5.

**Contexto**

Si está entre una consonante y una vocal<sup>5</sup>  
 Si es inicio de palabra y le sigue una consonante  
 Si está entre dos consonantes  
 Si es final de palabra y va precedida de una consonante

**Regla**

CjV = [i]  
 #jC = [i]  
 CjC = [i]  
 Cj# = [i]

Como en todas las reglas hay excepciones; en este caso, aparecieron dos palabras con grafías *y* y *j* con valor vocálico que no se ceñían a estas reglas; sin embargo, encajaban muy bien con las reglas mediante las cuales se asignaría un valor consonántico a estas grafías, veamos los ejemplos en 5 en los que se ilustran estas palabras:

- (5) a. este *testigo oja* cómo ella le dezía (*DLNE*, 55, 1576)  
 b. conservarse *yan* muchos años (*DLNE*, 1, 1525)

En (5a) podemos ver, por el contexto, que esta *j* en *oja* tiene valor vocálico; es decir, es [o·í·a] —pretérito imperfecto del verbo *oír*— y no [ó·xa] —*hoja*, sustantivo. Este contexto no recae en ninguna de las reglas descritas anteriormente (incluso, coincide con las reglas que asignan valor consonántico a *j* pues esta grafía aparece entre dos vocales, como se señala en §4.3.4.). Por ello, se le deberá asignar un valor vocálico [i] a esta grafía a través de una lista de excepciones.

Por otra parte, en el ejemplo (5b) tenemos un contexto en el que prototípicamente también se le asignaría un valor consonántico a la *y* (pues es inicio de palabra y enseguida hay una vocal, véase §4.3.5.), sin embargo, se asignará *ían* en el atributo **n** (neutralización) de la etiqueta *tok* mencionada en el capítulo 2. Este ejemplo deberá leerse [í·an] —futuro analítico *conservarse hían* (de *haber*)— y no [yán], palabra que no existe en la lengua española. También por medio de una lista de excepciones se daría valor vocálico a esta grafía en dicho contexto.

<sup>5</sup> Con excepción de los casos en la lista de excepciones del sonido [x].

#### 4.2.1.2. Vocal alta posterior [u]

La vocal alta posterior [u] se presenta en el corpus de tres diferentes maneras: la prototípica *u*, como *v* (sólo en las *Cartas de Diego de Ordaz* y en los *Procesos Inquisitoriales*, como se mencionó en §4.1.), y muy pocas veces, limitada a contextos determinados, con una *b*.

- (6)
- a. Preguntada si hera bautizada. Dixo *que* sí, y *que* la bautizó Luis/ Méndez, clérigo, en su casa del dicho Luis, puede *haver* siete años/ y medio *que* fue bautizada/ (*PI*, Caso de Diego Díaz, f. 5)
  - b. I esto me a cavzado la inchazón de las piernas. Pero, loado Dios, no lo tengo en nada, *que* cada día es menos (*DO*, Carta 1)
  - c. cerca /<sup>24</sup> de duzientas leguas desta çibdad de Tenustitan, halló los yndios de guerra (*DLNE*, 1, 1525)

En el ejemplo (6a) vemos graficada la [u] con una prototípica *u*, *bautizada*. En el ejemplo (6b) podemos apreciar la grafía *v* para el fonema vocálico [u], *cavzado*, y en el ejemplo (6c) observamos una de las pocas palabras, *çibdad*, que todavía en el siglo XVI conservaba la grafía *b* para representar a la vocal posterior alta [u]. Con este último ejemplo podemos apreciar que no existe, necesariamente, una correlación entre el sonido y la grafía; por ello, podemos afirmar que podía ser que existiera vacilación en la pronunciación dependiendo del contexto de comunicación. Es decir, podía ser que el hablante pronunciara, de hecho, [sib·dád].

Las grafías *v* y *b*, no prototípicas para este sonido vocálico [u], se presentan por diferentes motivos: la *v* como resultante de un calco del latín, que, en mayúsculas, que era como se escribía en esa lengua, no poseía una grafía U, sólo V, así que desde entonces había una confusión en las grafías como lo menciona Lathrop:

A partir del siglo I de nuestra era, encontramos la *v* clásica transcrita como beta [̢ = b]. En latín vulgar esta [b] se convirtió en la pronunciación normal tanto para *b* como para *v*, lo que originó una confusión total entre ambas a la hora de escribirlas (1984:30).

La *b*, en algunas palabras, se da como conservación de la grafía derivada del latín, antes de que el alófono velar agrupado se vocalizara. En el ejemplo (6c) atestiguamos un residuo gráfico del latín CIVITATE > *çibdad* cuya transcripción fonética podría ser ya [sjʉ·dád]. “Perduró también en la primera mitad del siglo XVI la conservación, muy decadente, de algunos grupos de consonantes que en el habla llana se habían simplificado o transformado (*cobdicia, cibdad*)” (Lapesa 1981:369).

A continuación proporciono la manera como quedaron establecidas las reglas de correspondencia arrojadas por el análisis para esta grafía.

### **Grafía *u* = [u]**

La *u* gráfica siempre tendrá una correspondencia con el sonido [u]:

**Contexto**  
Una *u* gráfica

**Regla**  
*u* = [u]

Las grafías *v* y *b* para su correspondencia al sonido vocálico no son sistemáticas, de manera que se incorporarán al transcriptor por medio de una lista de excepciones, primero, como se dijo arriba, porque en los textos no se presentan de la misma manera (en los *DLNE* la editora lo cambió a *u*) y segundo porque no son sistemáticas en el contexto referido, es decir, se presenta la misma palabra con *u* y *v* (en el ejemplo (6c) aparece *çibdad*, pero en los textos encontramos *ciudad* y en la siguiente lista de excepciones también se presenta *civdad*).

#### **Lista de vocablos con *v* cuyo valor vocálico es [u]**

|                 |                |
|-----------------|----------------|
| avdiencia       | revsado        |
| avn / avnque    | rrecavdo       |
| avsencia        | v (conjunción) |
| cavsado / cavsa | vn / vno / vna |
| çivdad          | vnçiones       |
| Chiavtla        | Vngría         |
| enbavquen       | vse / vsar     |
| Guaxovçingo     |                |

#### **Lista de vocablos con *b* cuyo valor vocálico es [u]**

|           |         |
|-----------|---------|
| Abdiencia | cibdad  |
| abnque    | debdas  |
| babtizó   | recabdo |

#### **4.2.2. Vocales medias**

En términos generales la correspondencia grafía-sonido de las vocales medias es bastante estable. Por ejemplo, para la vocal media anterior [e] la correspondencia gráfica siempre es *e*.

#### **Grafía *e* = [e]**

La *e* gráfica siempre tendrá una correspondencia con el sonido [e]:

#### **Contexto**

Una *e* gráfica

#### **Regla**

*e* = [e]

Sin embargo, en algunas ocasiones la vocal media anterior se presenta con doble grafía. Como sabemos, en las palabras compuestas que presentan doble *e*, fonológicamente se representan las dos, por ejemplo, *reencontrarse* se transcribe /*re·en·kon·trár·se*/. En el corpus se presentó la doble grafía de manera anómala, cuya correspondencia al sonido debería ser una sola [e], pues se encuentran en el corpus estas mismas palabras graficadas también con una sola *e*: *fe*, *ven*, *ver* y *veedor* tiene en varias ocasiones una de las *e* repuesta por el editor. También deberán incorporarse al transcriptor mediante una lista de excepciones:

#### **Lista de vocablos con *ee* cuyo valor es [e]**

|        |                            |
|--------|----------------------------|
| fee    | veen / vee / veer (de ver) |
| veedor |                            |

Para la vocal media posterior [o] no hay ninguna variación en cuanto a su representación, siempre se representa con una *o* gráfica. De tal suerte, se puede establecer una única regla de correspondencia entre esta grafía y su fonema correspondiente, como se ilustra a continuación.

### **Grafía *o* = [o]**

La *o* gráfica siempre tendrá una correspondencia con el sonido [o]:

**Contexto**  
Una *o* gráfica

**Regla**  
*o* = [o]

### **4.2.3. Vocal baja**

La vocal baja es la más estable en cuanto a su representación gráfica. Desde el latín, esta vocal era la que se mantenía de manera más sistemática aunque los segmentos adyacentes sufrieran cambios. En el corpus de trabajo siempre aparece graficada como *a* sin variar en ningún contexto. Es por ello que, al igual que la vocal media posterior, se hizo una sola regla para la correspondencia entre esta grafía y su sonido.

### **Grafía *a* = [a]**

La *a* gráfica siempre tendrá una correspondencia con el sonido [a]:

**Contexto**  
Una *a* gráfica

**Regla**  
*a* = [a]

### **4.2.4. Variación de vocales átonas.**

Lo que sí se presentó en el corpus de manera más o menos regular es la variación de vocales átonas. Se presenta entre vocales posteriores, alta [u] y media [o], y entre

vocales anteriores, también alta [i] y media [e]. En el transcriptor se respetará la vocal graficada y no se intentará cambiarlo, puesto que en el siglo XVI era una práctica de estatus social, que se daba entre escritores. Alatorre (1979/2003:316) menciona que la variación de vocales átonas más frecuentes eran de la *i* a la *e* como en *vanedad*, *enclinar*, *recebir*, y de la *u* a la *o*, como en *mormurar*, *mochacho*, *sofrir*. Estos desplazamientos se encuentran en todos los escritores del siglo XVI; los desplazamientos inversos, de la *e* a la *i* y de la *o* a la *u*, también eran frecuentes pero se limitaban al lenguaje coloquial, por ello también podremos encontrarlos en el corpus. Sin embargo, como menciona Lapesa (1981:368) estas vacilaciones de timbre en las vocales no acentuadas disminuyen en el transcurso del siglo XVI.

En el corpus de trabajo se encontraron reposiciones del editor como las siguientes: *ansimesmo*, *trespasar*, *monesterio*. Vemos que aquí está repuesta una *e* por el editor, por lo que el autor sugiere que era una *e* la que se pronunciaba en ese contexto, es decir, que si bien, hoy escribimos una *i* (por *así mismo*) o una *a* (*traspasar*, *monasterio*) en las reposiciones de los ejemplos, en ese momento histórico la forma prototípica y general, podría haber sido como lo repone el editor, pues había gran vacilación de timbre en vocales átonas.

### **4.3. Sonidos consonánticos**

En esta sección abordaremos el problema que representa la graficación de las consonantes en los textos del siglo XVI. Algunas de ellas (como las sibilantes) son las que presentan mayor vacilación gráfica dependiendo del amanuense (su grado cultural o educativo y su origen geográfico) y del año en que fue escrito el documento en cuestión, como lo vimos en el §4.1.

Al igual que en el caso de las vocales, se expondrán algunas opiniones teóricas de los estudiosos, la forma en la que se presentaron en el corpus de trabajo, las reglas de correspondencia grafía-sonido y las listas de excepciones según sea el caso. Además, dentro de algunos sonidos que presenten mayor grado de dificultad o que estén inmersos en cuestiones teóricas mayores, se abordarán temas específicos, como el caso de la confusión ortográfica entre *b* y *v*, el seseo, la aspiración, el yeísmo y se mencionará el andalucismo de manera general, pues ya se mencionó en el capítulo anterior, §3.3., y se remitirá al lector al parágrafo correspondiente.

#### 4.3.1. Consonantes oclusivas sordas

Las consonantes oclusivas sordas se mantienen bastante estables. La bilabial [p] y la dental [t] ofrecen ciertas variantes, sin presentar mayor dificultad. La velar [k] sí presentó alternancia significativa en sus graficaciones, algunas iguales a las que se usan actualmente para graficarla (*c*, *qu*), otras grafías como calcos del latín, como arcaísmos, que son residuo de las grafías griegas (*ch*, *x*), esto se explicará más adelante.

La consonante bilabial presentó dos maneras de graficación: una prototípica *p* y otra doble, *pp*, posiblemente residuo gráfico de la geminada en latín. Su correspondencia, en ambos casos, será siempre [p]. Veamos los ejemplos del corpus:

- (7)
- a. Y ansi a lo *que* dixere y firmare en esto y otra qualquiera cosa, se <sup>/24</sup> le podra dar credito, porque aca se **approbará** y cumplira, ultra de que él <sup>/25</sup> solo es bastante *para* cumplillo (*DLNE*, 46, 1574)
  - b. I al Enperador i a toda España a **pareçido** muy mal lo *que* esos señores an hecho en meter tanto la mano en las cosas del Marqués (*DO*, Carta 3)

En el ejemplo (7a) podemos ver la doble grafía *pp* en la palabra *approbará*. Como dije antes, esto como un calco del latín, pues en esa lengua se escribía *APPROBARE*. Las otras palabras que presentan *pp* en el corpus son *supplicar* del latín

SUPPLICARE y la palabra *apostólico(a)* del latín APOSTOLICUS que, aunque no lleva doble *p*, parece haber sido influenciada también por la geminada del latín.

En el ejemplo (7b) se pueden ver varios ejemplos de las prototípicas y generales *p*, que no representarán mayor problema para el transcriptor. A continuación presento las reglas que se implementarán a este.

### **Grafía *p* = [p]**

La *p* y *pp* gráficas siempre tendrán una correspondencia con el sonido [p]:

#### **Contexto**

Una *p* gráfica

Aparición de *pp* gráfica

#### **Regla**

*p*= [p]

*pp*= [p]

La única excepción para esta grafía será cuando *p* aparezca seguida de *h*, pues su correspondencia sonora será *ph* → [f] en palabras como *filosofía*, *Pharna* y *Joseph*. La lista de excepciones aparecerá en el apartado correspondiente al sonido labiodental fricativo sordo [f], en §4.3.4.

Para la consonante dental tampoco se encontraron vacilaciones significativas. Al igual que para la bilabial, apareció la doble grafía, *tt* y la prototípica *t*. Esta vez, sin embargo, también se encontró la grafía cultista *th*. En los tres casos su correspondencia sonora será [t]. Observemos los ejemplos siguientes extraídos del corpus:

- (8)
- a. traxo consigo nueve yndias con/sigo por mançebas, teniéndolas por tales en su casa, y/ dándoles lo que habían menester, y las tubo/ todo el dicho tienpo no las permytiendo confesar ni de/xar yr a la yglesia porque no le descubriesen sus dellittos/ y maldades (*PI*, Caso de Diego Díaz, f. 25 v.)
  - b. y de consentimjento de Albor /<sup>33</sup>noz, que como estava mal con el **thesorero** deseava vengarse dél, holgo de no gobernar, porque /<sup>34</sup> no governase el **thesorero** (*DLNE*, 7, 1529)
  - c. Devriase de contentar Monjaraz con lo que tiene, pues no mereçe tener indio, según los a tratado (*DO*, carta 5)

En el ejemplo (8a) vemos la palabra *dellittos* con doble grafía cuya transcripción deberá ser [t] simple. La grafía cultista *th* en el ejemplo (8b) se puede observar en la palabra *thesorero* proveniente del latín *THESAURARIUS*. Y en el ejemplo (8c) vemos la prototípica *t* simple.

Enseguida podemos ver como quedarán las reglas de correspondencia grafía-sonido para el caso de la consonante dental oclusiva sorda [t].

### **Grafía *t*, *tt*, *th* = [t]**

La *t*, *tt* y *th* gráficas siempre tendrán una correspondencia con el sonido [t]:

#### **Contexto**

Una *t* gráfica

Aparición de *tt* gráfica

Aparición de *th* gráfica<sup>6</sup>

#### **Regla**

*t* = [t]

*tt* = [t]

*th* = [t]

Observemos a continuación la cantidad de grafías que se usaron en el siglo XVI para representar el sonido [k].

El fonema /k/ presenta distintos signos: *c*, *q*, *qu*, *ch*. El uso de *c* y *qu* es idéntico al que presenta nuestra escritura hoy; el grafema *q* aparece en *quando*, *quatro*, y estas formas se atestiguan así escritas incluso en textos del siglo XVIII. Grafía de tipo cultista es *ch*, constante en todo el corpus para la voz *Chancillería*, y la *x* de *xristiano*, que transcribo con *ch*, *christiano* (García Carrillo 1988:35).

En este párrafo García Carrillo (1988) menciona, mediante el análisis de su corpus, las grafías con las cuales se representaba el sonido velar oclusivo sordo [k] en el español de México del siglo XVI. Este análisis coincide con el mío, como veremos más adelante en la exposición de reglas, listas de excepciones y ejemplos; en mi corpus también encontré la grafía *x* para representar este sonido. En (9) presento los ejemplos de graficación del sonido velar oclusivo sordo:

<sup>6</sup> Esta regla se puede eliminar, a juicio del encargado de hacer el transcriptor en un futuro próximo, siguiendo la regla de la *h* → ∅ que se verá en §4.3.4. también se puede eliminar implementando una lista de excepciones pues las palabras con este dígrafo son pocas.

- (9) a. La careza de esta tierra me a hecho mucho dagño en los gastos. *Que* si el agño respondiera como se esperava, yo *quedara* a dever muy poco (*DO*, carta 9)
- b. Fray Juan de Çumarraga, primero obispo de la dicha çuidad, e *Ynquisidor Appostólico* contra la/ herética pravedad e apostasia en ella, y en *todo* su obispado (*PI*, Caso 2, f. 132)
- c. Todos los *christianos que* con él fueron y a quatro o çinco mill yndios *que* lle /<sup>12</sup>vaba en su conpañja de los desta çibdad y tierra (*DLNE*, 1, 1525)

En los anteriores ejemplos podemos observar las grafías *c*, *qu* y *ch* que representan al sonido velar oclusivo sordo [k]. Estas tres grafías son sistemáticas en el corpus, es decir, no representan mayor complejidad y no presentan lista de excepciones. El ejemplo (9a) es un caso en el que la *c* va seguida de a, o, u, *careza* y *como*; el ejemplo (9b) presenta una *qu* prototípica del habla actual, *Ynquisidor*, y el ejemplo (9c) muestra un caso en el que *ch* suena [k], pues va seguida de *r*, *christianos*.

Con estas grafías, que son prototípicas para representar al sonido velar oclusivo sordo [k], se pudieron establecer reglas de aparición de manera sistemática. Enseguida presento las reglas que otorgó el análisis para estas tres grafías.

### **Grafía *c* = [k]**

La *c* gráfica tendrá una correspondencia con el sonido [k] en los siguientes casos:

#### **Contexto**

Si le sigue una vocal central baja [a]  
 Si le sigue una vocal media posterior [o]  
 Si le sigue una vocal alta posterior [u]

#### **Regla**

*ca*= [ka]  
*co*= [ko]  
*cu*= [ku]

### **Dígrafo *qu* = [k]**

El dígrafo *qu* tendrá una correspondencia con el sonido [k] en los siguientes casos:

#### **Contexto**

Si el dígrafo va seguido de vocal media anterior [e]  
 Si al dígrafo le sigue una vocal alta anterior [i]

#### **Regla**

*que*= [ke]  
*qui*= [ki]

## Dígrafo *ch* = [k]

El dígrafo *ch* tendrá una correspondencia con el sonido [k] en los siguientes casos:

### Contexto

Si el dígrafo va seguido de consonante alveolar vibrante simple [r]

### Regla

*chr* = [kr]

Se observaron en el corpus algunos contextos que no se ajustaron a las reglas anteriores; por ejemplo, apareció el dígrafo *qu* en el que sí suena la vocal alta posterior, es decir, se transcribirá [ku], por ejemplo, en el caso de *quando*, *quatro*. García de Diego (1970:53) menciona que la *qu* era normal en los casos etimológicos ante *a*, sin embargo, también se encuentran casos por confusión y menciona el ejemplo *blanquo*.

El uso de *qu*, en palabras como *quando*, *qual*, etc., no se normó hasta el siglo XVIII en el cual la Real Academia Española otorgó las reglas de ortografía para dichas grafías. En el siglo XVI, como podemos observar en el corpus, había confusión de estas grafías y se usaban indistintamente (aunque la frecuencia de aparición de la *qu* ante *a* es limitada).

Se observó también el dígrafo *ch* seguido de vocal central baja [a] que suena [k] y no [ç]. Y la grafía cultista *x* (por calco de grafías del alfabeto griego) para representar el sonido velar oclusivo sordo [k], que en los *DLNE* la editora lo representa como *ch* en cursivas, para indicar que es reposición suya *X<sup>tianos</sup>* > *christianos*. Presento enseguida los ejemplos en los que se observa lo anterior:

- (10) a. Más va la provisión del peñón. Esto hazed **querdamente**. Ya os digo: si vierdes *que* os lo darán i *que* no pornán enbaraço, presentalda;/i si otra cosa vierdes, guardalda, *que* presto será otro mundo (*DO*, carta 3)
- b. roge al tenjente /<sup>3</sup> del allcalde mayor de aquj, *que* se dize Pedro Borjes, *que* /<sup>4</sup> me hiziesse tanta **charj**dad *que* me truxesse aquj /<sup>5</sup> al monesterio al dicho don Joan Xancoi (*DLNE*, 27, 1555)
- c. Preguntada si conoçe a su padre de la dicha Petronylla. Dixo/ *que* no, más de que oyó dezir al dicho Diego Díez *que* hera su hija dél, y *que*/ por tal la tenía, e *que* hera hija de una yndia *que* se fue con un *xrisptiano*/ a Guatimala (*PI*, Caso de Diego Díaz, f. 9 v.)

En el ejemplo (10a) observamos la palabra *querdamente* que encaja en la regla *qu + e= [ke]* mencionada arriba, sin embargo, esta palabra debe transcribirse [kuer·da·mén·te] por lo que se implementará al transcriptor por medio de la lista de excepciones. En el ejemplo (10b) la palabra *charjdad* deberá transcribirse [ka·ri·dád] y no [çâ·ri·dád], también se anexará al programa por lista de excepciones. Y en el ejemplo (10c) la grafía *x* corresponde, en este contexto, al sonido [k], es decir, se transcribirá [kris·tiá·no]. A continuación presento las listas de excepciones para estos casos:

**Lista de vocablos con *qu* cuyo valor es [ku]**

|           |             |
|-----------|-------------|
| aquerdo   | quenten     |
| çinquenta | quenta      |
| pesqueços | querdamente |
| qual      | quidado     |

**Lista de vocablos con *ch* cuyo valor es [k]**

|              |          |
|--------------|----------|
| Chançilleria | charjdad |
| Chancilleria |          |

**Lista de vocablos con *x* cuyo valor es [k]**

|             |            |
|-------------|------------|
| Jesuxristo  | xrisptiano |
| Jhesuxristo | Xrisptóval |
|             | Xristoval  |

#### 4.3.2. Consonantes oclusivas sonoras

Al igual que las consonantes oclusivas sordas, las oclusivas sonoras se mantienen bastante estables. La bilabial [b] y la dental [d] ofrecen dos variantes, cada una, sin presentar mayor dificultad. La velar [g] sólo presenta una graficación pero con condiciones de contexto fónico, pues la misma grafía *g* también puede corresponder al fonema velar [x], según su contexto, como se verá en §4.3.4.

La bilabial presentó dos maneras de graficación: *b* y *v*. Como se vio en el capítulo anterior, en §3.1., existe una discusión teórica acerca del estatus fonológico del

fonema /b/. Para efectos prácticos, en el corpus de trabajo se transcribirá siempre como [b] una grafía *v*, de acuerdo con los teóricos que indican que sólo existió un solo sonido [b] y que en él confluyeron las dos grafías *b* y *v*.

Recordemos, sin embargo, que en §4.2.1.2. se asignó también un valor vocálico a *v*, de manera que las reglas de correspondencia se eliminarán entre sí; es decir, primero se aplicará la lista de excepciones en las que la grafía *v* tiene valor vocálico [u] y después todas las demás grafías *v* que se encuentren se transcribirán como [b].

Veamos algunos ejemplos del corpus:

- (11) a. y a dicho que es ynmortal, y *que/* ha **hablado** muchas vezes con el diablo de noche, e ha hecho y dicho otras muchas/ cosas contra *nuestra* santa fee cathólica (*PI*, Caso 2, f. 132)
- b. En esa carta *que va* con ésta os **escrivo** largo de lo *que conviene*. Ésta es *para vos* solo, i leída, la ronpáis (*DO*, carta 4)
- c. compró al dicho Martyn una viga grande de a pies, porque en su tierra/ no la tiene, y le dio por ella çient cueros de venados, y **llebando**, *que/* se los llevaron al dicho Martyn a su casa (*PI*, Caso 2, f. 132 v.)

En el ejemplo (11a) se aprecia la grafía *b* en vocablos como *hablado* y *diablo*.

En el ejemplo (11b) se observa la grafía *v* con valor consonántico [b] en palabras como *va*, *escrivo*, *conviene* y *vos*. El ejemplo (11c) ofrece una misma palabra con las dos graficaciones para este fonema, *llebando* y *llevaron*. Gracias a estos ejemplos constatamos que no hay mayor dificultad para la transcripción de este fonema. Las reglas que se implementarán al transcriptor las presento a continuación.

### **Grafía *b* = [b]**

La *b* gráfica siempre tendrá una correspondencia con el sonido [b]:

**Contexto**  
Una *b* gráfica

**Regla**  
*b* = [b]

### **Grafía *v* = [b]**

La *v* gráfica tendrá una correspondencia con el sonido [b], con excepción de los casos en los que dicha grafía tenga valor vocálico (véase §4.2.1.2.):

#### **Contexto**

Una *v* gráfica

#### **Regla**

*v* = [b]

La consonante dental oclusiva sonora no presenta mayores dificultades para su inclusión en el sistema de correspondencias grafía-sonido, se encontró graficada con *d* y la doble grafía *dd*. Veamos algunos ejemplos representativos que se hallaron en el corpus:

- (12) a. pareció **don** Joan, el cacique/ de Xiutepeque, y **dixo** que por **descargo de** su conciencia, él **quería/ dezir** lo que sabe **de** un hermano **de** Martín Uzelo, **que** se llama Miscoa/tle, y **de** otro hermano suyo, **que** se **dize** Tlaloc (*PI*, Caso 4, f. 183)
- b. El más obidiente hijo que *v. mddes.* tienen, que sus manos besa.\\ \\ Alonso de Salazar (*DLNE*, 71)

La grafía prototípica *d* se encuentra en varias palabras del ejemplo (12a), la doble grafía solamente se encontró en esta fórmula *v. mddes.* (ejemplo de 11b) que representaría en la normalización **n** de la etiqueta tok (mencionada en §2.2.) *vuestras mercedes* y que se encuentra presente en los *DLNE*. Esta fórmula seguramente se hallaba en proceso de lexicalización y no podemos determinar a ciencia cierta en qué etapa del proceso se encontraba, por lo tanto no se sabe qué transcripción fonológica sea reflejo real de la fórmula.

Como vimos en los ejemplos anteriores, la correlación entre grafía y sonido para [d] no representa mayor problema y las siguientes son las reglas que arrojó el análisis.

### **Grafía *d*, *dd* = [d]**

La *d* y *dd* gráficas siempre tendrán una correspondencia con el sonido [d]:

**Contexto**  
Una *d* gráfica  
Aparición de *dd* gráfica

**Regla**  
*d*= [d]  
*dd*= [d]

El sonido velar oclusivo sonoro [g] sólo se puede representar con la grafía *g* en determinados contextos:

- (13) a. este dicho Andrés, en qualquie/ra parte que va tiene una costumbre *que* comulga a la gente, y él mesmo co/mulga, y esto haze de unos honguillos que se llaman en su lengua na/nacatl, que es cosas endiablada (*PI*, Caso 4, f. 186)  
b. Ya me puedo engañar, i así quiera Dios. Esto del juro os encargo mucho, porque ruines no se vengen de mí, *que* an dicho *que*, pues vendo el juro, no tengo allá nada (*DO*, carta 8)

En el ejemplo (13a) podemos observar casos de *g* empleada como hoy en día y que corresponden al fonema velar oclusivo sonoro [g], es decir, seguida de *a*, *o*, *u*; o bien *gu* seguida de *i*, *e*. Esto en palabras ilustradas en el ejemplo (13a), como *comulga* y *honguillos*.

En el ejemplo (13b), sin embargo, encontramos una *g* con igual valor [g] pero sin el contexto adecuado para la ortografía actual. Sin asignarle su valor real por medio de una lista de excepciones, esta palabra caería en el contexto erróneo como se podrá ver en el §4.3.4. Por medio de esta lista se asignará la transcripción [bén·gen] y no la incorrecta [bén·xen].

En el corpus se observó que esta grafía sólo aparece tras algunas consonantes específicas: *r* y *l*. Veamos las reglas asignadas.

### **Grafía *g* = [g]**

La *g* gráfica tendrá una correspondencia con el sonido [g] en los siguientes contextos:

**Contexto**

Una *g* gráfica seguida de *a, o, u, ü*  
 El dígrafo *gu* seguido de *i, e*<sup>7</sup>  
 Una *g* gráfica seguida de las líquidas *r, l*

**Regla**

*g* = [g]  
*gu* = [g]  
*g* = [g]

Para casos como el ilustrado en el ejemplo (13b) se incorporará al transcriptor la siguiente lista de excepciones:

**Lista de vocablos con *g* sin *u* cuyo valor es [g]**

|          |           |
|----------|-----------|
| Diegito  | plégele   |
| Gerra    | Rodríguez |
| Gíe Page | Ruegen    |
| Higeras  | Vengen    |
| Migel    |           |

**4.3.3. Consonantes africadas sordas**

La grafía *ch* es la representación gráfica prototípica para representar un sonido africado, en este caso palatal /ç/. Se llama africado porque tiene un primer momento de oclusión y después uno de fricción. En el siglo XVI también se encontraba un sonido prepalatal africado sordo [ʃ̥] representado por la grafía *ç*, que provenía de la convivencia de dental o velar sordas con yod en latín.

Vimos en el capítulo anterior §3.1., que el sonido prepalatal africado sordo [ʃ̥] se encontraba aún vigente en el español que llegó a América. Sin embargo, en el corpus de trabajo alterna la grafía para este sonido [ʃ̥] con las grafías *s, c, z*, utilizadas también para representar el alveolar fricativo sordo [s], además de aparecer graficado con *tz*, y con esto podemos intuir que el proceso de confusión de estos fonemas se encontraba avanzado.

<sup>7</sup> Se encontró una excepción a esta regla en *verguença*, por la falta de diéresis en el documento original.

Por lo anterior y para efectos prácticos del transcriptor, la grafía  $\zeta$  se transcribirá como [s] excepto en los casos claros en los que el sonido está aún vigente por su utilización en léxico indígena, la mayor proporción en topónimos y antropónimos.

Veamos los ejemplos:

- (14) a. el sobredicho Miscoatle fue a un pueblo sujeto/ deste *que* depone y demando las minas tres cosas *para* hazerles/ dichas zerimonias e no *queriendo* darçelas fue uno del pueblo (*PI*, Caso 4, f. 183)
- b. Pasá luego a Guaçoçingo las yeguas i todas esas menudençias, i tened mucha vijilançia en el buen tratamiento de los indios, pues es rrazón i aquéllos lo mereçen (*DO*, carta 5)

En el ejemplo (14a) se ilustra una  $\zeta$  con valor de [s] en la palabra *darçelas*, el ejemplo (14b) muestra graficaciones con  $\zeta$  para [s] y también, en el topónimo *Guaçoçingo*, para [ʂ]. Las siguientes son las reglas arrojadas por el análisis para esta grafía.

### **Grafía $\zeta$ = [ʂ]**

La  $\zeta$  gráfica tendrá una correspondencia con el sonido [ʂ], en léxico indígena:

#### **Contexto**

Una  $\zeta$  gráfica en léxico indígena

#### **Regla**

$\zeta$  = [ʂ]

Esta regla se aplicará de manera general a todo el léxico indígena que aparezca en el corpus (si el corpus se amplía la misma regla deberá abarcar el léxico indígena que se incorpore), pero por el momento se tiene la siguiente lista de excepciones:<sup>8</sup>

#### **Lista de palabras de origen indígena con $\zeta$ cuyo valor es [ʂ]**

|                       |            |
|-----------------------|------------|
| Coaçacoalco           | Tepedçingo |
| Cuçamelco             | Tilançi    |
| Guaçoçingo/Guaxoçingo | Yçucar     |
| Matalçingo            |            |

<sup>8</sup> Se encontró la palabra *maçumarra* en el corpus y no pude rastrear su origen, así que, por lo pronto, la  $\zeta$  se transcribirá como [s].

Las dificultades que ofrece el léxico indígena, la permanencia temporal de algunos sonidos y la posible influencia del náhuatl en el sistema fonológico español se verá ampliamente en el §4.4.1. de este mismo capítulo.

La otra consonante africada sorda es la que corresponde al sonido palatal africado sordo [ç] graficado con el dígrafo *ch*. Esta consonante se muestra bastante estable desde que se origina en castellano medieval. En el corpus de trabajo siempre se muestra con ese dígrafo y suena [ç] con excepción de los ejemplos mostrados con valor de velar oclusiva sorda [k] mencionados anteriormente, en §4.3.1.

Veamos los ejemplos que se presentaron en el corpus para este sonido, en los cuales se apreciará la estabilidad de la correspondencia entre el dígrafo *ch* y su sonido:

- (15)
- a. I criad en Guaçoçingo muchos puercos, i muchas ovejas, i todo hazed de manera *que* los indios sean bien tratados, i hagáis en esto ventaja a *vuestros* vezinos (*DO*, carta 5)
  - b. *Preguntado* si éste *que* declara y el *dicho* Marcos si andan dis/famando los frayles, deziendo cosas deshonestas contra/ ellos, en espeçial diziendo *que* la corona de los frayles,/ *desque* se quitan a la capilla, paresçe a la natura de los hon/bres (*PI*, Caso 8, f. 144)
  - c. Algunos, çesarea *magestad*, de los *que* han estado en estas partes y experimentado las co /<sup>25</sup>sas dellas, y por lo *que* se pierden y vienen en dimjnuysion, dizen es porque las hallan /<sup>26</sup> fertiles y ricas y no cura njnguno, del mayor hasta el menor, sino tener ojo /<sup>27</sup> a *aprovecharse* y a *procurar* de aver con *que* se vaya en España (*DLNE*, 1)

En los tres ejemplos anteriores se observa el dígrafo *ch* con valor de [ç] en vocablos como *muchos*, *dicho* y *aprovecharse*. Esta correspondencia no ofrece mayor dificultad, por lo cual se asignó la siguiente regla de correspondencia entre el dígrafo y el sonido palatal africado sonoro [ç].

**Dígrafo *ch* = [ç]**

El dígrafo *ch* siempre tendrá una correspondencia con el sonido [ç]:

**Contexto**  
Aparición de *ch*

**Regla**  
*ch*= [ç]

#### 4.3.4. Consonantes fricativas sordas

Este grupo de consonantes es el que ofrece mayor dificultad para hacer reglas de correspondencia. Muchos de los sonidos que integran esta correlación se vieron envueltos en discusiones teóricas, como se mencionó en el capítulo anterior, por los problemas que representaba su graficación en el periodo del siglo XVI.

A grandes rasgos, los problemas que plantean son los siguientes: el labiodental [f] se aspiró [h] en palabras con *f*- inicial latinas y posteriormente desapareció Ø; en determinados contextos, se reforzó [x] y en otros se conservó<sup>9</sup>. El prepalatal [š] evolucionó en un sonido velar [x] (en algunos casos se confunde con el fonema originado por el refuerzo de la aspiración de *f*- inicial latina mencionado arriba) y, en algunos casos, sobre todo en léxico indígena, es alófono del alveolar [s]. Este último sonido se vio envuelto en la discusión acerca del fenómeno conocido como seseo, característico del español andaluz y, por lo tanto, del español atlántico.

Empezaré por el sonido labiodental fricativo sordo [f]. Este se encontró graficado con la prototípica *f*, con la doble grafía *ff* y con el dígrafo cultista *ph*. A continuación presento algunos ejemplos:

- (16)
- a. I acá pidieron los flaires *que* de allá vinieron vna çédula para vn valençiano *que* está en Guaçoçingo, para dalle çiertas tierras (*DO*, carta 3)
  - b. el/ dicho don Luys no pudo yr, y enbió a este testigo a ver qué es lo que quería el dicho Martyn, y/ al dicho Diego, e que después de ffecha la dicha fiesta, que podía haser çinco meses,/ poco más o menos, el dicho Martyn tomó a éstos que deponen, e los puso (*PI*, Caso 2, f. 132)
  - c. Para *que* los hijos de los caçiques y señores, muy poderoso señor, se ynstruyan en la fee, /<sup>9</sup> ay neçessidad nos mande *vuestra majestad* se haga un colegio donde les muestren a leer y gra /<sup>10</sup>matica y **philosophia** y otras artes (*DLNE*, 1, 1525)

---

<sup>9</sup> Véase la cita de Lathrop en la página 42.

En el ejemplo (16a) tenemos la grafía simple en *flaires*, palabra en la que se observa un trueque de líquidas que se verá en §4.2.2. de este mismo capítulo. El ejemplo (16b) muestra la doble grafía en *ffecha* y el ejemplo (16c) ofrece la grafía cultista en un término de origen griego *filosofía* en el que se conserva el dígrafo por calco etimológico.

Según García de Diego (1970:54) la duplicación de *f* en palabras compuestas era frecuente en esa época, sin embargo, vemos en el ejemplo (16b), extraído del corpus, que la doble grafía aparece en palabra simple y en posición inicial de palabra. Es decir, no cabe en la descripción que hace este autor.

Este mismo estudioso también nos habla sobre la utilización de la grafía cultista *ph* y menciona que “eran simples signos etimológicos, muchas veces empleados sin acierto” (García de Diego 1970:54). En el ejemplo (16c) del corpus la grafía cultista fue utilizada correctamente.

Ahora bien, algunas de las *f*- iniciales se transcribirán como sonidos aspirados pues se encontraban en proceso de desaparición (véase el §3.1.), esto se implementará al transcriptor mediante una lista de excepciones, misma que se exhibirá en la parte correspondiente al sonido glotal aspirado sordo [h].

En seguida ilustraré cómo fueron las reglas obtenidas con base en el análisis del corpus para transcribir estas grafías.

### **Grafía *f, ff* = [f]**

La *f* y *ff* gráficas tendrán una correspondencia con el sonido [f]:

**Contexto**

Una *f* gráfica<sup>10</sup>  
 Aparición de *ff* gráfica

**Regla**

*f* = [f]  
*ff* = [f]

**Dígrafo *ph* = [f]**

El dígrafo *ph* siempre tendrá una correspondencia con el sonido [f]:

**Contexto**

Aparición de *ph*

**Regla**

*ph* = [f]

Al sonido alveolar fricativo sordo [s] se le encuentra en el corpus de trabajo graficado de múltiples maneras: con *s*, con doble grafía *ss*, con *c* ante vocales anteriores, con *z*, y en algunos casos, por confusión, con *ç* y con *x*. A continuación vamos a ver ejemplos de cada grafía:

- (17)
- a. de allí fue con el muchacho este *testigo* a una/ cueba, donde halló muchos ydolos y máscaras con sangre, puestos en sus/ caxas y que tuxo de los dichos ydolos tres y algunas máscaras, y/ mucha ropa, y otras cosas de sacrificios (*PI*, Caso 1, f. 3 v.)
  - b. tengo mi *cassa* muy honrrada y proveyda de lo neçessario; y soy *cassado* con muger /<sup>5</sup> honrrada y virtuosa (*DLNE*, 75)
  - c. E tenido neçesidad *para* mi despacho de vender el juro, i e andado rodeando, por no lo vender, *porque* me era afruenta. I el padre Villagrà me a dado aquí mil i D o mil i DC *ducados*, i déxole el juro traspasado, *porque* goce de los çinquenta mil (*DO*, Carta 9)
  - d. con fuerça y biolençia havia fecho a Magda /<sup>23</sup>lena Mexia, mi hija, donzella honesta y recoxida, unas vezes por ruegos y prome /<sup>24</sup>zas y otras con amenazas y terrores, la avia corrompido y estupado, llevandole /<sup>25</sup> su virginidad (*DLNE*, 75, 1583)

En el ejemplo (17a) tenemos la palabra *máscaras*, con la grafía *s* prototípica para el sonido alveolar fricativo sordo [s] y también tenemos la misma palabra graficada con *x*, *máxcaras*. En el ejemplo (17b) tenemos las palabras *cassa*, *neçessidad* y *cassado*, escritas con la doble grafía; en el corpus se observó que la doble grafía apareció mayoritariamente en verbos conjugados en pretérito de subjuntivo con terminación en

<sup>10</sup> Con excepción de los casos derivados del verbo *hacer* > FACERE que se proporcionarán en la lista de [h].

-sse, como *mandasse, viesse, fuesse*, etc. y también en algunos superlativos terminados en *-issimo* como *christianissimo* y *hondissimo*.

El ejemplo (17c) nos ofrece la grafía *ç* y *c* seguidas de vocales anteriores, en *neçesidad, goce* y *çinquenta*, que estaban en avanzado proceso de confusión con el sonido alveolar fricativo sordo. Y, finalmente, el ejemplo (17d) muestra la grafía *z* en vocablos como *donzella, vezes, promezas* y *amenazas* para producir el mismo sonido. Podemos ver en la palabra *promezas* que la grafía *z* corresponde al sonido alveolar fricativo sordo [s] pero ya había confusión en las grafías.

Como vimos en el capítulo anterior, en §3.1., uno de los puntos importantes en el paso del castellano medieval al español moderno fue la pérdida de sonoridad en las sibilantes, y el ajuste en la zona y/o modo de articulación de algunas de éstas. Por la pérdida de sonoridad, el hablante no pudo establecer forma de diferenciar las grafías; por los ajustes de zona y modo de articulación, el hablante castellano, conservó diferencias: adelantó el sonido prepalatal africado sordo [ʃ̄] hasta realizar un sonido interdental fricativo sordo [θ], conservó su sonido alveolar fricativo sordo [s] y velarizó el prepalatal fricativo sordo [ʃ̄] → [x]. El hablante andaluz, canario y americano, (que constituyeron el dialecto que hoy se llama español atlántico), confundió el prepalatal africado sordo [ʃ̄] con el alveolar [s] y realizó la velarización del prepalatal fricativo de la misma manera que el hablante castellano. Con este proceso el hablante andaluz, canario y americano asignó la grafía *j* para representar el sonido velar fricativo sonoro [x] y se quedó con múltiples grafías (*s, ss, c, ç, z*) para representar solamente el sonido alveolar fricativo sordo [s]. Podemos intuir el problema ortográfico que se originó, pues se conservaron muchas grafías para un único sonido, y la pronunciación, que hasta ese entonces había otorgado pistas para hacer la debida distinción ortográfica, ya no

funcionó para ese propósito. El resultado, para algunos hablantes, es la total confusión de las grafías en la escritura; para los hablantes y amanuenses cultos, la distinción ortográfica provenía del conocimiento del origen etimológico de la palabra, y para aquellos no tan instruidos ocurría, o bien, la indistinción, o bien, las hipercorrecciones.

Todo este problema de la confusión de las grafías se originó en América por el español trasplantado de Andalucía como vimos en §3.2. El fenómeno en cuestión se conoce como seseo y se refiere al hecho de que como no llegó a América el sonido interdental [θ] proveniente del prepalatal africado sordo [ʃ̄], este último se confundió con el alveolar [s].

Al respecto Antonio Alatorre nos dice: “la variedad del español implantada en las Canarias y en el Nuevo Mundo fue la andaluza. En esos años iniciales del siglo XVI se desata, incontenible, el torrente de *faltas de ortografía*” (1979/2003:309).

Algunos autores rechazan el andalucismo americano pues proponen que en Andalucía existía seseo (confusión con /s/) y ceceo (confusión con /θ/), Alvar nos explica que:

Sin acabarse de consolidar el timbre seseante o ceceante del ceceo, el fenómeno pasó a Canarias y a América, donde se realizó como seseo. Nivelación que debe explicarse de algún modo que no sea el «autóctono». Porque si los andaluces llevaron a América un proceso en trance de realización, aunque no estabilizado, no cabe duda que el seseo no es autóctono (1982:138).

Con base en todo lo anterior y en el análisis del corpus, determiné las siguientes reglas de correlación entre el sonido alveolar fricativo sordo [s] y sus múltiples grafías.

### **Grafía s = [s]**

La s gráfica siempre tendrá una correspondencia con el sonido [s]:

**Contexto**  
Una s gráfica

**Regla**  
s= [s]

### **Grafía *ss* = [s]**

La *ss* gráfica siempre tendrá una correspondencia con el sonido [s]:

#### **Contexto**

Una *ss* gráfica

#### **Regla**

*s* = [s]

### **Grafía *c* = [s]**

La *c* gráfica tendrá una correspondencia con el sonido [s] en los siguientes contextos:

#### **Contexto**

Una *c* gráfica seguida de vocal alta anterior [i] (*i, j*)

Una *c* gráfica seguida de vocal media anterior [e]

#### **Regla**

*c* = [s]

*c* = [s]

### **Grafía *z* = [s]**

La *z* gráfica siempre tendrá una correspondencia con el fonema [s]:

#### **Contexto**

Una *z* gráfica

#### **Regla**

*z* = [s]

### **Grafía *ç* = [s]**

La *ç* gráfica, como se vio más arriba en §4.3.3., tendrá una correspondencia con el sonido [s] con excepción de los casos de utilización en léxico indígena, pues en estos su valor será de [ʃ]:

#### **Contexto**

Una *ç* gráfica

#### **Regla**

*ç* = [s]

También la grafía *x* se utilizó en el corpus para representar el sonido alveolar fricativo sordo [s]. Como esta grafía se utiliza prototípicamente para representar el sonido velar fricativo sordo [x] o el prepalatal fricativo sordo [ʃ], su aparición para representar al alveolar fricativo sordo [s] fue asistemática, de manera que se incorporará al transcriptor por medio de una lista de excepciones:

### Lista de palabras con *x* cuyo valor es [s]

|                       |           |
|-----------------------|-----------|
| caxco                 | máxcaras  |
| enxerta <sup>11</sup> | quixieren |

Para representar el sonido prepalatal fricativo sordo [š] se utilizó en el corpus la grafía *x*. Sin embargo, y como vimos en el apartado §4.3.4. del capítulo anterior, esta grafía también puede representar el sonido velar fricativo sordo [x] y la secuencia [ks]; de tal manera que su aparición con valor de prepalatal fricativo sordo [š] está sujeta a dos condiciones: que la palabra que contenga la grafía sea de origen indígena y que se ciña a las reglas propuestas en esta sección.<sup>12</sup>

Primero, veamos los ejemplos que ilustran el uso de esta grafía:

- (18) a. De /<sup>19</sup> esta gran çiudad de Tenuxtitan, Mexico, a xx /<sup>20</sup> de junjo de MDxxxij años (DLNE, 12)
- b. juntaronse /<sup>8</sup> unos tres prinçipales jndios con el Alonso Ortiz, que se /<sup>9</sup> dizen don Joan Xancol y don Hiermo Hastanhua y /<sup>10</sup> <y> don Po. Cacamatl (DLNE, 27, 1555)

Vemos en el ejemplo (18a) la palabra *Tenuxtitan* cuya grafía *x* corresponde al sonido prepalatal fricativo sordo [š], en seguida, también en (18a) vemos la palabra *Mexico* también de origen indígena con esta misma pronunciación. En el ejemplo (18b) vemos un antropónimo de origen indígena que, esta vez, corresponde a un nombre o apellido, *Xancol*. La única regla que se pudo obtener fue para palabras de origen indígena que tengan el siguiente contexto:

#### Grafía *x* = [š]

La *x* gráfica tendrá una correspondencia con el sonido [š]:

---

<sup>11</sup> En el contexto se puede ver que se refiere a *inserta* (DO, Carta 3).

<sup>12</sup> En las palabras de origen latino que muestran la grafía *x* puede observarse claramente que el sonido que representan es el alveolar fricativo sordo [s]: *caxco*, *enxerta*, *máxcaras*, *quixieren*.

**Contexto**

Si está ubicada entre vocal y consonante, en léxico indígena

Si está ubicada entre vocales, en léxico indígena

**Regla**

VxC= [ʃ]

VxV= [ʃ]

Partiendo de la opinión de García de Diego (1970:37) de que “en el siglo XVI se retrae la articulación de *x*, y en vez de la prepalatal de *dixo* y *quixada* se pronunció la actual *j* faucal, *dijo* y *quijada*, y a principios del XVII era ya general esta pronunciación para la de *hoja* y *jarro*”; se constata, como veremos más adelante, en las reglas de correspondencia entre [x] y *x, j*, que para este periodo las dos grafías alternaban en los mismos contextos, de manera que se infiere que la velarización estaba en un proceso avanzado.

Observé en el análisis del corpus que alternaban las grafías *x* y *j* en los mismos contextos (inicial de palabra y seguido por vocal), por lo que decidí que, para los fines del transcriptor, la grafía *x* seguida de vocal corresponderá al sonido prepalatal [ʃ] si se encuentra en la siguiente lista de palabras de origen indígena que claramente conservan el sonido:

**Lista de palabras con *x* cuyo valor es [ʃ]<sup>13</sup>**

Xancol  
Xicalango

El sonido velar fricativo sordo [x] se presenta en el corpus de trabajo graficado de tres formas diferentes: con *j*, con *g* seguida de vocales anteriores [e], [i] y con *x*. Ahora veremos claramente, por medio de las reglas arrojadas en este análisis, que las grafías *j* y *x* alternan en los mismos contextos, por lo que se puede ver que la

---

<sup>13</sup> Aunque se podría pensar que la grafía correspondiera a [x] en algunos contextos comunicativos. Se incorporarán a esta lista todos los topónimos y antropónimos de origen indígena que presenten esta grafía que puedan aparecer en documentos que se incorporen al corpus.

velarización mencionada en el §3.1. del capítulo anterior se encontraba en una fase bastante avanzada. Los ejemplos son los siguientes:

- (19)
- a. este *testigo* /<sup>30</sup> oja cómo ella le dezia: "¡puto, dexame!. Hazlo tú con tus braços, y bordonea con tus braços, /<sup>2</sup> que vengo harta de travaxar" (DLNE, 55, 1576)
  - b. dicen *que vuestro* marido es amigo de traer galas y de trabajar poco. /<sup>26</sup> Porque en esta tierra no ganan dineros sino quien lo trabaja /<sup>27</sup> muy trabajado (DLNE, 43, 1574)
  - c. Preguntado si *había* oydo *que* hera pecado grabísyimo y heregía casar/se doss bezes estando biba la primera muger (PI, caso 5)

En los ejemplos (19a) y (19b) se puede ver la misma palabra (el mismo contexto fónico) con diferentes grafías *travaxar* y *trabajar*, y con esto concluyo que ambas grafías *x*, *j*, correspondían al mismo sonido [x]. El ejemplo (19c) muestra la grafía *g* seguida de vocal anterior en *muger* y *heregía*, en las que el valor de la grafía *g* corresponde al sonido velar fricativo sordo [x]. El ejemplo de *muger*, (que ya no se presenta en el corpus como *muxer*), nos indica que ya había ocurrido la velarización del sonido [ʃ] y que sonaba como [x], pero el hablante no sabía de qué manera graficar este sonido, de manera que tomó la grafía velar por excelencia, es decir, una *g*.

En un estudio realizado por González (2005), se explica el proceso de ensordecimiento y posterior velarización del fonema prepalatal fricativo sonoro /ʒ/ → /ʃ/ → /x/ y la consecuente confusión gráfica:

Hasta el siglo XVI el idioma castellano constaba con dos fonemas, opuestos proporcionalmente: /ʒ/ dorsopalatal fricativo sonoro que [se] representaba ortográficamente *j*, *g*(*ge,gi*), y /ʃ/ dorsopalatal fricativo sordo, indicado en la escritura mediante la *x*. Sin embargo, en la evolución del idioma, al llegar a desaparecer la oposición que las diferenciaba, a favor del fonema sordo, y con el desplazamiento del punto de articulación a la zona velar, se creó vacilación ortográfica al dar un sonido único que podía ser representado por tres grafías “*j*”, “*g*” y “*x*” (González 2005:74).

Con base en lo anterior y en el análisis del corpus se determinaron las siguientes reglas de correlación entre las tres grafías y el sonido velar fricativo sordo [x].

### **Grafía *j* = [x]**

La *j* gráfica tendrá una correspondencia con el sonido [x]:

#### **Contexto**

Si está entre vocales  
Si es inicial de palabra y le sigue una vocal

#### **Regla**

V*j*V = [x]

#*j*V = [x]

### **Grafía *x* = [x]**

La *x* gráfica tendrá una correspondencia con el sonido [x] en los mismos contextos que una *j* gráfica:

#### **Contexto**

Si está entre vocales<sup>14</sup>  
Si es inicial de palabra y le sigue una vocal<sup>15</sup>

#### **Regla**

V*x*V = [x]

#*x*V = [x]

### **Grafía *g* = [x]**

La *g* gráfica tendrá una correspondencia con el fonema [x]:

#### **Contexto**

Si está seguida de vocales anteriores [e], [i]<sup>16</sup>

#### **Regla**

ge, ie = [x]

La siguiente es la lista de palabras con grafía *j* que deberán transcribirse con [x] aunque no recaigan en ninguna de las reglas anteriores:

#### **Lista de palabras con *j* cuyo valor es [x]**

injuria      perjudiciales  
Jorje      subjettos/subjetas/subjetos

Podemos observar en la anterior lista de palabras que *subjettos/subjetas/subjetos* conservan un grupo culto de consonantes *bj* que se ha perdido en el español moderno, una vez que esta *b* se perdió, el contexto fónico cambiaría, la *j* estaría entre vocales y

---

<sup>14</sup> Con excepción de las palabras que llevan esta grafía *x* que están en el mismo contexto fónico en las listas de excepciones correspondientes al sonido [s] y al conjunto [ks].

<sup>15</sup> También con excepción de las palabras con esta grafía *x* en la lista de palabras correspondientes al sonido [ʃ].

<sup>16</sup> Excepto en los casos que llevan esta grafía *g* proporcionados en la lista de excepciones correspondientes al fonema [g].



*de Ordaz*, lo cual nos da una pista de qué tan avanzado se encontraba el proceso de pérdida de dicho sonido. Este hecho también nos corrobora, como se explicó en §3.1. del capítulo anterior, que la pérdida del sonido avanzó de norte a sur en la Península Ibérica en época temprana, bien como sustrato vasco (Menéndez Pidal 1986, Pensado 1993), bien como pérdida temprana de la dentadura por la carencia de flúor en el agua de la zona como lo menciona Salvador (1982). Ya que, como vimos, Diego de Ordaz era originario del norte de la Península Ibérica.

Nótese en el ejemplo (20c) que las palabras que hoy día se escriben con *h*- inicial como *han* en las cuales el sonido [h] no existía, se escribían sin *h* pues en ese contexto no había un sonido. En cambio, en el caso de las palabras como *hacer*, con origen en palabras con *f*- inicial latinas, el sonido *f* sufrió un proceso de debilitamiento y, después, de pérdida. En ese momento, estaba en una fase de aspiración avanzada, tal vez en una etapa próxima de desaparición y por ello empezaba a usarse *h* para indicar ausencia de sonido.

La regla asignada para este sonido será, para la *f*- inicial de palabra, en vocablos provenientes del verbo *hacer*. Se incorporará al transcriptor por medio de la siguiente lista de excepciones:

**Lista de palabras con *f*- cuyo valor es [h]**

|             |              |
|-------------|--------------|
| faga        | fecho/fechos |
| faser/fazer | féchose      |
| faze        |              |

Para la grafía *h* la regla es hacerla una aspiración correspondiente al sonido glotal fricativo sordo [h] en posición inicial de palabra en vocablos provenientes de léxico con *f*- inicial latina:

## Grafía *h* = [h]

La *h* gráfica tendrá una correspondencia con el sonido /h/:

### Contexto

Si es inicial de palabra y proviene de una palabra latina con *f*- inicial

### Regla

#*h*= [h]

Explica García de Diego que “la *h* aspirada procedente de *f* se conserva hasta el siglo XVI” (1970:37). Con excepción de: *h*- inicial proveniente de *f*- inicial latina que corresponde al sonido fricativo glotal sordo [h], el grupo *ph* correspondiente al sonido labiodental fricativo sordo [f] y el grupo *ch* correspondiente al sonido palatal africado sordo [ç] o al sonido velar oclusivo sordo [k], todas las *h* que se encuentren en el corpus se transcribirán como cero fonético Ø.

### 4.3.5. Consonante fricativa sonora

Para graficar la consonante palatal fricativa sonora [y] se encontraron las grafías *y* y *ll*. Debemos recordar, como se mencionó en el §3.1. del capítulo anterior, que el yeísmo (pronunciación de *ll* como [y]) se ha documentado en América desde principios del siglo XVI, por lo que todas las grafías *ll* que se encuentren en el corpus se transcribirán como [y]. Veamos los ejemplos:

- (21)
- a. I hazed de manera *que*, con lo *que* arriba digo i de la casa i de esto i de lo demás *que* pudierdes, enbiéis todo lo más *que* podierdes, no os deshaziendo de las yeguas ni ovejas, sino lo *que* os pareçiere *que* conviene (*DO*, Carta 3)
  - b. dixo este testigo que sabe *que*/ el dicho don Diego tiene quatro mujeres, con las quales ha/ze vida maridable, y las dos *dellas* son hermanas, y de una/ tierra que se llama Tornacuxtle (*PI*, caso 6)

Vemos en el ejemplo (21a) una palabra graficada con *y*, *yeguas* y en el ejemplo (21b) las palabras *dellas* (ejemplo de sandhi)<sup>17</sup> y *llama* en las que se infiere, con base en

---

<sup>17</sup> Todo cambio fonético que se opera al principio o fin de una palabra, o, menos frecuentemente, en su interior, por influjo de los sonidos de la palabra vecina, en la frase (Lázaro Carreter 1953/1977).

los estudios que se explicaron en la parte correspondiente del capítulo anterior, que la pronunciación corresponde al sonido palatal fricativo sonoro [y].

García de Diego (1970) explica qué es el yeísmo y de qué manera surge a causa de la relajación de la pronunciación del sonido palatal lateral [ʎ]:

De la relajación articulatoria de la *ll* resulta el yeísmo. (...) Suele ser considerado el yeísmo como una absoluta confusión general de *ll* hecha *y* en que se identifican *poyo* con *pollo*, *halla* con *haya*. Se citan como regiones típicas de yeísmo total, Extremadura, Andalucía, Canarias y América, aunque la distinción de *ll*, *y* se conserva por las personas cultas y se defiende en América en Colombia, Ecuador, Perú, Chile y parte de la Argentina (1970:38).

Sin embargo, cuando menciona que las personas cultas conservan la distinción de sonidos y que en algunas zonas de América del sur se hace la distinción entre las dos grafías, obliga a un juicio erróneo ya que, afirma Parodi (1977:247), “la confusión de *ll* y *y* no aparece exclusivamente en textos redactados por personas de escasa cultura, pues se documenta en escritos de individuos cuyo oficio requiere un grado medio de cultura”.

Es posible que la razón por la que se distinguieran las dos grafías en la escritura obedecía, en las personas cultas, más a distinción etimológica que a pronunciación, de manera que un hablante culto en el siglo XVI que hiciera la distinción ortográfica correctamente por el conocimiento de la raíz de las palabras y no porque aún distinguiera los dos sonidos.

Aunque en el corpus de trabajo no se encontraron casos de confusión de *y* y *ll*, es decir, que no alternaban en los mismos contextos, se determinó, para fines prácticos del transcriptor, a partir de lo dicho anteriormente y de las diferentes teorías acerca de este fenómeno expuestas en el capítulo anterior, que las grafías *y* y *ll* se transcribirán como [y].

### **Grafía y = [y]**

La y gráfica siempre tendrá una correspondencia con el sonido [y]:

#### **Contexto**

Si es inicial de palabra y a continuación hay una vocal  
Si está entre dos vocales<sup>18</sup>

#### **Regla**

#yV= [y]  
VyV= [y]

### **Dígrafo ll = [y]**

La ll gráfica tendrá una correspondencia con el sonido [y]:

#### **Contexto**

El dígrafo ll siempre tendrá correspondencia con [y]<sup>19</sup>

#### **Regla**

ll= [y]

Se encontró un solo caso en el corpus en el que este sonido [y] se graficó con *j*, en la palabra *majores* que deberá transcribirse [ma·yó-res]. Para casos futuros de palabras que contengan este sonido [y] y esté graficado con *j* tendrán que incorporarse a la lista de filtrado y transcribirse conforme corresponda.

### **4.3.6. Consonantes nasales**

Las consonantes nasales se han mostrado, a lo largo de la historia del español, bastante estables en cuanto a su representación gráfica; el sonido bilabial nasal [m] y el alveolar nasal [n] se conservan desde el latín, mientras que el palatal nasal [ɲ] se originó en el periodo de nacimiento del castellano medieval (en el que, como se explicó en el §3.1. del capítulo anterior, hubo una proliferación en el surgimiento de sonidos palatales) por: la nasal alveolar geminada del latín *nn*, el grupo consonántico *gn, ng*, o la convivencia de nasal alveolar con yod *nj, ñ*.

---

<sup>18</sup> Se encontró en el corpus un único caso en el que la grafía *y* va doble, en *oyye*, y su correspondencia será un solo sonido [y], es decir [ó·ye].

<sup>19</sup> Excepto en las listas de excepciones en el apartado del sonido alveolar lateral [l], correspondientes a [l] y [rl].

Para los sonidos bilabial nasal [m] y alveolar nasal [n] se determinaron las siguientes reglas.

### **Grafía *m* = [m]**

La *m* gráfica siempre tendrá una correspondencia con el sonido [m]:

#### **Contexto**

Aparición de *m* gráfica

#### **Regla**

*m* = [m]

### **Grafía *n* = [n]**

La *n* gráfica siempre tendrá una correspondencia con el sonido [n]:

#### **Contexto**

Aparición de *n* gráfica

#### **Regla**

*n* = [n]

Estos son algunos ejemplos de estas grafías en el corpus:

- (22) a. Lo que el dicho Diego Díaz cometió grabe delito de ydola/tría, y lo regía en hazer el dicho ydolo con las dichas çirimonias,/ y querer hazer ydolatrar a los dichos yndios, y persuadir a los/ españoles que haziesen lo **mismo** (*PI*, caso 13)
- b. Mi marido es **onbre** muy delicado y muj **enfermo**, y teme mu /<sup>9</sup>cho la mar. Y asi, señor **ermano**, yo tengo determinado, junta /<sup>10</sup>mente con el señor mi marido y la señora mi tia, *que* vos os /<sup>11</sup> **vengays** a esta tierra (*DLNE*, 69, 1578)

Como vemos en los ejemplos anteriores, los dos sonidos [m] y [n] no ofrecen mayor dificultad para su transcripción automática. En el ejemplo (22b) vemos una palabra, *onbre*, que está graficada con una nasal alveolar [n] en lugar de una nasal bilabial [m]. Este grupo consonántico se respetará en el transcriptor.

Para el sonido palatal nasal [ɲ] se encontraron dos graficaciones distintas en el corpus: la prototípica *ñ* y el dígrafo *gñ* (sólo se encuentra en las *Cartas de Diego de Ordaz*).

Observemos algunos ejemplos:

- (23) a. Piadosamente se cree *que*, antes *que* se embarque, verná de la Nueva España algún navío *que* le hará detener las *merçedes* que su *Majestad* le a hecho (DO, Carta 2)  
b. las mu /<sup>33</sup>geres y niños son los *que* ay para ser enseñados y doctri /<sup>34</sup>nados en las cosas de Dios (DLNE, 15, 1533)

En los dos ejemplos de (23) se observan las diferentes graficaciones para el sonido palatal nasal [ɲ] y podemos ver que no existe dificultad alguna para asignar la regla de correlación, por lo que se determinó la siguiente.

### Grafía ñ = [ɲ]

La ñ gráfica siempre tendrá una correspondencia con el sonido [ɲ]:

#### Contexto

Aparición de ñ gráfica

#### Regla

ñ = [ɲ]

Las palabras graficadas con ñ para representar el sonido [ɲ] se incorporarán al transcriptor por medio de la siguiente lista de excepciones:

#### Lista de palabras con *gñ* cuyo valor es [ɲ]

|                                        |                   |
|----------------------------------------|-------------------|
| agño(s)                                | Maragñón          |
| castagño                               | nigños            |
| dagño                                  | pagño             |
| dogña                                  | pegñol/pegñón     |
| engagñar                               | segñalase/segñale |
| enpegño/enpegñar/enpegñado/desenpegñar | segñalado/segñaló |
| España/espagnol(es)                    | Villafagñe        |

### 4.3.7. Consonantes vibrantes

Las consonantes vibrantes también se muestran bastante estables en el corpus. El sonido vibrante simple [r] se grafica con *r* simple, el vibrante múltiple [r̄] se grafica con *r* simple y con la doble grafía *rr*. Veamos los ejemplos que nos ofrece el corpus:

- (24) a. por *qué* dixiste/ *que* yo *había* muerto a Francisco, el yndio, *dixeras/ que* lo *habías* muerto tú, o *dixeras que* lo *había/mos* muerto ambos, *ahora* me lo pagarás (PI, Caso de Diego Díaz, f. 44)
- b. Murieron *senos* tres *rreligiosos*, y *treynta* truxe a esta Nueva España. Y si truxera mill, fue /<sup>4</sup>ra *harto* menester para *rremediar* algo de lo mucho *que* está perdido (DLNE, 25, 1554)

En estos ejemplos de (24) vemos la graficación *r* para el sonido alveolar vibrante simple [r] y las dos graficaciones *r*, *rr* para el sonido alveolar vibrante múltiple [r̄].

Como estos sonidos no representan mayor complicación, determiné las siguientes reglas de correspondencia para transcribirse automáticamente.

### Grafía *r* = [r]

La *r* gráfica siempre tendrá una correspondencia con el sonido [r]:

#### Contexto

Aparición de *r* gráfica<sup>20</sup>  
 Si es final de palabra y va antecedida por vocal  
 Cuando está entre una vocal y una consonante

#### Regla

*r* = [r]  
 Vr# = [r]  
 VrC = [r]

### Grafía *r* = [r̄]

La *r* gráfica tendrá una correspondencia con el sonido [r̄] en los siguientes contextos:

#### Contexto

Cuando es inicial de palabra seguida de vocal  
 Cuando es antecedida por *n*  
 Cuando es antecedida por *l*

#### Regla

#rV = [r̄]  
 nr\_ = [r̄]  
 lr\_ = [r̄]

### Dígrafo *rr* = [r̄]

El dígrafo *rr* gráfico tendrá una correspondencia con el sonido [r̄]:

#### Contexto

Aparición de *rr* gráfica

#### Regla

*rr* = [r̄]

<sup>20</sup> Excepto cuando esta grafía es inicial de palabra, pues su correspondencia será [r̄].

#### 4.3.8. Consonante lateral

El sonido alveolar lateral [l] se encontró graficado en el corpus con la prototípica *l* y con la doble grafía *ll*, en algunas palabras. Esta doble grafía, sin embargo, ofrece dificultades porque en algunas ocasiones, como se vio en la parte correspondiente, obedece al sonido palatal central [ɣ] y otras deberá transcribirse como [rl] (en infinitivos + clítico *le*, *la(s)*, *lo(s)*, como *pedille* → [pe·dír·le]).

Veamos los siguientes ejemplos del corpus en los que se atestigua la dificultad que presentan algunas graficaciones de dicho sonido:

- (25)
- a. Después que *vuestra señoría* me dio su santa bendición en Sevilla, y nos hezimos a **la vela**, venimos *l<sup>2</sup>* con hartos trabajos, con los infortunios y tempestades del mar (*DLNE*, 25)
  - b. E después de lo susodicho, marte quinze días del mes/ de **jullio**, año susodicho de **mille** quinientos e treynta e/ nueve años, estando en *avdiencia del Santo Ofiçio*, su Señoría *Reverendísima*/ hizo paresçer ante sí al dicho son Carlos Chichimecatecotl (*PI*, caso 10)
  - c. Esto digo por la de Guaxoçingo, pero no serán tan locos *que* hagan otra cosa sino *conplilla*, porque la cosa *que* acá más a mal an tenido a sido no *conplir* lo *que* de acá va (*PI*, Caso de Diego Díaz, f. 44)

En el ejemplo (25a) vemos la grafía prototípica para el sonido alveolar lateral [l] en determinantes como *la* y *los*, en la contracción *del* y en el sustantivo *vela*. El ejemplo (25b) ofrece la doble grafía *ll* que corresponde a este sonido [l] en *jullio* y *mille*. Y, por último, el ejemplo (25c) muestra el caso de la doble grafía *ll* con valor del grupo consonántico [rl] en *conplilla* verbo en infinitivo *conplir* + el clítico *la*.

Con lo anterior se determinaron las siguientes reglas de correspondencia entre las grafías que se presentaron en el corpus de trabajo y el sonido alveolar lateral [l].

#### Grafía *l* = [l]

La *l* gráfica siempre tendrá una correspondencia con el sonido [l]:

#### Contexto

Aparición de *l* gráfica

#### Regla

*l* = [l]

La aparición asistemática de *ll* para el sonido alveolar lateral [l] generó la siguiente lista de excepciones:

**Lista de palabras con *ll* cuyo valor es [l]**

|              |                 |
|--------------|-----------------|
| alcalde      | jllustre        |
| humilldes    | jullio          |
| humillmente  | lljenciado      |
| inteligencia | mill/mille/mjll |

Para la doble grafía *ll* con valor de [rl] se llegó a la siguiente lista de excepciones:

**Lista de palabras con *ll* cuyo valor es [rl]<sup>21</sup>**

**con clítico lo(s)    con clítico la    con clítico le(s)**

|            |              |                   |
|------------|--------------|-------------------|
| avjsallo   | bella        | belles            |
| correxillo | conplilla    | castigalle        |
| enojallos  | descrevilla  | dalle             |
| negociallo | despachalla  | dezille           |
|            | gratificalla | hazelle / haselle |
|            | mostralla    | mudalle           |
|            | pagalla      | pedilles          |
|            | traella      | respondelle       |

**4.3.9. Grupo [ks]**

La grafía *x* también puede representar al grupo consonántico [ks]. Como ya vimos en el apartado correspondiente (§4.3.4.), esta grafía puede representar tanto [s], [š], como [x], de tal manera que su correspondencia con la secuencia consonántica [ks] se aplicará al transcriptor mediante la siguiente lista de excepciones:

**Lista de palabras con *x* cuyo valor es [ks]**

|                       |                       |               |
|-----------------------|-----------------------|---------------|
| <b>en VxC</b>         | <b>en VxV</b>         | <b>en Vx#</b> |
| exçeda/ exçediendo    | auxilio               | pax           |
| exçelente / excelente | examjnado             | rex           |
| excelentemente        | examjnar              |               |
| exçessivas/ excesivos | examinaron/examinando |               |
| excessjva             | exessivos             |               |
| expendera             | maxime                |               |

<sup>21</sup> Asigno el valor de [rl] a *ll* porque alterna la forma gráfica en los documentos entre *rl* y *ll*. No debemos descartar, sin embargo, que en algunos contextos se pronunciara [y] o [l], ya que así se atestigua en las rimas.

experiencia /experiência próximo /proximo /proxima  
experiencia / experjencia  
expressasse  
exsecutar  
extendido  
hexsaminados  
Mixteca / Tlaxcala  
Tornacuxtle / Texcuco

#### 4.4. Problemas específicos

Algunos de los problemas específicos que plantea el corpus son: el manejo de léxico indígena (sobre todo del náhuatl) y la consecuente permanencia temporal de sonidos que se estaban perdiendo en español, como lo mencioné en §3.1., y la presencia de grupos cultos de consonantes, unos que aún se conservaban como etimológicos, y otros en los que se daban trueques asistemáticos.

##### 4.4.1. Sonidos que permanecieron temporalmente por léxico indígena

Ya mencioné, en el capítulo anterior en §3.1.1., el reajuste de sibilantes que se dio como paso del sistema fonológico medieval al moderno en español; también mencioné en el mismo párrafo que los sonidos [ʃ] y [ʒ] que estaban en proceso de pérdida en el español se mantuvieron temporalmente en el español mexicano por su uso en léxico náhuatl. Al respecto, Lope Blanch menciona la influencia que tuvo el náhuatl sobre el español y corrobora esta idea:

Por lo que respecta al español normal, al habla común de la ciudad de México, he llegado a la conclusión de que las únicas peculiaridades que pueden atribuirse por ahora a la influencia del sustrato [en el campo fonológico] son las siguientes: existencia de un fonema /ʃ/ en voces de origen indígena (*xixi*), aunque de rendimiento fonológico mínimo, ya que normalmente actúa como alternante de /s/; aparición de un sonido [ʃ], en topónimos y antropónimos prehispánicos (*Atzompa*), que funciona como variante alofónica de /s/; articulación explosiva, licuante, de *t* seguida por *l* (*tl*), tanto en voces nahuas (*ix-tle*) como en palabras hispánicas (*a-tle-ta*) (1969b:12).

La idea que subyace en lo que dice Lope Blanch es que si un hablante español escuchaba alguna palabra indígena que tuviera un sonido parecido a [ʃ] o [ʃ̄], echaba mano de su propio sistema fonológico para reproducirla, sobre todo cuando tenían que utilizar léxico de nuevas cosas no existentes en su realidad.

Algunas veces, la solución de [ʃ̄] fue confundirse con [s]; otras su conservación mediante el sonido en proceso de pérdida [ʃ], aunque posiblemente no sería igual al fonema indígena en cuestión, pues “los españoles [...] sustituyeron los sonidos indígenas extraños por los más parecidos dentro de su sistema” (Arias 1997:42). Así, utilizaban la grafía *ç* que era la que reflejaba la correspondencia del sonido prepalatal africado sordo [ʃ̄] y que posteriormente, cuando el sonido se perdió, se desdobló en *tz*. Este sonido se usa en antropónimos y topónimos de origen náhuatl, lo vemos en el corpus en nombres de lugares como *Guaxoçingo*, *Coaçacoalco*, *Yçúcar* y en nombres de personas como *Tilançi*.

Su uso en español mexicano, nos explica Lope Blanch (1972), no se debe a conservación arcaizante del sonido del español antiguo [ʃ̄], sino a influencia de las lenguas indígenas. “En resumen, el sonido *ç* no tiene rendimiento fonológico, y se halla exclusivamente en algunos topónimos y antropónimos indígenas, por lo que hoy podría considerarse como un simple alófono de /s/; si embargo, no deja de ser un resto de la fonética propia de las lenguas prehispánicas” (Lope Blanch 1972:97).

Para el sonido prepalatal fricativo sordo [ʃ̄] son cuatro soluciones las que pueden verse frente a la evolución de este sonido: evolución normal a [x], es decir, velarización; confusión con [s], una despalatalización; conservación de [ʃ̄] y conversión en palatal

africada [ç]. Sin embargo, “los casos de conservación de /ʃ/ aparecen siempre en voces de origen indoamericano: *mixiote*, *xocoyote*, *axcan*” (Lope Blanch 1972:94).

En el corpus vemos este sonido [ʃ] graficado con *x* en topónimos y antropónimos como *Mexico*, *Guaxoçingo*, *Xancol*, *Xicalango*. Este sonido también se conservó en léxico indígena cotidiano como *mixiote* < METL + XIOTL, *jitomate* < XICTLI + TOMATL, *chocolate* < XOCOATL, *jacal* < XACALLI, que, por el carácter temático del corpus, no se encontró.

Debemos concluir que el español de México, en su sistema fonológico, tiene poca influencia de las lenguas indígenas, pues lo que toma de ellas, especialmente del náhuatl, es material léxico. En el plano fonológico lo que provoca es cierta permanencia de sonidos que ya existían en español y que estaban en proceso de pérdida. Su mantenimiento temporal fue motivado por lo que los españoles oían de la lengua náhuatl y así, lo adaptan a lo que ellos conocían, que eran los sonidos representados por *ç* y *x*.

#### **4.4.2. Grupos de consonantes**

Mencioné ya en el §4.3.8. que uno de los grupos consonánticos que se encontró en los *DLNE* fue el de *ll* con valor (posible, por la alternancia gráfica *rl* y *ll*) de [rl] en palabras compuestas de verbo en infinitivo seguido de clítico.

Otro grupo consonántico que se encontró en el corpus fue el formado por *ld* en palabras compuestas por imperativo (en segunda persona de plural, *vosotros*) con clítico de complemento indirecto o clítico de complemento directo. Adelante se presentarán ejemplos de este fenómeno. También se encontró el grupo *-nd* final de palabra. Se

encontraron, asimismo, trueques de líquidas y otros trueques asistemáticos que veremos a continuación.

Al juntarse el verbo en infinitivo con un clítico, ya sea de complemento directo o complemento indirecto, *corregir + la > corregirla*, los amanuenses del siglo XVI escribían en algunas ocasiones *rl* y en otras cambiaban el grupo *rl* por *ll* en la escritura. Esto nos indica que en algunos contextos o situaciones comunicativas se pronunciaba [rl] y en otros contextos [y] o [j]. Este grupo *ll* se cambiará a [rl] en el transcriptor automático, pues se encuentra en los *DLNE* y en *El habla de Diego de Ordaz*, pero no en los *Procesos Inquisitoriales*. En estos últimos se encuentra en su forma prototípica, como *rl*.

Al respecto dice Lope Blanch que “la asimilación de la *-r* en los infinitivos a la *l-* del pronombre átono siguiente es lo común en la fonética de Ordaz. Los casos de conservación de la *-r* son mucho menos frecuentes” (1998:49).

Se encontró una metátesis en el grupo consonántico *dl* en palabras compuestas por imperativo plural con clítico de complemento directo (*lo, la*) o complemento indirecto (*le, les*). Si debiera ser *guardad + la > guardadla*, se encontró *ld* por *dl*, *guardalda*. Este fenómeno ocurre con mayor frecuencia en *El habla de Diego de Ordaz*. Ninguna metátesis de este tipo se presentó en los *Procesos Inquisitoriales*; sólo cuatro ejemplos se encontraron en los *DLNE*: *enbialdo, decilde, encomendaldo, miraldo*. Y muchos en *El habla de Diego de Ordaz*: *dalde, guardalda, vendelda, mezclaldo, hazeldo, dezilde, meteldo, entregaldo, presentalda, daldes, mostraldes, meteldes, cobraldo, favoreçelde*, etc. En el transcriptor se convertirá a [dl]; sin embargo, no hay que descartar la posibilidad de que se pronunciara también [ld] en algunos contextos. Lope Blanch apunta que “como en el caso de la asimilación de *rl* en *ll*, esta metátesis [ld] se mantuvo con firmeza hasta bien entrado el siglo XVII” (2000:212).

Un trueque de líquidas bastante curioso es el presentado en *flaires* en el que la *l* tomó el lugar de la *r* y viceversa. Como es un trueque que no se presentó en el corpus más que en este caso, la palabra se transcribirá como [frái·les]. Sin embargo, García de Diego (1970) indica que la conversión de *r* en *l* y viceversa es frecuente en el periodo estudiado.

Encontré en el corpus de trabajo aféresis de *e* en palabras con *s-* inicial como cultismos del latín, pues en esa lengua iniciaban con *s líquida* (García de Diego 1970), los ejemplos son pocos, *screvi*, *scrivjo*. En el transcriptor se repondrán las *e* iniciales mediante la siguiente lista de excepciones, ya que la secuencia de tres consonantes juntas no es pauta para el español:

**Lista de palabras con *s-* inicial que se transcribirá como [es]**

|        |         |
|--------|---------|
| screvi | scrivió |
|--------|---------|

Igualmente se encontraron en el corpus paragoges con grupo *-nd* final en palabras como *segund*, *grand*, *ningund*, en las que se quitará la *-d* final para efectos prácticos del transcriptor (pero no debe desecharse la posibilidad de que esta *-d* final se pronunciara), pues ya eran arcaísmos como lo menciona Rafael Lapesa: “el idioma continúa despojándose del lastre medieval. Desaparece la alternancia gráfica de *t*, *d* finales, y apenas se ven sino formas con *d*, *antigüedad*, *voluntad*, *merced*” (1981:280). Lo cual indica que las palabras graficadas con *-t* final ya no se presentan sino con *-d* final y que esta misma *-d* ya era arcaica. También estos casos se incorporarán al transcriptor mediante la siguiente lista de excepciones:

**Lista de palabras con grupo *-nd* final que se transcribirá como [n]**

|         |        |
|---------|--------|
| grand   | segund |
| ningund |        |

En este capítulo se abordó el análisis del corpus de trabajo, confrontándolo con las opiniones y los estudios emprendidos por los teóricos y finalmente se estableció, para cada sonido, la regla correspondiente según sus variaciones gráficas. Así como se asignaron las reglas de correspondencia, se presentó, en aquellos sonidos que lo requirieran, la lista de excepciones correspondiente. Habrá que mencionar que la cantidad de las palabras que finalmente contendrá la lista de filtrado es mínima en comparación con la extensión total del corpus, por lo que las reglas obtenidas en el análisis fueron abarcadoras para el léxico del español en México en el siglo XVI. Como se ha mencionado anteriormente, la lista de filtrado final se encontrará en el Apéndice 3 de esta tesis.

## 5. CONCLUSIONES

Hemos visto, en el Capítulo 2 de esta tesis, las ventajas que ofrece trabajar con los corpus electrónicos anotados y cómo funciona la interfaz del CHEM. Hemos visto los trabajos en proceso para enriquecer este corpus y los trabajos que necesitarán incorporarse a futuro para mejorar el funcionamiento de sus herramientas.

Por el lado de la investigación filológica emprendida para determinar el estatus del sistema fonológico del español hablado en México en el siglo XVI, vimos, en el Capítulo 3, que estaban en proceso varios cambios que determinaron el paso del sistema fonológico del castellano medieval al español moderno. Algunos de estos cambios empezaron a gestarse desde el siglo XIV, por lo que para el siglo XVI se encontraban en una fase avanzada de cambio. Cuando ocurrió la llegada al Nuevo Mundo llegaron también, junto con los pobladores españoles, los hábitos de pronunciación del español.

En América, algunos de estos hábitos encontraron un ambiente propicio para finalizar su proceso (como el seseo andaluz) y otros lo contuvieron (como el proceso de pérdida de los sonidos prepalatales africado y fricativo que se mantuvieron temporalmente por influencia de los sonidos próximos del léxico indígena).

Vimos con suficientes datos que el español trasplantado a América, si bien incluía dialectos de toda la Península Ibérica, tenía mayoría de hablantes de origen andaluz, específicamente de Sevilla. Y por ello, se ha concluido (cf. Boyd-Bowman 1985) que el español trasplantado a América tiene origen andalucista o, mejor dicho, sevillano.

Con base en el análisis hecho para determinar las reglas de correlación entre sonido y grafía, y con base en investigaciones teóricas por parte de reconocidos

estudiosos, el estatus de cada sonido en el español que se hablaba posiblemente en México en el siglo XVI es el siguiente:

- La vocales medias (anterior y posterior) y la central baja se presentaron invariables, las altas presentaron diferentes graficaciones, para la anterior se utilizaron grafías como *i, j* y *y*, para la posterior *u, b* y *v*.
- La vacilación de timbre en las vocales átonas era una práctica común en el español del siglo XVI (Alatorre 1979/2003), por lo que se respetará la vocal graficada, atendiendo a las reglas de correspondencia establecidas.
- Las oclusivas sordas eran bastante estables, la única que presentó vacilaciones gráficas fue la velar [k], con las grafías *c, qu, ch* y *x*. La ortografía para este fonema no se normó sino hasta el siglo XVIII, con las reglas ortográficas de la Real Academia Española. Aún hoy representa problemas gráficos, pues se sigue utilizando *c* y *qu*, además de la recién usada en español *k*.
- La misma situación prevalece en las oclusivas sonoras. La bilabial [b] presentó vacilaciones gráficas con *b, v* y *u*. Sin embargo, se argumentó que el fonema representado con *v* no era labiodental, sino bilabial fricativo (Alatorre 1979/2003, Lathrop 1984/1992). Finalmente, se asignará el sonido bilabial oclusivo sonoro [b] para estas tres grafías (siempre que *u* tenga valor consonántico).
- La africada palatal [ç] se comportó invariable desde que surgió en castellano y su grafía siempre fue el dígrafo *ch*.
- Por otra parte, la prepalatal africada [ʃ] que estaba en proceso de confusión, en el español atlántico, con la alveolar fricativa sorda [s], frenó este proceso de pérdida al entrar en contacto en México con el náhuatl. Hecho que

mantuvo su uso en léxico náhuatl temporalmente hasta que estas palabras se adaptaron a la pauta del español, y, en palabras hispanas, siguió su camino natural: la desfonologización.

- La correlación de las fricativas sordas resultó complicada pues todos los sonidos presentaron diversas graficaciones. Veamos.
  - o El sonido [f] en inicial de palabra cambió desde muy temprano a [h] y posteriormente a Ø. Sin embargo en el siglo XVI aún se conservaba la grafía *f*- inicial de palabra aunque el hablante no la pronunciara o la aspirara, como una tendencia cultista. De igual manera, la *h*-graficada correspondía, ya a una aspiración [h], ya a un cero fonético Ø. Concluí, con base en el análisis del corpus, que las *f*- inicial de palabra latina que hayan cambiado esta grafía por *h*- en español actual se transcribirían como [h], pues el sonido derivado de *f*- inicial latina estaba en un claro proceso de aspiración; mientras las que se encuentren en otros contextos se transcribirán como [f]. Que las *h* se transcribirán como Ø, excepto en las combinaciones *ph*, *ch* y *chr*.
  - o El sonido [s] existía desde el latín, pero en el siglo XVI algunos otros sonidos se encontraban en una etapa avanzada de confusión con este. Que el hablante produjera grafías como *c*, *z*, *ç*, *ss* y *s* como [s] constituyó, en el español atlántico, el fenómeno conocido como seseo. Este fenómeno complica la asignación de sonidos cuando algunas de estas mismas grafías representan otros fonos (como *ç* representa [ʃ]). Y confunde al amanuense en el momento de redactar textos, por lo que se producen *faltas de ortografía*.

- El sonido [ʃ] graficado con *x* también se mantuvo temporalmente cuando la población española entró en contacto con la población indígena mexicana, después, las palabras de origen náhuatl que contenían este sonido y que se incorporaron al léxico general del español, se adaptaron a la pauta de esta lengua y entraron en el mismo proceso de velarización. Decidí, para los fines prácticos del transcriptor, dejar este sonido vigente únicamente en léxico indígena.
- Con relación al punto anterior, la grafía *x* también representaba el sonido velar fricativo sordo [x], pues la grafía alternaba en los mismos contextos con *j*. De manera que, en el transcriptor, *x* y *j* se transcribirán como [x], excepto cuando *x* se encuentre en léxico indígena.
- El sonido [y] era el que predominaba en la pronunciación americana. Según los teóricos, desde muy temprano se documentan casos de yeísmo en América, de manera que asigné el sonido [y] a graficaciones como *y* y *ll*. Sin embargo, nuestro corpus, quizá por la temática, no muestra yeísmo, es decir, no hay confusiones entre una y otra grafía, ni alternan en los mismos contextos. No puedo afirmar, consiguientemente, que no existiera el sonido [ʝ] en el español americano.
- Las nasales se muestran invariables, salvo la palatal, que muestra, aparte de la prototípica grafía *ñ*, el dígrafo con tendencia cultista *gñ* en las *Cartas de Diego de Ordaz*.
- Las líquidas no muestran casos de vacilación gráfica, salvo los casos especiales por error como la doble grafía inicial *rr* para representar el sonido

vibrante múltiple [r̄], o el caso de la duplicación de // por respeto a la ortografía latina de consonantes geminadas.

Hemos visto en estas conclusiones que las complicaciones en la representación gráfica se encuentran en las palatales. Es decir, tanto las vocales palatales, como las consonantes palatales que surgieron en la transición del latín al castellano medieval son las que presentan vacilaciones gráficas. Esta confusión se da, por el reajuste, en el paso del castellano medieval al español moderno, de sonidos palatales que surgieron en castellano y que eran muy próximos en su pronunciación.

Al realizarse ese reajuste en los sonidos palatales; al perderse unos y conservarse otros, se reduce el inventario de sonidos por contraste con el mantenimiento de formas gráficas desempleadas y por ello se cae en vacilaciones ortográficas.

Todo esto se refleja en opiniones de gramáticos, como Nebrija, que recomendaban que se escribiera tal y como se hablaba y que, por otro lado, trataban de mantener la graficación de las palabras atendiendo al origen etimológico de las mismas.

Para los fines prácticos del transcriptor automático que se regirá, para el siglo XVI, con mis reglas, traté de hacer las mayores simplificaciones posibles entre las opiniones de los teóricos y los resultados de mi análisis. Combinar estos dos puntos, no es fácil, por lo que no dudo que este trabajo tenga inexactitudes que podrán, siempre, superarse si se emprenden nuevas investigaciones en el futuro.

Se pudo observar, con base en el contraste de la cantidad de palabras contenidas en la lista de filtrado con la extensión del corpus de trabajo, que los resultados de la investigación fueron buenos, ya que esta lista no representa más que una mínima parte de las palabras contenidas en el corpus (poco más de ciento cincuenta palabras en la lista de filtrado contra la extensión total del corpus de trabajo que es de trescientas mil).

Esto significa que las reglas obtenidas fueron abarcadoras y generales para transcribir el léxico del siglo XVI. No podemos olvidar, sin embargo, que el periodo en cuestión presenta problemas por el reajuste de sonidos; por ello, se podrán encontrar muchas más palabras que no se ajusten a las reglas propuestas en esta investigación y que, posteriormente, la lista de filtrado podrá aumentar o las reglas podrán redefinirse.

Habrá que recordar, una vez más, que las reglas propuestas son para simplificar el trabajo del transcriptor automático y que de ninguna manera indican que los hablantes en realidad pronunciaran una palabra de la misma forma en todas las situaciones comunicativas. Es decir, si se toma una decisión, por ejemplo de transcribir *hazelle* > [a·sér·le] en lugar de [a·sé·ye] o [a·sé·je] indica que es la normalización **n** del atributo *tok*, del que ya se habló, que se utilizará para transcribir dicha palabra, pero no por ello debemos afirmar que el hablante del siglo XVI produjera en realidad lo que arroja dicha transcripción. Esto es, la lengua está en constante variación y movimiento, y los hablantes tienen diferentes realizaciones sonoras (al igual que sintácticas, léxicas, pragmáticas, etc.) dependiendo de cada situación comunicativa, de su entorno social, geográfico, temporal, etcétera.

## 6. BIBLIOGRAFÍA

### 6.1. Corpus

PI = BUELNA, MARÍA. 2006. *Los procesos inquisitoriales contra indígenas que realizó Fray Juan de Zumárraga en Nueva España (1536-1543)*, México: Universidad Autónoma Metropolitana.

DLNE = COMPANY, CONCEPCIÓN. 1994. *Documentos lingüísticos de la Nueva España. Altiplano central*, México: Universidad Nacional Autónoma de México.

DO = LOPE BLANCH, JUAN M. 1998. *El habla de Diego de Ordaz*, México: Universidad Nacional Autónoma de México pp. 193-228.

### 6.2. Referencias bibliográficas

ABAD, FRANCISCO. 1982. “Seseo y ceceo como problema de lingüística general”, en *Introducción plural a la gramática histórica*, Francisco Marcos Marín (rec.), Madrid: Cincel.

ALARCOS, EMILIO. 1950/1981. *Fonología española*, Madrid: Gredos.

ALATORRE, ANTONIO. 1979/2003. *Los 1001 años de la lengua española*, México: Fondo de Cultura Económica.

ALONSO, AMADO. 1952. *Historia del “ceceo” y del “seseo” españoles*, Bogotá: Instituto Caro y Cuervo.

———. 1953/1976a. *Estudios lingüísticos: temas hispanoamericanos*, Madrid: Gredos.

———. 1955/1976b. *De la pronunciación medieval a la moderna en español*, Madrid: Gredos.

ALVAR, MANUEL. 1982. “A vueltas con el seseo y el ceceo”, en *Introducción plural a la gramática histórica*, F. Marcos Marín (rec.), Madrid: Cincel.

ALZATE, JOSÉ DE. 1980. *Obras I-Periódicos, (1768, 1772-1773, 1787-1788)*, R. Moreno (ed.), México: Universidad Nacional Autónoma de México.

ARIAS, BEATRIZ. 1997. *El español en México en el siglo XVI. Estudio filológico de quince documentos*, México: Universidad Nacional Autónoma de México.

- BARTOLACHE, JOSÉ. 1983/1979. *Mercurio volante (1772-1773)*, R. Moreno (ed.), México: Universidad Nacional Autónoma de México.
- BLANCO, MARTA. 2006. *Aproximación a la cronología de las transformaciones funcionales de labiales y sibilantes del español*, Santiago de Compostela: Universidade, Servizo de Publicacións e Intercambio Científico.
- BUELNA, MARÍA. 2006. *Los procesos inquisitoriales contra indígenas que realizó Fray Juan de Zumárraga en Nueva España (1536-1543)*, México: Universidad Autónoma Metropolitana.
- BOYD-BOWMAN, PETER. 1985. *Índice geobiográfico de más de 56 mil pobladores de la América hispánica*, México: Fondo de Cultura Económica.
- . 1988. “Brotos de fonetismo andaluz en México hacia fines del siglo XVI”, *Nueva Revista de Filología Hispánica*, xxxvii, 1, pp. 75-88.
- CABRERA, CARLOS. 1992. “Las sibilantes medievales: Reflexiones sobre el problema de la desonorización”, *Medievalia*, 11, pp. 1-18.
- CANO, RAFAEL. 1988. *El español a través de los tiempos*, Madrid: Arco Libros.
- CATALÁN, DIEGO. 1982. “El fin del fonema /z/ [dz+ ~ z+] en español”, en *Introducción plural a la Gramática Histórica*, F. Marcos Marín (rec.), Madrid: Cincel.
- . 1989. *El español: Orígenes de su diversidad*, Madrid: Paraninfo.
- COCK, OLGA. 1969. *El seseo en el Nuevo Reino de Granada. 1550-1560*, Bogotá: Instituto Caro y Cuervo.
- COMPANY, CONCEPCIÓN. 1993. “Fonética novohispana a fines del siglo XVII” *Anuario de Letras*, xxxi, pp. 557-573.
- . 1994. *Documentos lingüísticos de la Nueva España. Altiplano Central*, México: Universidad Nacional Autónoma de México.
- CUERVO, RUFINO. 1901/1954. “El castellano en América”, en *Obras*, II, Bogotá: Instituto Caro y Cuervo, pp. 518-586.
- CUÉTARA, JAVIER. 2004. *Fonética de la ciudad de México. Aportaciones desde las tecnologías del habla*, tesis de maestría inédita, México: Universidad Nacional Autónoma de México.
- FONTANELLA DE WEINBERG, MARIA. 1976. *La lengua española fuera de España. América, Canarias, Filipinas, judeoespañol*, Buenos Aires: Paidós.
- FRAGO, JUAN. 1994. *Andaluz y español de América: Historia de un parentesco lingüístico*, Sevilla: Junta de Andalucía, Consejería de Cultura y Medio Ambiente.

- . 1996. “Formación del español de América”, en *Manual de dialectología hispánica. El español de América*, M. Alvar (dir.), Barcelona: Ariel, pp. 28-39.
- . 1999. *Historia del español de América*, Madrid: Gredos.
- GARCÍA, ANTONIO. 1988. *El español en México en el siglo XVI. Estudio lingüístico de un documento judicial de la Audiencia de Guadalajara (Nueva España) del año 1578*, Sevilla: Alfar.
- GARCÍA DE DIEGO, VICENTE. 1970. *Gramática histórica española*, Madrid: Gredos.
- GONZÁLEZ, MANUEL. 2005. *Las reformas a la ortografía del español: estudio histórico de los cambios más importantes*, tesis de licenciatura inédita, México: Universidad Nacional Autónoma de México.
- GRANDA, GERMÁN DE. 1994. “El proceso de koineización en el periodo inicial de desarrollo del español de América”, en *El español de América en el siglo XVI. Actas del simposio del Instituto Ibero-Americano de Berlín, 23 y 24 de abril de 1992*, J. Lüdtke (comp.), Madrid: Iberoamericana, pp. 87-108.
- GUITARTE, GUILLERMO. 1991. *Siete estudios sobre el español de América*, México: Universidad Nacional Autónoma de México.
- JAKOBSON, ROMAN. 1971. “Principios de fonología histórica”, en *El círculo de Praga*, Barcelona: Anagrama, pp. 104-129.
- LAPESA, RAFAEL. 1964. “El andaluz y el español de América”, en *Presente y futuro de la lengua española*, Madrid: Congreso de Instituciones Hispánicas, pp. 173-182.
- . 1981. *Historia de la lengua española*, Madrid: Gredos.
- LATHROP, THOMAS A. 1992. *Curso de Gramática Histórica Española*, Barcelona: Ariel.
- LÁZARO CARRETER, FERNANDO. 1953/1977. *Diccionario de términos filológicos*. Madrid: Gredos.
- LOPE BLANCH, JUAN M. 1969a. *La filología hispánica en México. Tareas más urgentes*, México: Universidad Nacional Autónoma de México.
- . 1969b. *El léxico indígena en el español de México*, México: El Colegio de México.
- . 1972. “La influencia del sustrato en la fonética del español de México”, en *Estudios sobre el español de México*, México: Universidad Nacional Autónoma de México, pp. 93-108.
- . 1988. “La labiodental sonora en el español de México”, *Nueva Revista de Filología Hispánica*, XXXVI, 1, pp. 153-170.

- . 1998. *El habla de Diego de Ordaz*, México: Universidad Nacional Autónoma de México.
- . 2000. *Español de América y español de México*, México: Universidad Nacional Autónoma de México.
- LÜDTKE, JENS. 1994. “Diferenciación y nivelación del español en la expansión a Canarias y al Caribe en el periodo de orígenes”, en *El español de América en el siglo XVI. Actas del simposio del Instituto Ibero-Americano de Berlín, 23 y 24 de abril de 1992*, J. Lüdtke (comp.), Madrid: Iberoamericana, pp. 39-56.
- LLOYD, PAUL M. 1987. *From Latin to Spanish. Vol. I: Historical Phonology and Morphology of the Spanish Language*, Pennsylvania: University of Pennsylvania.
- MARCHAL, BENOÎT. 2001. *XML con ejemplos*, México: Pearson Education.
- MARTINET, ANDRÉ. 1955/1974. *Economía de los cambios fonéticos. Tratado de fonología diacrónica*, Madrid: Gredos.
- MCENERY, TONY y ANDREW WILSON. 1996. *Corpus linguistics*, Edimburgo: Edinburgh University Press.
- MEDINA, ALFONSO. 2004. (Ms.) “Protocolo del Corpus Histórico del Español de México”, México: Universidad Nacional Autónoma de México.
- MEDINA, ALFONSO y CARLOS MÉNDEZ. 2006. “Arquitectura del Corpus Histórico del Español de México (CHEM)” en *Avances en la Ciencia de la Computación*, A. Hernández y J. Zechinelli (eds.), México: Sociedad Mexicana de Ciencia de la Computación, pp. 248-253.
- MENÉNDEZ PIDAL, RAMÓN. 1904/1940. *Manual de gramática histórica española*, Madrid: Espasa-Calpe.
- . 1926/1950. *Orígenes del español. Estado lingüístico de la Península Ibérica hasta el siglo XI*, Madrid: Espasa-Calpe.
- . 1962. “Sevilla frente a Madrid. Algunas precisiones sobre el español de América”, *Miscelánea homenaje a André Martinet*, pp. 99-165.
- MONDÉJAR, JOSÉ. 1991. *Dialectología andaluza*. Granada: Don Quijote.
- MORENO DE ALBA, JOSÉ G. 1972. *Historia de la lengua española*, México: Asociación Nacional de Universidades e Institutos de Enseñanza Superior.
- . 1988/2001. *El español en América*, México: Fondo de Cultura Económica.

- NAVARRO, TOMÁS. 1966-1967. “El alfabeto fonético de la *Revista de Filología Española*”, *Anuario de Letras*, VI, pp. 5-10.
- PARODI, CLAUDIA. 1977. “El yeísmo en América durante el siglo XVI”, *Anuario de Letras*, XXV, pp. 241-248.
- . 1995. *Orígenes del español americano*, México: Universidad Nacional Autónoma de México.
- PENNY, RALPH. 1993/2001. *Gramática histórica del español*, Barcelona: Ariel.
- PENSADO, CARMEN. 1993. “Sobre el contexto del cambio f > h en castellano”, *Romance Philology*, 47:2, pp. 147-176.
- QUILIS, ANTONIO. 1993/1999. *Tratado de fonología y fonética españolas*, Madrid: Gredos.
- REYNA, PALOMA. 2005. *El siglo XVIII en el altiplano central de México: Materiales para su estudio. Edición crítica, estudio filológico, introducción y notas*, tesis de licenciatura inédita, México: Universidad Nacional Autónoma de México.
- SALVADOR, GREGORIO. 1982. “Hipótesis geológica sobre la evolución de f- > h-”, en *Introducción plural a la gramática histórica*, F. Marcos Marín (rec.), Madrid: Cincel.
- TORRUELLA, JOAN y JOAQUIM LLISTERRI. 1999. “Diseño de corpus textuales y orales”, en *Filología e Informàtica. Nuevas tecnologías en los estudios filológicos*, J. M. Blecua, G. Clavería, C. Sánchez, J. Torruella (eds.), Barcelona: Universidad Autónoma de Barcelona-Editorial Milenio, pp. 45-77.
- VILLA-SEÑOR, ANTONIO DE. 1746. *Theatro americano. Descripción general de los reynos, y provincias de la Nueva-España, y sus jurisdicciones*, México: Imprenta de la Viuda de D. Joseph Bernardo de Hogal, Impressora del Real, y Apostolico Tribunal de la Santa Cruzada en todo el Reyno.
- WRIGHT, ROGER. 1982/1989. *Latín tardío y romance temprano en España y la Francia carolingia*, Madrid: Gredos.
- ZAMORA, JUAN y JORGE GUITART. 1982. *Dialectología hispanoamericana: Teoría, descripción, historia*, Salamanca: Almar.

### 6.3. Referencias electrónicas

Código ASCII (Wikipedia, consultado el 13 de noviembre de 2007 en: [http://es.wikipedia.org/wiki/American\\_Standard\\_Code\\_for\\_Information\\_Interchange](http://es.wikipedia.org/wiki/American_Standard_Code_for_Information_Interchange))

Unicode (Wikipedia, consultado el 13 de Noviembre de 2007 en:  
<http://es.wikipedia.org/wiki/Unicode>)

XML. (Wikipedia, consultado el 13 de Noviembre de 2007 en:  
<http://es.wikipedia.org/wiki/XML>)

## 7. APÉNDICE 1

| Alfabetos fonéticos (tabla de sonidos del siglo XVI) |           |                 |        |
|------------------------------------------------------|-----------|-----------------|--------|
| Sonido                                               | RFE       | AFI             | Mexbet |
| Bilabial oclusivo sordo                              | p         | p               | p      |
| Dental oclusivo sordo                                | t         | t               | t      |
| Velar oclusivo sordo                                 | k         | k               | k      |
| Bilabial oclusivo sonoro                             | b         | b               | b      |
| Dental oclusivo sonoro                               | d         | d               | d      |
| Velar oclusivo sonoro                                | g         | g               | g      |
| Prepalatal africado sordo                            | ŝ         | $\overline{ts}$ | ts     |
| Palatal africado sordo                               | ĉ         | $\overline{tʃ}$ | tS     |
| Labiodental fricativo sordo                          | f         | f               | f      |
| Alveolar fricativo sordo                             | s         | s               | s      |
| Prepalatal fricativo sordo                           | š         | ʃ               | S      |
| Velar fricativo sordo                                | x         | x               | x      |
| Glotal fricativo sordo                               | h         | h               | h      |
| Palatal fricativo sonoro                             | y         | ʒ               | Z      |
| Bilabial nasal                                       | m         | m               | m      |
| Alveolar nasal                                       | n         | n               | n      |
| Palatal nasal                                        | ɲ         | ɲ               | n~     |
| Alveolar vibrante simple                             | r         | r               | r(     |
| Alveolar vibrante múltiple                           | $\bar{r}$ | r               | r      |
| Alveolar lateral                                     | l         | l               | l      |
| Vocal alta anterior                                  | i         | i               | i      |
| Vocal media anterior                                 | e         | e               | e      |
| Vocal baja central                                   | a         | a               | a      |
| Vocal media posterior                                | o         | o               | o      |
| Vocal alta posterior                                 | u         | u               | u      |

## 8. APÉNDICE 2

### Condensación de reglas

| Vocal alta anterior [i]                                                         |                               |
|---------------------------------------------------------------------------------|-------------------------------|
| Contexto                                                                        | Regla                         |
| Una <i>i</i> gráfica                                                            | $i = [i]$                     |
| Si es una conjunción; es decir, si está entre dos espacios vacíos:              | $\emptyset y \emptyset = [i]$ |
| Si está entre una vocal y una consonante:                                       | $VyC = [i]$                   |
| Si es inicio de palabra y le sigue una consonante:                              | $\#yC = [i]$                  |
| Si es final de palabra, no importando si está precedida por vocal o consonante: | $y\# = [i]$                   |
| Si está antecedida por una consonante, sin importar qué grafía siga             | $Cy = [i]$                    |
| Si está entre una consonante y una vocal:                                       | $CjV = [i]$                   |
| Si es inicio de palabra y le sigue una consonante:                              | $\#jC = [i]$                  |
| Si está entre dos consonantes:                                                  | $CjC = [i]$                   |
| Si es final de palabra y va precedida de una consonante:                        | $Cj\# = [i]$                  |

| Vocal alta posterior [u]* |           |
|---------------------------|-----------|
| Contexto                  | Regla     |
| Una <i>u</i> gráfica      | $u = [u]$ |

| Vocal media anterior [e]* |           |
|---------------------------|-----------|
| Contexto                  | Regla     |
| Una <i>e</i> gráfica      | $e = [e]$ |

| Vocal media posterior [o] |           |
|---------------------------|-----------|
| Contexto                  | Regla     |
| Una <i>o</i> gráfica      | $o = [o]$ |

| Vocal central baja [a] |           |
|------------------------|-----------|
| Contexto               | Regla     |
| Una <i>a</i> gráfica   | $a = [a]$ |

| Consonante bilabial oclusiva sorda [p] |            |
|----------------------------------------|------------|
| Contexto                               | Regla      |
| Una <i>p</i> gráfica                   | $p = [p]$  |
| Aparición de <i>pp</i> gráfica         | $pp = [p]$ |

| <b>Consonante dental oclusiva sorda [t]</b> |                 |
|---------------------------------------------|-----------------|
| <b>Contexto</b>                             | <b>Regla</b>    |
| Una <i>t</i> gráfica                        | <i>t</i> = [t]  |
| Aparición de <i>tt</i> gráfica              | <i>tt</i> = [t] |
| Aparición de <i>th</i> gráfica              | <i>th</i> = [t] |

| <b>Consonante velar oclusiva sorda [k]*</b>                         |                   |
|---------------------------------------------------------------------|-------------------|
| <b>Contexto</b>                                                     | <b>Regla</b>      |
| Si le sigue una vocal central baja [a]                              | <i>ca</i> = [ka]  |
| Si le sigue una vocal media posterior [o]                           | <i>co</i> = [ko]  |
| Si le sigue una vocal alta posterior [u]                            | <i>cu</i> = [ku]  |
| Si el dígrafo va seguido de vocal media anterior [e]                | <i>que</i> = [ke] |
| Si al dígrafo le sigue una vocal alta anterior [i]                  | <i>qui</i> = [ki] |
| Si el dígrafo va seguido de consonante alveolar vibrante simple [r] | <i>chr</i> = [kr] |

| <b>Consonante bilabial oclusiva sonora [b]*</b> |                |
|-------------------------------------------------|----------------|
| <b>Contexto</b>                                 | <b>Regla</b>   |
| Una <i>b</i> gráfica                            | <i>b</i> = [b] |
| Una <i>v</i> gráfica                            | <i>v</i> = [b] |

| <b>Consonante dental oclusiva sonora [d]</b> |                 |
|----------------------------------------------|-----------------|
| <b>Contexto</b>                              | <b>Regla</b>    |
| Una <i>d</i> gráfica                         | <i>d</i> = [d]  |
| Aparición de <i>dd</i> gráfica               | <i>dd</i> = [d] |

| <b>Consonante velar oclusiva sonora [g]*</b>             |                 |
|----------------------------------------------------------|-----------------|
| <b>Contexto</b>                                          | <b>Regla</b>    |
| Una <i>g</i> gráfica seguida de <i>a, o, u, ü</i>        | <i>g</i> = [g]  |
| El dígrafo <i>gu</i> seguido de <i>i, e</i>              | <i>gu</i> = [g] |
| Una <i>g</i> gráfica seguida de las líquidas <i>r, l</i> | <i>g</i> = [g]  |

| <b>Consonante prepalatal africada sorda [ʃ]*</b> |                |
|--------------------------------------------------|----------------|
| <b>Contexto</b>                                  | <b>Regla</b>   |
| Una <i>ç</i> gráfica en léxico indígena          | <i>ç</i> = [ʃ] |

| <b>Consonante palatal africada sorda [ç]</b> |                 |
|----------------------------------------------|-----------------|
| <b>Contexto</b>                              | <b>Regla</b>    |
| Aparición de <i>ch</i>                       | <i>ch</i> = [ç] |

| <b>Consonante labiodental fricativa sorda [f]*</b> |                 |
|----------------------------------------------------|-----------------|
| <b>Contexto</b>                                    | <b>Regla</b>    |
| Una <i>f</i> gráfica                               | <i>f</i> = [f]  |
| Aparición de <i>ff</i> gráfica                     | <i>ff</i> = [f] |
| Aparición de <i>ph</i>                             | <i>ph</i> = [f] |

| <b>Consonante alveolar fricativa sorda [s]</b>                 |                |
|----------------------------------------------------------------|----------------|
| <b>Contexto</b>                                                | <b>Regla</b>   |
| Una <i>s</i> gráfica                                           | <i>s</i> = [s] |
| Una <i>ss</i> gráfica                                          | <i>s</i> = [s] |
| Una <i>c</i> gráfica seguida de vocal alta anterior [i] (i, j) | <i>c</i> = [s] |
| Una <i>c</i> gráfica seguida de vocal media anterior [e]       | <i>c</i> = [s] |
| Una <i>z</i> gráfica                                           | <i>z</i> = [s] |
| Una <i>ç</i> gráfica                                           | <i>ç</i> = [s] |

| <b>Consonante prepalatal fricativa sorda [ʃ]*</b>             |              |
|---------------------------------------------------------------|--------------|
| <b>Contexto</b>                                               | <b>Regla</b> |
| Si está, en léxico indígena, ubicada entre vocal y consonante | VxC= [ʃ]     |
| Si está ubicada entre vocales, en léxico indígena             | VxV= [ʃ]     |

| <b>Consonante velar fricativa sorda [x]*</b>   |              |
|------------------------------------------------|--------------|
| <b>Contexto</b>                                | <b>Regla</b> |
| Si está entre vocales                          | VjV= [x]     |
| Si es inicial de palabra y le sigue una vocal  | #jV= [x]     |
| Si está entre vocales                          | VxV= [x]     |
| Si es inicial de palabra y le sigue una vocal  | #xV= [x]     |
| Si está seguida de vocales anteriores [e], [i] | ge, i= [x]   |

| <b>Consonante glotal fricativa sorda [h]</b>                                     |              |
|----------------------------------------------------------------------------------|--------------|
| <b>Contexto</b>                                                                  | <b>Regla</b> |
| Si es inicial de palabra y proviene de una palabra latina con <i>f</i> - inicial | #h= [h]      |

| <b>Consonante palatal fricativa sonora [y]</b>              |                 |
|-------------------------------------------------------------|-----------------|
| <b>Contexto</b>                                             | <b>Regla</b>    |
| Si es inicial de palabra y a continuación hay una vocal     | #yV= [y]        |
| Si está entre dos vocales                                   | VyV= [y]        |
| El dígrafo <i>ll</i> siempre tendrá correspondencia con [y] | <i>ll</i> = [y] |

| <b>Consonante nasal bilabial [m]</b> |                |
|--------------------------------------|----------------|
| <b>Contexto</b>                      | <b>Regla</b>   |
| Aparición de <i>m</i> gráfica        | <i>m</i> = [m] |

| <b>Consonante nasal alveolar [n]</b> |              |
|--------------------------------------|--------------|
| <b>Contexto</b>                      | <b>Regla</b> |
| Aparición de <i>n</i> gráfica        | $n = [n]$    |

| <b>Consonante nasal palatal [ɲ]</b> |                       |
|-------------------------------------|-----------------------|
| <b>Contexto</b>                     | <b>Regla</b>          |
| Aparición de <i>ñ</i> gráfica       | $\tilde{n} = [\eta]$  |
| Aparición de <i>gñ</i> gráfica      | $g\tilde{n} = [\eta]$ |

| <b>Consonante vibrante simple [r]</b>            |              |
|--------------------------------------------------|--------------|
| <b>Contexto</b>                                  | <b>Regla</b> |
| Aparición de <i>r</i> gráfica                    | $r = [r]$    |
| Si es final de palabra y va antecedita por vocal | $Vr\# = [r]$ |
| Cuando está entre una vocal y una consonante     | $VrC = [r]$  |

| <b>Consonante vibrante múltiple [r̄]</b>      |                    |
|-----------------------------------------------|--------------------|
| <b>Contexto</b>                               | <b>Regla</b>       |
| Cuando es inicial de palabra seguida de vocal | $\#rV = [\bar{r}]$ |
| Cuando es antecedita por <i>n</i>             | $nr\_ = [\bar{r}]$ |
| Cuando es antecedita por <i>l</i>             | $lr\_ = [\bar{r}]$ |
| Aparición de <i>rr</i> gráfica                | $rr = [\bar{r}]$   |

| <b>Consonante lateral [l]</b> |              |
|-------------------------------|--------------|
| <b>Contexto</b>               | <b>Regla</b> |
| Aparición de <i>l</i> gráfica | $l = [l]$    |

\* Con su respectiva lista de excepciones.

## 9. APÉNDICE 3

### Lista de excepciones

| Forma gráfica       | Transcripción fonética |
|---------------------|------------------------|
| <i>abdiencia</i>    | [au·djén·sja]          |
| <i>abnque</i>       | [áun·ke]               |
| <i>agño(s)</i>      | [á·ño]                 |
| <i>allcalde</i>     | [al·kál·de]            |
| <i>anxeles</i>      | [án·xe·les]            |
| <i>aquerdo</i>      | [a·kwér·do]            |
| <i>auxilio</i>      | [au·ksi·ljo]           |
| <i>avdiencia</i>    | [au·djén·sja]          |
| <i>avjsallo</i>     | [a·bi·sár·lo]          |
| <i>avn</i>          | [a·ún]                 |
| <i>avnque</i>       | [áun·ke]               |
| <i>avsencia</i>     | [au·sén·sja]           |
| <i>babtizó</i>      | [bau·ti·só]            |
| <i>bella</i>        | [bér·la]               |
| <i>belles</i>       | [bér·les]              |
| <i>castagño</i>     | [kas·tá·ño]            |
| <i>castigalle</i>   | [kas·ti·gár·le]        |
| <i>cavsa</i>        | [káu·sa]               |
| <i>cavsado</i>      | [kau·sá·do]            |
| <i>caxco</i>        | [kás·ko]               |
| <i>cibdad</i>       | [sjü·dád]              |
| <i>çinquenta</i>    | [sin·kwén·ta]          |
| <i>çivdad</i>       | [sjü·dád]              |
| <i>coaçacoalco</i>  | [ko·a·şa·ko·ál·ko]     |
| <i>conplilla</i>    | [kon·plír·la]          |
| <i>correxillo</i>   | [ko·re·xír·lo]         |
| <i>cuçamelco</i>    | [ku·şa·mél·ko]         |
| <i>chancilleria</i> | [kan·si·ye·rí·a]       |
| <i>chançilleria</i> | [kan·si·ye·rí·a]       |
| <i>charjdad</i>     | [ka·ri·dád]            |
| <i>chiavtla</i>     | [çjáu·tla]             |
| <i>dagño</i>        | [dá·ño]                |

| <b>Forma gráfica</b>  | <b>Transcripción fonética</b> |
|-----------------------|-------------------------------|
| <i>dalle</i>          | [dár·le]                      |
| <i>debdas</i>         | [déu·das]                     |
| <i>descrevilla</i>    | [des·kre·bír·la]              |
| <i>desenpegñar</i>    | [de·sen·pe·ñar]               |
| <i>despachalla</i>    | [des·pa·çar·la]               |
| <i>dezille</i>        | [de·sír·le]                   |
| <i>diegito</i>        | [dje·gí·to]                   |
| <i>dogña</i>          | [dó·ña]                       |
| <i>enbavquen</i>      | [em·báu·ken]                  |
| <i>engagñar</i>       | [en·ga·ñar]                   |
| <i>enojallos</i>      | [e·no·xár·los]                |
| <i>enpegñar</i>       | [en·pe·ñar]                   |
| <i>enpegñado</i>      | [en·pe·ná·do]                 |
| <i>enpegño</i>        | [en·pé·ño]                    |
| <i>enxerta</i>        | [en·sér·ta]                   |
| <i>examinando</i>     | [ek·sa·mi·ná·do]              |
| <i>España</i>         | [es·pá·ña]                    |
| <i>español</i>        | [es·pa·nól]                   |
| <i>españoles</i>      | [es·pa·nó·les]                |
| <i>examinaron</i>     | [ek·sa·mi·ná·ron]             |
| <i>examjnado</i>      | [ek·sa·mi·ná·do]              |
| <i>examjnar</i>       | [ek·sa·mi·nár]                |
| <i>exçeda</i>         | [ek·sé·da]                    |
| <i>exçediendo</i>     | [ek·se·djén·do]               |
| <i>exçelente</i>      | [ek·se·lén·te]                |
| <i>excelente</i>      | [ek·se·lén·te]                |
| <i>excelentemente</i> | [ek·se·len·te·mén·te]         |
| <i>excesivos</i>      | [ek·se·sí·bos]                |
| <i>exçessivas</i>     | [ek·se·sí·bas]                |
| <i>excessjva</i>      | [ek·se·sí·ba]                 |
| <i>exessivos</i>      | [ek·se·sí·bos]                |
| <i>expendera</i>      | [eks·pen·de·rá]               |
| <i>experiencia</i>    | [eks·pe·rjén·sja]             |
| <i>experiência</i>    | [eks·pe·rjén·sja]             |
| <i>experjencia</i>    | [eks·pe·rjén·sja]             |
| <i>expiriencia</i>    | [eks·pi·rjén·sja]             |
| <i>expressasse</i>    | [eks·pre·sá·se]               |
| <i>exsecutar</i>      | [ek·se·ku·tár]                |
| <i>extendido</i>      | [eks·ten·dí·do]               |
| <i>faga</i>           | [há·ga]                       |
| <i>faser</i>          | [há·ser]                      |

| <b>Forma gráfica</b> | <b>Transcripción fonética</b> |
|----------------------|-------------------------------|
| <i>faze</i>          | [há·se]                       |
| <i>fazer</i>         | [há·ser]                      |
| <i>fecho</i>         | [hé·ço]                       |
| <i>fechos</i>        | [hé·ços]                      |
| <i>fêchose</i>       | [hé·ço·se]                    |
| <i>fee</i>           | [fe]                          |
| <i>gerra</i>         | [gé·ra]                       |
| <i>gie</i>           | [gí·e]                        |
| <i>grand</i>         | [gran]                        |
| <i>gratificalla</i>  | [gra·ti·fi·kár·la]            |
| <i>guaçoçingo</i>    | [gwa·šo·şín·go]               |
| <i>guaxoçingo</i>    | [gwa·šo·şín·go]               |
| <i>guaxovçingo</i>   | [gwa·şou·şín·go]              |
| <i>haselle</i>       | [a·sér·le]                    |
| <i>hazelle</i>       | [a·sér·le]                    |
| <i>hexsaminados</i>  | [ek·sa·mi·ná·dos]             |
| <i>higeras</i>       | [i·gé·ras]                    |
| <i>humilldes</i>     | [u·míl·des]                   |
| <i>humillmente</i>   | [u·mil·de·mén·te]             |
| <i>injuria</i>       | [in·xú·rja]                   |
| <i>intelligencia</i> | [in·te·li·xén·sja]            |
| <i>jesuxristo</i>    | [xe·su·krís·to]               |
| <i>jhesuxristo</i>   | [xe·su·krís·to]               |
| <i>jllustre</i>      | [i·lús·tre]                   |
| <i>yorje</i>         | [xór·xe]                      |
| <i>jullio</i>        | [xú·ljo]                      |
| <i>lljcenciado</i>   | [li·sen·sjá·do]               |
| <i>Maragñón</i>      | [ma·ra·ñón]                   |
| <i>matalçingo</i>    | [ma·tal·şín·go]               |
| <i>máxcaras</i>      | [más·ka·ras]                  |
| <i>maxime</i>        | [mák·si·me]                   |
| <i>migel</i>         | [mi·gél]                      |
| <i>mill</i>          | [mil]                         |
| <i>mille</i>         | [mí·le]                       |
| <i>mixteca</i>       | [miks·té·ka]                  |
| <i>mjll</i>          | [mil]                         |
| <i>mostralla</i>     | [mos·trár·la]                 |
| <i>mudalle</i>       | [mu·dár·le]                   |
| <i>negociallo</i>    | [ne·go·sjár·lo]               |
| <i>nigños</i>        | [ní·nos]                      |
| <i>ningund</i>       | [nin·gún]                     |

| <b>Forma gráfica</b> | <b>Transcripción fonética</b> |
|----------------------|-------------------------------|
| <i>pagalla</i>       | [pa·gár·la]                   |
| <i>page</i>          | [pá·ge]                       |
| <i>pagño</i>         | [pá·ño]                       |
| <i>pax</i>           | [paks]                        |
| <i>pedilles</i>      | [pe·dír·les]                  |
| <i>pegñol</i>        | [pe·ñól]                      |
| <i>pegñón</i>        | [pe·ñón]                      |
| <i>perjudiciales</i> | [per·xu·di·sjá·les]           |
| <i>pesqueços</i>     | [pes·kwé·sos]                 |
| <i>plégele</i>       | [plé·ge·le]                   |
| <i>proxima</i>       | [prók·si·ma]                  |
| <i>próximo</i>       | [prók·si·mo]                  |
| <i>proximo</i>       | [prók·si·mo]                  |
| <i>qual</i>          | [kwal]                        |
| <i>quenta</i>        | [kwén·ta]                     |
| <i>quenten</i>       | [kwén·ten]                    |
| <i>querdamente</i>   | [kwer·da·mén·te]              |
| <i>quidado</i>       | [kwj·dá·do]                   |
| <i>quixieren</i>     | [ki·sjé·ren]                  |
| <i>recabdo</i>       | [re·káu·do]                   |
| <i>respondelle</i>   | [res·pon·dér·le]              |
| <i>revsado</i>       | [reṽ·sá·do]                   |
| <i>rex</i>           | [rɛks]                        |
| <i>rodrígez</i>      | [ro·drí·ges]                  |
| <i>rrecavdo</i>      | [re·káu·do]                   |
| <i>ruegen</i>        | [rwé·gen]                     |
| <i>screvi</i>        | [es·kre·bí]                   |
| <i>scrivió</i>       | [es·kri·bjó]                  |
| <i>señalase</i>      | [se·ṽa·lá·se]                 |
| <i>señale</i>        | [se·ṽa·le]                    |
| <i>señalado</i>      | [se·ṽa·lá·do]                 |
| <i>señaló</i>        | [se·ṽa·ló]                    |
| <i>segund</i>        | [se·gún]                      |
| <i>subjetas</i>      | [sub·xé·tas]                  |
| <i>subjetos</i>      | [sub·xé·tos]                  |
| <i>subjettos</i>     | [sub·xé·tos]                  |
| <i>tepedçingo</i>    | [te·ped·şin·go]               |
| <i>texcuco</i>       | [teks·kú·ko]                  |
| <i>tilançi</i>       | [ti·lán·şi]                   |
| <i>tlaxcala</i>      | [tlaks·ká·la]                 |
| <i>tornacuxtla</i>   | [tor·na·kúks·tla]             |

| <b>Forma gráfica</b> | <b>Transcripción fonética</b> |
|----------------------|-------------------------------|
| <i>traella</i>       | [tra·ér·la]                   |
| v (conjunción)       | [u]                           |
| <i>valax</i>         | [bá·lax]                      |
| <i>vee</i>           | [be]                          |
| <i>veedor</i>        | [be·dór]                      |
| <i>veen</i>          | [ben]                         |
| <i>veer</i>          | [ber]                         |
| <i>vengen</i>        | [bén·gen]                     |
| <i>Villafagñe</i>    | [bi·ya·fá·ɲe]                 |
| <i>vn</i>            | [un]                          |
| <i>vna</i>           | [ú·na]                        |
| <i>vnçiones</i>      | [un·sjó·nes]                  |
| <i>vngría</i>        | [un·grí·a]                    |
| <i>vno</i>           | [ú·no]                        |
| <i>vsar</i>          | [u·sár]                       |
| <i>vse</i>           | [ú·se]                        |
| <i>xancol</i>        | [ʃan·kól]                     |
| <i>xicalango</i>     | [ʃi·ka·lán·go]                |
| <i>xrisptiano</i>    | [krisp·tjá·no]                |
| <i>xrisptóval</i>    | [krisp·tó·bal]                |
| <i>xristoval</i>     | [kris·tó·bal]                 |
| <i>yçucar</i>        | [i·ʃú·kar]                    |