Boletín de Filología, Tomo XLVIII
Número 1 (2013): 53 - 80

Textos chilenos del período colonial: ediciones, proyectos y perspectivas*

*Manuel Contreras Seitz***
Universidad Austral de Chile

Resumen

La documentación chilena del período colonial ha sido tratada por los diversos especialistas, sobre todo a partir de textos oficiales (decretos, reales cédulas, etc.) o literarios (crónicas, relatos, cartas de relación, autobiografías, por mencionar algunos), pero se ha dejado de lado, de manera sistemática, aquellos manuscritos que involucran la cotidianeidad y la inmediatez en la escritura. Desde el punto de vista de la crítica textual, parece plausible la hipótesis de que esta situación se correlaciona con las dificultades que se encuentran en el momento de efectuar la lectura desde el original. Abordaremos en esta comunicación los principales escollos para el trabajo paleográfico en los textos chilenos coloniales, las diversas ediciones que se han hecho sobre esta base documental, los proyectos que se relacionan con dicho material textual y las perspectivas metodológicas que pudieran contribuir a la construcción de una base sólida para mejorar y estandarizar las ediciones que se efectúen sobre un corpus manuscrito único.

* Este artículo se basa en los resultados del proyecto Fondecyt Nº 1100722, *Memoria documental en textos chilenos del período colonial. Transcripción y edición de manuscritos de los siglos XVI al XVIII*.

** Para correspondencia, dirigirse a: Manuel Contreras Seitz (manuelcontreras@uach.cl), Universidad Austral de Chile, Facultad de Filosofía y Humanidades, Instituto de Lingüística y Literatura, Campus Isla Teja s/n, Valdivia, Chile.

Palabras clave: filología chilena, edición de textos, Chile colonial.

Chilean Texts of Colonial Period: Editions, Projects and Perspectives

Abstract

The Chilean documentation of the colonial period has been treated by the diverse specialists, mainly starting from official (ordinances, royal decrees, etc.) or literary (chronicles, stories, relationship letters, autobiographies, to mention some) texts, but they have left aside, in a systematic way, those manuscripts that involve dailiness and the immediacy in writing. From the textual criticism point of view, it seems commendable the hypothesis that this situation is correlated with the difficulties that arise at the moment of reading from the original. In this communication we will approach the main obstacles for paleographic work on the Chilean colonial texts, the diverse editions that have been made on this documentary base, the projects that are related with this textual corpus and the methodological perspectives that could contribute to the construction of a solid base to improve and to standardize the editions that are made on a unique handwritten corpus.

Key words: Chilean philology, texts edition, colonial Chilean period.

Recibido: 04/03/13 Aceptado: 17/04/13

I. INTRODUCCIÓN

Una de las principales dificultades al momento de dirigir la vista hacia el pasado, aunque éste sea relativamente cercano y con una lengua común, como es el caso del periodo colonial, es la interpretación gráfica de los caracteres que conforman el acervo escriturario de un pueblo, sobre todo si las condiciones socioculturales que involucraron el desarrollo de la población no permitieron alcanzar –aun en el caso de órdenes religiosas de tanto prestigio y erudición como los jesuitas– un nivel educacional, y por ende una praxis escrituraria, que se mantuviese acorde con los parámetros dictados por las cancillerías regias, virreinales o por los grandes centros culturales europeos

o hispanoamericanos, como México y Lima. De esta complejidad se hace cargo Fernández (2009: 35) quien, al referirse a los aspectos gráficos de los textos que edita nos hace ver que:

> Una de las razones que han frenado el interés de lingüistas e historiadores por estos documentos es que la gran mayoría presenta dificultades para su lectura. Es frecuente encontrar copias de los escribanos del Consejo en los expedientes, en lugar de los originales, por alguna razón que desconocemos –quizá simplemente para facilitar su lectura o, tal vez, porque el papel en el que estuvieran escritas fuera de mala calidad o estuviera estropeado a causa del viaje–, pero también es frecuente hallar cartas escritas por individuos contratados para ello en América, que, sin llegar a ser escribanos, tenían conocimientos mínimos sobre las tradiciones epistolares, además de que tenían la capacidad de escribir, pero no todos esos escribientes contratados sabían hacerlo con la misma destreza, por lo cual la calidad gráfica y textual no siempre era la necesaria: de hecho, si el remitente había alcanzado tan sólo el primer grado de alfabetización y, a pesar de ello, se aventuraba a trasladar al papel lo que quería transmitir a sus familiares, la escasa calidad en la escritura de la correspondencia se incrementaba.

En los documentos chilenos coloniales esto es particularmente efectivo, sobre todo en la segunda mitad del XVI y primera del XVII. Es más, la historiografía chilena poco se refiere a este período sobre la base de fuentes primarias manuscritas más cotidianas, recurriendo con mayor propiedad a textos formales (reales cédulas, decretos, ordenanzas, etc.) que presentan menor dificultad en la lectura, pero que también entregan una visión sesgada de los acontecimientos, tanto sociales como lingüísticos.

Por ello, en este texto trataré de mostrar algunos aspectos del trabajo filológico, sobre todo enfocado hacia la constitución de un corpus documental colonial.

II. CORPUS DIACRÓNICOS: PROYECTOS Y EDICIONES

Al igual que los atlas lingüísticos, cuando se estudia la variación sincrónica del lenguaje, la constitución de un corpus diacrónico de la lengua es la base fundamental para el estudio histórico de los diversos aspectos de nuestro idioma, pues ya no solo se aprecia la óptica interna del estudio lingüístico,

sino que se va entretejiendo la red contextual que circunda y envuelve el texto que se está analizando[1].

Con todo, este ámbito propiamente filológico se ve actualizado y complementado en su metodología por medio de la *lingüística del corpus*, de larga data, pero que en tiempos recientes se ha redefinido y ha irrumpido en el ámbito hispánico de la investigación histórica, con sus aportes específicos en la constitución de corpus diacrónicos. Con anterioridad, en Finlandia, la Universidad de Helsinki fue una de las pioneras en el desarrollo de lo que podría llamarse un 'megacorpus' diacrónico, cuando en 1987 ya Ossi Ihalainen, Merja Kytö y Matti Rissanen daban cuenta de los avances del *Helsinki Corpus of English Texts*[2]; el proyecto se desarrolló entre 1984 y 1991, contando con un total de 1.572.800 palabras, con aproximadamente 450 textos (de entre 5.000 y 10.000 palabras) que abarcaron el milenio comprendido entre 710 y 1710.

A partir de este momento no solo se generaliza el término para referirse a la disciplina, sino que hay un nuevo impulso a ésta, producto del auge de la lingüística computacional, la mayor disponibilidad de corpus electrónicos y el desarrollo de tecnologías de reconocimiento óptico de caracteres (OCR) que facilita el ingreso de textos en las bases de datos. De hecho, el trabajo con corpus informatizados vs. el realizado con los impresos (en libros, esencialmente) permite otorgar una serie de ventajas que hasta ahora no habían sido consideradas en el ámbito de la historia del español de Chile.

En el ámbito de los estudios diacrónicos del español, claramente en Latinoamérica tenemos una deuda pendiente, sobre todo a partir de la necesidad, ya formulada por Guillermo Guitarte (1968)[3] de recurrir a

[1] O tal como lo expresa Jucker y Jacob (1995: 11) respecto de la pragmafilología: "Traditionally, historical linguists have spent most of their efforts on sound changes and on the phonology and morphology of historical texts. Syntax and semantics have always been less popular among the language historians. Pragmaphilology goes one step further and describes the contextual aspects of historical texts, including the addressers and addressees, their social and personal relationship, the physical and social setting of text production and text reception, and the goal(s) of the text".

[2] Para mayor información puede consultarse: http://www.helsinki.fi/varieng/CoRD/corpora/ HelsinkiCorpus/index.html

[3] Señala el autor (p. 159): "Me refiero al estudio de la lengua en los documentos guardados en los archivos de América. Poco se ha hecho a este respecto, y debemos confesar que, en este punto, la filología se halla en retraso frente a la historia, que hace tiempo utiliza las fuentes documentales para sus investigaciones. Para trazar la historia del español de América, en cambio, las fuentes han sido sobre todo los cronistas e historiadores de Indias; de manera adicional se han empleado textos literarios y gramáticas y, ocasionalmente, algunos documentos. Esto fue lo que ya hizo Cuervo, pero si en su tiempo era un procedimiento

los archivos para trazar la historia del español americano, aun cuando encontramos iniciativas relacionadas con la constitución de diccionarios históricos, que por su parte también tienen como base un determinado corpus diacrónico, aunque no siempre éste sea público o conocido. Por ejemplo, el *Léxico Hispanoamericano* de Peter Boyd-Bowman es una obra de gran ayuda para este ámbito, pero de la cual carecemos de acceso real a sus fuentes. En los últimos tiempos, en todo caso, muchos de estos proyectos han estado publicando referencias parciales o algunas de sus fuentes, con el fin de que lleguen a una mayor cantidad de investigadores e interesados en la materia. En tal sentido menciono los *Documentos para la Historia Lingüística de Hispanoamérica* (Fontonella 1993, Rojas 2000 y 2008 a y b), los *Documentos Lingüísticos de la Nueva España* (Company 1994; Melis et al. 2008) o el volumen *Ilegibilidad y Cotidianeidad* que recoge 102 documentos chilenos del período colonial, integrantes del CorDECh (Contreras 2005).

De todas maneras, no deja de ser desproporcionada la relación existente entre los corpus diacrónicos de la lengua española frente a los sincrónicos, principalmente orales.

Arellano y Rodríguez (1999) y Arellano y del Pino (2004) dan cuenta, también, de los derroteros no exentos de dificultades que han tenido que sortear las, en ocasiones escasas, ediciones de textos coloniales, particularmente las crónicas de las primeras épocas de la conquista hispana en tierras americanas[4]. Aportando a esta discusión, precisamente, Kordić (2006:

aceptable y representó un gran progreso, creo que hoy podemos ampliar la base de nuestros conocimientos". Alvar (1992) señalaba que: "Porque el conocimiento de la realidad americana desde comienzos del siglo XVI tendrá que hacerse por transcripción rigurosa de los textos y el conocimiento de la procedencia de los colonizadores. El primer motivo debe llevarnos a la formación de unas colecciones documentales tan rigurosas en sus lecturas como las que en su día hicieron Staaff, Menéndez Pidal y Navarro Tomás con referencia a los dialectos peninsulares".

[4] En este sentido, Arellano (1999) señala, en varios pasajes, en relación con las crónicas que: "En efecto, muy pocas de estas ediciones pueden ser consideradas ediciones críticas, y lo que es peor, muchas de ellas no son ni siquiera fiables. El texto que en ellas se ofrece es de procedencia más que dudosa, pues raramente se cotejan primeras ediciones y manuscritos originales. No se incluye, pues, ningún tipo de variantes. Además el texto se ha modernizado a capricho del editor" (p. 45), o como dice en otra parte: "Lo que sucede en la colección a que me refiero [Historia 16] es que la falta de competencia filológica de muchos editores provoca una modernización *sui generis*, que borra estructuras sintácticas o morfologías léxicas creyendo que se trata de incorrecciones del autor, creando generalmente verdaderas incorrecciones sustitutorias de redacciones originales aceptables" (p. 50). También Marrero-Fente (2004: 284) indica algo parecido: "En el caso de las crónicas iniciales de la conquista del Perú, no existe un criterio uniforme de edición, y la mayoría de las ediciones no hacen un análisis filológico del texto. /.../ La presencia de lecturas descuidadas incide en la falta de

198) da cuenta de las dificultades y de las características metodológicas para "la resolución de los pasos obscuros o lecturas inciertas" en los textos chilenos coloniales que, a diferencia del corpus peninsular, se refiere a documentos manuscritos originales. En este sentido, se concuerda con la opinión de la autora cuando dice que:

> El contar con testimonios manuscritos únicos, producidos por lo general por escritores de un nivel de educación formalizada baja, que no siguen normas gráficas claras o coherentes, permite considerar a estos documentos como fuentes lingüísticas históricas confiables y conservar sus rasgos característicos, y no sólo permite, sino que exige el abordamiento de estos aspectos para esclarecimiento del mensaje, cuestión que en muchas obras peninsulares resulta discutible, debido a la intervención de los impresores o a la contaminación entre distintos ejemplares, factores que pueden redundar en la deturpación absoluta de rasgos lingüísticos originales.

Desde este punto de vista, y en relación con el quehacer tradicional de la filología y de las ciencias historiográficas, el recurrir a la paleografía como herramienta esencial de apoyo al trabajo documental ha sido una característica indiscutida en el trabajo investigativo, por un lado, pero también un elemento esencial en la formación de especialistas en estas áreas. Esta disciplina desde sus comienzos ha sido, sobre todo, un poderoso complemento de la historia y, aun cuando tenga sus objetivos propios y una metodología específica, ha seguido estrechamente los derroteros de la misma, sobre todo si consideramos que es el primer paso que cualquier estudioso en la materia debe dar si quiere asomarse, por sus propios medios, a la historia cultural de un pueblo[5].

análisis comparados entre el texto base y su comentario crítico, en la escasez de cotejos de variantes, y en la falta de un criterio sistemático y uniforme en el tratamiento del texto, en especial la modernización ortográfica, elemento central en la fijación textual".

[5] En cuanto a la relación de la paleografía con la lingüística, Nuñez (1994: 26) señala que "la escritura es la fijación del lenguaje por medio de unos signos establecidos y, por lo tanto, está íntimamente ligada a los fenómenos que lo rigen. El filólogo necesita de la Paleografía primordialmente para establecer las leyes del desarrollo y del funcionamiento de la lengua. Lo que la palabra es a la Lingüística es la escritura a la Paleografía. Cuando la escritura alcanzó el alfabetismo se convirtió en un sistema cuyos elementos mínimos indescomponibles –letras o grafemas– representan los sonidos simples articulados por la voz del hombre; es decir, los morfemas".

2.1. El Corpus Histórico del Español de México (CHEM) y el CORDIAM

En el ámbito hispanoamericano, dada su tradición filológica, destaca la implementación del ***CHEM*** a cargo, como no, de la UNAM. Este corpus se ha puesto en marcha a partir del 2005, mediante un proyecto interno a cargo del Dr. Alfonso Medina Urrea, coordinado por el Instituto de Ingeniería[6]. Contempla la adición de materiales aportados por lingüistas, filólogos e historiadores, considerados claves para la representación del español de México entre los siglos XVI y XIX. La primera versión de este proyecto contempla la informatización de los *Documentos Lingüísticos de la Nueva España* (1994), 320 textos transcritos por Concepción Company, que abarcan todo el período colonial, desde 1525 a 1816. El corpus contempla calas históricas[7], una delimitación geográfica a la zona del Altiplano Central (ciudad de México y actuales Estados de Puebla, Querétaro, Tlaxcala, Guanajuato y Morelos; se han incluido en este corpus algunos documentos de las zonas de Oaxaca, Michoacán, Guerrero, Jalisco y de la sierra del Estado de Veracruz), así como criterios temáticos, donde se ha tratado de privilegiar el carácter coloquial de los textos. La autora recoge *cartas* –de emigrantes, misioneros, al Rey, al Consejo y, de menor frecuencia, las más coloquiales–, *denuncias y testimonios en juicios* –particularmente los más "populares", como aquellos por asesinato, brujería, blasfemia, ultrajes y despojos, además de juicios de residencia–, *inventarios y testamentos* –particularmente valiosos para el estudio léxico de los ámbitos de la vida cotidiana– y *peticiones e informes* –por la presencia de diversas copias y cómo aquello podría reflejar etapas del cambio fonológico, además de que en los primeros, sobre todo, puede encontrarse el escaso material relacionado con el español aprendido por indígenas.

[6] Incluye la participación de la UNAM (a través del Instituto de Ingeniería, Instituto de Investigaciones Filológicas, Facultad de Ingeniería y Facultad de Filosofía y Letras), del Instituto Mora y de la Brigham Young University. El patrocinio es de la DGAPA (Dirección General de Apoyo al Personal Académico) por medio del proyecto IN400905, 2005-2007.

[7] Señala Company (1994: 4) al respecto que: "Los materiales fueron seleccionados eligiendo un grupo considerable de documentos cada cincuenta años aproximadamente, en el supuesto de que en ese lapso, unas dos generaciones, los cambios pueden hacerse más fácilmente perceptibles en lengua escrita. Para cada corte cronológico consideré un margen de flexibilidad de búsqueda de unos quince años, y así se establecieron siete etapas que corresponden a seis cortes cronológicos, aproximadamente: 1525-1540, 1570-1585, 1620-1635, 1680-1695, 1735-1750, 1780-1795 y 1805-1820".

Por otra parte, Company incluye un criterio de "origen del autor", esto es, descontando el período en que la Nueva España es territorio poblado exclusivamente por españoles (los primeros 50-60 años), procura que el autor del documento haya nacido en suelo mexicano. Junto con ello, trata de incluir una representación "racial y social" que dé cuenta de la conformación de la sociedad mexicana de la Colonia (castizos, criollos, españoles, indios, mestizos, mulatos, negros, portugueses), aunque también incluye una buena cantidad de documentos de cuyo autor no se tiene referencia (44,7%).

La iniciativa más reciente, en el ámbito americano, es la correspondiente al ***CORDIAM*** (*Corpus Diacrónico y Diatópico del Español de América*), bajo la coordinación académica de Virginia Bertolotti (Universidad de la República, Uruguay) y de Concepción Company (UNAM, México) y al alero de la Academia Mexicana de la Lengua. Esta iniciativa, que ve la luz a fines del 2011, "tiene como objetivo general subsanar en gran parte las carencias en nuestro conocimiento de la diacronía y diatopía del español de América", como se señala en el documento de presentación del proyecto. Y a diferencia de otros corpus (CORDE, CHARTA, Corpus del Español, por ejemplo), solo se consignarán en él textos americanos, recogidos directamente de archivos, esto es, no incluirá en su acervo textos literarios ni periodísticos, para así contribuir a la realización de una dialectología histórica de la lengua española, centrada en el español americano, sirviendo, además, de punto de partida para el estudio de fenómenos de contacto lingüístico, problemas de gramática del español, así como el abordaje de problemas teóricos sobre el cambio lingüístico en sus diversos niveles de análisis.

Este corpus abarca 400 años de historia documental, entre 1494 y 1899, de 18 países hispanoparlantes, junto con documentación colonial hispana de Haití, Jamaica y Estados Unidos (California y estados sureños). Se justifica el margen cronológico al señalarse en su documento fundacional (CORDIAM 2011: 4) que:

> No obstante que la mayoría de los países hispanoamericanos logró su independencia de la Corona española en la segunda década del siglo XIX, el corpus abarca la totalidad de este siglo [XIX] porque ese periodo, ya independiente, es clave para entender la actual configuración dialectal de todos los países hispanoamericanos. De modo particular, es un siglo imprescindible para entender la diacronía del español de algunos países; por ejemplo, las variantes dialectales actuales de países como Argentina, Chile o Uruguay no pueden ser comprendidas sin los movimientos migratorios del siglo XIX.

A la fecha, este corpus está constituido por 3.448 documentos, de extensión variada, procedentes de 59 archivos de los países que conforman la muestra

documental del CORDIAM. Su principal característica tipológica es, precisamente, la de contar con todos los tipos que eran de uso habitual en la época, muchos de ellos, inclusive, pueden ser calificados de coloquiales e informales, claro está, en la medida en que la lengua escrita permea parte de aquella informalidad propia del uso oral. Si se revisa el contexto general, los documentos están compuestos por españoles, criollos, mestizos, indígenas, mulatos, etc., además de inmigrantes de nacionalidades diferentes a la hispana, que escribieron en español desde latitudes americanas.

2.2. La filología en Chile y el CorDECh

En tanto, la filología chilena ha pasado por diversos derroteros con preocupaciones, ya desde los inicios de la República, relacionadas con el desarrollo de los usos ortográficos (particularmente en la discusión sobre la norma que debería estandarizarlos), los estudios lexicológicos y lexicográficos (principalmente enfocados en la corrección idiomática y en los usos diferenciales), las obras dedicadas a las lenguas indígenas (con especial dedicación al 'araucano'), entre otros. Con todo, Rojas (1940: 171)[8], al referirse a los estudios sobre gramática y métrica históricas, ya señalaba en esta época que:

> Esta rama de los estudios filológicos, que bien pudiéramos considerar como fundamental, ha sido poco cultivada en nuestro país, no sabemos si porque ella exige una dedicación permanente y una profunda versación en asuntos lingüísticos, o por requerir un cabal conocimiento del latín y aun del griego, idiomas cuyo estudio se ha abandonado en Chile desde hace años.

Bien puede extrapolarse esta situación a las ediciones de textos, particularmente las del período colonial chileno, las que parecen haber experimentado parte de la animosidad con la cual se ha mirado, y aún se sigue observando, el periodo histórico vinculado con la Corona española en Hispanoamérica. Ha sido, más bien, desde el interés histórico que se

[8] Guillermo Rojas establece una guía crítica de los estudios filológicos en Chile, desde el período colonial hasta fines de la década de 1930, con lo cual permite una panorámica completa del conocimiento generado hasta ese entonces, refiriendo no solo las obras o autores de mayores connotación, sino todos aquellos que publicaron sus trabajos vinculados con la temática en estudio, obras que clasifica a partir de la lectura exhaustiva de las mismas.

rescataron algunos textos, en principio, como la *Colección de Documentos Inéditos para la Historia de Chile*, de José Toribio Medina; la *Crónica y Relación Copiosa y Verdadera de los Reinos de Chile*, en transcripción de Irving Leonard (1966); las *Fuentes para la Historia del Trabajo en el Reino de Chile*, de Álvaro Jara y Sonia Pinto (1982-1983), las *Cartas de Mujeres en Chile. 1630-1885*, de Sergio Vergara (1987), los *Protocolos de los Escribanos de Santiago*, de Álvaro Jara y Rolando Mellafe (1996), los *Testamentos de Indios en Chile Colonial*, de Julio Retamal (2000), o también el valioso esfuerzo que realizó la Editorial Universitaria, en su colección de *Escritores Coloniales de Chile*, en la cual se difundieron fragmentos de obras literarias de dicho periodo. Con todo, el esfuerzo de versiones filológicas recién comienza con el trabajo de Mario Ferreccio, quien inaugura la llamada *Biblioteca Antigua Chilena* (BACh)[9] con una edición crítica del *Purén Indómito*, donde se refieren exhaustivamente los criterios abordados para el estudio y la edición misma del manuscrito, continuando dicha labor, en la actualidad, la filóloga Raïssa Kordić, de la Universidad de Chile.

En la actualidad, sobre la base de textos no literarios, se viene implementando el proyecto del *CorDECh* (Corpus Diacrónico del Español de Chile)[10] el que, según las normas STEL (2007) es de carácter monolingüe (da cuenta de una lengua o variedad lingüística), grande (no tiene límite de palabras, o es muy elevado; no sigue criterios de representatividad o equilibrio), diacrónico (textos de diferentes etapas temporales; permite observar el desarrollo histórico de la lengua), textual (formado por textos enteros) y simple (de datos brutos, no anotados o codificados). Se ha conformado como un corpus simple, por cuanto la etapa de trabajo que se ha fijado en este momento apunta a la constitución de un volumen importante de textos, a fin de hacer visible una parte fundamental del período histórico de los siglos XVI y XVII, donde se pondrá especial énfasis, dado que es la etapa más compleja de leer en la documentación, debido al tipo de letra que la caracteriza. Para estimar una cifra, se supone que el CorDECh estaría

[9] La BACh tiene a su haber los siguientes títulos: *Purén Indómito*, Diego Arias de Saavedra; *Relación Autobiográfica*, Úrsula Suárez; *El Ignacio de Cantabria*, Pedro de Oña; *La Guerra de Chile,* Anónimo; *Cautiverio Feliz*, Francisco Núñez de Pineda y Bascuñán; *Testamentos Coloniales Chilenos*; *Epistolario de Sor Dolores Peña y Lillo (Chile, 1763-1769)*, los dos últimos en edición crítica de Raïssa Kordić, de quien pronto esperamos, además, la edición crítica de la *Crónica* de Jerónimo de Vivar.

[10] El CorDECh tiene su base en los proyectos Fondecyt Nº 1100722 (cfr. nota 1) y Nº 1040072 (*Pragmática histórica del español de Chile. Textualidades y contextos en el español de Chile del período colonial*), así como en las publicaciones asociadas de Contreras (2012, 2009a, 2009b, 2008, 2005, 2004, 2003, 2001).

conformado por alrededor de 60.000 imágenes (1 imagen = 1 página de texto).

La mayor parte de este corpus tiene un carácter inédito, ya que los textos se han transcrito directamente de los archivos, o bien, teniendo a mano una fotocopia o edición digitalizada de los mismos. En algunos casos, que se especifican, se han incorporado los textos chilenos que forman parte de los *Documentos para la Historia Lingüística de Hispanoamérica*, los cuales han sido editados teniendo a la vista la fotocopia de los originales, desplegando las abreviaturas y, en algunas ocasiones, corrigiendo las transcripciones efectuadas en su momento, pero siguiendo en su mayoría las normas adoptadas por la Comisión de Estudio Histórico del Español de América de ALFAL.

Los textos seleccionados tratarán de ser una muestra documental proporcionada del discurso informal, tanto oficial como no oficial, pues sobre todo en este último tipo de discursos el redactor tiende a preocuparse más por el contenido de tales documentos y mucho menos de la forma. Dentro de los textos transcritos, es probable que la mayor parte de ellos se caracterice por su naturaleza más bien oficial que no oficial, ya que la experiencia señala que siguen siendo muy pocos los encontrados que respondan a esta calidad.

Aun cuando se reconozca la supremacía de los textos privados sobre los oficiales, por los autores citados anteriormente, este criterio no siempre es aplicable, debido a que la documentación celosamente resguardada, atesoradora de los más invaluables y significativos manuscritos, que dan testimonio de las diversas actividades de la vida pública nacional, mal que le pese al investigador, solo dan cuenta de la vida privada y, junto con ello, del modo familiar o cotidiano de la expresión, cuando la emisión o recepción de ellos podía estar a cargo de algún personaje connotado o trascendente en el quehacer de la comunidad de la época. Esto dificulta aún más el poder lograr un continuo estilístico que represente a dicha sociedad, pues debe recordarse que el porcentaje de alfabetismo no era muy alto, con lo cual el nivel de quienes escriben estará, necesariamente, sesgado en lo social. Sin embargo, para suplir estas posibles deficiencias, se hará acopio de documentos lo suficientemente valiosos, aun cuando no se trate de correspondencia privada, para los fines de este estudio, siguiendo así lo expresado por Frago (1987: 70), en cuanto a que

> la conveniencia de recurrir a un equitativo análisis de corpus no literarios a la hora de investigar en las diferentes facetas históricas del español se acentúa, y aun se convierte en necesidad esencial, cuando del español americano se trata, ya que el período de su configuración como habla regional diferenciada (...) es de una literatura hispanoamericana de escasa

> identidad (...). En cambio, disponemos de un inabarcable acervo textual del todo válido para el más exigente análisis filológico (...).

En este sentido, dentro del marco informal, podremos distinguir aquellos que se hallen en los *textos privados*, correspondiente a cartas, testamentos propios, confesiones, etc.; así como los enmarcados dentro de los *textos públicos no oficiales*, como actas, declaraciones, juicios, inventarios de bienes, cartas poder, de compra-venta, etc. Esto nos permitirá distinguirlos –y compararlos– con aquellos textos manuscritos *oficiales*, tales como reales cédulas, pregones, mercedes, etc.

En todo caso, la base del CorDECh es la transcripción paleográfica estricta, dado que este tipo de transcripción sirve de sustento para cualquier tipo de estudio lingüístico, histórico o sincrónico, permite la proyección hacia diccionarios históricos, y se constituye como fuente para la difusión del material transcrito, reeditándolo para alcanzar públicos diversos. Por otra parte, la digitalización del documento presenta invaluables ventajas a la hora de transcribir y de conformar los abecedarios internos de cada documento y de cada autor. Este trabajo puede tener proyecciones, inclusive, para el desarrollo de *software* de reconocimiento óptico de caracteres manuscritos, en la medida en que, habiendo recopilado un corpus documental amplio, se puedan establecer los rasgos fundamentales de cada grafema –así como de su conjunto de alógrafos–, independientemente del individuo que los traza, por medio de las teorías específicas dentro del campo del modelamiento matemático.

III. NEOFILOLOGÍA Y DIGIPALEOGRAFÍA

Con el fin de salvaguardar los primeros aspectos en torno a la transcripción del manuscrito, en referencia específica a los textos chilenos coloniales, es que he querido trazar, al menos, una primera ruta por la que podría recorrerse este camino, en consideración a que las dificultades iniciales, precisamente de lectura y decodificación del texto, son las principales causas de la problemática filológica, específicamente en lo que toca al campo de la paleografía. No quiere decir esto que no se necesite la presencia de especialistas o de estudios paleográficos, muy por el contrario, lo que se pretende es esbozar una herramienta metodológica que permita facilitar la inserción de nuevas generaciones de filólogos en el cada vez más complejo

'arte' de la lectura y transcripción de textos coloniales, particularmente los de las primeras épocas.

En este sentido, es plausible pensar que combinando el conocimiento proporcionado por las disciplinas historiográficas, junto con los avances en el campo de la informática y el modelamiento matemático, específicamente en el ámbito del reconocimiento de patrones, puedan lograrse importantes avances en la configuración de un *software* que permitiese identificar con un grado más que aceptable de certeza (entre un 60% y un 80% al menos) tanto las complicadas grafías de una procesal encadenada como las más legibles de una bastarda italiana.

No es que la tarea sea novedosa en sí misma, pues el estudio de patrones manuscritos se ha venido desarrollando por, al menos, alrededor de 40 años. Sin embargo, lo que se propone tiene algunas características que pueden hacer de esta idea un centro de interés relevante: en primer lugar, la realización de un trabajo mancomunado entre filólogos, lingüistas, historiadores, con investigadores informáticos y matemáticos; en segundo término, el plantear una problemática para un tipo específico de carácter, el del documento latinoamericano (hispano y luso) del período de conquista, asociado a las grafías más complejas de aquella etapa –la procesal y la cadenilla o encadenada– con lo cual se podría dar cuenta de una parte importante de la documentación existente en nuestros archivos. Con todo, no hay que dejar de considerar que la diferencia esencial en este planteamiento es la complejidad existente a la hora de segmentar un texto continuo, en relación con una variante discreta como una grafía manuscrita de tipo imprenta o versalita, o la grafía del impreso (letra de computador, por ejemplo).

Como se mencionó, el problema mayor está en el reconocimiento 'offline' de caracteres manuscritos, esto es, la dificultad en procesar una imagen escaneada de un documento para identificar una escritura que no se está llevando a cabo simultáneamente, por medio de una herramienta digital construida y programada para el reconocimiento de patrones implicados (por ejemplo, un escáner de huella digital o un autentificador de firma electrónica), sino que se busca determinar, a partir de un trazo continuo de un texto, por una parte, la discrecionalidad de las unidades gráficas, y por otra, sus posibilidades combinatorias, sus alógrafos y la construcción de un léxico específico que debe ser programado. En este caso en cuestión, se añade el hito de que se trata de una conformación lingüística en muchos casos desconocida y que, cuando menos, es 'antigua'. El caso es que, desde la óptica de la informática y el modelamiento, el método más extendido para el proceso de reconocimiento de textos manuscritos 'offline' parece ser el uso de herramientas basadas en el *Hidden Markov Model* (HMM), que permite la constitución de una herramienta estadística potente que posibilita

el modelamiento de secuencias, caracterizada por un proceso subyacente que genera una secuencia observable. O en otras palabras, el HMM es un conjunto infinito de estados, cada uno de los cuales se encuentra asociado a una distribución de probabilidades (usualmente multidimensionales). Las transiciones entre los diversos estados se rigen, a su vez, por un conjunto de probabilidades (probabilidades de transición). En un estado en particular se puede generar una 'salida' u observación, según la distribución de probabilidad asociada. Con todo, aquello es solo el resultado y no el estado visible a un observador externo, por cuanto dichos estados permanecen 'ocultos' hacia el exterior, de donde se deriva el nombre del modelo.

En este mismo sentido, pero desde la inteligencia artificial y la ingeniería cognitiva, Brink et al. (2012) plantean ya una interesante colaboración con el campo de la paleografía cuantitativa, en el momento en que contribuyen al campo de la identificación de un autor de un documento por medio de la caracterización de los trazos de tinta y su direccionalidad. A su vez, en el proceso de reconocimiento de los caracteres de un texto, que implica no solo la segmentación de grafías individuales y sus variantes para 'leer' una palabra, sino también la disposición de una fase de reconstrucción de 'líneas de texto', pueden reconocerse tres fases (cfr. Bertolami y Bunke 2008: 3453):

(a) **pre-procesamiento**, donde se reduce el impacto de los diversos estilos escriturarios en la extracción de la información, para lo cual la imagen de la línea de texto manuscrita se normaliza en relación con su oblicuidad, inclinación, posición de la línea de base y ancho promedio de la letra. Una vez regularizada, la línea de texto manuscrito se convierte en una secuencia de un conjunto de vectores que utilizan, gráficamente, una ventana deslizante de un pixel de ancho y que se desplaza de izquierda a derecha sobre la imagen, un pixel por paso. Por cada posicionamiento, se extraen 9 figuras geométricas que determinan los pixeles del primer plano de la imagen, además de la primera y segunda capa de los mismos, la posición del contorno superior e inferior y su primer orden derivativo y el número de transiciones verticales en blanco y negro, junto con la densidad de pixeles entre el contorno superior e inferior.

(b) **reconocimiento**, en el que cada caracter es analizado de manera independiente, con un modelo basado en dos transiciones (una propia y otra para el estado siguiente), de acuerdo a los algoritmos pertinentes.

(c) **post-procesamiento**: en esta etapa lo fundamental es el cálculo del 'valor de confianza', esto es, si dicho valor excede un determinado umbral, el reconocimiento resulta aceptable, de lo contrario, se

> rechaza. Este cálculo implica que, mientras mayor sea el valor dado tendrá una mayor relevancia dentro de las posibilidades combinatorias totales de la línea de texto. Tales posibilidades pueden estar basadas en probabilidades o en propuestas alternativas.

Estas consideraciones, en todo caso, al decir de Stutzmann (2010), pasan por una restricción paleográfica previa de carácter 'grafético'[11] que no hay que menospreciar, por cuanto es la base estructural de los siguientes pasos de informatización, ya que el modelamiento solo puede generarse a partir de la correcta definición de los caracteres que conforman la composición escrituraria del documento. En este sentido, y como ya se mencionó con anterioridad, las investigaciones relacionadas con el tipo de grafía antigua, en cierta medida, trata de soslayar el problema de la segmentación, dado que existe una complicación cierta a la hora de tratar de responder a estos parámetros conjuntos (alógrafo y segmentación) para la correcta interpretación de un texto manuscrito y su posterior representación gráfica.

IV. EL ABORDAJE TEXTUAL DE LOS DOCUMENTOS: UNA EDICIÓN EN CAPAS

Después de todas estas consideraciones, no deja de ser interesante, entonces, plantearse la manera en que debiera desarrollarse un trabajo que satisficiera todas las necesidades que se producen al momento de tener que abordar un documento, o mejor, un corpus documental de una época de la cual tenemos poca referencia o el imaginario colectivo ha tratado de ocultar. Casassas (2011) presenta una interesante propuesta, basada en su experiencia con textos romances aljamiado-árabes, de *edición en capas*. El autor propone los siguientes niveles: reproducción digital del manuscrito, edición del texto manuscrito en su grafía (árabe) original, transliteración y transcripción, y versión modernizada del texto dirigida a un público amplio.

[11] Se distinguen cuatro niveles de transcripción: "«*graphic*» rendant toute la forme, «*graphetic*» distinguant chaque type de chaque lettre; «*graphemic*» préservant la suite de lettres; «*regularized*» unifiant les suites de lettres attestées à une forme normalisée" (Stutzmann 2010: 249-250).

En cuanto a la primera, se recomienda digitalizar cada folio del manuscrito de manera independiente, con el fin de trabajar en una reproducción de buena calidad, ojalá limpia de los posibles defectos, inclusive, del manuscrito, por medio de un procesamiento informático adecuado. En tanto, a la segunda capa se le atribuye la noción de "edición paleográfica del texto tal como aparece en el manuscrito sin enmendar ni corregir los posibles errores o descuidos del autor o copista". Tal como hemos señalado en ocasiones anteriores, se considera que ésta debe ser la base para todas las demás, y particularmente en aquellos manuscritos de difícil lectura. Casassas señala que en este nivel debe asentarse el estudio codicológico y paleográfico del texto, considerando que aquí estaría la edición propia del manuscrito ("lecturas dudosas, referencias a posibles errores del autor o copista, particularidades del texto en el manuscrito difíciles de reproducir fielmente en la edición, etc..."). Al respecto (2011: 4-5), señala:

> Una edición será considerada como definitiva cuando esté cotejada con el original, contrastada, por varios investigadores y se compruebe que no se detectan más errores. Aparte de los errores hay que tener en cuenta las palabras o pasajes que debido a la forma de la escritura o por estar el manuscrito deteriorado son difíciles de leer o interpretar. En estos casos es siempre enriquecedor el poder disponer de una forma paralela de las diferentes interpretaciones o lecturas de cada uno de los investigadores que se hayan enfrentado, o se enfrenten en un futuro, con el texto en cuestión.

Respecto de la transliteración y transcripción, el autor las considera como capas independientes (tercera y cuarta capas), recordándonos que no son técnicas equivalentes, aun cuando en numerosas ocasiones se utilicen estos conceptos como sinónimos. La *transliteración* "tiene como fin la sustitución de las unidades gráficas de un texto /.../ con las unidades de otro sistema gráfico /.../ según un criterio preestablecido de equivalencias biunívocas, y sin que sea integrado ningún elemento que no se encuentre presente en el texto original" (p. 9), esto es, pretende dar cuenta del contenido exacto del texto para todos aquellos que no cuenten con las competencias necesarias para realizar una lectura adecuada de las grafías del mismo, a fin "de posibilitar la reconstrucción del texto original de manera unívoca". En cuanto a la *transcripción*, ésta se entiende como la representación de fonemas por medio de símbolos o grafías que no existen en la escritura usual, por lo que no hay relación unívoca entre el original y el texto transcrito.

Cabe destacar que, en cuanto a los criterios de transcripción, éstos fueron acordados por la Comisión de Estudio Histórico del Español de América en 1989, para la conformación de normas uniformes en el estudio

del desarrollo histórico de la lengua española en Hispanoamérica. Estos criterios se adoptaron sobre la base de la metodología y alcances teóricos de las ciencias auxiliares de la historia (paleografía y diplomática) por lo que, tal como señala Fontanella (1993: 2):

> En lo que hace a las normas de transcripción, en la reunión mencionada [Bahía Blanca] también se acordó que esta fuera de carácter literal estricto, de tal modo que se respetaría fielmente la grafía original del texto, apartándose sólo en cuanto se separarían las palabras que aparecen unidas en el original, así como se unirían las letras de una misma palabra que estén escritas separadas. También se conservarían las abreviaturas y se mantendría la acentuación y la puntuación original.

Este plan se ha seguido, en general, en la colección de *Documentos para la Historia Lingüística de Hispanoamérica. Siglos XVI a XVIII* (1993, 2000 y 2008). En los casos de Company (1994), Contreras (2005), Melis et al. (2008) y Rivarola (2009), también se despliegan las abreviaturas, indicándolo en cursiva.

Como quinta capa, Casassas (2011: 7) sugiere la versión modernizada, cuyo propósito sea "el de dar versiones del contenido que sean fáciles de leer y puedan llegar a un público lo más amplio posible". Como se señaló, el texto de base es la transcripción paleográfica, por lo que se parte de una versión con un significado concreto, pues debemos hacernos cargo de que, en este caso, no solo bastaría con regularizar la ortografía y la sintaxis al uso actual de la norma, sino que también se requiere de una comprensión del discurso de fondo, y representarlo de una manera adecuada para que no le sea extraño al lector contemporáneo. En el siguiente nivel, dado que el autor trabaja con textos (romances aljamiado-árabes) que son traducciones de obras en lengua árabe, propone que se pueda localizar la obra original. En nuestro caso, por ser un texto cuya lengua de partida y de llegada es la misma, esta sección pudiera no ser pertinente, salvo que se encontrasen textos o partes de ellos en una lengua distinta –generalmente el latín– y que pudiese, en esos casos, rastrearse la referencia original que muchas veces queda en etapa de *interlengua*, sobre todo en los textos jurídicos coloniales chilenos.

Todas estas propuestas se refieren al proceso de *reproducción* o de *producción* de distintas versiones del texto original.

Junto con éstas, no debieran faltar, cuando ya determinamos el ámbito de la edición crítica, un *estudio de contenido*, que permita situar la comprensión del texto en el marco global de la cultura y sociedad de la época en la que se ha producido; imprescindible, también, resulta la elaboración, al menos, de un *glosario* del léxico del manuscrito que no sea el habitual en la variante lingüística meta. Una sugerencia práctica puede ser la de que se acompañe

el lexema en la grafía original, en lo posible con sus respectivas alografías. Los *estudios lingüísticos* propiamente tales son inherentes a una labor como ésta; entre otras cosas, son un aporte fundamental a la hora de distinguir los manuscritos de acuerdo al desarrollo histórico de la lengua según las diversas zonas geográficas, épocas o estilos, así como el generar información para la datación de los textos y manuscritos. En los casos en que esto sea pertinente, además, un *estudio literario* sería ampliamente recomendable, con el fin de mostrar "las relaciones de solidaridad entre los modos de significar del texto de que se trata y las propiedades del género al que el texto pertenece" (Morales 2001: 14).

En último término, el autor propone la edición crítica del texto. En este sentido, para el caso de la documentación chilena, es necesario seguir las advertencias que formula Kordić (2010: 285):

> *edición crítica* ha de entenderse como sinónimo de texto filológicamente procesado, editado con criterios rigurosos y científicos acorde a la realidad y las necesidades de la producción local.
> /.../ Esta lengua española poco formalizada, más arcaica y vulgar, propia de la soldadesca o de los religiosos autodidactas, plantea el imperativo de un cuidadoso trabajo de anotación de fenómenos lingüísticos múltiples, fenómenos algunos de ellos que en la tradición española se han confundido con errores de copista (Kordić 2006).

En este sentido es fundamental destacar que existe una importante tradición filológica en la edición de manuscritos que ha desplegado, a través del tiempo, la *Biblioteca Antigua Chilena* (BACh)[12]. Para las ediciones de sus obras se emplearon las siguientes normas textológicas que Kordić (2005: 27-28) sintetiza de la siguiente manera:

> La reducción fonografemática realizada contempla el principio básico del respeto y conservación de todo rasgo gráfico que implique efectiva o eventualmente la representación de un rasgo fónico diferencial, significativo; todo aquel recurso que, tras el examen del comportamiento gráfico del escriba, demuestre ser inoperante, se moderniza, con el fin de evitar en el texto editado la presencia de inútiles grafías exóticas que sobrecarguen visualmente el texto y confundan al lector.

[12] Como ya mencionábamos, gracias a los esfuerzos realizados, en primer lugar, por Mario Ferreccio Podestá y, junto con él y en la actualidad, la misma Raïssa Kordić Riquelme.

Se simplifican las letras geminadas cuyo valor sea idéntico al de sus simples: *abrill* > *abril*; *ffresadas* > *fresadas*; *missas* > *misas*; *rrelijiosos* > *religiosos*.

Se actualizan los siguientes usos gráficos: *quales* > *cuales*; *Christo* > *Cristo*; *cathedral* > *catedral*; *Joseph* > *Josef*; *forçosa* > *forzosa*; *obligaçion* > *obligación*; *ziudad* > *ciudad*; *oxo* > *ojo*; *lijitimas* > *ligítimas*; *muger* > *mujer*; *bestido* > *vestido*; *auer* > *haber*; *hultima* > *última*.

Las nasales también se representan según el uso actual: *enpezar* > *empezar*; *emferma* > *enferma*. Las tildes o signos suprasegmentales de nasal se transcriben como *n*: *ẽ*(*ė*)*tierro* >

entierro; en caso de omisión de la misma, si la voz no se presta a confusión, se restituye sin comentario: *acompanen* > *acompañen*; se restituye la vocal protética en caso de presencia de *s* en posición implosiva inicial: *scribano* > *escribano*.

Se conservan, sin indicación alguna, las contracciones habituales de la época como *deste*, *della*, etc.; las contracciones circunstanciales se indican con un apóstrofo en el lugar correspondiente: *qu'es*, *qu'está*, etc.

Se conservan tales cuales los latinismos hispanizados o vulgarizados: *u supra*, *partesnostes*, *proter numsias*, etc.

Se aplican, en la medida en que la confusa sintaxis lo permite, las normas actuales de puntuación, acentuación y uso de mayúsculas y minúsculas.

Se señala entre corchetes toda intervención realizada en el texto, con justificación de la misma en nota textual al pie; cuando la causa de la enmienda es una y la misma para todo el documento –por ejemplo, imposibilidad de lectura por pérdida de un tramo del folio en la costura apretada del volumen– sólo se hace advertencia una única vez.

Se utilizan comillas dobles (« ») para las citas textuales idénticas, y comillas simples, (`´) con indicación de confrontar (*cf.*), para valores sémicos aproximados registrados en las fuentes referidas. Cuando se estima necesario, se recoge, junto con el valor sémico de uso puntual del contexto, otro próximo complementario.

Finalmente, en algunos casos, sería interesante considerar una capa adicional a las propuestas para el texto y que es la de la ilustración, como un estudio complementario. Lo que se pretende es que, luego que un ilustrador profesional investigue e interprete los textos referidos y presentadas las propuestas a un equipo interdisciplinario, se decida el tipo de imagen que se plasmará en la edición final, sobre la base de criterios históricos, filológicos y pedagógicos, con el fin de poner de relieve los aspectos fundamentales para la mejor comprensión del manuscrito, siendo fieles al espíritu de las obras de la época en que se sitúe el documento, en cuanto a técnica de ilustración y modelo gráfico.

El trabajo de edición de un texto colonial chileno –dadas las particulares características de la zona, en cuanto a periferia del mundo hispanoamericano, como otras más en el Continente– requiere de especial atención, a fin de explicar cabalmente la lengua y cultura de la época, contextualizar los usos que aparecen en el escrito y disminuir al mínimo las posibles alteraciones o adulteraciones del original. Como bien señala Kordić (2007):

> no solo hay que determinar valores sémicos que con frecuencia no están recogidos en diccionarios, sino que hay que comprender exóticas formaciones morfológicas y acertar en la interpretación de un discurso sintácticamente embarullado y a veces desconcertante. Cuando no se tienen los criterios, los métodos y la formación en materia idiomática y textológica, las conclusiones pueden desembocar en magnas (y a veces tendenciosas) confusiones ilustradas. La labor hermenéutica historiográfica debe hermanarse con la filológica.

Para facilitar el trabajo actual y futuro, así como la labor docente en esta disciplina, sería recomendable que las ediciones del CorDECh, u otras similares, incluyeran un abecedario paleográfico, las abreviaturas y su respectivo despliegue, así como un estudio previo del conjunto textual que se publica, con especial referencia a los aspectos que puedan parecer más oscuros al lector actual, dada su escasa familiaridad con el contexto histórico o con el tipo de lengua que se presenta, tal como se ha hecho sistemáticamente con los textos editados en la BACh.

V. CONSIDERACIONES FINALES

La documentación colonial en el Archivo Nacional de Chile, en comparación con otras realidades hispanoamericanas, es relativamente modesta. Sin embargo, el bajo número de investigadores dedicados al estudio de fuentes primarias y, más aún, la escasa producción de ediciones críticas de este tipo de documentación, hace que sea necesario dimensionar la magnitud de una tarea pendiente en el trabajo de archivos documentales en Chile, cual es la de sacar a la luz un considerable corpus, poniendo especial énfasis en los siglos XVI y XVII, de los cuales poco sabemos y, menos aún, conocemos en su vasta tipología textual.

Para muestra de lo que decimos, solo algunos ejemplos. El fondo "Capitanía General", en el Archivo Nacional Histórico (ANH), comprende

1.069 volúmenes que abarcan la historia nacional entre 1591 y 1840; el fondo "Contaduría Mayor" está compuesto por 8.547 volúmenes, de los cuales corresponden a la Primera Serie, 4.831, mientras que a la Segunda, un total de 3.716 vols., entre los años de 1579 y 1896. Otro fondo de especial relevancia es el de "Escribanos de Santiago", que abarca el período comprendido entre 1579 y 1800, con una masa documental de 973 vols., o también el fondo "Real Audiencia", con 3.252 vols., que da cuenta de los expedientes vistos por este órgano colonial entre 1563 y 1817 (DIBAM 2009). Se mencionan estos fondos por su gran cantidad de volúmenes que comprenden los siglos XVI y XVII, precisamente. Si se calcula un promedio de 400 fojas por volumen, solo en este corpus documental tenemos un número aproximado de 5.536.400 fojas; y si a ello agregamos una proporción extremadamente optimista de 20 minutos para la transcripción de cada cara del folio, esto es, 3 folios por hora, la inversión de tiempo ideal ya se escapa de los márgenes de cualquier proyecto o parámetro de vida individual[13]. A ello ni siquiera se le añade la preparación de una edición crítica o anotada, solo el transcribir y digitar el respectivo texto.

Con la finalidad de aprovechar los avances que ha experimentado el campo del reconocimiento óptico de caracteres, particularmente en lo que se refiere a los manuscritos, se ha propuesto, en lo fundamental por parte de ingenieros informáticos y modeladores matemáticos, un conjunto de herramientas que podríamos situar dentro del campo de la paleografía digital, que destacan, en este sentido, dos aspectos relevantes en el proceso de producción del manuscrito: la presencia de patrones redundantes y de sus atributos visibles. En relación con ello, Wolf et al. (2011: 3545) proponen considerar las siguientes herramientas de análisis:

> (1) a handwriting matching tool that is used to join together fragments of the same scribe, and (2) a paleographic classification tool that matches a given document to a large set of paleographic samples. Both tools are carefully designed not only to provide a high level of accuracy, but also to provide a clean and conciese justification of the inferred results. This last requirement engender challenges, such as sparsity of the representation, for which existing solutions are inappropriate for document analysis[14].

[13] Solo por jugar con los números, habría que decir que se requeriría de un total de 1.845.467 horas, esto es, 76.894 días, o mejor, 211 años como mínimo, solo para transcribir estos volúmenes, sin contar con los demás existentes.

[14] Y continúa señalando, en este mismo sentido: "Paleography, the study of old handwriting, is the discipline that engages in the decipherment of ancient texts. In addition, it tries to ascertain when and where a given manuscript was written, and –if possible– by

El empleo de estas herramientas metodológicas permitiría, por tanto, caracterizar al autor de un texto manuscrito, pues, como propone Siddiqi y Vincent (2010), a través de la combinación de rasgos globales y locales, es posible generar un libro de códigos universal que permita la determinación de puntos del contorno de las grafías, de manera individual, y así poder identificarlas en el contexto del manuscrito, tal como lo grafican en las figuras expuestas a continuación:

División de la escritura

(a) división regular; (b) deslizamiento horizontal; y (c) división adaptativa

whom. Paleographers bring many skills and tools to bear on these questions, in what is often a complicated and laborious task, requiring reference to paleographic, linguistic and archaeological data. Because it is difficult to quantify the degree of certainty in the final readings and assessments, experts have begun to develop computerbased methods for paleographic research. So far, such methods have only been applied to small cases due to the high degree of labor involved. Moreover, these efforts have focused almost exclusively on scribal identity, and tend to use the computer as a "black box" that receives images of manuscripts and replies with a classification of the handwriting, which scholars may be reluctant to accept".

Ventanas posicionadas sobre un texto y su respectivo contorno de imagen y la contribución de una de las ventanas a la distribución local de recorrido

e

Direction (c)	Number of Pixels in Direction c	Percentage of Contour Length	Interval	Bin to Increment
5	4	9%	1	5,1
7	2	5%	1	7,1
6	36	86%	9	6,9
	42	100%		

f

El apoyo de herramientas de análisis informático de esta naturaleza permitiría una mayor certeza en la lectura de manuscritos y un menor grado de inseguridad en la resolución de dudas grafemáticas.

Nuevamente, la interacción con otras ciencias puede ayudar no solo a resolver problemas metodológicos en el trabajo filológico, sino que también permite una comprensión más holística del proceso escriturario que no solo es una cuestión gráfica, sino que involucra a toda la estructura humana del que, pluma en mano, ha dejado en aquellos folios toda una experiencia vital que hasta el día de hoy nos acompaña y nos entrega parte de nuestra identidad como pueblos.

Las actuales tendencias en Humanidades indican que se avanza firmemente hacia un aprovechamiento sostenido de la tecnología y la era digital. No solo en lo relacionado con la información de código abierto[15],

[15] La iniciativa de código abierto (open source initiative) surge como un proyecto que busca entregar programas liberados de pago, con alta calidad y efectividad, a toda la comunidad, sin restricciones, para lo cual debe se debe cumplir con algunos términos de referencia (distribución libre, acceso al código fuente, licencia de trabajos derivados, integridad de los autores del código fuente, no discriminación contra personas o grupos, no discriminación en función de la finalidad perseguida, distribución de licencia, licencia no específica al producto, la licencia no debe restringir otro programa, la licencia debe ser tecnológicamente neutral), cuya definición puede encontrarse en http://opensource.org.

sino que en el rescate, conservación y tratamiento digital de los documentos patrimoniales de cada nación, particularmente los que corren mayor riesgo de deterioro, aun cuando las especiales condiciones de custodia los preserven por más tiempo. Esto último, sin duda, tiene otro inconveniente, la conservación resguardada implica, a su vez, una disminución en la capacidad de transmisión y difusión del conocimiento y desarrollo cultural de los pueblos. Por ello, los desarrollos presentes de la codicología y de la paleografía se nutren cada vez más de la informática, el modelamiento matemático y la web. Fischer y Sahle (2010: xi) señalan, al presentar una serie de estudios sobre las disciplinas en comento, que: "These days, the starting point for manuscript research projects is often digital reproduction. Digital facsimiles convey a great number of the original features and characteristics and can be easily provided and shared. Carried out on a large scale, digital reproduction is the cheapest way of making entire collections of manuscripts accesible". Y más adelante (2010: xii-xiii), se refieren a la relación entre la catalogación y la digitalización de manuscritos en los siguientes términos:

> Converting knowledge from analogue to digital is not just a technical issue of how to do this as quickly and effectively as possible. Rather, it sets its own methodological agenda. One of the changes ocurring in codicology, in comparison to traditional print cataloguing, is the relationship between the codex (or its visual digital surrogate) and its description /.../ and therefore the description itself. Moreover, there is a tendency towards open forms of collaboration in providing information and access to the manuscript heritage. /.../ The mere existence of digital reproductions and online catalogues prompts us to consider connecting catalogues by bringing all the available documentation together in comprehensive portals.

Es más, los autores terminan señalando que "In an ideal digital world, all knowledge of the handwritten tradition would be collected, connected, enriched and accesible from a single point of entrance" (p. xiii).

En este contexto, entonces, a lo que el CorDECh aspira no es solo a constituirse como una colección documental de textos transcritos con cierta rigurosidad, sino que también a abrir caminos de comprensión de un pasado rico en tradiciones, formas de apreciar la sociedad y la relación con el entorno, diverso en sus relaciones interculturales y, por qué no decirlo, una vía de acceso para fortalecer el impulso de la filología chilena, de la mano de la docencia y de la extensión cultural, a veces poco consideradas a las luces de la investigación científica y del vértigo de la periodicidad en la difusión del conocimiento.

REFERENCIAS BIBLIOGRÁFICAS

Alvar, Manuel. 1992. La investigación del español de América: Proyectos inmediatos, ponencia plenaria, en *Actas del Congreso de la Lengua Española*, Sevilla, disponible en http://cvc.cervantes.es/obref/congresos/sevilla/plenarias/ponenc_alvar.htm [Consulta 01/07/2007].

Anónimo. s.XVII. *La guerra de Chile*, edición crítica de Mario Ferreccio Podestá y Raïssa Kordić Riquelme, estudios preliminares de José Miguel Barros Franco y Osvaldo Silva Galdames. Universidad de Chile, Facultad de Filosofía y Humanidades, Departamento de Ciencias Históricas, Seminario de Filología Hispánica; 1996 (BACh, 4), Santiago: Arte Gráfico.

Arellano, Ignacio y J.A. Rodríguez Garrido (eds.). 1999. *Edición y anotación de textos coloniales hispanoamericanos*. Universidad de Navarra, Madrid: Iberoamericana/Vervuert.

Arellano, Ignacio y Fermín del Pino (eds.). 2004. *Lecturas y ediciones de crónicas de Indias. Una propuesta interdisciplinaria*. Universidad de Navarra, Madrid: Iberoamericana/Vervuert.

Arias de Saavedra, Diego. s. XVII. *Purén indómito*, prólogo y edición crítica de Mario Ferreccio Podestá, estudio preliminar de Mario Rodríguez Fernández. Biblioteca Nacional – Universidad de Concepción, Seminario de Filología Hispánica, 1984 (BACh, 1), Concepción: La Noria.

Bertolami, Roman y Horst Bunke. 2008. Hidden Markov model-based ensemble methods for offline handwritten text line recognition. *Pattern Recognition* Nº 41, Cambridge:Elsevier; pp.3452-3460. Disponible en http://www.sciencedirect.com/science/article/pii/S0031320308001349

Brink, A.A.; J. Smit; M. L. Bulacu y L. R. B. Schomaker. 2012. Writer identification using directional ink-trace width measurements. *Pattern Recognition*, Cambridge:Elsevier; pp. 162-171. Disponible en http://www.sciencedirect.com/science/article/pii/S0031320311002810

Casassas Canals, Xavier. 2011. La edición de manuscritos romances aljamiado-árabes: una propuesta de edición por capas. *Mélanges offerts au Prof. Mikel de Epalza*, Abdeljelil Temimi (ed.), Tünez, pp. 177-196. Disponible en http://www.casassas.net/articles/edicion_manuscritos_%20romances_aljamiado_arabes_mudejares_moriscos.pdf) [Consulta 03/05/2012].

Company, Concepción. 1994. *Documentos lingüísticos de la Nueva España (Altiplano Central)*. Instituto de Investigaciones Filológicas, Centro de Lingüística Hispánica, México D.F.: UNAM.

Contreras Seitz, Manuel Eduardo. 2012. O Corpus Diacrónico del Español de Chile (CORDECH). Aspectos teóricos y metodológicos. *Crítica Textual e Edição de Textos. Interagindo com outras ciências*. Curitiba: Prismas, pp. 193-217.

——— 2009a. Hacia la constitución de un corpus diacrónico del español de Chile. *RLA. Revista de Lingüística Aplicada* 47(2), II Sem., Concepción:Universidad de Concepción; pp.111-134. Disponible en http://www.scielo.cl/pdf/rla/v47n2/ART_07.pdf

——— 2009b. Bases para un estudio de la sintaxis histórica del español de Chile. *Estudios Filológicos* 44. Valdivia: Universidad Austral de Chile, pp. 27-51. Disponible en http://www.scielo.cl/pdf/efilolo/n44/art02.pdf

——— 2008. Cómo editar textos coloniales. *Estudios Filológicos* 43. Valdivia:Universidad Austral de Chile, pp. 63-82. Disponible en http://www.scielo.cl/pdf/efilolo/n43/art05.pdf

——— 2005. *Ilegibilidad y Cotidianeidad. Paleografía y colección diplomática de documentos chilenos del período colonial (1548-1651)*. Programa de Estudios y

Documentación en Ciencias Humanas, Colección Fuentes Documentales, Osorno: Universidad de Los Lagos.

——————— 2004. *El Español de Chile en el Período Colonial. Fonética*. Programa de Estudios y Documentación en Ciencias Humanas, Colección Monográficos, Osorno: Universidad de Los Lagos.

———————— 2003. ¿Por qué escriben las mujeres? Documentos femeninos del período chileno colonial. *Estudios Filológicos* 38. Valdivia: Universidad Austral de Chile; pp. 61-92. Disponible en http://www.scielo.cl/scielo.php?pid=S0071-17132003003800005&script=sci_arttext

——————— 2001. Apuntes filológicos sobre la *Nueva Obra y Breve en Prosa y Metro sobre la Muerte del Ilustre Adelantado Don Diego de Almagro*. *Estudios Filológicos* 36. Valdivia: Universidad Austral de Chile, pp. 99-115. Disponible en http://www.scielo.cl/scielo.php?pid=S0071-17132001003600007&script=sci_arttext

CORDIAM. 2011. *Corpus Diacrónico y Diatópico del Español de América*, documento de trabajo, Virginia Bertolotti y Concepción Company (coord. acad.), 14 págs.

Dirección de Bibliotecas, Archivos y Museos (DIBAM). 2009. *Guía de fondos del Archivo Nacional Histórico. Instituciones coloniales y republicanas*, Centro de Investigaciones 'Diego Barros Arana', Santiago: Versión.

Fernández Alcaide, Marta. 2009. *Cartas de particulares en Indias del siglo XVI. Edición y estudio discursivo*, Madrid: Iberoamericana/Vervuert. Disponible en vista parcial en http://books.google.cl/books?id=C5nGf8o-rQ4C&pg=PA371&lpg=PA371&dq=Una+introducci%C3%B3n+filol%C3%B3gica+a+la+documentaci%C3%B3n+del+Archivo+General+de+Indias&source=bl&ots=Kv7hwxNKEb&sig=fFqzxdnqqHL1dCh23wO3-agKCvA&sa=X&ei=i7wzUMC9H9S_0QG80oHIDg&ved=0CDwQ6AEwCA#v=onepage&q=Una%20introducci%C3%B3n%20filol%C3%B3gica%20a%20la%20documentaci%C3%B3n%20del%20Archivo%20General%20de%20Indias&f=false

Fischer, Franz y Patrick Sahle. 2010. Introduction. Into the Wide – Into the Deep: Manuscript Research in the Digital Age. *Kodikologie & Paläographie im Digitalen Zeitalter 2 / Codicology & Palaeography in the Digital Age 2*. Franz Fischer, Christiane Fritze y Georg Vogeler (eds.). Schriften des Instituts für Dokumentologie und Editorik 3, Norderstedt:Book on Demand GmbH.

Fontanella de Weinberg, Mª Beatriz (comp.). 1993. *Documentos para la historia lingüística de Hispanoamérica. Siglos XVI a XVIII*, Vol. I, Anejo LIII del Boletín de la Real Academia Española; ALFAL, Comisión de Estudio Histórico del Español de América, Madrid: Espasa-Calpe.

Frago Gracia, Juan Antonio. 1987. Una introducción filológica a la documentación del Archivo General de Indias. *Anuario de Lingüística Hispánica*, 3. Valladolid: Universidad de Valladolid, pp. 67-97.

Guitarte, Guillermo. 1968. Para una historia del español de América basada en documentos: El seseo en el Nuevo Reino de Granada (1550-1650). *Actas de la 5ª Asamblea Interuniversitaria de Filología y Literatura Hispánicas*, Héctor Ciocchini y Josefa Zamudio (coords.). Bahía Blanca: Universidad Nacional del Sur, pp. 158-165.

Ihalainen, Ossi; Merja Kyto y Matti Rissanen. 1987. The Helsinki Corpus of English Texts: Diachronic and dialectal report on work in progress, en Corpus Linguistics and Beyond. Proceedings of the Seventh International Conference on English Language Research on Computerized Corpora, edited by Willem Meijs. Costerus, New Series, Volume 59, Amsterdam: Rodopi; pp. 21-32.

Jara, Alvaro y Sonia Pinto. 1982-1983. *Fuentes para la Historia del Trabajo en el Reino de Chile: legislación 1546-1810*. Santiago: Andrés Bello. Tomo I disponible en http://www.memoriachilena.cl/archivos2/pdfs/MC0038858.pdf; Tomo II disponible en http://www.memoriachilena.cl/archivos2/pdfs/MC0038849.pdf

Jara, Álvaro y Rolando Mellafe. 1996. *Protocolos de los Escribanos de Santiago: primeros fragmentos, 1559 y 1564-1566*. 2 vols., DIBAM, Centro de Investigaciones 'Diego Barros Arana'. Santiago: Andros. Tomo I disponible en http://www.memoriachilena.cl/archivos2/pdfs/MC0038859.pdf; tomo II en http://www.memoriachilena.cl/archivos2/pdfs/MC0038871.pdf

Jucker, Andreas y Andreas Jacob. 1995. The historical perspective in pragmatics. *Historical Pragmatics*, Andreas Jucker (ed.). Amsterdam: John Benjamins; pp. 3-33.

Kordić Riquelme, Raïssa. 2005. *Testamentos coloniales chilenos*, prólogo y edición crítica, estudio preliminar de Cedomil Goić (BACh, 6). Universidad de Navarra, Centro de Estudios Indianos. Madrid: Iberoamericana/Vervuert.

——————— 2006. La crítica textual hispanoamericana: algunas especificaciones metodológicas. *Onomázein* Nº 13/1. Pontificia Universidad Católica de Chile, Instituto de Letras. Santiago: Andros; pp. 191-202. Disponible en http://www.onomazein.net/13/13_13.pdf

——————— 2007. Chile colonial. Filología e historia: Las cartas de Sor Dolores Peña y Lillo. en *Suplemento Artes y Letras*, El Mercurio. Santiago, domingo 3 de junio.

——————— 2008. *Epistolario de Sor Dolores Peña y Lillo (Chile, 1763-1769)*, (BACh, 7). Universidad de Navarra, Centro de Estudios Indianos. Madrid: Iberoamericana/Vervuert.

——————— 2010. Falsas variantes en las ediciones de textos. *Boletín de Filología*, Tomo XLV, Número 2, Universidad de Chile. Santiago: LOM; pp. 279-286. Disponible en http://www.scielo.cl/pdf/bfilol/v45n2/art10.pdf

Leonard, Irving. 1966. *Crónica y relación copiosa y verdadera de los reynos de Chile*, (edición facsimilar). Santiago: Fondo Histórico y Bibliográfico José Toribio Medina.

Marrero-Fente, Raúl (ed.). 2004. *Perspectivas transatlánticas. Estudios coloniales hispanoamericanos*. Madrid: Verbum.

Medina Zavala, José Toribio. 1888-1902. *Colección de Documentos Inéditos para la Historia de Chile, desde el viaje de Magallanes hasta la Batalla de Maipú (1518-1818)*, 30 volúmenes. Santiago: Imprenta Ercilla (vols. 1-5). Imprenta Barcelona (vols. 6-7). Imprenta Elzeviriana (vols. 8-30). Disponibles en http://archive.org/search.php?query=colecci%C3%B3n%20de%20documentos%20in%C3%A9ditos%20para%20la%20historia%20de%20Chile

Melis, Chantal; Agustín Rivero Franyutti y Beatriz Arias Álvarez. 2008. D*ocumentos lingüísticos de la Nueva España (Golfo de México)*. Instituto de Investigaciones Filológicas, Centro de Lingüística Hispánica. México D.F.: UNAM.

Morales, Leonidas. 2001. *La escritura de al lado. Géneros referenciales*. Serie Ensayo. Santiago: Cuarto Propio. Disponible en http://www.memoriachilena.cl/archivos2/pdfs/MC0029865.pdf

Núñez Contreras, Luis. 1994. *Manual de paleografía*. Madrid: Cátedra.

Núñez de Pineda y Francisco Bascuñán. 1607-1682. *Cautiverio feliz*. 2 tomos, edición crítica de Mario Ferreccio Podestá y Raïssa Kordić Riquelme, estudio preliminar de Cedomil Goić. Universidad de Chile, Seminario de Filología Hispánica, 2001 (BACh, 5). Santiago: RIL Editores.

Oña, Pedro de. 1629. *El Ignacio de Cantabria, poema sacro*. Edición crítica a cargo de Mario Ferreccio Podestá, Gloria Muñoz Rigolett y Mario Rodríguez Fernández. Biblioteca Nacional – Universidad de Concepción, 1992 (BACh, 3). Santiago: Universitaria.

Retamal Ávila, Julio. 2000. *Testamentos de indios en Chile colonial (1564-1801)*. UNAB. Santiago: RIL Editores. Disponible en vista previa en http://books.google.cl/books?id=X9Xn9mwhL6wC&printsec=frontcover&hl=es#v=onepage&q&f=false

Rivarola, José Luis (2009). *Documentos lingüísticos del Perú. Siglos XVI y XVII. Edición y comentario*. Anejos de la Revista de Filología Española Nº 103. Madrid: CSIC. Disponible

en vista previa en http://books.google.cl/books?id=ZTLpkBLVWTQC&printsec=frontcover&hl=es#v=onepage&q&f=false

Rojas Carrasco, Guillermo. 1940. *Filología Chilena. Guía bibliográfica y crítica*. Ediciones de la Universidad de Chile. Santiago: Imp. y Lit. Universo S. A.

Rojas Mayer, Elena M. (comp. y ed.). 2000. *Documentos para la historia lingüística de Hispanoamérica. Siglos XVI a XVIII*. Vol. II, Anejo LVIII del Boletín de la Real Academia Española. ALFAL, Comisión de Estudio Histórico del Español de América.Madrid: Espasa-Calpe.

——————(comp. y ed.). 2008a. *Documentos para la historia lingüística de Hispanoamérica. Siglos XVI a XVIII*. Vol. III, Anejo LX del Boletín de la Real Academia Española. ALFAL, Comisión de Estudio Histórico del Español de América, Madrid: Espasa-Calpe.

——————(comp. y ed.). 2008b. *Documentos para la historia lingüística de Hispanoamérica. Siglos XVI a XVIII*. Vol. IV, Anejo LXI del Boletín de la Real Academia Española. ALFAL, Comisión de Estudio Histórico del Español de América, Madrid: Espasa-Calpe.

Servei de Tecnologia Linguistica (STEL). 2007. *Introducción a la lingüística del corpus*. http://www.ub.edu/stel/esp_support.htm [consulta: julio de 2007].

Siddiqi, Imran y Nicole Vincent. 2010. Text independent writer recognition using redundant writing patterns with contour-based orientation and curvature features. *Pattern Recognition* 43. Cambridge: Elsevier, pp. 3853-3865. Disponible en http://www.sciencedirect.com/science/article/pii/S0031320310002438

Stutzmann, Dominique. 2010. Paléographie statistique pour décrire, identifier, dater... Normaliser pour coopérer et aller plus loin? *Kodikologie & Paläographie im Digitalen Zeitalter 2 / Codicology & Palaegraphy in the Digital Age 2*. Schriften des Instituts für Dokumentologie und Editorik 3, Franz Fischer, Christiane Fritze y Georg Vogeler (eds.), pp. 247-277. Norderstedt: BoD GmbH.

Suárez, Úrsula. 1666-1749. *Relación autobiográfica*. Prológo y edición crítica de Mario Ferreccio Podestá, estudio preliminar de Armando de Ramón. Biblioteca Nacional, 1984 (BACh, 2). Santiago: Universitaria. Disponible en http://www.memoriachilena.cl/archivos2/pdfs/MC0013671.pdf

Vergara, Sergio. 1987. *Cartas de mujeres en Chile. 1630-1885*, estudio, selección y notas. Santiago: Andrés Bello. Disponible en vista previa en http://books.google.cl/books?id=yJf7uUVlvroC&printsec=frontcover&hl=es#v=onepage&q&f=false

Vivar, Jerónimo de. 1558. *Crónica y relación copiosa y verdadera de los Reynos de Chile hecha por Gerónimo de Bibar natural de Burgos*. Transcripción paleográfica del Prof. Irving Leonard según el manuscrito original, propiedad de The Newberry Library. Introducción de Guillermo Feliú Cruz. Edición facsimilar y a plana del Fondo Histórico y Bibliográfico José Toribio Medina, 2 tomos, 1966. Santiago: Instituto Geográfico Militar. Disponible en http://www.memoriachilena.cl/archivos2/pdfs/MC0008847.pdf

Wolf, Lior; Liza Potikha y Nachum Dershowitz. 2011. Computerized paleography: tools for historical manuscripts. *Proceedings of International Conference on Image Processing (ICIP 2011)*. Bruselas, Bélgica, pp. 3545-3548. Disponible en http://www.genizah.org/professionalPapers/ComputerizedPaleography.pdf