

República de Cuba



Facultad de Humanidades

Departamento de Letras

Trabajo de diploma en opción al título de Licenciado en Letras

Título: Corpus de documentos notariales del Departamento oriental de Cuba

(S. XVII-XIX).

Autora: Danay Liset González Ferrás.

Tutoras: Dr.C. Irina Bidot Martínez

Dr.C. Celia María Pérez Marqués

Santiago de Cuba

2018

*A mi abuela Chelo, por su confianza, fe y amor, aunque no
estés presentes sigues en mi corazón.*

Agradecimientos

A la Revolución Cubana y al MININT, por darme la oportunidad de cumplir mi sueño de ser una profesional.

A mi mamá, por su amor, preocupación y desvelo por que hiciera todo bien. Por su ejemplo de mujer.

A mi papá, por estar siempre a mi lado y por el cariño que esconde detrás de su armadura.

A mi tía, mi segunda madre, por su complicidad, amor, dedicación y por estar siempre junto a mí, a pesar de la distancia.

A mi hermano, por tantas noches de desvelos y por nunca dejar de creer en mí.

A mi abuela, por todo su apoyo y cariño.

A mis mejores amigas: Debo, Elita, Katy, Rosy y Yisi por hacer de estos cinco años en Santiago una experiencia única.

A mis compañeros de aula Julio y Daniela por las largas horas de estudio y por compartir su arte.

A mi tutora Irina por su paciencia y dedicación, por hacer que me enamorara de cada una de estas páginas.

A mi tutora Celia por su apoyo y sabiduría.

A todos los profesores del departamento de Letras por compartir sus conocimientos y por contribuir a que sea una mujer de bien.

A mis amigos de Las Tunas, Andrea, Náyade, Yeli y Danel por nunca olvidarse de mí, a pesar de la distancia.

A los arquitectos del cuarto 138, Tin, Osbelito, Ale, Chino, Fermín, por acogerme, por tantas risas compartidas.

A los mecánicos del cuarto 102, Mario, Robert, Chino, Michel, Raúl y Isael por hacer de mis últimos días en Santiago inolvidables.

A todos muchas gracias.

Resumen

El trabajo “Corpus de documentos notariales del Departamento oriental de Cuba (S.XVII-XIX)” realiza una propuesta para la conformación de un corpus de un total de 77 documentos notariales, ubicados en los archivos históricos provinciales de Holguín, Santiago de Cuba y Guantánamo. Aporta el rescate de los documentos antes mencionados con una triple presentación: paleográfica, facsimilar y guía comentada, organizados según la tipología textual y siglos. Presenta una nueva guía para realizar las transcripciones de los documentos según los criterios de edición de la red internacional CHARTA. Dichos resultados se insertan en el Proyecto Internacional de investigación “Las ciencias sociales, humanísticas y la arquitectura frente a los retos del desarrollo social en Santiago de Cuba y la Región Este del país. Potenciando el Patrimonio”.

Abstrac

The research “Corpus of the notarial documents of the eastern Departament of Cuba (S. XVII-XIX)” makes a proposal for the creation of a corpus of a total of 77 documnets, located in the historical archives of Holguín, Santiago de Cuba y Guantánamo. It provides the rescue of the afore mentioned documents with a triple presentation: paleografhic, facsímile and annotated guide, organized according to the editing criteria of the international CHARTA network. The seresults are inserted in the international Research Project “Social sciences, humanities and architecture facing the challengs of social develoment in Santiago de Cuba, and he Eastern Region of the country. Enhancing the heritage”.

Índice

	PAG
Introducción	----- 1
Capítulo I “Fundamentos teóricos para la elaboración de un corpus lingüístico de documentos notariales”	----- 8
1.1 Lingüística de corpus	----- 8
1.2 Los corpus lingüísticos, características generales	----- 12
1.3 Característica formales y estilísticas de los documentos notariales	----- 24
Capítulo II “Características del corpus de documentos notariales del oriente de Cuba (S. XVII- XIX)”	----- 32
2.1 Referente metodológico para la elaboración del corpus: la red CHARTA	----- 32
2.2 Procedimientos para la conformación científica del corpus de documentos notariales del oriente de Cuba	----- 36
2.2.1 Caracterización de la muestra y adecuaciones de los criterios de edición de la Red CHARTA	----- 36
2.2.2 Estructura del corpus y modos de búsqueda	----- 43
Conclusiones	----- 45
Recomendaciones	----- 47
Bibliografía	
Índice de Anexos	

Introducción

El estudio de la modalidad cubana de la lengua española es el objeto de investigación fundamental dentro de la ciencia lingüística en Cuba, el cual ha tomado gran importancia y ha sido motivo de interés para los lingüistas, profesores y estudiantes universitarios en nuestro país, principalmente a partir del triunfo revolucionario.

En la época actual, muchos son los investigadores formados en las universidades del país y el extranjero que integrados a los cuerpos docentes y de centros de investigación han contribuido con sus resultados a los avances de esta ciencia; de tal manera que puede afirmarse que los estudios lingüísticos en Cuba van a la par del desarrollo internacional.

Estas investigaciones, al ahondar en múltiples aspectos de la modalidad cubana del español y tomando como referente los distintos niveles de la lengua, ofrecen una caracterización del español hablado en Cuba. Pero, por otra parte, también se encuentran aquellos estudios que van hacia el estudio de su historia, los cuales aún resultan menores con relación con los de la lengua en uso.

A esta ardua tarea de corte historicista dentro de los estudios lingüísticos se ha unido la carrera de Letras de la Universidad de Oriente, la cual, a partir de la incorporación de profesores y estudiantes al Proyecto VLIR¹ “Las ciencias sociales y las humanísticas frente a los retos contemporáneos del desarrollo local en Santiago de Cuba y la región este de Cuba. Potenciando el patrimonio”, específicamente en el tópico “Fuentes documentales en la región oriental de Cuba y su contribución al rescate, conservación y promoción del Patrimonio”, se ha

¹Proyecto desarrollado, a partir de 2013, de conjunto entre la Universidad de Oriente y un grupo de universidades belgas, en el cual se han tenido como resultado un conjunto de artículos científicos, trabajos de curso y de diploma elaborados por profesores y estudiantes miembros de dicho proyecto.

propuesto determinar por siglos los rasgos tipificadores del español de la isla en el antiguo Departamento oriental.

La búsqueda y revisión de documentos para dicha tarea se extiende a las instituciones que poseen información fiable para elaborar estos estudios con la calidad requerida (Cfr. Bidot y Guerra, 2015)².

El proyecto ha propiciado la defensa de los siguientes trabajos de diploma de la carrera de Letras:

- *Caracterización lingüística de documentos manuscritos de Santiago de Cuba siglo XVII. (2014)*, de Odette Acevedo Medina.
- *Estudio lingüístico de documentos notariales de Jiguaní en la segunda mitad del siglo XIX. (2015)*, de Sucel Paula Segura Oduardo.
- *Análisis lingüístico-formal de cartas de asunto religioso de la segunda mitad del siglo XVIII en Santiago de Cuba. (2016)*, de Dallamy Liliana Escalona Zambrano.
- *Caracterización lingüístico- formal de cartas de venta del siglo XVIII en Holguín. (2016)*, de Geisy Labrada Hernández.
- *Caracterización lingüístico-formal de textos manuscritos de Baracoa en la segunda mitad del siglo XIX. (2016)*, de Rebeca Durán García.
- *Caracterización lingüística de documentos manuscritos de Guantánamo en el siglo XIX. (2017)*, de Ailen Robinson Maure.
- *Conectores argumentativos en testamentos del siglo XIX de Holguín y Santiago de Cuba. (2017)*, de Geilys Zaday Pérez Leyva.
- *Caracterización lingüística de documentos manuscritos en la segunda mitad del siglo XVIII de Santiago de Cuba. (2017)*, de Karen Safonst Caballero.

² Igualmente se debe destacar el interés suscitado por la Dr. C. Marlen Domínguez de la Universidad de la Habana, quien junto a profesores de Historia de La lengua Española en el país, en el 2012 creó un grupo de investigación a través del cual se pretende la búsqueda del surgimiento y evolución de los elementos tipificadores del español hablado en Cuba.

Estos trabajos han ayudado al rescate y transcripción paleográfica de documentos manuscritos, los cuales han sido analizados a partir de comentarios gráfico-fonético, morfosintáctico y léxico- semántico, y según las características de las tipologías textuales a las que corresponden, en su mayoría documentos notariales con sus especificidades tipológicas que han sido elaborados por individuos representativos de estratos sociales con una determinada instrucción, capaces de llevar a cabo esa tarea.

A la información aportada por estas tesis se incorporan los resultados de los trabajos de curso de Historia de la Lengua, que igualmente han ofrecido información.

En la tesis de Ailen Robinson Maure, se recomienda que:

A partir del resultado, este puede unirse a los trabajos de diploma anteriormente realizados en Baracoa, Santiago de Cuba, Jiguaní, Holguín y al que se hará en Camagüey para confeccionar un corpus lingüístico de la región oriental en el periodo colonial. (Robinson, 2017:70)

De esta Recomendación surge el presente Trabajo de Diploma que tiene como

problema de investigación:

Inexistencia de un corpus lingüístico de documentos notariales, representativos del Departamento oriental de Cuba en los siglos XVII-XIX que contribuya a los estudios de la modalidad cubana de la lengua española en el periodo.

Para dar solución a esta necesidad se trazó el siguiente **objetivo:**

Elaborar un corpus lingüístico con documentos notariales representativos del Departamento oriental entre los siglos XVII – XIX, que contribuya a los estudios de la modalidad cubana de la lengua en el periodo.

El cumplimiento de este objetivo permitirá dar solución al problema investigativo declarado, sobre la base de la siguiente **hipótesis**:

Si se organizan los documentos notariales representativos del Departamento oriental, entre los siglos XVII-XIX, teniendo en cuenta su distribución por tipologías textuales y siglos, se puede conformar un corpus lingüístico contentivo de esta información que contribuya al estudio de características de la modalidad cubana de la lengua.

Para la realización del corpus creado a partir de esta investigación se escogieron de los trabajos de diplomas, mencionados anteriormente, documentos que se enmarcan en lo que se ha dado en llamar el “universo de los discursos jurídico-administrativos” (Oesterreicher, 2004: 738, *Apud* Ramírez, 2011: 3).

Los textos jurídicos son aquellos producidos en el proceso de administración de justicia y en su redacción es utilizada la norma culta de la lengua, y los notariales se convierten en instrumentos públicos una vez emitidos o autorizados por un notario, por lo que cuentan de una oficialidad, que garantiza su fiabilidad para los estudios de la lengua.

Ello hace que los textos tomados para la conformación del corpus se caractericen por su formalidad, aspecto que tendrá que ser tomado en cuenta al interpretar los datos lingüísticos que se puedan extraer de él.

Para la realización de este corpus se escogieron específicamente los protocolos notariales y, dentro de estos, las escrituras públicas con características que se corresponden con un registro formal. Estos se encuentran en los trabajos de diploma de Acevedo (2014), Labrada (2016), Robinson (2017), Pérez (2017) y Safonst (2017).

También se seleccionaron documentos analizados en los trabajos de curso de Historia de la lengua por estudiantes de tercer año de los cursos 2014- 2015 y 2015- 2016³.

Todos han servido de muestra para la realización del corpus. Precisamente aquí es donde radica la **novedad** de la investigación, porque hasta donde se tiene investigado, es el primer corpus que recogerá muestras de documentación del período colonial en el antiguo Departamento Oriental.

La realización de este corpus se adscribe a los presupuestos teórico-metodológicos de la RED CHARTA (Corpus Hispánico y Americano en la red: Textos Antiguos), con lo cual se evidencia la **actualidad** de la investigación.

La RED CHARTA se concibe como un proyecto global para la edición y análisis lingüístico de textos archivísticos en español de los siglos XII al XIX. Contiene documentos de Europa, América y Asia, entendidos estos en sentido amplio, pues no solo acoge piezas oficiales de la cancillería, la administración civil, la de justicia, la Inquisición o contratos de compraventa, sino también cartas particulares y aún billetes y notas sueltas.

Los documentos se presentan en edición digital fiable, comprobable y se pueden citar directamente tanto en estudios científicos como en trabajos divulgativos, o simplemente para satisfacer la curiosidad por el pasado, incluso local y familiar. (Corpus CHARTA. Red Internacional).

Los métodos empleados en la realización de este trabajo son generales y específicos de la ciencia lingüística. Dentro de los primeros se señala la

³En ambos cursos se trabajaron diferentes tipologías textuales ubicadas en el Archivo Histórico Provincial de Santiago de Cuba (AHPSC) en los siglos XVIII y XIX. No se tuvieron en cuenta los trabajos del curso 2016-2017, en los que se transcribió y analizó un pleito legal del siglo XVIII resguardado en el Archivo General de Indias (AGI), ni los del actual, 2017-2018, en los que se trabajó con una variedad de documentos del siglo XIX situados en el AHPSC y el Archivo Histórico Nacional (AHN), pues en ambos casos se repetían las tipologías textuales ya seleccionadas y se incorporaban textos de otros archivos que si bien resguardan documentos del Departamento oriental, no se encuentran ubicados en este territorio.

observación científica, que contribuyó a la determinación de la muestra llevada al corpus y el inductivo-deductivo y el analítico-sintético, en función de determinar los referentes fundamentales que le sirvieron de base y sustento.

La investigación se sustenta en los presupuestos de la lingüística de corpus, a través de la cual se podrá conformar el corpus.

Las técnicas utilizadas fueron el análisis bibliográfico, que aportó el sustento teórico y metodológico para la investigación, y el trabajo en archivos, que sirvió de base para completar la muestra llevada al corpus, así como las herramientas que aporta la lingüística de corpus para la elaboración del principal aporte de la tesis.

El trabajo está estructurado en introducción, dos capítulos, conclusiones, recomendaciones, bibliografía y anexos.

En el capítulo I, **Fundamentos teóricos para la elaboración de un corpus lingüístico de documentos notariales**, se ofrecen las características de los corpus lingüísticos, así como los aspectos formales de los documentos notariales que agrupan las diferentes tipologías textuales que conforman el corpus.

En el capítulo II, **Características del corpus de documentos notariales del oriente de Cuba (S. XVII- XIX)**, se hace referencia a los presupuestos teórico-metodológicos de la Red internacional CHARTA, los cuales han servido de base, con determinadas adecuaciones que también se exponen, para la conformación del corpus que se presenta. Por otra parte, se explican las especificidades de la muestra y la manera en que se ha estructurado el corpus.

Y se ofrecen los elementos fundamentales que tiene el corpus que se presenta como resultado.

Las conclusiones constituyen la síntesis de los principales resultados, las recomendaciones se vuelven pilares para la continuación del presente estudio, y la bibliografía se convierte en una guía de estudio en la cual se exponen los textos consultados durante todo el proceso investigativo. Los anexos resultan de vital importancia ya que complementan la investigación y recogen el corpus lingüístico digital.

Como aportes de esta investigación se encuentran:

- Presentación de una concepción teórico-metodológica para la elaboración de un corpus lingüístico de documentos notariales a partir de los presupuestos ofrecidos por la red CHARTA, con las adecuaciones correspondientes al contexto en que se enmarcan las investigaciones realizadas en la Universidad de Oriente.
- Corpus lingüístico de documentos notariales representativos del Departamento oriental entre los siglos XVII – XIX.

Capítulo I “Fundamentos teóricos para la elaboración de un corpus lingüístico de documentos notariales”.

En el capítulo se recogen los aspectos generales con respecto a las principales características de los corpus lingüísticos. Posteriormente se exponen las peculiaridades formales de los documentos notariales que conforman la muestra de este trabajo.

1.1 Lingüística de corpus

La ciencia lingüística en la actualidad ha tomado gran auge en nuestra región, debido al creciente interés por el estudio del uso del español en su modalidad hispanoamericana. En este contexto, la lingüística se auxilia de otras ciencias como la sociolingüística, la cual parte del desarrollo de técnicas que permiten la recolección y almacenamiento de grandes cantidades de producciones lingüísticas reales, que permiten la observación en profundidad del comportamiento lingüístico de unos hablantes en concreto (Ávila, Lasarte y Villena, 2008: 30-31).

Para este propósito la sociolingüística se ha beneficiado del desarrollo de la lingüística de corpus (LC), entendida como el estudio empírico de la lengua a partir de datos que proporciona el análisis de ejemplos reales de producciones lingüísticas (orales o escritas) almacenadas en un ordenador (Pérez y Moreno, s.f: 71).

La lingüística de corpus fundamenta su nacimiento en el proceso de acumulación de datos lingüísticos. Definir si la lingüística de corpus es una metodología o una teoría, ha sido objeto de debate desde su aparición en la década de los años 60 del pasado siglo.

Existe un gran número de autores que abordan este asunto, tal es el caso de Giovanni Parodi en su artículo “Lingüística de corpus: una introducción al ámbito” (2008), quien menciona una serie de autores que ofrecen varias consideraciones acerca de la definición de lingüística de corpus, tales como: Svartivk, McEnery & Wilson, Tognini-Bonelli⁴ (Parodi, 2008: 95).

La lingüística de corpus puede ser entendida como un fin en sí mismo o como una herramienta de apoyo al lingüista. En los dos casos se ha convertido en herramienta básica para las investigaciones sociolingüísticas. La principal ventaja de esta propuesta de trabajo es el almacenamiento de gran cantidad de textos reales, lo que puede facilitar los estudios de diversas lenguas (Ávila, Lasarte y Villena, 2008: 31-32).

La lingüística de corpus en su versión actual constituye un enfoque metodológico para el estudio de las lenguas, el cual revela oportunidades revolucionarias para la descripción, análisis, y enseñanza de discursos de todo tipo, además de brindar una base empírica para el desarrollo de materiales educativos y metodológicos de diversa índole, así como para la construcción de gramáticas, diccionarios y otros, tanto de discursos generales como especializados, orales y escritos (Parodi, 2008: 95).

También se sostiene que la lingüística de corpus constituye un conjunto o colección de principios metodológicos para estudiar cualquier dominio lingüístico y que se caracteriza por brindar sustento a la investigación de la lengua en uso, a partir de corpus lingüísticos con sustrato en tecnología computacional y programas informáticos *ad hoc* (Ídem).

⁴ Para más información, los datos editoriales de estos autores son: Svartivk, J. (Ed.). 1992. *Directions in corpus linguistics: proceeding of Nobel symposium*. Berlín: Mouton de Gruyter; McEnery, T. & Wilson, A. 1996. *Corpus linguistics*. Edinburg: Edinburg University Press; Kennedy, G. 1998. *An introduction to corpus linguistics*. New York: Longman; Stubbs. M. 1996. *Text and corpus analysis. Computer-assisted studies of language and culture*. Massachusetts: Blackwell; Stubbs, M. 2001. *Words and phrases: corpus studies of lexical semantics*. Oxford: Blackwell; y Tognini- Bonelli, E. 2001. *Corpus linguistics at work*. Amsterdam: Benjamins.

Por lo que en este sentido se puede entender que la lingüística de corpus no es una rama o un área específica de la lingüística, sino que como un método de investigación puede ser empleado en todas las ramas o áreas de la lingüística y desde enfoques teóricos diferentes. Propone una metodología para la investigación de las lenguas y el lenguaje, la cual permite llevar a cabo investigaciones prácticas en contextos reales y regidas estrictamente por presupuestos reguladores (*Ídem*).

Desde este punto de vista estudia una variada información lingüística compilada a través de corpus, dado que la LC apoya la indagación de datos de unidades de sentido y con propósitos comunicativos específicos (*Ibíd.*: 96).

De esta forma queda expuesta la interdisciplinariedad de la LC, al permitir explorar cualquier área o dominio de la lingüística, aporta el estudio de corpus textuales digitales, en su mayoría de tamaño amplio y con soporte en tecnología computacionales variadas.

De aquí que también se relacione con la llamada lingüística computacional, campo científico de carácter interdisciplinar, vinculado a la lingüística y a la informática, cuyo fin fundamental es la elaboración de modelos computacionales que produzcan distintos aspectos del lenguaje humano y que faciliten el tratamiento informatizado de las lenguas (Pérez y Moreno, sf :68).

La Lingüística computacional posee dos campos de acción fundamentales: la *lingüística computacional teórica* y la *lingüística computacional aplicada*. La primera de estas líneas, la lingüística computacional teórica, se centra en tres objetivos complementarios: la elaboración de modelos lingüísticos en términos formales e implementables; la aplicación de dichos modelos a los diferentes niveles de descripción lingüística y la comprobación computacional del modelo y de sus predicciones (*Ídem*).

En cuanto a la lingüística computacional aplicada, es la más relacionada con la tecnología y, específicamente, se centra en el diseño de sistemas informáticos capaces de gestionar, comprender, producir y traducir enunciados orales y escritos en lenguaje natural, para lo que desarrolla aplicaciones informáticas.

Referente al campo de trabajo de la lingüística computacional Pérez y Moreno (*Idem*) explican que se caracteriza por la aplicación de los ordenadores a la investigación lingüística, o al estudio científico del lenguaje, aplicación que es válida para todas las disciplinas lingüísticas que utilizan herramientas informáticas para sus estudios, especialmente en el caso de la lingüística de corpus.

La lingüística de corpus también guarda relación con la terminología, porque un corpus relevante de textos dará las explicaciones y definiciones que aparecen con ese término y las paráfrasis que lo describen durante su negociación. (Martin, s.f. :1029)

Como se ha dicho con anterioridad la lingüística de corpus puede aplicarse a diversas esferas de la lingüística y confiere, en el ámbito de la Lexicografía, nuevas posibilidades que apuntan hacia una nueva concepción del lenguaje, del significado y de la elaboración de diccionarios (*Ibíd.*: 1031).

En esta esfera destaca el Collins COBUILD English Language Dictionary⁵, primer diccionario originado a partir de un corpus, específicamente, el corpus general del inglés más grande que existe con unas 18.3 millones de palabras.

Otro tema importante que es necesario destacar de la LC, es la distinción que se hace entre dos tipos generales de análisis del corpus: cualitativo, en el que se hace una descripción detallada y completa de un fenómeno lingüístico o del comportamiento de una palabra o grupo de palabras, y cuantitativo, en el que se

⁵Esta obra se diseñó a mediados de los 70 y se publicó en 1987 bajo la dirección de John Sinclair. La historia de este proyecto se encuentra en la obra *Looking up: an account of the COBUILD project in lexical computing*.1987, editada por Sinclair.

asignan índices de frecuencia a los fenómenos lingüísticos observados en el corpus y estos pueden servir para construir modelos estadísticos más complejos, que expliquen la evidencia hallada en el texto (Pérez y Moreno, s.f: 80).

Estos dos tipos de análisis son considerados complementarios, al ofrecer, por un lado, el análisis cualitativo, una gran riqueza y precisión en las observaciones realizadas; en tanto los fenómenos poco frecuentes pueden recibir igual atención que los muy frecuentes. Por otro lado, el análisis cuantitativo puede ofrecer información que sea estadísticamente significativa y resultados que pueden considerarse generalizables (McEnery y Wilson, 1996: 63 *Apud Ídem*), por lo que, en las investigaciones en la actualidad, es muy frecuente que se combinen ambos análisis.

La lingüística de corpus separa de manera clara la frontera de la materia propia de los quehaceres del lingüista y promueve la participación activa de los datos que provienen de evidencias empíricas.

Promueve el uso de las herramientas informáticas propias de los avances tecnológicos de la época en la que vivimos y ofrece al lingüista la metodología para las investigaciones donde la lengua real es el objeto fundamental de estudio, y en muchas ocasiones ha posibilitado la creación de corpus lingüísticos que registren dicha lengua.

1.2 Los corpus lingüísticos, características generales.

Definir el concepto de corpus, tal y como se usa hoy en día en el ámbito de la lingüística o lexicografía de corpus, o en la lingüística computacional en general, no es tan sencillo como podría parecer a primera vista.

Según el *Gran Diccionario de Uso del Español Actual*, corpus es el «conjunto de textos, procedentes del lenguaje oral o escrito o de ambos, recopilados de fuentes variadas y *ordenados* y clasificados según determinados criterios, de tal manera que, sobre ese conjunto, es posible realizar estudios e investigaciones lingüísticas o literarias.» (Sánchez, 2001: *Apud* Hrušková, 2008:7)

También en la investigación de Hrušková se encuentran definiciones con mínimas diferencias, como la que propone el *Diccionario de la Lengua Española de la Real Academia*: «conjunto lo más extenso y ordenado posible de datos o textos científicos, literarios, etc., que pueden servir de base a una investigación» (Ídem). En este caso se marca la extensión como una de sus características, lo cual no resulta del todo válido pues no necesariamente debe ser “lo más extenso posible”, según la bibliografía revisada.

Por otro lado, se ha revisado la definición que propone J. Sinclair “*chosen to characterize a state or variety of a language*”⁶ (*Apud* Hrušková, 2008: 7).

Esta definición es, en comparación con las anteriores expuestas, algo general referido más al uso u objetivo de los córpora.

Otros conceptos exponen que la representatividad del corpus puede ser o no absoluta y resaltan su aplicación informática:

Un corpus es un conjunto homogéneo de documentos lingüísticos de cualquier tipo (orales, escritos, literarios, coloquiales, etc.) que se toman como modelo de un estado o nivel de lengua predeterminado, al cual representan o se pretenden que representen. Ese conjunto de enunciados se someterá a un tratamiento informático cuyos resultados

⁶Escogido para caracterizar a un estado o variedad de un idioma.

permiten el mejor conocimiento de las estructuras lingüísticas de la lengua representada. (Alvar y Villena, 1994: 9-10)

En otros trabajos se aclara que un corpus “es una colección de elementos lingüísticos seleccionados y ordenados de acuerdo con criterios lingüísticos explícitos con la finalidad de ser usado como muestra de la lengua” (J. Pérez Guerra, 1998: 43).

Mientras que en palabras de Gerardo Sierra Martínez, “un corpus lingüístico consiste en la recopilación de un conjunto de textos de materiales escritos y/o hablados, agrupados bajo un conjunto de criterios mínimos, para realizar ciertos análisis lingüísticos” (2008: 451).

Cada una de estas definiciones enfoca un aspecto distinto del concepto de corpus. En este trabajo se tomaron en consideración sus principales elementos para definir que:

Un corpus lingüístico es un conjunto de textos almacenados en formato electrónico, y agrupados con el fin de estudiar una lengua o una determinada variedad lingüística. Su objetivo es constituirse en elemento de referencia para el estudio de una frase concreta o un cierto aspecto de una lengua.

En el caso del trabajo que se presenta, el corpus que lo sustenta servirá a los estudiosos de la modalidad cubana de la lengua española durante el periodo colonial, específicamente en el territorio oriental, y se centra en un estrato sociocultural elevado, pues los documentos que recoge fueron elaborados por notarios, funcionarios públicos que trasladaban a las escrituras, con exactitud y celeridad, las intervenciones orales de un tercero (Godoy, s.f.: 5)⁷.

⁷La procedencia socio-racial de los notarios estuvo estrictamente vigilada por las autoridades mediante una serie de requisitos personales y profesionales de obligatorio cumplimiento. Con esto, no solo se esperaba que entrara al gremio el varón mayor de edad, “de mente sana, vidente y

Las características básicas del corpus lingüístico, según Jitka Hrušková (2008: 8) son: representatividad, tamaño final, formato y anotaciones. En su trabajo explica cada una de estas características, las cuales serán referidas a continuación (Ibíd.:8-9).

La representatividad es una cualidad necesaria para cada corpus lingüístico. Un corpus lingüístico que es representativo puede ser útil en una investigación sobre la lengua. El corpus que no es representativo es una mera conjunción de textos que no nos dirá mucho sobre el estado de la lengua. Una cosa importante en relación con la representatividad es, como menciona Sinclair, el objetivo con el que se crean los corpus. (Hrušková, 2008: 8)

Otros investigadores resaltan de igual forma la importancia que tiene para un corpus la representatividad. Refieren que los corpórea deberían representar el estado de una lengua, por lo que lo ideal sería que incluyeran todos los enunciados de la misma, lo cual es prácticamente imposible, tanto por razones lingüísticas como técnicas; precisamente es esta cuestión la que le confiere el carácter de representatividad al corpus.

Así este debe incluir lo más representativo de una lengua para poder ser objeto de investigaciones lingüísticas y de otra índole.

En cuanto al tamaño final Hrušková (*op. cit*) refiere que es otra característica importante de los corpus y explica que, al crear un nuevo corpus se hace

oyente y constituido en íntegra fama (...) que tenga pleno conocimiento del arte notarial o tabeliano”, sino el que pudiese demostrar un origen familiar legítimo, justificando sus ascendientes como “cristianos viejos de limpia sangre y generación, libre de toda mala raza de moros, judíos, mulatos (...) ni penitenciado por el santo oficio de la inquisición, ni castigado por otro algún tribunal que les haya causado mala nota de su familia”. (Perera y Meriño, 2009: 104)

imprescindible establecer su extensión, aunque el resultado sea algo distinto.
(*Ídem*)

De acuerdo con el tamaño, existen varias clasificaciones para los corpus. Esta división no es totalmente absoluta, por lo que un corpus puede pertenecer a más de una clasificación:

1. **Corpus de referencia.** Es un corpus inmenso que debería proporcionar una visión completa de la lengua. Estos corpus se usan para la creación de diccionarios y sirven como punto de referencia para las gramáticas. Ejemplos: British National Corpus, Českýnárodní korpus (SYN2000), CREA.
2. **Corpus de fragmentos textuales.** Es característico de este corpus que no se incluyen textos completos, sino solo muestras. Las muestras normalmente tienen el mismo tamaño. Ejemplo: Brown Corpus.
3. **Corpus monitor.** Es un corpus que tiene un tamaño constante y al que se añaden textos nuevos y a la vez se eliminan textos antiguos en una cantidad equivalente. Ejemplo: CREA.
4. **Subcorpus.** Es una porción del corpus; según Atkins, Clear y Ostler puede ser cualquier parte de un corpus mayor. Según Sinclair, también el subcorpus debería tener las características de un corpus (por ejemplo la representatividad). Ejemplo: El subcorpus oral de banco de datos CREA-CORDE.
5. **Corpus especial.** Se crea con un objetivo diferente que un corpus de referencia: se centra solo en el estudio del lenguaje de un grupo social, de una zona, etc. Suele ser pequeño y no debe cumplir la condición de

representatividad. Ejemplo: CLUVI (Corpus Lingüístico de la Universidad de Vigo)

6. **Corpus bilingüe.** Normalmente se crea con el objetivo de la traducción automática. Existen dos tipos de corpus bilingües:

- a) el *corpus paralelo*, que es un conjunto de textos y traducciones a otra lengua. Ejemplo: CLUVI (Corpus Lingüístico de la Universidad de Vigo)
- b) el *corpus comparable*, que se compone de textos con características similares, pero no se trata de traducciones. Ejemplo: BCRC (Bilingual Comparable Retrieval Collection; corpus comparable de inglés y búlgaro., ejemplo elegido con ciertas restricciones). (Ibíd: 9)

El corpus que se presenta en este Trabajo de Diploma responde a las características de los corpus especiales, en tanto se crea con un objetivo diferente que un corpus de referencia (*Vid* objetivo), se centra solo en el estudio del lenguaje de un grupo social y de una zona, ya declarados (*Vid* Introducción).

En cuanto al formato, se dice que para que el texto pueda formar parte del corpus, es necesario que esté en forma electrónica. (*Ibíd*:10)

Atendiendo a los documentos manuscritos, que es el caso que nos ocupa, el método para transferirlos a la forma electrónica es el de transcribirlos en el teclado, o adaptarlos, si ya está en forma electrónica.

Las anotaciones, que son otro paso imprescindible en la realización de los corpus, según Jitka Hrušková (*op. cit*) son clasificaciones e interpretaciones del texto. Gracias a ellas se puede hacer búsquedas más precisas en el corpus y se dividen en:

Intratextuales, que ofrecen información sobre el texto entero: el título de la obra, el nombre del autor, la fecha de publicación, etc.

Intertextuales, que son clasificaciones en el nivel de las palabras. Se trata de clasificaciones gramaticales (*tag*), de las que las más frecuentes son las morfológicas (las categorías lexicales) o las sintácticas (la estructura).

En el caso del corpus que se presenta ofrece anotaciones intratextuales como: la tipología textual a la que responde cada documento, el archivo en que se conservan, la región geográfica y el periodo histórico en que fueron elaborados, así como el escribano que lo redactó.

Atendiendo a lo expuesto con anterioridad, es necesario resaltar dos de las primeras características de un corpus:

Ha de ser limitado, aunque no por ello debe dejar de ser un fiel reflejo de la lengua.

Los documentos que constituyen el corpus deben guardar entre sí, y en todo momento, las mismas proporciones que existen en el conjunto del estado o nivel de la lengua representados. (Alvar y Villena, 1994: 10)

Existe también una denominación atendiendo a la extensión de los documentos recogidos para el corpus. Cuando se toman íntegramente los documentos para la confección del corpus, estamos en presencia de un *corpus textual*. En cambio, cuando se seleccionan fragmentos de los documentos, se denomina *corpus de referencia*. (Ídem)

El corpus aportado con esta investigación se puede considerar de referencia, en tanto si bien la mayoría de los documentos que lo integran aparecen íntegros, no se ha conformado con la totalidad de los existentes en los archivos consultados.

Otra clasificación que se hace de los corpus es atendiendo a la especificidad de los textos que los componen. A partir de este parámetro se definen seis tipos (Torruella y Llisterri, 1999: 11):

- 1- **Corpus general:** corpus que, al pretender reflejar la lengua común en su ámbito más amplio, se interesa por recoger cuantos más tipos de género mejor. Este tipo de corpus es útil para describir la lengua común de una colectividad, el lenguaje que utilizan los hablantes en situaciones comunicativas normales.
- 2- **Corpus especializado:** recoge textos que pueden aportar datos para la descripción de un tipo particular de lengua.
- 3- **Corpus genérico:** condicionado por el género de los textos que contiene, solo se interesa por algunos de ellos, como una recopilación de textos de revistas científicas especializadas o la selección de textos poéticos.
- 4- **Corpus canónico:** está formado por todos los textos que configuran la obra completa de un autor, independientemente de los géneros.
- 5- **Corpus periódico o cronológico:** recoge los textos de unos años determinados o de unas épocas concretas.
- 6- **Corpus diacrónico:** incluye textos de diferentes etapas temporales sucesivas en el tiempo, con el fin de poder observar evoluciones en la lengua.

El corpus que se presenta puede considerarse especializado, en tanto agrupa textos que se corresponden con el estilo funcional jurídico- administrativo; pero también diacrónico, al recoger y ordenar los documentos seleccionados para su conformación teniendo en cuenta su correspondencia a los siglos XVII al XIX.

La aplicación práctica de los corpóras lingüísticos tiene una gran variedad, y según Alvar y Villena (1994: 10) existen diversos usos en diferentes ramas de las ciencias lingüísticas.

La utilización más fácil y frecuente de un corpus se realiza en *lexicografía*, al ser muy útiles en la confección de diccionarios. A partir de él se pueden determinar una variedad de aspectos que ayudarían a que un diccionario no contenga voces o acepciones desusadas, ni ejemplos falsos o expresiones forzadas. (*Ídem*)

En el campo de la *lexicología y la semántica* también tienen aplicación práctica los datos contenidos en un corpus. A partir de estos se pueden hacer análisis en cuanto a la formación de palabras, cambios de significados, evolución del léxico, para documentar la presencia de formas no empleadas antes, la aparición de neologismos, la introducción de barbarismos, etc.

En la *fonética* sobresale su uso como fuente considerable de información en lo relacionado con la síntesis del lenguaje y la generación del habla. También resulta de gran utilidad su aplicación en la comunicación oral con las máquinas. (*Ibíd.:* 11)

Los corpóras son útiles para todo lo relacionado con la *lingüística computacional*, en la confección de diccionarios-máquinas de diversos tipos, en la composición electrónica para las imprentas: divisores silábicos, correctores ortográficos, diccionarios de sinónimos, etc.

Su uso es igualmente válido para la creación de los llamados diccionarios electrónicos, en la traducción asistida por ordenador y en cualquier aspecto del

tratamiento del lenguaje, resaltando todo lo que facilite la comunicación con las máquinas en el lenguaje natural.

A la par, los corpóra resultan de gran utilidad en la *estadística lingüística* para establecer índices de frecuencias de formas: sonidos, letras, palabras y obtener informaciones sobre las reglas combinatorias de los signos, el grado de viabilidad de los elementos de la formación de palabras, la frecuencia de aparición de las voces⁸. Por otro lado, proporcionan información sobre los diferentes niveles de la lengua, en auge para los estudios de la *sociolingüística* moderna.

En lo referido a la enseñanza de lenguas, un corpus puede aportar materiales de gran eficacia, aceptando su representatividad para la confección de gramáticas de las lenguas, en las que se describa el funcionamiento de la lengua en sus niveles, poniendo al alcance de los alumnos los elementos de aprendizaje y comprobación necesarios.

Por último, fuera del ámbito de la filología, tienen aplicación básica en los procesos industriales que necesitan de aplicaciones lingüísticas basadas en los conocimientos que pueden proporcionar los corpóra⁹.

Como se puede apreciar, el uso de un corpus es muy variado y, específicamente el que pretende elaborar esta investigación, recoge textos manuscritos desde el siglo XVII hasta el XIX.

Su objetivo primordial es el de reunir los documentos notariales, dispersos en los trabajos de diploma mencionados en la introducción, para facilitar a los investigadores de la modalidad cubana de la lengua española, el acceso rápido y

⁸Revisar, sobre esta aplicación, el libro de Manuel Alvar y Juan Andrés Villena, *Estudios para un corpus del español*, 1994, pág.11

⁹Sobre esta aplicación, revisar el libro de Manuel Alvar y Juan Andrés Villena, *Estudios para un corpus del español*, 1994, pág.11

confiable a los documentos, tanto de su copia facsimilar como de la transcripción paleográfica.

Varias son las investigaciones dedicadas a trazar una metodología para la elaboración de un corpus. En uno de sus trabajos Torruella, J y Llisterri, J (1999) proponen una serie de elementos necesarios para la elaboración de los córpora, donde resaltan la finalidad concreta y los límites que tendrá el corpus, como una de las características principales que permitirá determinar la validez de los trabajos que se realicen a partir de su explotación.

En cuanto a la finalidad, se dice que es el primer aspecto que se debe definir cuándo se empieza a diseñar un corpus, ya que servirá de base para tomar las diferentes decisiones en los demás puntos a definir con posterioridad. El siguiente paso a seguir, después de establecer la finalidad, es determinar los límites geográficos, temporales y/o lingüísticos. Para lograr dicho objetivo se deberá marcar una fecha de inicio y otra de final.

Del mismo modo, es preciso especificar las lenguas que el corpus va a incluir y/o el área geográfica que abarcará. Estos dos aspectos no son necesariamente absolutos, sino que puede estar sujetos a cambios o variaciones, según el tipo de corpus que se va realizar. Es decir que, entre un corpus y otro, tanto los límites geográficos y los temporales, pueden ser diferentes, en tanto respondan a la investigación que se realiza.

Sobresalen ejemplos como el corpus del español que se está realizando en el King's College de Londres, donde se recogen textos posteriores a 1990; el corpus de Biblograf, que recoge textos publicados a partir de 1950; el Longman Lancaster English Language, corpus que recoge textos posteriores al año 1899; el Corpus Textual Informatizat de la Llengua Catalana del Institut d' Estudis Catalans que empieza la recolección de textos a partir de 1833. (Torruella y Llisterri, 1999:16).

De igual forma, también varían los límites geográficos, donde cada corpus puede contener textos del mismo idioma de diversas zonas geográficas. Con estas características sobresalen los corpus del español CREA (Corpus de referencia del español actual) y CORDE (Corpus diacrónico del español).

Estos corpus presentan textos tanto de España como de Hispanoamérica. El CREA tiene una proporción de 50% de España y 50% de Hispanoamérica¹⁰. En cuanto al CORDE se encuentran alrededor de 58 millones de palabras textuales de Hispanoamérica y casi 237 millones de palabras textuales de España para una proporción de 21% y 79% respectivamente. (Hrušková, 2008: 19)

¹⁰Estos datos se presentan en los “rasgos generales de CREA”, ofrecidos en su página web: http://corpus.rae.es/ayuda_c.htm#_Toc30228258. Según Hrušková (*op. cit*), si se realiza una búsqueda más exhaustiva, el usuario puede percatarse de que en realidad hay aproximadamente 84 millones de palabras procedentes de recursos españoles y 67 millones hispanoamericanos, formando una proporción de 56% y 44%. (Hrušková, 2008: 19)

1.3 Características formales y estilísticas de los documentos notariales.

Como la muestra que forma parte del corpus que se presenta en su totalidad está conformada por documentos notariales, se hace necesario presentar las principales características de estos.

Un documento notarial se define como la configuración formal escrita de una actuación jurídica trascendente. Una actuación realizada por un autor u otorgante que da el contenido o materia; una escrituración, que le da la forma y que es la puesta por escrito de la actuación, y como tal creación.

También exige un realizador de tal tarea, un autor. Se reviste de una serie de formalidades extrínsecas e intrínsecas y de las garantías necesarias para dar valor y credibilidad: jurídica, administrativa, histórica, tanto al mensaje, contenido y texto de que es portador, como al propio soporte documental. (Pérez, 2011: 38).

Este tipo de documentos son redactados por notarios, que son personas con una cultura general vasta, lo cual evidencia su formalidad en el uso de la lengua. Esta característica dota a los textos escritos por notarios, de un alto valor documental para los estudios de la lengua española. Para su redacción los notarios utilizan la norma culta de la lengua o un estilo formal.

En los documentos notariales se evidencia el uso del lenguaje administrativo, que es la lengua empleada por los órganos de la administración del estado.

Se aprecia una estrecha relación entre los textos jurídicos y los administrativos. El lenguaje jurídico se puede definir como la lengua empleada por los órganos de la Administración de Justicia en sus relaciones con la colectividad o con las personas físicas y jurídicas, es decir, como un tipo de lenguaje administrativo específico.

El lenguaje jurídico y el administrativo comparten los recursos lingüísticos (gramaticales y léxicos) y coinciden también en los factores extralingüísticos que

los caracterizan (el canal, el emisor, el receptor y la finalidad o intención comunicativa). Se utilizan en textos legales o de contenido jurídico-legal y en los documentos oficiales. (El lenguaje jurídico administrativo, s.d.e. *Apud.* Labrada, 2016: 24-25).

Estos documentos se manifiestan de forma escrita, a través de variados tipos, entre los que destacan:

Testamento: Acto jurídico, individual, libre y revocable por el que una persona regula su sucesión nombrando uno o más herederos. La práctica más habitual y segura es el testamento ante notario, en su modalidad de testamento abierto¹¹. Existen otras formas, como el testamento cerrado ante notario, el ológrafo, o el militar y marítimo o ante testigos, pero se usan poco. (Notaría Modelo)¹²

Capitulaciones matrimoniales: Acuerdo entre cónyuges cuyo objetivo esencial es pactar un régimen económico conyugal, así como liquidar el anterior. También las pueden otorgar aquellos no casados, que en el año siguiente vayan a contraer matrimonio.

Poder: Facultad que se da a otra persona para realizar y ejecutar determinados actos jurídicos y materiales en nuestro nombre. El apoderado no tiene que aceptar el poder, es un negocio unilateral del poderdante. Puede no conocer que se le ha otorgado el poder, porque basta que una persona acuda a la notaría y lo otorgue. El poder es revocable por parte del poderdante, por medio de otra escritura posterior, quedando sin efecto, por tanto, la primera. (*Ídem*)

Acta: En este documento el notario hace constar, por ejemplo, la relación de lo acontecido durante la celebración de una asamblea, congreso, sesión, vista

¹¹El testamento abierto es el que se realiza cuando el testador manifiesta su última voluntad en presencia de las personas que deben autorizar el acto, quedando enteradas de lo que en él se dispone. (Sánchez, 1989: 52 *Apud.* Pérez, 2016: 28)

¹² Documento disponible en <http://notariasantiagodecompostela.wordpress.com/los-documentos-notariales>. Consultado el 3 de marzo de 2018.

judicial o reunión de cualquier naturaleza y de los acuerdos o decisiones tomados.
(Ídem)

Compraventa: Contrato por virtud del cual uno de los contratantes se obliga a transferir la propiedad de una cosa o derecho y el otro, a su vez, se obliga a pagar por ello un precio convenido. Puede ser de inmuebles, muebles o derechos.
(Ídem)

Declaración de obra nueva: es la manifestación en escritura pública del propietario o de todos los copropietarios de una finca de haberse proyectado, comenzado o concluido edificios, nuevas plantas o mejoras en los mismo, es decir, la edificación (o la modificación de la misma) que se ha construido o se está construyendo sobre un terreno.

Constitución de sociedades mercantiles: Acto de fundación de una sociedad donde se definen sus elementos básicos: capital inicial, domicilio, régimen jurídico, etc. Las sociedades mercantiles se constituyen de acuerdo con la legislación mercantil, independientemente de que su finalidad sea comercial. Lo son las sociedades anónimas y las sociedades limitadas, entre otras.

Declaración de herederos abintestato: Es heredero abintestato aquella persona (o personas) que, al no haber heredero nombrado en testamento, lo es por establecerlo la ley. Para ser nombrado heredero abintestato es necesario formalizar un expediente de declaración, que será notarial o judicial, según el parentesco que se tenga con el fallecido.

Póliza: Documento mercantil, redactado por una de las partes. Si la interviene un notario tiene unos efectos privilegiados, fundamentalmente a la hora de exigir por vía judicial lo pactado en ella. Puede no recurrirse a un notario, pero entonces carecerá de esos efectos.

Protesto: Documento donde se hace constar la negativa de aceptar o pagar una letra de cambio, pagaré o cheque para no perjudicar o disminuir los derechos y acciones de las personas que han intervenido.

Préstamo hipotecario: Préstamo que tiene como garantía de pago un bien inmueble o varios, sean viviendas, locales, garajes, solares, fincas rústicas, etc. (*Ídem*)

El corpus que aporta esta tesis recoge como tipologías textuales: testamentos, poderes, cartas de venta y cartas de libertad; por ser estos los tipos de documentos que más se han trabajado en las investigaciones que le han servido de base (*Vid* Introducción).

La estructura formal del documento notarial quedó establecida desde el siglo XIII, cuando los fundadores del *Ars Notaria* adaptaron las viejas figuras contractuales del derecho romano, que aseguraban la perfección jurídica de los negocios, para establecer normas y modelos que facilitarían la formulación de los negocios y las cláusulas más comunes, delimitándose así tanto la estructura general del documento notarial -que todavía puede ser reconocida en la documentación contemporánea-, como los diferentes tipos documentales; los cuales aunque desde luego no cubren todas las posibilidades negociables, si satisfacen, incluso en nuestros días, las necesidades más frecuentes. (Ramírez, 1997: 79)

En relación con esto, toda escritura notarial, independientemente de que se trate de un poder, una venta, un testamento o inclusive de un contrato sin nombre específico, comparte una misma estructura documental y un lenguaje básicos, que están dados en función de la eficacia de los negocios escriturados, y son los que le dan al documento seguridad legal y su fuerza de prueba jurídica. (*Ídem*)

Desde el punto de vista de su estructura diplomática, la escritura notarial consta de tres grandes partes que son: el protocolo, el cuerpo o centro del documento y el escatocolo o protocolo final. Cada una de estas partes desempeña una diferente

función y se subdivide a su vez en diferentes tipos de cláusulas que en conjunto forman el tenor o discurso documental (*Íbid.*: 79-80).

El primero y el último contienen en conjunto las fórmulas legales que dan al escrito su perfección de prueba legal-por ejemplo, el lugar y la fecha en que se hizo la escritura, los nombres de los testigos o las validaciones del escribano. Por su parte, el centro o cuerpo del documento contiene diferentes tipos de cláusulas mediante las cuales se perfecciona el negocio jurídico en sí. (*Ídem*)

En cuanto a su forma, los documentos notariales aparecen casi siempre agrupados en protocolos, que no es sino la agrupación de todas las escrituras y actas firmadas ante un notario en un año, aunque, por extensión, se considera habitualmente protocolo a cualquier agrupación de escrituras ordenadas cronológicamente en un volumen, con independencia del período de tiempo que abarquen las escrituras. (Notaría Modelo)

Dentro de cada protocolo se encuentran las escrituras notariales, una detrás de otra sin guardar más relación de contenido entre sí que la de haber sido firmadas y redactadas, ante un mismo notario.

Por ejemplo, en los dos folios que se presentan a continuación, se puede apreciar cómo culminan los documentos que se refieren en el corpus (06 y 43) e inicia, seguidamente después de las firmas, otro redactado con la misma letra, lo cual evidencia que han sido elaborados por la misma persona, aunque sean tipologías diferentes que responden a protocolos distintos.



Carta de libertad



Poder

Doc. 06b



Carta de venta



Poder

Doc. 43b

Cada una de las escrituras notariales es el reflejo de un negocio jurídico, un acto en el cual intervienen una o varias personas que requieren de la presencia del notario para que quede una pública constancia de ello. Por eso se denomina fe pública a la función que notarios han ejercido a lo largo del tiempo (*Ídem*).

En cuanto a sus características estilísticas, se encuentran:

- uso de arcaísmos, cultismos, latinismos, tecnicismos jurídicos, fórmulas léxicas; la abusiva creación de eufemismos innecesarios
- sustitución de verbos simples por perífrasis léxicas; el uso del futuro de indicativo, con significado imperativo y normativo (ej.: se sancionará a..) y el del futuro de subjuntivo en lugar del presente o pretérito perfecto del mismo modo (ej.: sucediere y hubiere sucedido en lugar de sucede y haya sucedido)
- uso frecuente del gerundio y de la voz pasiva
- utilización de la tercera persona del singular o de la primera del plural con valor impersonal
- abundantes construcciones con participio
- incisos explicativos, aclaraciones, referencias y citas
- vocabulario abstracto
- constante uso de sinónimos y de repeticiones de palabras
- utilización de párrafos de gran extensión, lo que dificulta muchas veces la comprensión del contenido (*Ídem*).

La situación comunicativa está marcada por un notable formalismo; el tratamiento de la información es objetivo y el lenguaje es muy conservador (*Ídem*).

La prosa notarial debe ser sobria, diáfana y su estilo ajustado al acto que se instrumenta. La terminología empleada en la construcción de cláusulas obligacionales debe ser adecuada y concreta, y nunca rebuscada, ni caprichosa; la presentación del instrumento será limpia y debidamente emplanada, como corresponde a un título que ha de representar un derecho por un tiempo ilimitado (Verdejo, 1988: 89).

Los documentos tomados como muestra para la realización de este corpus se insertan dentro de los Fondos Protocolos Notariales de los diferentes archivos de

los que fueron seleccionados; es decir, responden a las características del estilo funcional jurídico administrativo esbozadas con anterioridad.

Capítulo II “Características del corpus de documentos notariales del oriente de Cuba (S. XVII- XIX)”.

En este capítulo se hace referencia a los presupuestos teórico-metodológicos de la Red internacional CHARTA, los cuales han servido de base, con determinadas adecuaciones que también se exponen, para la conformación del corpus que se presenta. Por otra parte, se explican las especificidades de la muestra y la manera en que se ha estructurado el corpus.

2.1 Referente metodológico para la elaboración del corpus: la red CHARTA.

La elaboración del Corpus de documentos notariales del Departamento oriental de Cuba (S. XVII-XIX) corpus (DONOC) se adscribe a los presupuestos teórico-metodológicos proporcionados por la red internacional CHARTA.

Esta Red CHARTA es un proyecto de edición de un corpus de documentos antiguos en el que se aprovechan la metodología de la filología y los avances de la informática con el fin de publicar en Internet un corpus de textos y documentos antiguos de los siglos XII al XIX escritos en España e Hispanoamérica (Diez del Corral y Martín, 2014: 288).

El origen de este proyecto se sitúa en una reunión celebrada en 2005, en Deusto, donde investigadores de las universidades de Alcalá, Deusto, Murcia, Valladolid, Las Palmas de Gran Canaria y del Consejo Superior de Investigaciones Científicas se reunieron para debatir sobre el aprovechamiento del lenguaje informático de marcación en las tareas de la edición digital.

La falta de un estándar en el campo de la edición de textos antiguos motivó a estos investigadores a trabajar en conjunto con el objetivo de proponer unos criterios que pudieran ser válidos para toda la comunidad científica (*Ídem*).

La presentación en público de esta iniciativa tuvo lugar en el Primer congreso de la Red, celebrado en el Consejo Superior de Investigaciones Científicas (Madrid, noviembre de 2009), bajo el rótulo “Tradición e innovación: nuevas perspectivas para la edición y el estudio de documentos antiguos”.

En él, los miembros de los diferentes grupos de trabajo pusieron en común sus tareas en el campo de la edición y estudios lingüísticos (*Ibíd.*: 289).

En septiembre de 2011, siguiendo el camino iniciado dos años atrás, se celebró el Segundo congreso, esta vez en la Universidad de Neuchâtel (Suiza), la cual concluyó con la acogida de nuevos integrantes.

El principal objetivo de este grupo ha sido establecer una metodología común para la edición de textos, ya que hasta el momento han coexistido y coexisten diferentes criterios que atienden a los intereses dispares desde los que se acercan al texto los investigadores –historiadores, diplomáticos, paleógrafos, etc. (*Ibíd.*: 290).

En segundo lugar, se ha realizado una propuesta de presentación gráfica para la transcripción paleográfica y la edición crítica, atendiendo a criterios filológico-lingüísticos rigurosos y homogéneos (*Ídem*).

También forma parte de las tareas de este grupo internacional, elaborar un amplio corpus en red de fuentes documentales que abarque la variedad geográfica del español, tanto de España como de América, en un arco cronológico amplio y con muestras de textos pertenecientes a tradiciones discursivas muy dispares, todo ello valiéndose de las herramientas informáticas para el análisis lingüístico de los textos (*Ídem*).

Los documentos se presentan en edición digital fiable, comprobable y que se puede citar directamente tanto en estudios científicos como en trabajos

divulgativos, o simplemente para satisfacer la curiosidad por el pasado, incluso en un ámbito local y familiar (Red Internacional CHARTA)¹³.

De cada uno de los documentos se ofrece: transcripción paleográfica, presentación crítica y facsímil. Ello permite llevar a cabo estudios de paleografía, historia de la escritura, grafemática, fonética histórica, morfología y sintaxis, historia del léxico, configuración discursiva, así como de diplomática, historia general y de las mentalidades y de la vida privada, etc. (*Ídem*).

Esta edición múltiple se justifica por la imposibilidad de proporcionar con una sola toda la información que el estudioso demanda. El triple acceso permite llevar a cabo estudios de la más variada índole; los tres son complementarios, y permiten el exhaustivo análisis del documento en su integridad (Criterios de edición: 2013)¹⁴.

El facsímil hace posible, aparte de comprobar las lecturas, abordar estudios diplomáticos y paleográficos. La conjunción de este con la transcripción paleográfica permitirá poner de relieve el sistema gráfico y plantear la evaluación fonética de las grafías.

La presentación crítica¹⁵ facilita la lectura; es la “edición” adecuada para el estudio de la morfología, de la sintaxis y del léxico, al tiempo que favorece los estudios históricos.

Este triple acceso tiene particular aplicación en la edición digital, aunque también se ha ensayado en el formato tradicional de libro (*Ídem*).

¹³ Este documento puede consultarse en el sitio <http://www.redcharta.es/>

¹⁴ Criterios de edición de documentos hispánicos (Orígenes-siglo XIX) de la Red Internacional CHARTA. Versión de abril de 2013. Documento digital, recuperado de <http://www.redcharta.es/criterios-de-edicion/> [consultado el 20 de enero de 2018]

¹⁵ En el caso de este trabajo, se presenta, en lugar de la Edición crítica, la Guía comentada o Ficha de trabajo en archivo que contiene información sobre cada documento. Esto responde al contexto en que se enmarcan las investigaciones en la Universidad de Oriente relacionadas con este aspecto, las cuales se insertan en el Proyecto de investigación VLIR ya mencionado (*Vid* Introducción).

Los criterios que presenta el grupo CHARTA son fruto de una reflexión prolongada entre todos los miembros y grupos de investigación que lo integran.

La principal ventaja que presentan frente a los criterios individuales es que han sido muy meditados y ensayados en la edición de textos variados, tanto desde un punto de vista cronológico como desde un punto de vista tipológico (Diez del Corral y Martín, 2014: 302).

Estos criterios presentan otro rasgo primordial que es su universalidad. Se ha intentado que los criterios sean válidos para todo el ámbito hispánico, con el objetivo de poder editar la inmensa cantidad de series documentales que se custodian en los diferentes archivos de los países hispanohablantes.

La validez de los criterios que ofrece la red CHARTA podría incluso ampliarse a otros textos en lenguas romances.

La idea de la edición múltiple es perfectamente aplicable y recomendable para la edición de cualquier documento y los criterios de la transcripción paleográfica también servirían, aunque quizá se necesitaría –al igual que para la presentación crítica– una revisión integral de todos ellos para ver su adecuación a los textos que se quieren editar y, quizá, la propuesta de criterios específicos para cada lengua, e incluso para cada época histórica (*Ibíd.*: 303).

El grupo de trabajo de CHARTA considera y ofrece sus criterios de edición como una base que cualquier editor de textos en el ámbito hispánico puede tomar como referencia, a pesar de que en última instancia el mismo pueda exponer y tomar sus propias soluciones (*Ídem*), lo cual es lo que se ha hecho en el corpus (DONOC)

2.2 Procedimientos para la conformación científica del corpus de documentos notariales del oriente de Cuba.

2.2.1 Caracterización de la muestra y adecuaciones de los criterios de edición de la Red CHARTA.

Para la conformación del corpus (DONOC) se escogieron documentos que responden a la clasificación de *Protocolos Notariales* en los Archivos en que se ha trabajado: el Archivo Histórico Provincial de Holguín (AHPH), el Archivo Histórico Provincial de Santiago de Cuba (AHPCS) y el Archivo Histórico Provincial de Guantánamo (AHPG); y que cubren los siglos XVII al XIX.

Los documentos están clasificados, dentro del corpus, por tipologías textuales, distribuidos de la siguiente forma:

- cartas de libertad o manumisión: 8
- cartas de ventas: 44
- poderes: 9
- testamentos: 17

Para un total de 77 documentos.

La distribución de la muestra según los siglos a que pertenecen, la región del Departamento oriental a la que pertenecen y el archivo en que se encuentran resguardados los documentos ha quedado de la siguiente forma:

Siglo	Región	Archivo	Tipología textual
XVII (4)	Santiago de Cuba	AHPSC	Cartas de venta (2) Testamento (1) Poder (1)
XVIII (38)	Santiago de Cuba Holguín	AHPSC AHPH	Cartas de venta (32) Testamentos (2) Poder (3) Carta de libertad (1)
XIX (36)	Santiago de Cuba Holguín Guantánamo	AHPSC AHPH AHPG	Cartas de venta (10) Testamentos (14) Poder (5) Cartas de libertad (7)

Cuadro 1. Distribución de las muestras

De cada uno de estos documentos se hacen tres presentaciones para facilitar su estudio posterior.

Primeramente, se presenta su copia paleográfica (transcripción), en segundo lugar, su copia facsimilar en formato JPG y, en tercer lugar, una Guía comentada o Ficha de trabajo en archivo, en lugar de la Edición crítica establecida por CHARTA (*Vid supra* Nota 15), y sustentada tal decisión con la propuesta del grupo de investigación de la posibilidad de presentar distintas soluciones, según el tipo de investigación (Criterios de edición: 2013).

A continuación, se explican los diferentes elementos que componen la ***cabecera del documento***.

Cada documento debe presentar en la cabecera información referente a los datos más importantes que describen el documento, como la identificación del grupo de investigación, corpus y documentos dentro del corpus.

En el caso de este corpus solo se pondrán el número de orden del documento dentro del corpus, seguido en línea aparte por las siglas del nombre de dicho corpus.

Ejemplo:

01
Corpus DONOC

Para la identificación del documento se señalan las siglas del archivo que lo conserva y, en algunos casos, la signatura del archivo¹⁶ y el folio.

Ejemplo:

AHPSC, Protocolo notarial I, 45r-47v

La data y la localización se expresan según el orden: año, mes y día (sin comas ni otros signos).

Ejemplo:

1750 mayo 20

Cuando la data no aparece o está incompleta, se procede del siguiente modo:

- la falta de todos los datos se codifica con las siglas s.f. (sin fecha)
- la falta de alguno de ellos con s.a (sin año), s.m. (sin mes) y s.d. (sin día).

Cuando el año puede reconstruirse con certeza se sitúa entre corchetes ([]); si la reconstrucción es aproximada, se emplea *ca.* ante el año:

¹⁶ En el caso de la signatura del archivo, solo se pondrán en algunos documentos, porque la muestra que conforma el corpus se encuentran en trabajos de diplomas anteriores (vid introducción) y por cuestiones ajenas a esta investigación, algunos de los manuscritos no poseen todos los datos que se requieren para añadirlos al corpus.

Ejemplo:

s.f [ca.1750]
s.a. [1750] mayo 20
s.a. [ca. 1750] mayo s.d.

Los datos referentes al registro contienen la indicación del tipo de documento, emisor, destinatario y breve resumen del contenido. Todo esto sin sobrepasar los 500 caracteres. Puede incluirse la expresión del tipo de soporte (pergamino o papel).

También se inserta, si es necesario, información sobre el estado de conservación de la pieza.

Por último, se agregan los nombres de los investigadores responsables de la edición, que se presentan alineados a la derecha, en columna. Tras el nombre del transcriptor del documento, se indican los nombres de los revisores, si los hubiera.

Ejemplo:

Transcriptor del documento
1er. revisor
2º revisor

Un ejemplo completo de cabecera de un documento sería:

01¹⁷

Corpus DONOC

AHPSC, Protocolo Notarial, ff. 242r- 242v

1783 diciembre 30 (Santiago de Cuba, Cuba)

Carta de manumisión en favor de Candelaria, esclava de Doña Juana Ferrer.

¹⁷Esta cabecera es la que corresponde al documento número 01 de la muestra (*Vid corpus*).

Luis Agustín Sierra

Papel

Regular estado de conservación

Karen Safonst Caballero.

Danay González Ferrás

En cuanto a los *elementos codicológicos*, se han tenido en cuenta las siguientes adecuaciones, a partir de lo estipulado por CHARTA.

Sobre la numeración de hoja y línea¹⁸:

- El número de hoja y de línea se consignan entre llaves y en el lugar que corresponda, aunque sea en mitad de una palabra (sin espacios)
- Las hojas corresponden al texto editado, por lo que siempre se empezará por la hoja 1, aunque en el original la numeración sea otra
- La indicación de recto (“r”) y vuelto (“v”) se hace inmediatamente tras el número de hoja, sin espacio.

Ejemplo:

hoja número 3: {h 3}

línea 7: {7}

Los párrafos no se numeran y para indicar el cambio de línea en las acotaciones marginales se emplea la línea vertical (|).

¹⁸ Dentro de los criterios de CHARTA, también se consignan las columnas, pero debido a que los documentos que conforman el corpus en cuestión no poseen columnas, solo se escoge la numeración de las líneas y las hojas.

En cuanto al deterioro del original

- Cuando exista la certeza de que hay caracteres en un espacio deteriorado por roto, doblez o mancha, en la transcripción paleográfica se emplearán asteriscos (*): si se sabe el número concreto de letras ilegibles, se emplea un asterisco por cada letra; si, por el contrario, no se sabe el número exacto de letras, se emplean tres asteriscos separados entre sí por un espacio y recogidos entre corchetes

Ejemplo:

d ** vez * nos
dos v [* * *]

- Cuando no se puede leer o falta texto, se expresa en la transcripción paleográfica la causa del deterioro en cursiva y entre corchetes, en el lugar que le corresponda.

Ejemplo:

[*roto*]
[*doblez*]
[*mancha*]

Signos o elementos especiales

- Se emplean los corchetes y la cursiva para indicar la presencia de sellos de papel timbrado, signos o elementos especiales:

Ejemplo:

[*sello*] de cera colgante, de plomo, de placa
[*crismón*]
[*cruz*]
[*signo*] y [*signo: texto*]

- Se distingue entre firma (cuando hay texto) y rúbrica (trazos que se añaden al nombre al firmar). El texto de la firma se consigna dentro de los corchetes y tras (:).

Ejemplo:

[firma: nombre]

[firma: nombre] [rúbrica]

Transcripción paleográfica

- Las abreviaturas en la transcripción paleográfica, se mantienen tal y como se representan en el documento original.
- En cuanto a las grafías, se reproducen las letras y no las distintas formas que pueda adoptar cada letra. Es decir, se mantienen como en el original.
- En cuanto al uso de las mayúsculas, minúsculas, la acentuación y la puntuación; también se mantienen tal como aparecen en el original.
- Para la unión y separación de las palabras y ruptura de palabra a final del renglón, se diferencian las palabras unidas de las palabras enlazadas mediante trazo cursivo.
- La ruptura de palabra a final de renglón se marca donde corresponda en el manuscrito con el número de línea correspondiente inserto en la palabra y sin espacios.

Ejemplo:

perte{6}nencia

cas{23}amiento

2.2.2 Estructura del corpus y modos de búsqueda.

El corpus se estructura en tres carpetas (transcripciones, copia facsimilar y guía comentada), que corresponden a un criterio de búsqueda general. Cada una de las carpetas organiza la información internamente según la tipología textual y, dentro de cada tipología textual, según el siglo (por orden cronológico).

En la carpeta principal, llamada Transcripciones, se encuentran los textos transcritos en soporte digital, de los cuales, mediante herramientas de búsqueda, se puede extraer información lingüística, ya sea léxica, morfológica o sintáctica. Cada texto se identifica externamente con un número corrido (01- 77).

En el primer y en el último renglón de cada texto tecleado aparece una palabra que tiene creado un hipervínculo. Al clicar sobre la palabra con hipervínculo del primer renglón, se carga una carpeta con imágenes del documento original, que corresponden a las copias facsimilares de la transcripción.

Y por su parte, al clicar sobre la palabra asociada al hipervínculo que se encuentra en el último renglón, aparecerá un texto con información más detallada sobre el documento, así como, los datos del archivo histórico en que se encuentra¹⁹.

En la carpeta llamada Copia facsimilar, se recogen todos los textos fotografiados, organizados por orden numérico (en correspondencia con el número de orden de los documentos, seguidos de las letras por orden alfabético en dependencia de la cantidad de fotos existentes).

Ejemplo:

54a

54b

¹⁹Esto es lo que en el corpus se llama Guía comentada o Ficha de trabajo en archivo, documento concebido dentro del Proyecto de investigación al cual pertenece esta investigación (*Vid* Introducción).

54c

Así pues, estos ficheros pueden ser abiertos directamente o ser cargados al clicar sobre el hipervínculo que se encuentra en el primer renglón de cada texto tecleado de forma digital.

Lo mismo ocurre con la carpeta llamada Guía comentada, en la que se recogen todas las Guías comentadas, organizadas también por orden numérico, con el mismo número que los textos transcritos a los que corresponden. Esos ficheros, al igual que en el caso de los que se encuentran en la carpeta de las copias facsimilares, pueden ser abiertos de forma directa o ser cargados al clicar sobre el hipervínculo que le corresponde en cada texto tecleado de copia digital.

Como ya se explicó con anterioridad, en el margen superior izquierdo de cada hoja transcrita se colocó un código que posibilita identificar el texto, de esta manera se garantiza que el corpus esté bien organizado.

El número del texto permite relacionar el documento fotografiado con el transcrito digitalmente y con la Guía comentada que lo acompaña; de manera tal que los tres documentos se interrelacionan y permiten tener un criterio general lo más completo posible de cada una de las muestras seleccionadas para la conformación del corpus presentado en este Trabajo de Diploma.

Conclusiones

En la presente investigación se logró comprobar la validez de la hipótesis planteada ya que, una vez organizados los documentos notariales representativos del Departamento oriental, entre los siglos XVII-XIX, teniendo en cuenta su distribución por tipologías textuales y siglos, se pudo conformar un corpus lingüístico contentivo de esta información, que contribuye al estudio de las características de la modalidad cubana de la lengua.

Con ello se dio solución al problema de investigación declarado y se cumplió el objetivo propuesto al elaborarse el corpus de documentos notariales del oriente de Cuba (S. XVII- XIX) (corpus DONOC).

Resulta indiscutible la valía de los documentos archivísticos de siglos pasados, como fuente imprescindible para los estudios de la variante cubana de la lengua. El caso específico del corpus DONOC, constituye una valiosa herramienta de trabajo que organiza estos documentos, de forma tal que posibilita una mejor accesibilidad de los mismos para los investigadores.

El criterio de búsqueda establecido, según las tipologías textuales y los siglos, permite una rápida localización de los documentos, lo cual garantiza una interfaz asequible para cualquier usuario del corpus, pues solo exige un conocimiento básico de computación.

Las transcripciones paleográficas realizadas a los documentos que conforman el corpus, a partir de los criterios establecidos por la red CHARTA, están en consonancia con las nuevas investigaciones que se realizan en el mundo sobre documentos antiguos del español.

De esta manera dichas transcripciones cumplen con los requerimientos fundamentales determinados metodológicamente por CHARTA, con lo cual se actualiza esta labor dentro del Proyecto de investigación que sustenta la tesis.

Los resultados de este estudio, constituyen un aporte para la ciencia lingüística en Cuba y en particular para la lingüística de corpus, al ser el primer corpus, que se tenga conocimiento, que recoge muestras de documentación del periodo colonial en el antiguo Departamento oriental. Demuestran la importancia y novedad de investigaciones como esta, para estudios, no solo lingüísticos, sino también de otras áreas de las ciencias sociales y humanísticas que necesiten acercarse al valioso legado documental que se atesora en nuestros archivos.

Recomendaciones

A partir de la investigación realizada para la conformación del corpus DONOC, se recomienda:

- Continuar incrementando el corpus con otras tipologías textuales y documentos procedentes de otros archivos y regiones del país.
- Elaborar un programa informático que posibilite la sistematización de toda la información incorporada y por incluir en el corpus, en aras de su generalización y aplicabilidad.

Bibliografía

- Acevedo Medina, Odette. 2014. Caracterización lingüística de documentos manuscritos de Santiago de Cuba siglo XVII. Trabajo de Diploma en opción al título de Licenciada en Letras. Santiago de Cuba: Universidad de Oriente.
- Alvar Esquerro, Manuel y Villena Ponsoda, J. A. 1994. Estudios para un corpus del español. Málaga: s.e.
- Ávila Muñoz, Antonio M; Lasarte Cervantes, María de la Cruz y Villena Ponsoda, Juan Andrés. 2008. El español hablado en Málaga II. Málaga. Editorial Sarría.
- Bidot Martínez, Irina. 2015. *Estudio de documentos manuscritos del siglo XVIII de la región oriental y su contribución a la historia de la variante cubana de la lengua*. Lengcom. Revista científica y divulgativa de lenguaje y comunicación, nro. 5.2. Publicado Viernes, 6 de febrero de 2015. <http://www.lengcom.es/revista-lengcom/>.
- Bidot Martínez, Irina y Acevedo Medina, Odette. 2015. Caracterización gráfico-fonética de documentos manuscritos del siglo XVII de Santiago de Cuba, en Comunicación social: Retos y perspectivas. Actas. XIV Simposio Internacional de Comunicación Social, Centro de Lingüística Aplicada, Vol. I, (150- 154), Santiago de Cuba.
- Cañizares, Fernando Diego. 1979. Teoría del Derecho. Guantánamo: Poligráfico "Juan Marinello".
- Castro Vida, Matilde (ed). 2008. El español hablado en Málaga I. Málaga: Editorial Sarría.

- CHARTA. s.f. Recuperado el 20 de octubre de 2017. Disponible en <http://www.redcharta.es>.
- Corpus.CHARTA s.f. Recuperado el 20 de octubre de 2017. Disponible en <http://www.corpuscarta.es>.
- Criterios de edición. s.f. Recuperado el 20 de octubre de 2017. Disponible en <http://www.redcharta.es/criteios-de-edicion>.
- Díez del Corral Areta, Elena y Martín Aizpurru, Leyre. 2014. Sin corpus no hay historia: la red CHARTA como un proyecto de edición común. Cuadernos de lingüística, vol2.
- Durán García, Rebeca. 2016. Caracterización lingüístico-formal de textos manuscritos de Baracoa en la segunda mitad del siglo XIX. Trabajo de Diploma en opción al título de Licenciada en Letras. Santiago de Cuba: Universidad de Oriente.
- El derecho notarial y sus perspectivas para el siglo XXI. Consultado 17 de abril de 2016. Disponible en [http:// www.prodigyweb.net.mx/bieletto](http://www.prodigyweb.net.mx/bieletto).
- El lenguaje jurídico y administrativo. Recuperado el 21 de abril de 2018 de, <http://educativa.catedu.es/44700165/aula/archivos/repositorio//4000/40/html/2ellenguajejurdicoyadministrativo.html>.
- Escalona Zambrano, Dallamy L. 2016. Análisis lingüístico-formal de cartas de asunto religioso de la segunda mitad del siglo XVIII en Santiago de Cuba. Trabajo de Diploma en opción al título de Licenciada en Letras. Santiago de Cuba: Universidad de Oriente.
- Investigación. s.f. Recuperado el 20 de octubre de 2017. Disponible en <http://redcharta.es/investigacion>.

- Jitka Hrusková. 2008. Los corpus CREA y CORDE en el contexto de los corpus lingüísticos. Tesis de diploma, Masarykova Univerzita.
- Labrada Hernández, Geisy. 2016. Caracterización lingüístico- formal de cartas de venta del siglo XVIII en Holguín. Trabajo de Diploma en opción al título de Licenciada en Letras. Santiago de Cuba: Universidad de Oriente.
- Martín Herrero, Cristina. s.f. Aproximación a ciertas perspectivas en lingüística de corpus. Salamanca: s.d.e
- Mijares Ramírez, Ivonne. 1997. Escribanos públicos y escrituras públicas en el siglo XVI. El caso de la Ciudad de México. Recuperado el 24 de abril de 2018. Disponible en <http://www.historicas.unam.mx/publicaciones/publicadigital/libros/escribanos/escribanos.html>.
- Notaria modelo. s.f. Recuperado el 1 de abril de 2018. Disponible en <http://naveiromodelonotarios.com/el-notario/los-documentos-notariales>.
- Objetivos. Recuperado el 20 de octubre de 2017. Disponible en <http://redcharta.es/que-es-charta-y-cuales-son-sus-objetivos>.
- Parodi, Giovanni. 2008. Lingüística de corpus: una introducción al ámbito, en revista de Lingüística Teórica Aplicada. Chile: s.d.e.
- Pérez Celis, Fernando. 2011. Catálogo de las escrituras notariales del siglo XVIII. Notarias 22, 25 y 352 de fondo antiguo, Sección Ordinaria, del Acervo Histórico del Archivo general de Notarias de la ciudad de México. Trabajo de diploma. Universidad Nacional Autónoma de México.

- Pérez Hernández, Chantal y Moreno Ortiz, Antonio. s.f. lingüística computacional y lingüística de corpus. Potencialidades para la investigación textual. Málaga: s.d.e
- Pérez Leyva, Geilys Zaday. 2017. Conectores argumentativos en testamentos del siglo XIX de Holguín y Santiago de Cuba. Trabajo de Diploma en opción al título de Licenciada en Letras. Santiago de Cuba: Universidad de Oriente.
- Propuesta de tipología textual. CHARTA. s.f. Recuperado el 20 de octubre de 2017. Disponible en <http://redcharta.es/tipologias-textuales>.
- Puñal Fernández, Tomás. 2002. Modelos diplomáticos de cartas de venta según algunos ejemplos extremeños del siglo XIV. España. Universidad de Extremadura.
- Qué tipos de documentos elaboran los notarios. Recuperado el 17 de abril de 2018. Disponible en <http://enayuda-notariosynotarias.com>.
- Ramírez Luengo, José Luis. 2011. *Un corpus para la historia del español en Nicaragua: edición de documentos oficiales del siglo XVIII (1704-1756)*. Moenia 17, 2011 (en prensa).
- Reglamento notarial (2013). Consultado 17 de abril de 2018. Disponible en <http://www.notariosyregistradores.com>.
- Safonst Caballero, Karen. 2017. Caracterización lingüística de documentos manuscritos en la segunda mitad del siglo XVIII de Santiago de Cuba. Trabajo de Diploma en opción al título de Licenciada en Letras. Santiago de Cuba: Universidad de Oriente.
- Sainz Padrón, Loisi. s.f. Estudio lingüístico de testamentos escritos en Cuba en los siglos XVI, XVII y XVIII. Estructura sintáctica de la voz pasiva. La Habana. Memorias de la VII Conferencia Internacional Lingüística 2011. La Habana: Instituto de Literatura y Lingüística.

- Sánchez, Aquilino. 2001. Gran Diccionario de uso del español actual. Madrid, SGEL.
- Segura Oduardo, Sucel Paula. 2015. Estudio lingüístico de documentos notariales de Jiguaní en la segunda mitad del siglo XIX. Trabajo de Diploma en opción al título de Licenciada en Letras. Santiago de Cuba: Universidad de Oriente.
- Torruella y Casañas, Joan. s.f. Los ejes principales en el diseño de un corpus diacrónico: el caso de CICA.s.d.e.
- Torruella, J y Llisterri, J.1999. Diseño de corpus textuales y orales, en Blecua, J.M; Clavería, G; Sanchez, C; Torruella, J (eds). Filología e informática. Nuevas tecnologías en los estudios filológicos. Barcelona. Editorial Milenio. Recuperado el 20 de noviembre de 2017. Disponible en http://liceu.aub.es/~joaquim/publicacions/torruella_llisterri_99.pdf.
- Verdejo Reyes, Pedro. 1983. Derecho notarial. Ciudad de La Habana: Pueblo y Educación.
- Yáñez, Rosa María. 1983. Historia General del Estado y el Derecho. La Habana. Ediciones Enspes.

Índice de Anexos

Anexo 1: Mapa con la ubicación de los territorios seleccionados.

Anexo 2: Tabla con los datos generales del corpus.

Anexo 3: Corpus de documentos notariales del oriente de Cuba (S XVII-XIX). Ver en el CD (soporte digital)

Anexo1



Leyenda

Rojo: Santiago de Cuba

Verde: Guantánamo

Azul: Holguín

Anexo 2

Datos generales del Corpus (DONOC)

Nro. del doc.	Tipología textual	Siglo	Archivo	Referentes generales
01	Carta de manumisión	XVIII	AHPSC	Carta de manumisión a favor de Candelaria, esclava de Doña Juana Ferrer
02	Carta de libertad	XIX	AHPG	Carta de libertad a favor de la negra Zamira, esclava de Simon Leguen
03	Carta de libertad	XIX	AHPG	Carta de libertad a favor del negro Salomon, esclavo de Simon Leguen
04	Carta de libertad	XIX	AHPG	Carta de libertad de la esclava mulatica Rosa por Juan Bautista Acebedo
05	Carta de libertad	XIX	AHPG	Carta de libertad a favor de la morena Ma Juana, esclava de Don Angel Gregorio Malletá.
06	Carta de libertad	XIX	AHPSC	Carta de libertad en favor de la negra María Asunción esclava de Gonzalo Blanco.
07	Carta de libertad	XIX	AHPSC	Carta de libertad a favor del negro Feliz Argáin, esclavo de Antonio de Herrera
08	Carta de libertad	XIX	AHPSC	Carta de libertad a favor de la negra Dolores, esclava de María Gertrudis Sorrilla.
09	Carta de venta	XVII	AHPSC	Carta de venta de tierras al Capitán Don de Castañeda por Esteban Ramos
10	Carta de venta	XVII	AHPSC	Carta de venta de un solar de Alonso Fernández Tames a Miguel Castellanos
11	Carta de venta	XVIII	AHPSC	Venta de dos esclavos del cirujano Josef Antonio Rodríguez a Don Ramon Correoso
12	Carta de venta	XVIII	AHPSC	Venta de solar de Francisco a Maria del Rosario Samor
13	Carta de venta	XVIII	AHPH	Carta de venta de una negra, María de Jesus, del Capitan Don Rodrigo González de Rivera a María de la Cruz
14	Carta de venta	XVIII	AHPH	Venta de del negro esclavo Francisco Joseph, propiedad de Pedro Jorge Moreno a el Capitan Don Baltasar de Fuentes
15	Carta de venta	XVIII	AHPH	Venta del hatu de las Cuebas de Doña María de la Cruz Echavarría a Don Calletano de Almaguer
16	Carta de venta	XVIII	AHPH	Venta del corral de Chaparra de Doña María Bello de Ávila al sargento mayor Don Miguel Calderón
17	Carta de venta	XVIII	AHPH	Venta de un hatu de Doña Ramira al Sargento Christobal Pupo

18	Carta de venta	XVIII	AHPH	Venta del negro Pedro, esclavo del regidor alferes mayor Don Diego de la Torre Chavarria, a Don Cristobal Rodríguez y González
19	Carta de venta	XVIII	AHPH	Venta del negro Antonio, esclavo de Roque de Aquino Miranda a Patricio Reynaldo
20	Carta de venta	XVIII	AHPH	Venta de la mulata Manuela de Doña Belnardinadlatore a Pablo Ramirez
21	Carta de venta	XVIII	AHPH	Venta de la negrita Juana esclava de Francisco Sablon a Pablo Ramirez
22	Carta de venta	XVIII	AHPH	Venta de negrito Julian, esclavo del teniente Don Diego de la Torre Chavarria, a su hija Doña Maria de la Cruz Chavarria
23	Carta de venta	XVIII	AHPH	Venta de la mulatica BarbaraCandiidaJsepha, esclava de Doña María del Prado a Don Lorenzo Castellanos Cisneros
24	Carta de venta	XVIII	AHPH	Venta de la mulata Catarina esclava de ChristobalSimon De Albaria a Esteban Rondón
25	Carta de venta	XVIII	AHPH	Venta de la negra Maia de los Dolores, esclava de Joseph de Leiva a Pablos Ramirez
26	Carta de venta	XVIII	AHPH	Venta del negro Juan, esclavo de LuizaRuis de Aguilera a Luis de Leiva
27	Carta de venta	XVIII	AHPH	Venta del hato de Managuaco de Don Miguel González de Rivera a Luis de Leiva
28	Carta de venta	XVIII	AHPH	Venta de parcela en el hato de Cacocum de Doña María de los Dolores Sarmiento a Joseph de Gaston
29	Carta de venta	XVIII	AHPH	Venta de parcela despoblada en el hato de Cacocum, de Juan Rodriguez a Gabriel Baptista
30	Carta de venta	XVIII	AHPH	Venta de una parte de hato de Holguín, de María de los Reyes Baptista a Joseph Ricardo
31	Carta de venta	XVIII	AHPH	Venta de una parcela del hato de Auras y San Marcos de Sebastian de la Cruz a Don Luis de Leyva
32	Carta de venta	XVIII	AHPH	Venta de unos colgadizos de Joseph de Leyva a Joseph de Leyva su hijo
33	Carta de venta	XVIII	AHPH	Venta del ingenio de Manuel Ricardo a Juan Ricardo, su hermano
34	Carta de venta	XVIII	AHPH	Venta del negro Francisco, esclavo de Christobal Rosales, a Joseph Gutierrez
35	Carta de venta	XVIII	AHPH	Venta de la negaMaria Guadalupe, esclava de Don Phelipe de Medina a Don Salvador Barsaga
36	Carta de venta	XVIII	AHPH	Venta de parecela en el hato de San Marcos de Luis Gonzalez De Rivera a Pedro de Leiva
37	Carta de venta	XVIII	AHPH	Venta del negro Nicolas, esclavo del Capitan Don Baltazar de Fuentes a Don Basilio González de la Peña
38	Carta de venta	XVIII	AHPH	Venta de la negra Josefina, esclava de Baltazar de Fuentes a Basiilio de Peña
39	Carta de venta	XVIII	AHPH	Venta del mulatico Vicente, esclavo de Don Luis González de Rivera a Don Antonio de Alvares

40	Carta de venta	XVIII	AHPH	Venta del hatu del Almirante de Diego y Luis Baptista al Sargento Don Miguel Calderón
41	Carta de venta	XVIII	AHPH	Venta de casa de Don Joseph de Peña al Capitan don Francisco Bermudez
42	Carta de venta	XVIII	AHPH	Venta del negro Pedro, esclavo de Doña YnesBapista, al teniente Don Miguel Cardel
43	Carta de venta	XIX	AHPG	Venta a Pedro Tomay Bon del negro Lorenzo Cartemanuel, propiedad de D. José Ant. Catá
44	Carta de venta	XIX	AHPG	Venta de un solar de D. Juan Bautista Soler y D. Juan Cómas a Catalina Cadena
45	Carta de venta	XIX	AHPG	Venta de parte de un almacén de don Bartolomé Dihins, a don Augusto Samanos
46	Carta de venta	XIX	AHPG	Venta de las esclavas Petrona, Florencia y Salustina, de Santiago Fournier, a Matea Fournier
47	Carta de venta	XIX	AHPSC	Carta de venta un Solar Yermo de José Ant ^o . de la cuesta a José Dias
48	Carta de venta	XIX	AHPSC	Carta de venta de la venta de un colgadiso de madera y Texa de Bernardina Margarita y José Trinidad de Medina a Juan Crisostomo Ramos
49	Carta de venta	XIX	AHPSC	Carta de venta de una negra nombrada Maria del Carmen casta lucumi de Andres de Rovina a D ^a . Manuela Collazo
50	Carta de venta	XIX	AHPSC	Carta de venta de una negra nombrada Rita criolla de Sto. Domingo de D ^a . Fra ^{ca} . Estévez al convento de Ntra. Sra. de Velem
51	Carta de venta	XIX	AHPSC	Carta de venta de mulato Manuelcriollo de Ma Magdalena Mediaseja a D José Herrera
52	Carta de venta	XIX	AHPSC	Carta de venta de negro criollo de Ma Magdalena Mediaseja a José Franco
53	Poder	XVII	AHPSC	Poder al Lic. Don Manuel Cabral
54	Poder	XVIII	AHPSC	Poder de D. ⁿ NicolasAnt. ^o de Rubalcaba, Paz, Vargas a D. ⁿ Fran. ^{co} Xav. ^{er} de RubalcavaSanches de Car Mona, su hijo
55	Poder	XVIII	AHPSC	Poder de Antonio Palacios Saldiertum a D. ⁿ Luis de Leiba
56	Poder	XVIII	AHPSC	Poder de D. ⁿ Luis Nasario Ferrer al cav. ^o Rex. ^{of} Depocit. ^o gral.cap. ⁿ d. ⁿ Miguel Ferrer, al Abog. ^{do} d. ⁿ Miguel Ferrer, y a D. ⁿ Ant. ^o de Roxas
57	Poder	XVIII	AHPSC	Poder de Micaela Ferrera a Fran ^{co} Antonio Caballero
58	Poder	XVIII	AHPSC	Poder de María Guadalupe del Castillo, viuda del Sarg ^{to} . Seg ^{do} .de GranadinosManu ^l . del Barrio al Sub ^{te} . D Mariano Ferrer
59	Poder	XIX	AHPG	Poder de D Jose M ^a Escalera a D. NicolasSustrillo
60	Poder	XIX	AHPG	Poder general de D José Bertran y Miret a D. José Antonio Cabrado

61	Poder	XIX	AHPG	Poder especial de Dn Juan Francisco Carmillerí al Lic.Dn Francisco Sorzano Abogado del Ilustre colegio de Santiago de Cuba
62	Testamento	XVII	AHPSC	Testamento de Pedro de Oliva y su mujer Ynes de avila
63	Testamento	XVIII	AHPSC	Testamento de Agustin Rodríguez y María Manuela Arías, marido y mujer, morenos libres
64	Testamento	XVIII	AHPSC	Testamento de D.ª Gerbacia de Sevedes
65	Testamento	XIX	AHPG	Testamento de don Salvador Rosell
66	Testamento	XIX	AHPG	Testamento de Da Manuela Parra
67	Testamento	XIX	AHPSC	Testamento de Nicolas Rigores y María Josefa Sanchez
68	Testamento	XIX	AHPH	Testamento de Cayetano Aysen, moreno libre
69	Testamento	XIX	AHPH	Testamento de Doña Maria Luisa Gonzales
70	Testamento	XIX	AHPH	Testamento de Bernardina de la Concepcion Alvares
71	Testamento	XIX	AHPH	Testamento de Dña Maria del Rosario Ynfante
72	Testamento	XIX	AHPSC	Testamento de la morena Ana Robert y Teixo
73	Testamento	XIX	AHPSC	Testamento de la morena María Josefa Llanyaque sin otro apellido
74	Testamento	XIX	AHPSC	Testamento de Francisca Guillet, sin otro apellido
75	Testamento	XIX	AHPSC	Testamento de D Francisco Salvador Marfulli
76	Testamento	XIX	AHPSC	Testamento del moreno JoaquinPortuondo Peralta
77	Testamento	XIX	AHPSC	Testamento de Luisa duverger

AHPSC: Archivo histórico provincial de Santiago de Cuba

AHPH: Archivo histórico provincial de Holguín

AHPG: Archivo histórico provincial de Guantánamo

Nota: se han respetado las grafías de los nombres propios que aparecen en el corpus.

Anexo 3

Corpus de documentos notariales del oriente de cuba (DONOC). Ver en el CD (soporte digital).