

令和元年5月29日現在

機関番号：12601

研究種目：基盤研究(C) (一般)

研究期間：2016～2018

課題番号：16K02657

研究課題名(和文)古スペイン語文書の相対年代学

研究課題名(英文)Relative chronology of old Spanish documents

研究代表者

上田 博人(Ueda, Hiroto)

東京大学・大学院総合文化研究科・特任研究員

研究者番号：20114796

交付決定額(研究期間全体)：(直接経費) 2,400,000円

研究成果の概要(和文)：多変量統計手法を用いてスペイン語史の諸相を実証的に明らかにする。資料はスペイン、アルカラ大学の研究グループ(Pedro Sanchez-Prieto Borja代表)が作成した「1700年以前のスペイン語文書コーパス」Corpus de Documentos Espanoles Anteriores a 1700(計1500文書)と、上田が作成したデジタル資料「手稿・印刷本の文字」Letras Espanoles en Manuscritos e Impresos(計28冊)を用いる。年代ごとに抽出された言語特徴はスペイン語史研究のための基礎的な資料になる。

研究成果の学術的意義や社会的意義

従来のスペイン語研究の成果には、特定の資料について、その言語特徴を網羅的に集めたものが多い。幅広く言語史を俯瞰する研究もあるが、ここでは特定の特徴を例としてあげるだけで、その統計的な検証は行われていない。頻度や分散さえも提示されていないこともある。また、資料については、多くの場合、すでに刊行された本や資料集が使われているため、その表記の基準が異なり、厳密な比較を困難にしている。

本研究では、従来の研究の方法と資料の扱い方の不備を補うべく、方法に情報科学と数理科学の成果を取り入れ、厳密な転写方式に統一した、新しい文献学の資料を駆使し、研究成果を順次発表していく。

研究成果の概要(英文)：We study the different aspects of the history of the Spanish language with multivariate statistical methods. We use the materials of two corpora: "Corpus of Spanish Documents Prior to 1700"; (research group of the University of Alcalá; directed by Pedro Sanchez-Prieto Borja) and "Spanish Letters in Manuscripts and Printed Documents" (Hiroto Ueda). The features extracted from these materials according to the chronological order serve as fundamental data for studies of the Spanish language.

研究分野：スペイン語学、言語データ分析

キーワード：スペイン語学 言語データ分析 古スペイン語文書 相対年代学

## 様式 C - 19、F - 19 - 1、Z - 19、CK - 19 (共通)

### 1. 研究開始当初の背景

従来のスペイン語研究の成果には、特定の資料について、その言語特徴を網羅的に集めたものが多い。幅広く言語史を俯瞰する研究もあるが、そこでは特定の特徴を例としてあげるだけで、その統計的な検証は行われていない。頻度や分散さえも提示されていないこともある。

一方、現在では各種の大型言語コーパスが作成されており、これら进行处理する言語データ分析プログラムを作成すれば、数量的な側面を統計的に分析して言語の実態をより適切に示すことができるはずである。

### 2. 研究の目的

多変量統計手法を用いてスペイン語史の諸相を実証的に明らかにする。資料はスペイン、アルカラ大学の研究グループ(Pedro Sanchez-Prieto Borja 代表)が作成した「1700年以前のスペイン語文書コーパス」Corpus de Documentos Espanoles Anteriores a 1700(計1500文書)と、上田が作成したデジタル資料「手稿・印刷本の文字」Letras Espanoles en Manuscritos e Impresos(計28冊)を用いる。前者は公証文書であり、発行年・発行地が明示されているものが多く、後者は文学、法律、文法書などの書籍であり、写本で伝わっている本の中には年代が確定できないものがある。そこで、前者から年代ごとの偏在的言語特徴を抽出し、統計的に後者の相対年代を推定する。年代ごとに抽出された言語特徴はスペイン語史研究のための基礎的な資料になる。

また、資料については、多くの場合、すでに刊行された本や資料集が使われているため、その表記の基準が異なり、厳密な比較を困難にしている。

本研究では、従来の研究の方法と資料の扱い方の不備を補うべく、方法に情報科学と数理科学の成果を取り入れ、厳密な転写方式に統一した、新しい文献学の資料を駆使し、研究成果を順次発表していく。

資料については、2015年度に、先述のスペイン・アルカラ大学の「1700年以前のスペイン語文書コーパス」と、東京大学・上田研究室の「手稿・印刷本の文字」が完成したので、これらを使用する。

分析法では、正規表現を拡張した高度な検索法を導入する。出力には、絶対頻度、相対頻度だけでなく、さまざまな統計的手法を応用し、グラフを使って視覚化する。

スペイン語言語史を通じて、ほとんど全文書に存在する以下の言語特徴を検出し、そこから、言語年代推定のための変数表を作成する。

偏在言語特徴として、以下の書記法を取り上げる(C:子音、V:母音、<...>略字): (1) 語頭の rr- / R- / r-, (2) ff- / f- / h- / 消失, (3) ss- / s-, (4) -ss- / -s-, (5) x / j, (6) u / v / b, (7) i / j / y, (8) ç / c / z, (9) 語尾 e の過度の脱落とその復元形(part / parte), (10) pora / para, (11) qua- / cua, (12) -t / -d, (13) [mnu] + j / [mnu] + i, (14) nn / <n>n / n<n> / n, (15) <n>C / nC, (16) q<uV> / quV, (17) C<VrV> / C(VrV), (18) <e>l / el, (19) d<e> / de, (20) gue-gui / gue-gui, (21) et / & / e / y, (22) non / no<n> / no, (23) (aldelen)+定冠詞, (24) aver / haber (変化形を含む)。

本研究では、このような年代変遷をすべて探し出し、その頻度表を作成する。その後、これに多変量解析を適用して各作品の相対的年代を推定し、それに従って新たな年代推移を記述する。同時に、言語特徴を年代順に並べ、その順次的変遷を辿りながら、分布の連続性・平滑性・単峰性・正規性の度合いを確認する。さらに、これまで独自に開発してきた数量分析法の有効性を確認し、言語科学に限らず他分野への応用可能性について考察する。

「手稿・印刷本の文字」は次の資料を含む: Anonimo: Poema de mio Cid. 1207. Biblioteca Digital Hispanica. (1r-12v) [BNE: Vitr/7/17] // Fazienda de Ultramar. ca1200 (1r-7r) // Fuero de Alcala. ca1230. (1r-17v) // Alfonso X: General Estoria. c.1270. (1r-4v) // Libro de Alexandre. c1300 (1r-10v) // Anonimo. Historia troyana. [BNE: ms 10.146] (63r-63v, 68r-69v, 72r-74v) // Arcipreste de Hita: Libro de buen amor. 1389. Real Academia Espanola. (1r-12v) // Don Juan Manuel: El Conde Lucanor. ca.1330. Copia del S. XV. (26ra-131vb) // Rabi Moše Arragel: Biblia de Alba. 1433. (26ra-33vb) // Especulo de los legos. 1450 [BNE. ms.94] (1r-8v) // Antonio de Nebrija: Gramatica de la lengua castellana. Salamanca, 1492. (1r-7v) // Fernando de Rojas: La Celestina. Burgos, 1499. (A1r-B1v) // Gonzalo Fernandez de Oviedo. 1514. Sumario de la natural historia de las Indias. (3r-7v) // Juan de Valdez, 1535?. Dialogo de la lengua. (1r-19v.) // Anonimo: Lazarillo de Tormes. Medina del Campo, 1554. (A2r-C1v) // Fray Luis de Leon. 1583. La perfecta casada. Salamanca. [BNE: U352(2)] (2r-13r) // Miguel de Cervantes: Don Quixote de la Mancha, Madrid, 1605. (1r-9r) // Francisco de Quevedo: Historia de la vida del Buscon. Zaragoza, 1626. (1r-12r) // Baltasar Gracian: El criticon. 1651. (p.1-26) // Pedro Calderon de la Barca. Autos sacramentales alegoricos y historiales. (1672; 1677) (p.399-411) // Francisco Bances Cardamo. El Austria en Jerusalem. 1704. (p. lxi-lxx) // Comedia famosa, 1747, El picarillo en Espana de don Joseph Canizares (p.1-11) // Gaspar Melchor de Jovellanos. 1787. Delincuente honrado. (p.1-37) // Ortografia de la lengua castellana, compuesta por la Real Academia Espanola. 1815. (p.1-14) // Jose de Espronceda, El diablo mundo. (p.1-24) // Pedro Antonio de Alarcon, El sombrero de tres picos. 1874. (p.2-13) // Benito Perez Galdos, Dona perfecta, (1876; 1899) (p.5-28).

### 3. 研究の方法

大規模言語コーパス「1700年以前のスペイン語文書コーパス」(CODEA 計1500文書)と、上田が作成した手稿・印刷本デジタル資料「手稿・印刷本の文字」(LEMI 計28冊)の内容を精査

し、発行地域と年代を確認する。CODEA における偏在言語特徴の分布を調べ、年代推定のための変数(パラメータ)を求める。言語特徴と年代の相関関係を見る。多変量解析を応用した新しい相対年代を推定する。方法論と適用の結果を発表する。

古文書の文法タグ付与(動詞・形容詞などの変化形から代表形を導出する方法)を完成する。

偏在言語特徴として、以下の書記法を取り上げる(C:子音、V:母音、<...>略字):(1) 語頭の rr-/R-/r-, (2) ff-/f-/h-/ 消失, (3) ss-/s-, (4) -ss-/s-, (5) x/j, (6) u/v/b, (7) i/j/y, (8) ç/c/z。

各種統計量の逸脱の程度を確率で示し、その有意性を確認する。従来の方法では、平均差や相関係数の有意性を示す確率が扱われてきたが、本研究で提案する、単峰性・連続性・正規性などや各種の限定値についても、確率によって有意性を示す。

偏在言語特徴として、以下の書記法を取り上げる(C:子音、V:母音、<...>略字):(7) i/j/y, (8) ç/c/z, (9) 語尾 e の過度の脱落とその復元形(part / parte), (10) pora / para, (11) qua- / cua, (12) -t / -d, (13) [mnu] + j / [mnu] + i, (14) nn / <n>n / n<n> / n, (15) <n>C / nC, (16) q<uV> / quV。

成果を国際学会・国内研究会・インターネットで発表する。

#### 4. 研究成果

古スペイン語のそれぞれの語形を古文書形、校訂形、現代語形に対応させたファイルを作成し、現代語形から辞書の見出し語にあたる代表形で検索できるデータベースの資料作成方法を開発した。これによってこれまで、個別に検索していた言語特徴を全体的に一律に対応関係を出力させることができるようになった。これを旧カスティージャ地方の文書について実行し、インターネットで公開した。この方法は汎用性が確保されるので、現在はマドリード自治体の文書に活用させている。

#### 5. 主な発表論文等

[雑誌論文](計 10 件)

2019a. Francisco Moreno Fernandez / Hiroto Ueda. "Cohesion and Particularity in the Spanish Dialect Continuum", in *Open Linguistics*.

<https://www.degruyter.com/downloadpdf/j/opli.2018.4.issue-1/opli-2018-0035/opli-2018-0035.pdf>

2018a. Hiroto Ueda. "Tratamiento linguistico y matematico de textos digitales espanoles. Presentacion del Programa LEXIS-web", *Actas del IX Congreso de la Asociacion Asiatica de Hispanistas (Bangkok, 2016)*, pp. 617-630

[http://www.sinoele.org/images/Revista/17/monograficos/AAH\\_2016/AAH\\_2016\\_hiroto\\_ueda.pdf](http://www.sinoele.org/images/Revista/17/monograficos/AAH_2016/AAH_2016_hiroto_ueda.pdf)

2018b. Hiroto Ueda. "Letras espanolas y letras japonesas. Distribucion de frecuencias e intenciones semanticas", en Toshihiro Takagaki (ed.) *Exploraciones de la linguistica contrastiva espanol-japones*, UAM Ediciones, pp. 169-190.

2017a. Hiroto Ueda. "Un esbozo historico de las formas abreviadas espanolas con indices de continuidad y de suavidad" *Actas del II Congreso Internacional sobre el espanol y la cultura hispanica del Instituto Cervantes de Tokio (2015)*, pp.24-44.

[http://cvc.cervantes.es/ensenanza/biblioteca\\_ele/publicaciones\\_centros/PDF/tokio\\_2015/06\\_ueda.pdf](http://cvc.cervantes.es/ensenanza/biblioteca_ele/publicaciones_centros/PDF/tokio_2015/06_ueda.pdf)

2017b. 上田博人「del と al の歴史的形成」『日本イスパニヤ学会会報』24,巻頭言, pp. 2-3. ("Formacion historica de «del» y «al»", *Boletin de la Asociacion Japonesa de Hispaista*, 24, pp. 2-3.)

2017c. Hiroto Ueda. "Two statistical treatments of Spanish vocabulary: composite indices of frequency and dispersion and principal component analysis applied to ordinal frequencies", *Dialectologia, special issue*, 7, pp.187-227.

<http://www.publicacions.ub.edu/revistes/dialectologiaSP2017/>

2017d. Hiroto Ueda. "Formacion historica de «del» y «al» (version ampliada)", *90 Anos de la Academia Boliviana de la Lengua*, La Paz, pp. 147-151.

2016a. Hiroto Ueda. "Analizador linguistico comun con reglas gramaticales y diccionario, preparados por el usuario: Una aplicacion para el analisis tipologico del lexico espanol", *Linguistica Espanola Actual*, 37/2, pp. 241-264.

2016b. Hiroto Ueda. "El nacimiento de la letra jota como grafia consonantica", en Johannes Kabatek (ed.) *Linguistica de corpus y linguistica historica iberorromana*, Walter de Gruyter, Nördlingen (Germany), pp. 299-321, con Maria Jesus Torrens.

2016c. Hiroto Ueda. "La grafia doble <ff> en el castellano medieval: Similitud paleografica y linguistica con la doble <ss> alta", *Revista de Historia de la Lengua Espanola*, 10, pp. 105-132.

[学会発表](計 24 件)

2019a. Hiroto Ueda. "Caracteristicas estadisticas de las letras japonesas, hiragana, katakana y kanji: en

- compracion con el alfabeto espanol", Conferencia en la Universidad Autonoma de Madrid, 25 de febrero de 2019
- 2019b. Pedro Sanchez-Prieto Borja / Hiroto Ueda. "El proyecto ALDICAM-CM: del archivo al atlas", en *Jornadas. La lengua de Madrid, ayer y hoy*, Universidad de Alcalá, 1 de marzo de 2019.
- 2018a. Antonio Moreno Sandoval / Hiroto Ueda. "Union y separacion de preposicion y articulo definido del espanol. Observaciones con la frecuencia probabilistica en el analisis de Pareto", X Congreso Internacional de Linguistica de Corpus, Caceres, Espana, 10, mayo, 2018.  
<https://lecture.ecc.u-tokyo.ac.jp/~cueda/kenkyu/rekisi/prep-art/prep-art-slide.pdf>
- 2018b. Hiroto Ueda. "Origen de 'del', 'al' y 'el' del espanol. Distintas frecuencias en el analisis de Pareto", Seminario del proyecto UAM-Santander. Universidad de Estudios Extranjeros de Tokio, 21 mayo, 2018  
<https://lecture.ecc.u-tokyo.ac.jp/~cueda/kenkyu/rekisi/prep-art/del-al-el-slide.pdf>
- 2018c. Hiroto Ueda. "Metodos de linguistica espanola de corpus en linea. Medidas de diversidad lexica con evidencias documentales", con Antonio Moreno Sandoval, CANELA. Confederacion Academica Nipona, Espanola y Latinoamericana, Universidad Doshisha, Kioto, 26/mayo/2018.  
<https://lecture.ecc.u-tokyo.ac.jp/~cueda/kenkyu/goi/diversidad-lexica/diversidad-lexica-slide.pdf>
- 2018d. Hiroto Ueda. "Measures of frequency and dispersion in individual texts. Applied to the morphology of future of Spanish verbs in space and time", IXth Congress of the International Society for Dialectology and Geolinguistics, 23-27 July 2018, Vilnius, Lithuania.  
<https://lecture.ecc.u-tokyo.ac.jp/~cueda/kenkyu/rekisi/pondre-frec/pondre-frec-slide.pdf>  
<https://lecture.ecc.u-tokyo.ac.jp/~cueda/kenkyu/rekisi/pondre-frec/pondre-frec.pdf>
- 2018e. Leyre Martin Aizpuru / Hiroto Ueda. "Fonema oclusivo velar sordo y sus grafias correspondientes en espanol. Variaciones y cambios de 'k', 'c', 'ch', 'qu' en espacio, tiempo y tipos textuales", XI Congreso Internacional de Historia de la Lengua Espanola, Pontificia Universidad Catolica del Peru, 6-10 de agosto, 2018.  
<https://lecture.ecc.u-tokyo.ac.jp/~cueda/kenkyu/rekisi/qua/lima2018.pdf>
- 2018f. Hiroto Ueda. "Grandes rasgos historicos de grafias espanolas. Metodos de cuantificacion y visualizacion de frecuencias de los datos de ALDICAM", *III Congreso Internacional sobre el Espanol y la Cultura Hispanica en Japon*, Instituto Cervantes de Tokio, 6 de octubre de 2018.  
<https://lecture.ecc.u-tokyo.ac.jp/~cueda/kenkyu/rekisi/rasgos-historicos/rasgos-historicos-slide.pdf>  
<https://lecture.ecc.u-tokyo.ac.jp/~cueda/kenkyu/rekisi/rasgos-historicos/rasgos-historicos.pdf>
- 2018g. Hiroto Ueda. "Formacion historica de la conjuncion espanola 'y'. Adaptacion al modelo de los diptongos |i| + vocal y su latencia en siglos medievales", *Circulo de Estudios Linguisticos Hispanicos de Tokio*, Universidad de Estudios Extranjeros de Tokio, 10 de noviembre de 2018  
<https://lecture.ecc.u-tokyo.ac.jp/~cueda/kenkyu/ekisi/e-y/e-y-slide.pdf>  
<https://lecture.ecc.u-tokyo.ac.jp/~cueda/kenkyu/rekisi/e-y/e-y.pdf>
- 2018i. Hiroto Ueda. "Variacion geografica en los cambios historicos del espanol. La conjuncion en Madrid y la apocope extrema en Espana ", *Circulo de Estudios Linguisticos Hispanicos de Tokio*, Universidad de Estudios Extranjeros de Tokio, 10 de noviembre de 2018  
<https://lecture.ecc.u-tokyo.ac.jp/~cueda/kenkyu/rekisi/geografia/geografia-slide.pdf>
- 2018j. Hiroto Ueda. "Cambios historicos observados en los documentos de Madrid. Cuantificacion y visualizacion de los datos linguisticos de ALDICAM" *III Jornadas Madrid en sus documentos. Proyecto «Atlas Linguistico Diacronico e Interactivo de la Comunidad de Madrid»*, Universidad de Alcalá, 13 de diciembre de 2017  
<https://lecture.ecc.u-tokyo.ac.jp/~cueda/kenkyu/rekisi/cambios/cambios-slide.pdf>
- 2017a. Hiroto Ueda. "Recursos informaticos en web para el analisis y la ensenanza de espanol. Analisis estadisticos", Instituto Cervantes en Boston, Estados Unidos, 23 marzo, 2017.
- 2017b. Hiroto Ueda. "Queismo y dequeismo observados en los datos de VARIGRAMA en el sistema LYNEAL. Aproximacion a la realidad de variacion sintactica espanola a traves de analisis multivariados pluridimensionales", *Spanish Dialect Syntax*, Universidad Complutense de Madrid, 21 de abril de 2017.
- 2017c. Hiroto Ueda. "Union y separacion de preposicion y articulo definido del espanol. Su estructura y frecuencia en espacio y tiempo", Taller/Workshop PROGRAMES: Gramaticalizacion, lexicalizacion y analisis del discurso desde una perspectiva historica (cuestiones metodologicas e instrumentales), Universidad Complutense de Madrid, 5 de mayo de 2017.
- 2017d. Hiroto Ueda. "Metodos linguisticos basados en evidencias documentadas. DEL y AL en espanol", 28 de mayo de 2017, CANELA (Confederacion Academica Nipona, Espanola y Latinoamericana), Instituto Cervantes en Tokio.
- 2017e. Hiroto Ueda. "Las grafias <u>, <v> y <b> a lo largo de la historia del espanol. Analisis separado de frecuencias y analisis conjunto multivariante", V Congreso Internacional de la Red

- CHARTA. 14 de junio, 2017. Universidad de Lausana.
- 2017f. Hiroto Ueda. "Unilateral Correspondence Analysis. Applied to Spanish linguistic data in time and space", Sixteenth International Conference on Methods in Dialectology (METHODS XVI) National Institute for Japanese Language and Linguistics, Tokyo, 10 August, 2017.  
<http://lecture.ecc.u-tokyo.ac.jp/~cueda/kenkyu/chiri/correspondence/correspondence.pdf>
- 2017g. Leyre Martin Aizpuru / Hiroto Ueda. "Parametros multivariantes para el tratamiento de datos linguisticos digitales (LYNEAL): puntuacion manuscrita con funcion discursiva en documentacion cancilleresca (siglo XIII)", 3er Congreso de Humanidades Digitales Hispaicas, 18 al 20 de octubre de 2017, Malaga, Espana.
- 2016a Hiroto Ueda. "Tratamiento linguistico y matematico de textos digitales espanoles: Presentacion del Programa LEXIS-web", IX Congreso Internacional de la Asociacion Asiatica de Hispanistas, Universidad de Chulalongkorn, Bangkok, Tailandia, del 22 al 24 de enero de 2016  
<a href="zyoho/tratamiento-slide.pdf">
- 2016b Hiroto Ueda. "Metodos de geografia linguistica con estadistica multivariante. Analisis de asociacion", Charla en la Universidad de Barcelona (14 de marzo de 2016)
- 2016c. Hiroto Ueda. "Reglas de relacion entre fenomenos linguisticos y parametros espacio-temporales" Jornada del Proyecto del Corpus Linguistico UAM, U. de Tokio y TUFS, Universidad de Estudios Extranjeros de Tokio. 2016/5/12.
- 2016d. Hiroto Ueda. "LYNEAL: Presentacion de la herramienta de analisis de corpus", Jornada sobre Linguistica y Filologia Digitales, Universidad Autonoma de Madrid, 2016/9/14.
- 2016e. Hiroto Ueda. "Aplicacion del software LYNEAL a la lexicologia medica", Carlos Herrero Zorita, Hiroto Ueda, Antonio Moreno Sandoval, Jornada sobre Linguistica y Filologia Digitales, Universidad Autonoma de Madrid, 2016/9/14.
- 2016f. Antonio Moreno Sandoval / Hiroto Ueda. "Presentacion de programas en web para analisis linguisticos generales: LETRAS-web y NUMEROS-web", TechLing'16. Primeiro Congresso Internacional em Linguas, Linguistica e Tecnologia. Univerisidade do Minho. 2016/10/12.
- 2016g. Hiroto Ueda. "Archivo linguistico integral y corpus CODEA y CHARTA", Encuentro Cientifico de la Red Internacional CHARTA. Universita di Padova. 2016/10/15.

〔図書〕(計 件)

〔産業財産権〕

○出願状況(計 件)

名称：  
発明者：  
権利者：  
種類：  
番号：  
出願年：  
国内外の別：

○取得状況(計 件)

名称：  
発明者：  
権利者：  
種類：  
番号：  
取得年：  
国内外の別：

〔その他〕

ホームページ等

## 6 . 研究組織

(1)研究分担者

研究分担者氏名：

ローマ字氏名：

所属研究機関名：

部局名：

職名：

研究者番号（8桁）：

(2)研究協力者

研究協力者氏名：

ローマ字氏名：

科研費による研究は、研究者の自覚と責任において実施するものです。そのため、研究の実施や研究成果の公表等については、国の要請等に基づくものではなく、その研究成果に関する見解や責任は、研究者個人に帰属されます。