

# Acercamiento desde la lengua vasca al estudio del variacionismo lingüístico cuantitativista

Approach from the Basque language to the study of  
quantitativist linguistic variationism

Juan Abasolo Isasa

*Universidad del País Vasco / Euskal Herriko Unibertsitatea, España*  
<https://orcid.org/0000-0002-1911-4118>

e-mail: [juan.abasolo@ehu.eus](mailto:juan.abasolo@ehu.eus)

Ariane Ensunza Aldamizetxebarria

*Universidad del País Vasco / Euskal Herriko Unibertsitatea, España*  
<https://orcid.org/0000-0003-0908-7467>

e-mail: [ariane.ensunza@ehu.eus](mailto:ariane.ensunza@ehu.eus)

Recibido: 15/09/2021  
Aprobado: 22/03/2022

## RESUMEN

Mucho antes de la creación de la Universidad del País Vasco (UPV/EHU) ya se habían realizado tantísimos acercamientos al estudio del euskera, lengua vasca, y sus dialectos, Wilhelm von Humboldt puede ser citado; y es de obligada citación Louis Loucien Bonaparte (1863). Pero es durante el final del siglo XX, y más aún, en el corriente nuevo milenio que, acompañada por un nuevo estatus legal, la lengua vasca conoce un gran desarrollo en la academia, tanto en el estudio acerca de la misma lengua desde diferentes perspectivas de las ramas de la lingüística, como también un significativo desarrollo del conocimiento de corte más científico o social desarrollado en esta lengua (Artetxe-Sanchez, 2011; Agirre, 2021). Todo este desarrollo, principalmente acompañado por la UPV/EHU, también fructifica en otras universidades del territorio vasco, separado entre dos estados y presente en tres entes administrativos principales. En este marco es donde encuentra fermento el estudio del variacionismo dialectal (Aurrekoetxea et al., 2013) desde una perspectiva objetiva y científica. Se realizan diferentes estudios basados en datos recogidos desde la época post-napoleónica (Aurrekoetxea & Videgain, 2009) pasando por otros recogidos en fechas recientes (Aurrekoetxea et al., 2019), hasta los últimos recibidos de boca de los estudiantes de magisterio, futuros formadores de los niños vascos, de todo el territorio (Iglesias et al., 2020). El desarrollo metodológico permite que hoy en día parte de estos estudios sea referente en sucesivos estudios de otras lenguas (Clua & Lloret, 2015; Planas et al. 2015) y también en la perspectiva variacionista de carácter sociolingüístico (Unamuno & Aurrekoetxea, 2013), en el que se estudia la influencia de la variedad estándar/escolar en las características de las hablas locales.

*Palabras clave:* variacionismo, educación superior, diglosia, bilingüismo, lengua vasca.

**ABSTRACT**

Long before the creation of the University of the Basque Country (UPV/EHU), many approaches to the study of Euskera, the Basque language, and its dialects had already been made, Wilhelm von Humboldt can be cited; and Louis Loucien Bonaparte (1863) is also worth mentioning. But it is during the end of the 20th century, and even more so, in the current new millennium that, accompanied by a new legal status, the Basque language underwent a great development in the academy, both in the study of the language itself from different perspectives of the branches of linguistics, as well as a significant development of knowledge of a more scientific or social nature developed in this language (Artetxe-Sanchez, 2011; Agirre, 2021). All this development, mainly accompanied by the UPV/EHU, also bears fruit in other universities in the Basque territory, separated between two states and present in three main administrative entities. It is in this framework that the study of dialect variationism (Aurrekoetxea et al., 2013) finds ferment from an objective and scientific perspective. Different studies are carried out based on data collected from the post-Napoleonic period (Aurrekoetxea & Videgain, 2009) through others collected recently (Aurrekoetxea et al., 2019), to the latest received from the mouths of teacher training students, future trainers of Basque children, from all over the territory (Iglesias et al., 2020). The methodological development allows that nowadays part of these studies is a reference in successive studies of other languages (Clua & Lloret, 2015; Planas et al. 2015) and also in the variationist perspective of sociolinguistic character (Unamuno & Aurrekoetxea, 2013), in which the influence of the standard/school variety on the characteristics of local speech is studied.

*Keywords:* variationism, higher education, diglossia, bilingualism, Basque language.

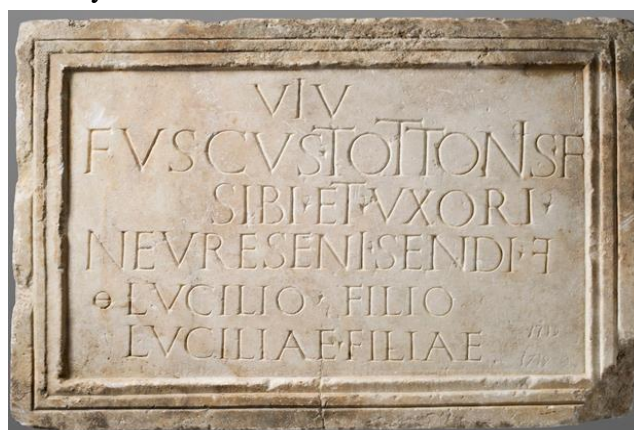
**INTRODUCCIÓN****LA HISTORIA**

El euskera o lengua vasca es el idioma hablado por los vascos y en el País Vasco, situado en ambas vertientes de los Pirineos y a la orilla del mar cantábrico.

Historiadores y geógrafos griegos y romanos (Estrabón, Ptolomeo, Plinio) del S. I nos hablan de la existencia de tribus en los territorios del País Vasco actual que no se asemejan a los demás pueblos ni física ni lingüísticamente (Estrabón, Geografía).

Las primeras palabras testimoniadas del euskera datan del S. I-II-III, donde encontramos unos 400 nombres propios y 70 nombres de deidades escritas en estelas funerarias romanas. Por ejemplo el de la figura 1, se halló en Arrien-en-Bethmale en la Aquitania histórica, hoy en día departamento de Ariège, Francia. Allí puede leerse NEVRE SENI SENDI que es transparente y perfectamente interpretable desde el euskera<sup>1</sup>.

**Figura 1:** Inscripción vasca y romana data del 27 AC



*Anonimus & François-Louis Pons (fotógrafo),  
Musée Saint-Raymond. CC BY-SA 4.0*

<sup>1</sup> También se han encontrado estelas similares al sur-oeste de los Pirineos, como en Navarra, Guipúzcoa, Álava, Vizcaya, Rioja, Soria, Aragón y Lleida (Cataluña). Luego tenemos 6 siglos de oscuridad y silencio.

Las primeras frases escritas en euskera datan del S. X-XI, que son dos frases subordinadas muy cortas encontradas en un manuscrito de San Millán de la Cogolla (Rioja, España). El manuscrito está escrito en latín, pero tiene dos notas aclaratorias en los márgenes escritos en euskera: *izioqui dugu* [nos alegramos] y *guez ajutuezdugu* [nosotros no tenemos adecuado].

En el siglo XII encontramos el primer listado de palabras en el diccionario *Liber Sancti Jacobi* (1140) que preparó el francés Aymeric Picaud para los peregrinos que iban a Santiago de Compostela.<sup>2</sup> Tres siglos más tarde (S. XV) el caballero alemán Arnold von Harff también redactó un diccionario corto al visitar el mismo santuario. Sin embargo, lo que se considera el primer diccionario vasco llega en 1562 de la mano del italiano Nicolao Landuchio, que escribió el *Dictionarium linguae cantabricae* que desafortunadamente no vio la luz hasta el S. XX (1958, in Michelena et al., 2011). Cabe destacar la procedencia de los autores: francia, alemania, sicilia, toscana.

Por otro lado, también se cuentan con algunas cartas escritas en vasco entre los S. XIV y XVI<sup>3</sup>. Se puede observar aquí que el vasco no solamente se utilizaba para comunicaciones familiares, sino también para comunicaciones de índole estratégica.

Con el Concilio de Trento (1545-1563) se impulsa la enseñanza de la doctrina en los idiomas locales y se observa un auge de obras escritas en vasco. De hecho, el primer libro impreso del euskera data del S. XVI (*Linguae Vasconum Primitiae*, 1545), y lo publicó el cura Bernart Etxepare en Burdeos. Recoge 15 poemas, de los cuales dos son religiosos, diez poemas de amor, un poema autobiográfico y dos poemas de apología a la lengua vasca, que se publicaba por primera vez: *Berze jendek uste zuten ezin eskriba zaitaien, orai dute forogatu enganatu zirela* o dicho en castellano "las otras gentes pensaban que no se podía escribir [en euskera], pero ahora [con la impresión del libro] han visto que se confundieron". Esta idea refleja claramente que en aquellos tiempos existía en cierta parte de la sociedad la idea difundida de que era imposible de escribir, fundamento y argumento también de discriminación de la lengua vasca en varios espacios; idea rebatida en el S XIV mediante la impresión de aquel primer libro.

## ESTANDARIZACIÓN

El S. XVIII fué un tanto de luz como de sombra para la lengua vasca. Por un lado, en este siglo llega la Ilustración y se publican las primeras gramáticas del euskera, en el que destaca el trabajo del guipuzcoano Padre Larramendi: *El imposible vencido* (1729). Una vez más el título de una obra muestra que en esa época sigue arraigada la idea de la imposibilidad intrínseca a la lengua, argumentación que se seguía utilizando en favor de las élite conocedora de las hablas románicas. El título reclama que se ha conseguido superar lo que se creía imposible, es decir, sistematizar la gramática del euskera. Además de la gramática, Larramendi también publicó el enorme *Diccionario trilingüe del Castellano, Bascuence y Latín* (1745).

Por otro lado, durante la historia de la lengua vasca han sido numerosas las políticas que han rechazado el uso del euskera tanto en la calle como en las universidades o en cualquier nivel de educación (ver *El libro negro del euskera*, J. M. Torrealdai:1998). Por ejemplo, en

---

<sup>2</sup> En el *Liber Sancti Jacobi* aparecen unas 15 palabras vascas: *saia* (difícil), *urcia* (cielo), *ogui* (pan), *ardum* (vino), *aragui* (carne), *arraign* (pescado), *echea* (casa), *iaona* (señor), *andrea* (señora), *elicera* (a la iglesia), *gari* (trigo), *ur[ic]* (agua), *erreguia* (rey).

<sup>3</sup> Por ejemplo, tenemos una carta de 1415 escrita por el secretario del rey de Navarra, o las 19 cartas de espionaje que la señora de Urtubia escribió al gobernador civil de Guipuzcoa (S. XVI).

1730 por primera vez se documenta por escrito el uso del "anillo" en una escuela de Beasain (Guipúzcoa), práctica para disminuir el uso del euskera en las escuelas.<sup>4</sup>

La Revolución Francesa (1789) agravó aún más la situación, reprimiendo cualquier idioma que no fuera el francés, por lo que el vasco, el bretón, el occitano y el catalán que se hablaban en zona francesa fueron prohibidos: *Jargons barbares et ces idiomes grossiers qui ne peuvent plus servir que les fanatiques et les contre-révolutionnaires*; y aún a día de hoy se menciona a estas lenguas como *patois*, consideración claramente peyorativa. Las sombras de principios del XIX en la Vasconia peninsular incluyen que quedaban prohibidas las representaciones teatrales que no fueran íntegramente en castellano.

La segunda mitad del S. XIX marca una diferencia respecto a la primera mitad, y parece que el euskera se recupera con el Renacimiento Vasco<sup>5</sup>. En 1848 el suletino Agosti Xaho promueve publicar la primera revista vasca: *Uskal-Herriko Gaset*, en 1853 Anton de Abbadia organiza los juegos florales, *Euskal Lore Jokoak*, en Labort siguiendo las corrientes similares de Europa. En el País Vasco peninsular, Juan Antonio Mogel publica el libro *Peru Abarca* y Arturo Campión la *Gramática de los cuatro dialectos literarios de la lengua euskara*. Finalmente, en 1885 se comenzó a publicar en Los Ángeles (California) la primera revista semanal, *Escualdun gazeta*; en 1887 se empieza a publicar el semanario *Eskualduna* en el País Vasco continental que durará hasta 1944; y en 1896 el político vizcaíno Sabino Arana escribe *Lecciones de ortografía del euskera bizkaino*. El Renacimiento de la lengua vasca se alarga hasta 1936, año en el que el Gobierno Vasco (Eusko Jaurlaritza) nombra el euskera lengua oficial –junto con el castellano– mediante Estatuto.

Sin embargo, en 1936 dio comienzo a la Guerra Civil Española (1936-1939), que paró todos los avances que se dieron hasta entonces en los ámbitos de las publicaciones, la educación y la política vasca. En 1939 terminó la Guerra Civil Española, pero empezó la dictadura de Francisco Franco (1939-1975, 36 años), que reprimió, prohibió y persiguió toda la cultura de la península que no fuera la española, es decir, tanto la vasca, como la catalana o la gallega. La represión no fue únicamente cultural, sino también lingüística y política. De todas maneras, hacia 1960 y los años consiguientes, la tendencia se revierte y florece la educación en euskera (Clark, 1979:149), tema que trataremos en el apartado 4.

En este contexto, tal y como apunta Salaburu (2018), los intentos y las propuestas de unificación no tuvieron éxito.<sup>6</sup> De hecho, Euskaltzaindia o la Academia de la Lengua Vasca se creó en 1918, con el objetivo de fomentar, investigar y cuidar el idioma vasco. Sin embargo, por razones histórico-políticas que ya hemos mencionado no fue hasta medio siglo después (1968) cuando se asentaron en Aránzazu las bases para la unificación del euskera y se fijó un plazo de diez años para analizar el éxito o fracaso de la propuesta. En 1978 se celebró el Congreso de Bergara y se evaluaron los resultados, concluyendo que, a diferencia del período previo, la propuesta mostraba éxito<sup>7</sup>, no exento de discusiones y controversias<sup>8</sup>. Gracias a la colaboración de lingüistas y filólogos en esos años se promulgaron varias normas de escritura,

---

<sup>4</sup>Se trata de que cada día el niño que primero habla en vasco en la escuela es marcado con un anillo, del cual debe deshacerse durante el día pasándoselo a otro compañero que escuche hablar en euskera, y así, el niño que posea el anillo al final del día será castigado.

<sup>5</sup> Euskal Berpizkundea (Renacimiento Vasco)

<sup>6</sup> En 1901 se hizo una junta sobre la ortografía unificada en Hendaia, y al año siguiente en Hondarribia se hizo la Reunión de los Vascófilos (Euskaltzaleen Biltzarra). En 1905 el lingüista Resurrección María Azkue publicó el *Diccionario Vasco-Español-Francés*.

<sup>7</sup> El 66,8% de los escritores habían adoptado el modelo estándar. Tan solo un 4,6% quiere volver hacia atrás y que cada cual escriba como mejor le parezca. El 90% de los libros publicados siguen el modelo de unificación (Salaburu 2018:62)

<sup>8</sup> Para más información sobre las controversias generadas en las decisiones tomadas para el vasco unificado, se puede consultar Salaburu 2018.

se publicó el *Hiztegi Batua* [Diccionario Unificado] y también varios tomos de gramática vasca (Salaburu 2018:50-51).

En los años posteriores se fue avanzando y perfilando la propuesta hasta llegar al modelo que conocemos hoy en día. No obstante, los profesores de las escuelas aplicaron desde el comienzo las normas comunes, dado que la nueva herramienta permitía aligerar y solventar sus dudas. Con ello, desde las escuelas salían nuevas promociones educadas en el modelo estándar y así llegó a la universidad; adelantamos siguiendo a Salaburu (2018:62-63) que en la facultad de Ciencias de la UPV/EHU comenzaron a impartirse las primeras clases en vasco, y también en Magisterio; este tema que se desarrolla en la subsección dedicada a la educación. Poco a poco el euskera se hizo lugar en otros ámbitos como la administración, los medios de comunicación y las nuevas tecnologías.

## OFICIALIDAD Y USO

Previamente a hablar de la perspectiva lingüística de la oferta educativa y el alumnado, es deber dar luz sobre los diferentes estatus legales que tiene hoy en día la lengua. La Vasconia tradicional está dividida por dos administraciones estatales y tres instituciones administrativas, cada una con sus propio ordenamiento jurídico:

- La zona bajo administración española está dividida en dos grandes administraciones llamadas, más algún enclave:
  - Comunidad Autónoma Vasca (CAV), también llamada País Vasco o Euskadi, formada por los territorios vizcaínos, guipuzcoanos y alaveses. En este último se encuentra el enclave de Treviño, bajo administración de la Provincia de Burgos.
  - La Comunidad Foral de Navarra (CAN), que está formada por la alta Navarra
- La zona bajo administración francesa. Ésta forma parte de la Región de Aquitania y el Departamento de los Pirineos Atlánticos. Recientemente articuló la comunidad de municipios *Euskal Hirigune Elkargoa* (EHE), *Communaute Pays Basque* [Comunidad del País Vasco]. Está formada por los territorios de Labort, la baja Navarra y Sola.

El vasco o euskera es cooficial –junto al español– en todo el territorio de la CAV, no lo es en la zona bajo administración burgalesa y tiene una oficialidad estratificada en la Comunidad Autónoma de Navarra, de acuerdo con la Ley Foral del Euskera (BOE-A-1987-1257), que subdivide el territorio en tres regiones: vascófona, mixta y no vascófona. En todo el territorio francés, aunque se reconoce la realidad plurilingüe de éste, la única lengua oficial es la parisina *Langue d’Oil*, que hoy llamamos *francés*. Sería relevante declarar también que históricamente la lengua de contacto con el vasco no fue el francés hasta épocas recientes, debido al retroceso de la lengua occitana.

Según el Instituto estadístico vasco *Eustat*, hoy en día los territorios vascos constan de una población de unos 3 millones de personas, que según la encuesta sociolingüística llevada a cabo en 2016 responde a las siguientes caracterizaciones lingüísticas:

*60.7% of the population aged 2 years and over had some knowledge of Basque 895,942 were Basque speakers, that is, people who understand and speak Basque well, and 391,897 were Quasi-Basque speakers, with an average or good level of understanding but with difficulty speaking.* (Eustat, 2018)

El vasco padeció, y en parte de su territorio todavía padece, de un constante retroceso en uso y conocimiento; que gracias a nuevas políticas lingüísticas y de educación en particular parecen reconducir. Así y todo, no más de un tercio de la población de estos territorios es vascoparlante, territorios en los que desde recientes épocas es imposible encontrar adultos monolingües vascos.

## EDUCACIÓN

Cuando hablamos de educación, hablamos de que hoy en día en las escuelas el euskera existe como lengua vehicular y de enseñanza en una gran cantidad del territorio vasco en las dos vertientes de los Pirineos, siempre de formas diferenciadas enmarcadas en su realidad legal.

**Figura 2:** Aula de la ikastola Berrostitxueta, Bilbao, 1933, gestionada por mujeres de Emakume Abertzale Batza



*Ezkurdia et al., 2017*

Pero no debe creerse que la educación con la lengua vasca como lengua vehicular vaya a ser una novedad. Reseñemos que para el año 1923 el 64% de las estudiantes de la escuela normal de Bilbao también cursaban estudios de idioma vasco (Ezkurdia et al., 2017), con intención de impartir docencia en esa lengua. En ese mismo año se prohibió por Real Decreto la impartición de clases en euskera, que recién se retomó ya cuando el gobierno de España fue de estructura republicana. De esa época data la fotografía de la figura 2 tomada en una de las ikastolas (escuelas basadas en la lengua vasca) organizadas por el movimiento de mujeres nacionalistas vascas, Emakume Abertzale Batza.

Así mismo, en el territorio bajo administración francesa solamente puede ser utilizada la lengua vasca como vehicular en todas las etapas dentro de un pequeño círculo de escuelas de titularidad privada, en las que se ofertan todos los niveles educativos y obligatorios. También existen algunas pocas experiencias piloto en las escuelas nacionales republicanas, en las que la lengua vasca se utiliza, pero nunca más de un 50% de las horas lectivas, en todos los niveles.

Dentro del territorio vasco bajo administración española también existen diferentes modelos, pero hay mayor implementación de la enseñanza en vasco en la red pública de educación. Así, en el territorio que hoy conforma la CAV, la gran mayoría de la oferta educativa se hace en lengua vasca, para todo el alumnado independientemente de su origen lingüístico, vasco, mixto o foráneo. En la parte sur de Navarra, que según el ordenamiento territorial hoy conforma la CAN, la citada Ley Foral del Euskera demarca los territorios en los que la oferta educativa pública dispone o no educación con el vasco como lengua vehicular. En la zona legislada como no-vascófona también existe, igualmente, oferta dentro de la red privada en la que sí se utiliza como lengua vehicular. Por último citamos que en el enclave de Treviño, de administración burgalesa y sin oficialidad del euskera, existen dos pequeños

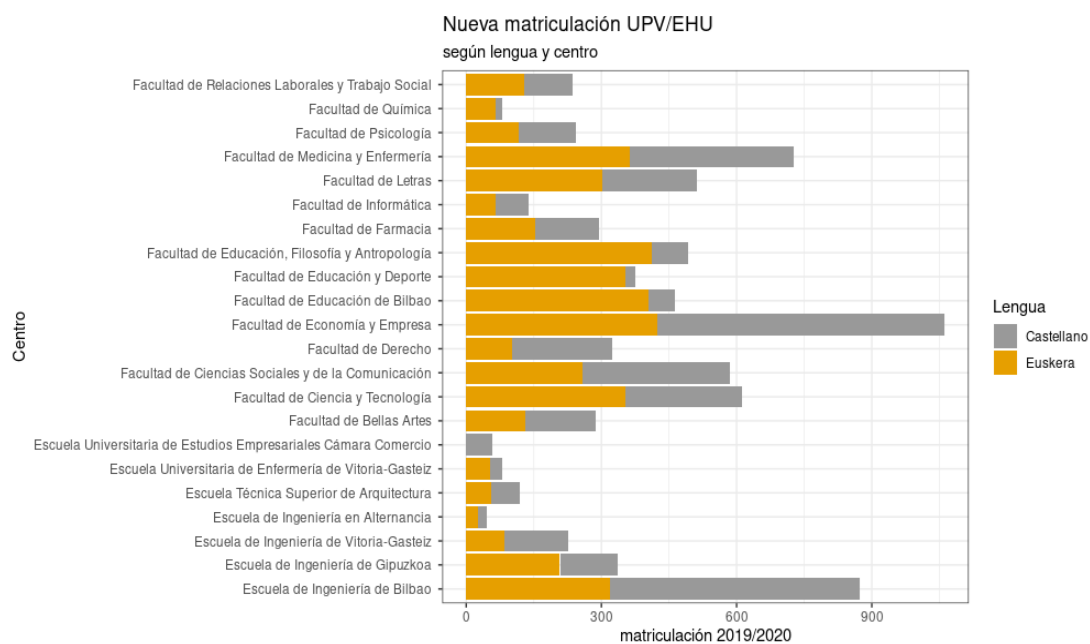
centros: uno público sin presencia del idioma vasco, y otro de gestión cooperativa y privada en el que el euskera es la lengua vehicular.

### EL EUSKERA Y LA UNIVERSIDAD

Más allá de la educación obligatoria, nos encontramos frente a la institución universitaria. La primera inclusión del vasco como objeto de estudio no se dio en territorio vasco, sino que fue en la Universidad de Burdeos, en el año 1948, en plena postguerra. Así, hasta que se creó la Universidad de Navarra y ésta su cátedra de Lengua Vasca, en Pamplona en el 1964, no se dictaron clases universitarias en lengua vasca dentro de su territorio.

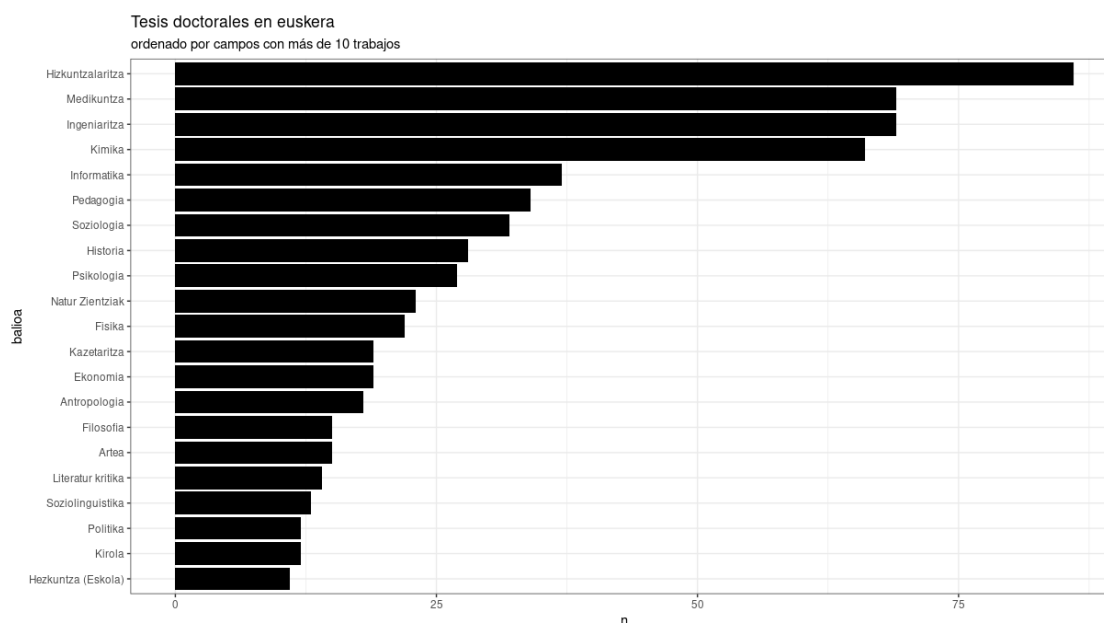
Desde aquellos comienzos tardíos y haciendo propios los reclamos que en filas vasquistas, de las dos orientaciones antagónicas principales de la Vasconia peninsular, españolista nacional, como Urmeneta (1967) o desde una perspectiva nacionalista vasca como la formación de la UEU (1973), la sociedad vasca recién consiguió darle su sitio al euskera entre las lenguas de transmisión del conocimiento académico, con la creación de la Universidad del País Vasco/Euskal Herriko Unibertsitatea, UPV/EHU, en 1980, como ordenadamente presentó Artetxe Sánchez (2011) en la sección referente a la historicidad en la universidad.

**Figura 3:** Alumnado nuevo en la UPV/EHU



UPV/EHU

Hoy en día dentro de la UPV/EHU se pueden cursar la mayoría de sus titulaciones de grado en lengua vasca, así como gran cantidad de estudios de postgrado; como se aprecia en la figura 3 en un solo centro no hubo alumnos de nueva matriculación en euskera, así como también se aprecia que una gran proporción de las matrículas en estudio de grado se hacen en esta lengua. En las universidades de titulación privada de Deusto (UD) o Mondragon Unibertsitatea (MU) o las públicas de Navarra (UPNA) o del País de Pau y del Adour (UPPA) se pueden seguir estudios en lengua vasca, tanto concernientes a la propia lengua, como carreras técnicas o científicas.

**Figura 4:** Áreas con mayor número de tesis doctorales presentadas en lengua vasca

Elaboración propia, fuente: *inguma.eus*

Prosiguiendo con la perspectiva de estudios académicos, es necesario hacer mención a la producción de tesis doctorales escritas en la lengua que nos atañe. Así, luego de la primer tesis leída en lengua vasca en el año 1976, que versaba sobre temática médica (cardiología) hoy en día las tesis leídas se acercan al millar, en las que, como se puede apreciar en la figura 4 en primer lugar se encuentra la temática lingüística, seguida por la medicina, ingeniería, química e informática.

Y, siendo que los autores nos debemos a la investigación lingüística y variacionista, desde esa perspectiva nos acercaremos a la investigación del campo que nos atañe.

### VARIACIONISMO CUANTITATIVISTA

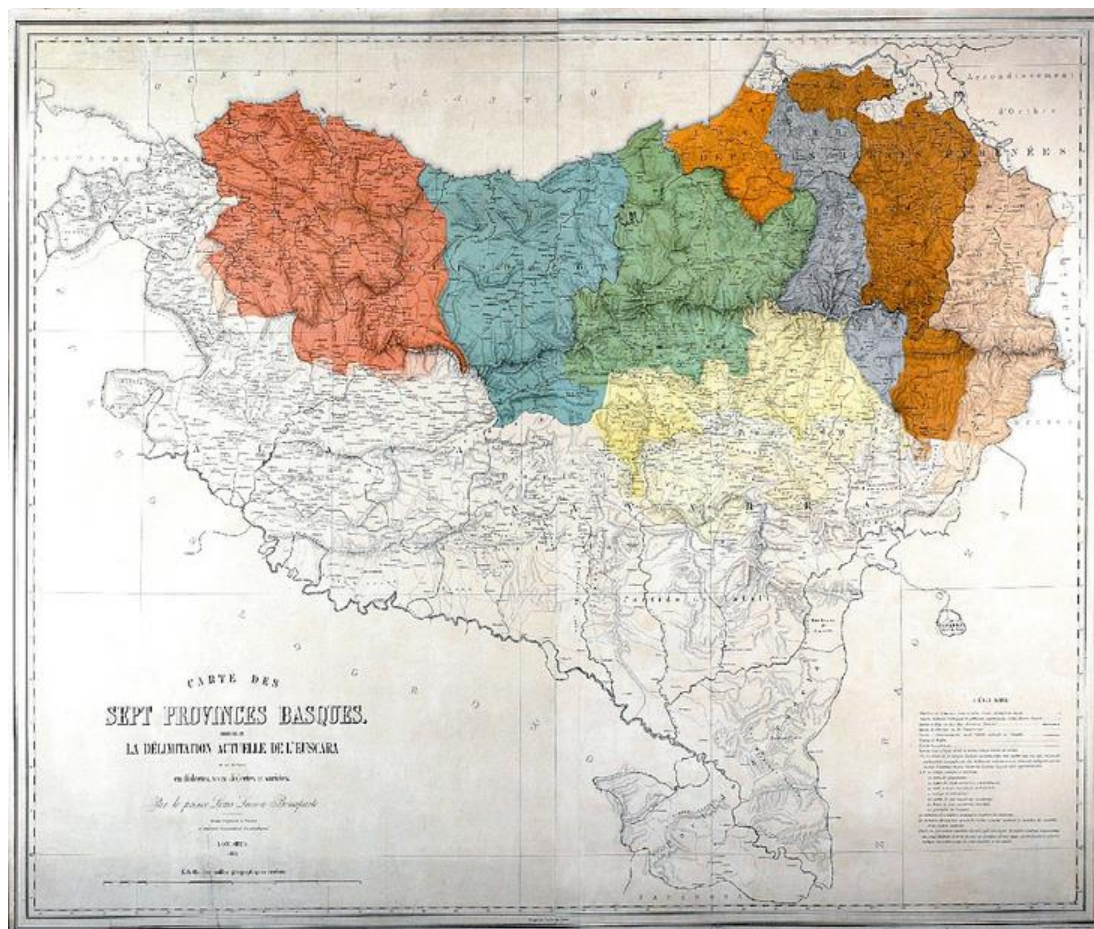
En lo que se refiere a la investigación del variacionismo lingüístico cabe mencionar que los estudios pueden abordarse desde diferentes perspectivas: la geográfica, la sociocultural, la contextual, o la histórica serían buenos ejemplos. Estos estudios, por otro lado, pueden realizarse mediante métodos tanto cualitativos como cuantitativos, o también con la combinación de ambos.

El interés por investigar el idioma vasco lo fomentó Wilhelm von Humboldt tras su primer viaje a tierras vascas en 1799. Después de él, llegaron las primeras investigaciones acerca de las variantes dialectales. La figura 5 muestra el primer mapa dialectal vasco basado íntegramente en criterios lingüísticos, realizado en 1863 por el Príncipe Louis Lucien Bonaparte.

La mayoría de las clasificaciones de los dialectos vascos realizadas hasta el S. XX se han abordado desde la perspectiva de la dialectología tradicional, que se basa en el estudio de la variación geográfica con análisis íntegramente cualitativos. Pero con el S. XXI se abre el horizonte a las clasificaciones dialectológicas cuantitativas basadas en las técnicas de agrupamiento jerárquico o *clustering* (Aurrekoetxea et al. 2019), siguiendo los pasos de Séguy (1971) y Hans Goebel (1981), precursores del *clustering* en los análisis de la variación dialectal.



**Figura 5:** Mapa de las Siete Provincias Vascas mostrando la delimitación actual del euskera y sus divisiones en dialectos y variedades



Bonaparte, 1863

*¿De qué hablamos al mencionar la dialectometría?*

Allá por el año 1971 se propuso una medida objetiva de distancia lingüística (DL), que Séguy llamó *distancia lexical*; en su proposición de relacionar esa, otrora, nueva medida con la distancia espacial. Esa primera formulación se desarrolló desde una perspectiva léxica, basada en los datos previamente recogidos en diversos atlas lingüísticos.

La metodología, a grandes rasgos y explicada con objetivo didáctico, pretende apuntar a la inter-inteligibilidad de los puntos en los que se recoge la información lingüística de los elementos a estudiar. Así, se pueden generar grandes matrices de datos en las que se verifica el uso de las formas lingüísticas en las diferentes localidades o puntos de información, como se observa en la tabla 1.

**Tabla 1:** Perspectiva cualitativa del uso de formas lingüísticas de cada localidad

	1	2	3	4
Ahatsa	-aren ganean	anhitz	arroltze	aski
Ahetze	-aren ganean	(h)anitz	arroltze	aski
Aiherra	-aren ganean	(h)anitz	arroltze	aski
Ainharbe	-an ganean	hanitx	arraultze	aski
Ainhize	-Ø ganean	(h)anitz	arraultze	aski
Ainhoa	-an ganean	(h)anitz	arroltze	aski
Aintzila	-an ganean	anhitz	arraultze	aski
Aiziritze	-Ø ganean	anizko	arroltze	aski
Aldude	-Ø ganean	(h)anitz	arroltze	aski
Aloze	-an ganean	hanitx	arraultze	aski

*Aurrekoetxea et al., 2004.*

En la tabla 1 se presentan las primeras diez localidades y los primeros cuatro ítems del atlas lingüístico realizado por Bourciez y luego editado y dialectometrizado por integrantes del grupo de investigación EUDIA (Aurrekoetxea et al., 2004, 2005; Aurrekoetxea & Videgain, 2009; Ormaetxea, 2005).

Esos datos lingüísticos, léxicos en esta muestra específica, necesitan ser llevados a una dimensión matemática, para ser cuantificados. Así, de esa muestra primigenia, se traspone una matriz de presencia/ausencia, representada en la tabla 2:

**Tabla 2:** Reelaboración de los datos desde perspectiva cuantitativa de presencia (1)/ ausencia (0) de datos léxicos en cada localidad

	1 -aren gaínean	1an gaínea n	1 Ø gaínea n	2 anhitz	2 (h)anit z	2 hanitx	2 anizko	2 anhitz	3 arrolt ze	3 arrault ze	4 aski
Ahatsa	1	0	0	1	0	0	0	0	1	0	1
Ahetze	1	0	0	0	1	0	0	0	1	0	1
Aiherra	1	0	0	0	1	0	0	0	1	0	1
Ainharbe	0	1	0	0	0	1	0	0	0	1	1
Ainhize	0	0	1	0	1	0	0	0	0	1	1
Ainhoa	0	1	0	0	1	0	0	0	1	0	1
Aintzila	0	1	0	0	0	0	0	1	0	1	1
Aiziritze	0	0	1	0	0	0	1	0	1	0	1
Aldude	0	0	1	0	1	0	0	0	1	0	1
Aloze	0	1	0	0	0	1	0	0	0	1	1

*Elaboración de los autores*

Como se puede apreciar a tabla 2 presenta los mismos datos que la tabla 1, pero con una perspectiva claramente cuantificable, que permite el cálculo de una matriz de semejanzas o diferencias, basada en las correspondencias de las características lingüísticas medidas mediante el relevamiento de datos del atlas que nos atañe, como se puede observar en la tabla 3:

**Tabla 3:** Cálculo de matriz de distancias basadas en semejanzas/diferencias

	Ahats a	Ahetze	Aiher ra	Ainhar be	Ainhize	Ainho a	Aintz ila	Aiziritze	Aldude	Aloze
Ahatsa	0	25	25	75	75	50	50	50	50	75
Ahetze	25	0	0	75	50	25	75	50	25	75
Aiherra	25	0	0	75	50	25	75	50	25	75
Ainharbe	75	75	75	0	50	50	50	75	75	25
Ainhize	5	50	50	50	0	50	75	50	25	75
Ainhoa	50	25	25	50	50	0	50	50	25	50
Aintzila	50	75	75	50	75	50	0	75	75	50
Aiziritze	50	50	50	75	50	50	75	0	25	75
Aldude	50	25	25	75	25	25	75	25	0	75
Aloze	75	75	75	25	75	50	50	75	75	0

*Elaboración de los autores*

En la matriz de distancias de la tabla 3 se presenta una reducción matemática de las cualidades lingüísticas presentadas precedentemente. Siendo que todos los datos se refieren a sólo cuatro ítems, las diferencias posibles en base a un 100% de diferencia, son de 0%, 25%, 75% y un posible 100%, que no se da, ya que el cuarto ítem en todas las localidades relevadas recibe la misma respuesta.

Esa matriz de datos que se presenta muy simplificada en este artículo. En las publicaciones en las que se utilizó representa un total de 138 localidades y 246 ítems con un total de 33943 respuestas recogidas y estandarizadas. Esas cantidades, que necesariamente exceden el espacio de este artículo, pueden ser imaginadas por el lector. Con esas cantidades cobra sentido la utilización de técnicas de estadística multivariante, que son las que producen resultados relevantes que, a su vez, permiten avances en la dialectología hasta ahora imposibles con tal nivel de objetivación.

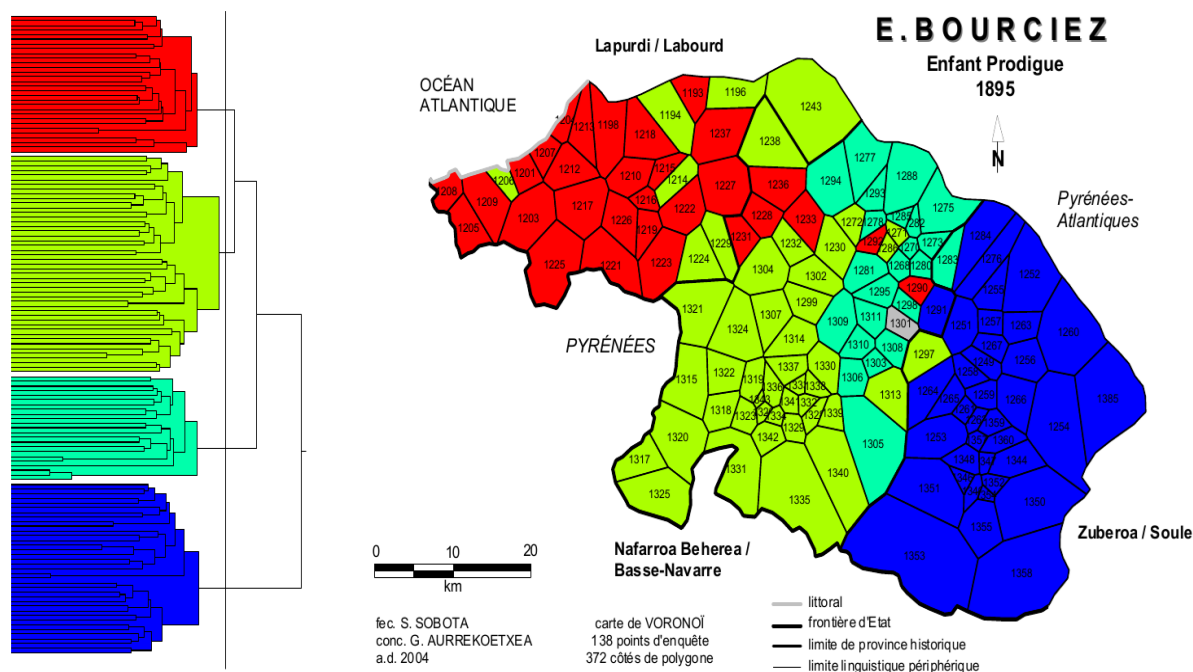
### TÉCNICAS, CORPUS Y TRABAJOS RECIENTES Y ACTUALES

En el grupo de investigación EUDIA se desarrolló un software web, especializado en cálculos multivariantes e información geográfica que responde a las necesidades propias del investigador de la variación (geo)lingüística. Este software, de nombre DiaTech (Aurrekoetxea et al., 2013) es de acceso y uso libre y responde a las técnicas desarrolladas dentro de la dialectometría. Desde su diseño se planteó la importancia de que pudiera ser registrada la respuesta múltiple, realidad sociolingüística a la que los desarrollos informáticos anteriores no podían responder. Los dos softwares de uso anterior sirvieron de base teórica y pragmática en el desarrollo de DiaTech; tanto el programa VDM, Visual Dialectometry<sup>9</sup>, desarrollado entre los años 1997 y 2000 por Edgar Haimerl de la Universidad de Salzburgo (Goebel, 2006) como el programa GabMap (Nerbonne et al., 2011), planteado, para hallar las distancias fonéticas interlocales dentro del mismo lema, no eran capaces de realizar cálculos multivariantes cuando

<sup>9</sup><http://ald.sbg.ac.at/dm/Engl/VDM/default.htm>

las respuestas al mismo ítem pertenecían a diferentes ítems. El programa Diatech<sup>10</sup>, como se mencionó en las líneas anteriores, es una aplicación web que funciona en línea, de acceso libre previo registro, tiene una interface multilingüe (inglés, francés, castellano y euskera), orientada al usuario final que no necesita más que conocimientos mínimos de usuario en informática, debido a sus características visuales. Hoy en día da acceso a diversos análisis estadísticos basados en diferentes perspectivas dialectométricas con sus correspondientes cálculos.

**Figura 6:** Estructuración geolingüística de los datos agrupados según el algoritmo de Ward



*Aurrekoetxea & Videgain, 2009:105*

En los siguientes párrafos presentamos algunos de los trabajos de análisis de la variabilidad lingüística que se se realizaron o se están realizando dentro del grupo de investigación EUDIA, del que se somos partícipes ambos autores, basándose en los cálculos de las matrices de distancias lingüísticas obtenidos mediante el uso de DiaTech.

Tomando como base a analizar los datos recogidos en 150 localidades por Edouard Bourciez en 1895 se realizó un proceso de estandarización de las versiones (Aurrekoetxea & Videgain, 2004), que luego se utilizó para construir un atlas lingüístico en dos tomos (Aurrekoetxea et al., 2004, 2005), el primero referente al léxico y el segundo a la gramática.

Esos datos permitieron hacer un análisis geolingüístico con herramientas propias de la dialectometría (Aurrekoetxea & Videgain, 2009), y, entre otros resultados, se presentó un análisis matemático de la estructuración de las distancias lingüísticas representadas en dicho atlas, la primera proyección científica y reproducible de un conjunto de datos con vistas al análisis de las variedades con objeto de la representación de los dialectos<sup>11</sup>. La figura 6 da cuenta de este trabajo basado en las hablas vascas del territorio francés:

En la parte izquierda de la figura 6 se representa una conglomeración de los 137 puntos de encuesta, en función de sus interdistancias lingüísticas. En esta aglomeración los criterios que se utilizan para considerar cómo se van conformando los conglomerados, son los que estableció Ward (1963) y ya era anteriormente implementado en el cito VDM. Los datos del

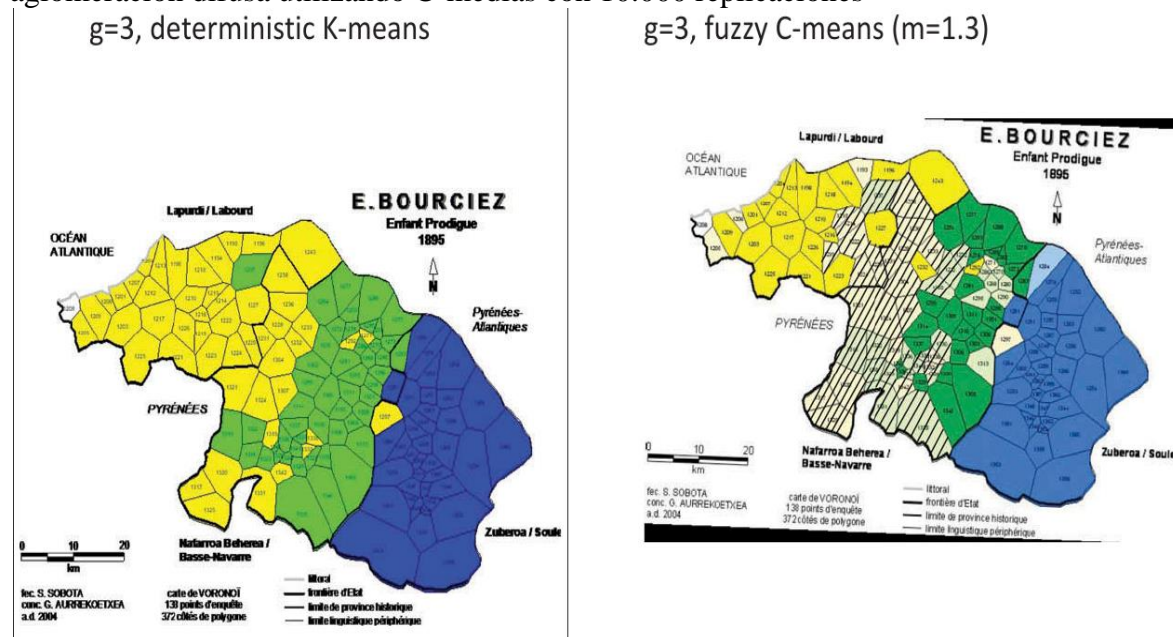
<sup>10</sup><http://eudia.ehu.es/diatech/index>

<sup>11</sup>Existen dos trabajos anteriores, pero no hay disposición pública de los datos analizados.

gráfico de la izquierda, llamado dendrograma, se van agrupando en base a su nivel de similitud y, a su vez, esos grupos se pueden agrupar con otros elementos, que a su vez pueden ser grupos o elementos todavía sin agrupar, en función de su distancia lingüística con el resto de los grupos.

Esas agrupaciones se representan mediante las líneas verticales de unión, que cuanto más a la derecha aparecen, representan mayor distancia entre las unidades que agrupan. Así las cosas, existe una decisión subjetiva en la que el investigador decide cuál es el número de agrupaciones que desea considerar. En ese trabajo Aurrekoetxea y Videgain optaron por cuatro grupos, acorde con una de las representaciones tradicionales de la dialectología vasca aplicada en el territorio vasco continental. Los datos del dendrograma pueden ser volcados sobre una representación geográfica, tal el mapa de la parte derecha de la imagen, en la que se aprecia coherencia entre la distribución geográfica y la agrupación matemática, a la vez que también se entreveran más las localidades, a medida que son más cercanas las uniones entre las agrupaciones. Se aprecia claramente que la zona de Sola (en azul) representa una unidad, que es la que primero se desprende del grupo primigenio del total del atlas estudiado y también es el primer subgrupo que se forma de entre los cuatro que se consideran, según puede verse en el dendrograma.

**Figura 7:** Representación cartográfica de la clasificación dialectal del corpus de “Bourciez”. A la izquierda representación determinística por medio de K-medias y a la derecha aglomeración difusa utilizando C-medias con 10.000 replicaciones



Aurrekoetxea et al., 2020

La perspectiva que fue descrita hasta estas líneas, llamada determinista, busca señalar a qué dialecto pertenece el habla de cada una de las localidades. Esta perspectiva de análisis, si bien es correcta para muchos objetivos, no responde a la intuición de que las variedades locales específicas, si bien pertenecen un algún grupo determinado, pueden tener componentes característicos de otros grupos, en diferente medida dentro del mismo grupo.

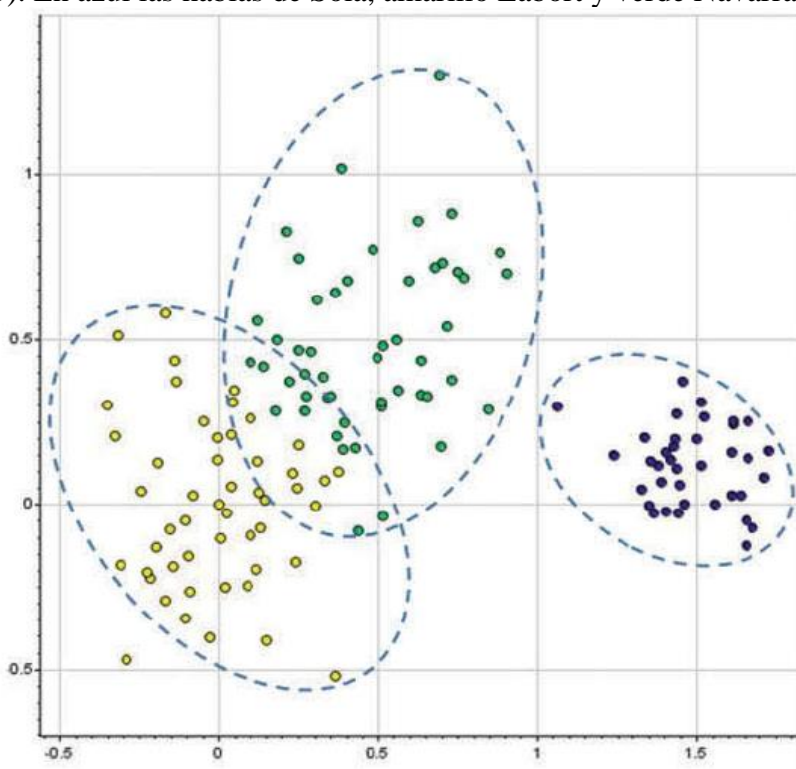
A este extremo le viene a responder el análisis no determinista, que señala ya no si pertenece o no a determinado grupo, si no que lo que señala es el porcentaje de pertenencia a los grupos. Esto se realiza utilizando las técnicas llamadas de aglomeración difusa o *fuzzy cluster* en inglés con difundido uso también en castellano de la forma inglesa. En la siguiente representación cartográfica de la figura 7 se representan tres grupos dialectales. Esos tres

grupos que en la izquierda son claramente identificables, en la representación de la derecha se puede ver que están conformados por gran cantidad de hablas en espacio de transición.

La representación de las diferentes variedades locales también se puede realizar utilizando técnicas de reducción de las dimensiones, escalado multidimensional (MDS por sus siglas en inglés), en el que se representa cada habla como un punto en el plano, distante del resto de los puntos, en función de la distancia lingüística que los separa. De esta manera, en la figura 8 se observa que un grupo es más compacto entre sí que los otros dos que se definen.

Las variedades representadas en el espacio común a los grupos de Labort y Navarra dentro de la proyección MDS figura 8, coinciden con aquellas que en el mapa de la figura 7 izquierda se representan como hablas de transición.

**Figura 8:** Representación según técnicas de reducción de la dimensionalidad (MDS) de la clasificación dialectal del corpus ‘Bourciez’, obtenida utilizando el método de K-medias con tres grupos (g=3). En azul las hablas de Sola, amarillo Labort y verde Navarra continental

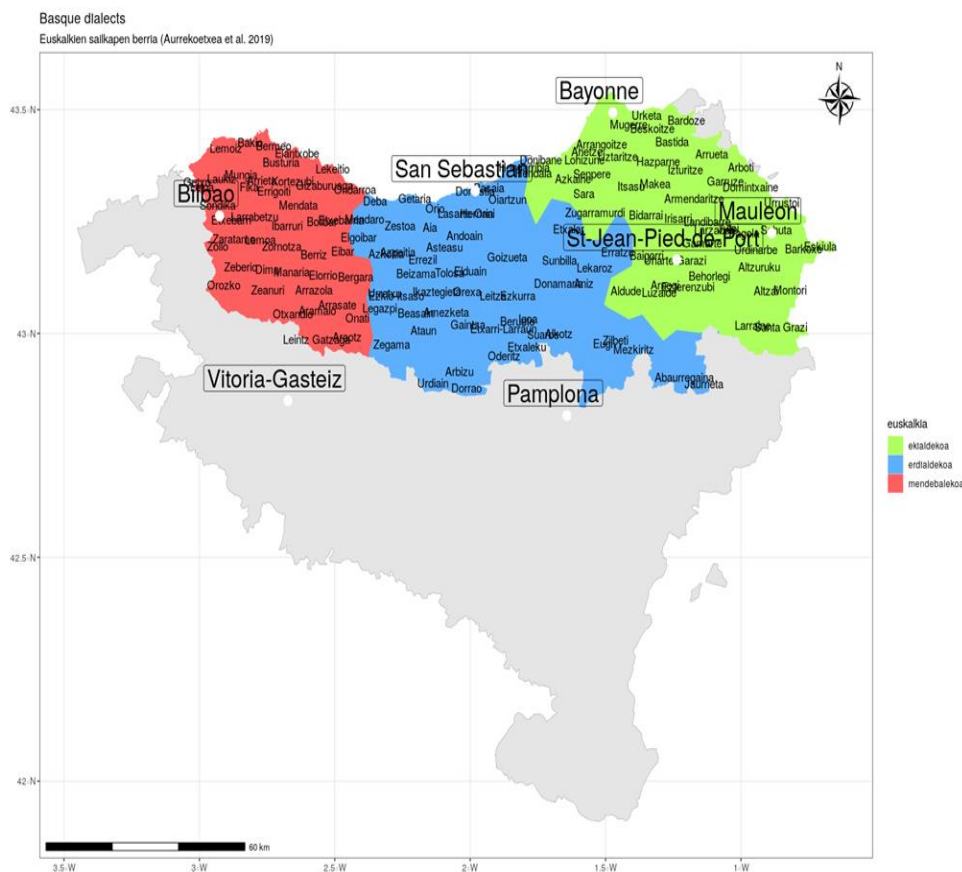


*Aurrekoetxea et al., 2020*

Sin duda, el más grande y exhaustivo atlas que se haya realizado sobre la lengua vasca es el Euskararen Herri Hizkeren Atlasa [Atlas de las hablas locales del vasco](Euskaltzaindia, 2008, 2010, 2012, 2014). En ese atlas se recogen las voces en 150 puntos de toda la geografía del euskera mediante encuesta. Así, hasta el día de hoy se publicaron once tomos con las respuestas positivas a estas. Originalmente, las respuestas se presentaban solamente en alfabeto fonológico, con intención de ser tan fiel como se podía al dato recogido.

Posteriormente a un análisis del conjunto de datos total, se determinaron los ítems que podrían representar al uso en el total de la zona de habla vasca y se procedió a una reindización y lematización de la que adolecían, para así poder generar la primera clasificación científica de los dialectos del euskera que se presentó a la sociedad vasca a finales del año 2019 (Aurrekoetxea et al., 2019), datos que se representan en la figura 9.

**Figura 9:** División de los dialectos vascos



Aurrekoetxea et al. (2019) - Elaboración propia

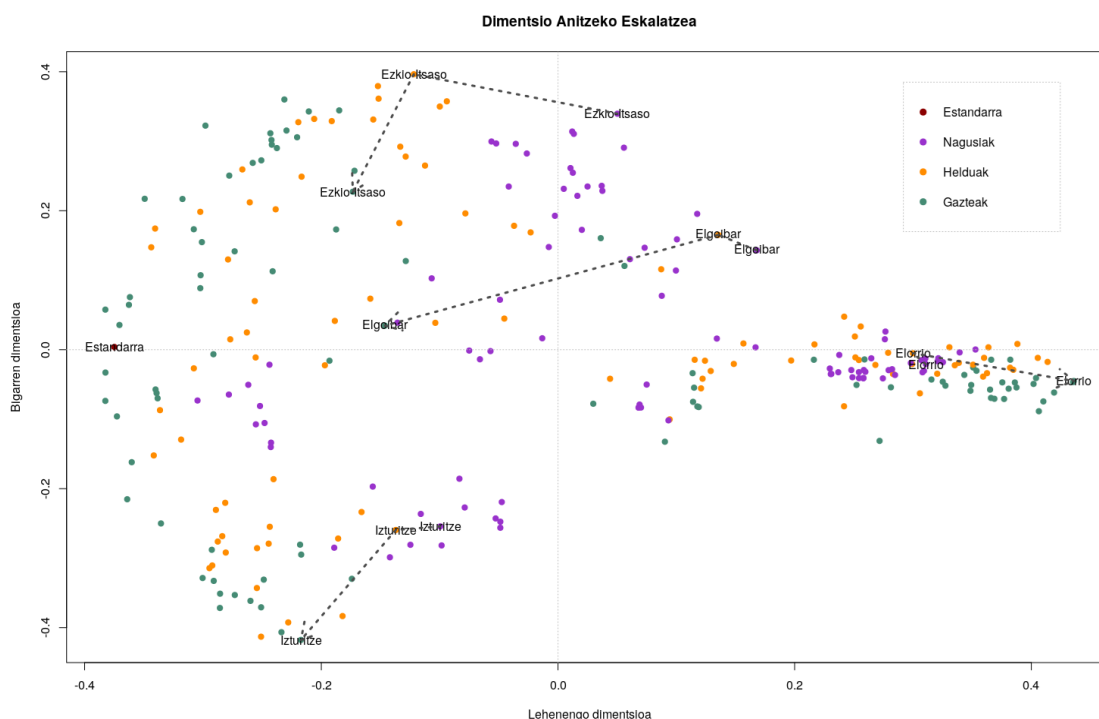
Hoy en día, dentro del grupo de investigación EUDIA se están desarrollando técnicas que van más allá de la dialectometría estándar. Se desarrollaron o están desarrollando y aplicando técnicas de dialectometría inversa (Aurrekoetxea et al., 2021), por ejemplo, en la que se explora un corpus dialectometrizado para revelar cuáles son las relaciones entre los ítems y sus variaciones. Se han aplicado técnicas de Análisis de Componentes Principales (Aurrekoetxea et al., 2016) o Análisis de Concordancias Múltiples y sus relaciones con la reducción de dimensionalidad (Abasolo & Ensunza, 2022); así como actualmente el análisis de las relaciones entre la distancia lingüística en sus diferentes dimensiones con la distancia geográfica (Aurrekoetxea et al., 2021).

En el análisis de la variabilidad lingüística, utilizando las técnicas antes mencionadas validadas en la dialectología científica, se están analizando los datos del Euskararen Atlas Soziolinguistikoa, EAS, [Atlas Sociolingüístico Vasco], en el que los informantes que aportan los datos pertenecen a las generaciones posteriores a los del EHHA. Estos informantes están divididos en dos generaciones, una alfabetizada en idioma vasco ya en su vida adulta y la otra escolarizada en idioma vasco. Estos datos permiten que se esté estudiando la variabilidad dentro de las hablas locales según la variable de la edad, *factor* en léxico de Lavob, como se muestra una escalación multidimensional de las distancias entre la morfología verbal de 100 localidades en tres generaciones diferentes (Unamuno et al., 2019) representados gráficamente en la figura 10. También se estudió la influencia del proceso de escolarización (Ensunza, 2016, 2019; Iglesias et al., 2020; Abasolo et al., 2020) y cómo se da el proceso de nivelado hacia las formas estandarizadas (Unamuno et al., 2021). Para profundizar acerca de los estudios sobre la nivelación entre dialectos y estándar vasco, puede consultarse el trabajo de Aurrekoetxea, Unamuno et al. (2020).



En el sector derecho del centro de la figura 10 se aprecia que algunas variedades, en su morfología verbal, convergen hacia el estándar; pero también se puede apreciar el caso de Elorrio, perteneciente al dialecto occidental, que en la misma comparación intergeneracional parece que se aleja del estándar en proceso de dialectización.

**Figura 10:** Dispersión de las características de la morfología verbal en tres generaciones

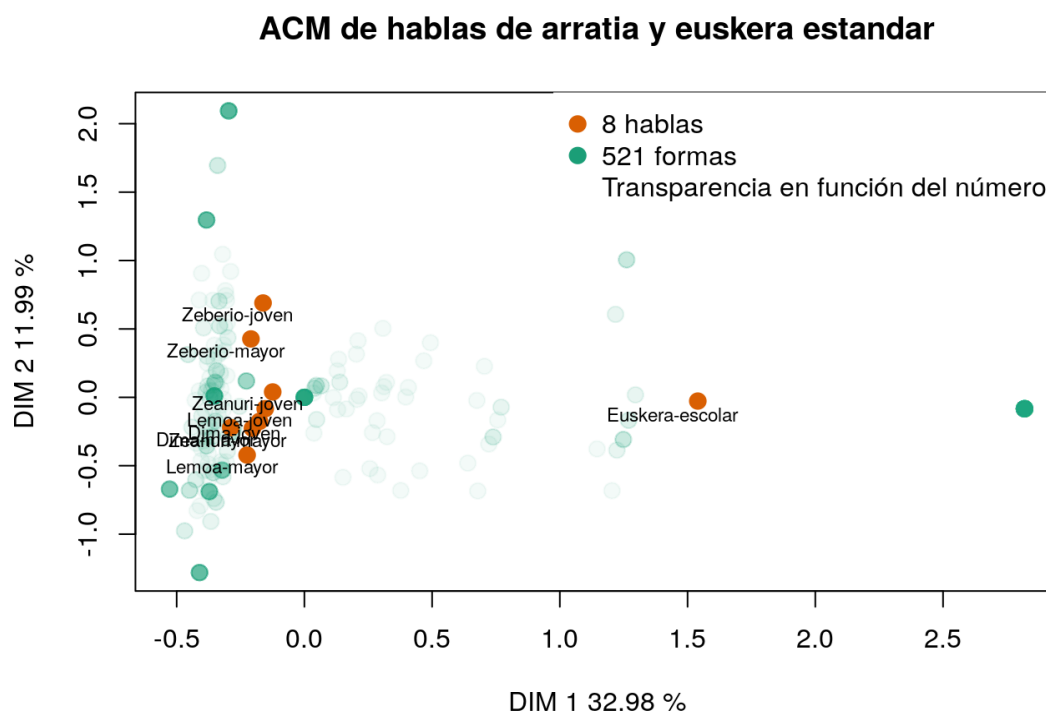


*Unamuno et al., 2019*

Un paso más allá de las técnicas propias de la dialectometría, en las que la abstracción matemática abre el camino a las citadas técnicas, también pueden ser utilizadas otras estrategias de análisis multivariante en las que los datos no son abstraídos matemáticamente, aunque sí es necesario un alto poder de computación para ser aplicado. Así, se investigó la influencia de la alfabetización en euskera en el valle de Arratia (Abasolo et al., 2020; Iglesias et al., 2020) utilizando técnicas de Análisis de Concordancias Múltiples aplicado en lenguaje R (Lê et al., 2008). Esta técnica, además de presentar las más importantes dimensiones de variación, permite recobrar la información de cada uno de los elementos que la conforman, tal como cuál es cada una de las formas lingüísticas que construye la dimensión y cuál es su peso en ésta.

Con el uso de esta técnica, no solo se señaló la mayor cercanía de las hablas de los jóvenes a la lengua estándar, sino que en este trabajo también se pudieron determinar las características que van variando su uso en dirección a la lengua estándar. Esos elementos en la figura 11 aparecen representados por los puntos que están entre las hablas locales y el euskera escolar, y pueden ser recobrados y listados.

**Figura 11:** Resumen exploratorio del análisis que representa el 44% de la variación del corpus, utilizando ACM



*Iglesias et al., 2020*

## CONCLUSIONES

Indudablemente la historia del euskera o la lengua vasca goza de gran tradición, documentándose ya en los primeros siglos de nuestra era. Aunque tengamos textos, fragmentos y diccionarios que documentan el idioma, no es hasta el siglo XVI donde empieza a resurgir de entre las sombras y hacerse sitio en la imprenta. Aun así, tras siglos oscuros de prohibición y persecución, no es hasta fechas recientes que se asientan las bases de lo que hoy en día es el estándar o el euskera unificado. Con este logro, el euskera consigue extenderse del uso privado al uso público, como a la enseñanza en la universidad, a los medios de comunicación y la administración.

En cuanto a la dialectología, puede decirse que siempre ha suscitado interés por su carácter exótico y desconocido. Ya en el siglo XIX podemos encontrar las primeras clasificaciones de los dialectos vascos. En las últimas décadas, las investigaciones se han extendido y ampliado a nuevos horizontes, pasando de utilizar métodos tradicionales a métodos técnico-científicos.

Para finalizar presentaremos algunos paralelismos que encontramos entre el euskera y el guaraní que creemos que puedan ser de interés, principalmente desde una perspectiva sociolingüística.

El primer paralelismo: edición popular en el S XIX. Ambos empezaron a generar prensa, es decir, existía el movimiento para publicar prensa en los respectivos idiomas, durante la

Guerra de la Triple Alianza, *Cacique Lambaré*<sup>12</sup> (1867) en guaraní y *Escualdun Gazeta*<sup>13</sup> (1885) en California (América) en euskera.

El segundo paralelismo: el uso que se le da al idioma. Siendo que no es inteligible desde las lenguas vecinas, en los enfrentamientos, por ejemplo deportivos. Es natural que los jugadores de un equipo utilicen el vasco para comunicarse entre ellos, inclusive con los compañeros que no conoce bien el vasco, ya que los contrarios no lo van a entender. Suponemos que lo mismo ocurrirá con los equipos paraguayos y seguramente con la propia selección de fútbol, así como con otros equipos del globo terráqueo.

Tercer paralelismos: la prohibición. Los hablantes de los otros idiomas no lo entienden; es decir, el guaraní no se entiende si no sos guaraní así como el vasco no se entiende si no sos vasco. Continuando con el ejemplo del fútbol, hace poco sucedió aquí que en un partido que un árbitro castellano prohibió el uso del euskera en los partidos. Podrán imaginarse el revuelo que causó este hecho. En Paraguay no hay que hurgar en la historia antigua para encontrar prohibición del uso del guaraní en determinados entornos<sup>14</sup>.

Cuarto paralelismo: uso en la guerra. Recordemos las situaciones que se dan en las guerras. Tanto las de la Triple Alianza como la Guerra del Chaco, como en el País Vasco la Resistencia a los nazis de un lado de la frontera y en la Guerra Civil Española del otro, como también en el frente del Pacífico dentro del ejército de los Estados Unidos; la inteligibilidad de la lengua fue también puesta en valor.

Último paralelismo: Supuesta dificultad. Destacaremos que se ha mencionado que ambos idiomas son difícilísimos de aprender. Se dice del vasco que es una lengua difícil. Claramente no es inteligible, no es transparente desde las lenguas vecinas, como ya se puso en relevancia. Se dice del guaraní que es el noveno idioma más difícil de aprender<sup>15</sup>. Sin embargo, los niños que hablan euskera o guaraní aprendieron igual de fácil su propio idioma, como lo hace un chino, un español, un alemán, o cualquier persona que aprende su primer idioma.

Creemos que los métodos científicos desarrollados por el grupo de investigación EUDIA y mencionados en este artículo pueden ofrecer diversas oportunidades aplicables al estudio de los diversos continuum de variedades de los guaraní – jopará – castellano paraguayo, o ese llamado *portuñol*, aún poco definido, que puede señalar algún habla de Ciudad del Este o de Misiones o también de Rivera.

Tal y como declaró la Unesco en *Language Vitality and Endangerment* (Drude, 2003:1), la diversidad lingüística es esencial para el patrimonio humano y el conocimiento de un solo idioma puede ser la clave para responder preguntas fundamentales del futuro. Por ello, defendemos que es de vital importancia estudiar los idiomas como entes variables e irrepetibles de cada época, y testimoniar lo que nos ofrecen en la actualidad, para poder llegar a entender su desarrollo en el futuro.

En ese arduo trabajo son de gran ayuda las metodologías cuantitativas que ofrece la dialectometría, ya que permiten comparar y correlacionar objetivamente grandes cantidades de datos lingüísticos. En este aspecto, el software DiaTech desarrollado por el grupo de investigación EUDIA proporciona una herramienta para la gestión de bases de datos lingüísticos, incluyendo respuestas múltiples, con la que el investigador puede cartografiar

---

<sup>12</sup>[http://bibliotecanacional.gov.py/hemeroteca/cacique\\_lambara\\_1867-1868/](http://bibliotecanacional.gov.py/hemeroteca/cacique_lambara_1867-1868/)

<sup>13</sup><http://www.liburuklik.euskadi.eus/handle/10771/28841>

<sup>14</sup><https://www.lanacion.com.py/pais/2021/03/01/mediante-comunicado-spl-y-snc-repudian-prohibicion-de-hablar-en-guarani-en-una-estancia/> y <https://www.elnacional.com.py/nacionales/2021/03/03/la-academia-de-la-lengua-guarani-repudia-la-prohibicion-del-uso-del-guarani/>

<sup>15</sup><https://www.mre.gov.py/la/index.php/noticias/el-guarani-el-idioma-de-confianza-de-los-paraguayos>

mapas para mostrar la variación geolingüística, además de realizar análisis estadísticos basados en diferentes perspectivas dialectométricas con sus correspondientes cálculos.

Mediante estas herramientas es posible representar la estructuración de las diferentes variedades locales y mostrar objetivamente la distancia o semejanza entre cada grupo lingüístico. Todo ello, podría ser de ayuda en una clasificación científica objetiva de los dialectos del guaraní o los romances sudamericanos, entre otros. Asimismo, los nuevos análisis de dialectometría inversa permiten analizar la variabilidad dentro de las hablas locales según factores sociales. Esto permite estudiar tanto el proceso de nivelación entre hablas locales y formas estandarizadas como determinar las características que van variando su uso en dirección a la lengua estándar.

## REFERENCIAS

- Abasolo, J., Iglesias, A., Eguskiza, N., & Unamuno, L. (2020). Estandarraren eragina aztertzeko metodologia eta inplikazio didaktikoak. *Ikastorratza. e-Revista de Didáctica*, 25, 1–22. [https://doi.org/10.37261/25\\_alea/1](https://doi.org/10.37261/25_alea/1)
- Abasolo, J. & Ensunza, A. (2022). MCA teknikak euskal herri hizkeretan, zer darakuskute? [en imprenta] anonymous, & Pons, F.-L. (27AC.). Musée Saint-Raymond, Ra 385label QS:Len,"Musée Saint-Raymond, Ra 385" [Marblemedium QS:P186,Q40861]. Musée Saint-Raymond. <https://commons.wikimedia.org/wiki/File:MSR-Ra385-FLP.tif>
- Artetxe Sánchez, K. (2011). Euskararen aurrerabidea zientzia eta goi mailako esparruan: Azterketa bibliometrikoa. *Eusko Jaurlaritzaren Argitalpen Zerbitzu Nagusia = Servicio Central de Publicaciones del Gobierno Vasco*. [https://www.basquecountry.eus/contenidos/documentacion/cuaderno\\_sociologico\\_vasco\\_28/eu\\_cu\\_soc28/adjuntos/csv28.pdf](https://www.basquecountry.eus/contenidos/documentacion/cuaderno_sociologico_vasco_28/eu_cu_soc28/adjuntos/csv28.pdf)
- Aurrekoetxea, G. (2003). Euskalkiak estandarren uholdepean (Arratiako kasua). In *Ahozkotasuna aztergai* (167–177 pp). Mendebalde Kultur Elkarte.
- \_\_\_\_\_. (2010). The correlation between morphological, syntactic and phonological variation in the Basque language. In B. Heselwood (Ed.), *Proceedings of Methods XIII: Papers from the Thirteenth International Conference on Methods in Dialectology, 2008* (207-218 pp). Peter Lang.
- \_\_\_\_\_. (2011). EHHA-tik euskalkien mapa berrirantz. *Pirinioetako hizkuntzak: oraina eta lehen*: Euskaltzaindiaren XVI. biltzarra, 1173–1194. <https://dialnet.unirioja.es/servlet/articulo?codigo=3855925>
- \_\_\_\_\_. (2016a). Distantzia geografikoaren eta hizkuntza distantziaren arteko korrelazioa. In G. Aurrekoetxea, J. M. Makazaga Eizagirre, & P. Salaberria (Ed.), *Hire bordatxoan: Txipi Ormaetxea omenduz* (55–72pp). Universidad del País Vasco = Euskal Herriko Unibertsitatea.
- \_\_\_\_\_. (2016b). Analysis of the morphological variation of Basque. *Dialectologia et Geolinguistica*, 24(1). <https://doi.org/10/gfpqb2>
- Aurrekoetxea, G., & Videgain, C. (2009). Le projet bourcier: Traitement géolinguistique d'un corpus dialectal de 1895. *Dialectologia: revista electrónica*, 2, 81–111.
- Aurrekoetxea, G., & Videgain, X. (2004). Haur prodigoaren parabola Ipar Euskal Herriko 150 bertsiotan. *Eskuizkribua. Anuario del Seminario de Filología Vasca «Julio de Urquijo»*, 0(0). <http://www.ehu.eus/ojs/index.php/ASJU/article/view/8789>
- Aurrekoetxea, G., Abasolo, J., & Videgain, X. (2021). Corrélation entre la différenciation linguistique et la distance géographique en basque. In M. Nevaci, I. Floarea, & I.-M. Farcas (Ed.), *Ex Oriente Lux in Honorem Nicolae Saramandu*. Edizioni dell'Orso.
- Aurrekoetxea, G., Abasolo, J., Clua, E., Usobiaga, I., & Salicrú, M. (2021). Hizkuntza-aldagaien aldakortasuna: Lehen hurbilketa. *Uztaro: giza eta gizarte-zientzien aldizkaria*.
- Aurrekoetxea, G., Fernandez-Aguirre, K., Modroño, J. I., & Rubio, J. (2016). Principal Components and clustering on lexical distances from standard Basque to local varieties. *Journal of Research Design and Statistics in Linguistics and Communication Science*, 3(1), 5–22. <https://doi.org/10/gfxfwq>
- Aurrekoetxea, G., Fernandez-Aguirre, K., Rubio, J., Ruiz, B., & Sanchez, J. (2013). «DiaTech»: A new tool for dialectology. *Literary and Linguistic Computing*, 28(1), 23–30. <https://doi.org/10/gfpqb3>
- Aurrekoetxea, G., Gaminde, I., Ormaetxea, J. L., & Videgain, X. (2019). Euskalkien Sailkapen Berria. ["Nueva clasificación de los dialectos vascos"] UPV/EHU, Argitalpen Zerbitzua.

- Aurrekoetxea, G., Unamuno, L. & Ensunza, A. (2020). Dialectos y estándar vasco. In. Emili Boix-Fuster & Maria-Pilar Perea (Ed), *Llengua i dialectes: esperances per al català, el gallec i el basc* (151-167). Universitat de Barcelona.
- Aurrekoetxea, G., Iglesias, A., Clua, E., Usobiaga, I., & Salicrú, M. (2020). Analysis of Transitional Areas in Dialectology: Approach with Fuzzy Logic. *Journal of Quantitative Linguistics*.  
<https://www.tandfonline.com/doi/abs/10.1080/09296174.2020.1732765>
- Aurrekoetxea, G., Videgain, X., & Iglesias, A. (2004). «Bourciez» bildumako euskal atlasa (BBEA): 1. Lexikoa. *Anuario del Seminario de Filología Vasca Julio de Urquijo: International journal of basque linguistics and philology*. <https://dialnet.unirioja.es/servlet/articulo?codigo=2654688>
- Aurrekoetxea, G., Videgain, X., & Iglesias, A. (2005). «Bourciez» bildumako euskal atlasa (BBEA): 2. Gramatika. *Anuario del Seminario de Filología Vasca Julio de Urquijo: International journal of basque linguistics and philology*, 39(1), 1–277.
- Biscailuz, M. V. (1885). *Escualdun Gazeta*. Los Ángeles, California.
- Bonaparte, LL. (1863). *Carte des sept provinces basques montrant la delimitation actuelle de l'euscara et sa division en dialectes, sous-dialectes et varietes*, Standford's Geographical Establishment, London. [Berrarg. [1990], in *Opera Omnia Vasconice Euskaltzaindia*.
- Drude, S. (2003). Language vitality and endangerment. Document submitted to the International Expert Meeting on UNESCO Programme Safeguarding of Endangered Languages. Paris, 10-12 March.
- Ensunza, A. (2016). The sociolinguistic variation of palatalization: the case of Gernika-Lumo. *Dialectología: revista electrónica* (16), 71-91.
- \_\_\_\_\_. (2019). The influence of standard basque in Busturialdea. Part 1: Demonstrative pronouns. *Dialectología: revista electrònica*, 0(22), 17-29–29.
- Estrabón. (1991-2015). *Geografía*. Editorial Gredos.
- Etxepare, B. (1545). *Linguae Vasconum Primitiae*. Bordeaux
- Euskaltzaindia. (2008). *Euskararen Herri Hizkeren Atlas*, I, Lexikoa. Euskaltzaindia.
- \_\_\_\_\_. (2010). *Euskararen Herri Hizkeren Atlas*, II, Lexikoa. Euskaltzaindia.
- \_\_\_\_\_. (2013). *Euskararen Herri Hizkeren Atlas*, V, Izen morfologia. Euskaltzaindia.
- \_\_\_\_\_. (2014). *Euskararen Herri Hizkeren Atlas*, VI, Aditz morfologia. Euskaltzaindia.
- Eustat. (2018). Three out of five people living in the Basque Country had some knowledge of the Basque language in 2016. *eustat Euskal Estatistika Erakundea - Instituto Vasco de Estadística*.  
[https://www.eustat.eus/elementos/ele0014600/not0014683\\_i.html](https://www.eustat.eus/elementos/ele0014600/not0014683_i.html)
- Ezkurdia, G., Bilbao, B., & Pérez, K. (2017). Emakume Abertzale Batza (EAB) Bilbon. *Euskalingua*, 31, 14–27.
- Goebel, H. (1981). *Eléments d'analyse dialectométrique (avec application à l' AIS)*. In: *RLiR* 45, 349-420.
- \_\_\_\_\_. (2006). Recent Advances in Salzburg Dialectometry. *Literary and Linguistic Computing*, 21(4), 411–435. <https://doi.org/10.1093/lc/fql042>
- Iglesias, A., Abasolo, J., Etxebarria, A., & Ensunza, A. (2020). Influencia del euskera estándar entre dos generaciones, alfabetizada y no alfabetizada, del valle de Arratia. *Huarte de San Juan. Filología y Didáctica de la Lengua*, 20, 107–126. <https://doi.org/10.48035/rhsj-fd.20.5>
- Larramendi, M. (1729). *El imposible vencido*.
- \_\_\_\_\_. (1745). *Diccionario trilingüe del Castellano, Bascuence y Latín*.
- Lê, S., Josse, J., & Husson, F. (2008). FactoMineR: A package for multivariate analysis. *Journal of Statistical Software*, 25(1), 1–18. <https://doi.org/10/gfjkpd>
- Michelena, L., Lakarra, J. A., & Michelena, L. (2011). *Textos arcaicos vascos. N. Landuchio. Dictionarium linguae Cantabrigiae* (1562). UPV/EHU.
- Nerbonne, J., Colen, R., Gooskens, C., Kleiweg, P., & Leinonen, T. (2011). Gabmap—A web application for dialectology. *Dialectología: Revista Electrònica*, 65-89–89.
- Ormaetxea, J. L. (2005). *Euskal lexikografiari ekarri berria: Bourciez-en euskal testuetako lexikoa*. *Lapur dum. Euskal ikerketen aldizkaria | Revue d'études basques | Revista de estudios vascos | Basque studies review*, 10, 153–168. <https://doi.org/10/c7mgqz>
- Salaburu, P (2018). “El proceso de estandarización de la lengua vasca”, En J. Giralt & F. Nagore (eds.), *Lenguas minoritarias en Europa y estandarización*, 46-75.
- Séguy, J. (1971). “La relation entre la distance spatiale et la distance lexicale”, *Revue de Linguistique Romane* 35, 335-57.
- Torrealdai, J. (1998). *El libro negro del euskera*. Donostia: Ttartalo.
- Unamuno, L., Abasolo, J., & Eguskiza, N. (2021). *Euskara estandarra eta euskalkien arteko levelling edo berdintze-fenomenoa. Hizkuntzalari Euskaldunen V. Topaketa: Kimu berriak Euskal Hizkuntzalaritzan, Eibar*.
- Unamuno, L., Abasolo, J., Eguskiza, N., & Iglesias, A. (2019). "Euskararen atlas sozio-geolinguistikoa" euskara estandarretik so. *FLV*, 50 urte: metodo eta joera berriak (euskal) hizkuntzalaritzan, Pamplona.

- Urmeneta, M. J. (1967, diciembre). Navarra y el Vasconce. Principe de Viana. Suplemento mensual de la revista, destinado al fomento del vasconce, 21, 10.
- Ward, J. H. (1963). Hierarchical Grouping to Optimize an Objective Function. *Journal of the American Statistical Association*, 58(301), 236–244. <https://doi.org/10.1080/01621459.1963.10500845>