Dialectologia. Special issue, 13 *(2024), 5-24.*
ISSN: 2013-2247

Received 20 January 2024

Accepted 19 May 2024

Published November 2024

DOI: 10.1344/DIALECTOLOGIA2024.2024B.2

ANÁLISIS DE LAS VARIEDADES CENTRALES DE LOS DIALECTOS

Juan ABASOLO & Gotzon AURREKOETXEA [*]

Universidad del País Vasco

juan.abasolo@ehu.eus / gotzonaurre@gmail.com

ORCID: 0000-0002-1911-4118 / 0000-0002-9562-4806

Resumen

La geolingüística cuantitativa o dialectometría proporciona información mucho más detallada y fiable sobre la variación geolingüística que los procedimientos tradicionales. Este estudio se centra en el análisis de las variedades centrales de los dialectos vascos. Para ello en primer lugar se ha procedido a la clasificación de las variedades por medio de la clasificación aglomerativa para posteriormente centrarse en las variedades céntricas de cada dialecto.

Palabras clave: geolingüística cuantitativa, clasificación dialectal, variedades céntricas, vasco

ANÀLISI DE LES VARIETATS CENTRALS DELS DIALECTES

Resum

La geolingüística quantitativa o dialectometria proporciona informació molt més detallada i fiable sobre la variació geolingüística que els procediments tradicionals. Aquest estudi se centra en l'anàlisi de les varietats centrals dels dialectes bascos. Amb aquesta finalitat, en primer lloc s'ha procedit a classificar les varietats mitjançant la classificació aglomerativa i posteriorment s'han analitzat les varietats cèntriques de cada dialecte.

Paraules clau: geolingüística quantitativa, classificació dialectal, varietats cèntriques, basc

[*] Bilboko Hezkuntza Fakultatea, Hizkuntzaren eta Literaturaren Didaktika, UPV/EHU University of the Basque Country Sarriena auzoa, z/g, 48940, Leioa, Bizkaia.

© Author(s)

ANALYSIS OF CENTRAL DIALECT VARIETIES

Abstract

Quantitative geolinguistics or dialectometry provides much more detailed and reliable information on geolinguistic variation than traditional procedures. This study focuses on the analysis of the central varieties of Basque dialects. To do so, we have firstly classified the varieties by means of cluster classification and then focused on the central varieties of each dialect.

Keywords: quantitative geolinguistics, dialect classification, central varieties, Basque

1. Introducción

La geolingüística cuantitativa ha revolucionado radicalmente el análisis de la variación geolingüística en distintas facetas que se pueden agrupar en dos grupos: en el uso de grandes cantidades de datos y en el uso de métodos cuantitativos. Todo análisis dialectométrico [1] se basa en grandes cantidades tanto de datos o características lingüísticas como de localidades a analizar. Todo ello repercute en una mayor objetividad de los análisis y, por consiguiente, en una mayor fiabilidad de los resultados obtenidos.

La objetividad en la recogida de datos es uno de los requisitos de la ciencia. Toda recogida de datos tiene que estar basada en procedimientos contrastados por el uso mayoritario de la comunidad científica siguiendo una serie de pautas consideradas como objetivas. El uso de grandes masas de datos aleja el riesgo de una selección subjetiva del dialectólogo y acerca más el estudio de la variación geolingüística a los parámetros utilizados en la investigación científica. No se debe obviar sin embargo, que el solo uso de grandes masas de datos no es sinónimo de objetividad; los datos han de tener el rasgo de representatividad de los hechos lingüísticos. Esta condición es indispensable cuando, como es el caso, la investigación no se puede basar en el universo de los datos. En lingüística, como en otras ciencias, se asume que es imposible el uso de todos los datos. Por ello, tenemos que basarnos en una muestra que sea representativa.

[1] Entendemos aquí la dialectometría como sinónimo de geolingüística cuantitativa.

Otro de los rasgos que objetivan la investigación dialectométrica es el uso de métodos cuantitativos y estadísticos. Una vez conformada la base de datos de la investigación, la labor del dialectólogo se circunscribe a argumentar la selección de la unidad de distancia lingüística que será utilizada, por una parte; y por otra, el uso de las técnicas estadísticas que han de ser aplicadas.

El uso de técnicas estadísticas también posibilita al investigador ir más lejos en el análisis. Su investigación ya no se detiene en la detección de las fronteras o la extensión que pueda tener un rasgo lingüístico, sino que permite la jerarquización de las fronteras y las distancias lingüísticas, la visualización de la distancia lingüística utilizando diversas técnicas, el análisis de la correlación entre la distancia geográfica y la diferenciación lingüística, etc.

Estas técnicas, además, permiten reproducir la investigación utilizando los mismos o diferentes procedimientos y técnicas, ya sea por el mismo dialectólogo o por otro, lo que conlleva a la confrontación de los resultados o en su caso a la posibilidad de corregir los pasos equivocados o dudosos, que hayan podido producir resultados diferentes. Como consecuencia de ello, los análisis y resultados son más fiables, más concretos y más exactos, con lo cual gana la investigación dialectológica o geolingüística.

El punto de que se trata en esta contribución incide en avanzar en el conocimiento de la variación, en este caso, dentro de un mismo dialecto; es decir, introducimos técnicas para discernir computacionalmente las variedades centrales dentro del propio de dialecto.

2. Metodología y desarrollo computacional

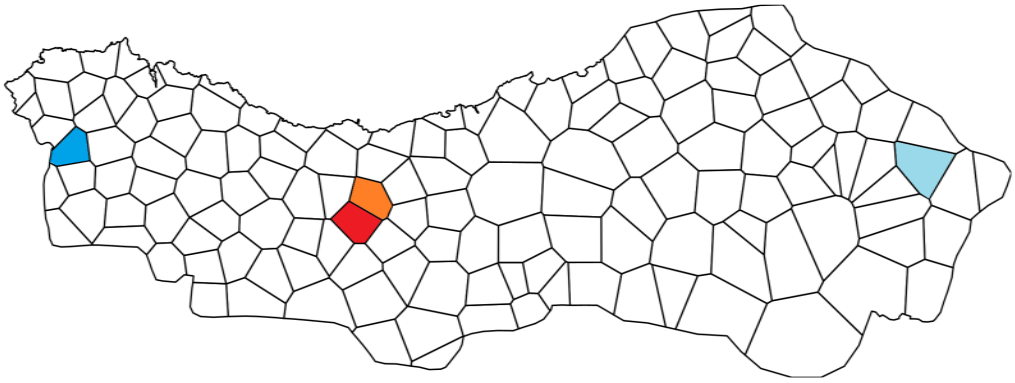
La metodología seguida en este estudio está relacionada con la clasificación aglomerativa jerárquica de las variedades. Posteriormente se ha introducido una nueva técnica computacional para averiguar cuáles son las variedades centrales de

cada dialecto atendiendo al criterio del promedio de las distancias lingüísticas entre ellas.

2.1 Metodología

La base de datos que se ha utilizado es el *Euskararen Sailkapen Berria-ESB* (Aurrekoetxea, Gaminde, Ormaetxea & Videgain 2019) que es el corpus utilizado para la clasificación dialectal del vasco, por medio de procedimientos dialectométricos. Se trata de un corpus entresacado de una muestra de los datos recogidos en el proyecto *Euskararen Herri Hizkeren Atlasa-EHHA* o atlas lingüístico vasco, llevado a cabo bajo los auspicios de Euskaltzaindia (Aurrekoetxea & Videgain 2010-2020). La muestra entresacada de los datos del atlas fue recogida en una base de datos después de una depuración y una lematización (en el léxico) o tipificación (en las características gramaticales (que corresponde *grosso modo* a las *taxats* en la nomenclatura dialectométrica de H. Goebl). El paso inmediatamente posterior a la *taxación* o etiquetación y lematización de los datos fue la aplicación de la distancia de similitud-RIV (Goebl 1983: 115, 1987: 70, 2013: 145), mediante la cual se logró la matriz de distancias lingüísticas entre las localidades analizadas. Esta matriz de distancias expresa la diferenciación lingüística entre las localidades en cifras de modo que la diferencia o distancia lingüística entre dos localidades se expresa en %, puesto que la distancia 0 equivale a identidad perfecta y 100 a diferenciación total.

Esta matriz proporciona ya datos de distancias lingüísticas entre variedades. A modo de ejemplo se puede citar que la distancia más pequeña de las 145 variedades analizadas en el estudio se da entre las localidades de Errezil (en naranja) y Beizama (en rojo) con un 13,65% de distancia o diferenciación lingüística; y la mayor distancia, en cambio, se da entre las localidades Sohüta (en azul claro) y Etxebarri (azul oscuro) con un 66,07% de diferenciación (Mapa 1).



Mapa 1: Localidades más cercanas lingüísticamente (en rojo) y más alejadas (en azul)

Esta matriz de distancias tiene una gran importancia; se trata de la primera matriz que proporciona a los dialectólogos distancias cuantitativas entre 145 variedades dialectales del vasco. Este hecho supera *de facto* la fase de suposiciones, creencias, opiniones, etc. sobre las diferencias lingüísticas entre variedades. Ya no hay necesidad de hacer hipótesis sobre la magnitud de diferenciación lingüística. Esa fase de la dialectología está ya superada. Los datos cuantitativos indican con precisión la distancia de una variedad a otra, de un dialecto a otro, cualquiera que sea el punto de partida o de llegada. Se podrá defender o si una distancia determinada pueda ser considerada como grande o como pequeña; en otras palabras, que la distancia lingüística máxima entre variedades (el 66%) es grande o pequeña, mayor o menor de la que se suponía o se esperaba, teniendo en cuenta la literatura anterior sobre el tema, cuando se consideraba que hay grandes distancias entre los dialectos en la lengua vasca. El dialectólogo cuenta con una unidad de distancia y una matriz; dicho lo cual, creemos que dicha matriz puede ser mejorada y que en un futuro se podrían llegar a medir las distancias lingüísticas con procedimientos o técnicas incluso más precisas.

Es cierto que en el vasco se dan unas distancias lingüísticas que no se dan en otras latitudes y otros idiomas. No es momento de profundizar en las razones de este

hecho. Pero uno de los factores más importantes de dichas diferencias creemos que hay que encontrarlos en la tardía estandarización del vasco (el vasco se comenzó a estandarizar en los años 60 del siglo XX).

Sin duda la historia de la lengua y del pueblo vasco es un factor determinante tanto en las distancias lingüísticas entre las variedades como en los datos de correlaciones entre la distancia geográfica y la diferenciación lingüística (Aurrekoetxea 2010, Aurrekoetxea, Abasolo & Videgain 2021a,b). A primera vista parece que hay una confrontación entre la alta correlación encontrada entre la distancia geográfica y la diferenciación lingüística, por una parte, y las mencionadas distancias lingüísticas tan grandes entre localidades, por otra. Pero no hay tal confrontación, sino, como primera hipótesis, un poso de muchos siglos, en los que las variedades se han ido paulatina y constantemente diferenciando, sin grandes interrupciones o migraciones, entre localidades y regiones que componen la tierra del vasco.

2.2 Desarrollo computacional

Se han desarrollado dos técnicas computacionales en la búsqueda de las zonas o variedades céntricas de los dialectos: el análisis de conglomerados y el análisis de las distancias entre las variedades dentro de cada dialecto.

El análisis de conglomerados *(cluster analysis)* se ha llevado a cabo con la aplicación *Diatech*[2] utilizando la unidad de distancia RIV (ver por ejemplo Goebl 1992: 436-438) y el algoritmo Ward.

Para determinar la medida de centralidad de las variedades dentro de los ya determinados dialectos, se utilizaron submatrices de la matriz de distancias originales; una por dialecto. La medida de centralidad o periferia de cada variedad viene determinada por la media matemática de la distancia de cada variedad al resto de variedades del dialecto y su posterior clasificación, en función de estas medidas: los resultados se expondrán mediante diagramas y mediante mapas. Se utilizarán los

[2] http://eudia.ehu.eus/diatech/index.

dendrogramas, por una parte, y por otra, los mapas de las variedades dialectales del vasco.

3. Resultados

Primeramente, se llevará a cabo la clasificación jerárquica de los dialectos, para, partiendo de ella, mostrar los resultados de las variedades centrales de cada uno de ellos. Los mapas y dendrogramas de las siguientes secciones dan cuenta de los resultados obtenidos.

3.1 Dialectos del vasco según la clasificación jerárquica

A lo largo de la historia de la dialectología vasca ha habido diversas clasificaciones dialectales atendiendo a criterios y marco teóricos diferentes (Ensunza & Unamuno 2023). La última de ellas (Aurrekoetxea, Gaminde, Ormaetxea & Videgain 2019) ha sido elaborada utilizando grandes masas de datos (649 características lingüísticas en 145 localidades, lo que da un corpus de alrededor de 100.000 ítems, teniendo en cuenta también las respuestas múltiples recogidas en muchas localidades) y mediante técnicas dialectométricas. Según esta clasificación, el euskera cuenta con tres dialectos (Figura 1 y Mapa 1): occidental, central y oriental.

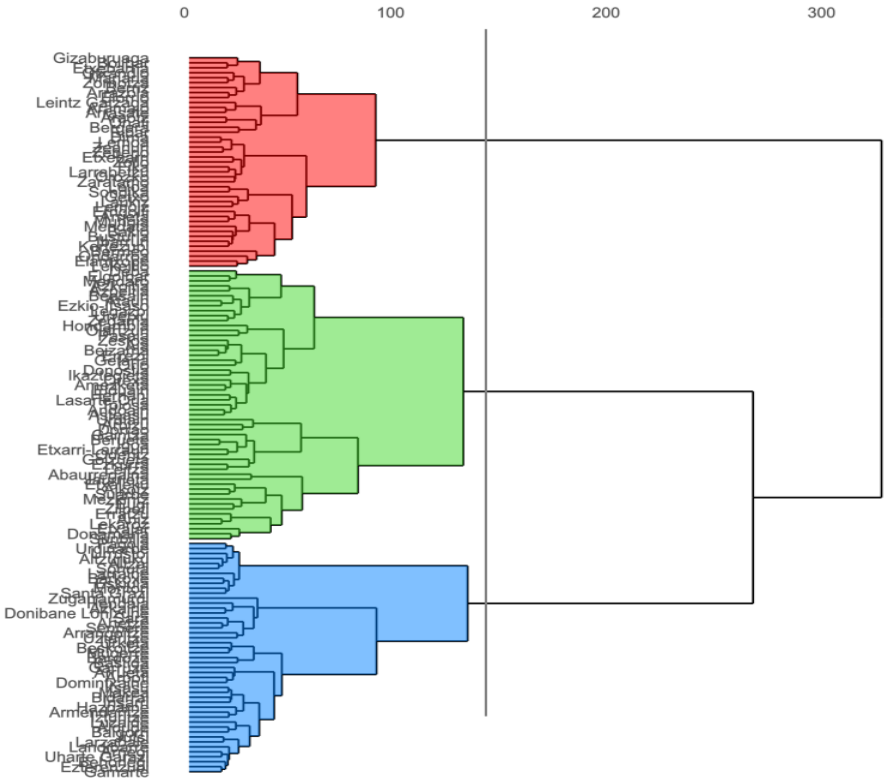


Figura 1. Dendrograma de los dialectos vascos (Aurrekoetxea *et al*. 2019: 126)

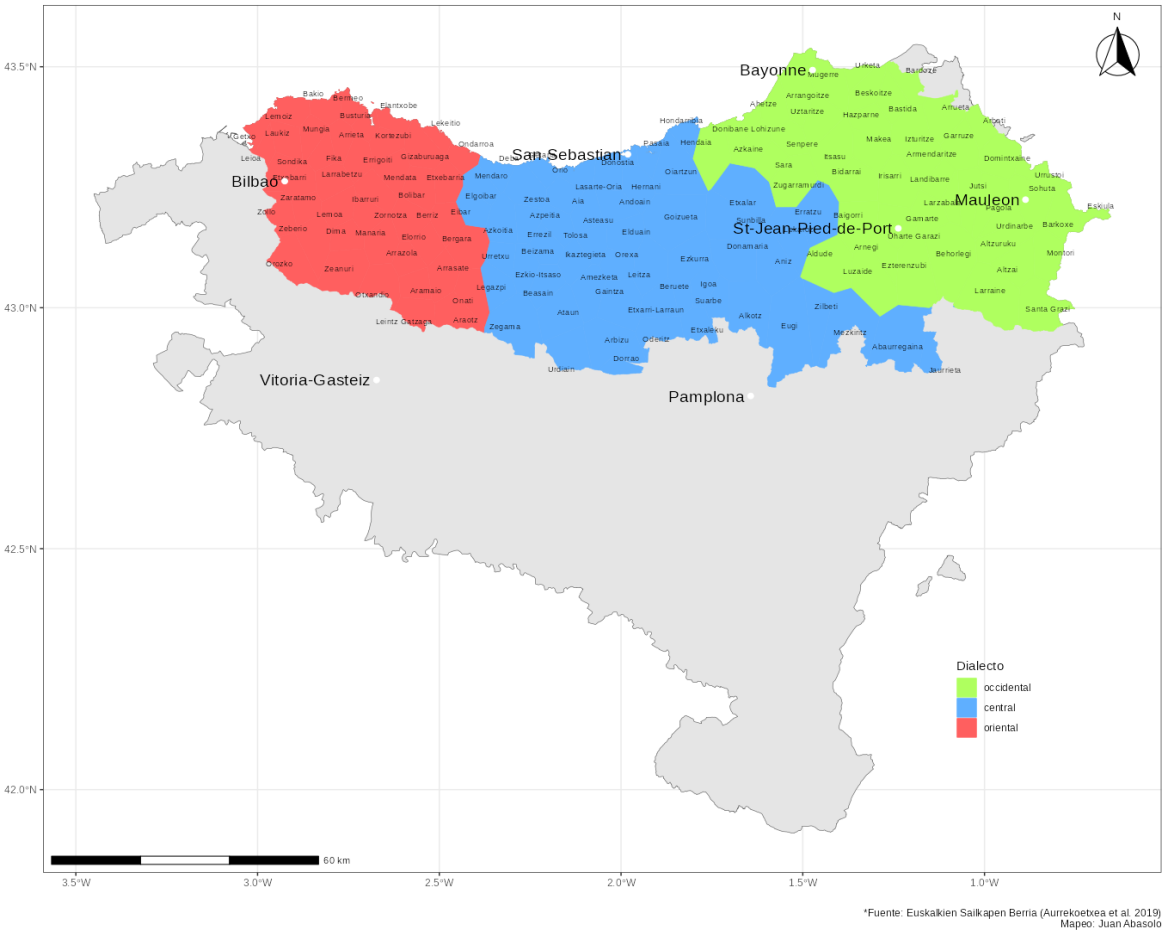
La delimitación de los dialectos es un tema no resuelto aún hoy en día. A pesar de que con la utilización de grandes masas de datos y recurriendo a técnicas estadísticas como clasificación aglomerativa jerárquica se logren resultados de mayor fiabilidad que por medio de recursos tradicionales, faltan todavía decisiones compartidas sobre las técnicas a utilizar para el corte del dendrograma y que puedan ser asumidas por la comunidad científica. En diversas aplicaciones dialectométricas el análisis de la clasificación jerárquica proporciona un dendrograma en el que las variedades analizadas se agrupan según su mayor o menor homogeneidad interna y un mapa en la que se representan los dialectos. Por muy estimable que sea la ayuda que proporciona la automatización de los procesos y la estadística informatizada, es labor del dialectólogo en todo momento la toma de decisiones. Y entre ellas se encuentra la decisión de la determinación del nivel de diferenciación entre dialectos. En este caso, como ayuda, la estadística descriptiva usa criterios que pueden ser aplicados también

a datos lingüísticos. Uno de ellos es que en la creación del dendrograma las ramas del árbol resultan más largas cuanto mayor es distancia lingüística que representan. En el dendrograma de la Figura 1 se pueden distinguir tres grandes ramas, cada una de ellas representando un grupo dialectal. Es criterio universalmente asumido que cuando se trata de dar un corte al árbol jerárquico representado en el dendrograma ese corte se ha de llevar a cabo en niveles en los que las ramas sean de la mayor longitud posible. En nuestro caso el corte del dendrograma se ha llevado a cabo en el nivel 160 de la escala (del total de 360), lo cual representa el 44,4 % de la variación expresada la matriz de distancias (ver la línea vertical en la Figura 1). Esta decisión ha sido tomada por criterios puramente estadísticos, obviando incluso las variedades que puedan estar incluidas o no en algunos de los grupos.

Entendemos también deseable y necesaria la discusión sobre la magnitud de la variación que debe reunir como mínimo un área dialectal para ser considerada como dialecto. Es una cuestión que debe abordar sin dilatar la geolingüística cuantitativa. Mientras no se dé esta discusión consideramos que, al menos en el caso del vasco, se puede considerar idóneo cualquier cifra por encima del 40% de toda la variación lingüística contenida en la base de datos.

3.2 Mapa de los dialectos del vasco

El primer paso es la determinación de los dialectos del vasco. Transportando la clasificación proporcionada por el dendrograma al mapa se logra el mapa de los dialectos vascos (Mapa 2).



Mapa 2. Dialectos vascos (Fuente: mapa de J. Abasolo, partiendo de los datos de Aurrekoetxea *et al* 2019: 127)

El euskera cuenta con tres dialectos: occidental (en rojo) que alberga 43 localidades analizadas, central (en azul) con 55 localidades analizadas y oriental (en verde) con 47 localidades. Los dialectos occidental y central ocupan el territorio peninsular del área vasca, mientras que el dialecto oriental se asienta en la parte continental, en la zona septentrional de los Pirineos.

Esta metodología permite el análisis a diferentes niveles de la escala que proporciona el dendrograma para medir la distancia lingüística. Utilizando esta misma escala, ¿a qué nivel habría que cortar el dendrograma para poder hablar de subdialectos o, dicho de otra manera, qué nivel de similitud entre las variedades sería tolerable para hablar de un subdialecto? En la investigación referida más arriba para poder hablar de subdialectos habría que bajar hasta el nivel 85 (Figura 2). A este nivel

de la escala se recoge el 25% de la distancia lingüística. El corte llevado a cabo a dicho nivel presenta 7 grupos dialectales. En concreto el dialecto occidental se partiría en dos, otro tanto el dialecto central y el oriental se divide en tres grupos (con variedades que tienen distancias lingüísticas inferiores al 25% entre sí).

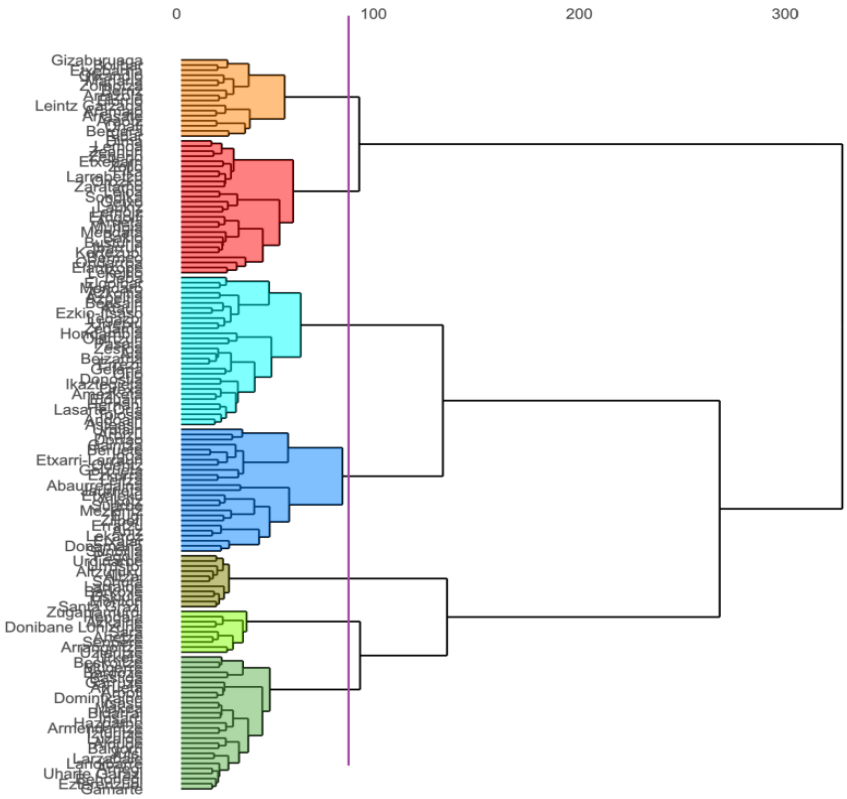
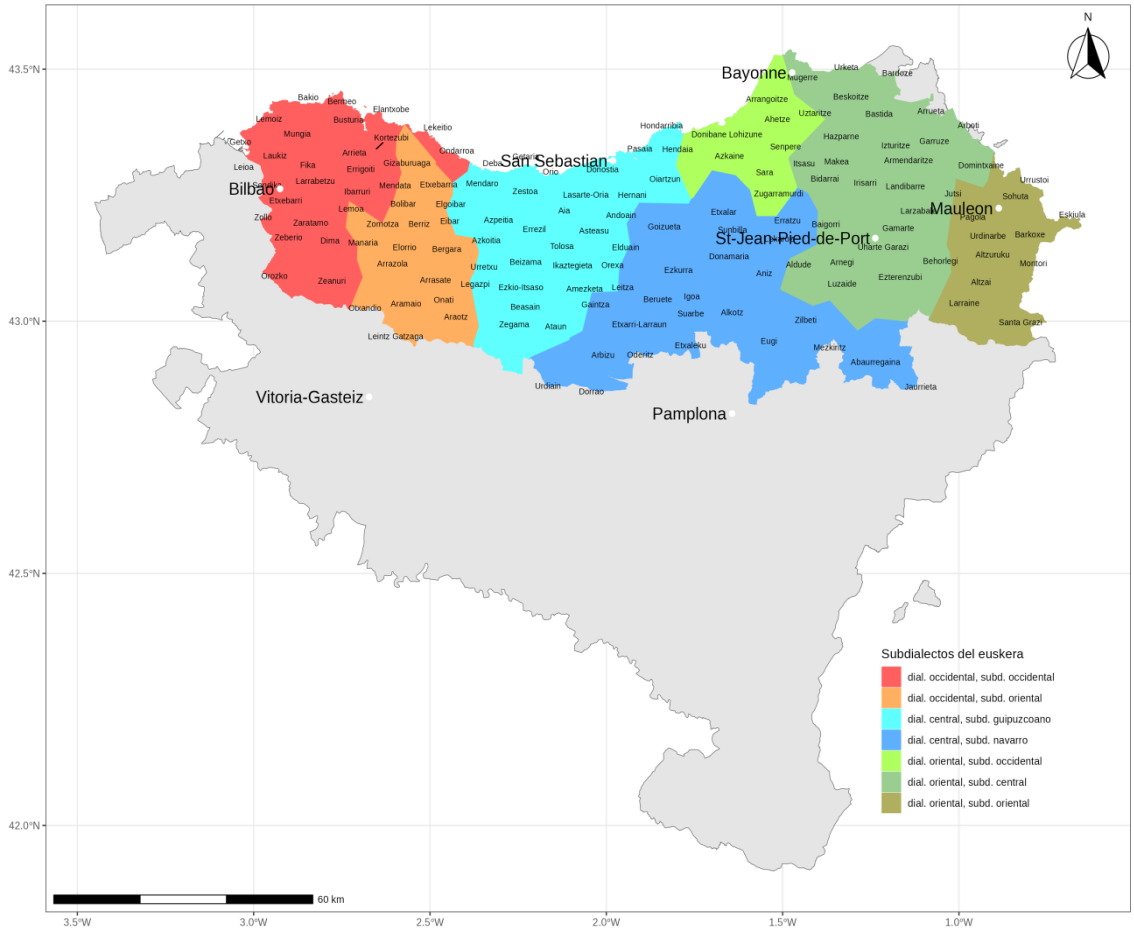


Figura 2. Dendrograma de los subdialectos vascos (Fuente: Aurrekoetxea *et al*. 2019: 138)

El dialecto occidental se divide en dos subdialectos, siendo el subdialecto occidental el más numeroso en cuanto a las variedades que alberga (27 localidades por 16 el subdialecto oriental), aunque su predominio no sea tan evidente geográficamente. El dialecto central también cuenta con dos subdialectos: el guipuzcoano (en azul claro) y el navarro (en azul intenso), que agrupa todas las localidades navarras del dialecto central. El dialecto oriental se divide en tres subdialectos: el occidental, el central y el oriental. *Groso modo,* estos tres subdialectos son los tres dialectos en la clasificación

de Azkue (1905). Es de resaltar el grado de singularidad que tiene el subdialecto oriental (en verde oscuro); es el grupo dialectal más singular y es el de mayor entidad propia, que se une a los subdialectos central y occidental del mismo dialecto al mismo nivel que se unen los subdialectos guipuzcoano y navarro del dialecto central y presenta menor distancia intragrupal entre sus variedades.



Mapa 3: subdialectos vascos (fuente: Aurrekoetxea *et al*. 2019: 139)

Esta nueva configuración del mapa de los subdialectos y dialectos del vasco tiene semejanzas y desigualdades con respecto a los mapas dialectales anteriores: el dialecto occidental ocupa casi el mismo espacio geográfico que dicho dialecto ocupa en otras clasificaciones, aunque a nivel de subdialecto difiere sobre todo en la parte septentrional central del dialecto. El dialecto central ocupa el espacio que ocupan dos

o tres dialectos en clasificaciones anteriores (dos en el caso de la clasificación de Zuazo (2008) y tres en la de Bonaparte 1863). En el dialecto oriental ocurre otro tanto (dos en el caso de la clasificación de Zuazo y cuatro en la de Bonaparte).

Este hecho nos aboca a un sinfín de cuestiones. Sin ninguna intención de tratar todas ellas y centrándonos únicamente en una, nos podríamos preguntar sobre el grado de diferenciación que debe haber entre dos grupos dialectales para que puedan ser denominados dialecto o subdialecto. Siguiendo el procedimiento de este artículo el porcentaje de aproximadamente del 25% a 40% de la distancia lingüística correspondería al subdialecto, de ahí en adelante al dialecto. Todo ello a modo de propuesta para el planteamiento de una discusión en profundidad sobre el grado de distintividad que debe tener un grupo dialectal para poder ser tenido como dialecto o subdialecto. Como se desprende, esta metodología da por superada la basada en isoglosas o dialectología isoglótica (Aurrekoetxea, Ensunza, Ščofic & Van de Velde 2023).

3.3 Variedades céntricas de los dialectos

La dialectología cuantitativa es operativa no solo para la medición de la distancia lingüística entre variedades —la clasificación jerárquica—, sino incluso para analizar tanto las variedades o zonas de transición, como las variedades céntricas de cada dialecto o área dialectal, entre otras. Ahondando en el tema, creemos que es interesante y necesario, desde nuestra posición teórica, analizar las variedades céntricas de los dialectos y determinar la distancia lingüística desde esas variedades céntricas a las variedades fronterizas. Entendemos también que esta cuestión está relacionada con la idea de la difusión lingüística, aunque no tenga el mismo punto de partida. Entendemos como variedad céntrica aquella que logra mejores promedios de distancia lingüística con relación a todas las variedades de un mismo dialecto. Y concebimos como mejores promedios las distancias más pequeñas entre variedades. Como hipótesis propugnamos que las zonas céntricas de cada dialecto o zona dialectal

se sitúan en zonas geográficamente centrales de cada dialecto y reúnen características lingüísticas que se han ido expandiendo paulatinamente a lo largo del mismo. Esta hipótesis defiende que cada característica haya podido expandirse de modo autónomo e independientemente tanto de la velocidad de expansión como del vigor e intensidad. Como resultado del mismo, aquellas características más vigorosas y que hayan experimentado una velocidad de expansión más alta habrían podido divulgarse a zonas más alejadas.

De acuerdo a la matriz de distancia utilizada en esta investigación, es la variedad de Mañaria la que ostenta el título de la variedad más céntrica del dialecto occidental; es decir, la variedad que tiene mejores relaciones lingüísticas o más afinidad lingüística con las restantes variedades del dialecto occidental. El promedio de las distancias lingüísticas de esta variedad con las restantes de su dialecto es 25,128%. Las siguientes variedades son Larrabetzu e Ibarruri. En el otro extremo, la variedad que peor afinidad lingüística registra en este dialecto es Bergara, variedad de frontera con el dialecto central. En el dialecto central la variedad céntrica es Beizama, seguida de Errezil y Ataun. En este dialecto el promedio de las distancias que la variedad más céntrica tiene con todas las demás variedades del dialecto es 27,869%. Es decir, dos puntos y medio más que en el caso del dialecto occidental; lo cual viene a decir que en el dialecto central la dispersión de las variedades es mayor que en el occidental. En cuanto se refiere al dialecto oriental la variedad más céntrica es Landibarre, seguida de Gamarte y Uharte Garazi. Landibarre tiene un promedio de 25,442%, muy parejo con el dialecto occidental.

La variedad de Mañaria se sitúa en el subdialecto oriental del dialecto occidental, la variedad de Beizama en el subdialecto guipuzcoano del dialecto central y la variedad de Landibarre en el subdialecto central del dialecto oriental.[3]

Este análisis quedaría cojo si no se analizaran los mayores promedios de cada dialecto: en efecto, en el dialecto occidental el promedio mayor corresponde a Bergara que tiene un 34,991%. Teniendo en cuenta que el promedio más bajo es el logrado por

[3] Las variedades céntricas se representan en blanco en los mapas correspondientes.

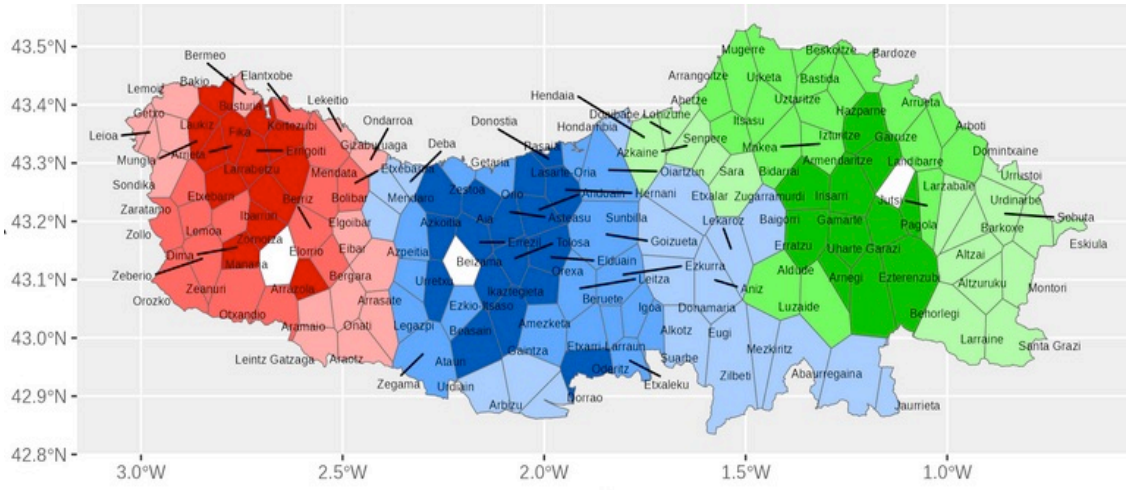
Mañaria (24,589%), la dispersión de los promedios es el 10,393% en el dialecto occidental. En el dialecto central el promedio menor corresponde a Beizama y el mayor a Jaurrieta (42,666%), con una dispersión del 14,797%. Por último, en el dialecto oriental el promedio más bajo lo detiene Landibarre y el más alto corresponde a Zugarramurdi con 37,711%, con una dispersión que se sitúa en 12,269%. Por lo tanto, de acuerdo con estos datos, el dialecto más compacto es el occidental y el que más dispersión contiene el dialecto central.

3.4 Estructuración de los dialectos

Además de la determinación de la variedad céntrica de cada dialecto, la geolingüística cuantitativa posibilita el análisis de la estructuración de las variedades en cada dialecto. Logrado el listado de todas las variedades de cada dialecto según promedios de distancia lingüística que cada variedad haya logrado con respecto a las restantes variedades, se pueden agrupar atendiendo a los índices logrados. De acuerdo con ello, las variedades se pueden agrupar en grupos diferentes (aquí se han agrupado en tres grupos) atendiendo a los valores de las distancias de cada uno de ellos o agrupando equitativamente el número de variedades contenidas en cada grupo.

3.4.1 Estructuración del dialecto según número de variedades

Se han agrupado las variedades de cada dialecto en tres grupos configurando cada uno de ellos con igual número de variedades. En el dialecto occidental, las localidades con porcentajes de similitud mayores han sido coloreadas con rojo intenso, y las de menor similitud con rojo pálido. En el central y el oriental se ha procedido de la misma forma con el azul y el verde (Mapa 4).

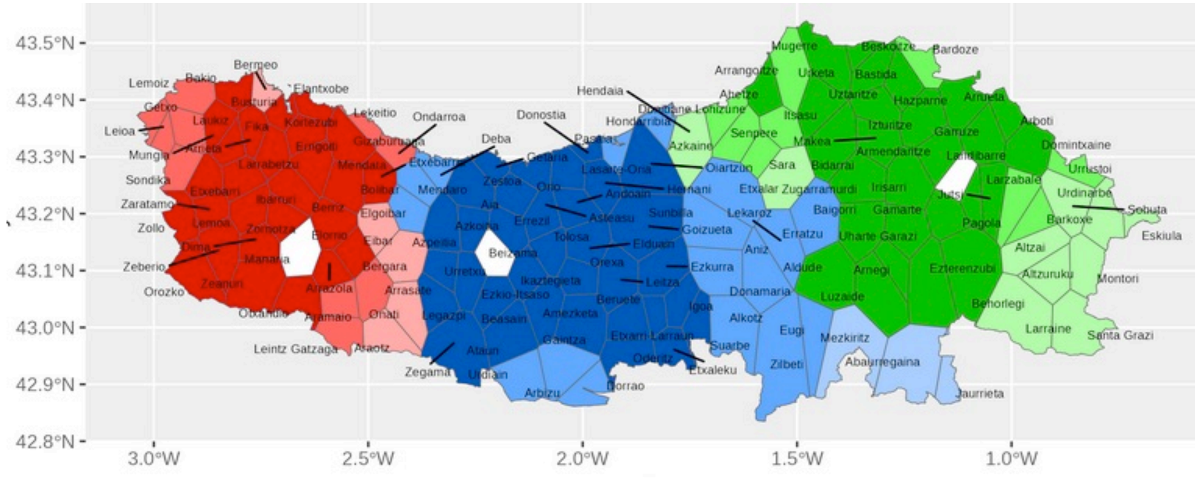


Mapa 4. Estructuración de los dialectos según número de variedades

La estructuración del dialecto occidental es geográficamente equilibrada, con una región céntrica y dos regiones laterales a cada lado. Es de resaltar que todas las variedades pertenecientes a Gipuzkoa se hallan situadas en el grupo de variedades con distancias lingüísticas más altas (Eibar, Bergara, Arrasate, Oñati, Aramaio y Araotz). En el dialecto central las variedades céntricas se sitúan en la provincia de Gipuzkoa, excepto Oderitz, y se hallan escoradas hacia el occidente; consecuentemente, las variedades con mayores distancias lingüísticas del dialecto se sitúan en la zona oriental de Nafarroa. Por último, en el dialecto oriental se puede apreciar un equilibrio geográfico, tanto en cuanto a las variedades más céntricas como a las que computan mayores distancias lingüísticas; es de reseñar que todas las variedades analizadas del subdialecto oriental excepto Domintxaine pertenecen al tercer grupo.

3.4.2 Estructuración del dialecto según la distancia lingüística

También se ha llevado a cabo agrupaciones atendiendo a la distancia lingüística. Para ello se han creado 3 grupos teniendo en cuenta que cada uno de ellos agrupara la misma distancia lingüística (Mapa 5).



Mapa 5. Estructuración de los dialectos según la distancia lingüística

De las 43 variedades analizadas del dialecto occidental, 14 se hallan en el segundo y tercer grupos (el 32,56%), mientras que el 67,44% pertenecen al grupo más céntrico (29 variedades). En el dialecto central de las 55 variedades de las que está compuesta, 20 pertenecen al segundo y tercer grupos (36,37%) y 35 al grupo más céntrico (63,63%). En lo que concierne al dialecto oriental, de las 47 variedades que la componen 19 variedades se sitúan en los grupos segundo y tercero (40,43%) y 28 al grupo más céntrico (59,57%).

Según estos datos el dialecto más cohesionado es el occidental, con un grupo central que alcanza los porcentajes más altos. En segundo lugar, se encuentra el dialecto central y en tercer lugar el dialecto oriental.

En cuanto a la situación geográfica tanto el dialecto occidental como el oriental cuentan con una distribución más equitativa que el dialecto central que se encuentra escorada hacia la parte occidental.

4. Conclusiones

La geolingüística cuantitativa o dialectometría proporciona posibilidades hasta ahora inconcebibles en el análisis de la variación. El análisis de las variedades céntricas de los dialectos se ha llevado a cabo partiendo de una clasificación jerárquica que proporcionan los programas informáticos para análisis dialectométricos, y utilizando herramientas de R Statistical Software (v4.3.0; R Core Team 2023) tal como (Kassambara 2023), con las que se puede avanzar con procedimientos no recogidos en dichos programas.

El objetivo de este estudio ha sido la determinación de las variedades céntricas de los dialectos. La estructuración de los dialectos y el estudio de las variedades céntricas amplia el conocimiento de cada dialecto delimitando las variedades con mejor interactividad lingüística. Estimamos que los resultados obtenidos son interesantes, puesto que, entre otras cosas, permiten visualizar la estructuración geolingüística de cada dialecto. Con ello se abre una nueva hipótesis basada en que las variedades céntricas del dialecto como poseedoras de los porcentajes más bajos de distancia lingüística, poseen también, hipotéticamente, la mayor cantidad de características propias de dicho dialecto. Esta hipótesis discurre del ámbito de la cuantificación al aspecto cualitativo que es en definitiva lo verdaderamente importante en dialectología. Por lo tanto, este análisis cuantitativo queda incompleto si no se llega a un análisis cualitativo en la que se estudien las características lingüísticas más extendidas y que más influencia tienen en la configuración de cada grupo dialectal. Es un trabajo que debe ser abordado en un futuro próximo.

Referencias

AURREKOETXEA, Gotzon (2010) “The correlation between morphological, syntactic and phonological variation in the Basque language”, in B. Heselwood & C. Upton (eds.),

Proceedings of Methods XIII. Papers form the Thirteenth International Conference on Methods in Dialectology, 2008, Frankfurt and Main: Peter Lang, 207-218.

AURREKOETXEA, Gotzon, Juan ABASOLO & Xarles VIDEGAIN (2021a) "Corrélation entre la différenciation linguistique et la distance géographique en basque", en Manuela Nevaci, Irina Floarea & Ioan-Mircea Farcas (eds.), *Ex Oriente Lux in Honorem Nicolae Saramandu.* Alessandria: Edizioni dell'Orso, 127-140.

AURREKOETXEA, Gotzon, Juan ABASOLO & Xarles VIDEGAIN (2021b) "Hizkuntza-distantziaren eta distantzia geografikoaren arteko korrelazioa", in Naia Eguskiza & Aitor Iglesias (eds.), *Hezkuntzari buruzko ekarpenak (EUDIA-9),* Bilbao: UPV/EHU Servicio Editorial, 23-40.

AURREKOETXEA, Gotzon, Ariane ENSUNZA, Jožica ŠKOFIC & Hans Van de VELDE (2023) "An introduction to dialect classifications in Europe", *Dialectología*, Special issue X, 1-17.

AURREKOETXEA, Gotzon & Xarles VIDEGAIN (dir.) (2010-2010) *Euskararen Herri Hizkeren Atlasa-EHHA,* 11 tomos, Bilbao: Euskaltzaindia.

AZKUE, Resurrección M. de (1906) *Diccionario Vasco-Español-Francés—Dictionnaire Basque-Espagnol-Français*, Paris: Paul Geuthner.

BONAPARTE, L.L. (1863) *Carte des sept provinces basques, montrant la délimitation actuelle de l'euscara et sa division en dialectes, sous-dialectes et variétés,* Londres: Stanford's Geographical Establishment.

ENSUNZA, Ariane, Lorea UNAMUNO (2023) "Basque dialect classifications", *Dialectologia,* Special Issue X, 69-91.

GOEBL Hans (1992) "Problèmes et méthodes de la dialectométrie actuelle (avec application à l'AIS)", in G. Aurrekoetxea & X. Videgain (eds.), *Actas del congreso internacional de dialectologia (Bilbao1991-X-21/25),* Bilbao: Euskaltzaindia, 429-476.

GOEBL, Hans (1983) "Matrices, distances et interpoints. Etude dialectométrique sur les isoglosses quantitatives", in Marie-Rose Simoni-Aurembou (ed.), *Cahier des Annales de Normandie n° 15. Dialectologie et littérature du domaine d'oïl occidental: actes du colloque tenu à l'Université de Caen en février 1981,* 113-137. DOI: 10.3406/annor.1983.3899

GOEBL Hans (1987) "Points chauds de l'analyse dialectométrique: pondération et visualisation", *Revue de linguistique romane*, 51, 63-118.

GOEBL Hans (2013) "La dialectometrización del ALPI: rápida presentación de los resultados", en E. Casanova Herrero y C. Calvo Rigual (eds.), *Actas del XXVI Congreso Internacional de*

Lingüística y de Filología Románicas (Valencia 2010), Berlin/Boston: Walter de Gruyter, vol. VI, 143-154.

KASSAMBARA, Alboukadel (2023) *rstatix: Pipe-friendly framework for basic statistical tests [Manual]* <https://CRAN.R-project.org/package=rstatix>

R CORE TEAM (2023) *R: A language and environment for statistical computing [Manual],* R Foundation for Statistical Computing. <https://www.R-project.org/>

ZUAZO, Koldo (2008) *Euskalkiak, euskararen dialektoak,* San Sebastián: Elkar.