Names | A Journal of Onomastics

Surname Distribution in Galicia: Population Structure, Dialects, and History

Xulio Sousa
Universidade de Santiago de Compostel, SPAIN

Rosa María Crujeiras
Universidade de Santiago de Compostel, SPAIN

María José Ginzo Villamayor
Universidade de Santiago de Compostel, SPAIN

ans-names.pitt.edu
ISSN: 0027-7738 (print) 1756-2279 (web) Vol. 73 No. 2, Summer 2025 DOI 10.5195/names.2025.2699

Abstract

The study of the geographic distribution of surnames is a useful source of information for probing the structure of populations and their links with history, and language. The relationship between the transmission of surnames and the dissemination of some genetic markers explains why family names continue to be a rich depository of information for geneticists, human biologists, and anthropologists. In Western societies, family names are distributed in space according to patterns determined by the origin of their bearers, the mobility of populations, and the social and cultural characteristics of communities. Consequently, studying the distribution of surnames is one way of reconstructing the history of populations. This study considers the use of surnames as a basis to regionalize the surname space of Galicia, a cultural, linguistic, and historical community in northwest Spain. We studied the isonymy structure using the distribution of single surnames in 315 administrative divisions to discover traces of historical phenomena. The analysis was conducted according to procedures that have already been used in previous studies, but with some modifications that enhance the value of the information analyzed. The results allow us to identify onomastic regions that show interesting correlations with dialect and historical divisions of the Galician territory.

Keywords: family names, surname regions, isonymy, population change, dialects, anthroponymy, Galician language

1. Introduction

1.1. Background

The studies on the distribution of surnames among populations from different parts of the world have, for over half a century, shown how useful onomastic analysis is as a complement of genetics in order to probe more deeply into the structure of populations and their links with the history, language, and culture of human communities (Guppy 1890; Lasker 1980; Piazza et al. 1987; Boattini et al. 2012; Cheshire 2014; Kandt et al. 2016; Sousa & Ginzo-Villamayor 2020; Caramelli et al. 2021). Surnames are historical signs of identity in social groups and cultural traits whose transmission bears strong similarity with that of some biological features. If high-quality onomastic data are available, it becomes feasible to use them as a source of information to reveal the interaction between culture, language, history, and biology. Surnames function as cultural and biological markers that help to identify ethno-linguistic groups living in the same area. The most recent research shows that surnames can be used as an interesting complement in the study of modern populations (Jobling 2012; Solé-Morata et al. 2015; Boattini et al. 2021; Kamel et al. 2021).

Over the last few decades, the studies on the distribution of surnames in Western societies have shed light on the distant history of communities and served to help recognize the link between this history and language diversity, both at the macro- and micro-space levels (Goebl 1996; Manni & Barrai 2001; Manni et al. 2008; Scapoli et al. 2005; Mateos et al. 2011; Boattini et al. 2012). The European populations were the object of several investigations of this kind, covering both extensive areas (Cheshire et al. 2011; Kandt et al. 2020) and smaller natural and administrative regions (Morelli et al. 2002; Zei et al. 2003; Boattini et al. 2021). The geographic area of the Iberian Peninsula, comprising Spain and Portugal, was the focus of several studies over the last few decades, showing the richness of personal onomastic data to shed some light on the history of its populations (Mateos & Tucker 2008; Calderón et al. 2015; Rodríguez Díaz 2015; Martínez-Cadenas et al. 2016; Hanks et al. 2016; Sousa & Ginzo-Villamayor 2020). The aim of this paper is to contribute to the discovery of these interesting connections between data of such different natures.

ISSN: 0027-7738 (print) 1756-2279 (web) Vol. 73 No. 2, Summer 2025 DOI 10.5195/names.2025.2699

1.2. Surnames, Society and History

The dynamics of human population structure is driven by two processes, drift and migration, and these can be identified and described through isonymy analysis. Drift and migration determine differently the typical characteristics of the structure of the populations in countries with different histories of emergence and development. In areas of America with important migration flows and population movements that were significant and long-lasting, migration had a direct impact on their population structure (Piazza et al. 1987). In contrast, in western European countries, where a large part of the population remained stable for centuries, drift was the key factor in the organization of populations (Martínez-Cadenas et al. 2016).

In Western Europe, surnames started to be passed along patrilineal lines in the Middle Ages. This system was officially established in several countries starting in the late 18th century (Scott et al. 2004). The latest investigations show that in many European societies the patrilineal system had been used by a large part of the population from time immemorial (Winney et al. 2012).

In Spain, the Civil Code of 1870 included the obligation to keep a uniform family name and to pass it down from parents to children. The state administration required every individual to be identified by a name and two surnames, the first one inherited from the father and the second one from the mother. In this way, a custom that had been in use for centuries, and which was characteristic of the Spanish population from the perspective of surname structure, acquired official status (Rodríguez-Larralde et al. 2003; Salinero & Testón Núñez 2010). Onomastic investigations confirm that the origins of this system date back to the late 10th century, when a two-name system (name + surname) started to be used, which would eventually spread and become established in the different peninsular areas from the 12th century onwards (Kremer 2003). Through comparison of genetic and onomastic data, researchers also determine that the oldness of surnames among the Spanish population can vary between 200 and 800 years (Martínez-Cadenas et al. 2016; Ferreiro et al. 2021).

Surnames in Spain and Portugal often come from one of three categories: (i) first name of the father (patronymics); (ii) origin names (place names); and (iii) lexically derived names. These three basic categories gave rise to most surnames of the current peninsular population, and they were also the ones taken to America through the Spanish and Portuguese colonization processes (Kremer 2003). Surnames from patronymic origin (for example, *García*, *Fernández*, *González*) are the most frequent ones among the Spanish population, followed by those deriving from nouns and those from toponymic origin (Faure et al. García 2009). However, the proportion of these frequencies is not the same throughout the different linguistic areas of the Iberian Peninsula (Moreu-Rey 1993; Boullón Agrelo 2008; Rodríguez-Díaz et al. 2017). In Catalonia, occupational surnames are quite common (for example, *Ferrer* 'blacksmith', *Sabater* 'shoemaker', *Teixidor* 'weaver'), whereas in Galicia and in the Basque Country, surnames from toponymic origin are far more numerous (for example, *Castro*, *Souto*, *Vila*, *Ibarra*, *Zabala*, *Agirre*).

The Iberian population, analyzed overall from the onomastic perspective, seems to be the most uniform among the European populations. This uniformity is a result of several causes, the two main ones being: (i) low external migration flows for centuries; and (ii) late internal migration movements, a result in turn of late industrialization and urbanization processes (Lopez-Gay 2014). A method of patronymic surname creation that was common throughout the peninsular lands from the Middle Ages has been successfully spread, and this is a crucial element in characterizing the population: patronymic surnames (first name + *-ez*/*-es*, *Fernando* > *Fernández*). A consequence of the establishment of this system is the abundance of this type of family names across the different linguistic and cultural areas of Spain and Portugal (for example, *Fernández*, *Fernandes*, *Rodríguez*, *Rodrigues*, *González*, *Gonçalves*).

From the point of view of the abundance of surnames, the Spanish population appears to be the most uniform among the European populations studied to date. This uniformity is typical of the central regions (Castilla, León, and La Mancha), and of the northern regions of the Peninsula (Galicia, Asturias, and the Basque Country). The most common surnames in Spain are also the most common ones in every individual Spanish region (for example, *González*, *Rodríguez*, *Gómez*). In addition, it must be borne in mind that the northern part of the Peninsula is the area considered by some scholars to be the birthplace of most of the surnames that later spread across Spain and Portugal during and after the Reconquest (Mir de la Cruz 1981). A result of this historical process is that in the southern Peninsula there is a very low rate of monophyletic surnames, and that the most common ones among them mostly have a Castilian-Leonese origin (Calderón et al. 2015).

DOI 10.5195/names.2025.2699 Vol. 73 No. 2, Summer 2025 ISSN: 0027-7738 (print) 1756-2279 (web)

The present study aims to analyze the relationship between surname geographical distribution, dialect regions, and historical borders in Galicia. The Galician autonomous community displays characteristics that make it particularly interesting for a study of this kind. Firstly, it has some cultural, historical, social, and linguistic traits that are partly singular and partly shared with other northern areas of the Peninsula. Secondly, the historical changes in the distribution of the population over the last few centuries have been well-documented and well-studied (Bycroft et al. 2019). In addition, the administrative division of the territory in small municipalities allows us to have a finer detail picture of the family name frequency distribution. This study focuses on the analysis of surname isonymy and on the identification of surname regions. However, the sheer amount of the data analyzed, and the methodology used yield results that we deem to be more robust and reliable.

2. Materials and Method

2.1. Data Source

The study of surnames in Spanish territory over the last few decades has been used as a source of information to probe deeper into the structure of the population, both present and past. Most of the early studies used telephone directories as the source of data (Scapoli et al. 2007; Mateos & Tucker 2008), owing to the difficulty of accessing more comprehensive population databases. Fortunately, the contributions that have been published over the last few decades normally use data provided by government agencies like censuses or municipal registers (Román-Busto & Fuster 2015; Rodríguez-Díaz et al. 2015, 2017). Telephone directories are less and less reliable sources because they provide only partial pictures of the onomastic landscape, particularly of rural areas.

The onomastic information used in this investigation is more comprehensive than in preceding investigations of the same type conducted on the Iberian Peninsula as a whole. Surname data here were taken from the register of inhabitants of Galicia of the year 2011, provided by the Instituto Galego de Estatística (IGE). The database has onomastic information about 2,430,512 people with reference to each of the 315 municipalities. The total number of single surnames contained in the base is 20,754. As a geographical area of reference, most of the studies on surname distribution carried out on Spain or on Spanish autonomous communities use the province, an administrative entity that is much larger than the municipality (for example, Rodríguez-Díaz et al. 2015). The municipalities are the local level of administration, and they allow us to attain a more in-depth knowledge of the geographical distribution of surnames. The entries of the database contain detailed information about every individual: first surname, second surname, municipality of birth and year of birth. This information makes possible the geo-referencing of each single surname and to reveal the movements of diffusion of surnames throughout the territory.

2.2. Study Area, Demography, and History

The geographic area analyzed in this investigation is the administrative territory of Galicia, a Spanish autonomous community located in the northwest of the Iberian Peninsula (figure 1). It borders Portugal to the south and the regions of Asturias and Castile and León the east, while the north and west face onto the Atlantic Ocean. Today the region is divided into four administrative provinces (A Coruña, Lugo, Ourense and Pontevedra) and 313 municipalities (315 until 2016). Galicia covers a total of 29,574.4 km^2 and its population, according to the latest census of 2021, is 2,695,645 citizens.

ISSN: 0027-7738 (print) 1756-2279 (web) Vol. 73 No. 2, Summer 2025 DOI 10.5195/names.2025.2699

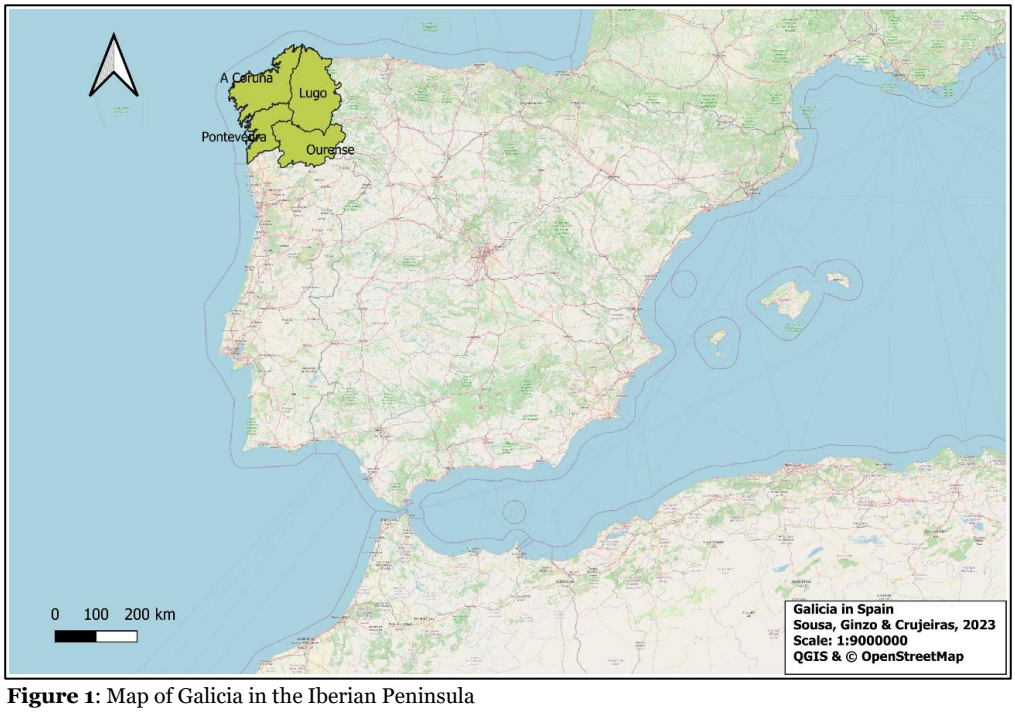

Figure 1: Map of Galicia in the Iberian Peninsula

The demographic history of Galicia is characterized by a dense, continuing human occupation, particularly in its rural landscape (Murphy et al. 2014). Until the mid-20th century, Galicia was also distinguished by the eminently rural nature of its population and the highly scattered nature of its settlements: 31,000 inhabited entities and two-thirds of the population lived in villages with a population of fewer than 1,000.[1] As is the case with a large part of the population in the Iberian Peninsula, the Galician territory is characterized by low immigration and high external migration flows (López Taboada 1996; Romero Valiente 2003; Zuniga et al. 2012). Galicia was a mainly agrarian society for centuries, and this meant that internal population movements were scarce and their significance was quite limited until the 20th century. The first noteworthy movements started only in the second half of 20th century, when a strong push towards urbanization started, a process that was particularly intense between 1960 and 1981 (Dubert García 1998; Aldrey Vázquez 2006). This exodus of people from the countryside to the city had a strong impact on Galicia's historical development. The consequences this development brought about in Galicia with regard to the birth locations of couples are similar to those brought about a couple of centuries earlier by the industrial revolutions in the population structure of several European countries (Kaplanis et al. 2018).

Galicia shares some historical, cultural, and social traits with the other northern peninsular regions. The Galician language evolved from the Latin spoken in the area, which later, as the Reconquest progressed—as was the case with the rest of the Romance varieties spoken in the north of the Peninsula—would spread southwards, giving rise to European Portuguese (Castro 2006; Venâncio 2019; Bycroft et al. 2019). Nowadays, the totality of the Galician population claim to be able to understand and speak Galician, which has the status of an official language, along with Spanish, in the autonomous community of Galicia (Dubert & Galves 2016).

DOI 10.5195/names.2025.2699 Vol. 73 No. 2, Summer 2025 ISSN: 0027-7738 (print) 1756-2279 (web)

From an onomastic perspective, Galicia shares some traits with the rest of European areas, but it also displays some singular characteristics. The surname system is based on transmission from parents to children (Boullón Agrelo 2008; Feschet 2009). Until the mid-20th century, Galicia was an eminently rural society where cities functioned quite independently from their surrounding territory. This situation affected directly on the dissemination of surnames in Galicia, as the population movements caused by urbanization brought about a parallel change in the territorial distribution of family names (Rodríguez-Díaz et al. 2015).

As in other European societies, in Galicia many surnames have a local distribution, which is determined either by their toponymic origin, or by the regional dissemination of the words from which the surnames derive (for example, occupational names, nicknames). These concentrated distribution forms have been preserved for a long time and they were only affected by recent internal migrations. Therefore, taking these change factors into account will allow us to draw a more robust onomastic regionalization in the search for connections between the distribution of surnames and other characteristics of the population.

2.3. Data Processing

2.3.1. Filtering Data

This investigation analyses the first surnames taken from the 2011 census across the 315 municipalities for the purpose of obtaining surname regions of the Galician population. The onomastic data used in this investigation were processed to obtain a more robust picture of the regional distribution of surnames and, above all, to probe into the historical structure of the Galician population. Our intervention consisted of filtering the original database in three ways:

(a) The most frequent surnames, that is, those that appear in many municipalities (more than 54), were also excluded. Among these are many patronymic surnames with the *-ez* 'son of' ending (for example, *Rodríguez*, *López*, *Fernández*, *González*, *Vázquez*, *Álvarez*), based on the given name of the father, therefore being polyphyletic surnames, most of which are ranked among the 100 most frequent.

(b) Those surnames that were recorded only in one municipality, which therefore do not have a noteworthy territorial diffusion, were excluded (some examples are *Fateira*, *Labanda*, *Painceiro*, *Rabanos, Sabariego*, *Touceiro*), along with a huge number of surnames that clearly have a foreign origin (for example, *Aabid*, *Fahde*, *Markowicz*, *O'Connor*, *Unruh*).

c) Finally, those surnames whose bearers were born after 1965 were also excluded from the database. With this reduction, we want to attenuate the consequences of the intensification of the urbanization process that Galicia underwent from the 1960s onwards. This process involved population movements from rural areas to the cities and it had a direct impact on the territorial distribution of surnames. As an example of these movements, consider the distribution of the surname *Outomuro* in three different moments of the 20th century (figure 2). This surname is derived from the place name *Outomuro* (black dot on map).

ISSN: 0027-7738 (print) 1756-2279 (web) Vol. 73 No. 2, Summer 2025 DOI 10.5195/names.2025.2699

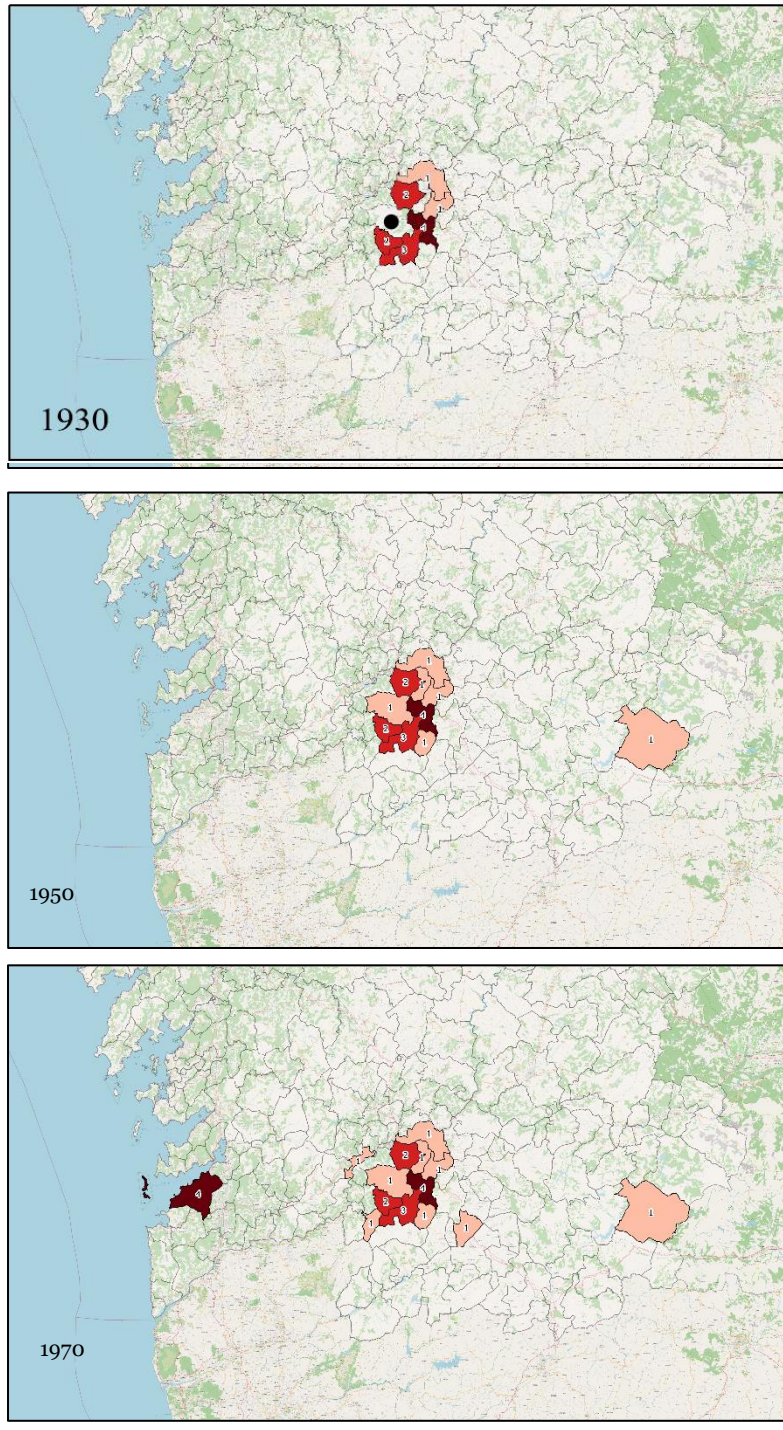

Figure 2: Diffusion of Surname *Outomuro* from 1930 to 1970

DOI 10.5195/names.2025.2699 Vol. 73 No. 2, Summer 2025 ISSN: 0027-7738 (print) 1756-2279 (web)

The top map in figure 2 shows the distribution of people bearing this surname who were born before 1931; the surname is recorded in 4 municipalities, all adjacent and located within a 25 km radius from the place from which the family name derives (a small village called *Outomuro*). The middle map shows the distribution for bearers born before 1951—the surname appears in eight municipalities, one of them being non-adjacent to the rest and located in the eastern end of Galicia (Viana do Bolo), 105 km away from the original village. The bottom map shows the distribution of surnames of bearers born before 1970—the surname *Outomuro* is documented in 12 municipalities, among them the one that includes Galicia's most populous city, Vigo. In this city, which is far from the area where the surname originated, the records show there is a high concentration of bearers of the surname *Outomuro* in this period.

Some previous works on the analysis of surname regions have already applied some of these filters in order to obtain a more robust and reliable data set (Cheshire 2014). However, it is not very common to include the time variable when selecting the data analyzed for the regionalization of surnames. In countries that have reliable historical censuses, it is easier to obtain a picture of the onomastic structure of the population at different moments of the past (Darlu et al. 2011; Kandt et al. 2020). In the case of Galicia, and in the rest of Spain, it is even more difficult to have access to historical censuses with comprehensive information, and consequently, it becomes necessary to use indirect methods like the one we applied here in order to fully exploit the historical value provided by onomastic information.

2.3.2. Isonymy and Index of Regional Surname Structure

Once the filtering process was completed, the resulting number of surnames to analyze was 6,150 (table 1). This data set was treated statistically to calculate the surname similarities between regions. Isonymy is one of the most frequently used indexes for the analysis of the distribution of surnames. Isonymy refers to the possession of the same surname: people with the same surname are more likely to share the same family lineage. In the nineteenth century, George H. Darwin (2009) suggested using surnames to determine the number of marriages between cousins. In the mid-20th century, the analysis of isonymy was extended as a procedure for the study of the structure of populations (Shaw 1960). In recent years, the filtering methods of the data analyzed have been refined (Cheshire 2014), although the most commonly used procedures remain essentially the same.

The isonymy within a region is defined as $I_i = \sum_{k \in S_i} p_{ki}^2$ where $I_i = \sum_{k \in S_i} p_{ki}^2$ denotes the relative frequency of surname k in region i, with S_i being the collection of surnames in the region. A population with few surnames has high values of isonymy, and a population with a large number of surnames has low values of isonymy. The isonymy within a region reveals the aspect of within-population structure, while the isonomy between two regions uncovers the between-population similarity.

There are also two common methods of calculating a measure of surname differentiation, and we have computed both Nei, Euclidean, and Lasker distances. Lasker distances can be interpreted as a measure of similarity or difference between two areas, namely I and j, where large distances show less similarity in surname composition. The Lasker distance is given by $L = -log(I_{ij})$, where I_{ij} denotes the isonymy between regions i and j, defined as

$$I_{ij} = \sum_{k \in S_{ij}} p_{ki} p_{kj}$$

S_{ij} being the collection of surnames present in both regions i and j. Euclidean and Nei's distances have been developed for genetic data, but they can be utilized for the frequencies of surnames (Mikerezi et al. 2013). Euclidean distance is defined by

$$E = \sqrt{1 - \sum_{k \in S_{ij}} \sqrt{p_{ki} p_{kj}}}$$

and Nei's distance is defined by

$$N = -log\left(\frac{I_{ij}}{\sqrt{I_i I_j}}\right)$$

Entropy:

$$H\_i = -\sum_{k \in S_i} \frac{p_{ki}}{\log_2(p_{ki})}$$

with i=1, ..., n.

The meaning of this observed entropy of the surname distribution is easily understood from its limiting cases. Additionally, to detect isolation by distance between locations i and j, the linear correlation of surname distances (Lasker's, Euclidean, and Nei's) with their geographic distances can be computed.

ISSN: 0027-7738 (print) 1756-2279 (web) Vol. 73 No. 2, Summer 2025 DOI 10.5195/names.2025.2699

Table 1: Surname Data and Filtering Process

	Original Data	1º Reduction	2º Reduction	3º Reduction
People	2,430,512	646,858	626,035	121,4191
Different single surnames	20,754	20,268	9,557	12,070
%	100	97.65	46.05	58.16

Data Visualization

Recently, several methods have been tested for the determination of the isonymy of onomastic data and for the graphic representation of the results. Through the analysis of the regional distribution of surnames, Cheshire, Longley, and Singleton (2010) found a strong link between the population's surnames and geographic location in Great Britain, on the basis of the calculation of Lasker distances. Cheshire and Longley (2012) turned to kernel density estimation with the aim of creating heat maps that discover those regions in Great Britain where specific surnames are more concentrated. In Italy some researchers (Boattini et al. 2010, 2012, 2021) analyzed the geographic distribution of different Italian surnames applying neural networks, which allowed them to distinguish between monophyletic and polyphyletic surnames. The surname space of the Czech Republic was studied by Novotný and Cheshire (2012), who found correlations between their network representation and ethno-cultural boundaries in that country.

In this investigation, once the measures of similarity are obtained, the final step is to create a visualization of the different surname clusters obtained through Multivariate Analysis; then clusters of surname zones can be constructed. This is normally done by representing the clusters with dendrograms formed from the matrices of Lasker's distances. The basic information of splitting or merging clusters is the similarity or isonymic distance between areas, and this distance can be obtained through different methods, such as complete linkage or Ward's procedure (Kaufman & Rousseeuw 1990). We used Ward's method, a hierarchical agglomerative clustering procedure successfully applied in linguistic variation analysis and for other purposes (Goebl 1996; Moisl 2015; Strauss & Maltitz 2017). The main goal of this method is to join clusters in which the variation within them has not increased significantly, so that the new clusters formed are more homogeneous. Ward's method is right for quantitative variables and usually creates compact and even-sized clusters.

DOI 10.5195/names.2025.2699 Vol. 73 No. 2, Summer 2025 ISSN: 0027-7738 (print) 1756-2279 (web)

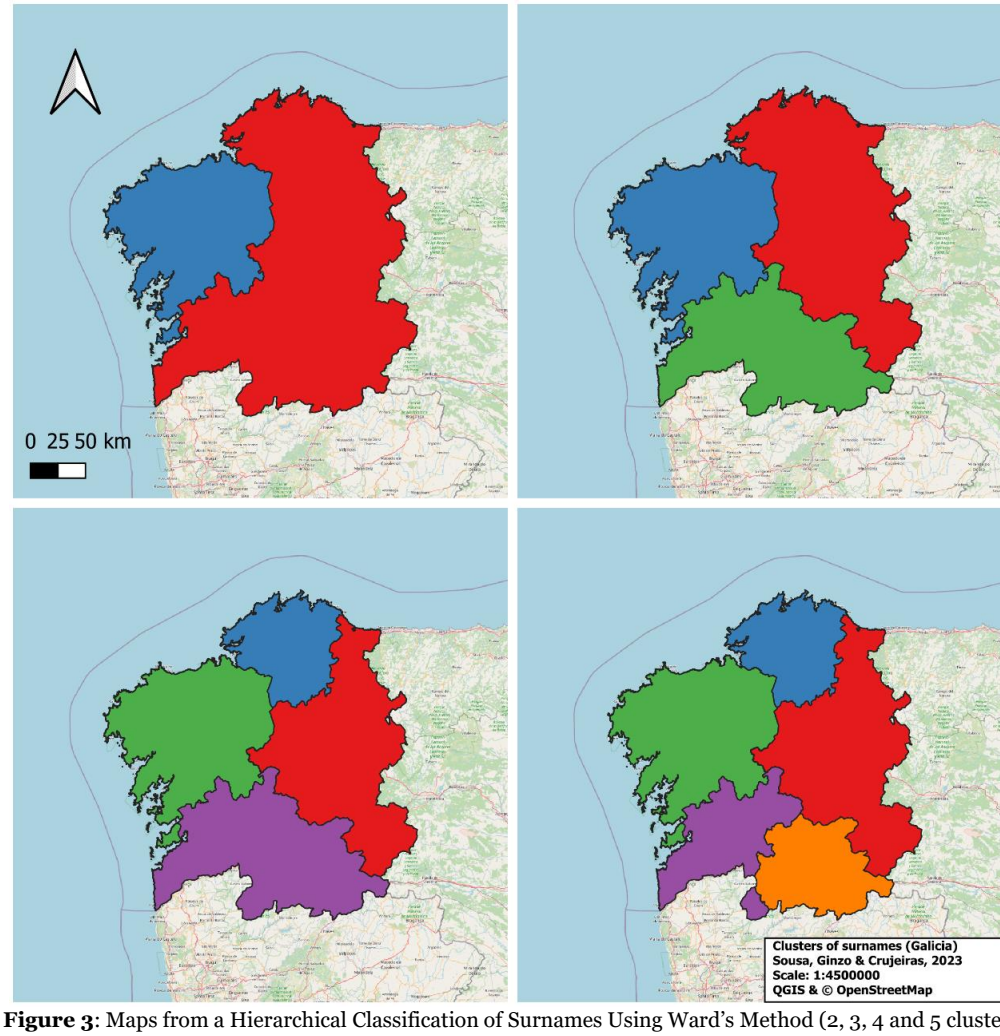

Figure 3: Maps from a Hierarchical Classification of Surnames Using Ward's Method (2, 3, 4 and 5 clusters)

Figure 3 shows the maps resulting from the analysis of the data from Galicia. The colors of the clusters correspond to the colors of the regions on the maps. Four visualizations are shown, with divisions between two and five clusters, where the regional distribution of similarities and differences according to the onomastic data can be seen.

3. Results

The works published on the analysis of surnames in Spain all point out the singular character of this country in Europe as a whole on account of its high levels of isonymy and the low diversity of surnames, according to the records. Rodríguez-Días, Manni and Blanco-Villegas (2015) suggest that the high rates of isonymy in Spain are linked to the population size at the time when the surnames were introduced and a sparse population during the expansion of the Christian kingdoms in the Middle Ages. Despite the specificity of Spanish onomastics as a whole, some differences between the different cultural and linguistic areas can be discerned. The isonymy rates are lower in the eastern areas (Catalonia, Valencia, Aragón) and higher in the western areas, particularly in Galicia, Asturias, and León. An investigation by Rodríguez Larralde et al. (2003) on the isonymy structure of Spain analyses 24 Galician towns and cities, and Galicia as a whole is

ISSN: 0027-7738 (print) 1756-2279 (web) Vol. 73 No. 2, Summer 2025 DOI 10.5195/names.2025.2699

characterized by a great degree of uniformity with regard to the abundance of surnames and by a relatively high isolation by distance.

Our investigation allows us to know in detail the internal structure of the Galician population because the quantity of data is higher and the administrative divisions that were considered, the municipalities, yield more fine-grained results about space organization. An additional data table shows the results of the isonymy analysis for all Galician municipalities on the basis of the data analyzed.[2] The high rates of isonymy in Galicia as a whole are confirmed and some differences can be discerned in inland areas. Out of the total 315 municipalities, only 11 have an α index higher than 100; the municipalities of A Coruña (α=672) and Ourense (α=413) display a higher diversity of surnames; Fisher's alpha index (α) is a measure of species diversity in a sample, used here to measure the abundance of surnames. It is also noteworthy that there is a set of municipalities with a higher F_{ST} (above 0.02), which implies that there is a higher degree of onomastic homogeneity possibly as a result of a high level of geographic isolation. These municipalities form a continuous area in the province of Ourense, particularly in the eastern half, and in the southeastern part of Lugo.

As depicted above, we cluster the isonymy matrix using Ward's clustering algorithm of municipality groups for 2011 with population weights that allow us to identify five main clusters (figure 3). When the number of clusters is below five, they are less compact and therefore less interesting when trying to find similarities with other territorial distributions. The dendrogram is a helpful way to show the whole taxonomy of areas, showing the distances at which specific areas are assigned to iteratively increasing clusters. Examining the resulting classification of areas for their geographical attributes, we can note that a strong regional grouping of areas by isonymy appears from this classification procedure. The most compact and solid cluster (it remains invariable in successive divisions) comprises a large part of the municipalities of the province of A Coruña and the north end of the province of Pontevedra (figure 3).[3] The next fragmentations occur in the groupings of municipalities comprising the north end of A Coruña, the provinces of Lugo and Ourense and the larger part of the province of Pontevedra.

Surnames and Dialects

The onomastic regions that can be observed in the division in five clusters do not match the current administrative division of Galicia in four provinces, nor do they display many similarities with the overall division of Galician language varieties in three main dialect areas. The three main dialectal varieties identified by experts for Galician language are distributed in three strips running from north to south (Fernández Rei 1990); this classification is based on an analysis of the geographic distribution of phonetic and morphosyntactic traits. However, the latest dialect studies, conducted with quantitative methods and based on a larger corpus of data, yield a territorial distribution of varieties where one can discern similarities with the maps of surname regions shown on figure 3. Sousa's work (2017a) on the lexical areas of the Galician linguistic domain identifies clusters of varieties that are grouped differently from the divisions drawn according to phonetic and morphosyntactic data (Álvarez Blanco et al. 2006; Dubert 2011). The map of lexical areas that appears on figure 4 was produced with the results of the study of the dialectometric analysis of 136 lexical variables. The procedure used to obtain the lexical clusters is the same that was used to obtain the surname regions (Sousa 2017a). The two maps on figure 4 show the similarity in the organization of the four clusters for both kinds of data—onomastic and lexical. The similarities in the distribution of the areas are particularly interesting, since lexical variation in dialectology studies is considered to follow patterns that are different from those observed in other domains of linguistic analysis—the number of variants corresponding to each variable is usually higher and there are fewer correlations in the regional distribution of variants. These similarities can be linked to studies of lexical areas in the Iberian Peninsula that show the existence of historical, demographic, cultural, and geographic factors determining the way in which lexical forms are distributed in space (Cintra 1995; Fernández-Ordóñez 2012).

Regarding the correlations between traditional dialect classifications and the regionalization of surnames, several previous studies have shown that the onomastic and linguistic regions do not necessarily have similar organizations. Manni, Heeringa, and Nerbonne (2006) analyzed onomastic and linguistic data in the Netherlands and they did not find obvious correlations between the distribution of the two types of data. However, they do suggest that they have found some similarities between the boundaries of surname regions and the boundaries separating Catholic from Protestant areas. Rodríguez-Díaz, Manni, and Blanco Villegas (2015), in their study on the surnames of 47 Spanish provinces, apart from discovering some similarities between linguistic and surname variables, suggest that there is a correlation between the geography of contemporary surnames and the political geography at the end of the Middle Ages. The spatial distribution of contemporary surnames allows them to recognize in their study the boundaries between the Christian kingdoms of the north and their expansion southwards (Bycroft et al. 2019).

DOI 10.5195/names.2025.2699 Vol. 73 No. 2, Summer 2025 ISSN: 0027-7738 (print) 1756-2279 (web)

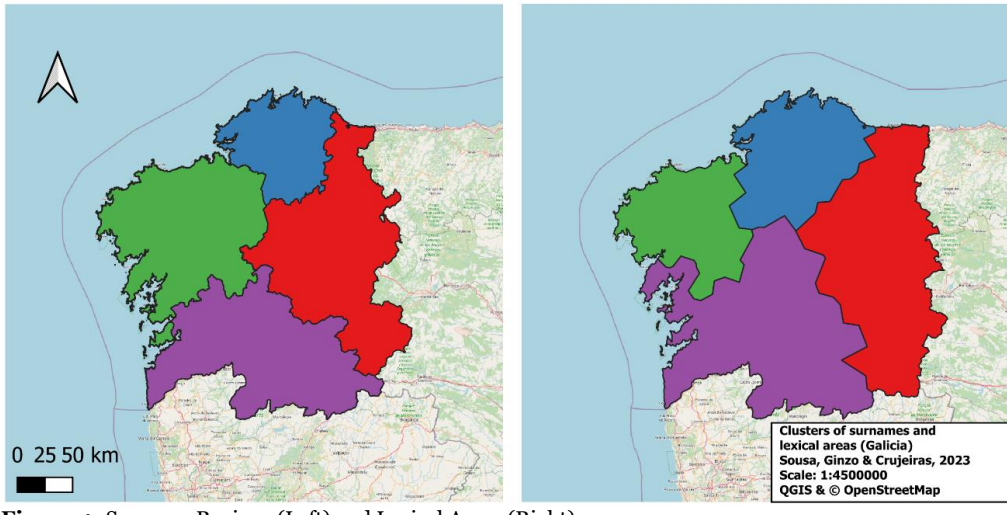

Figure 4: Surname Regions (Left) and Lexical Areas (Right).

Surnames and History

In the territorial organization of Galician surnames in five regions, some traces of ancient territorial divisions can be identified. The map on the right-hand side in figure 5 shows the ecclesiastical organization in dioceses in present-day Galicia and the similarities with the map on the left with the surname regions are easy to recognize.

Today, dioceses are ecclesiastical areas under a bishop's control. The ecclesiastical organization of Galicia is divided in six dioceses, five of them with their capitals in Galicia and one with its capital in Astorga, in the autonomous community of León and Castile. This division in dioceses was established from the 4th century onwards and is based on the administrative divisions of the Roman Empire (López Alsina 2009; Sánchez Pardo 2014). Over time, they underwent several modifications as a result of different historical developments, like the arrival of the Suebi and the Muslims, which had an impact on the consolidation of their boundaries until the Middle Ages. Among them, the area that is considered one of the oldest is the current diocese of Mondoñedo-Ferrol, stretching along the northern part of the provinces of A Coruña and Lugo. By comparing the two maps, we can see that this is the division that displays more similarities in terms of their extension. From a linguistic perspective, this area displays some features that distinguish it from the rest of Galician linguistic varieties (Santamarina 1982). Historians are at one in recognizing that dioceses must have been based on previous territorial organizations (López Alsina 2009; Sánchez Pardo 2014; Díaz Martínez 2015). Furthermore, they suggest that dioceses must have acted as agglutinative elements of peasant communities in a Galicia that has remained essentially rural, with only a handful of cities and large towns. This means that the religious organization of the territory must have had an impact on the social organization of the community.

The similarities between the population structure identified through the study of surnames and these ecclesiastical divisions must be interpreted carefully by historians and demographers. With their help, we will be able to confirm if the correlation between these two kinds of data in terms of territorial organization is linked to some determining ancient historical factors. In any event, it is obvious that the surname regions that were discovered are a consequence of the close, long-running interrelationships between the inhabitants of these areas.

ISSN: 0027-7738 (print) 1756-2279 (web) Vol. 73 No. 2, Summer 2025 DOI 10.5195/names.2025.2699

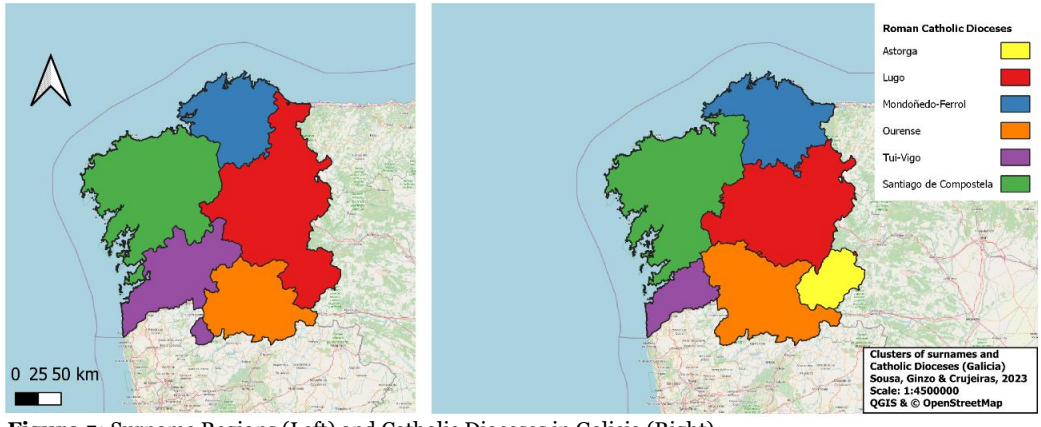

Figure 5: Surname Regions (Left) and Catholic Dioceses in Galicia (Right)

4. Concluding Remarks

Our original purpose when we started this investigation was to probe into the existence of similarities between dialect fragmentation and surname regions, following the lead of some previous investigations conducted in other European countries. In previous study on the territory of Asturias, it was discovered that the distribution of surname regions is very similar to that of Asturias' dialect varieties (Pardiñas et al. 2012; Sousa & Ginzo-Villamayor 2020).

Our investigation proves the usefulness of availing oneself of a database with comprehensive information about the entire population (for example, population census, municipal registers) and of disaggregating this information by geographic divisions. Furthermore, the method applied here allowed us to recognize the usefulness of filtering the data to help us probe into special distribution patterns in detail. Because of the high frequency with which patronymic surnames feature in the peninsular territory, along with the fact that we are dealing mainly with polygenetic surnames, it is convenient to introduce some corrections that contribute to enhancing the value of the distributional data. Apart from this correction affecting both the most frequent and the less frequent surnames, it is advisable also to employ another filter of a temporal kind, which, in this case, helps to introduce corrections with regard to the diffusion of surnames as a result of migration movements. In the case of Galicia, whose population has remained remarkably stable throughout the centuries, when compared to the populations of the rest of the European countries, this correction becomes particularly relevant and it can be easily applied owing to the fact that internal migration movements on a significant scale only took place once Galicia was well into the 20th century (Dubert García 1998).

The application of the method we put forward in this investigation yields results that are particularly useful for the linguistic and historical investigation of the peninsular populations. Galicia is characterized by the almost continuing preservation of a scattered population system and by a lower dependence of the city-countryside relation system than it was common in other peninsular areas. We are now in a position to suggest that the results found in the study of onomastic regionalization of the Galician territory will prove very useful to identify the social relations prevailing in primitive communities. The similarities between the organization of dioceses, surnames and lexical variants contribute to confirm the conclusions reached by historical studies about the relevance of ecclesiastical divisions of Galicia in the Middle Ages. On the one hand, we can surmise that the system of distribution of dioceses was superimposed on primitive systems of population divisions. On the other hand, it becomes possible to hypothesize that this organization system was an important one, because, as the onomastic and linguistic data show, it was preserved throughout centuries and had a direct impact on social relations and linguistic diversity.

Our investigation is based on a documentation of Galician surnames that allowed us to conduct an analysis in terms of both time and space. The results allow us to establish for the first time a close relationship between surnames, dialects, and history. The relevance of past patterns of social organization becomes apparent when analyzing the present-day distribution of surnames and recognizing the impact they had on the diffusion of language variants. Surnames may mirror aspects of human history and linguistic diversity both in large areas and in small territories, as is the case we studied in this investigation. Linguists can use the

DOI 10.5195/names.2025.2699 Vol. 73 No. 2, Summer 2025 ISSN: 0027-7738 (print) 1756-2279 (web)

analysis of surnames to probe into the causes of dialect variation. Historians find in surnames an essential source in order to probe into the survival of ancient social structures and study the prevailing forms of social relations throughout time. The results of the investigation confirm the need to continue researching further into the relationships between the geography of surnames and other genetic, social, historical and cultural traits of societies. Finally, this study demonstrates the utility of isonymy as a method for reconstructing the history of human populations. This method was used to study one of the Iberian Peninsula's most remote locations, which has a fascinating biological and demographic history.[4]

Acknowledgments

The authors would like to thank the Instituto Galego de Estatística (IGE, Xunta de Galicia, Spain), which provided us with the onomastic data of the population census of 2011.

Notes

[1] In 2010, 15.6% of the Galician population lived in scattered villages, a singular fact in Spain, where only 3.4% of the population live in this kind of environment (Fernández and Meixide 2013).

[2] The data that support the findings of this study are openly available in Zenodo at http://10.5281/zenodo.12774282.

[3] Sousa (2017b) offers some data that are helpful to understand the onomastic singularity of this area. It partly coincides with the area where patronymic surnames with *-ez* ('son of') endings are less frequent, and where the family names with suffixes *-eiro*/*-eira* are more frequent. This northwestern area of Galicia is usually considered to be a conservative one from a linguistic perspective, as it displays some traits that are not shared with the rest of the Galician linguistic domain.

[4] For the preparation of this work, Xulio Sousa conceived of the study and led the discussion and the writing of the first draft of the manuscript. Rosa María Crujeiras and María José Ginzo Villamayor performed the analysis and produced all the figures in the main text and extended data figures. All of the authors contributed to the method design, data analysis, and manuscript review.

Disclosure Statement

No potential conflict of interest was reported by the authors.

Funding Details

This work was funded by FEDER/Ministerio de Ciencia, Innovación y Universidades – Agencia Estatal de Investigación (PGC2018-095077-B-C44) and by the Competitive Reference Groups 2021-2024 (ED431C 2021/24) from the Xunta de Galicia. The second and third author were supported by Grant PID2020-116587GB-I00 funded by MCIN/AEI/10.13039/501100011033 and the Competitive Reference Groups 2021-2024 (ED431C 2021/24) from the Xunta de Galicia.

References

Aldrey Vázquez, José Antonio. 2006. *A poboación galega, 1900–2005: Cambio demográfico e implicacións territoriais* [The Galician Population, 1900-2005: Demographic Change and Territorial Implications]. Vigo: Ir Indo.

Álvarez Blanco, Rosario, Francisco Dubert, and Xulio Sousa, eds. 2006. *Lingua e territorio* [Language and Territory]. Santiago de Compostela: Consello da Cultura Galega.

Boattini, Alessio, Eugenio Bortolini, Roland Bauer, Marta Ottone, Rossella Miglio, Paola Gueresi, and Davide Pettener. 2021. "The surname structure of Trentino (Italy) and its relationship with dialects and genes". *Annals of Human Biology* 48, no. 3: 260–269.

Boattini, Alessio, Antonella Lisa, Ornella Fiorani, Gianna Zei, Davide Pettener, and Franz Manni. 2012. "General method to unravel ancient population structures through surnames, final validation on Italian data". *Human Biology* 84, no. 3: 235–270.

Boattini, Alessio, Maria Elena Pedrosi, Donata Luiselli, and Davide Pettener. 2010. "Dissecting a human isolate: novel sampling criteria for analysis of the genetic structure of the Val di Scalve (Italian Pre-Alps)". *Annals of Human Biology* 37, no. 4: 604–609.

Boullón Agrelo, Ana Isabel. 2008. "Historia interna del gallego: onomástica". [Internal history of Galician: onomastics] *Romanische Sprachgeschichte: Ein Internationales Handbuch Zur Geschichte Der Romanischen Sprachen* [Romance Language History: An International Handbook on the History of Romance Languages]. Edited by Gerhard Ernst. Berlin: de Gruyter, 3168–3177.

Bycroft, Clare, Ceres Fernandez-Rozadilla, Clara Ruiz-Ponte, Inés Quintela, Ángel Carracedo, Peter Donnelly, and Simon Myers. 2019. "Patterns of genetic differentiation and the footprints of historical migrations in the Iberian Peninsula". *Nature Communications* 10, no. 1: 551. https://doi.org/10.1038/s41467-018-08272-w

Calderón, Rosario, Candela L. Hernández, Pedro Cuesta, and Jean Michel Dugoujon. 2015. "Surnames and Y-chromosomal markers reveal low relationships in Southern Spain". *PloS One* 10, no. 4. https://doi.org/10.1371/journal.pone.0123098

Caramelli, David, Cosimo Posth, and Olga Rickards. 2021. "Reconstruction of the human peopling of Europe: A genetic insight". *Annals of Human Biology* 48, no. 3: 175–178.

Castro, Ivo. 2006. *Introdução à história do português*. [An introduction to the history of Portuguese] 2ª. Lisboa: Colibri.

Cheshire, James A. 2014. "Analysing surnames as geographic data". *Journal of Anthropological Sciences* 92: 99–117.

Cheshire, James A., and Paul A. Longley. 2012. "Identifying spatial concentrations of surnames". *International Journal of Geographical Information Science* 26, no. 2: 309–325.

Cheshire, James A., Paul A. Longley, and Alex D. Singleton. 2010. "The surname regions of Great Britain". *Journal of Maps* 6, no. 1: 401–409.

Cheshire, James A., Pablo Mateos, and Paul A. Longley. 2011. "Delineating Europe's cultural regions: population structure and surname clustering". *Human Biology* 83, no. 5: 573–598.

Cintra, Luís Filipe Lindley. 1995. *Estudos de dialectologia portuguesa*. [Studies in Portuguese dialectology] Lisboa: Sá Da Costa.

Darlu, Pierre, Guy Brunet, and Dominique Barbero. 2011. "Spatial and temporal analyses of surname distributions to estimate mobility and changes in historical demography: the example of Savoy (France) from the eighteenth to the twentieth century". *Navigating time and space in population studies*. Edited by Emily R. Merchant, Glenn D. Deane, Myron P. Gutmann and Kenneth M. Sylvester. Dordrecht: Springer, 199–113.

Darwin, George H. 2009. "Marriages between first cousins in England and their effects". *International Journal of Epidemiology* 38, no. 6: 1429–1439.

DOI 10.5195/names.2025.2699 Vol. 73 No. 2, Summer 2025 ISSN: 0027-7738 (print) 1756-2279 (web)

Díaz Martínez, Pablo de la Cruz. 2015. "La organización del espacio y el control del territorio en la Galicia germánica" [The Organisation of Space and the Control of Territory in Germanic Galicia]. *El territorio en la historia de Galicia: organización y control, siglos I-XXI*. [Territory in the history of Galicia: organisation and control, 1st–21st centuries]. Edited by Antonio C. Pereira Menaut and Ermelindo Portela Silva. Santiago de Compostela: Universidade de Santiago de Compostela, 37–95.

Dubert, Francisco. 2011. "Developing a database for dialectometric studies: the ALGa phonetic data. Dialectometrical analysis of 230 working maps". *Dialectologia et Geolinguistica* 19, no. 1: 23–61.

Dubert, Francisco, and Charlotte Galves. 2016. "Galician and Portuguese". *The Oxford Guide to the Romance Languages*. Edited by Adam Ledgeway and Martin Maiden. Oxford: Oxford University Press, 411–446.

Dubert García, Isidro. 1998. "Mundo urbano y migraciones campo-ciudad en Galicia, Siglos XVI-XIX" [Urban world and rural-urban migrations in Galicia, 16th-19th Centuries]. *Revista de Demografía Histórica-Journal of Iberoamerican Population Studies* 16, no. 2: 39–86.

Faure, Roberto, María Asunción Ribes, and Antonio García. 2009. *Diccionario de apellidos españoles*. [Dictionary of Spanish surnames] 5ª ed. Madrid: Espasa Calpe.

Fernández Rei, Francisco. 1990. *Dialectoloxía da lingua galega*. [Dialectology of the Galician language] Vigo: Xerais.

Fernández-Ordóñez, Inés. 2012. "El norte peninsular y su papel en la historia de la lengua española" [The northern peninsular and its role in the history of the Spanish language]. *Estudios sobre tiempo y espacio en el español norteño*. [Studies of Time and Space in Northern Spanish]. Edited by Sara Gómez Seibane and Carsten Sinner. San Millán de la Cogolla: Cilengua, 23–68.

Ferreiro, David, Bernabé Núñez-Estévez, Mateo Canedo, Catarina Branco, and Miguel Arenas. 2021. "Evaluating causes of current genetic gradients of modern humans of the Iberian Peninsula". *Genome Biology and Evolution* 13 no. 4. https://doi.org/10.1093/gbe/evab071

Feschet, Valérie. 2009. "The surname in Western Europe liberty, equality and paternity in legal systems in the twenty-first century". *L'Homme* 20, no. 1: 63–73. https://shs.hal.science/halshs-00927696/file/The_Surname_in_Western_Europe_L_Homme_ZFG_PDF_-_copie.pdf

Goebl, Hans. 1996. "La convergence entre les fragmentations géo-linguistique et géo-génétique de l'Italie du nord" [The Convergence of Geo-Linguistic and Geo-Genetic Fragmentation in Northern Italy]. *Revue de Linguistique Romane* 60: 25–49.

Guppy, H. 1890. *Homes of Family Names in Britain*. London: Harrison and Sons.

Hanks, Patrick, Richard Coates, and Peter McClure. 2016. *The Oxford Dictionary of Family Names in Britain and Ireland*. New York: Oxford University Press.

Jobling, Mark A. 2012. "The impact of recent events on human genetic diversity". *Philosophical transactions of the Royal Society of London. Series B, Biological sciences* 367, no. 1590: 793–799.

Kamel, Célia, Bérengère Saliba-Serre, Marie-Hélène Lizee, Michel Signoli and Caroline Costedoat. 2021. "Surnames in South-Eastern France: structure of the rural population during the 19th century through isonymy". *Journal of Biosocial Science*, 1–16.

Kandt, Jens, James A. Cheshire, and Paul A. Longley. 2016. "Regional surnames and genetic structure in Great Britain". *Transactions* 41, no. 4: 554–569.

Kandt, Jens, Justin van Dijk, and Paul A. Longley. 2020. "Family name origins and intergenerational demographic change in Great Britain". *Annals of the American Association of Geographers* 110, no. 6: 1726–1742.

Kaplanis, Joanna, Assaf Gordon, Tal Shor, Omer Weissbrod, Dan Geiger, Mary Wahl, Michael Gershovits et al. 2018. "Quantitative analysis of population-scale family trees with millions of relatives". *Science* 360, no. 6385: 171–175.

Kaufman, Leonard, Peter J. Rousseeuw. *Finding Groups in Data: An Introduction to Cluster Analysis*. New York: Wiley.

Kremer, Dieter. 2003. "Spanish and Portuguese family names". *Dictionary of American Family Names*. Edited by Patrick Hanks. New York: Oxford University Press, XV–XIX.

Lasker, G. W. 1980. "Surnames in the study of human biology". *American Anthropologist* 82, no. 3: 525–538.

López Alsina, Fernando. 2009. "Da protoparroquia ou parroquia antiga altomedieval á parroquia clásica en Galicia" [From the proto-parish or ancient early medieval parish to the classical parish in Galicia]. *A parroquia en Galicia* [The Parish in Galicia]. Edited by Fernando García Pazos. Santiago de Compostela: Xunta de Galicia, 57–75.

ISSN: 0027-7738 (print) 1756-2279 (web) Vol. 73 No. 2, Summer 2025 DOI 10.5195/names.2025.2699

López Taboada, Xosé Antonio. 1996. *La población de Galicia: 1861-1991: Crecimiento, movimiento natural y estructura de la población de Galicia* [The Population of Galicia: 1861-1991: Growth, Natural Movement and Structure of the Population of Galicia]. A Coruña: Fundación Caixa Galicia.

Lopez-Gay, Antonio. 2014. "Population growth and re-urbanization in Spanish inner cities: the role of internal migration and residential mobility". *Revue Quetelet* 2, no. 1: 67–92.

Manni, Franz and Italo Barrai. 2001. "Genetic structures and linguistic boundaries in Italy: a microregional approach". *Human Biology* 73, no. 3: 335–347.

Manni, Franz, Wilbert Heeringa, and John Nerbonne. 2006. "To what extent are surnames words? Comparing geographic patterns of surname and dialect variation in the Netherlands". *Literary and Linguistic Computing* 21, no. 4: 507–527.

Manni, Franz, Wilbert Heeringa, Bruno Toupance, and John Nerbonne. 2008. "Do surname differences mirror dialect variation". *Human Biology* 80, no. 1: 41–64.

Martínez-Cadenas, Conrado, Alejandro Blanco-Verea, Barbara Hernando, George B. J. Busby, Maria Brion, Angel Carracedo, Antonio Salas and Cristian Capelli. 2016. "The relationship between surname frequency and Y chromosome variation in Spain". *European Journal of Human Genetics* 24, no. 1: 120–128. https://doi.org/10.1038/ejhg.2015.75

Mateos, Pablo, Paul A. Longley, and David O'Sullivan. 2011. "Ethnicity and population structure in personal naming networks". *PloS One* 6, no. 9: e22943.

Mateos, Pablo, and Ken Tucker. 2008. "Forenames and surnames in Spain in 2004". *Names* 56, no. 3: 165–184.

Mikerezi, Ilia, Endrit Xhina, Chiara Scapoli, Guido Barbujani, Elisabetta Mamolini, Massimo Sandri, Alberto Carrieri, Álvaro Rodriguez-Larralde, and Italo Barrai. 2013. "Surnames in Albania: a study of the population of Albania through isonymy". *Annals of Human Genetics* 77, no. 3: 232–243.

Mir de la Cruz, Rafael. 1981. "Estadística y antroponimia (V)" [Statistics and anthroponymy]. *Estadística Española* 91: 63–73.

Moisl, Hermann. 2015. *Cluster Analysis for Corpus Linguistics*. Berlin: de Gruyter.

Morelli, Laura, Giorgio Paoli, and Paolo Francalacci. 2002. "Surname analysis of the Corsican population reveals an agreement with geographical and linguistic structure". *Journal of Biosocial Science* 34, no. 3: 289–301.

Moreu-Rey, Enric. 1993. *Antroponímia: història dels nostres prenoms, cognoms i renoms* [Anthroponymy: History of our names, surnames and nicknames]. Barcelona: Universitat de Barcelona.

Novotný, Josef, and James A. Cheshire. 2012. "The surname space of the Czech Republic: examining population structure by network analysis of spatial co-occurrence of surnames". *PloS One* 7, no. 10. https://doi.org/10.1371/journal.pone.0048568

Pardiñas, Antonio F., Agustín Roca, Eva García-Vazquez, and Belén López. 2012. "Assessing the genetic influence of ancient sociopolitical structure: micro-differentiation patterns in the population of Asturias (Northern Spain)". *PloS One* 7, no. 11: e50206. https://doi.org/10.1371/journal.pone.0050206

Piazza, A., S. Rendine, G. Zei, A. Moroni, and L. L. Cavalli-Sforza. 1987. "Migration rates of human populations from surname distributions". *Nature* 329, no. 6141: 714–716.

Rodríguez Díaz, Roberto. 2015. *La población española. Aproximación a la estructura biológica, genética y poblacional a partir de los apellidos* [The Spanish Population. An Approach to the Biological, Genetic and Population Structure Based on Surnames]. Salamanca: Universidad de Salamanca.

Rodríguez-Díaz, Roberto, María José Blanco-Villegas and Franz Manni. 2017. "From surnames to linguistic and genetic diversity: five centuries of internal migrations in Spain". *Journal of Anthropological Sciences* 95: 249–267.

Rodríguez-Díaz, Roberto, Franz Manni, and María José Blanco-Villegas. 2015. "Footprints of Middle Ages kingdoms are still visible in the contemporary surname structure of Spain". *PloS One* 10, no. 4: e0121472. https://doi.org/10.1371/journal.pone.0121472

Rodríguez-Larralde, Álvaro, Antonio González-Martin, Chiara Scapoli, and Italo Barrai. 2003. "The names of Spain: a study of the isonymy structure of Spain". *American Journal of Physical Anthropology* 121, no. 3: 280–292.

Román-Busto, Jorge, and Vicente Fuster. 2015. "Secular trends in the relationship between surnames in a population: study of a border town". *Homo: internationale Zeitschrift fur die vergleichende Forschung am Menschen* 66, no. 2: 158–164.

DOI 10.5195/names.2025.2699 Vol. 73 No. 2, Summer 2025 ISSN: 0027-7738 (print) 1756-2279 (web)

Romero Valiente, J. M. 2003. “Migraciones” [Migrations]. *Tendencias demográficas durante el siglo XX en España* [Demographic trends during the 20th century in Spain]. Edited by Andrés Arroyo Pérez. Madrid: Instituto Nacional de Estadística, 209–253.

Salinero, Gregorio and Isabel Testón Núñez. 2010. *Un juego de engaños: movilidad, nombres y apellidos en los siglos XV a XVIII*. [A Game of Deception: Mobility, Names, and Surnames in the 15th-18th Centuries] Madrid: Casa de Velázquez.

Sánchez Pardo, José Carlos. 2014. “Organización eclesiástica y social en la Galicia Tardoantigua. Una perspectiva geográfico-arqueológica del parroquial suevo” [Ecclesiastical and Social Organisation in Late Antique Galicia. A Geographic-Archaeological Perspective on the Swabian Parish]. *Hispania Sacra* 66 no. 134: 439–480.

Santamarina, Antón. 1982. “Dialectoloxía galega: historia e resultados” [Galician dialectology: history and results] *Tradición, actualidade e futuro do galego*. [Tradition, Present and Future of the Galician Language]. Edited by Ramón Lorenzo and Dieter Kremer. Santiago de Compostela: Xunta de Galicia, 153–187.

Scapoli, Chiara, H. Goebl, S. Sobota, E. Mamolini, Álvaro Rodriguez-Larralde, and Italo Barrai. 2005. “Surnames and dialects in France: population structure and cultural evolution”. *Journal of Theoretical Biology* 237, no. 1: 75–86.

Scapoli, Chiara, Elisabetta Mamolini, Alberto Carrieri, Álvaro Rodriguez-Larralde, and Italo Barrai. 2007. “Surnames in western Europe: a comparison of the subcontinental populations through isonymy”. *Theoretical Population Biology* 71, no. 1: 37–48.

Scott, James C., John Tehranian, and Jeremy Mathias. 2004. “Government surnames and legal identities”. *National Identification Systems: Essays in Opposition*. Edited by Carl Watner and Wendy McElroy. Jefferson: McFarland, 34–38.

Shaw, Richard F. 1960. “An index of consanguinity based in the use of the surname in Spanish speaking countries”. *Journal of Heredity* 51, no. 5: 221–230.

Solé-Morata, Neus, Jaume Bertranpetit, David Comas, and Francesc Calafell. 2015. “Y-chromosome diversity in Catalan surname samples: insights into surname origin and frequency”. *European Journal of Human Genetics* 23, no. 11: 1549–1557.

Sousa, Xulio. 2017a. “Aggregate analysis of lexical variation in Galician”. *Language Variation - European Perspectives VI*. Edited by Isabelle Buchstaller and Beat Siebenhaar. Amsterdam/Philadelphia: Benjamins, 71–84.

Sousa, Xulio. 2017b. “Alcune riflessioni sulla geonomastica personale” [Some reflections on personal geonomastics]. *Di nomi e di parole. Studi in onore di Alda Rossebastiano*. Edited by Elena Papa and Daniela Cacia. Roma: Società Editrice Romana, 387–400.

Sousa, Xulio, and María José Ginzo-Villamayor. 2020. “Surname regions and dialectal variation in the Asturian linguistic space”. *Journal of Linguistic Geography* 8, no. 2: 102–114.

Strauss, Trudie, and Michael Johan von Maltitz. 2017. “Generalising Ward's method for use with Manhattan distances”. *PloS One* 12, no. 1: e0168288.

Venâncio, Fernando. 2019. *Assim nasceu uma lingua: Sobre as origens do Português*. [This is how a language was born: On the origins of Portuguese] Lisboa: Guerra & Paz.

Winney, Bruce, Abdelhamid Boumertit, Tammy Day, Dan Davison, Chikodi Echeta, Irina Evseeva, Katarzyna Hutnik et al. 2012. “People of the British Isles: preliminary analysis of genotypes and surnames in a UK-control population”. *European Journal of Human Genetics* 20, no. 2: 203–210.

Zei, Gianna, Antonella Lisa, Ornella Fiorani, Chiara Magri, Lluis Quintana-Murci, Ornella Semino and A. Silvana Santachiara-Benerecetti. 2003. “From surnames to the history of Y chromosomes: the Sardinian population as a paradigm”. *European Journal of Human Genetics* 11, no. 10: 802–807.

Zuniga, Maria, Angel Pueyo, and Jose-Luis Calvo. 2012. “The Spanish population during the twentieth century and beyond”. *Journal of Maps* 8, no. 4: 386–391.

ISSN: 0027-7738 (print) 1756-2279 (web) Vol. 73 No. 2, Summer 2025 DOI 10.5195/names.2025.2699