ZEITSCHRIFT FÜR DIALEKTOLOGIE UND LINGUISTIK 87, 2020/2, 285–306
DOI 10.25162/ZDL-2020-0010

XULIO SOUSA / FRANCISCO DUBERT GARCÍA

Measuring Language Contact in Geographical Space: Spanish Loanwords in Galician*

Sprachkontakt im Geographischen Raum messen: Spanische Lehnwörter im Galicischen

ABSTRACT: The quantitative analysis of linguistic data has been used in variational linguistics to reveal relationships between varieties and distribution patterns of linguistic variants that have often been hidden from traditional methodologies. This research approach helps to understand the spatial organization of varieties in a more comprehensive way, as well as the similarities and differences between them, regardless of their classification as languages or regional varieties. Nevertheless, neither modern methodology nor traditional dialectology has yet given much attention to the analysis of the lexical transfer that occurs between varieties that are in close geographical contact, be they varieties of the same or two different languages.
The purpose of this article is to show how dialectometric techniques can be used to analyse the contact between linguistic varieties, as well as to identify the distributional patterns of loan words. The data analysed are taken from a Galician linguistic geography project – a Romance variety spoken in north-western Spain – carried out in the 1970s ("Atlas Lingüístico Galego" – "Galician Language Atlas"). The lexical variables studied contain Galician and Spanish variants. The dialectometric methods used make it possible to identify geographical distribution patterns for the Spanish variants, to identify areas that are more resistant to the inclusion of loan words, and to evaluate the influence that extralinguistic factors can have on the distribution of loan words. Finally, the paper shows the usefulness of quantitative methods to provide a more comprehensive description of contact-induced language change.
Keywords: language contact, lexical borrowing, dialectometry, geolinguistics, Galician, Spanish

KURZFASSUNG: Die quantitative Analyse linguistischer Daten wurde in der Varietätenlinguistik verwendet, um Beziehungen zwischen Varietäten und Verbreitungsmustern sprachlicher Varianten zu offenbaren, die den traditionellen Methodiken häufig verborgen blieben. Dieser Forschungsansatz hilft dabei, die räumliche Organisation der Varietäten auf umfassendere Weise zu verstehen, ebenso wie die zwischen ihnen bestehenden Gemeinsamkeiten und Unterschiede, und zwar unabhängig von ihrer Klassifikation als Sprachen oder regionale Varietäten. Trotzdem hat sich bisher weder die moderne Methodik noch die traditionelle Dialektologie eingehend mit der Analyse des lexikalischen Transfers befasst, der sich zwischen Varietäten

* This work was funded in part with the help of the Ministerio de Economía y Competitividad (project FFI2015–65208-P), the Ministerio de Ciencia, Innovación y Universidades of the Spanish Government (project PGC2018–095077-B-C44), the Galician Government and the European Union (GRC2013/40).

ergibt, die in engem geographischem Kontakt stehen, seien dies nun Varietäten derselben Sprache oder zweier verschiedener Sprachen.
Dieser Artikel möchte zeigen, auf welche Art dialektometrische Verfahren auch für die Analyse des Kontakts zwischen linguistischen Varietäten verwendet werden können, ebenso wie für die Erkennung der Verbreitungsmuster von Lehnwörtern. Die analysierten Daten sind einem Projekt der Sprachgeografie des Galicischen – einer romanischen Varietät, die im Nordwesten Spaniens gesprochen wird – entnommen, das in den 70er Jahren des vergangenen Jahrhunderts verwirklicht wurde („Atlas Lingüístico Galego" – „Galicischer Sprachatlas"). Die untersuchten lexikalischen Variablen enthalten galicische und spanische Varianten. Die angewandten dialektometrischen Verfahren ermöglichen es, geografische Verbreitungsmuster für die spanischen Varianten zu erkennen, Gebiete zu identifizieren, die der Eingliederung von Lehnwörtern eher widerstehen, und auch den Einfluss zu bewerten, den extralinguistische Faktoren auf die Verbreitung von Lehnwörtern haben können. Schließlich zeigt die Arbeit den Nutzen quantitativer Methoden, um eine umfassendere Beschreibung von kontaktbedingtem Sprachwandel zu geben.
Schlagworte: Sprachkontakt, Lehnwörter, Dialektometrie, Geolinguistik, Galicisch, Spanisch

1. Introduction

In Romance linguistics there exists a tradition of contrastively studying the spatial continuity of different cognates along the Romance languages in order to explain the split of Vulgar Latin into the different Neo-Latin languages (LÜDTKE 1968, HERMAN 1990, CONTINI 1991). ROHLFS (1960) studied the geolinguistic distribution of the Romance denominations for 51 concepts in the Iberian Peninsula. He showed that the old forms of present-day French words like *manger* 'to eat', *arriver* 'to arrive', *guérir* 'to cure/treat', *raisin* 'grape', *prier* 'to beg/pray', épaule 'shoulder', *oublier* 'to forget', *trouver* 'to find', *fromage* 'cheese' spread from their original Galo-Romance territory to Catalonia and Italy. They superseded the corresponding old Romance forms directly inherited from Latin in those territories. ROHLFS attributed this spreading to the moment when "Italy is included in the scope of the Frankish Empire. From then on, no longer was Italy that decided the linguistic revolutions, but rather, new currents full of vigour and born in France acted upon neighbouring Romance language countries." (ROHLFS 1960: 63, our translation).

After analysing the geographic distribution of the denomination of those concepts, ROHLFS concludes by dividing the Romance speaking territory in four parts: Iberia, Galia, Italia and Dacia. The space called Iberia includes Spanish (Sp.) and Portuguese (Pt.) (and excludes Catalan), since these languages share some specific cognate forms (Sp. *manzana* ~ Pt. *maçã* 'apple', Sp. *pierna* ~ Pt. *perna* 'leg', Sp. *quemar* ~ Pt. *queimar* 'to burn' etc.). The Ibero-Romance varieties (particularly, Castilian, Asturian-Leonese, Galician and Portuguese) share some important linguistic features that make them a group and separate them from the other Romance languages. Some of those forms are the relative/interrogative pronouns Sp. *quien* ~ Pt. *quem* 'who', the indefinite Sp. *alguien* ~ Pt. *alguém*

'somebody', three verbal conjugations (*amar* 'to love', *temer* 'to fear', *partir* 'leave', a series of demonstrative pronouns and deictic adverbs opposing three degrees of distance (Pt. *este* 'this', *ese* 'that', *aquel* 'he/it over there'; Pt. *aquí* 'here', *aí/ahí* 'there', *alí/allí* 'over there'). These Iberian languages have also lost the clitic pronouns descending from Lat. *inde* and *ibi* (preserved in Aragonese, Catalan, French, and Italian). Thus, these facts suggest an intimate relation between Spanish, Asturian-Leonese, Galician, and Portuguese (see ENTWISTLE 1973, DUBERT GARCÍA 2017).

Following CINTRA's study as regards Portuguese (1983), the aim of the current analysis is to delve into the geographical distribution of the denominations of eight Galician and Portuguese concepts mostly linked to domains of traditional life such as *soro* 'whey', *año* 'lamb' or *ubre* 'udder' (DUBERT GARCÍA/SOUSA 2002, cf. SOUSA 2017a). This analysis aims to show the historical influence of the dialects of the centre of the Peninsula (Asturian-Leonese and Spanish), in both Galician (Gal.) and Portuguese (Pt.): the central denomination *ordeñar* 'to milk' has penetrated in south-eastern Portuguese; and the central denomination *machorra* 'sterile (cow)' has penetrated Galician. Similarly, FERNÁNDEZ-ORDÓÑEZ (2011), studying the formation of the Spanish lexicon throughout 27 lexical and 5 syntactic items taken from the "Atlas Lingüístico de la Península Ibérica", discovered the exchanging of denominations and syntactic constructions between Spanish and the different Iberian Romance languages.

The studies of lexical loans between languages, however, tend to be associated with historic, geographic, sociolinguistic visions of language contact. Thus, from the comparison of the geographical distribution of the different denominations of some concepts across different languages, they show that the historical exchanging of lexical items is a common feature of neighbouring languages. In fact, neighbouring languages may exchange words, but it is well known that they may also exchange grammatical constructions or phonological segments. Hitherto, language contact situations have been studied with a main focus on phonetics and grammar, and only secondarily on vocabulary, despite frequent statements to the effect that in language contact situations vocabulary is always the level of language that is affected first, and most extensively (HOUT/MUYSKEN 1994, HASPELMATH/TADMOR 2009, TADMOR/HASPELMATH/TAYLOR 2010, THUN 2010, WINFORD 2010, FRANCO et al. 2019).

The aim of this paper is to show how quantitative geolinguistic methods help to identify patterns in loanwords' spatial distribution, and whether the same methods can be used to identify the factors that lead to the spread of lexical items across language varieties in two linguistic domains. Survey data used come from a Galician language atlas from the late twentieth century. As for methodology, the current study applies the dialectometric method, a quantitative analysis of geolinguistic data that helps to find geographic patterns hidden behind a long series of dialectal maps.

Traditionally geolinguistics has been concerned with lexical transfers between varieties associated with a single language, with special attention to the regional and diachronic spread of specific forms (BRITAIN 2010, 2018). Less often, the discipline examines lexical transfers between varieties attributed to different languages and the spread of new forms over a linguistic area. The study of lexical loans between languages, how-

ever, tends to be associated with sociolinguistics and analyses of language contact. The current research combines the study of lexical borrowing between two languages with an analysis of the geolinguistic diffusion of lexical items in order to determine whether there are geographical patterns in the distribution of loanwords.

In the current paper the term "loanword" will be used following HASPELMATH (2009: 36): "a word that at some point in the history of a language entered its lexicon as a result of borrowing (or *transfer*, or *copying*)". As such, loanwords are opposed to native words. In borrowing situations, the language from which the loanword is taken (in this case, Spanish) is the donor language and the language that adopts the loanword (in this case, Galician) is the recipient language. Some loanwords enter the language together with new elements of the world (Sp. *tenedor* 'fork'; standard Galician *garfo*); others replace Galician forms (Sp. *rodilla* 'knee', Gal. *xeonllo*). Some loanwords have undergone phonetic or morphological adaptation and have even entered the standard variety (for example Gal. *xamón*, from Sp. *jamón* 'ham'; *pantalóns*, from Sp. *pantalones* 'trousers'); others have been taken on without such adaptation (*gemelo* 'twin'; standard Galician *xemelgo*).

2. Galician and Spanish in contact

Galician is a Romance language spoken in the North-West of Spain (Galicia, western Asturias and Leon and north-western Zamora), and historically related to Portuguese. Throughout all its history, Galician has been in constant contact with Spanish and has progressively been diverging from Portuguese. Certainly, some differences between Galician and Portuguese have their roots in individual internal changes of each one of the languages (for example, DUBERT GARCÍA/GALVES 2016). Other differences can be attributed to a special relation between Galician and Asturian-Leonese (DUBERT GARCÍA 2010, 2017). However and finally, other differences were produced by the influence of Spanish on Galician. Despite the early use of Galician in literature and in official settings, Spanish has played an increasingly central role as the high language (as defined by FERGUSON 1959) in Galicia since the thirteenth century. This process reached a point where Spanish was the only language used in schools, religion, the courts and writing, while Galician became the low language (MONTEAGUDO 1999, MARIÑO 1998); in the terminology used by MULJAČIĆ (1991), Galician was (and still is) a roofed language, and Spanish was (and still is) its roof language.

This situation was only partially reversed in the final decades of the 20th century, when Galician gained official protection, acquired co-official status and developed a standard variety that has come to be used in mass media and taught at schools and universities (RAMALLO/REI-DOVAL 2015).

On the other hand, Galician has been in direct geographic contact with Asturian-Leonese, another Romance language developed and placed between Galician and Spanish (FERNÁNDEZ-ORDÓÑEZ 2009). Since Asturian-Leonese was also roofed by Spanish and is a neighbouring language, probably Galician and Asturian-Leonese were in a situ-

ation ranging from mere lexical borrowing to moderate structural diffusion, according to WINFORD's classification (2010).

As a consequence of this contact with Spanish, Galician has converged with the high language to varying degrees (MONTEAGUDO 1999, DUBERT GARCÍA 2005). Hitherto there have not been many empirical historical studies providing data as to the consequences of this unceasing contact. However, because of the adoption into Galician of hundreds of Spanish loanwords in the course of its history (GARCÍA 1976) it is commonly claimed that the lexicon and, more specifically, the content words are the language component most affected by contact. This impressionistic claim is likely to be accurate and coincides with the results of typological analyses concerned with patterns of borrowing (HASPELMATH 2008, TADMOR 2009). The Galician-Spanish contact is a very old one. However, it should be noted that not only Galician, but also Portuguese, has been importing words from Spanish. Portuguese has taken hundreds of words from Spanish, for instance *castelhano* 'Castilian', *cavalheiro* 'knight', *gatilho* 'trigger', *ensimesmar* 'to be rapt', *despojar* 'strip'; VENÂNCIO 2013). While in the Portuguese cultured elites there seems to be no problem with these loanwords, the penetration of Spanish words in Galician, much more intense than in Portuguese, has been perceived as a problem by the cultured elites (according to WINFORD 2010; MONTEAGUDO 2017).

Besides this, the degree of contact of Galician with Spanish seems to be an increasing phenomenon that even affects inflectional morphology (DUBERT GARCÍA/ACUÑA-FARIÑA 2018).

The reasons for the growing Spanish influence on Galician are related to the extension and improvement of compulsory education, the strengthening of literacy, the power of contemporary mass media (radio, television and cinema), etc. Until the beginning of the 1980s all these activities had been carried out exclusively in Spanish.

3. Data and method

The data for the present study comes from the "Atlas Lingüístico Galego (ALGA 1990–)", based on a survey carried out between 1974 and 1976. This linguistic geography project was developed along the lines of traditional dialectology. The places surveyed were villages in rural Galicia and adjoining Galician speaking parts of Asturias, León and Zamora. The subjects interviewed were native speakers of Galician and archetypal NORM informants (Non-mobile Old Rural Males). The questionnaire consisted of 2 711 questions, more than 80 % of which focused on lexical issues. The main method used was indirect questioning. It is important to bear in mind that informants were asked for the Galician word for things. It was assumed therefore that the replies represent words fully integrated into the informant's Galician. This data gathering procedure ensures that none of the words given arose from code switching (HASPELMATH 2009).

In the period when the survey was carried out, Galician still lacked official recognition as the language of Galicia (which was only obtained in 1981). Furthermore, Galician had no standard variety (the standard dates from 1982), was not taught in schools

(teaching of Galician in schools began in 1979), and was still excluded from the administrative, educational and official domains, even though it was the habitual language of most of the population. The picture that emerges represents the result of the coexistence of Galician and Spanish without any influence from a Galician standardisation process (MARIÑO 1998, MONTEAGUDO 1999, RAMALLO/REI-DOVAL 2015).

The material discussed in the present study comes from the project's database (SOUSA 2017b). In general terms, it was selected in accordance with the following criteria, in order to have a coherent set of working maps: a) content words were taken into account, not function words;[1] b) the loans included in our database have replaced Galician words, so that they are not names for novel concepts, but "words that more or less duplicate already existing words" (MYERS-SCOTTON 2002); c) the words considered have not undergone a phonetic adaptation; d) these loanwords are not yet found in current Galician dictionaries; e) all the items chosen present a Spanish word form that is used to some extent in Galician: working maps representing questions for which none of the answers contain at least one Spanish loanword were excluded. 68 questionnaire items that meet these criteria were selected for the study. This places some limits on the study, which thus focuses on a significant number of vocabulary items pertaining to a specific period in the language's recent history, and on speakers with a certain kind of socioeconomic profile, place of origin, age and so on.

The selected words were included in a database for treatment using a software package called "Visual DialectoMetry (VDM)".[2] This programme enables a statistical analysis of geolinguistic information according to the principles and methods used by the Salzburgschool of dialectometry. Dialectometry (DM) was developed as a method of quantitative exploration of data found in linguistic atlases and methodologically responds to the following formula: "DM = geolinguistic + numerical taxometry, or classification" (GOEBL 2010: 64). This method involves starting from a selection of working maps of a linguistic atlas with two-dimensional information (locality and linguistic variant). This information is condensed in a data matrix with the dimensions working map x enquiry point. The VDM programme processes this information and establishes a similarity between localities or enquiry points. The similarity measure used by default in the programme is the "Relative Identity Value (RIV)", which measures "the percentage of pairwise matchings between the discrete nominal (or qualitative) types of those linguistic features that are registered in the data matrix" (GOEBL 2010: 65). Finally, from this value a similarity matrix is obtained from which the cartographic visualisations are created, depending on the aims of the research. In the current study the VDM programme is used to analyse the quantitative distribution of loanwords across the Galician language area, to produce maps and to examine the existence of distribution patterns.

1 According to TADMOR (2009) the proportion of borrowable content words is much higher than of function words. The same paper gives data which show the word class most likely to be borrowed is the noun.

2 "Visual DialectoMetry" is a computer programme created by EDGAR HAIMERL in Salzburg between 1997 and 2000. Using data taken from a database, it enables automatic calculations of indices of similarity, skewness, etc., and the creation of different kinds of dialectometrical maps. See <www.dialectometry.com>; last accessed 4th April 2019.

The maps in Figures 8–11 are created with this programme. For the maps in Figures 1–7 QGIS software, an open-source and free geographic information system (GIS) application, was used.

The answers to each questionnaire item were grouped into taxates (SOUSA 2008), reducing them to just two possibilities: 1 = a Galician form (native word); 2 = a Spanish form (loanword). Note that the Galician words that are replaced by Spanish forms are not always the same ones throughout the territory. In some cases, there are simply phonological variations in the form of a lexeme, for example Gal. *morcego*, *moricego*, *mouricego*, *murucego*, *morcengo* 'bat' may be replaced by Sp. *murciélago*, *murciégalo*, *murciégano* (all the originally Galician forms are subsumed under 1, and the Spanish forms under 2). In other cases, a variety of different Galician words are replaced by one Spanish word, for example the dialectal Gal. variants *pedraza*, *pedrazo*, *pedras*, *sarabia*, *saraiba*, *escarabana* 'hail' are replaced by Sp. *granizo* (again, all the Galician forms are grouped together under 1). We have ignored phonological variations that, from the point of view of the lexicon, are considered minor or easily predictable from geographical location. Normally, a single Spanish form replaces several Galician forms (SOUSA 2008).

It must be pointed out that sometimes and in certain places two replies were given in the questionnaire, one a native word and the other a loanword. Given that VDM only allows us to process one reply per question in any one place, we opted to register the Spanish loan, in order to be able to map the extent of Spanish influence even when this is only incipient. Furthermore, when an item on the questionnaire lacks a reply it has been codified as 0, which is treated as a different reply. Thus, it is also necessary to point out that the data of 30, 35 and 40 (in the south of the centre of the province of A Coruña) are not reliable and should be ignored, since they lack answers for many questions in the questionnaires of the ALGa.

4. Analysis

This article examines aspects of the contact between Galician and Spanish. Firstly, Galician dialects shall be classified taking into account the number of loanwords they have adopted from Spanish and Leonese. Secondly, the similarity among Galician dialects shall be analysed taking into account the loanwords they share.

4.1 Degree of borrowings in the Galician dialects

As Table 1 shows, the frequency of loanwords varies widely from item to item. The Spanish words for 'foot', 'crown (of the head)', 'chin' and 'shame' occur very rarely, whereas for other items (such as 'knee' and 'eyelid') the Spanish terms are overwhelmingly more common.

A set of analysed words refers to parts of the body: *xeonllo* 'knee', *testa* 'forehead', *pálpebra* 'eyelid', *pé* 'foot', *cocote* 'nape', *queixo* 'chin'. There is an interesting situation with

Item	Localities with Spanish words	Localities with Galician words	Localities with no answer
xeonllo 'knee'	157	10	0
testa 'forehead'	138	29	0
pálpebra 'eyelid'	131	28	8
pé 'foot'	2	165	1
cocote 'crown'	3	164	0
queixo 'chin'	6	161	0
andoriña 'swallow'	46	119	12
vergonza 'shame'	7	158	2
orfo 'orphan'	121	37	9

Table 1: Example of data used in the study

names for joints in the human body: the Spanish word *nudillo* 'knuckle' occurs at 8 localities in contrast to the Galician words *cotelo*, *cotobelo*, etc. which is used in 158 places, whereas Sp. *rodilla* 'knee' occurs in 157 places as opposed to only 10 localities where Galician words were supplied. Again, among parts of the face, Sp. *párpado* 'eyelid' occurs at 131 localities on the map, contrasting with only 28 places with Galician names (*pálpebra*, *capela*, etc.); whereas Sp. *barbilla* 'chin' only occurs in 6 localities, while 161 localities use Galician forms such as *queixo*, *papo*, etc. Spanish influence predominates in the names of the days of the week: 105 of the 167 localities covered in ALGa use the Spanish word *lunes* 'Monday', 149 use Sp. *jueves* 'Thursday' and 154 use Sp. *Viernes* 'Friday'. Even an adverb like *lonxe* 'far' is replaced by Spanish equivalents at 49 localities on the map.

Looked at separately, Figures 1–4 (see appendix) do not reveal any distribution pattern of the Spanish loanwords. In fact, each Spanish form appears to tell a story of its own, represented by differing geographical distributions. In Figure 1, the native form *xeonllo* 'knee' appears to be mainly limited to peripheral areas. Figure 2 shows that Spanish forms like *pañuelo* 'handkerchief' were recorded in scattered localities. Figures 3 and 4 reveals that Spanish forms *golondrina* 'swallow' and *desayunar* 'to (have) breakfast' appear in the peripheral areas.

Sometimes the Spanish loans seem to form compact clusters. For example, Sp. *ganado* 'livestock' (at 36 localities) appears mainly to be a central-eastern form; Sp. *muela* 'molar' (78) forms fairly compact enclaves; Sp. *martillo* 'hammer' (81) mainly shows up in the west; Sp. *codo* 'elbow' (108) extends through the centre and south, with Galician forms mainly in north-western and north-eastern areas. There are other Spanish loans, such as Sp. *nuez* or *nudo* 'Adam's apple', Sp. *nudillo* 'knuckle', Sp. *ingle* 'groin', that appear to be scattered without any perceptible geographical patterning.

Therefore, when we look at the maps one by one, it is hard to perceive any kind of areal patterning. However, by quantifying the data and transferring the contents of the quantified database to a map, it is possible to discover patterns of distribution. Figure 5 was generated by calculating the presence of taxate 2 (the Spanish loanwords) and clas-

sifying the localities in five groups depending on the value of this count. The colour scale (from very light blue to very dark blue) indicates the occurrence of Spanish loans: light blues are used for lower values, and medium and dark blues for higher values, so that the darker the blue, the higher the indices of the presence of Spanish loanwords.

Places with over 20 % of loanwords are located in areas which, although not completely continuous, are mostly in the eastern half of the Galician area, which means that proximity to dialects of Asturian-Leonese and, indirectly, also to Castilian dialects seems to have played a relevant role in the spread of Spanish words. In fact, it must be noted that the language collected in the most southeasternmost locality in the net (149, one of the localities coloured in the darkest blue in Fig. 5) is really an Asturian-Leonese dialect.[3] It is to be noted that Asturian-Leonese and Spanish share many phonological developments (for example retention of Latin intervocalic -L- and -N-, palatalization (BrE) of Latin -LL- and -NN-, and diphthongisation (BrE) of Latin stressed -Ĕ- and -Ŏ-), so that some loanwords of this area could have their origins in Asturian-Leonese and not in Spanish.

Thus, Figure 5 (see appendix) shows that the eastern half of the territory constitutes an area with stronger Spanish influence than the rest, with a large, compact zone with colours ranging from medium to very dark blues. Within this area, the highest percentages occur where south-eastern Lugo, north-eastern Ourense and southern León meet. Conversely, much of the centre and west of Galicia presents lower percentages of Spanish borrowing. A conspicuous white patch occurs in the western part of the province of A Coruña (12, 16, 17, 18, 22, 27, 33, 38, 30, 42, 45) with values ranging from 5 % to 14 % of questionnaire items exhibiting Spanish loans. This small area is part of one of the linguistically most isolated Galician-speaking zones, the Galician Northwest, which retains many other conservative linguistic elements. This area is described in ÁLVAREZ/DUBERT GARCÍA/SOUSA (2006), DUBERT GARCÍA (2011) and DUBERT GARCÍA (2012) as a zone that is resistant to the introduction of innovations and the adoption of features of other areas. One interesting feature of Figure 5 is the existence of localities with low rates of loanwords (light blue) in direct, immediate, contact with places with high rates (dark blue), without any kind of transition (medium blue). This seems to be more common in some northern and southern extreme areas; a more detailed study is needed to try to explain this distribution.

However, a kind of gradience can be detected: western very light blue areas, central light blue areas, eastern medium and dark blue areas. All those areas are sometimes sprinkled by splashes in different blues (dark in the light area and vice versa), both on the coast and inland. Strips of medium blue and dark appear along the coast of the province of Pontevedra (61, 62, 77, 80 and 82), at isolated localities in the west of A Coruña province (24 and 34), the northern Coruña coast (7, 8, 15, 10 and 11), further to the

3 This Asturian-Leonese speaking locality is Le.5. This place is encoded on the map as Le.5 (Pombriego, in León province). The ALGa locations near this place also display features of Asturian-Leonese, although to a minor extent.

south (23 and 31), and at the meeting point of south-eastern A Coruña province and north-eastern Pontevedra province.

Figures 6 and 7 (see appendix) help to understand the factors that affect the geographical distribution of loanwords. Figure 6 includes the areas with the highest indices of Spanish loans, the major cities of Galicia, and the main traffic routes. The map in Figure 7 shows the density of the Galician population in the Census of 1970 (please remember that the data of the "Atlas Lingüístico Galego" were gathered in the mid-seventies). Three patterns may be found:

A. In the eastern area, a low population density coincides with a moderately high degree of borrowing from Spanish. So, the presence of Spanish loanwords cannot be accounted for by the existence of large urban areas. This area of moderately high Spanish influence is located along the border between Galician and Asturian-Leonese/Spanish. The Spanish influence may be due to waves arriving from the centre of the Iberian Peninsula through horizontal language contact among different dialects. The idea of a point of convergence between Galician and Spanish located in the eastern part of Galicia is supported by the distributions of other linguistic features, including *gheada* and *seseo* (FERNÁNDEZ REI 1990)[4], phenomena that are well preserved in western Galician.
B. In the western localities displaying moderately high borrowing from Spanish, the population density tends to be higher, which seems to indicate an influence from the presence of Spanish in towns and cities. This pattern is observed in localities 7, 34, and 61, with relatively high numbers of Spanish loans and high population densities. This Spanish influence may be better accounted for by the social position of Spanish, exerting a vertical, top-down effect.
C. Pathways for the spread of Spanish lexical items turn out to be traffic routes with areas outside Galicia: the road from León to Pontevedra, the Lugo coastal road which starts in Asturias, and to some extent, the road from Zamora to the South of Galicia. These routes, which are nowadays motorways, have served as paths of communication for people and goods for centuries, so it is hardly surprising that they have played an important part in the introduction and spread of Spanish lexical forms in Galicia.

In different degrees, such top-down penetration must also have been taking place in the eastern geographical area, that of horizontal contact, since Spanish was/is the high, dominating, language in all the Galician speaking territory. It would be interesting to see whether Spanish loanwords entering through both channels (horizontal versus top-down contact) are the same ones or different ones, and whether there is a lexical distribution associated with different forms of penetration.

4 The westernmost Galician dialects preserved as an alveolar consonant an old lamino-alveolar fricative one, which the central and eastern dialects transformed into a dental consonant; the preservation of the alveolar feature is called *seseo*. The western dialects transformed an old voiced velar stop into a glottal continuant; this phenomenon is called *gheada*; most of the eastern dialects preserved the old velar or are recovering it. Both *seseo* and *gheada* characterise western Galician vis-à-vis Asturian-Leonese and northern Peninsular Spanish (DUBERT GARCÍA/GALVES 2016).

4.2 Interdialectal similarity with respect to borrowing

Since the main concern of dialectometry is not individual maps or features, but rather entire masses of numeric data, we cannot analyse the spreading of individual words or loanwords. We can, however, take a look at the degree of linguistic similarity between different localities in the ALGa's net as regards the loanwords they have adopted. In doing so, we aim to find out if patterns can be detected in their geographical distribution. Figure 8 was obtained from the database of SOUSA (2017a), who analysed the lexical variation of the ALGa. This database includes 136 working maps selected for having mainly native words and without taking into account the presence of loanwords in the sample. Figure 8 classifies the localities of the ALGa's net by the mean of similarity each locality reaches. To calculate the mean of similarity of each locality (DUBERT GARCÍA 2011, GOEBL 1981, 1993), first, VDM compares the answers of each locality with the answers of the others, in such a way that it obtains 166 inter-punctual values of similarity for each locality (the ALGa's net has 167 localities). Next, VDM calculates the mean of the 166 values for each locality: this is the mean of similarity or mean similarity of each locality. Finally, VDM divides the range of the means of the 167 localities into groups according to their values and gives warm colours to the localities with higher values and cold colours to the points with lower values; VDM uses different algorithms to break the continuum of values and assign them the corresponding colour (see GOEBL 1987 for the different algorithms).

Figure 8 (see appendix) shows that the central and western dialects exhibit higher means of similarity than the oriental and south-western dialects. GOEBL (1981: 389) considers that "dans une perspective communicative [...], la moyenne arithmétique d'une distribution de similarité peut être utilisée à en évaluer numériquement la position central au sein du réseau examiné".

Figure 9 (see appendix) was obtained by using the same method with our present database. VDM also compares all the localities with each other, so that finally a mean similarity index is obtained for each locality. However, in this case, what is compared is the similarity of the localities with respect to the individual loanwords they have introduced: each locality is compared with the rest in terms of the loanwords they share. Figure 9 shows a slightly different geographic profile from Figure 8, with warm and cold colours distributed in a different way; this signifies that the profiles of similarity of Figures 8 and 9 are different: loanwords and general lexicon seem to create different geolinguistic structures. However, if we focus on the values of the means of similarity obtained as presented in the legend, the figures of the indices are noticeably higher in Figure 9 than in Figure 8. Generally speaking, Galician dialects seem to be more homogeneous if only the loanwords are taken in account. This makes sense bearing in mind that introducing loanwords diminishes dialectal diversity, since one and the same loanword (*granizo* 'hail') can eliminate different dialectal native variants (*pedraza*, *sarabia*, *escarabana*).

However, the most interesting feature of Figure 9 is that the area with the highest levels of similarity is the occidental one, i. e., the one with the least loanwords. Conversely, the area with the lowest levels of similarity is the oriental one, i. e., the one with the

most loanwords. This could indicate that within the western area, the loanwords tend to be the same across the different local varieties. In the oriental area, on the other hand, the loanwords introduced in the variety of each locality may be slightly different. This distribution could be consistent with the horizontal source of loanwords in the eastern varieties, where each Spanish or Asturian-Leonese word follows its own different path.

Figure 10 (see appendix) shows the distribution of dialects according to the values of the standard deviation corresponding to the means of Figure 9. To prepare Figure 10, we asked VDM to calculate the standard deviation of the means of similarity that it obtained for each locality in Figure 9. Again, an algorithm divides the range of values, in this case, in four groups and assigns warm colours to high indices and cold colours to low indices. It is useful to look at both Figures 9 and 10 together so that we observe both the mean similarity index and the standard deviation index for each locality (for the interpretation of standard deviation in dialectometry, see DUBERT GARCÍA 2011). This information is complementary, because a mean may either reflect a compact concentration of values that are close to each other (i. e. a low standard deviation, with greater similarity within the sample) or conceal dispersed values that differ substantially from the mean (i. e. a high standard deviation, with big differences within the sample). From this comparison we may see again that in general terms western Galician dialects opposes eastern dialects. The areas with more loanwords tend to be the same as those with low mean similarities; and their low means tend to go with a low standard deviation. Low means of similarity (cold colours in Figure 9) and low standard deviations (cold colours in Figure 10) signify that the means of each locality are reliable in the sense that each locality consistently has lower indices of similarity with the other localities of the net. If that is the case, then each local variety of this area has had a certain autonomy in the selection of loanwords it has taken. Occidental areas, however, present high means of similarity (warm colours in Figure 9) and high indices of standard deviation (warm colours in Figure 10). This signifies than in the western areas we find dialects that are very similar to one another because they share the same loanwords, and that are very different from dialects of the other areas because they have different loanwords.

Finally, Figure 11 (see appendix) shows the degree of correlation between geographical proximity and linguistic similarity. It is possible to suppose that the linguistic similarity between two dialects increases as geographical proximity between them does; conversely, linguistic difference between two dialects may increase as geographical distance does. Since VDM enables the calculation of the degree of correlation between two distributions of data, we can explore in what degree the linguistic and geographical distance[5] are related in the ALGa's net. This method is known as "correlative dialectometry" (for a detailed explanation of correlative dialectometry, see DUBERT GARCÍA 2012, 2013). Thus, we are going to analyse in what degree the similarity of dialects in terms of the loanwords they share is related to the geographic distance that separates them. The statistical metric that VDM uses to show the degree of correlation is the "Bravais-Pear-

5 It should be noted that the distance measured is the distance in a map, ignoring topography, rivers, mountains, valleys, roads, etc.

son correlation coefficient". The VDM programme calculates this coefficient for every locality on the net. The values of this coefficient range from -1 to +1. If the value obtained is -1, we have a perfect inverse linear relation: when X (proximity) increases, Y (similarity) decreases proportionally; if the value is 0 we have no linear correlation between X and Y, and proximity and similarity seem to be autonomous; when the value is +1, we find a direct linear relation: when X (proximity) increases, Y (similarity) increases proportionally. Warm colours are applied to the higher values of the coefficient (those closer to 1); cold colours are applied to lower values (those farther from 1).

As Figure 11 shows, western Galician, the one with less loanwords and more internal uniformity, tends to show direct linear correlation between similarity and proximity (or conversely, between difference and distance): the similarity in the loanwords taken increases with nearness. On the contrary, in eastern Galician, the one with more loanwords and less internal uniformity, proximity and similarity (or, conversely, distance and difference) tend to be autonomous. This is consistent with the fact that eastern dialects have less similarity in the loanwords they have taken. Only Asturian Galician, to the northeast, is an exception: high indices of loanwords and tendency to direct linear correlation between proximity and similarity. Probably this is due to the fact that Asturian Galician is a characteristic dialect spoken outside Galicia, in close contact with Asturian varieties and Asturian Spanish. The other localities in the Galician net with many loanwords are surrounded by areas which have better preserved their native lexicon (Fig. 5).

5. Conclusions

This study is based on a selection of lexical items gathered from the ALGa's database. We have carried out two kinds of analyses: first (§ 4.1), we have classified the Galician dialects bearing in mind the number of loanwords they have adopted from Spanish and Leonese; secondly (§ 4.2), we have studied the degree of similarity among the Galician dialects based on the loanwords they have acquired or the native lexicon they have preserved.

Given the exploratory nature of this paper, its conclusions can only be tentative pending analysis of further items. However, the aggregating quantitative method of dialectometry yields interesting results through which to explore a language contact situation, at least on the lexical level. Two areas are distinguishable: an eastern area with Spanish influence through geographical contact with Spanish and Leonese, and a western area with less Spanish influence overall, but which contains foci of influence potentially linked to districts with higher levels of population density and with traffic routes. We also found localities with high numbers of loanwords in direct contact with localities that have preserved their native lexicon to a greater extent, without a transition area; a more exhaustive analysis could explain the existence of these spots.

The western dialects are also more homogeneous as regards the loanwords they have taken and the native lexicon they have preserved; those dialects also tend to show a cor-

relation between linguistic and geographic proximity. The eastern dialects, by contrast, are less homogeneous as regards the loanwords they have adopted and tend to be more autonomous in the loanwords they have assimilated.

References

ALGA (1990–): Atlas Lingüístico Galego. A Coruña – Santiago de Compostela: Fundación Barrié – Universidade de Santiago de Compostela.

ÁLVAREZ, ROSARIO / FRANCISCO DUBERT GARCÍA / XULIO SOUSA (2006): Aplicación da análise dialectométrica aos datos do *Atlas Lingüístico Galego*. In: ÁLVAREZ, ROSARIO / FRANCISCO DUBERT GARCÍA / XULIO SOUSA (ed.): Lingua e territorio. Santiago de Compostela: ILG–CCG, 461–493.

BRITAIN, DAVID (2010): Contact and Dialectology. In: HICKEY, RAYMOND (ed.): The Handbook of Language Contact. Oxford: Blackwell Publishing, 208–229.

BRITAIN, DAVID (2018): Dialect contact and New Dialect Formation. In: BOBERG, CHARLES / JOHN NERBONNE / DOMINIC WATT (ed.): The Handbook of Dialectology. Hoboken, NJ: John Wiley & Sons, 143–158.

CINTRA, L. F. LINDLEY (1983): Áreas lexicais no território português. In: CINTRA, L. F. LINDLEY (ed.): Estudos de dialectologia portuguesa. Lisboa: Sá da Costa, 55–94.

CONTINI, MICHEL (1991): L'Atlas Linguistique Roman. In: Nazioarteko Dialektologia Biltzarra. Agiriak 7, 339–356.

DUBERT GARCÍA, FRANCISCO (2005): Interferencias del castellano en el gallego popular. In: Bulletin of Hispanic Studies 82 (3), 271–292.

DUBERT GARCÍA, FRANCISCO (2010): As formas *camín* e *chen* en galego e asturiano: solucións supra-rexionais, solucións rexionais e contínuum xeolectal. In: Revista de Filología Románica 27, 29–58.

DUBERT GARCÍA, FRANCISCO (2011): Developing a database for dialectometric studies: the ALGa phonetic data. Dialectometrical analysis of 230 working maps. In: Dialectologia et Geolinguistica 19, 23–61.

DUBERT GARCÍA, FRANCISCO (2012): Geographic proximity and Linguistic similarity. A correlative analysis of Galician phonetic data. In: ÁLVAREZ PÉREZ, XOSÉ AFONSO / ERNESTINA CARRILHO / CATARINA MAGRO (ed.): Proceedings of the International Symposium on Limits and Areas in Dialectology (LimiAr), Lisbon 2011. Lisboa: Centro de Linguística da Universidade de Lisboa, 71–91.

DUBERT GARCÍA, FRANCISCO (2013): An analysis of Galician Dialects in Correlative Dialectometry. In: ÁLVAREZ PÉREZ, XOSÉ AFONSO / ERNESTINA CARRILHO / CATARINA MAGRO (ed.): Current Approaches to Limits and Areas in Dialectology. Newcastle upon Tyne: Cambridge Scholars Publishing, 171–198.

DUBERT GARCÍA, FRANCISCO (2017): Sobre a *Gallaecia Magna* e as relacións históricas e xeolingüísticas entre galego, portugués e asturiano. In: Estudis romànics 39, 43–69.

DUBERT GARCÍA, FRANCISCO / CARLOS ACUÑA-FARIÑA (2018): Restructuring and complexification of inflectional morphology under linguistic contact: the case of a Galician dialect. In: CUZA, ALEJANDRO / PEDRO GUIJARRO-FUENTES (ed.): Language Acquisition and Contact in the Iberian Peninsula. Berlin: Mouton de Gruyter, 187–213.

DUBERT GARCÍA, FRANCISCO / CHARLOTTE GALVES (2016): Galician and Portuguese. In: LEDGEWAY, ADAM / MARTIN MAIDEN (ed.): The Oxford Guide to the Romance Languages. Oxford: Oxford University Press, 411–446.

DUBERT GARCÍA, FRANCISCO / XULIO SOUSA (2002): Áreas lexicais galegas e portuguesas. A proposta de Cintra aplicada ó galego. In: ÁLVAREZ, ROSARIO / FRANCISCO DUBERT GARCÍA / XULIO SOUSA (ed.): Dialectoloxía e Léxico, Santiago de Compostela: Instituto da Lingua Galega / Consello da Cultura Galega, 193–222.

ENTWISTLE, WILLIAM J. (1973): Las lenguas de España: castellano, catalán, vasco y gallego-portugués. Madrid: Itsmo.

FERGUSON, CHARLES (1959): Diglossia. In: Word 15 (2), 325–340.

FERNÁNDEZ-ORDÓÑEZ, INÉS (2009): Los orígenes de la dialectología hispánica y Ramón Menéndez Pidal. In: VIEJO, XULIO (ed.): Cien Años de Filoloxía Asturiana (1906–2006): Actes del Congresu Internacional, 11–41.

FERNÁNDEZ-ORDÓÑEZ, INÉS (2011): La lengua de Castilla y la formación del español. Madrid: Real Academia Española.

FERNÁNDEZ REI, FRANCISCO (1990): Dialectoloxía da lingua galega. Vigo: Xerais.

FRANCO, KARLIEN / DIRK GEERAERTS / DIRK SPEELMAN / ROELAND VAN HOUT (2019): Maps, Meanings and Loanwords: The Interaction of Geography and Semantics in Lexical Borrowing. In: Journal of Linguistic Geography 7 (1), 14–32.

GARCÍA GONZÁLEZ, CONSTANTINO (1976): Interferencias lingüísticas entre gallego y castellano. In: Revista de la Sociedad Española de Lingüística 6 (2), 327–343.

GOEBL, HANS (1981): Éléments d'analyse dialectométrique (avec application à l'AIS). In: Revue de Linguistique Romane 45, 349–420.

GOEBL, HANS (1987): Poins chauds de l'analyse dialectométrique: pondération et visualisation. In: Revue de Linguistique Romane 51, 63–118.

GOEBL, HANS (1993): Dialectometry: A Short Overview of the Principles and Practice of Quantitative Classification of Linguistic Atlas Data. In: KÖHLER, REINHARD / BURGHARD B. RIEGER (ed.): Contributions to Quantitative Linguistics. Dordrecht: Kluwer, 277–315.

GOEBL, HANS (2010): Dialectometry: Theoretical prerequisites, practical problems, and concrete applications (mainly with examples drawn from the "Atlas Linguistique de la France", 1902–1910). In: Dialectologia. Special Issue I, 63–77.

HASPELMATH, MARTIN (2008): Loanword typology: Steps toward a systematic cross–linguistic study of lexical borrowability. In: STOLZ, THOMAS / DIK BAKKER / ROSA SALAS PALOMO (ed.): Aspects of language contact: New theoretical, methodological and empirical findings with special focus on Romancisation processes. Berlin: Mouton de Gruyter, 43–62.

HASPELMATH, MARTIN (2009): Lexical borrowing: concepts and issues. In: HASPELMATH, MARTIN / URI TADMOR (ed.): Loanwords in the World's Languages: A Comparative Handbook. Berlin: Mouton de Gruyter, 35–54.

HASPELMATH, MARTIN / URI TADMOR (2009): The Loanword Typology project and the World Loanword Database. In: HASPELMATH, MARTIN / URI TADMOR (ed.): Loanwords in the World's Languages: A Comparative Handbook. Berlin: Mouton de Gruyter, 1–34.

HERMAN, JÓZSEF (1990): Du latin aux langues romanes: études de linguistique historique. Berlin/Boston: Mouton de Gruyter.

HOUT, ROELAND VAN / PIETER MUYSKEN (1994): Modelling lexical borrowability. In: Language Variation and Change 6, 39–62.

LÜDTKE, HELMUT (1968): Geschichte des romanischen Wortschatzes. Freiburg: Rombach.

MARIÑO PAZ, RAMÓN (1998): Historia da lingua galega. Santiago de Compostela: Sotelo Blanco.

MONTEAGUDO, HENRIQUE (1999): Historial social da lingua galega. Idioma, sociedade e cultura a través do tempo. Vigo: Galaxia.

MONTEAGUDO, HENRIQUE (2017): A lingua no tempo, os tempos da lingua. O galego, entre o portugués e o castelán. In: NEGRO ROMERO, MARTA / ROSARIO ÁLVAREZ / EDUARDO MOSCOSO MATO (ed.): Gallæcia. Estudos de lingüística portuguesa e galega. Santiago de Compostela: USC, 17–60.

MULJAČIĆ, ŽARKO (1991): Per un approccio relativistico al rapporto: lingua nazionale dialetti. In: PELLEGRINI, GIOVAN BATTISTA (ed.): XVIII Congresso di studi dialettali italiani "Fra dialetto e lingua nazionale: realtà e prospettive". Padova: Unipress, 247–259.

MYERS-SCOTTON, CARON (2002): Contact linguistics: Bilingual encounters and grammatical outcomes. Oxford: Oxford University Press.

RAMALLO, FERNANDO / GABRIEL REI-DOVAL (2015): The standardization of Galician. In: AMMON, ULRICH / JEROEN DARQUENNES / SUE WRIGHT (ed.): Die Standardisierung von Minderheitensprachen. The standarization of minority languages. La standarisation des langues minoritaires. Berlin/Boston: Mouton de Gruyter (Sociolinguistica. Internationales Jahrbuch für europäische Soziolinguistik. 29), 61–81.

ROHLFS, GERHARD (1960): Diferenciación léxica de las lenguas romances. Madrid: Publicaciones de la Revista Española de Filología.

SOUSA, XULIO (2008): Criterios de clasificación de datos léxicos en el análisis dialectométrico. In: BLAIKNER-HOHENWART, GABRIELE (ed.): Ladinometria. Festschrift für Hans Goebl zum 65. Geburtstag. Salzburg/Bozen: Universität Salzburg / Freie Universität Bozen / Istituto Culturale Ladino, 123–139.

SOUSA, XULIO (2017a): Aggregate analysis of lexical variation in Galician. In: BUCHSTALLER, ISABELLE / BEAT SIEBENHAAR (ed.): Language Variation. European perspectives VI. Amsterdam: John Benjamins, 71–83.

SOUSA, XULIO (2017b): From field notebooks to automatic mapping: the Atlas Lingüístico Galego database. In: Dialectologia et Geolinguistica 25, 1–22.

TADMOR, URI (2009): Loanwords in the world's languages: Findings and results. In: HASPELMATH, MARTIN / URI TADMOR (ed.): Loanwords in the World's Languages: A Comparative Handbook. Berlin: Mouton de Gruyter, 35–54.

TADMOR, URI / MARTIN HASPELMATH / BRADLEY TAYLOR (2010): Borrowability and the notion of basic vocabulary. In: Diachronica 27 (2), 226–246.

THUN, HARALD (2010): Variety complexes in contact: A study on Uruguayan and Brazilian Fronterizo. In: AUER, PETER / JÜRGEN ERICH SCHMIDT (ed.): Language and Space: An International Handbook of Linguistic Variation. Volume 1: Theories and Methods. Berlin: Mouton de Gruyter (Handbücher zur Sprach- und Kommunikationswissenschaft. 30.1), 706–723.

VENÂNCIO, FERNANDO (2013): Atitudes portuguesas face ao castelhano. In: GUGENBERGER, EVA / HENRIQUE MONTEAGUDO / GABRIEL REI-DOVAL (ed.): Contacto de linguas, hibridade, cambio: contextos, procesos e consecuencias. Santiago de Compostela: Consello da Cultura Galega, 315–326.

WINFORD, DONALD (2010): Contact and Borrowing. In: HICKEY, RAYMOND (ed.): The Handbook of Language Contact. Oxford: Blackwell Publishing, 170–187.

Appendix

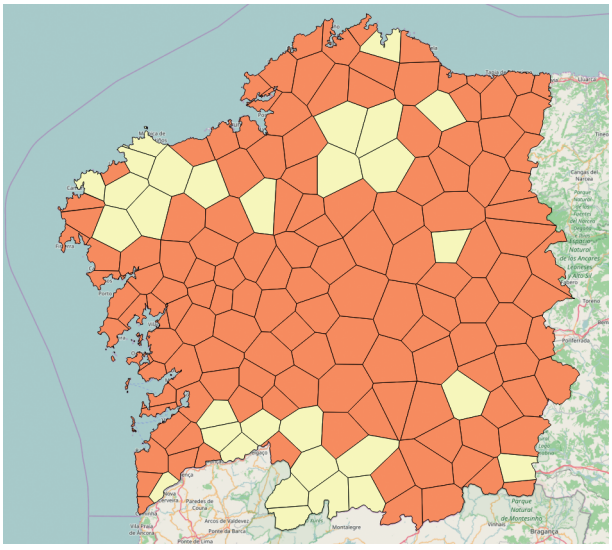

Fig. 1: 'Knee': Gal. *xeonllo* (yellow) / Sp. *rodilla* (orange)

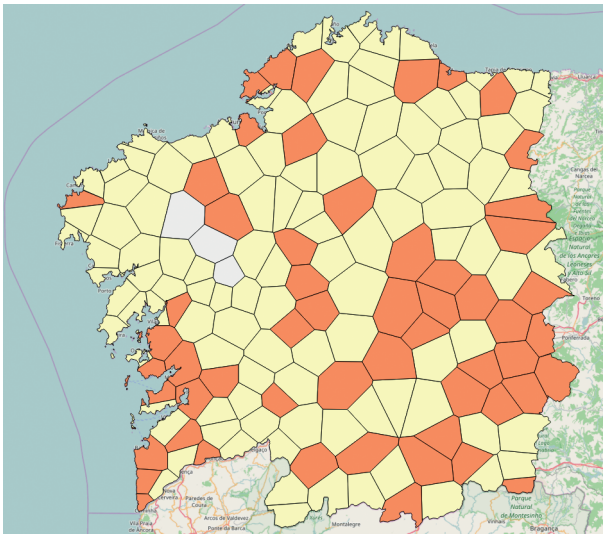

Fig. 2: 'Handkerchief': Gal. *pano da man* (yellow) / Sp. *pañuelo* (orange)

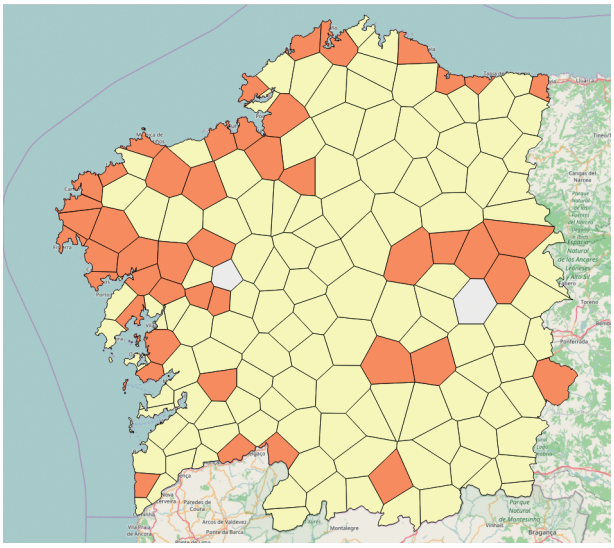

Fig. 3: 'Swallow': Gal. *andoriña* (yellow) / Sp. *golondrina* (orange)

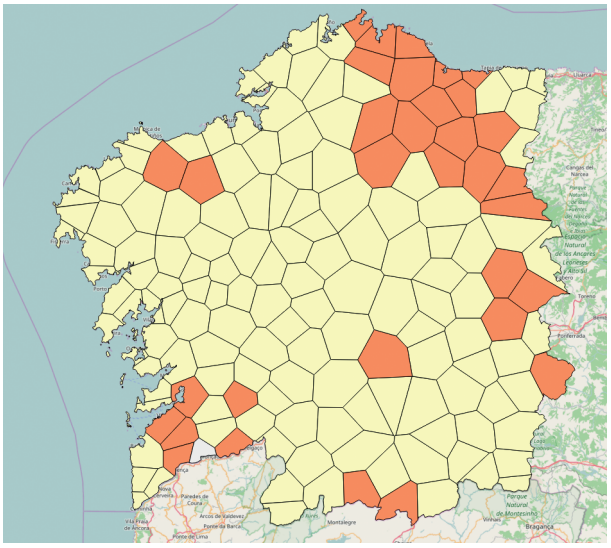

Fig. 4: 'To breakfast': Gal. *almorzar* (yellow) / Sp. *desayunar* (orange)

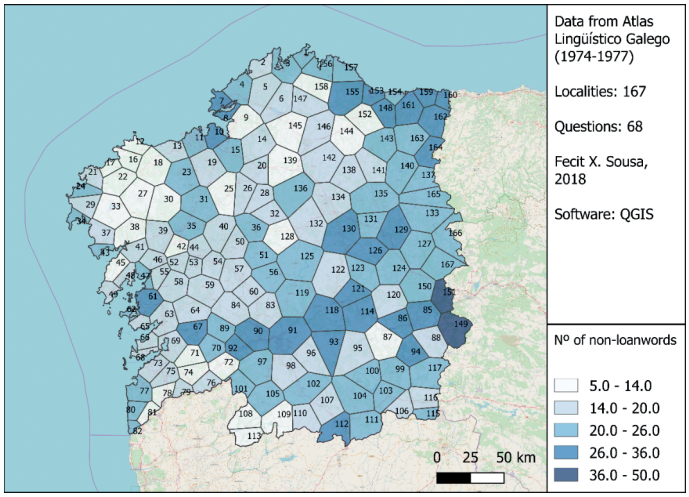

Fig. 5: High index of loans

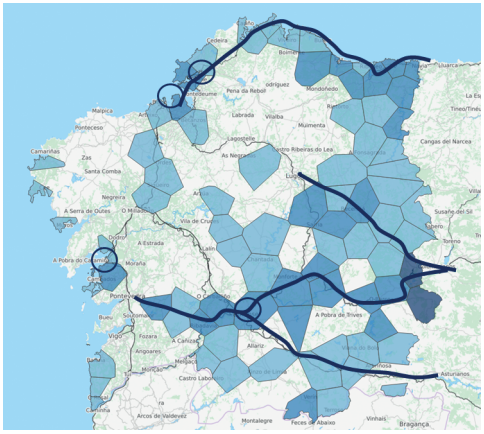

Fig. 6: Loanwords, roads and major cities

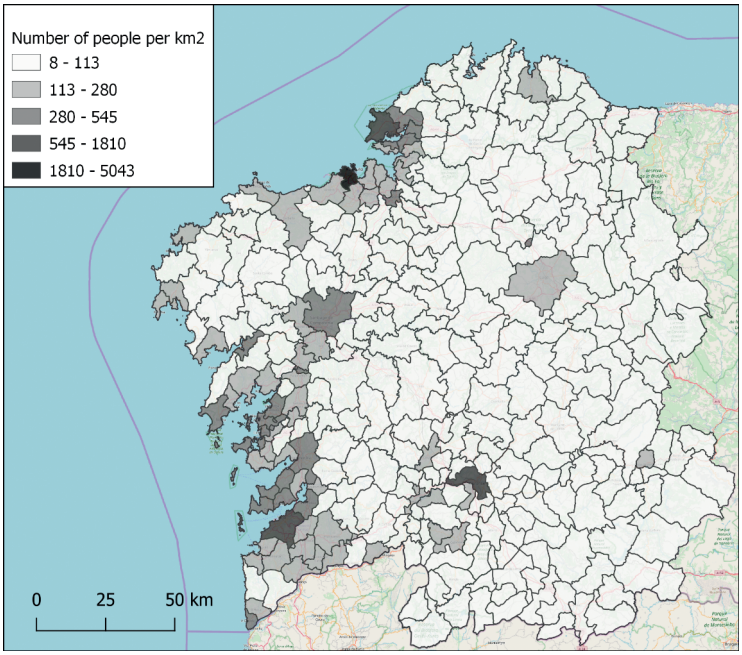

Fig. 7: Population density map based on Census 1970 data

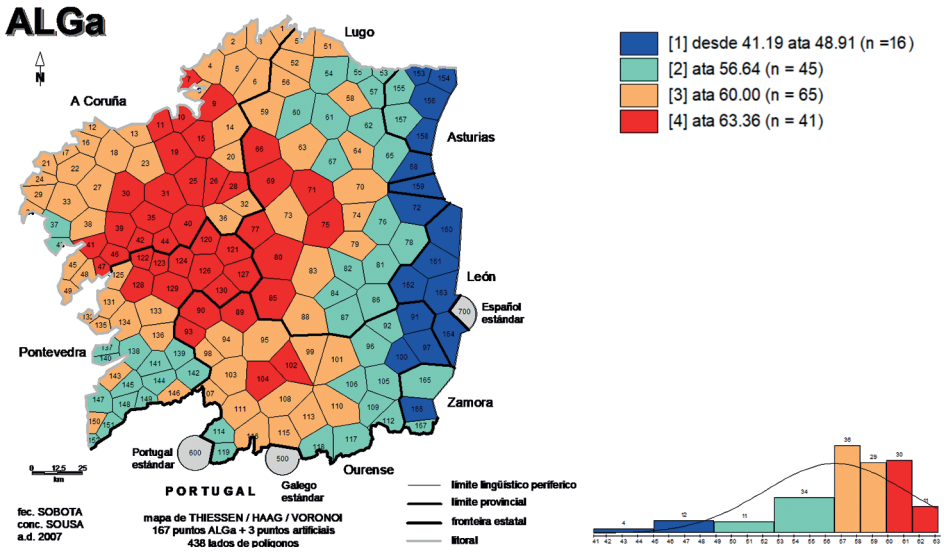

Fig. 8: Mean similarity indices for each locality with respect to the general lexicon of the ALGa

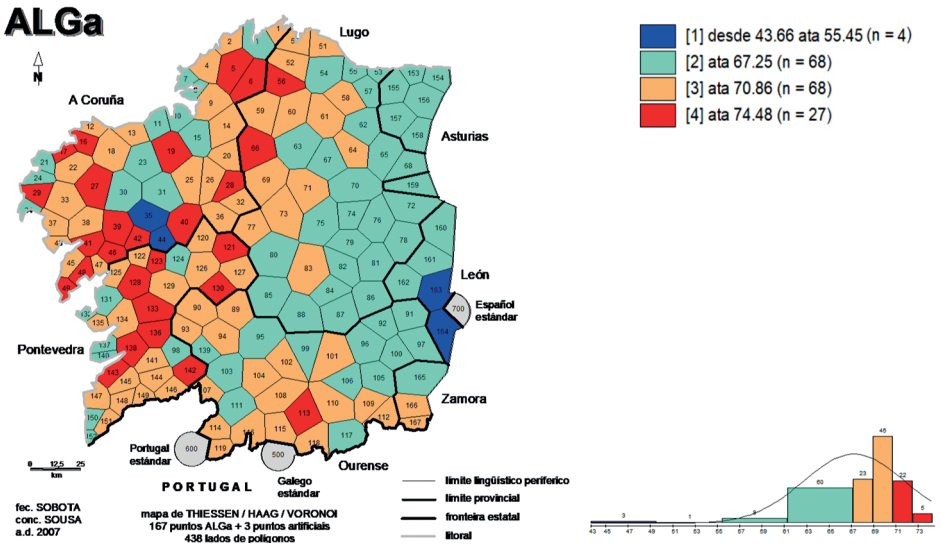

Fig. 9: Mean similarity indices for each locality as regards the loanwords taken

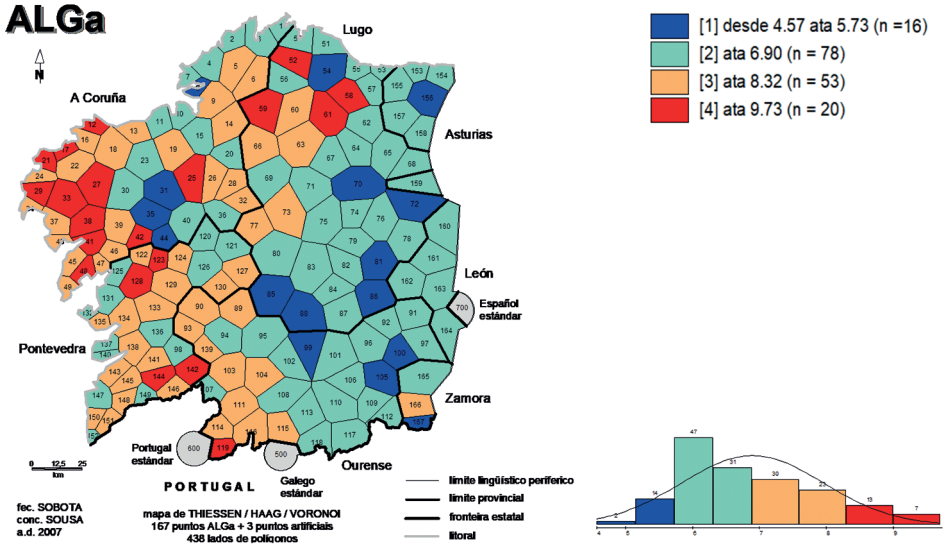

Fig. 10: Choropleth map of the standard deviation (similarity index)

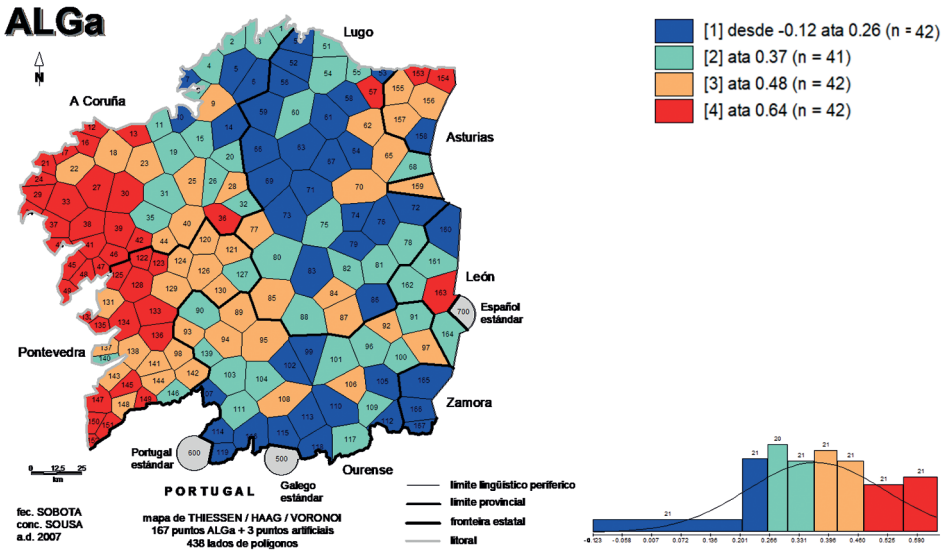

Fig. 11: Choropleth map of the correlation between linguistic and geographical distances

XULIO SOUSA
Facultade de Filoloxía, Universidade de Santiago de Compostela, 15782 Santiago de Compostela, SPAIN
e-mail: <xulio.sousa@usc.es>
FRANCISCO DUBERT GARCÍA
Facultade de Filoloxía, Universidade de Santiago de Compostela, 15782 Santiago de Compostela, SPAIN
e-mail: <francisco.dubert@usc.es>