

Article

Semantic Fields and Castilianization in Galician: A Comparative Study with the *Loanword Typology* Project

María Álvarez de la Granja  and Francisco Dubert García * 

Instituto da Lingua Galega, Universidade de Santiago de Compostela, 15782 Santiago de Compostela, Spain; maria.alvarez.delagranja@usc.gal

* Correspondence: francisco.dubert@usc.gal

Abstract: This study examines the correspondence between the borrowability indices from the *Loanword Typology* (LWT) project and Castilianization indices from the *Atlas Lingüístico Galego* (ALGa) across seven semantic fields. To this end, we identified all Castilianisms in the ALGa and conducted a quantitative analysis to compare these indices. Results obtained indicate a mismatch between the rankings of the LWT project and the ALGa. For example, the field ‘The body’ has the highest level of Castilianization according to the ALGa but the lowest borrowed score in the LWT project. Moreover, Castilianization levels in the ALGa show greater dispersion than borrowability levels from the LWT project. In fact, in each semantic field, many concepts (52.2%) have low levels of Castilianization, between 0% and 10%, and only a few concepts have high levels. A more detailed analysis of three semantic fields (‘The body’, ‘Agriculture and vegetation’, and ‘The physical world’) suggests that explanations based solely on semantic criteria (such as the existence of an unalterable central lexicon) are insufficient; other factors such as prestige, urbanization, cultural modernity, frequency of word usage, and perhaps other intralinguistic factors should be taken into account.

Keywords: language contact; loanwords; semantic fields; Galician; Castilian Spanish



Citation: Álvarez de la Granja, María, and Francisco Dubert García. 2024. Semantic Fields and Castilianization in Galician: A Comparative Study with the *Loanword Typology* Project. *Languages* 9: 244. <https://doi.org/10.3390/languages9070244>

Academic Editors: Xosé Luís Regueira and Elisa Fernández Rei

Received: 6 April 2024

Revised: 29 June 2024

Accepted: 2 July 2024

Published: 9 July 2024



Copyright: © 2024 by the authors. Licensee MDPI, Basel, Switzerland. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC BY) license (<https://creativecommons.org/licenses/by/4.0/>).

1. Introduction

In a study on the linguistic influence of Castilian Spanish on Galician,¹ when discussing lexical Castilianisms, Dubert García (2005, p. 289) concluded the following:

Aunque es común escuchar que existen campos de la vida que presentan un léxico más castellanizado que otros (el léxico referido al mundo de la Iglesia o de la escuela estaría más castellanizado que el referido al mundo del trabajo en el campo), personalmente juzgo difícil, si no imposible, señalar qué requisitos puede cumplir una palabra gallega para poder ser sustituida por una castellana equivalente. Como ya he indicado, este tipo de interferencia incluso crea diferencias dialectales y sociolectales: hay dialectos en que se ha introducido *huevo* o *hueso* mientras que se ha mantenido *avó*; y dialectos que conservan *ovo* u *óso* pero que han admitido *abuelo*. Supongo que son precisas más investigaciones y más profundas para poder encontrar un sentido o un patrón.

This work aims to be a contribution in this regard. To this end, it determines the level of correspondence between borrowability indices of a selected group of semantic fields and concepts from the *World Loanword Database* (WOLD) of the *Loanword Typology* project (LWT; Haspelmath and Tadmor 2009a, 2009b)² and the Castilianization indices obtained for the corresponding concepts in the *Atlas Lingüístico Galego* (ALGa; García and Santamarina 1990).³ We also intend to compare the degree of Castilianization in the Galician language across different semantic fields. Therefore, this work enriches the reflection on linguistic contact with a quantitative analysis of linguistic data.

The LWT project includes 41 languages, but not Galician, making comparisons impossible. This prevents us from verifying to what extent Galician’s patterns of lexical

borrowing align with common linguistic behaviors or are idiosyncratic. Until now, no comparative study of this kind has been conducted on Galician, and the processes of transfer it underwent, specifically those derived from Spanish contact, have usually been treated as symptoms of abnormal behavior. Notably, only a few studies, such as those by [Negro Romero \(2013\)](#) and [Sousa and Dubert García \(2020\)](#), have quantitatively analyzed the extent of lexical Castilianization across Galician-speaking regions. Furthermore, no research has yet investigated the susceptibility of specific semantic fields to Spanish lexical transfers. Our study aims to address this significant gap by analyzing how Galician interacts with Castilian influences and comparing these patterns to those documented in the *LWT* project's languages. This comparative analysis is essential to discern whether Galician exhibits typical or anomalous borrowing behaviors and to evaluate the relative frequency of its loanwords. To effectively measure the degree of Castilianization in Galician, a broad, balanced, and representative data source is needed. The *ALGa* stands out as the only work that meets this requirement. It provides a dataset obtained uniformly, that is, using the same methodology, gathered at the same historical moment, and from informants of similar social status. This consistency allows for comparisons of potential territorial differences in Castilianization. For this reason, we work with data from this atlas, representing the degree of Castilianization that could be found in the popular language in the 1970s, before the implementation of Galician in education and the creation and dissemination of the standard variety (Section 2). It should be noted, however, that some recent studies ([Rodríguez Lorenzo 2022](#), p. 272 ff.) show that the degree of Castilianization in popular varieties has tended to increase over time in such a way that more up-to-date data would likely show higher indices of Castilian influence.

The analysis of *ALGa* data indicates that, in most cases, Castilianization in Galician does not typically involve the absolute replacement of Galician words by Castilian ones. More commonly, both linguistic forms coexist within the same area or with their distribution varying by location. This results in varying percentages of Castilianization across different concepts, as we will discuss in more detail in the methodology section (Section 3). Hence, by comparing the borrowed scores for different concepts in the *LWT* project with the degree of Castilianization found in Galician for those same meanings, we can identify divergent and convergent trends within the data found in the set of 41 languages studied in the aforementioned project.

To achieve the indicated objectives, we will first present the contact situation between Galician and Castilian (Section 2); then, we will explain the methodology and sources of the data used to carry out the comparison (Section 3), the results obtained (Section 4), and a discussion (Section 5); finally, we present some conclusions (Section 6).

2. The Contact Situation between Galician and Spanish

Language contact between Galician and Spanish is a long-standing phenomenon, as demonstrated by, for example, [Monteagudo and Santamarina \(1993\)](#), [Monteagudo \(1999\)](#), or [Mariño Paz \(2008\)](#). This contact is even evident in the earliest written texts in Galician from the mid-13th century, where there are signs of possible Castilianisms ([Boullón Agrelo and Monteagudo 2009](#)). As is well known ([Monteagudo and Santamarina 1993](#); [Monteagudo 1999](#); [Mariño Paz 2008](#)), a series of political changes that took place between the mid-14th and late 15th centuries led to a gradual increase in the presence of Castilian in Galicia, as well as the loss of contexts of use for Galician, which ceased to be employed in literary, cultural, religious, and administrative domains, and was practically no longer written. This process of Castilianization had two related aspects: an increase in the number of speakers of Spanish due to language shift processes and a consequent increase in the presence of linguistic transfers from Castilian to Galician. However, until the 18th century, Castilianization primarily affected the higher strata of society since the wider population continued to use only Galician. Over time, Spanish became associated with power and prestige, while Galician was linked to the most disadvantaged social classes and underwent a growing process of stigmatization. The expansion of education from the 18th century

onwards, with Castilian as the sole language of instruction and study, further reinforced this dichotomy, promoting not only a language shift but also various linguistic transfer processes. Crucially, as Galician was marginalized from prestigious functions, it often did not directly receive many of the lexical innovations from learned sources (Latin, classical Greek, French, etc.), receiving many such elements only through Castilian. [Monteagudo \(2003, p. 201\)](#) refers to this phenomenon where Castilian mediated the lexical renewal of Galician as *interposition*:

Desde o século XV por vía da regra o galego non recibiu inmediatamente latinismos, anglicismos ou italianismos, mais só castelanismos, isto é, vocábulos previamente peneirados polo idioma interposto e adaptados a el, aínda que estes á súa vez resultasen ser no castelán, latinismos, anglicismos ou italianismos de orixe.

In the 19th and 20th centuries, de-Galicianization became evident among the new urban classes, the middle class, and the educated population. By the mid-20th century, as social mobility and the influence of mass media increased, Spanish became popularized and extended to all social strata. Today, Castilian is known and used by most of the population, either as a first or second language.⁴ Thus, as noted by [Dubert García \(2005, p. 272\)](#), the expansion of the use of Castilian as a spoken language and as a mother tongue in Galicia is a modern phenomenon, which has led to a significant presence of Castilian transfers in Galician, on account of the increased degree of contact between the two languages. The number of lexical transfers is particularly dominant in the working classes, to which the *ALGa* speakers belong.

Galician was not introduced as a subject in early childhood, primary, and secondary education until the school year 1979–1980, that is, after the dictatorship of General Franco, in a period when the social distance between the two languages was even more reinforced. In the 1980s, a normative code for orthography, morphology, and lexicon was developed with the aim of establishing a standard variety ([Ramallo and Rei-Doval 2015](#)). This variety was then employed not only in education but also in administration, the media, and political discourse. However, at the time the *ALGa* surveys were conducted, this variety had not yet been established, and therefore left no trace on the informants. For this reason, the *ALGa* represents the common Galician as it was spoken in the 1970s.

3. Materials and Methods

3.1. The Loanword Typology Project

The *Loanword Typology* project, coordinated and developed by Martin Haspelmath and Uri Tadmor between 2004 and 2008, consists of a total of 41 languages. The project analyzed 1460 meanings (cf. the *LWT* meaning list) across 24 semantic fields, selected based on their genealogical, geographical, typological, and sociolinguistic diversity ([Haspelmath and Tadmor 2009a, p. 3](#)). Each of the contributors responsible for the 41 languages was required to provide the lexical counterparts for the different concepts,⁵ indicating whether such forms were or were not borrowings and providing, in case they were, the word of origin, its meaning, the language of provenance, as well as other data about the circumstances of the transfer process. They could also individually add other concepts, but these were not considered in the statistical calculations.

In the *LWT* project, a word, compound, or phraseological expression is considered a borrowing in a given language when it has been taken from another (including a substrate language) at any point in its history. This definition considers that proto-languages are phases of the same language ([Haspelmath and Tadmor 2009a, p. 12](#)). It should be noted that only established borrowings incorporated into the lexicon of the recipient language are included, thus excluding nonce borrowings ([Haspelmath and Tadmor 2009a, p. 12](#)). Furthermore, the *LWT* project only considers direct lexical borrowings, excluding other types of linguistic transfers such as semantic and structural calques ([Haspelmath and Tadmor 2009a, p. 14](#)).

In the *LWT* project, the researchers involved had to choose among five options when qualifying the expressions associated with each concept: “no evidence for borrowing” (0), “very little evidence for borrowing” (0.25), “perhaps borrowed” (0.5), “probably borrowed” (0.75), and “clearly borrowed” (1). Each of these options was assigned the numerical values indicated in parentheses. Based on these values, the *borrowed score* was calculated for each concept, which therefore ranged from 0 to 1 and could be expressed as a percentage.⁶ The average borrowing rate for the set of 1460 concepts was 24.2% (Tadmor 2009, p. 55). The main findings of the *LWT* project can be consulted in Tadmor (2009) and on the *WOLD* website (Haspelmath and Tadmor 2009b).

3.2. *The Atlas Lingüístico Galego*

The *ALGa* project was initiated at the Instituto da Lingua Galega in 1974 under the direction of Constantino García and Antón Santamarina. Surveys were conducted by three researchers (Rosario Álvarez, Francisco Fernández Rei, and Manuel González González) between 1974 and 1977 at 167 points (152 localities in Galicia, 7 in Asturias, and 8 in the provinces of León and Zamora). The questionnaire contained 2711 questions covering 148 phonetic, 240 morphological, 139 syntactic, and 2184 lexical aspects, distributed among 17 thematic areas.⁷ The majority of informants adhered to the *NORM* pattern: speakers from rural areas with a low level of education who had lived most of their lives in their birthplace, coinciding with the survey point (Sousa 2017, pp. 322–25). It is, therefore, a traditional linguistic atlas whose data correspond to oral vernacular varieties. The language used in the interviews by the researchers (who were native fluent speakers) was always and only Galician.

Although the *ALGa* is currently in the process of being edited, most of the data are available in a restricted access database for project team members. So far, seven of the planned twelve volumes have been published. These include *Verbal Morphology* (1990), *Non-Verbal Morphology* (1995), *Phonetics* (1999), *Lexicon: Weather and Chronological Time* (2003), *Lexicon: The Human Being (I)* (2005), *Lexicon: Earth, Plants, and Trees* (2015), and *Lexicon: The Human Being (II)* (2020). The volumes have been published in print, although at the time of writing this work, a digital resource is being developed that will allow access to the information contained in them through a freely accessible website.

3.3. *Selection of Semantic Fields and Concepts from the LWT Project*

First, we carried out a selection from the 24 semantic fields used in the *LWT* project. This selection was driven by our interest in examining how thematic areas might influence loan processes, acknowledging that not all meanings from the *LWT* project could be included due to the constraints of this paper’s scope (Section 3.6). To manage scope and complexity, we decided to focus exclusively on the analysis of nouns,⁸ which have the highest presence in the *LWT* project (Haspelmath and Tadmor 2009a, p. 8). For this reason, we chose semantic fields with a minimum of 50 concepts with nominal expression, excluding the field ‘Modern World’, as most of the meanings contained therein have no entry in the *ALGa*. Consequently, we worked with concepts realized through nouns from seven semantic fields: ‘The physical world’, ‘Kinship’, ‘Animals’, ‘The body’, ‘Food and drink’, ‘Clothing and grooming’, and ‘Agriculture and vegetation’.⁹

3.4. *Correspondence between LWT Project Concepts and ALGa Questions*

For each of the concepts identified from the *LWT* project after the mentioned filtering process, we searched for correspondence entries in the *ALGa* survey questions. Thus, for example, the concept 3.817 ‘the ant’ from the *LWT* project corresponds to question 1377 “Formiga” ‘ant’ from the *ALGa*; or concept 1.72 ‘the wind’ matches with question 585 “Vento, aire” ‘wind’.¹⁰ However, it should be noted that we did not require total identification between the concept from the *LWT* project and the one from the *ALGa* to include them in our analysis. Thus, while the meaning ‘the weather’ is not listed in the Galician data source, there is question 634 “Bo tempo” ‘good weather’ in *ALGa*, which predominantly

includes the concept ‘weather’ in responses such as *bo tempo*, *bon tempo*, and *buen tiempo*. Accordingly, we included and analyzed the responses to this *ALGa* question, considering the expression of the noun and not the adjective.¹¹

We have prioritized *ALGa* questions focused on the lexical domain, but if the concept from the *LWT* project is not covered by them, we resort to questions about phonetic or grammatical aspects. For example, although the concept ‘the woman’ is not directly listed in the *ALGa*, question 469, which investigates the Galician translation of the Spanish phrase “La mujer se ha vuelto a casa” ‘the woman has returned home’ for grammatical study, provided us with the necessary data. We use the Galician equivalents for *mujer* ‘woman’ in the responses to this question to analyze the degree of Castilianization of the concept ‘the woman’. Similarly, the concepts ‘the milk’ and ‘the nut’ are asked in the phonetic questions 28 and 143, respectively, from which we extract the responses.

It should also be noted that we occasionally split a concept from the *LWT* project if it corresponded to more than one question in the *ALGa*. For example, for the meaning ‘the spring or well’, explained with the gloss ‘natural (spring) or artificial (well) source of water’ in the *LWT* project, we assigned two *ALGa* questions: 729 “Manancial, fontenla” ‘natural source’ and 899 “Fonte” ‘artificial source’.¹² Conversely, in a few cases, we merged multiple concepts from the *LWT* project because they corresponded to a single concept in Galician. This is the case of ‘the father-in-law (of a man)’ and ‘the father-in-law (of a woman)’: the responses are taken from question 2365 “Sogro” ‘father-in-law’, as the distinction between the two meanings is not lexicalized in Galician.

However, not all concepts with nominal expression from the seven selected semantic fields were subject to analysis, as some had to be discarded. In Section 3.6, we will explain the reasons for exclusion, but first, in Section 3.5, we will describe the treatment given to the concepts that were considered in the final analysis.

3.5. Analysis of the Responses Linked to Each Concept

3.5.1. Calculation and Annotation of Responses in Our Database

For each of the *ALGa* questions corresponding to the concepts from the *LWT* project that were selected, we calculated, on one hand, the number and percentage of responses that did not result from a process of borrowing from Castilian, and on the other hand, the number and percentage of responses that entered Galician from Spanish. The data from this last category were compared with the *borrowed score* of the concept in the *LWT* project. At this point, it is important to emphasize that only Castilian borrowings were counted (Section 3.5.2) while the *LWT* project considers borrowings from any language (including, for example, substrate languages).

We obtained responses to each question from an internal application used in the *ALGa* project, which records the forms provided by the informants for the different questions of the atlas. It should be noted that such responses might not exactly match the forms shown in the maps published in the various volumes, as these result from later editing work and might also combine responses from various questions (or even from other sources).

In most cases, Galician words and borrowings coexist. For example, in question 578 “Lúa” ‘the moon’, the Galician responses *lúa* and *llúa* are recorded alongside the Spanish *luna*, as shown in Table 1. However, there are cases where only Galician words are recorded, such as in question 1057 “Fariña” ‘the flour’, and others, less common, where 100% of the words came from Castilian, as in question 1888 “Misto” ‘the match’.

Table 1. Responses to question 578 “Lúa” ‘the moon’ in the *ALGa*.

Responses from the <i>ALGa</i>	Number of Responses
<i>Lúa</i>	76
<i>Llúa</i>	2
<i>Luna</i>	126
Total	204
Borrowing score	61.8%

Additionally, we often encounter different synonyms, quasi-synonyms, or words with similar meanings as responses to the same question in the *ALGa*, whether of Galician or Spanish origin. For example, in responses to question 1975, regarding the concept ‘the breakfast’, both the adaptations of the Castilian word *desayuno* and the Galician forms *almorzo*, *parva*, and *mañá* are recorded, as shown in Table 2. Similarly, for the concept ‘the centipede’ (question 1380 “Milpernas, cempés”), there are, among many other forms, *cempés*, *ciempiés*, *cempernas*, *cobra de cen patas*, *rapacarallas*, etc.

Table 2. Responses to question 1975 “Almorzo” ‘the breakfast’ in the *ALGa*.

Responses from the <i>ALGa</i>	Number of Responses
<i>almorzo</i>	144
<i>desaúno</i>	17
<i>parva</i>	11
<i>desalluno</i>	8
<i>desaiuno</i>	7
<i>mañá</i>	1
Total	188
Borrowing score	17%

Following the same criterion as for the *LWT* project, where multiple lexical counterparts are allowed for a single meaning in each language,¹³ we computed all these items. We distinguished between forms that entered from Spanish and those that did not, entering these data in a spreadsheet that formed our database, accessible at <https://zenodo.org/records/12583583> (accessed on 28 June 2024). In the first four columns (see Table 3), we provide information related to the *LWT* project as found in the *WOLD*. Column A specifies the semantic field associated with the concept; column B lists the *WOLD* code for that concept; column C gives the concept’s name in English as listed in the *WOLD*; and column D shows the borrowed score, which is the average probability that this concept is expressed through a loanword in the set of languages that provided data for that meaning in the *LWT* project, presented as a percentage to facilitate comparison with the data for Galician (field M in Table 4).

Table 3. Information extracted from *WOLD* entered into our database.

A	B	C	D
<i>WOLD</i>			
Food and drink	5.42	The breakfast	26%
Food and drink	5.55	The flour	35%

Table 4. Information extracted from *ALGa* entered into our database.

E	F	G	H	I	J	K	L	M	N
<i>ALGa</i>									
Almorzo	1975	188	almorzo (144), parva (11), mañá (1)	156	83%	desaúno (17), desalluno (8), desaiuno (7)	32	17%	
Fariña	1057	168	fariña (155), farina (4), faría (3), óleo (3), mestura (1), millaras (1), remillas (1)	168	100%	--	0	0%	<i>Farina</i> is considered a non-Spanish-origin response because it was collected at the points of Eastern Galician indicated in Álvarez de la Granja/Dubert (2024), Section 3.5.3

Columns E–N in Table 4, which is a continuation of Table 3, provide data from the *ALGa*. Column E lists the name and column F the number of the question in the atlas database. Column G indicates the total number of responses from the *ALGa* considered for that concept (Section 3.5.2); among these, Column H details the specific responses that did not enter from Castilian, noting (in parentheses) the number of occurrences for each form. Column I shows the total number of responses not transferred from Spanish, and column J the percentage they represent of the total. Columns K, L, and M provide analogous data for responses that are loanwords that entered from Castilian. Finally, column N includes observations.

It should be noted that, as can be seen in column G in Table 4, the number of total responses available may vary from concept to concept. This variation can be attributed to several factors: the potential removal of some responses as outlined in Section 3.5.2, the absence of responses at some of the surveyed points for certain meanings, or the variability at a single point where different responses might be recorded depending on the concept—ranging from Galician words and Castilianisms to synonyms and terms of varying specificity.

In our database, each of the seven semantic fields is labeled with the name of the semantic field followed by “_analyzed”. To the right of each of these sheets is another, documenting the concepts excluded for the reasons stated in Section 3.6, now appending “_excluded” to the semantic field name.

3.5.2. Criteria for Selection and Presentation of Responses

Not all entries listed in the *ALGa* were included in our database. Below, we explain the various criteria applied for selecting responses.

Firstly, we considered the responses provided in 166 out of the 167 survey points, as we excluded those corresponding to the municipality of Benuza, Le-5, where Leonese is spoken (Krüger 1965, p. 271).

Secondly, we excluded responses interpreted as free combinations of a head and a modifier rather than as lexicalized units. It should be noted that such interpretation is not always clear-cut. For example, for question 715 “Regato, rielo” ‘the river or stream’, we excluded responses like *río pequeno* ‘small river’; or for question 733 “Terreo pantanoso, boedo” ‘the swamp’, we excluded forms like *terreno pantanoso* or *terra pantanosa* ‘swampy land’. This criterion aligns with the one established in the *LWT* project:

Phrasal expressions were only to be given if they were fixed and conventionalized. Contributors were specifically asked not to provide descriptions or explanations of the meaning as counterparts. For example, for *LWT* 4.393 (the feather) a language may have had ‘hair of bird’ as the best equivalent, but if this was not

a fixed expression, it could not be used as the language's counterpart, and the entry should have been left unfilled (Haspelmath and Tadmor 2009a, p. 11).

Thirdly, we excluded forms that the *ALGa* notes explicitly indicated as being used only in toponymy, such as *Fervencedo* in question 723 "Abanqueiro, breixa, ruxidoira" 'the waterfall'.

Fourthly, some responses offer lexical forms that clearly diverge from the investigated concept. For example, in question 1810 "Abella" 'the bee', one of the responses is *avespa* 'wasp', and in question 1 "Aguia" 'the eagle', responses included *azor* 'goshawk', *garza* 'heron', or *gavilán* 'sparrow hawk'. We discarded such responses when we reasonably believed that they did not correspond to the concept in question. However, given the dialectal nature of the *ALGa*, deciding whether to exclude or include a term can be challenging, as lexical variation is common. For instance, a term for 'bean' might denote 'lentil' or 'chickpea' in different dialects. In any case, given that we are dealing with unique or minimally representative responses, they barely alter the results.

Finally, on account of our interest in exclusively studying nouns, responses that are not nouns, nominal compounds, or noun phrases are excluded from the analysis. For example, for the concept 'the waterfall', responses like *ruxir a auga* 'to roar the water' are excluded; for the concept 'the swamp', adjectives like *pantanosos* 'swampy' or *fangosos* 'muddy' are left out. Exceptions are made for cases where the noun can be extracted from the recorded expression. For example, in the concept 'the waterfall', the response *facen cachón* 'to make waterfall' is counted as an instance of *cachón* 'waterfall'.

Apart from these cases, we do not seek absolute concordance between the concept and the response, following the approach established in the LWT project:

Thus, in our project there was often less than complete identity between LWT meanings (labeled in English) and their counterparts in the various project languages. The semantic scope of the counterparts could be broader or narrower than that of the LWT meaning, or a more complex semantic relationship between them could obtain (Haspelmath and Tadmor 2009a, p. 9).

For example, for the concept 'the earring' under question 2329 "Pendientes" in the *ALGa*, we include *aros* or *aretas* 'hoop earrings', even though they denote specific types of earrings.

When multiple responses are collected at the same point sharing the same root, even if they present phonetic or morphological differences (such as *coiro* and *cuiro* collected in Vimianzo or *coiro* and *couracha* collected in Porto in question 53 "Coiro" 'the leather'), we count them as a single response, and in field H (Table 4), we only include the first form provided by the informants (in this case, *coiro* in both instances).¹⁴ However, exceptions are made for cases where learned and inherited forms of the same root coexist at the same point. For example, we counted two different responses, *esplanada* and *llanada*, for question 752 "Chaila, chaira" 'the plain' in Rodeiro. In cases of synonymous expressions with the same head, such as *pano* and *pano da nariz* in As Pontes de García Rodríguez or *pano do bolsillo* and *pano dos mocos* in Cerceda, from question 2331 "Pano de man" 'the handkerchief/rag', we count only one response.

This criterion applies unless the variants are Galician words and Castilian loanwords, in which case both are counted separately, such as *couro* and *cuero* in question 53 "Coiro" 'the leather' from Salvaterra do Miño, where the former is Galician, and the latter is Castilian. Similarly, if the responses collected at the same point have different roots, as is the case with the Galician words *coiro* and *pelello* in Boal in the same question, they are counted separately, and all are recorded in our database without distinguishing between first and subsequent responses.

Furthermore, in the case of several responses collected at the same point, where one is a generic term and the other is a sub-specification or there are different sub-specifications of the same concept, we count a single response whenever they share a root in the head, the general designator of the queried meaning. This happens, for example, with the responses *lagarto*, *lagartiño*, and *lagarto da agua* collected in Moaña in question 1392 "Lagarto" 'the

lizard’, where only the first response *lagartiño* is noted and counted. The same criterion is applied to responses like *paloma da casa*, *paloma do monte*, and *paloma rula* collected in Ribeira in question 1806 “Pomba” ‘the dove’: a single loan case is counted, and only *paloma da casa*, the first response, is noted and counted. Conversely, if the responses offered at the same point have different roots in the head, as is the case in Entrimo with *pomba* and *paloma brava*, we count and record all distinct forms.

3.5.3. Categorizing Responses as Spanish Transfers

Identifying borrowings from Castilian is a more complex task than might initially appear, and not all decisions made in this regard allow for absolute certainty. In any case, as we indicated, the data, including our classification, are available at <https://zenodo.org/records/12583583> (accessed on 28 June 2024) and can be accessed and modified by other researchers.

The LWT project has already addressed difficulties in identifying loanwords, requiring the language coordinators to choose among five options when qualifying words as transfers, each of which was assigned a certain score (Section 3.1): “clearly borrowed” (1), “probably borrowed” (0.75), “perhaps borrowed” (0.5), “very little evidence for borrowing” (0.25), and “no evidence for borrowing” (0). However, in our analysis, since we often deal with a large number of entries for each response, we do not work with this continuum because it would overly complicate this study and because the words that pose insurmountable doubts are relatively few and usually infrequent.¹⁵ Thus, the words collected in the ALGa must be classified necessarily into the group of forms transferred from Castilian or into the group of those that are not, although in some cases, as we indicated before, there may be doubts in this regard (for example, *aceña* ‘water mill’). Therefore, we categorize all entries strictly into two groups: those that are transfers from Castilian, equivalent to the WOLD’s “clearly borrowed” (counted as 1 point), and those that are not, aligning with “no evidence for borrowing” (counted as 0 points).

It must be emphasized once again that this work focuses exclusively on quantifying the influence of Spanish contact on Galician, which could be called *vertical influence*. Importantly, loanwords from Astur-Leonese, such as *coldo* ‘the elbow’ and *farina* ‘the flour’ recorded in eastern Galician, are counted alongside Galician forms (for the identification of Astur-Leonese words, the *Diccionario General de la Lengua Asturiana*, (García Arias 2024), was decisive). These loanwords, reflective of a dialect continuum rather than the imposition of a more prestigious language like Spanish, are predominantly observed in eastern Galicia and not generalized across the Galician-speaking territory, unlike Castilianisms. Forms that retain the Latin intervocalic -L-, the palatalization of intervocalic -LL- and the Astur-Leonese diphthongization of Ę/Ŏ in Asturias, as well as in the points of León Le-1 (Candín), Le-2 (Vilafranca do Bierzo), and Le-4 (Carracedelo), are particularly contentious. Furthermore, we do not consider the retention of the intervocalic -L- in points of Lugo like A Fonsagrada (L-16), Nogueira de Muñiz (L-19), and Navia de Suarna (L-23) as indicative of Castilian influence. Conversely, in the rest of the points, we interpret words with these traits as Castilianisms. It follows, then, that when classifying the data, we take into account the internal isoglosses of Galician, such that forms considered transfers in some points are not in others: for example, nine cases of *telar* ‘loom’ are classified as Castilianisms, while another nine collected in Asturias, in Le-4, and L-19, are not.

The historical marginalization of Galician distanced it from elevated spheres, where the interposition of Castilian (Section 2) further prevented direct contact of Galician with other languages. As a result, some loanwords from geographically distanced languages, such as the Italianism *pantano* ‘swamp’ or the Gallicism *pantalón* ‘trousers, pants’, are considered to have reached Galician through Castilian. For this reason, they are counted as direct borrowings from Spanish. This aligns with the criteria established in the LWT project, where the immediate source of loanwords is specified (Haspelmath and Tadmor 2009a, p. 16). The same applies to a few learned forms recorded in the ALGa (such as *mandíbula* ‘jaw’ or *pómulo* ‘cheekbone’), which probably also reached Galician through Castilian, as this was

for a long time the only language of culture. In any case, if any of these forms are found in medieval sources, we consider them Galician words (as is the case with *catarata* ‘waterfall’).

Following also the criterion established in the *LWT* project,¹⁶ we interpret not only the direct transfers from Spanish as loanwords, such as *conejo* ‘rabbit’, but also the words adapted from Spanish to Galician, such as *conexo*.¹⁷ Similarly, we consider as loans those words that, while not being direct transfers from Spanish or clear adaptations, seem to show influence from Castilian, often because they represent a mixture between Galician and Castilian forms, as is probably the case with the word *murciego* ‘bat’ recorded in Cambados.¹⁸

A different case is that of lexically or syntactically complex forms created in Galician, but which include Castilianisms. Using the same criterion employed in the *LWT* project,¹⁹ these forms are not counted in the group of Spanish loans. This is the case, for example, of the Galician term *vella cenando* ‘firefly’, which includes the Spanish loanword *cenando* ‘having dinner’—in contrast to the Galician *ceando*—or *blandoeira* ‘swamp’, which is built on the Castilianism *blando* ‘soft’ as opposed to the Galician *brando* or *mol*. An exception is made for complex expressions built with a Castilianism that was recorded as a response for the same concept. For example, for the meaning ‘the bull’, we have, among other responses, *toro* and *toro marón*. We count not only *toro* as a Castilianism but also *toro marón*, even though the latter expression could be a Galician creation. Conversely, neither semantic extensions of Castilianisms that occur in Galician, such as *sombrilla* ‘mushroom’, literally ‘parasol’, nor forms created from these semantic extensions, such as *sombrilla de sapo*, literally ‘parasol of toad’, are counted as loanwords.

3.6. Excluded Concepts

As indicated in Section 3.4, not all concepts with nominal expression from the seven selected semantic fields were included in the analysis. There are three different reasons for their exclusion. Details about the excluded concepts, along with the specific reasons for each exclusion, can be consulted at <https://zenodo.org/records/12583583> (accessed on 28 June 2024).²⁰

A first reason for exclusion is very simple: the concept is discarded because it does not appear in the *ALGa*, as is the case, for example, with meanings such as ‘the divorce’, ‘the linen’, or ‘the pepper’. The absence of concepts from the *LWT* project in the *ALGa* can be due to different reasons. On the one hand, it should be noted that the *ALGa* owes its debt to the research of the Romance world in which it is inserted. Thus, due to the comparative nature of this type of work, the *ALGa* authors took into account items that appeared in already published Romance atlases, and frequently omitted other questions absent in this tradition. It should be noted that among the concepts present in the *ALGa*, those linked to the traditional world predominate, and modern world realities are scarce (see note 7 for the thematic areas considered). On the other hand, concepts from the *LWT* project that designate realities foreign to Galicia (for example, animals like ‘the crocodile or alligator’ or foods like ‘the manioc bread’) or not lexicalized in Galician (such as ‘the married man’) are absent from the *ALGa*. Finally, many questions in the *ALGa* were formulated primarily to obtain phonetic or grammatical data rather than lexical information; the words that provide these data do not necessarily coincide with the concepts sought in the *LWT* project. The absence of a concept from the *ALGa* was the first criterion for exclusion considered, so if the concept is missing in the atlas, we mark it in our database as “Absent from the *ALGa*” and no further criteria are considered for its exclusion.

In addition, due to the proximity between Galician and Spanish, a considerable number of meanings are realized through cognates with the same form in both languages.²¹ In these cases, it does not make sense to analyze borrowing processes. These concepts are labeled with “Coincidence between the Galician and Spanish forms” in our database. This is the case of concepts like ‘the towel’, ‘the oil’, or ‘the arm’, which are realized in both languages as *toalla*, *aceite*, and *brazo*, respectively. We exclude the concept regardless of whether, in addition to the coincident forms, there might be some secondary Spanish loanwords in the *ALGa* responses. Thus, for example, the concept ‘the donkey’ is excluded, as the most

common words in both Spanish and Galician (and in the *ALGa*), *burro* and *asno*, coincide in both languages, although two examples of the Spanish loanword *pollino* (a less frequent term in Spanish than the indicated synonyms) also appear in the responses of the *ALGa*. Determining the percentage of Castilianization of ‘the donkey’ based on these two examples of *pollino* would be misleading, given the dominance of the cognate terms *burro* and *asno* in both languages. Consequently, the comparison with the percentage of Castilianization of other concepts where the most common words in Spanish do not coincide with those in Galician (for example, ‘the rabbit’: Spanish *conejo*, Galician *coello/coenllo*) is meaningless. Thus, those concepts whose most frequent expression in Spanish coincides with a Galician form are excluded,²² unless the *ALGa* data feature some Castilian loanword with a higher number of responses than the coinciding expression.

Obviously, we also exclude cases where terms common to both Galician and Spanish coexist with unique Galician terms not found in Spanish. This is the case of the concept ‘the storm’, where alongside the coincident terms *tormenta* or *tronada*, other Galician forms absent in Spanish, such as *treboada* or *trebón*, are recorded in the *ALGa*. In this regard, we must indicate that our analysis does not explore the phenomenon of contact typical of less prestigious languages, which consists of favoring synonyms coinciding with the prestigious language and avoiding dissimilar words, which may lead to the modification of their meaning or the reduction in their use (González Seoane 1994, p. 96; Kabatek 2000, pp. 31–33).

Finally, we excluded six concepts (‘the vulture’, ‘the skirt’, ‘the fruit’, ‘the tongue’, ‘the thumb’, and ‘the fox’), which account for 2.62% of the sample if included, due to the difficulties encountered in determining whether the forms recorded in the *ALGa*, coinciding with Spanish terms *buitre*, *falda*, *fruta*, *lengua*, *pulgar*, and *zorro*, should be regarded as loanwords from Castilian. Our doubts stem from the etymological origins of the word forms or the phonetic changes shared by Galician and Spanish. These forms are labeled “Difficulties in identifying Castilian loanwords” in our database. It should be noted that if these words are not transfers from Spanish, the corresponding concepts should be excluded due to their frequent use in both languages and their linguistic coincidence. We find this option preferable to their dubious interpretation as loanwords from Castilian.

4. Results

4.1. Analysis of *ALGa* Data

The total number of concepts from the *LWT* project analyzed in this study is 223. Table 5 collects the data grouped by semantic field in column A. Column B indicates the number of concepts in the *WOLD* for each field. Column C shows the number of concepts with nominal expression in this database, and column D shows how many of these are present in the *ALGa*. In column E, we indicate how many of them were analyzed once we excluded those whose main denomination in Spanish coincides with a Galician form (Section 3.6). Column F shows the total number of *LWT* concepts studied, once some were excluded due to the difficulties we encountered in determining whether the word coinciding with Spanish is or is not a Spanish loanword in Galician (Section 3.6). In the same column, separated by a slash, we also show the total number of *ALGa* questions considered. The figures do not always coincide due to the existence of some groupings and splits, as explained in Section 3.4: in ‘Kinship’, the concepts ‘the father-in-law (of a man)’ and ‘the father-in-law (of a woman)’ were unified into one, since the same expression is used in Galician in both cases. The same applies to ‘the mother-in-law (of a man)’ and ‘the mother-in-law (of a woman)’, ‘the son-in-law (of a man)’ and ‘the son-in-law (of a woman)’, and ‘the daughter-in-law (of a man)’ and ‘the daughter-in-law (of a woman)’. In ‘The physical world’, the concepts ‘the spring or well’ and ‘the stone or rock’ were split into two, resulting in a total of 31 Galician concepts and not 29. In ‘Animals’, the concept ‘the calf’ was split into two, based on the age of the animal, according to the questions established in the *ALGa*: “Tenreiro”, for calves less than a year old, and “Becerro, cuxo”, for those around two years old. Consequently, the total number of Galician is 49 and not 48. Finally, column

G indicates the percentage that the number of *LWT* concepts studied represents over the total number of *WOLD* meanings (column B) in each field. The fields are ordered from those with the lowest number of concepts analyzed to those with a higher number.

Table 5. Number of *WOLD* concepts and *ALGa* questions.

A	B	C	D	E	F	G
Semantic Field	<i>WOLD</i> Concepts	Nominal <i>WOLD</i> Concepts	<i>ALGa</i> Questions	Concepts with Different Expressions in Galician and Spanish	Final <i>WOLD</i> Concepts/ <i>ALGa</i> Questions Analyzed	% Final <i>WOLD</i> Concepts Analyzed
Clothing and grooming	59	54	26	21	20/20	33.9%
Kinship	85	71	36	25	25/21	29.4%
The physical world	75	68	37	29	29/31	38.7%
Food and drink	81	57	39	31	30/30	37%
Agriculture and vegetation	74	66	41	30	30/30	40.5%
The body	158	110	68	43	41/41	25.9%
Animals	116	114	66	50	48/49	41.4%
Total	648	540	307	229	223/222	34.4%

Table 6, along with overall data, presents detailed statistical metrics for each semantic field in the *ALGa*, including the average percentage of Castilianization, standard deviation, range, and the most frequently occurring percentage range. The fields are ordered from lowest to highest degree of Castilianization. As indicated, the percentage of Castilianization of a concept indicates the proportion of words or expressions recorded in the *ALGa* that have been borrowed from Castilian, relative to the total number of words or expressions registered in this database to express that concept.

Table 6. Data of Castilianization by semantic field.

Semantic Field	Percentage of Castilianization	Standard Deviation	Range	Most Repeated Percentage Range
Kinship	15.7%	27.3%	0–85.5%	0–10% (15/21)
Agriculture and vegetation	16.9%	24.6%	0–100%	0–10% (18/30)
Animals	19.1%	27.6%	0–98.4%	0–10% (30/49)
The physical world	24.5%	27.8%	0–100%	0–10% (14/31)
Clothing and grooming	28%	36.6%	0–99.4%	0–10% (12/20)
The body	29.2%	27.6%	0–100%	0–10% (15/41)
Food and drink	29.1%	30%	0–92.9	0–10% (12/30)
Total	23.3%	28.6%	0–100%	0–10% (116/222)

Table 6 shows that the average percentage of Castilian influence in the *ALGa* data is 23.3%, ranging from 0% to 100%, with a standard deviation of 28.6%. The most repeated

percentage range is 0–10%, with a mode at 0%, comprising 44 concepts without any loanwords (19.8% of the total meanings). It should be noted that the standard deviation is very high: it indicates a very high range of dispersion in the indices, rendering these averages only relatively meaningful.

Figure 1 presents the distribution of concepts by percentage ranges of borrowability, clearly illustrating how most meanings fall within the lower range, while the remaining ones offer fairly similar figures.

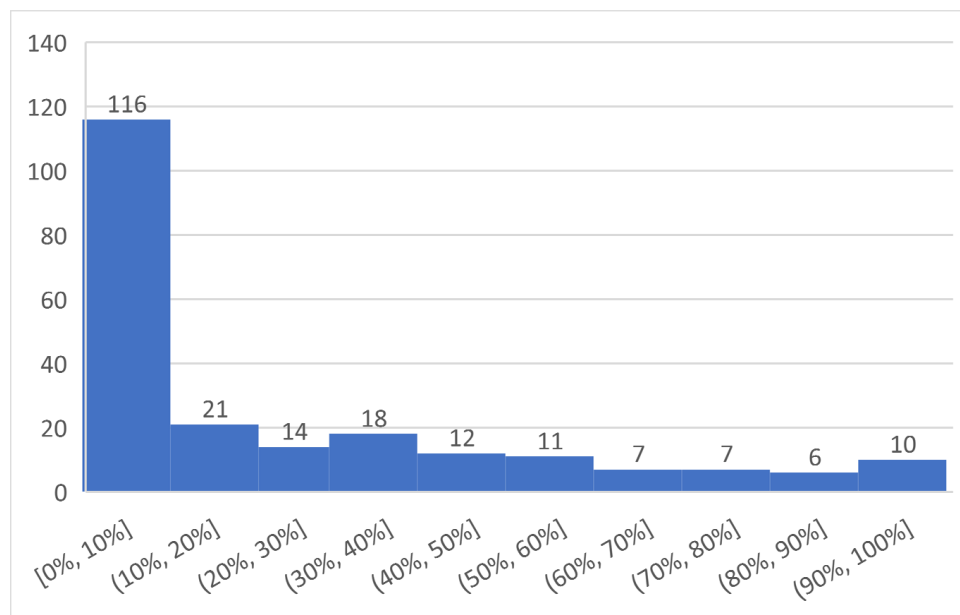


Figure 1. Distribution of concepts by percentage ranges of borrowability.

In the spreadsheet available on <https://zenodo.org/records/12583583> (accessed on 28 June 2024), the level of Castilianization for the different concepts studied can be consulted, organized by semantic fields.

4.2. Comparative Analysis with the WOLD

The average percentage of Castilianization in the *ALGa* concepts analyzed across the seven selected fields is only 0.8 points lower than the average percentage of loanwords for the same meanings in the *LWT* project. Table 7 compares the percentage of loanwords in the *WOLD* (considering only the analyzed concepts) and the percentage of Castilianization in the *ALGa*. The fields are ordered based on the differences found, starting from those with a higher presence of loanwords in the *WOLD* than in the *ALGa* to those where the opposite is true. We also provide data corresponding to the standard deviation as well as the percentage of loanwords in the total concepts of each field in the *WOLD* to determine to what extent the analyzed meanings are representative of the whole.²³

Table 7. Comparison between *WOLD* and *ALGa*.

Semantic Field	% of Loanwords in the <i>WOLD</i> (All Concepts in the Field)	% of Loanwords in the <i>WOLD</i> (Only Analyzed Concepts)	Standard Deviation in the <i>WOLD</i> (Only Analyzed Concepts)	% of Castilianization in the <i>ALGa</i>	Standard Deviation in the <i>ALGa</i>	Difference in Loanwords Percentage
Agriculture and vegetation	30.5%	28.8%	14.5%	16.9%	24.6%	11.9%
Clothing and grooming	37.8%	37.4%	19.2%	28%	36.6%	9.4%

Table 7. Cont.

Semantic Field	% of Loanwords in the WOLD (All Concepts in the Field)	% of Loanwords in the WOLD (Only Analyzed Concepts)	Standard Deviation in the WOLD (Only Analyzed Concepts)	% of Castilianization in the ALGa	Standard Deviation in the ALGa	Difference in Loanwords Percentage
Food and drink	30.8%	36.8%	21.3%	29.1%	30%	7.7%
Animals	27.9%	23%	11.5%	19.1%	27.6%	3.9%
Kinship	15.2%	18.2%	8%	15.7%	27.3%	2.5%
The physical world	21.5%	20.4%	11.9%	24.5%	27.8%	−4.1%
The body	15%	12.5%	6%	29.2%	27.6%	−16.7%
Average	23.9%	24.1%	15.8%	23.3%	28.6%	0.8%

Table 8 compares the position of each semantic field in the borrowability ranking of the ALGa with those from the LWT project, where 1st represents the field with the highest number of loanwords and 7th the field with the least transfers, considering the selected concepts. This comparison reveals that, in Galician, the first four fields are above the overall average of 23.3% of Castilianisms; the remaining three fields fall below this average.

Table 8. Ranking of semantic fields in the ALGa and WOLD according to the borrowability index.

Semantic Field	Position in the ALGa	Position in the LWT Project
The body	1st (29.2%)	7th (12.5%)
Food and drink	2nd (29.1%)	2nd (36.8%)
Clothing and grooming	3rd (28%)	1st (37.4%)
The physical world	4th (24.5%)	5th (20.4%)
Animals	5th (19.1%)	4th (23%)
Agriculture and vegetation	6th (16.9%)	3rd (28.8%)
Kinship	7th (15.7%)	6th (18.2%)

5. Discussion

5.1. Global Discussion

Firstly, Table 7 clearly shows the standard deviation indices in each semantic domain analyzed in the ALGa are very high, always higher than those in the WOLD domains. This indicates that in our sample there is a lot of internal dispersion in the Castilianization indices within all semantic fields. Furthermore, as shown in Table 6 and Figure 1, 116 out of the 222 lexical items analyzed (52.2%) have indices between 0 and 10%, with 47 items having a 0% index (21.2% of the total).²⁴ Conversely, only 41 out of 222 lexical items show a Castilianization level over 50% (18.5%). Therefore, only a minority of words significantly increase the Castilianization indices within each semantic field. In our opinion, this fact can only be interpreted as “each word has its own history”: with these data, we can conclude that it is not possible to assume uniform semantic factors across the board influencing Castilianization indices. Instead, the data suggest that semantic fields may not be the sole or most reliable predictor of Castilianization indices. Consequently, other factors will need to be considered to explain these variations fully. Table 6 reveals that the least Castilianized semantic fields in the ALGa, among the seven studied, are ‘Kinship’ (15.7%), ‘Agriculture and vegetation’ (16.9%), and ‘Animals’ (19.1%). These fields fall below the Castilianization average (23.3%). Conversely, the most Castilianized fields, which exceed the average, are ‘The physical world’ (24.5%), ‘Clothing and grooming’ (28%), ‘Food and drink’ (29.1%), and

'The body' (29.2%). According to Table 8, the four most Castilianized semantic fields do not present great percentage variations among them, with a range of just 4.7 points (from 24.5% to 29.2%). The last three fields also present similar indices, differing only 3.4 points. The contrast between the least and most Castilianized groups is more pronounced than the internal differences within these blocks.

The most remarkable aspect of Table 8 is that the final ranking of semantic fields in *ALGa* does not coincide with the ranking derived from *WOLD* data. However, in general, the differences between the degree of Castilianization of the thematic domains in the *ALGa* and the borrowed score of such domains in the *WOLD* are not very high, as they only exceed 10% in two cases: in 'Agriculture and vegetation', 11.9%, and in 'The body', 16.7%. Furthermore, in borrowability ranking by fields, the difference between each field in the *LWT* project and in the *ALGa* is always one or two positions, except precisely in the fields 'Agriculture and vegetation' (*ALGa*: 6th/*LWT*: 3rd) and 'The body' (*ALGa*: 1st/*LWT*: 7th), with a difference of three and six positions, respectively. In any case, the data in Table 8 indicate that the pattern is not consistent across the two databases. This is to be expected given that we are comparing changes that occur between languages with sometimes quite different social and linguistic backgrounds and that the data from the *LWT* project represent averages of very different situations, evidenced by a standard deviation of 15.8% (28.6% in Galician). This is an indication that perhaps it is not always possible to explain the data in terms of general linguistics. Furthermore, given the connections between the vocabulary and the cultural universe of each community, community-focused explanations could be necessary rather than broad linguistic analyses.

It should be noted that the *LWT* project includes loanwords received by a language from any other, across various historical periods (including loanwords taken from a substrate language), and considers proto-languages phases of the same language (Section 3.1). Thus, the *LWT* project adopts a long-term temporal perspective. In contrast, in our analysis, we focus exclusively on transfers from Castilian to Galician. Given this broader scope of the *LWT* project, it might be expected that the borrowed score should be higher than in Galician. Indeed, this occurs in all semantic fields, except for 'The physical world', where the *ALGa* is 4.1 points higher and, notably, in 'The body', where it is 16.7 points higher. It will always be possible to wonder how much the indices in the *ALGa* database would increase if we counted all documented loanwords from any language, but this is not the scope of this study. In any case, it is noteworthy that, considering that we are measuring only the presence of Castilianisms in Galician (and ignoring other sources, such as substrate languages, Arabisms, Germanisms, etc.) and the limited historical time they had to enter Galician (roughly from the Late Middle Ages to the end of the 20th century), the Castilianization indices are rather high.

Furthermore, there are many points in the *ALGa* where the presence of two responses is verified for the same concept: the Galician word and the Castilianism that replaces it. This is the case with concepts like 'the bowl', with 259 responses, of which 99 are Castilianisms (38.2%), and with several instances where the Galician word *cunca/conca* coexists with the Castilianism *taza*; 'the spoon', which has 231 responses with 112 (48.5%) Castilianisms and several instances where the Galician word *culler* or variants are recorded alongside the Castilian word *cuchara*; or 'the forehead', which has 275 responses with 140 (50.9%) cases of the Castilianism *frente*, which frequently coexist with the Galician word *testa*. In our view, this suggests that the lexical Castilianization observed could be a recent and quite rapid development, in line with the history of the social penetration of Castilian in Galicia (Section 2). In this context, it is noteworthy that in the semantic field 'The body', 14 out of 41 concepts have 200 or more responses (when the expected would be 166); the same happens with 'Food and drink': 8 out of 30 concepts have more than 200 responses; or with 'Agriculture and vegetation', 9 out of 30 did the same. These figures seem to show a very lively state of variation at the time the surveys were conducted. In any case, the analysis of the correlation between the number of responses per concept and the Castilianization index yields a Pearson coefficient of 0.051 (practically 0). This indicates an absence of correlation

between a high number of responses and a high index of Castilianisms: the dual responses may be attributed to the presence of Castilianisms coexisting with Galician forms or the coexistence of synonyms within Galician itself.

All these indices suggest that Galician speakers did not seem to hold purist ideas or taboos against the lexical Castilianization of their language at the time the *ALGa* survey was conducted. They also reflect the intensity of contact. And finally, the high index of lexical Castilianization is also a consequence of cultural and linguistic similarities. In the context of the *LWT* project, languages with different typological and sociocultural backgrounds may be in contact; in the case of the *ALGa*, the languages in contact are not only linguistically close, but also characteristic of practically identical cultures and societies. This situation greatly facilitates the penetration of Castilianisms into Galician.

Due to space constraints, we cannot provide a more detailed analysis of the seven selected semantic fields selected for this study. In the following pages, we will focus on the three fields that offer the most disparate or least predictable results compared to the *WOLD* data. These fields are 'The body' and 'The physical world', which, contrary to expectations, have a higher borrowability index in the *ALGa* than in the *LWT* project, and 'Agriculture and vegetation', which after 'The body', presents the highest percentage differences with respect to the *WOLD* data.

5.2. The Body

Among the seven semantic fields selected for study, 'The body' has the lowest borrowed score in the *LWT* project, at 12.5%. It is also one of the fields with the lowest borrowability index across all 24 fields examined in this project, with only 'Sense perception', 'Spatial relations', and 'Miscellaneous function words' having a lower borrowed score. [Tadmor \(2009, p. 65\)](#) attributes the low presence of loanwords in this field to its universality and the lack of a need to borrow terms from another language:

The semantic fields at the other extreme, comprising those least amenable to borrowing, are no less interesting. They consist of concepts that are universal and shared by most human societies. Practically every language can be expected to have indigenous words for such concepts, and therefore has no need to borrow them. These fields consist of Sense perception, Spatial relations, The body, and Kinship, which have a borrowing rate of just 10–15%.

The existence of a basic vocabulary is a common assumption in linguistics, though not without controversy. It is asserted that "there exists a basic or core vocabulary which is universal and relatively culture-free, and thus is less subject to replacement than other kinds of vocabulary" ([Campbell 2020, p. 423](#); see also [Matras 2009, p. 166](#)). In line with this idea, the *LWT* project leaders compiled the Leipzig–Jakarta list of basic vocabulary, i.e., the vocabulary less susceptible to be expressed through borrowing ([Tadmor 2009, p. 68 ff.](#)). To this end, they took into account not only the unborrowed score of the different concepts but also the representation score (number of languages in the *LWT* sample that had at least one counterpart for this meaning), the simplicity score (the average simplicity of all words for this meaning, based on whether the expressions are analyzable, non-analyzable, or semantically analyzable), and the age score (the average age of all words for this meaning). A lower borrowability index, a greater presence in languages, greater simplicity, and older age contribute to the classification of vocabulary as basic. In this list of 100 items ([Tadmor 2009, pp. 69–71](#)), 25 elements corresponding to the semantic field 'The body' appear:

Body parts constitute the most prominent group. Items from the semantic field The body make up only about a tenth of all the items on the 1460-item *LWT* meaning list, but fully a quarter of the items on the Leipzig–Jakarta list of basic vocabulary. Most items represent external organs expected to be known to any normal speaker in any society: mouth, ear, nose, eye, arm/hand, leg/foot, and many others. ([Tadmor 2009, p. 71](#))

In the same line, Carling et al. (2019, p. 2) consider that body parts are part of the basic lexicon of a language. As such, all languages should have words for body part-related concepts and would not need to take them from others. Specifically, they are considered “notions inherent to the environment, which are normally not changed by cultural activity”, so they should have “a lower borrowability” than other notions “which imply adaptation and change.”

Contrary to these predictions, in the *ALGa* data, ‘The body’ has the highest number of transfers from Spanish (29.2%). It is followed by ‘Food and drink’ (29.1%, which was expected to be among those that accept more loanwords, by virtue of its relationship with cultural changes, Carling et al. 2019, p. 2). In 26 of the 41 body-related concepts in the *ALGa*, the percentage of loanwords from Spanish is higher than the borrowed score of the *WOLD*.

Furthermore, our analysis of ‘The body’ in the *ALGa* reveals that five out of the thirteen concepts listed in the Leipzig–Jakarta list have a borrowed score exceeding 25%.²⁵ In four of these cases, the borrowed score surpasses 50%, and notably, the concept of ‘the blood’ shows a 100% Castilianization rate across all sampled dialects. Table 9 lists these thirteen concepts from the Leipzig–Jakarta list, detailing the percentage of Castilianization in the *ALGa* and the borrowed score in the *WOLD*, ordered from the highest to the lowest degree of borrowing in Galician.

Table 9. Concepts from the semantic field ‘The body’ selected in our study and present in the Leipzig–Jakarta list of basic vocabulary.

Concept	Borrowed Score <i>ALGa</i>	Borrowed Score <i>WOLD</i>
the blood	100.0%	10%
the knee	84.4%	9%
the thigh	59.7%	9%
the skin or hide	58.3%	11%
the nose	28.1%	3%
the navel	15.7%	12%
the neck	4.8%	10%
the liver	1.2%	13%
the eye	0.6%	10%
the foot	0.0%	14%
the tooth	0.0%	12%
the hand	0.0%	15%
the ear	0.0%	10%

While the data from our study show some deviations, these are not entirely unprecedented: “Basic vocabulary is in general resistant to borrowing [...]. Of course, basic vocabulary can also be borrowed—though this happens less frequently so that its role as a safeguard against borrowing is not fool-proof” (Campbell 2020, p. 329). Indeed, 7 out of the 41 languages studied in the *LWT* project have a borrowed score higher than the degree of Castilianization observed in Galician in this field: Selice Romani (57.2%), Gurindji (42.4%), Romanian (39.2%), Saramaccan (35.2%), Thai (30.4%), Tarifiyt Berber (29.5%), and Japanese (29.2%).

The question that arises next is why Galician presents these high indices of Castilianization in the semantic field ‘The body’. Among the reasons for introducing loanwords into a language, two stand out: necessity, when speakers lack expression for a new concept and thus need to fill a lexical gap; and prestige, when speakers adopt an unnecessary loanword from the language used by the socially more powerful community to gain approval and social status (Carling et al. 2019, p. 1; Matras 2009, pp. 149–53). It is important to recognize

that, in the 1970s, Castilian was the prestigious language in the Galician-speaking territory (Monteagudo 1999). Consequently, Castilian served a donor language of loanwords for Galician, both on account of necessity and prestige. Under the necessity category, since Galician was not used in high-functioning societal roles (Section 2), lexical innovation primarily occurred through borrowing from Castilian, the closest language, rather than through the creation of new native terms. Regarding prestige, the dynamic typically follows that “a less prestigious language borrows from a more prestigious one” (Carling et al. 2019, p. 29).

Related to prestige and the social value of forms, other pragmatic factors play a role in shaping lexical choice, as indicated by Pattillo (2021) or Matras (2009): taboo, euphemism, register, connotations of words related to sex or genital organs. In relation to all these factors, it can be noted that sometimes the Galician name for a body part becomes specialized for animals, while the loanword is dedicated to a human body part (Negro Romero 2013, p. 229). This is the case with ‘the kidney’ or ‘the back’, for example, where the Galician form *ril* is used to designate the kidneys of animals and the Castilianism *riñón* is used for humans. Similarly, the Galician word *lombo* refers to the back of animals and the Castilian word *espalda* is used for humans. However, we cannot forget that in the ALGa questionnaire, many of these responses were collected indiscriminately for body parts of humans and animals. It would be interesting to know what would happen if they were asked separately for each.

The prestige factor alone is not sufficient to explain why, for example, ‘the thigh’ or ‘the skin’ are Castilianized, but ‘the foot’, ‘the hand’, ‘the eye’, or ‘the ear’ are not (although their borrowability indices are equal to or above 10% in the LWT project). It also does not account for the Castilianism *pelexo* ‘skin of an animal’ replacing the Galician form *pelello* (58.3%), considering that the concept, which has only an 11% borrowability index in the WOLD, refers to animals.

The frequency of use of a concept is indeed a commonly cited factor that can impact the likelihood of its denomination being borrowed (Pagel et al. 2007; Carling et al. 2019; Pattillo 2021). High usage levels seem to hinder lexical substitution. For example, the denominations of ‘the hand’, ‘the foot’, ‘the tooth’, and ‘the ear’ are not Castilianized, but ‘the molar’, ‘the elbow’, ‘the eyebrow’, and ‘the eyelid’ are. However, usage frequency does not seem to explain the Castilianization of ‘the blood’. The presence of synonyms can also play a role in facilitating the entry of loanwords. According to Pattillo (2021, p. 389), synonyms constitute an open door to loanwords (see also Vejdemo and Hörberg 2016). Carling et al. (2019) and Pattillo (2021) also mention the existence of other factors that we cannot explore here, such as word length, widespread use in phraseology, or the existence of metonymic or metaphorical extensions. Consider the Galician word *ollo* ‘eye’ for ‘attention’; *ollo da pechadura* ‘keyhole’; *ollo da agulla* ‘eye of the needle’, etc.).

Thus, everything seems to indicate that while prestige is probably at the root of all loanwords taking place in this field, there are likely several explanatory factors that interact to determine why only some concepts are replaced or displaced by Castilian loanwords. Finally, we must add that, in principle and in general, Galician and Castilian share a similar repertoire of concepts related to body parts, which facilitates transference. This shared conceptual framework means that both languages segment and categorize body parts in comparable ways—a stark contrast to languages like Russian, where a single term might encompass what in Galician and Castilian are two distinct concepts (e.g., the Russian word for ‘foot’ and ‘leg’ is the same).

5.3. Agriculture and Vegetation

According to Carling et al. (2019, p. 18), this semantic field should be more open to loanwords because it belongs to the cultural framework, an area that changes over time and is susceptible to external influences due to the introduction of new crops, techniques, tools, uses or forms of ownership, etc. These concepts are, unlike those related to body parts, prone to “adaptation and change”, making their lexical evolution more a matter of necessity than prestige. However, precisely for this reason, this part of the lexicon “is more

complex to investigate”, since changing socio-historical conditions must be considered. For example, there may be several creative strategies for naming new realities other than borrowing as languages can repurpose existing words or create new expressions to describe novel concepts. In Galician, this has been the case, for example, with ‘corn’: when this cereal entered Galicia, it received two dialectal naming strategies (Álvarez 2002). In the first dialectal strategy, the word *millo*, firstly used to refer to ‘millet’, came to be used for the new concept of ‘corn’; as a result, the name of the old ‘millet’ changed from *millo* (now ‘corn’) to *millo miúdo* ‘millet’, literally ‘small millet’. In the second, the form *maíz* ‘corn’ was taken, imported from Amerindian languages through Castilian.

Therefore, concepts that are more intertwined with cultural activities, such as grazing, livestock farming, and agriculture, are more susceptible to borrowing due to their dependence on specific cultural interactions. However, we see that in the case of Galician in relation to Castilian, this trend does not seem to be fulfilled, as the field ‘Agriculture and vegetation’ is one of the most conservative and admits a relatively low level of Castilian loanwords (16.9%), despite the importance of temporary immigration of Galician peasants to Castilla in the Modern Age. In fact, 25 out of the 30 concepts analyzed in the *ALGa* from this field have borrowing indices below those observed in the *WOLD* (28.8%). Many of the concepts with zero borrowing refer to natural entities, such as ‘the oak’ (29% loan in the *WOLD*), ‘the grass’ (17% in the *WOLD*), ‘the hay’ (15% in the *WOLD*), or ‘the leaf’ (10% in the *WOLD*). However, a cultural object, such as ‘the sickle’, has 0.4% in the *ALGa* and 52% in the *WOLD*; also, ‘the threshing-floor’ has 0% in the *ALGa* and 18% in the *WOLD*.

In any case, the concepts in this field can be classified according to Carling et al. (2019, p. 6) based on the “intensity of cultural involvement and labor”. For them, the likelihood of borrowing increases according to the “increased labor intensity”. Thus, they consider three possibilities: “(1) Indoor, garden and small-scale farming zone, (2) Large farming zone, and (3) Technology and industry zone”. According to this framework, a farming culture takes fewer loanwords than an industrial culture, and small-scale farming takes fewer loanwords than large-scale farming. Regarding the settlement sphere, words used within the family or the indoor environment take fewer loanwords than those used in the outdoor environment: “It appears that cognitive proximity, familiarity, routine, intimacy, and frequency are all features that may compete with social motivations to borrow word-forms” (Matras 2009, p. 171).

The agricultural context in Galicia during the time of the *ALGa* surveys, characterized by subsistence farming and self-consumption rather than market-oriented agriculture (Villares 2016), undoubtedly influenced the linguistic landscape; or at least, that was the way of life represented by the NORMs selected as informants (Section 3.2). This indoor, garden, and small-scale farming system seems to have promoted linguistic conservatism in Galician data, both in the field of vegetation (such as plant names) and in that of cultural entities (such as tools). And yet it is still difficult to explain why ‘the hoe’ has a loanword index of 10% in the *ALGa* and 29% in the *WOLD*, or why ‘the pitchfork’ has 28% in the *ALGa* and 41% in the *WOLD*, or, conversely, why ‘the harvest’ has 44% in the *ALGa* and only 27% in the *WOLD*. Once again, other explanatory factors seem necessary.

5.4. The Physical World

As previously noted, the field ‘The physical world’ presents a borrowing score of 24.5% in the *ALGa* data, with a standard deviation of 27.8%. In contrast, the *LWT* project data present a borrowing index of 20.4% with a standard deviation of 11.9%. Therefore, there are higher borrowing rates in the *ALGa* and greater dispersion.

In this field, the highest rate of Castilianization is found in the concept ‘the match’, with 100%. This is not surprising considering that it is a relatively recent cultural object and not a natural entity. It is, therefore, a Castilianism born out of necessity. It is illustrative that in Portuguese it is called *fósforo*, through a Latinism (which the *ALGa* does not record).

However, unnecessary Castilianisms referring to natural, non-cultural objects abound in the *ALGa* data: for the expression ‘the dust’ there is an 88% Castilianization rate, com-

pared to 23% in the *WOLD* data. Similarly, the concept ‘the moon’ is Castilianized at 62% versus 28% in the *WOLD*. The ‘fog’ and ‘the hill’ also show higher Castilianization rates in the *ALGa* at 35% and 26%, respectively, compared to 14% and 17% in the *WOLD*. In contrast, ‘the snow’ and ‘the weather’ have minimal Castilian influence in the *ALGa*, both at 1%, despite higher rates in the *WOLD* of 24% and 36%, respectively. Much of this lexicon refers to common natural elements and are closely related on a semantical level: ‘the fire’ has only 1% of Castilianization, ‘the ash’, 5%, while ‘the flame’ has 38%. ‘The sky’ is at 6%, but ‘the star’ is at 59%. Surprisingly, ‘the valley’—another natural but outdoor element—has a 55% Castilianization rate, which might be more understandable given its exposure to external influences.

It seems evident that the entry of many of these loanwords is explained by prestige, and in some cases, factors like frequency of use: the concept ‘the flame’ can be expressed in Galician through *chama, lapa, labarada, larada*. . . , while ‘the fire’ is only expressed by *lume*.²⁶ The presence of so many synonyms for ‘the flame’ could indicate an openness of the concept to receiving loans (Section 5.2). In any case, once again, it seems necessary to resort to different explanatory factors, linked to specific expressions, and not to the general content of the field.

6. Conclusions

The main objective of this study was to assess the level of correspondence between the borrowability indices attributed to a selection of concepts from the *LWT* project and the indices obtained for those same concepts in the *ALGa* data. We also aimed to compare the degree of Castilianization in the Galician language across the different semantic fields studied. After briefly describing the sociohistorical situation of Galician, the projects we relied on, and the methodology, we presented and discussed the main results.

One of the findings reveals relatively high rates of Castilianization in the *ALGa* data. This demonstrates the strength of the contact between Galician and Castilian, the higher social prestige of Spanish, the cultural similarities among the various speech communities, and the proximity of the linguistic varieties.

Although the difference in the borrowing index between the overall averages obtained in the *WOLD* and the *ALGa* is only 0.8 points, a detailed analysis showed that the rankings of loanwords in the *LWT* project do not align perfectly with those of the *ALGa*. This variation is to be expected considering that each speech community experiences and is influenced by different historical and linguistic circumstances.

In the *ALGa*, the standard deviation of the means of the Castilianization indices for each concept is higher than in the *LWT* project. This indicates that the range of variation in the values are higher in Galician; in fact, in each semantic field, many concepts (52.2%) have low Castilianization indices, between 0% and 10%, and only a few concepts have high indices. From this, we can conclude that there are no universal semantic factors (semantic fields) that definitively condition the Castilianization indices.

The more detailed analysis of three semantic fields (‘The body’, ‘Agriculture and vegetation’, and ‘The physical world’), which yielded somewhat unpredictable results, shows that the influential factors are the prestige of Castilian, the degree of urbanization, modernity within the cultural universe of the semantic field, and perhaps the frequency of word usage: in general, with some exceptions, concepts that are used more frequently seem to be more resistant to borrowing. In any case, the reader should consider that the kind of informants surveyed in the *ALGa* are NORMS. The sociolinguistic evolution of the region since the mid-seventies of the last century has surely influenced the degree of lexical Castilianization in Galician.

Moreover, other intralinguistic factors that, according to some scholars, may influence the reception of loanwords remain to be explored: word length, phonological similarity, connotative values, the existence of synonyms, etc. Further studies are needed that employ individualized, quantitative, and qualitative analysis of concepts and their expressions to shed some light on these facts.

Author Contributions: Conceptualization, methodology, validation, formal analysis, investigation, writing—original draft, writing—review and editing, visualization: M.Á.d.I.G. and F.D.G. All authors have read and agreed to the published version of the manuscript.

Funding: This research was funded by the Ministerio de Cultura (Government of Spain).

Institutional Review Board Statement: Not applicable.

Informed Consent Statement: Not applicable.

Data Availability Statement: The original contributions presented in the study are included in the article. Further inquiries can be directed to the corresponding author/s.

Conflicts of Interest: The authors declare no conflicts of interest.

Notes

- 1 From now on, and with the aim of simplification, we will use the terms Spanish or Castilian interchangeably.
- 2 Borrowability can be defined as “The relative likelihood that words with particular meanings would be borrowed” (Haspelmath and Tadmor 2009a, p. 1).
- 3 See, respectively, <https://wold.cldd.org/> (accessed on 28 June 2024) and <https://ilg.usc.es/gl/proxectos/atlas-linguistico-galego-alga> (accessed on 28 June 2024).
- 4 According to the *Mapa Sociolingüístico de Galicia* (the study temporally closest to the surveys of the ALGa that provides information on the linguistic acquisition and usage in Galicia), in 1992, 62.4% of people had Galician as their *initial language*; 25.6% had Spanish; 11.4% had both languages; and 0.6% were in other situations (Fernández Rodríguez and Rodríguez Neira 1994, p. 39). Regarding the *habitual language*, in 1992, 38.7% spoke only Galician; 29.9% spoke more Galician than Spanish; 10.6% spoke only Spanish; and 20.8% spoke more Spanish than Galician (Fernández Rodríguez and Rodríguez Neira 1995, p. 49). A comparison with the data from the latest survey by the Instituto Galego de Estatística (2019), conducted in 2018, shows a significant setback in the intergenerational transmission of Galician and a considerable increase in the use of Spanish: 42.19% of the Galician population indicates that Galician is their *initial language*; 31.72% indicate Spanish; 23.68% indicate both; and 2.41% indicate other situations. Regarding the *habitual language*, 30.33% say they always speak Galician; 21.55% speak more Galician than Spanish; 23.14% speak more Spanish than Galician; 24.71% always speak in Spanish; and 0.77% indicate other situations.
- 5 Archi, Bezhta, Ceq Wong, Dutch, English, Gawwada, Gurindji, Hausa, Hawaiian, Hup, Imbabura Quechua, Indonesian, Iraqw, Japanese, Kali’na, Kanuri, Ket, Kildin Saami, Lower Sorbian, Malagasy, Manange, Mandarin Chinese, Mapudungun, Old High German, Oroqen, Otomi, Q’eqchi’, Romanian, Sakha, Saramaccan, Selice Romani, Seychelles Creole, Swahili, Takia, Thai, Tarifit Berber, Vietnamese, White Hmong, Yaqui, Wichí, and Zinacantán Tzotzil (Haspelmath and Tadmor 2009a, pp. 3–4).
- 6 Tadmor (2009, p. 66) talks about “unborrowed score” and the numerical assignment is reversed (from 1 to “no evidence for borrowing” to 0 for “clearly borrowed”). However, we use the “borrowed score” listed in the WOLD.
- 7 1. The time and weather, 2. Topographical accidents, 3. Agriculture, 4. Wine, oil, flour, bread, wool and linen, 5. Plants, 6. Insects, birds, wild animals, 7. Fishing and hunting, 8. Pastoral life, 9. Domestic animals, 10. The home. Domestic occupations, 11. The human body. Movements and actions, 12. Dress and footwear, 13. The family. Human life, 14. The spiritual world, 15. Games and amusements, 16. Trades, 17. Weights and measures (García et al. 1977, p. 14).
- 8 Some research establishes a relationship between the borrowability index and grammatical category (Haspelmath 2009, p. 35).
- 9 We leave out the concepts added individually by those responsible for some languages of the LWT project, which, as indicated in Section 3.1, were not considered in the statistical calculations of this project.
- 10 The concepts of the LWT project are written between single quotation marks and with an article, as they appear in WOLD, while the semantic fields are also written between single quotation marks, but with an initial capital letter, also according to WOLD’s presentation. To keep them differentiated, the questions from ALGa are presented without an article and between double quotation marks, while our English glosses of Galician words are written between single quotation marks, with a lowercase initial letter and without an article.
- 11 Question 635 “Mal tempo” ‘bad weather’ is also included in the ALGa, but we chose “Bo tempo” because it has a slightly higher number of responses. In any case, the percentage of Castilianization is similar in the responses to both questions.
- 12 We excluded question 897 of the ALGa “Pozo” ‘well’ because its expression is mostly identical in Galician and Spanish (see Section 3.6 for more information on this issue).
- 13 “The World Loanword Database addresses this problem by allowing several (indeed, an unlimited number of) counterparts per meaning” (Haspelmath and Tadmor 2009a, p. 9).
- 14 The ranking of responses is established in the ALGa through letters (A, B, C. . .).
- 15 As we will discuss in Section 3.6, in a few highly ambiguous cases where the classification of a form as a loanword significantly influenced the results, we opted to exclude the concepts from the analysis.

- 16 “Borrowing a word often entails a certain modification of the source word, required for the integration of the word into the recipient language” (Haspelmath and Tadmor 2009a, p. 16).
- 17 Except as a possible realization of the phenomenon known as “gheada” (Fernández Rei 1990, pp. 163–89), Galician does not have the voiceless velar fricative sound [x], represented by <j> in *conejo* ‘rabbit’. However, there are many cognates where the Castilian [x] corresponds to the Galician [j], represented by <x> (*gente/xente, cojo/coxo*, etc.). For this reason, in some instances, *conejo* has been adapted in Galician as *conexo*.
- 18 Since it coexists at this point with *murciélago* and also with *murciégalo*, response A, only the latter form is computed and recorded in our database.
- 19 “Excluded from the class of loanwords are neologisms (=productively created lexemes) which consist partly or entirely of foreign material, because they are created in the recipient language, and not transferred from a donor language” (Haspelmath and Tadmor 2009a, p. 13).
- 20 The concepts that do not have a nominal character were also included in the sheets labeled as “_excluded” accompanied by their part of speech tag.
- 21 Phonetic phenomena such as *gheada* or *seseo* (Fernández Rei 1990), which do not have lexical restrictions, are set aside.
- 22 The most common expression in Spanish is determined through corpora in those cases where there are both different and coincident synonyms with Galician.
- 23 The data for all WOLD concepts presented in the second column were extracted from the website itself, selecting only those meanings marked as “True”. These were the only ones considered for statistical calculations in this project. Note, however, that the overall results provided on the project website include all concepts, both those marked with “True” and “False”, and therefore do not match with our findings. Conversely, the data we provide does match with those presented in Table 6 of Tadmor (2009, p. 64), with minor differences likely due to decimal adjustments.
- 24 In 44 cases, the percentage is 0%, and in three cases, it is 0.4%. These are added together since only the whole number is considered for analysis.
- 25 Twelve concepts from the semantic field ‘The body’ listed in this inventory were not analyzed in our study due to their absence from the ALGa or because they coincide with Spanish.
- 26 In the ALGa, *fogo/fougo* is recorded, but only in six locations.

References

- Álvarez, Rosario. 2002. Viño novo en odres vellos: Os nomes do millo. In *Dialectoloxía e Léxico*. Edited by Rosario Álvarez, Francisco Dubert García and Xulio Sousa. Santiago de Compostela: Consello da Cultura Galega, Instituto da Lingua Galega, pp. 69–94. Available online: <http://hdl.handle.net/10347/9769> (accessed on 5 April 2024).
- Boullón Agrelo, Ana Isabel, and Henrique Monteagudo. 2009. *De Verbo a Verbo: Documentos en Galego Anteriores a 1260*. Santiago de Compostela: Universidade de Santiago de Compostela. [CrossRef]
- Campbell, Lyle. 2020. *Historical Linguistics. An Introduction*, 4th ed. Cambridge: The MIT Press. [CrossRef]
- Carling, Gerd, Sandra Cronham, Robert Farren, Elnur Aliyev, and Johan Frid. 2019. The causality of borrowing: Lexical loans in Eurasian languages. *PLoS ONE* 14: e0223588. [CrossRef] [PubMed]
- Dubert García, Francisco. 2005. Interferencias del castellano en el gallego popular. *Bulletin of Hispanic Studies* 82: 271–91. [CrossRef]
- Fernández Rei, Francisco. 1990. *Dialectoloxía da Lingua Galega*. Vigo: Xerais.
- Fernández Rodríguez, Mauro A., and Modesto A. Rodríguez Neira. 1994. *Lingua Inicial e Competencia Lingüística en Galicia*. A Coruña: Real Academia Galega.
- Fernández Rodríguez, Mauro A., and Modesto A. Rodríguez Neira. 1995. *Usos Lingüísticos en Galicia. Compendio do II volume do Mapa Sociolingüístico de Galicia*. A Coruña: Real Academia Galega.
- García Arias, Xosé Lluís. 2024. *Diccionario General de la Lengua Asturiana*. Available online: <http://mas.lne.es/diccionario/p/introduccion> (accessed on 17 March 2024).
- García, Constantino, and Antón Santamarina, dirs. 1990. *Atlas Lingüístico Galego*. A Coruña: Fundación Pedro Barrié de la Maza, Conde de Fenosa. Santiago de Compostela: Universidade de Santiago de Compostela, vol. 7.
- García, Constantino, Antón Santamarina Fernández, Rosario Álvarez Blanco, Francisco Fernández Rei, and Manuel González González. 1977. O Atlas Lingüístico Galego. *Verba* 4: 5–17.
- González Seoane, Ernesto. 1994. Variedade e empobrecemento do léxico. *Cadernos de Lingua* 10: 89–102. [CrossRef]
- Haspelmath, Martin. 2009. Lexical borrowing: Concepts and issues. In *Loanwords in the World’s Languages: A Comparative Handbook*. Edited by Martin Haspelmath and Uri Tadmor. Berlin: Walter de Gruyter, pp. 35–54. [CrossRef]
- Haspelmath, Martin, and Uri Tadmor. 2009a. The Loanword Typology project and the World Loanword Database. In *Loanwords in the World’s Languages: A Comparative Handbook*. Edited by Martin Haspelmath and Uri Tadmor. Berlin: Walter de Gruyter, pp. 1–34. [CrossRef]
- Haspelmath, Martin, and Uri Tadmor, eds. 2009b. *World Loanword Database*. Leipzig: Max Planck Institute for Evolutionary Anthropology. Available online: <http://wold.cldd.org> (accessed on 21 March 2024).

- Instituto Galego de Estatística. 2019. *Enquisa Estrutural a Fogares. Coñecemento e Uso do Galego*. Santiago de Compostela: Instituto Galego de Estatística. Available online: https://www.ige.eu/web/mostrar_actividade_estadistica.jsp?idioma=gl&codigo=0206004 (accessed on 5 April 2024).
- Kabatek, Johannes. 2000. *Os Falantes como Lingüistas. Tradición, Innovación e Interferencias no Galego Actual*. Vigo: Xerais.
- Krüger, Fritz. 1965. Aportes a la fonética dialectal de Sanabria y de sus zonas colindantes. *Revista de Filología Española* XLVIII: 251–82. [CrossRef]
- Mariño Paz, Ramón. 2008. *Historia de la Lengua Gallega*. Munich: Lincoln Europa.
- Matras, Yaron. 2009. *Language Contact*. Cambridge: Cambridge University Press. [CrossRef]
- Monteagudo, Henrique. 1999. *Historia Social da Lingua Galega*. Vigo: Galaxia.
- Monteagudo, Henrique. 2003. Sobre a norma léxica do galego culto: Da prosa ficcional de Nós ao ensaio de Galaxia. In *A Estandarización do Léxico*. Edited by María Álvarez de la Granja and Ernesto González Seoane. Santiago de Compostela: Consello da Cultura Galega, Instituto da Lingua Galega, pp. 197–254. Available online: <http://hdl.handle.net/10347/9819> (accessed on 5 April 2024).
- Monteagudo, Henrique, and Antón Santamarina. 1993. Galician and Castilian in contact: Historical, social, and linguistic aspects. In *Trends in Romance Linguistics and Philology. Volume 5: Bilingualism and Linguistic Conflict in Romance*. Edited by Rebecca Posner and John N. Green. Berlin and New York: De Gruyter Mouton, pp. 117–74. [CrossRef]
- Negro Romero, Marta. 2013. Contacto galego-castelán e cambio no léxico do corpo humano. In *Contacto de Linguas, Hibrididade, Cambio: Contextos, Procesos e Consecuencias*. Edited by Eva Gugenberger, Henrique Monteagudo and Gabriel Rei-Doval. Santiago de Compostela: Consello da Cultura Galega, Instituto da Lingua Galega, pp. 221–43. Available online: <http://hdl.handle.net/10347/9482> (accessed on 5 April 2024).
- Pagel, Mark, Quentin D. Atkinson, and Andrew Meade. 2007. Frequency of word-use predicts rates of lexical evolution throughout Indo-European history. *Nature* 449: 717–20. [CrossRef] [PubMed]
- Pattillo, Kelsie. 2021. On the borrowability of body parts. *Journal of Language Contact* 14: 369–402. [CrossRef]
- Ramallo, Fernando, and Gabriel Rei-Doval. 2015. The standardization of Galician. *Sociolingüística* 29: 61–81. [CrossRef]
- Rodríguez Lorenzo, David. 2022. Variación e cambio lingüístico en tempo real. Un estudo sobre o galego con base en materiais xeolingüísticos. Ph.D. thesis, Universidade de Santiago de Compostela, Santiago de Compostela, Spain. Available online: <http://hdl.handle.net/10347/29352> (accessed on 5 April 2024).
- Sousa, Xulio. 2017. Documenting and mapping geolinguistic variation: The linguistic database of the Atlas Lingüístico Galego. In *Gotzon Aurrekoetxea Lagunarterik Hara*. Edited by Aitor Iglesias Chaves and Ariane Ensunza Aldamizetxebarria. Bilbao: Universidad del País Vasco/Euskal Herriko Unibertsitatea, pp. 321–36. Available online: <http://hdl.handle.net/10347/15575> (accessed on 5 April 2024).
- Sousa, Xulio, and Francisco Dubert García. 2020. Measuring language contact in geographical space: Spanish loanwords in Galician. *Zeitschrift für Dialektologie und Linguistik* 87: 285–306. [CrossRef]
- Tadmor, Uri. 2009. Loanwords in the world's languages: Findings and results. In *Loanwords in the World's Languages: A Comparative Handbook*. Edited by Martin Haspelmath and Uri Tadmor. Berlin: Walter de Gruyter, pp. 55–75. [CrossRef]
- Vejdemo, Susanne, and Thomas Hörberg. 2016. Semantic factors predict the rate of lexical replacement of content words. *PLoS ONE* 11: e0147924. [CrossRef] [PubMed]
- Villares, Ramón. 2016. *Historia de Galicia*. Vigo: Galaxia.

Disclaimer/Publisher's Note: The statements, opinions and data contained in all publications are solely those of the individual author(s) and contributor(s) and not of MDPI and/or the editor(s). MDPI and/or the editor(s) disclaim responsibility for any injury to people or property resulting from any ideas, methods, instructions or products referred to in the content.