THE ROLE OF LANGUAGE IN INTERNATIONAL TRADE: HOW DOES LANGUAGE AFFECT THE CHOICE OF FOREIGN TRADING PARTNERS?

Maria João Ferro
ISCAL and NOVA CLUNL
(mjferro@iscal.ipl.pt)

Sandra Ribeiro
Universidade Autónoma de Lisboa/OBSERVARE and ISCAL
(sribeiro@autonoma.pt)

ABSTRACT

In an increasingly global and networked economy, companies have a wider market at their disposal but also a number of barriers they have to overcome in order to expand their business internationally. When the decision of trading with another country is made, one of the key issues companies have to deal with is language: language barriers can hinder international trade, while language similarities can boost it due to the weight of communication costs. The economics of language is an interdisciplinary field of study that aims to bring together reflections on the role of language in the economy and conversely on how the economy affects language choices.

In this study we focus on the relationship between international trade and language barriers by analysing Portuguese exports of goods and find that language does play an important role in the choice of foreign trading partners although sharing Portuguese as the common language with a group of countries does not imply exporting more to those countries as one might expect. We therefore argue for the increase of multilingual policies in education and more investment in reaping the benefits that Portuguese as an international language might bring to our country.

Keywords: International Trade, Language Barriers, Multilingualism, Globalization.

JEL Classification: F140, F620.

1. INTRODUCTION

Globalization has been increasingly interconnecting people with a profound effect on the way companies do business. We know this is not an entirely new phenomenon - Friedman (2007), for example, goes as far as calling the phase from 2000 onwards 'globalization 3.0', considering that the difference in this new state of affairs lies in the power of the individual both in terms of global cooperation and global competition and not solely on the global nature of social interactions. However, the truth is that the world has become a global village (McLuhan, 1962) which thrives due to the construction of networks.

> Networks are becoming the nervous system of our society, and we can expect this infrastructure to have more influence on our entire social and personal lives than did the construction of roads for the transportation of goods and people in the past (van Dijk, 2006 [1991]: 2).

Van Dijk talks about 'information highways' and Castells (2001), who has thoroughly provided a theoretical framework for the network society, stated at the beginning of the 21st century that the Internet has become the very tissue of our lives and that the new information technologies underlay the global economy. Finding trading partners in distant parts of the world is now much easier, which has obviously opened up the global market.

Having a wider market, however, also implies that companies have a number of barriers they have to overcome in order to expand their business internationally. Trade costs can assume a variety of forms, such as transportation costs, currency conversion, customs costs, information acquisition, communication costs, and many others. In this paper, we focus on communication costs, more specifically on language barriers: while trading with a partner that shares the same language (or a very similar one) can boost trade because it decreases communication costs, trading with a partner with whom there is no shared language will imply some sort of intermediary (e.g. translators and interpreters) that will

forcibly increase those costs. Language skills are necessary from the first stage of internationalization of a company, starting with the search for trading partners, and all throughout the negotiation stages and even later, if contracts are breached and legal actions must be taken. Ease of communication between trading partners is therefore of the utmost importance for the establishment of the relationship and for the maintenance of the commercial ties. As Fidrmuc & Fidrmuc (2016: 32) put it,

> Speaking the same language facilitates communication and makes transactions easier and more transparent. In this, the effect of language is similar to that of common culture, legal norms or units of measurement: Engaging in mutually beneficial exchange is possible without them, but it is generally more costly and the outcome is less predictable.

Sharing a common language and the ease of communication it entails, although not indispensable, is therefore an important determinant to stimulate trade between two countries. Considering the current structure of society - namely the fact that people and companies alike try to establish profitable networks and many countries are 'eager to reap the benefits that such networks can bring, including trade- and investment-linked technological spillovers and stronger employment demand in manufacturing' (Arvis *et al.*, 2013: 3) - we posited two hypotheses that jumpstarted our study:

H1 Portuguese exports are higher to countries that share the same language.

H2 Portuguese exports are higher to countries that share a similar language.

Working within the field of the economics of language, we analysed trade data from the top 56 countries to which Portuguese companies export to test the two hypotheses. We found that, as expected, language is not the sole determinant to foreign trade and that although it does play a part in the choice of trading partners, the concepts of 'same language' and 'similar language' are not enough and foreign language acquisition must also be factored into this study.

We believe our study importantly contributes to the literature on Portuguese international trade by focusing on the relationship between Portuguese exports of goods and the language(s) spoken in the country of destination. Furthermore, we depart from commonly used language proximity indexes and base our language variable on etymological proximity, by classifying languages according to the linguistic family they belong. Although not entirely new - for example, Adserà and Pytliková (2015) use a combination of three indexes, including the level of proximity between two languages within the linguistic family tree to explore the importance of language in explaining international migration flows - our measure is based strictly on linguistic criteria and does not rely on purely statistical data, such as the widely used Levenshtein's distance (see Isphording and Otten, 2013) that calculates the minimum number of changes necessary (additions, deletions, and substitutions) to transform one word into another (Wichmann, Müller & Velupillai, 2010).

The remainder of the paper is organized as follows: section 2 provides a very brief overview of the theoretical framework of the economics of language; section 3 explores the concepts of official and common spoken languages; section 4 describes the method we have used; in section 5 we present and discuss the findings of our study; and section 6 concludes the paper and provides some future research avenues.

2. LANGUAGE AND THE ECONOMY

The relationship between economy and language, albeit not very obvious at first sight, was established in the 1960s by Jacob Marschak (1965) who coined the term 'economics of language'. For Marschak, language was a fundamental tool for men to carry out their economic activities, and hence its definition as a resource with all it entails. The economics of language fundamentally studies the interconnection between language and the economy: how one influences the other and vice-versa. Studies on the economics of language range from immigration, to foreign direct investment (FDI), and international bilateral trade.

Several authors have shown that proficiency in the language spoken in the country of destination is part of the human capital of each immigrant worker (Breton, 1978; Chiswick & Miller, 1995, 2014). Hence the investment in acquiring the linguistic capital of the host country will be profitable for immigrants and will consequently decrease the gap between their personal income and the income of a native worker with similar characteristics. Language is necessary not only to find jobs, but also to communicate with employers, co-workers and, depending on the job,

customers as well. Therefore, proficiency in the host country's language naturally influences the choice of destination, since it reduces the costs associated with migrating and increases the probability of success in the destination country's labour market. Moreover, based on a large-scale study using data for thirty OECD countries between 1980 and 2010, Adserà & Pytliková (2015) show that when immigrants do not master the language of the host country, the degree of proximity between their native tongue and the host country's will be a decisive factor when the time comes to select a destination. This means that, given the choice, people who take the decision to leave their country in search for a better life will tend to choose a host country whose language they master or is sufficiently similar to their own language for them to acquire it fairly easily. Portuguese migration flows clearly mirror this choice: until the 1970s, Portuguese workers mostly moved to France and Luxembourg (taking advantage of the similarity between Portuguese and French, two Romance languages, and in a time when French was the dominant foreign language learned in Portugal), while currently they are mainly directed to Brazil, the United Kingdom, and Switzerland (Marques, 2010; Peixoto, 2012). The current choice of host country is a good illustration of our findings that we will explore later on this paper: Portugal shares the same language with Brazil and consequently the move implies virtually no costs in terms of language acquisition; while English is currently (and has been for some time) the most widely studied foreign language in Portugal - according to Eurostat (2015), 34.9% of the Portuguese children learn English at primary level and 93.4% learn it at lower secondary level - which clearly facilitates the move to the country without a large initial investment in its linguistic capital.

Similarly, the choice of a country to invest is also influenced by language, among other factors. Focusing on four major languages - English, French, Spanish, and Arabic - Oh, Selmier & Lien (2011) show that a common language is determinant for FDI, in fact even more so than in international trade. Kim *et al.* (2015) also find a significant correlation between language and FDI, but they tie it not only with language as a means of communication, but also as a vehicle for culture, which had already been suggested by Guiso, Sapienza, & Zingales (2009) who explored cultural biases, using the commonality between two languages as a proxy for common culture. Lameli *et al.* (2014) also followed this route, testing dialect interference in the choice of trading partners in Germany, a country with only one *de jure* and *de facto* language - German - whose ancient dialectal differences still matter today in terms on domestic trade flows. The authors interpret this correlation as a measure of cultural ties, implying that language is an expression of culture.[196] Basing their study on Switzerland, a multilingual country with very specific characteristics, given that the speakers of the four official languages - German, French, Italian, and Romansh (by order of speakers) - are located within well-established boundaries and are usually proficient in the other languages besides their own native one, Egger and Lassmann (2015) correlate common native language with an expression of common culture and find that it has an effect on extensive margins of trade rather than on intensive ones.

While these three aspects - immigrants' language skills, FDI, and international bilateral trade - are interconnected, empirical studies usually focus on one. In this paper, we analyse the influence of language in international trade, concentrating solely on Portuguese exports.

2.1 The influence of language in international trade

Half a century after Marschak first introduced the concept, language is now usually included in bilateral trade studies. Trade depends on the interaction between individuals and is therefore influenced by the ease of communication between them, which depends not only on technological factors but also, and more importantly, on sharing a language in which they are able to communicate. In order to trade, two individuals from two different countries can follow one of five strategies regarding language: (i) they can share the same language and obviously use it to communicate; (ii) they can each speak in their own language and be understood by the other (a phenomenon known as intercommunication and very widespread, for example, in Scandinavian countries); (iii) they can choose one of their two languages provided that one of them speaks their counterpart's language; (iv) they can choose a foreign language that both of them can speak; (v) they can hire the services of an intermediary, such as a translator or an interpreter. These solutions range from the least costly to the most costly in terms of the transaction, since sharing the same language will entail no costs whatsoever, while hiring the services of a third party can not only be costly but also introduce noise into the communication process.

Communication costs have significantly decreased since the last quarter of the 20th century with the development of new information technologies. The Internet, in particular, has brought together potential trading partners who would have been too geographically distance in the past to even be aware of each other's existence. A common shared language, much similarly to the effect of sharing the same currency explored by Rose (2000), can decrease the fixed costs of trade and thus influence the choice of export location, although Helpman, Melitz, & Rubinstein (2008) have found that it does not influence the volume of exports once that decision has been made.

[196] In a recent study involving Portuguese exporters and Angolan distributors, Alves, Raposo, & Antunes (2012) have found that sharing the language and the culture it entails is an determinant for bilateral trade between these two countries, since the respondents to their survey explicitly mention 'same language' and 'same culture' as important factors for the success of the relationship.

Gravity models used to explain bilateral trade flows frequently include some sort of language measure, typically represented by the country's official tongue. Some authors have found a correlation between common language and trade volumes (see, for example, Helliwell, 1998; Melitz, 2008; Egger & Lassmann, 2012), while others have shown that defining common language based solely on the country's official language is not enough to capture the several nuances of the influence of language in international trade (notably Melitz & Toubal, 2014). Therefore, some sort of measure of the linguistic similarity between languages must be found, because sharing a similar language may have a considerable positive impact on trade flows, although admittedly not as important as sharing a common one.

2.2 Measuring linguistic similarity

Trying to capture the role of language similarity as a facilitating factor for bilateral trade in a more detailed fashion than the simple variable 'common official language' could generate, some authors have devised forms to measure linguistic similarity between languages. Four methods of measuring linguistic similarity between two languages have gained particular support: the results of second language acquisition tests taken by American students (Chiswick & Miller, 2005), Language Barrier Index (Lohman, 2011), Levenshtein's distance (Isphording & Otter, 2013), and open-circuit and direct-communication languages (Melitz, 2008).[197] We shall look very briefly into each of this methods before describing the method we have used in our study.

Chiswick & Miller (2005: 1) define linguistic distance as 'the extent to which languages differ from each other' and base their findings on the results of tests made by English-speaking American learners of forty-three different foreign languages that were submitted to two written assessments: the first test was written sixteen weeks after the course had started, and the second six weeks later. Following from the results of these assessment tests, the authors created a table of linguistic distances between English (the native tongue of the students assessed) and the forty-three foreign languages they were learning. According to the ranking presented by the authors, Afrikaans, Norwegian, Romanian and Swedish are the closest languages to English, while Korean and Japanese are the most distant ones. Besides purely linguistic issues that we will not go into here, Chiswick and Miller's classification is of limited applicability since it cannot be extrapolated to language pairs that do not include English (and even so only in relation to the forty-three languages surveyed).

Johannes Lohman created a Language Barrier Index (LBI) that 'quantifies international language barriers by measuring the dissimilarity between the main languages of trading partners' (Lohman, 2011: 159). The LBI was built based on the similarity between two languages, using linguistic data obtained from the *World Atlas of Language Structure* (WALS) (Dryer & Haspelmath, 2013), which presently provides detailed data on 2,678 languages, including the description of up to 144 linguistic features for each language (although not every feature is available for each language). Although based on linguistic criteria, and therefore more promising than Chiswick & Miller's proposal, the WALS is not an entirely reliable source when we look at the classification of Portuguese, due to the several inaccuracies we have found (Ferro & Costa, 2016 forthcoming).

In keeping with Lohman, Ishphording & Otten (2012) resort to linguistic data developed for the reconstruction of language families by applying lexicostatistics, which provides the quantitative comparison of lexical cognates, i.e., a word that has the same linguistic derivation of another word. This methodology was developed by the Max Planck Institute for Evolutionary Anthropology and uses a specific software entitled Automatic Similarity Judgement Program (ASJP). The main purpose of the ASJP is the automatic reconstruction of relationships between languages (Bakker *et al.*, 2009). As said before, the Levenshtein's distance measures the minimum number of additions, deletions, and substitutions to transform a word into another one (Wichman *et al.*, 2010), which, among other things, does not take into consideration diachronic change and does not account for loans, onomatopoeias, or any random changes.

Melitz (2008) does not suggest a measure to calculate linguistic distance as such, but classifies languages depending on the channels of influence of a common language, distinguishing direct-communication languages from open-circuit ones. Melitz introduces the important notion of a 'widely spoken' language in a given country, i.e., used by at least 20% of the population, and does not base his calculations solely on the official language(s) of a country.

We suggest a method for classifying linguistic similarity based on linguistic criteria, specifically etymological ones, and organized languages according to the linguistic family they belong. Since our set of data refers to Portugal exports only, we had a small number of languages to work with and found that the simple classification of languages into Romance, Germanic, and Other would be enough for our purposes. Moreover, building on Melitz's work, we also included in our study two common spoken languages in Portugal, considering that cultural and other incentives for learning a language seem to be more important than its similarity to Portuguese.

[197] For a comprehensive overview of these methods and a critique regarding their linguistic accuracy see Ferro & Costa (2016, forthcoming).

3. OFFICIAL LANGUAGES AND COMMON SPOKEN LANGUAGES

Traditionally, gravity models relied on the official language of the country for their common language variable, sometimes counting up to three official languages depending on the linguistic diversity of the country. However, over time, researchers found that basing their models on the concept of official language alone would not yield accurate results, because the linguistic capital of a nation includes much more than that - we have to consider the possible influence of significant immigrant communities who possess a different language, together with the stock of foreign languages learned in that country.

Immigrants have an important influence on bilateral trade between the host country and their country of origin. They establish informal networks through which it is easier to find trading partners in both countries. The importance of the immigrants' language is not negligible when we consider the aggregate linguistic capital of a country. Fidrmuc & Fidrmuc (2016) include Russian in the set of the most widely spoken languages in the European Union (EU), although it is not an official language of any EU member-state and therefore it is not an official language of the EU. The high number of Russian speakers in some EU countries has to do with migrant communities and historical reasons. In a recent Special Eurobarometer survey co-ordinated by the Directorate-General for Communication of the European Commission (2012), respondents from all the 27 member-states at the time provided information on a variety of topics related to their language proficiency. Interestingly, a significant part of the respondents from former East-block countries reported that Russian was their first language - e.g. in Latvia (27%) and Estonia (19%). In fact, in those two countries, only 71% and 80% of the population, respectively, are likely to use the official language of the country.

Proficiency in a foreign language is therefore an important determinant to trade and not only the official language(s) of the countries (Melitz & Toubal, 2014; Egger & Toubal, 2015; Fidrmuc & Fidrmuc, 2016). Basing their studies on the 2005 edition of the Eurobarometer Survey on languages in Europe, Fidrmuc & Fidrmuc (2016) show that not only is the widespread capacity to communicate in a foreign language an important determinant to foreign trade, but also that taking into consideration all the EU member-states English plays a particularly important role in conducting business. The authors define what they call 'conduit languages' in the EU as languages spoken at least in three different countries by at least 10% of the population in each, which leaves them with English, German, French and Russian. While including a language that is not official in the EU (Russian), this measure paradoxically leaves out Spanish and Portuguese, which although widely used outside Europe - 2nd and 6th respectively largest languages worldwide in terms of first-language speakers (Lewis, Simons, & Fennig, 2013) - have a relatively small number of speakers within the EU.

3.1 Portuguese as an international language

Portuguese is a pluricentric language, i.e. it possesses various standard versions, and is currently spoken by 208,525,450 people. It is the official language in nine countries (Angola, Brazil, Cape Verde, East Timor, Equatorial Guinea, Guinea-Bissau, Mozambique, Portugal, São Tomé and Príncipe), that span four continents. Apart from these nine countries that make up the Community of Portuguese Language Countries (CPLP), CIA's *World Factbook* also lists speakers of Portuguese in seven other states: Andorra, Bermuda, Gibraltar, Jersey, Luxembourg, Macau, and Switzerland - probably connected with migrant communities or former colonial ties, and lists Portuguese as the sixth most widely spoken language in the world, after Mandarin Chinese, Spanish, English, Hindi and Arabic.

Portuguese is therefore an important language in the global world, which led us to question whether that importance would be reflected in the volume of Portuguese exports. However, as we have stated before, the role of the foreign languages spoken in a country is also relevant, and therefore we also included that in our study.

3.2 Foreign language learning policies

Learning a language implies time, effort and often money. The harder a language is to learn, the higher those transaction costs will be. Within the framework of the game theory, Selten & Pool (1991) construct a game in which learners choose their language portfolio based on a cost/benefit analysis endeavouring to maximize benefits while minimizing costs. Building on that work, Ginsburgh, Ortuño-Ortín, & Weber (2007) conclude that the larger the population that speaks a given language, the less those speakers are inclined to learn a new one and simultaneously, due to network externalities, the more speakers a language has, the more additional learners it will attract.

The most widely learned foreign language in Portugal is currently English. And although language learning is a dynamic affair that depends on the strategic view of the government of the country, based on what the respondents to the Eurobarometer (EC, 2012) stated, English will continue to be the first option as regards foreign language learning in Portugal. Portuguese respondents answered that English is the most useful language for their personal development (53%), while 87% of the respondents stated that English is the most useful language for children to learn for their future. The latter question is interesting because French (the most widely studied language in Portugal three generations ago) still comes second (32%), followed by Spanish (10%), German (5%), and Chinese (4%).

However, linguistic competences do not come high in the priorities of Portuguese respondents - 61% answered that they do not speak any foreign language well enough to have a conversation. From those that answered that they do speak at least one foreign language, the ranking is the same as above: English (27%), French (15%), Spanish (10%), and German (1%). These linguistic skills are not as developed as in other countries (Portugal comes third from the bottom together with the UK when considering the percentage of people that are able to maintain a conversation in at least one foreign language) but need to be taken into account when we consider the role of language in international trade from Portugal's standpoint.

4. METHOD

For this study, and in line with several authors, we chose the most widely used econometric tool to study international trade, i.e., the gravity model. This model has been used since the 1960s, initially by Tinbergen (1962), but subsequently improved over the years and expanded with several variables that intend to explain trade flows between two countries. This model initially considered that exports between two countries are positively associated with the size of their economies and negatively related with factors that indicate the existence of barriers to trade, most prominently the distance between them. Thus, the basic gravity model relates the volume of exports between two countries T_{ij} with the economic weight of those two countries, measured using the GDP of exporter and importer ($GDP_i GDP_j$) and the cost of trade between them, represented by the distance between them D_{ij} (models (1) and (2)), where i and j indicate countries.

Therefore, the initial model was represented by:

$$T_{ij} = f\left[\frac{(GDP_i\ GDP_j)}{D_{ij}}\right] \qquad (1)$$

$$T_{ij} = \beta 0\ (GDP_i \,.\, GDP_j)^{\beta}1 \quad .\ D_{ij}^{\ \beta 2} \quad .\ e^{\varepsilon} \qquad (2)$$

Initially the theoretical foundations for the gravity model were explored in the work of Anderson (1979), Helpman & Krugman (1985), and Kalirajan (1999). Over time, the basic variables considered by the model (GDP and distance) were augmented with other variables, such as population, GDP per capita (Bergstrand, 1990), country size and set of countries (Azevedo, 2004). Binary variables that intended to represent the specific characteristics of the countries have also been added, such as cultural proximity, language (Endoh, 1999; Breuss, & Egger, 1999; Nitsh, 2000; Feenstra, 2002), cultural similarity, belonging to the same trade bloc - in the 1990s this effect was also considered as existence of preferential trade agreements (Breuss & Egger, 1999) - common borders, colonial relationship (Glick & Rose, 2002), among others.

Most international trade models based on the generalized gravity model study existing exports between two countries as a function not only of the variables considered on the basic model but also of their incomes (measured by GDP), their population, the distance between them, as well as a set of dummy variables in order to measure 'qualitative factors'.

Considering that the equation always implies a log-log transformation, we present the possible representation for the augmented gravity model:

$$Ln(T_{ij}) = \beta_0 + \beta_1\ Ln\ (GDP_i\ GDPj) + \beta_2\ LnD_{ij} + \beta_3\ Lang_{ij} + \beta_4\ Cont_{ij} + \beta_5\ RTA_{ij} + \beta_6\ ComCol_{ij} + \varepsilon_{ij} \qquad (3)$$

Where i and j indicate countries and the variables are defined as follows:

T – trade volume (either imports or exports and imports) between two countries:
GDP – real GDP
D – Distance
Lang – dummy variable that is 1 when i and j share a common language and 0 otherwise
Cont – dummy variable that is 1 when i and j share a common land border and 0 otherwise
RTA – dummy variable that is 1 when i and j belong to a free trade agreement area and 0 otherwise
Comcol – dummy variable that is 1 when i and j had a colonial relationship and 0 otherwise

There are currently many theoretical studies and empirical analyses on this topic. Most studies work with the analysis of trade volumes between pairs of countries in a NxN format and few apply the model to a specific country.

This type of study is fairly recent: Wall (1995) first studied trade flows between the US and 85 countries from 1994 to 1996 in order to estimate the costs of protectionism. Sohn (2005) later used the gravity model to explain South Korean trade flows and the idea of using the model to study a particular country took hold.

In line with these authors, we based our study on the data available for international trade flows from 2014, i.e., real data regarding 2013. We use the multiple linear regression of the gravity model since it has been extensively used in the past forty years and has shown to have empirical robustness and explanatory power (Kepaptsoglou *et al.*, 2010). The Ordinary Least Squares (OLS) method is the most usual technique for estimating the coefficients of the gravity model specification in its log-log form. The present study uses OLS considering the variable explained as an economic variable translating the logarithm of export volume between Portugal and a trading partner.

The variables used in the regression are economic (such as export volume between Portugal and a trading partner), linguistic (official language of the country, language family and/or language proximity), and geographical (distance between Portugal and a trading partner). Based on these data we study the relationship between the volume of exports from Portugal to its 56 main trading partners worldwide in 2013 and the language family/language used in those countries in order to find evidence for the hypotheses stated before. We chose the 56 main trading partners because in this study we wanted to analyse solely the relations that Portugal establishes with its main trading partners. Since we needed to have a sufficient number of observation to obtain a fair degree of quality of the model and, at the same time, analyse only the main partners we chose 56 out of the 211 that represent the total number of countries to which Portugal exports.

We also considered distance in order to add more quality to the analysed relationship. This is represented by the kilometres that separate Lisbon and the capital of the country. Since we consider the variable logarithm, this will represent the elasticity of trade regarding an absolute geographical distance. We expect the coefficient of this variable to be negative since it constitutes a barrier to trade given that the larger the distance between the countries the larger the barrier to the commercial relationship. Building on Lohmann (2011), we expect to find empirical evidence of the inverse effect of language commonality or similarity.

5. RESULTS AND DISCUSSION

We posited that Portuguese exports are higher to countries that share the same language (H1). To draw a conclusion on that, we studied the following model:

$$Ln(T_{ij}) = \beta_0 + \beta_1 P_j + \beta_2 LnD_{ij} + \varepsilon_{ij} \quad (4)$$

Where T_{ij} represents exports between Portugal and country j, P_j is a dummy variable that is 1 when country j has Portuguese as an official language and D_{ij} is the distance between Portugal and country j.

Table 1 - Results of the estimation of model (4)

Explanatory variables	Ln Exports	
	OLS Coefficient	Standardized coefficient (Beta)
Constant	18.667 (1.601)	------
P_j	0.403 (0.519) 0.095	
LnD_{ij}	-0.764 (0.197)	-0.476
F = 7.599		
R^2 = 0.22		

Notes:
Numbers in parentheses are standard deviations
Significance level 5%

Table 1 shows that in spite of the model being explanatory overall, only the variable D_{ij} has explanatory capacity. This result is consistent with the fundamental hypotheses of the gravity model highlighting a decrease in trade with a given country due to an increase in distance. Thus, we found no support for H1 given that the variable P_j has no

explanatory capability, i.e., there is no direct relationship between the volume of Portuguese exports and the fact that the destination country has Portuguese as an official language. This conclusion is hardly surprising in part, we believe, because companies might not be taking full advantage of the network of Portuguese-speaking countries to expand their business.

Extending the analysis and organizing countries according to the language family of the official language of the country, we studied model (5). Combining a threefold approach to the influence of language in trade, we grouped the 56 countries according to their language families. The criteria underlying our classification were:

a) linguistic criteria: languages were classified according to an etymological principle, based on their language family;
b) similarity between languages: given that Portuguese is a Romance language we decided to include the languages belonging to this family in our analysis to account for the similarity between them;
c) foreign languages: the most common foreign language studied in Portugal is currently English, a Germanic language, followed by two Romance languages (French and Spanish) and then another Germanic language, German; this led us to include in our analysis Germanic languages as well.

Since at this stage we were interested in isolating these two language families - Romance and Germanic languages - we classified all the remaining languages as belonging to the Other group. We considered the dummy variables R_j, G_j, and O_j to identify respectively Romance, Germanic and Other language families. Our aim was to analyse whether belonging to each of these language families has a direct relationship with Portuguese exports for that country. Thus, the variable R_j is 1 when country j has a Romance official language (and 0 otherwise) and G_j when country j as a Germanic official language. In case one of these variables is 1, the variable O_j is 0, conversely the latter would be 1 when the country has an official language that does not belong to any of these two families. However, econometrically this cannot be used together with the other two variables because it would entail multicollinearity given the linear relation that exists between the three independent variables.

$$Ln(Tij) = \beta 0 + \beta 1\ Rj + \beta 2\ Gj + \beta 3\ LnDij + \varepsilon ij \qquad (5)$$

Table 2 - Results of the estimation of model (5)

Explanatory variables	Ln Exports	
	OLS Coefficient	Standardized coefficient (Beta)
Constant	17.999 (1.611)	------
R_j	0.682* (0.391)	0.243
G_j	0.447 (0.382)	0.163
LnD_{ij}	-0.723 (0.192)	-0.450
F = 6.026		
R^2 = 0.258		

Notes:
Numbers in parentheses are standard deviations
Significance level 5%
* Significance level 10%

Table 2 shows that the model is explanatory overall and only variable G_j does not have explanatory capacity. Thus we found support for H2 (Portuguese exports are higher to countries that share a similar language) given that R_j has explanatory capacity, i.e., there is a direct relationship between Portuguese volume of exports and the fact that the destination country has a Romance official language. Since this is also Portugal's language family, this result was expected given that when the countries share the same language the linguistic barrier is erased and when they share a similar one communication costs tend to be lower.

In order to strengthen the analysis of H2, we introduced a new variable named ProxLing that intends to capture the language proximity between two countries. We defined this variable taking into consideration the official language of the destination country, which had to be Portuguese, Spanish, or English. Our aim was to capture a threefold

effect: with this variable we identified the countries that share a common language with Portugal but also included those that have Spanish as their official language to reflect language similarity, and those that have English as their official language to capture the effect of the most widely studied and spoken foreign language in Portugal.

Based on European Commission (2012), French is the second most widely spoken foreign language in Portugal and therefore we initially included French in our model as well. However, the analysis did not provide any statistical relevance to that fact, which we understood as being a consequence of the shift that happened some decades ago from French into English as the first foreign language studied by Portuguese children. Nowadays, although older generations still speak French, younger generations, those currently in charge of establishing commercial relationships with foreign partners, will most probably be fluent in English and not French.

Thus, the dummy variable $ProxLing_j$ is 1 when country j has Portuguese, Spanish, or English has its official language and 0 otherwise.

$$Ln(Tij) = \beta 0 + \beta 1\ ProxLingj + \beta 2\ LnDij + \varepsilon ij \quad (6)$$

After reaching a conclusion regarding the impact of language on Portuguese exports and solely with the aim of improving the quality of or model, we added the fact that a country might belong to the EU. Both effects together form the variable UEProxLing, which intends to combine the effect of a country simultaneously having one of the three aforementioned languages as its official language and belonging to the EU. We would like to highlight the fact that although Norway and Switzerland do not belong to the EU, they were considered as such given the free trade agreements that exist between Portugal and these countries.

$$Ln(Tij) = \beta 0 + \beta 1\ UEProxLingj + \beta 2\ LnDij + \varepsilon ij \quad (7)$$

Table 3 presents data on the analysis of these models.

Table 3 - Results of the estimation of models (6) and (7)

Explanatory variables	Ln Exports (6)		Ln Exports (7)	
	OLS Coefficient	Standardized coefficient (Beta)	OLS Coefficient	Standardized coefficient (Beta)
Constant	18.107 (1.546)	------	15.578 (1.835)	
$ProxLing_j$	0.726 (0.318)	0.243	--------	
$UEProxLing_j$	------	-------	0.757 (0.271)	
LnD_{ij}	-0.722 (0.187)	-0.450	-0.450 (0.210)	
F = 10.475			F = 12.144	
R^2 = 0.283			R^2 = 0.314	

Notes:
Numbers in parentheses are standard deviations
Significance level 5%

After analysing Table 3, we can conclude that both models are explanatory overall and that every variable has explanatory capacity. We would like to stress that the quality increases with the introduction of the fact that the country belongs to the EU together with what we called language proximity. Thus, we strengthened the positive effect of language proximity on Portuguese exports, which increases when both countries belong to the EU.

6. CONCLUSION

Trade costs are an important determinant of a country's capacity to benefit from regional and global production and distribution networks, and consequently bear a great importance from a policy perspective (Arvis *et al.*, 2013). As we have seen, language barriers can impose significant costs on bilateral trade between countries that do not share some sort of common spoken language - either official or acquired foreign languages. Kim *et al.* (2015) argue that governments have the capacity to manipulate the population's linguistic skills regarding foreign languages, by implementing language policies that either introduce a different language in the educational systems, virtually erasing

the native tongue from school and teaching all subjects in the new language. Or then can take a less effective step, but also much more common since it offers less controversy, introducing the desired foreign language in the school curriculum. Portuguese governments have decided to take the latter path and have gradually introduced English into the school curriculum from an increasingly earlier age. Spanish and Chinese (Mandarin) have lately been introduced as well, undoubtedly in the expectation of spurring a generation of potential speakers of those languages, who are expected to boost international trade and FDI. Foreign languages undoubtedly should play an important role in the school curriculum.

Fidrmuc & Fidrmuc (2016) argue for a causality relation between trade and language proficiency that can work either way: if enough people in two countries can communicate easily in a given language, trade will be more intensive between those two counties; but if two countries trade a lot, they will have an incentive to learn each other's language. Looking at Portuguese data, and the total volume of exports to Spain (the 1st country in terms of greatest exports from Portugal), one would think that Portuguese and Spaniards alike would have an incentive to learn each other's languages. If that holds true for Portugal, since Portuguese respondents to the Eurobarometer Survey on languages (European Commission, 2012) place Spanish as the third most useful foreign language currently and in the future and believe that they speak it well enough to have a conversation (10%), it does not for Spain, given that Portuguese does not appear in the Spanish factsheet at all. However, Italian a much less spoken language throughout the world comes sixth and seventh in the Spaniards' language preference, respectively regarding the present and the future.

Multilingualism is still not widely promoted in Portugal. Although younger generations are more language-aware much has to be done to stir the population into effectively investing in linguistic skills. In 2013, only 35% of primary pupils studied a foreign language in Portugal, a ridiculously low number when compared with the 81.7% average of the EU (Eurostat, 2015). The literature states that significant gains can arise from learning foreign languages, both to the individual, and the society at large by means of facilitating bilateral trade. Fidrmuc & Fidrmuc's (2016) research suggests that if EU countries had a higher level of multilingualism - similarly to the Netherlands, for example, where 37% of the population speak at least three foreign languages, as opposed to Portugal at only 4% - trade would be 30-60% higher than what is attributable to geographic and economic factors.

Conversely, more should be done to promote Portuguese learning abroad. With dismal figures in the EU - according to the Special Eurobarometer, Portuguese is spoken as an additional language (i.e. not first-language) by only 1% of EU population (European Commission, 2012) - Portuguese surprisingly ranks 12 in the number of foreign languages learned in the United States (Reto, 2014) and was identified as one of the ten language for the future in the United Kingdom (British Council, 2013), which constitutes a measure of the importance attributed to Portuguese.

Regarding future work, we intend to expand our analysis to all the 211 countries to which Portugal exports, and also to consider other languages besides the official language of the destination countries. For the time being we are only considering exports of goods, but we intend to include total exports and imports as well and thus provide a comprehensive analysis of Portuguese commercial relations.

ACKNOWLEDGEMENTS

This research has been partially financed by Portuguese national funding through FCT – Fundação para a Ciência e a Tecnologia as part of the project Centro de Linguística da Universidade Nova de Lisboa - UID/LIN/03213/2013.

REFERENCES

Adsera, A., & Pytlikova, M. (2015). The role of language in shaping international migration. *The Economic Journal, 125(586)*, F49-F81.

Alves, G., Raposo, M., & Antunes, A. (2012). The influence of relationship quality and culture in the relations between exporter and distributor: The case of Portuguese exports to Angola. DOI: http://hdl.handle.net/10400.22/1186

Anderson, J. E. (1979). A theoretical foundation for the gravity equation. *American Economic Review, 69*, 106-116.

Arvis, J. F., Duval, Y., Shepherd, B., & Utoktham, C. (2013). Trade costs in the developing world: 1995-2010. *Policy Research Working Paper*, *6309*. The World Bank. Available at SSRN: http://ssrn.com/abstract=2198265 [25/03/2016].

Azevedo, A. (2004). O efeito do Mercosul sobre o comércio: Uma análise com o Modelo Gravitacional. *Pesquisa e Planejamento Econômico, 34*, 307-339.

Bakker, D., Müller, A., Velupillai, V., Wichmann, S., Brown, C. H., Brown, P., ... & Holman, E. W. (2009). Adding

typology to lexicostatistics: A combined approach to language classification. *Linguistic Typology, 13(1)*, 169-181.

Banco de Portugal (2014). Lusofonia. Available at http://d3f5055r2rwsy1.cloudfront.net/images/pdfs/cardinal_lusofonia2014.pdf [25/03/2016].

Bergstrand, J. (1990). The Heckscher-Ohlin-Samuelson Model, the Linder Hypothesis and the Determinants of Bilateral Intra-industry Trade. *Economic Journal, 100(403),* 1216-29.

Breton, A. (1978). Nationalism and language policies. *The Canadian Journal of Economics, 11(4)*, 656-668.

Breuss, F., & Egger, P. (1999). How Reliable Are Estimations of EastWest Trade Potentials Based on Cross-Section Gravity Analyses? *Empirica, 26(2)*, 81-95.

British Council (2013). Languages for the future: Which languages the UK needs most and why. Available at https://www.britishcouncil.org/sites/default/files/languages-for-the-future-report-v3.pdf [25/03/2016].

Castells, M. (2001). *La galaxia internet*. Barcelona: Areté.

Chiswick, B. R., & Miller, P. W. (1995). The endogeneity between language and earnings: International analyses. *Journal of Labor Economics, 13(2)*, 246-288.

Chiswick, B. R., & Miller, P. W. (2005). Linguistic distance: A quantitative measure of the distance between English and other languages. *Journal of Multilingual and Multicultural Development, 26(1)*, 1-11.

Chiswick, B. R., & Miller, P. W. (2014). International migration and the economics of language. *IZA Discussion Paper 7880*. Bonn: Institute for the Study of Labor (IZA).

Dryer, M. S., & Haspelmath, M. (Eds.) (2013). The world atlas of language structure online. Leipzig: Max Planck Institute for Evolutionary Anthropology. Available online at http://wals.info [07-10-2015].

Egger, P. H., & Lassmann, A. (2012). The language effect in international trade: A meta-analysis. *Economics Letters, 116(2)*, 221-224.

Egger, P. H., & Lassmann, A. (2015). The causal impact of common native language on international trade: Evidence from a spatial regression discontinuity design. *The Economic Journal, 125(584)*, 699-745.

Egger, P., & Toubal, F. (2015). Common Spoken Language and International Trade. Research gate. DOI: 10.13140/RG.2.1.1165.8727

Endoh, M., 1999. Trade Creation and Trade Diversion in the EEC, the LAFTA and the CMEA: 1960-1994. *Applied Economics, 31*, 207-216.

European Commission (2012). *Europeans and their languages*. Special Eurobarometer 386. Directorate-General for Communication. Available at http://ec.europa.eu/public_opinion/archives/ebs/ebs_386_en.pdf [29-03-2016].

Eurostat (2015). *News release, 164/2015*. Eurostat Press Office. Available at http://ec.europa.eu/eurostat/documents/2995521/7008563/3-24092015-AP-EN.pdf [25/03/2016].

Feenstra, R. (2002). Border Effects and the Gravity Equation: Consistent Methods for Estimation. *Scottish Journal of Political Economy*, 49(5), 491–506.

Ferro, M. J., & Costa, R. (2016, forthcoming). The economic value of language (Ch. 10). In O. Gomes & H. F. Martins (Eds.). *Advances in applied business research: The L.A.B.S. Initiative* (pp. xxx-xxx). New York: Nova Publishers.

Fidrmuc, J., & Fidrmuc, J. (2016). Foreign languages and trade: Evidence from a natural experiment. *Empirical Economics, 50(1)*, 31-49.

Friedman, T. (2007). *The world is flat: A brief history of the twenty first century*. 3rd rev. ed. New York: Picador.

Ginsburgh, V., Ortuño-Ortín, I. & Weber, S. (2007). Learning foreign languages: Theoretical and empirical implications of the Selten and Pool model. *Journal of Economic Behavior & Organization, 64(3-4)*, 337-347.

Glick, R., & Rose, A. (2002). Does a currency union affect trade? The time-series evidence *European Economic Review 46*

Guiso, L., Sapienza, P., & Zingales, L. (2009). Cultural biases in economic exchange? *The Quarterly Journal of Economics, 124(3)*, 1095-1131.

Helliwell, J. F. (1998). *How much do national borders matter?* Washington: Brookings Institution.

Helpman, E., & Krugman, P., (1985). *Market Structure and Foreign Trade – Increasing Returns, Imperfect Competition, and the International Economy*_MIT Press

Helpman, E., Melitz, M., & Rubinstein, Y. (2008). Estimating trade flows: Trading partners and trading volumes. *The Quarterly Journal of Economics, 123(2)*, 441-487.

Isphording, I. E., & Otten, S. (2013). The costs of Babylon: Linguistic distance in applied economics. *Review of International Economics, 12(2),* 354-369.

Kalirajan, K. (1999). Stochastic varying coefficients gravity model: An application in trade analysis. *Journal of Applied Statistics*, *26(2)*, 185-193.

Kepaptsoglou, K., Karlaftis, M., & Tsamboulas, D. (2010). The gravity model specification for modeling international trade flows and free trade agreement effects: A 10-year review of empirical studies. *The Open Economics Journal, 3*, 1-13.

Kim, M., Liu, A. H., Tuxhorn, K. L., Brown, D. S., & Leblang, D. (2015). Lingua mercatoria: Language and foreign direct investment. *International Studies Quarterly, 59(2)*, 330-343.

Lameli, A., Nitsch, V., Südekum, J., & Wolf, N. (2015). Same Same But Different: Dialects and Trade. *German Economic Review, 16(3)*, 290-306.

Lewis, M. P., Simons, G. F. S & Fennig, C. D. (eds.). (2013) *Ethnologue: Languages of the World*. 17th ed. Dallas: SIL International. Available at http://www.ethnologue.com [07/03/2013].

Lohmann, J. (2011). Do language barriers affect trade? *Economics Letters, 110(2)*,159-162.

Marques, J. C. (2010). A emigração portuguesa em tempos de imigração. *Polígonos, 20*, 115-129.

Marschak, J. (1965). The Economics of Language. *Behavioral Science, 10(2)*, 135-140.

McLuhan, M. (1962). *The Gutenberg galaxy: The making of the typographic man*. Toronto: University of Toronto Press.

Melitz, J. (2008). Language and foreign trade. *European Economic Review, 52(4)*, 667-699.

Melitz, J., & Toubal, F. (2014). Native language, spoken language, translation and trade. *Journal of International Economics, 93(2)*, 351-363.

Nitsch, V. (2000). National borders and international trade: Evidence from the European Union. *Canadian Journal of Economics, 33*, 1091-1105.

Oh, C. H., Selmier, W. T., & Lien, D. (2011). International trade, foreign direct investment, and transaction costs in languages. *The Journal of Socio-Economics, 40(6)*, 732-735.

Peixoto, J. (2012). A emigração portuguesa hoje: O que sabemos e o que não sabemos. *SOCIUS Working Paper 5/2012*. Lisbon: Centro de Investigação em Sociologia Económica e das Organizações.

Reto, L. (Coord.) (2014). *O ensino da língua portuguesa nos EUA*. Fundação Luso-Americana. Available at http://www.flad.pt/wp-content/uploads/2014/09/Ensino-do-Português-nos-EUA-FINAL2.pdf [25/03/2016].

Rose, A. (2000). One money, one market: The effect of common currencies on trade. *Economic Policy, 15(30)*, 8-45.

Selten, R. & Pool, J. (1991). The distribution of foreign language skills as a game equilibrium. In R. Selten, *Game Equilibrium Models* (pp. 64-84), vol. 4. Berlin: Springer-Verlag.

Sohn, C. (2005). Does the gravity model explain South Korea's trade flows? *Japanese Economic Review*, 56 (4), 417-430

Tinbergen, J. (1962). *Shaping the world economy: Suggestions for an international economic policy*. New York: The Twentieth Century Fund.

van Dijk, J. (2006 [1991]). *The network society: Social aspects of new media*. 2nd ed. Thousand Oaks: Sage Publications.

Wall, H. J. (1999). Using the gravity model to estimate the costs of protection. *Federal Reserve Bank of St. Louis Review*, *81(1)*, 33-40.

Wichmann, S., Müller, A., & and Velupillai, V. (2010). Homelands of the world's language families: A quantitative approach. *Diachronica, 27(2)*, 247-276.

APPENDIX

Table A1 - Data for the Gravity Estimation

Exports (thousands euros)	2013	LnExport	Distance (km)	Ln Distance
Algeria	527404	13.1757221	2285	7.7341213
Angola	3112688	14.9509972	6240	8.7387354
Argentina	89533	11.4023625	10329	9.2427107
Australia	90067	11.4083091	16150	9.6896753
Austria	257034	12.4569636	2037	7.6192334
Belgium	1343285	14.1106286	1583	7.3670770
Brazil	738946	13.5129801	7486	8.9207898
Bulgaria	55343	10.9213054	2831	7.9483852
Canada	213133	12.2696716	6918	8.8418819
Cape Verde	201995	12.2159982	3020	8.0130121
Chile	77421	11.2570133	10570	9.2657750
China	657484	13.3961757	9157	9.1222738
Czech Republic	285491	12.5619657	2190	7.6916568
Denmark	314982	12.6602707	2280	7.7319307
Egypt	65520	11.0901107	3859	8.2581633
Equatorial Guinea	65758	11.0937366	5000	8.5171931
Finland	216751	12.2865045	3391	8.1288801
France	5496752	15.5196679	1138	7.0370276
Germany	5508688	15.5218370	1951	7.5760973
Gibraltar	342334	12.7435421	442	6.0913098
Greece	192696	12.1688691	2576	7.8539930
Guinea-Bissau	69787	11.1532030	3144	8.0532511
Hong Kong	130726	11.7808588	10904	9.2968849
Hungary	181123	12.1069316	2389	7.7786301
India	116801	11.6682269	8339	9.0286985
Ireland	154050	11.9450325	1558	7.3511582
Israel	98986	11.5027337	3982	8.2895394
Italy	1564826	14.2632851	1771	7.4792996
Japan	139006	11.8422723	10993	9.3050139
Korea	87058	11.3743298	10448	9.2541658
Kuwait	57577	10.9608784	5152	8.5471402
Luxemburg	67562	11.1208009	1617	7.3883278
Mexico	196456	12.1881937	8717	9.0730304
Morocco	732595	13.5043483	852	6.7475865
Mozambique	327778	12.7000918	7894	8.9738582
Netherlands	1892131	14.4532142	1786	7.4877337
Nigeria	61390	11.0250022	3766	8.2337687
Norway	106897	11.5796210	2613	7.8682542

Poland	440110	12.9947799	2518	7.8312202
Romania	301725	12.6172712	2787	7.9327210
Russia	263046	12.4800842	7310	8.8969985
São Tomé and Príncipe	50344	10.8266347	4614	8.4368504
Saudi Arabia	151612	11,9290799	5250	8.5659833
Senegal	53133	10.8805534	2835	7.9497972
Singapore	57092	10.9524192	11782	9.3743282
Slovakia	89151	11.3980868	2440	7.7997533
South Africa	160894	11.9885010	8419	9.0382463
Spain	11176719	16.2293435	300	5.7037824
Sweden	440625	12.9959494	2964	7.9942949
Switzerland	419110	12.9458886	1565	7.3556411
Tunisia	166195	12.0209170	1631	7.3969486
Turkey	381111	12.8508459	3709	8.2185175
United Arab Emirates	101711	11.5298907	6067	8.7106195
United Kingdom	2612563	14.7758422	2100	7.6496926
USA	1997743	14.5075286	5974	8.6951720
Venezuela	190114	12.1553791	6863	8.8338999