

Economics and (Minority) Language: Why is it so hard to save a threatened language?

Stefan Sperlich* José-Ramón Uriarte†

June 29, 2015

Abstract

Linguistic and cultural diversity is a fundamental aspect of the present world. It is therefore important to understand how this diversity could be sustained. Certainly, the education system, the justice system, and the economic decisions will play an important role. Our view is that a key agent in keeping diversity is the minority language speaker. Thus, we focus on the bilinguals' language choice behavior in societies with two official languages: A , spoken by all, and B , spoken by the bilingual minority. This kind of bilinguals is thought of as a population playing repeatedly a game that decides the language they will use in each interaction. We make the hypothesis that bilinguals have reached a population state with strong stability properties. Then we take the evolutionary stable mixed strategy Nash equilibrium of the game to build an economic model of linguistic behavior. It is shown that model-based predictions fit well the actual use data of Basque, Irish and Welsh. Some language policy prescriptions are provided to increase the use of B .¹

Keywords: economics of language, language use game, evolutionary stability, threatened languages.

JEL codes: Z1, H89, C72, C57

*Geneva School of Economics and Management, Université de Genève, Bd du Pont d'Arve 40, CH-1211 Genève, Suisse. E-mail address: stefan.sperlich@unige.de.

†Corresponding author: University of the Basque Country, Departamento de Fundamentos del Análisis Económico I, Avenida Lehendakari Aguirre 83, E-48015 Bilbao, Basque Country-Spain. E-mail address: jr.uriarte@ehu.es

¹This joint work started while José-Ramón Uriarte was visiting the Humbolt-Universität zu Berlin and the University of Geneva. He would like to thank the generous hospitality of both universities. We acknowledge helpful comments from the participants of the CESifo workshop *The economics of language policy* in Venice 2013. Financial support provided by the Basque Government and the Ministerio de Economía y Competitividad (ECO2012-31626) are gratefully acknowledged.

1 Introduction

When Anderlini and Felli (2004) reviewed Rubinstein’s book on economics and language (Rubinstein, 2000), they rightly claimed that ”The lack of economic analysis of the natural language that characterizes human economic behavior is certainly a large and visible hole”. Ten years later, the size of the hole remains similar. And this is astonishing because economists are interested in the study of competitive situations and, indeed, natural languages with their speech communities compete for speakers, very much like firms compete for a market share, see Nelde (1987,1995). It is hard to understand why this competitive situation, so intensely studied by sociolinguists (e.g. Fishman, 1991,2001), has not received the attention it deserves from game theorists. It did attract the attention of mathematicians, physicists and engineers interested in complex systems: Abrams and Strogatz (2003) were the first to show formally that, in the long run, the outcome of two competing languages in a given population is that one of them will fade away. This work gave rise to a substantial body of research published, mostly, in journals of Physics, see Patriarca et al. (2012) for an overview.

One reason could be that economists were interested in cheap talk games. In that class of games, what matters is any type of communication device players share: noises, signs, codes or words, with meanings only the players involved know, can all be considered ”language”, see Crawford and Sobel (1982), Farrel and Rabin (1996), Demichelis and Weibull (2008) or Heller (2014). More recently, Blume and Board (2013) weakened the assumption of a perfectly shared language by assuming that individuals speak the same language, but have different language competence.

It can safely be said that the bulk of the literature on the use of game-theoretic tools in the economics of language is limited to *semantics* (the study of meaning) and *pragmatics* (the study of meaning in context), though the area of language-learning is a notable exception; a seminal paper in this area is Selten and Pool (1991) but see also Church and King (1993), Ginsburg et al. (2007) or Gabszewicz et al. (2011).² Grin et al. (2010) and Ginsburgh and Weber (2011) are two recent surveys of the most relevant research lines developed in the economics of language.

The present work deals with highly developed multilingual societies endowed with two official languages: A , denoting the language spoken by all individuals, and B , denoting the language spoken by the bilingual minority³. Since the two languages

²In a different line, Chen (2013) has shown how the corpus of certain natural languages may shape the intertemporal economic behavior of members of its speech community.

³In the present paper, a monolingual speaker does not become bilingual by learning any second language. It should be clear from the outset that we are referring only to bilingual speakers in the

are used by the same group of people, the bilingual minority, it is said that there is *language contact*. Language contact is the most extreme form of competition between languages. The pressure of the competition is felt by the social support of language *B*, the bilingual speakers. The contact situation will influence the language choice behavior of this minority, the actual use they make of *B* in the interactions between them, their demand and supply of language *B* related goods and services, and the role they play in the transmission of *B*. Thus, the survival of language *B* and its related culture, and, therefore, the diversity of the society, depends solely on the bilinguals. Hence, what is at stake in this competitive situation is the society's language and cultural diversity. UNESCO's (2002), Article 1, says 'As a source of exchange, innovation and creativity, cultural diversity is as necessary for humankind as biodiversity is for nature'. Thus, we have a relevant competitive situation here. And yet, using the words of Anderlini and Felli, quoted above, the economic analysis of a natural minority language is a very "large and visible hole".

The type of multilingual societies we consider has enough resources to design and implement a well-articulated language policy in favor of *B*. This amounts to the existence of resources devoted to schools, teachers, textbooks, editing houses, media, institutions, that support the teaching and transmission of language *B* and its related culture, and markets where language related goods are traded. It should be noted that the minority language *B* has become an official language because, mainly through voting, people have revealed their linguistic preferences for *B* and their rights are formally satisfied by the equality, by law, of *A* and *B*.

Typical examples of this type of societies are the Basque Country, Ireland, Wales, and Scotland. In the Basque Country, the official languages are Basque and French in the French part, and Basque and Spanish in the Spanish part; in Ireland it is Irish and English; in Wales Welsh and English; in Scotland, Gaelic and English.⁴ Of course, there are more examples of highly developed multilingual societies with a minority language, but it is typically hard to get data which allow a deeper insight into the daily use of language *B*. The idea is that these societies will set a kind of benchmark in the set of all societies with minority and threatened languages contemplated by Fishman (2001).

There are some misunderstandings that are common when dealing with minority languages. One has to do with the decision to learn language *B* and the other with its use. In the multilingual societies under consideration, the learning of official two 'internal' official languages *A* and *B*.

⁴We do not include the case of French in Quebec because it is obvious that the fate of French and its related culture is not exclusively in the hands of the French speaking people of Quebec. Furthermore, while French is a minority in overall Canada, it is not so inside Quebec. Outside Quebec, French is hardly used in Canada.

language B is ensured by the education system, see MERCATOR (2008,2014). Thus, the driving force that leads parents to choose an education in the minority language B for their children is not the communicative benefits, as assumed in the Selten and Pool (1991) modeling of the decision to learn a non-native language. That is, the increasing returns to language community size plays a minor role here, for obvious reasons. Rather, the decision to learn B has to be explained by political and cultural reasons: by the support of cultural recovery and cultural diversity, and, since language is viewed as a marker of identity, by the sense of belonging to a cultural group. Then the variations of α , the proportion of bilinguals in the society, are due to factors that are exogenous to the model. This coincides with the empirical finding of Clinginsmith (2015) that language population growth is independent of language size for languages with more than a certain minimum size.

On the other hand, the use of B is restricted by statistical reasons. You need the matching of at least two bilinguals to open the possibility that B be used. But bilinguals need to recognize each other too. And, if that is the problem, the mistake is to assume that "talk is cheap": a bilingual only needs to display the "flag" or whatever sign could be used to inform the, probably bilingual, unknown interlocutor and problem solved; two bilinguals would speak in B if they want to. But the actual situation is quite the contrary, as it is well known (see Tables 1 to 3 in the next section). The idea that *talk is cheap* in a context of contact of languages does not take into account that the possibility of conflict is always present, see Nelde (1987, 1995). In particular, when the random matching is between a bilingual and a monolingual (a more likely event than the matching between two bilinguals⁵), if the bilingual signals his type and manifests the desire to speak B , then the monolingual is forced to confess his ignorance and lack of knowledge of an official language. This might generate embarrassment and introduce tension in the interaction and could hinder the dialog alignment between the interactants.

The purpose of this work is the study of the determinant of the minority language use in the type of multilingual societies mentioned above. One would think that steady increases in the proportion of bilingual speakers, α , would imply similar steady increases in the social use of B . What actually happens in those societies is that increases in the knowledge of the minority language are accompanied by smaller rates of increase in the use of B , and, in some cases, by an almost a constant use. This is what the data about the use of Basque, Irish and Welsh are telling us. We seek to explain this, so to speak, paradox, which can be formulated as follows:

⁵Since it is assumed that the proportion of bilinguals is smaller than that of monolinguals: $\alpha < 1 - \alpha$.

Why is it that having the political system, the legal instruments⁶ to facilitate the use of B , the resources, and the education system to implement a language policy in favor of B , and - even more important - the people's support and preference for the language, there is such a low use of B ?

To study this issue we use the working tools of economics; essentially, econometrics and a game of language use where bilinguals have essentially two pure strategies: *reveal your bilingual type by using B* and *hide your bilingual type by using A* .

Bilingual speakers face frequent language choice situations. We have here a nice real-life example of a population, the bilinguals, playing a game we name the Language Use Game. Thus, it is natural that they build linguistic conventions which would serve them to facilitate the choice of language and minimize the frictions associated with their language coordination. This is captured by the (interior) evolutionary stable⁷ mixed strategy Nash equilibrium of the Language Use Game, denoted x^* , they are assumed to be playing. In the context of that game, an evolutionary stable equilibrium would be a *linguistic convention* built by the bilinguals (Weibull, 1995). This equilibrium will tell us how the bilingual population is partitioned into two groups: those who choose the strategy *Reveal* and those who choose *Hide*. A bilingual in the former group will use B always whenever is matched to any other bilingual, and those in the latter group will use A between them.

We take this equilibrium as a model of the fraction of bilinguals who in real life situations use B in their interactions. To convert this working hypothesis into something operational, several transformations have to be done. First, we develop a theory to convert the Nash equilibrium x^* into a function depending on just one exogenous parameter, the proportion α of bilingual speakers: $x^*(\alpha)$. This function associates to each α the corresponding equilibrium. We then convert the Nash function into empirical models of B use. And, since the data about the use of Basque, Irish and Welsh are obtained with different methods⁸, we must build an empirical model of B use specific for each language. For Basque we build two models of *street use*; Model 1 captures the concept of *strong R(evealing) player*, and Model 2, the concept of *Weak R player*. Those who in the census answer affirmatively to use *daily* Irish/Welsh, are thought to be strategy *R players*; and those who answer

⁶The Gaelic Language Act of 2005, in Scotland; the Law of Normalization of Basque's Use of 1982, in the Basque Country; and the Welsh Language Measure of 2011, which gave the Welsh official status in Wales.

⁷In the associated single population replicator dynamics.

⁸Data about the use of Basque are obtained anonymously, from conversations heard in the streets of the main municipalities of the Basque Country; we call it Street Use: it measures the percentage of interactions executed in B. Data of Irish and Welsh are obtained from the census; we call it Daily Use: it refers to the percentage of people who at least once per day use B.

no are thought to be *H players*.

The three empirical models of the Nash equilibrium as a function of α , are increasing and convex functions; that is, the predictions they make capture the empirical fact that the use of *B* increases with α , as one can observe in the data of the three languages. Further, the model of *daily use* fits very well the observed data of Welsh for 2005 (the only year with reliable data). Similarly, the model fits very well too the observed data of Irish. For Basque we have the most reliable data that covers almost two decades. We show that for the *street use* of Basque, the data fits provided by Model 2 are better than those of Model 1. Thus, it seems that the actual bilinguals' linguistic behavior conforms to the linguistic politeness theory (Brown and Levinson, 1987) in which Model 2 is based. The resulting functional forms of the three models are compared with a non-parametric fit of the language use data. We show that they are astonishingly close. Furthermore, this closeness property holds true over time. This intertemporal stability might be interpreted as if the bilingual population in the three regions have reached an evolutionary stable equilibrium or, equivalently, behave according to linguistic conventions.

Finally, the linguistic behavior of bilingual speakers is characterized by, what is common in human language, a tendency to the optimization of the difference between communication benefits and costs⁹, guided by an economizing attitude and principles of least effort, see Selten and Warglien (2007). What is specific in the human bilingual behavior in a language contact situation, is that the linguistic conventions used by bilinguals are based on the less effort demanding strategy *Hide*, which is the source of language *B* coordination failures, and, therefore, of the weakening of the use of bilinguals' preferred language. This would be our answer to the question posed by Fishman (2001): *Why is it so hard to save a threatened language?*

The rest of the paper is structured as follows. In Section 2, we present the concept of 'street use' measure *KE* (making reference to the *Kale Erabilera* index) defined for Basque, and the *Daily Use* measure *DU* defined for Irish and Welsh. Section 3 introduces the 'Language Use Game' (LUG) from which we elaborate a theoretical framework for the *KE* and *DU*. In Section 4 we relate the theoretical analogs to a Nash equilibrium function that depends on the level of bilingual speakers. In Section 5 we estimate the model for the Basque Country, Ireland and Wales. We compare our empirical results based on our model with nonparametric analogues as a kind of model check, and study the results over time and countries. The last section concludes indicating some language policy measures to promote the use of a minority language *B*.

⁹The costs attached to the speech production in one of the two active language system (memory, word length, utterance length, frequency and other factors), see Frank and Goodman (2012).

2 The 'Street Use' and 'Daily Use' of a Minority Language

We shall consider societies with two languages: A , denotes the language spoken by all members of the society, and B , the language spoken by a bilingual minority¹⁰. Thus, there are, essentially, two linguistic groups: the monolingual speakers, those who speak just language A , and the minority of bilingual speakers, those who speak both languages, A and B . Since A and B are used by the same social group, we have, what in sociolinguistics is called, a *language contact* situation (Nelde, 1987, 1995, and Winford, 2003). In this setting, the minority language B is exposed to a direct competition with the majority language.

In this society, people, who are interacting at a certain time and place, could use either A , B , or even a mixture of both languages. Out of the total conversations that one could register at random, one could count the conversations in each of the two languages and know the proportion of people who took part in them. From there one may infer the proportion of the bilingual population of the observed place who use B in their interactions.

Definition of the Street Use Measure (KE): *Using random samples of anonymously registered conversations in the streets at a given time and place (say, a municipality or sociolinguistic zone), the Street Use Measure (KE) of minority language B shows the number of individuals observed in conversations speaking language B out of the total number of individuals observed in the place. Dividing the KE index of the given municipality by the proportion α of bilingual speakers of that municipality, KE/α , we would obtain what is often called 'Efficiency Index' (EI).¹¹*

We have data about KE only for the case of Basque. A municipality with a high EI means that the proportion of the bilingual population who actually uses language B is high. Since this efficiency measure has been used often by the sociolinguists of the Basque Country (Altuna and Basurto, 2013), we will keep the same notation in the present paper. However, we also introduce an additional index which we believe to be more informative and, as a by-product, reveals a (in our case, 'minor') problem with the KE index or any other language use information. More specifically, we assume that the street use of language B refers to conversations of 'random matches'. Therefore, an index that measures the 'efficiency of B ' should relate the observed street use of B to the probability of randomly observing a conversation composed by

¹⁰It is assumed that B is spoken *only* in the concerned society.

¹¹To our knowledge, the methodology for measuring the street use of a minority language based on anonymous observations has been developed by the group *Soziolinguistika Klusterra - the Sociolinguistic Cluster*, who operates in the Basque Country (see Altuna and Basurto, 2013).

two bilinguals. This probability is certainly $\Pr[\text{bilingual}] \cdot \Pr[\text{bilingual}] = \alpha^2$. However, in small communities, the assumption of imperfect information and random matches is rather unlikely to be satisfied always. For this reason, we call this the 'hypothetical efficiency index' $HEI = KE/\alpha^2$.

For Ireland and Wales we are only provided with the less informative index *Daily Use (DU)*. This should refer to the use of Irish and Welsh, respectively, outside home and, if possible, outside the educational system (mainly schools or colleges). We will see later that *DU* has to be modeled somewhat differently from *KE*; it does not refer to the percentage of interactions executed in *B* but just to the percentage of people who at least once per day use *B* in one of these interactions. Nonetheless, it is clear that both, the expected *KE* and the expected *DU* are functions of the same factors.

Tables 1 to 3 show the α , *KE* or *DU*, *EI*, and *HEI* for the case of Basque¹² (on the Spanish side), Ireland,¹³ and Wales¹⁴ for different years. Some features of the numbers are worth mentioning: We seem to observe a structural break for Ireland between 2002 and 2006. In fact, this is simply due to the fact that only since 2006 we have information about the daily use of Irish outside the educational system. In other words, the *DU* observations are measured with a serious error before 2006. We should have in mind that in these tables only the aggregates are given; for the empirical study we will use the data taken on province or small area level because for larger aggregations, α hardly exhibits variation around the country-mean of bilinguals.

Note that where we have observations over time, 1991 to 2011 for the Basque Country, 1996 and 2002 for Ireland, and 2006 and 2011, also for Ireland. We see that the efficiency index *EI* is quite stable, except for 2011 in the Basque Country where 4 percentage points are lost. When looking at the hypothetical efficiency index *HEI*, we observe super-efficiency, i.e. values above 100%. In other words, we observe, by far, larger proportions of bilingual conversations than one would expect if all conversations were random matches with incomplete language information. A plausible explanation is that this assumption might sometimes fail, what would explain these high values for *HEI*. Having this in mind, we can conclude from the tables that the efficiency index of interest, *HEI* - the one accounting not just for the proportion of bilinguals but for the proportion of conversations composed of two bilinguals -

¹²Data can be found on <http://www.soziolinguistika.org>. It is also given a detailed description of the definition and measuring of the *Kale Erabilera* index.

¹³Data are taken from <http://www.cso.ie/en/census/index.html>.

¹⁴All official statistics we have found, state the language knowledge, but no information about its use is given. The data used here are taken from Jones (2012) and counter checked with different reports released by the Welsh Language Board in Cardiff.

Table 1: Evolution of Knowledge α and Street Use (KE) of Basque in the Basque Country. First two columns give the percentage of bilinguals in the group of 16 years and over and the year this number was recorded by the Sociolinguistic Survey in the Basque Country. Columns three and four give the street use index (KE) and the year it was measured by the Cluster of Sociolinguistics.

Year	100 α	Year	100 KE	EI	HEI
1991	22.30	1993	11.80	0.53	2.37
1996	24.40	1997	13.00	0.53	2.18
2001	25.40	2001	13.30	0.52	2.06
2006	25.70	2006	13.70	0.53	2.07
2011	27.00	2011	13.30	0.49	1.82

Table 2: Evolution of Knowledge α and Daily Use (DU) of Irish in the Republic of Ireland. First two columns give the percentage of bilinguals in the group of 3 years and over and the year this number was recorded by the different Census in Ireland. Columns three and four give the DU and the year measured.

Year	100 α	Year	100 DU	EI	HEI
1996	41.10	1996	10.16	0.25	0.60
2002	41.88	2002	09.05	0.22	0.52
2006	40.83	2006	02.10	0.05	0.13
2011	40.60	2011	02.15	0.05	0.13

has been steadily decreasing. This is especially surprising for the Basque Country where the percentage of bilinguals has experienced a significant increase (from 1993 to 2011 by 5 percentage points, that is, by more than 20%).

Any theoretical model intended to have a certain descriptive power for the minority

Table 3: Knowledge α and Daily Use (DU) of the Welsh in Wales. First two columns give the percentage of bilinguals in the group of 3 years and over and the year this number was recorded by the Population Survey in Wales. Columns three and four give the DU and the year it was estimated from the Language use surveys.

Year	100 α	Year	100 DU	EI	HEI
2005	26.60	2005	15.38	0.58	2.17

language street use should capture and reflect these findings and aspects of the KE and DU , corresponding to the language under investigation.

3 A Model for the Street Use of a Minority Language.

In a *language contact* situation, bilinguals must make frequent language choices and are, in a natural way, involved in a game of language use coordination, seeking to maximize their linguistic preferences and communication efficiency. In this section we show how the bilingual speakers could be thought of as a population playing a game of language use. Let α denote the proportion of bilingual speakers, and $1 - \alpha$ the proportion of monolingual speakers; assume $\alpha < 1 - \alpha$.

Languages A and B satisfy the following assumption:

A.1. Languages: the languages with official status, A and B , are linguistically distant.

Under this assumption, successful communication is only possible when the interaction takes place in one language (see Ginsburgh and Weber, 2011, for a survey on how to compute the distance between two languages; see also Desmet et al., 2009, or Ispording and Otten, 2013). Thus, passive bilinguals, those who understand B but do not speak it, are not allowed. Therefore when a monolingual interacts with a bilingual, the interaction will necessarily take place in the majority language A . The assumption shows that language choice is not a trivial one.

This assumption (and its implicit political system) is satisfied, among others, by the Basque Country, Ireland and Wales. Basque is a preindoeuropean language in contact with the two Romanic languages, Spanish and French. Irish and Welsh, two Celtic languages, are in contact with the Germanic language English. All these are modern multilingual societies. In fact, we want to deal with this kind of societies. More specifically, we want to address the use of minority language B in multilingual societies which are economically developed. Thus, we will be dealing with competitive societies, both in the economic domain and the linguistic domain, with sufficient resources to implement well articulated language policies to protect and transmit B through the education system, which is mostly public. The idea is to set, with these societies, a kind of benchmark in the set of all societies with minority and threatened languages. If minority languages will be shifted here, things would not go better in the less developed multilingual societies.

Information in modern multilingual societies

We seek to know the conditions under which B might be used in the dynamic parts of these modern societies: in the urban side of them, by the bilingual population working in the core industries of those economies. We think that the survival of B , not as a museum piece to be admired by scholars and tourists, depends on the rate of use of B by the bilingual population linked directly and indirectly with these parts of the society. What kind of information about linguistic types (bilingual or monolingual) is more realistic to assume in that context?

(i) In modern societies, particularly in the areas mentioned previously, there is a great mobility (both social and geographical) of the work force. In this situation, people participate often in anonymous interactions. This implies that bilinguals take part in frequent interactions not knowing the linguistic type (bilingual or monolingual) of the interlocutor; that is, often, bilinguals interact with people whose linguistic type is private information.

(ii) In these wealthy societies, second and third generations of immigrants may choose to learn the minority language B in the school system. Thus, ethnic features do not reveal the knowledge of B .

(iii) Language contact is also a relevant element affecting information. As remarked by Nelde (1995) "neither contact nor conflict can occur between languages; they are conceivable only between speakers of languages and between the language communities". That is, contact, competition and conflict occurs between the bilingual speakers of B and those who only speak A . A continuous contact and interaction with speakers of language A eliminates signals or traits of language B speakers; for instance, accents are erased. The accents are signals that could reveal who speaks B and who does not, but the contact situation makes both bilinguals and monolinguals to have a similar accent, shaped by the dominant language. For example, in the Spanish side of the Basque Country people of any linguistic type have Spanish accent, while in the French side they have French accent.

(iv) For Nelde (1987,1995) "contact between languages always involves an element of conflict". We could observe the subtleties of this potential conflict by means of *linguistic politeness theory*, see Brown and Levinson (1987). The matching between a bilingual and a monolingual occurs more often than the bilingual-bilingual matching. Then if a monolingual is addressed in B , or observes in the interactive partner a display of markers signaling the desire to speak in B , he would be forced to reveal his type and confess his ignorance of the official language B . This might create feelings of uncertainty of his status. In terms of *linguistic politeness theory*, both the *positive face* (the desires to be liked, admired, ratified, and related to positively) and the *negative face* (i.e. the freedom of choice and freedom from imposition) of the monolingual would be damaged. But also, forcing the conversation in language A ,

the bilingual's *negative face* would be damaged too. And this hurting each other's *face* could hinder the minimal alignment of interests needed for an interaction to follow its natural path (to a common ground in which the interactants may bargain and compete). Thus, we may say that, messages conveying support of B or a display of preference for language B or starting the conversation speaking in B , could be harmful for both sides. In other words, *talk is not cheap*; for more discussion on this, see Uriarte (2015).

(v) In face-to-face interactions people try to avoid conflicts, and to this end they develop specific strategies. Politeness based strategies are just a behavior built to avoid or minimize confrontation. In our context, this affects particularly to bilinguals because they may choose language. To complete the process of gradual elimination of linguistic informative signals, the bilinguals themselves develop uninformative linguistic strategies (see below the strategy *Hide*) to avoid any possibility of upsetting the (unknown) monolingual. Essentially, that kind of strategy would be the following: if you are in the role of speaker, start the conversation in language A ; if you are in the role of listener and you are addressed in language A , respond in language A . Use language B only if your (unknown) speech partner speaks to you in B . Then if two bilinguals play this strategy they would fail to coordinate in language B .

(vi) Finally, imitation behavior so important in modern societies, fueled mainly by the media, shape external aspects and tastes narrowing down possible differences between bilinguals and monolinguals, so that one would not distinguish traits that would reveal a B speaker.

To conclude, in the type of societies we are dealing with, it is safe to say that, in many relevant domains and interactions, there is asymmetric information about the linguistic type of the interactive partners; in a different context, see how Blume and Board (2013) apply this assumption to language competence. Thus, we think it is realistic to assume that bilinguals' choice of language is made under *imperfect information*. On the other hand, since bilinguals and monolinguals have access to the sociolinguistic statistics, it seems realistic to assume that bilinguals know the average proportion α of bilinguals in the society. Thus, for the measured street use percentage we assume:

A.2.Imperfect information: *The participants of an interaction do not have, ex-ante, any information about the linguistic type (bilingual or monolingual) of each other. They only know the average proportion of bilingual and monolingual speakers, α and $(1 - \alpha)$ respectively, of the society, being $\alpha < (1 - \alpha)$.*

Democracy allows, mainly through voting, to know the individual preferences which are the basis for collective decisions. The fact that language B is official reflects

not only the linguistic rights of the minority language speakers, but, also, their preference (weak or strong) for language B . Thus, more formally, we assume:

A.3. Language preference: *bilingual speakers prefer to speak B .*

On a first glimpse this assumption might be disputable to a certain extent. However, it will automatically be relaxed by a proper choice of the payoffs which reflect how strong or weak this preference really is.

A monolingual speaker cannot make language choices; thus, assuming that every individual has the same language competence and communication skills in A , a monolingual speaker will always get a sure (communication benefit) payoff, say, n . Since language choices are made under imperfect information, a bilingual may choose the majority language A ; in that case, we will assume that he will get, as a monolingual, the payoff n , because this was a voluntary choice. Bilingual speakers will get the maximum payoff, m , when they coordinate in their preferred language B . However, $(n - c) > 0$ would be the payoff to a bilingual speaker who, having chosen B , is matched to a monolingual and, therefore, is forced to speak A . Then c denotes the frustration cost felt by this bilingual. To sum up, the following rule for the payoffs' ordering applies:

A.4. Payoffs: *It holds $m > n > c > 0$. Further, the frustration cost should be smaller than the weighted benefits, i.e. $c < (m - n)\frac{\alpha}{(1-\alpha)} = b(\alpha)$.*

Notice that a bilingual may either hide his linguistic type by choosing to speak language A , or reveal it by choosing to speak B . Then, under this set of assumptions the bilinguals' language strategic behavior is captured fairly well by the following pure strategies which are based on how B is used.

Pure strategies:

R: *reveal your bilingual type always.*

H: *hide your bilingual type and reveal it only when matched with a R player.*

With R you risk to get the minimum payoff, $n - c$, if matched with a monolingual. Choosing R also demands to lead the language coordination process; you need to signal your type, but, as mentioned in Section 3, this may damage the reputation of the monolingual and introduce tension in the dialog alignment. So you might want to avoid any type of risk, material or emotional, by choosing the less demanding strategy H which recommends to be completely aligned with the interlocutor's language. That is, you would use A and switch to B only if the interlocutor plays R . Or, as a responder, you would simply use the language spoken by the speaker. Therefore, H is not a material payoff risky choice, because you would get, either n or m and avoid possible tensions and interlocutor's *face* damaging. The language

outcome of a strategy profile composed of R and H would be B . This captures the real-life situation in which a bilingual speaks A but then code switches to B when the interlocutor happens to be an R player who answer him in B . Therefore, the language coordination problem bilinguals have, under any type of information, would be led by the one who chooses R . Strategy H players have a passive attitude in the coordination, and the profile (H, H) causes a language coordination failure since bilinguals end up speaking the less preferred language A . For this model, Figure 1 shows the language associated to each strategy profile.

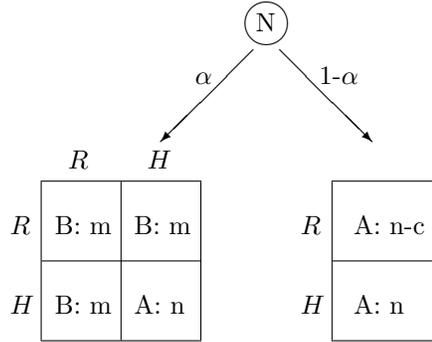


Figure 1: The Language Use Game: Capital letters indicate the language spoken and small letters the payoffs associated to each strategy profile.

Figure 1 explains the Bayesian game in which bilinguals are involved: A bilingual expects to meet another bilingual with probability α and play the game described by the payoff matrix in the left side of Figure 1. A bilingual expects to meet a monolingual with probability $1 - \alpha$ and, depending on the strategy chosen, will get the corresponding payoff shown by the column on the right side of Figure 1. The monolinguals will always get n .

This strategic interaction is a real-life example of a one-player population game: a game played by the population of bilingual speakers. Let N denote the population of bilinguals. At each time, two randomly chosen bilingual agents interact; that is, interactions are modeled as pairwise random matching between agents of N . Let x denote the proportion of bilingual speakers who play the pure strategy R .

It is easy to see that the game has two unstable equilibria, (R, H) and (H, R) , and an interior mixed strategy Nash equilibrium, where the percentage of R players is

$$x^* = 1 - \frac{c(1 - \alpha)}{\alpha(m - n)} \tag{1}$$

This equilibrium is evolutionary stable in the associated one-population Replicator Dynamics. The equilibrium could be thought of as an optimal partition of the bilingual population $N = Nx^* \cup N(1 - x^*)$. In the context of the LUG, this partition

is a 'linguistic convention' built, in the long run, by the bilingual speakers¹⁵; see Weibull (1995).

In Tables 1 to 3 it is distinguished between those who know the minority language and those who actually use it. The distinction between *knowledge* and *use* of language B is well captured by the present game. The equilibrium of the game is telling us the percentage of bilingual speakers who play R , those in the group Nx^* , and thus speak B when they interact with a bilingual speaker (R or H player), and the bilinguals who play H , those in the group $N(1 - x^*)$, speaking A between them.¹⁶

From the three equilibria of the game, it seems clear that it is with the stable equilibrium x^* that we could build an empirical model of language use. But first, let us state our main hypothesis.

Hypothesis: *The mixed strategy Nash equilibrium x^* could be thought of as a theoretical representation (that is, a model) of the fraction of bilingual speakers who, in real-life situations, use language B in their interactions.*

Actually, we know the precise number of those who have the knowledge of B , represented by the percentage α , and we have observations of KE (or DU), the measures of actual use. Hence, in the next section, we shall model x^* as a function depending on the proportion of bilingual speakers, $\alpha \in (0, 1)$, and take the KE (or DU) as observations with an expectation that is a function of α and x^* .

The goal now is first to build an empirical model and then to confront this hypothesis with the empirical evidence.

4 The Street Language Use Measure as an Endogenous Nash Equilibrium function

To develop the idea contained in the hypothesis introduced in the previous section, we must account for several issues. In particular, we should admit that the equilibrium, x^* , as written in (1), is inadequate as a model of real life language use.

¹⁵Indeed, the language spoken by each group in equilibrium is the following:

1. The subpopulation Nx^* consists of bilingual speakers who play the pure strategy R . Members of this group will speak B when they interact with any other bilingual.

2. The subpopulation $N(1 - x^*)$ consists of those who play the pure strategy H . In the interactions between bilingual speakers belonging to this group the language spoken is A . Members of this group will only speak B when they interact with those in Nx^* .

¹⁶Note that $(1 - x)^2$ is the matching probability of two bilingual H players and $(2x - x^2)$ all the matchings in which the bilingual R players participate. Thus, in equilibrium, the percentage of bilinguals speaking in A will be higher than those speaking in B for values of $x^* \in (0, 0.293)$.

4.1 Minority language learning decisions and their impact on x^*

The driving force to decide to invest in the learning of the minority language B are not the communicative benefits as in the Selten and Pool (1991)'s model. In the context of this work we doubt it is correct to assume that the communicative benefit of an individual from learning B increases with the number of people he may communicate using B ; he can always communicate with them in A , and, in most cases, the speech community is just a small minority. Further, trade has little or no influence in the learning of B because, in a language contact situation a non-native has much larger incentives to learn A , instead of learning B . In other words, increasing returns to scale to language size plays a minor role here.

This is not to say that communicative benefits do not play a certain role. In fact, we acknowledge certain influence in the assumption about payoffs $m > n$. But this is rather due to the fact that bilinguals recognize each other when they interact in language B , i.e. their belonging to a cultural group. For Fishman (1991), the culturally specific language of a society "is more than just a tool of communication for its culture. [...] Such a language is often viewed as a very specific gift, a marker of identity and a specific responsibility vis-a-vis future generations".

Language B has become official language because a majority of the multilingual society is in favor of language rights, language justice and maybe even a positive discrimination in favor of B . Thus, in the type of societies we are dealing with, adults learn B , fathers send their children to schools where contents are taught in B , and university students may choose some of their lectures in B , because these people support the policy of cultural recovery, of maintaining cultural diversity, of national pride and, also, the social benefits of bilingualism. For these reasons, we will treat the decision of language learning investment and therefore the proportion α of those who have the knowledge of B , as *exogenously given*. Thus, in our modeling, we are assuming that the growth of the speech community of B is independent of its size. This coincides with the empirical finding of Clinginsmith (2014) that language population growth is independent of language size for languages with more than a certain minimum size¹⁷.

4.2 The empirical model of language B use

In x^* , as defined in (1), we have further m , n and c as exogenously given parameters. If we take α as the only variable in the model, then we would get the mixed

¹⁷Clinginsmith (2014) quantifies this minimum amount in 35,000.

strategy Nash equilibrium as a function of the proportion of bilinguals in the society $x^*(\alpha)$. We may think that those who use language A (i.e. the monolinguals and the bilinguals who voluntarily choose it) as a means of communication will get a certain payoff level n which, essentially, remains constant and independent of α . This payoff n could be thought of as the natural payoff level that one might obtain from using the language spoken by all members of the society. With m and c kept constant, $x^*(\alpha)$ would be an increasing and concave function of α , as it happens with the model of language use of Iriberry and Uriarte (2012). But it is unrealistic to assume that m and c do not change when the proportion of bilingual speakers in the society (α) change.

It seems more natural to assume that the parameters m and c change with α . When α reaches a certain higher level, B would then be perceived, mainly by its speech community, not as an endangered language. As α keeps increasing, bilingual speakers would tend to feel that there are no reasons for exceptional (high) levels of payoffs and would be inclined to assign smaller ones to the (now a much more frequent) event of coordinating in language B . In other words, as α increases, m should decrease and approach the payoff level, n , of the normalized language A . By the same reason, as α increases, the event of failing to coordinate in B would be less frequent, and so the frustration cost c would decrease too tending to 0. Hence, there must exist some crucial level for alpha at which the convergence of m to n and of c to 0 will occur.

Nevertheless, one would still expect that $x^* = x^*(\alpha)$ be an increasing function of α . That is, as α increases, the equilibrium proportion of bilingual speakers who would play strategy R , $x^*(\alpha)$, must also increase. But, one should choose functional forms for $m(\alpha)$ and $c(\alpha)$, as decreasing functions, which would give rise to an increasing equilibrium function $x^* = x^*(\alpha)$ compatible with α^2 .

Even though the ratio $\frac{\alpha}{(1-\alpha)}$ is increasing in α , for certain functional forms of $m(\alpha)$ and a given value of n , the weighted profit $b(\alpha) = (m(\alpha) - n)\frac{\alpha}{(1-\alpha)}$ should be decreasing. Let us suppose that $m(\alpha)$ has the simple form $m(\alpha) = \frac{K}{\alpha}$, where $K > 0$ is a constant.

Let α^* denote the proportion at which bilingual speakers perceive that B has reached the status of a socially normalized language in the sense that $m(\alpha^*) = n$. For any $\alpha < \alpha^*$, $m(\alpha) > n$, and a bilingual speaker would get a positive net profit $m(\alpha) - n$ whenever he is able to coordinate in language B .

As said above, the frustration cost $c(\alpha)$ is decreasing in α . In order to allow the use of B in the equilibrium, we need to assume $0 < c(\alpha) < b(\alpha)$ for any $\alpha < \alpha^*$. We get

$$c(\alpha) = (m(\alpha) - n)\frac{\alpha}{(1 - \alpha)} - R(\alpha) \geq 0 , \quad (2)$$

for some $R(\alpha) > 0$ being the net benefit. Now, inserting equation (2) in (1) we obtain

$$x^*(\alpha) = \frac{\alpha(m(\alpha) - n) - c(\alpha)(1 - \alpha)}{\alpha(m(\alpha) - n)} = \frac{R(\alpha)(1 - \alpha)}{\alpha(m(\alpha) - n)}.$$

And with $m(\alpha) = \frac{K}{\alpha}$,

$$x^*(\alpha) = \frac{R(\alpha)(1 - \alpha)}{K - n\alpha}. \quad (3)$$

Equation (3) shows how the equilibrium proportion of the bilingual population playing R changes as the proportion of bilingual speakers, α , changes. Note that the denominator of (3) is greater than zero because $m(\alpha) = \frac{K}{\alpha} > n$ for all $0 < \alpha < \alpha^*$. Clearly, $K = n\alpha^*$.

Hence the function $x^* = x^*(\alpha) > 0$, will be increasing in α if $R(\alpha)$ is either increasing or at least not decreasing faster than $\frac{1-\alpha}{\alpha(m(\alpha)-n)}$ is increasing in α .¹⁸ Therefore one would suppose that the net benefit function

$$R(\alpha) = b(\alpha) - c(\alpha) = (m(\alpha) - n)\frac{\alpha}{(1 - \alpha)} - c(\alpha)$$

is a function of α with a well defined first derivative. The perception that B has reached the status of a normalized language will normally occur before a 100% of the population becomes bilingual, making $m(\alpha^*) - n = 0$, at some $\alpha^* \leq 1$.

4.3 Reference points

However, people may have different perceptions about when the minority language B could be said to be normalized. Given a certain linguistic context (say, a certain municipality or linguistic zone), the perception an individual may have about the normalization of B would be conditioned by that language environment. The present proportion of bilingual speakers in that environment will define the minority language reference point for each individual living in that context. As it happens with the perception of attributes such as wealth, the frequency with which an individual experiences the event of meeting bilingual speakers in the past and in the present will determine an adaptation level or reference point. Then people will perceive and quantify α^* depending of that reference point. Quoting Kahneman and Tversky (1979): "The same level of wealth, for example, may imply abject poverty for one person and great riches for another, depending on their current assets". Similarly, in a linguistic context where the number of bilingual speakers is relatively small, people would assign a value to α^* smaller than the value assigned by those who live in a context with a relatively high proportion of bilingual speakers.

¹⁸It can easily be checked that $g(\alpha) = \frac{1-\alpha}{\alpha(m(\alpha)-n)}$ is increasing by calculating its first derivative, inserting $m(\alpha) = K/\alpha$ and using that $K/\alpha^* = n$ with $\alpha^* \leq 1$.

Let $\alpha_L(\alpha)$ be the function that assigns to each α of a linguistic context L the value α^* at which the B speakers of that area perceive that the language B is already normalized. We shall assume that $\alpha_L(\alpha)$ is a concave and increasing function with negative second derivative, as the 'value function' of Kahneman and Tversky (1979). Thus, $\alpha_L(\alpha) = \alpha^*$, with $\alpha^* > \alpha$, would be the value at which $m(\alpha^*) = n$. Each linguistic context L will have a specific α^* or perception as to which is the proportion of language B speakers that converts B into a normalized language.

Then, at the convergence point, $m(\alpha^*) = \frac{K}{\alpha^*} = n$, so the constant $K = n\alpha^*$. Making the substitution in the denominator of (3), we get

$$x^*(\alpha) = \frac{R(\alpha)(1 - \alpha)}{n(\alpha^* - \alpha)} . \quad (4)$$

Note that since $x^*(\alpha)$ is an interior mixed strategy equilibrium, that is, $x^*(\alpha) \in (0, 1)$, then $R(\alpha)(1 - \alpha) < n(\alpha^* - \alpha)$, where $1 > \alpha^* > \alpha$, and hence $n > R(\alpha)$.

4.4 Three empirical models

As mentioned above, we have data obtained with different methods.¹⁹ This requires different modeling of the data: expected language B use in the street (KE) for Basque, and daily use (DU) for Irish and Welsh.

Model 1: Having in mind Figure 1, the model of expected (street use) KE is obtained by simply adding up all the probabilities of random matchings in which an R player does participate; that is, the probability of a random match where an R player meets another R player ($\alpha^2 x^{*2}$), plus the probability that an R player meets an H player ($\alpha^2 x^*(1 - x^*)$), plus the probability that an H player meets an R player ($\alpha^2(1 - x^*)x^*$). In sum,

$$E[KE|\alpha, x^*] = \alpha^2 (2x^* - x^{*2}) . \quad (5)$$

Model 2: Model 1 is the result of an R player who takes always the lead in the language coordination process. Let us call him *strong R-player*. In real life we may find a bilingual R player who not always leads the language coordination process. Speaking always B , as recommended by strategy R , is very demanding as it was mentioned in Section 3 to justify the assumption about the type of information.

¹⁹In the Basque Country, every four years during some week of the month of October, people in the main streets of each municipality do not know that there are surveyors gathering information about the language they are using in their conversations, as well as the number of people participating in them. In Ireland and Wales, the data are obtained from the Census, from answers to questions like "how often do you use B outside the educational system, with friends...".

Notice too that the choice of always speaking B is equivalent to adopting the role of leader in solving the bilingual speakers' language coordination problem. Finally, time is an important factor, and choosing B is more time consuming than speaking, without hesitation, in A with unknown people.

We may conclude that "speaking B always" is risky, because the bilingual may suffer often a frustration, requires to be persistent, ready to cope with possible misunderstandings and arguments; and to be considered rude and impolite (Brown and Levinson, 1987). It is therefore very plausible that (with time) a "speak B always" bilingual might get tempted to avoid tensions, misunderstandings and time consuming, and adopt a more passive behavior. Contrary to the *strong R-player*, this type of R player will take into account any prior information he might get. We shall call him *weak R-player*.

Without loss of generality, let us assume that 50% of the time he takes the leading role by opening the conversation, certainly with B . The other 50% he is led (that is, he adopts the role of responder), and then he accepts the language used by the one who is leading the conversation.²⁰ The resulting expected KE is the sum of the probability of a random match where an R player meets another R player ($\alpha^2 x^{*2}$), plus the probability that an R player meets an H player ($\alpha^2 x^*(1 - x^*)$), i.e.

$$E[KE|\alpha, x^*] = \alpha^2 x^{*2} + \alpha^2 x^*(1 - x^*) = \alpha^2 x^* . \quad (6)$$

Model 3: For modeling the expected daily use, DU , it is assumed that almost all individuals playing strategy R will answer (in the census or survey) that they use B every day, whereas almost all individuals playing strategy H will answer that they do not. Deviations from these rule should cancel out in average. Consequently, the expected DU is simply:

$$E[DU|\alpha, x^*] = \alpha x^* . \quad (7)$$

We may conclude with the following:

Corollary 4.1. *Let α denote the proportion of bilingual speakers in a certain sociolinguistic context. Then, $\alpha^2(2x^* - x^{*2}) = PKE_1(\alpha)$ is the predicted street use of B (PKE) in that context if players of the R strategy are strong (Model 1); if the R players are weak, then it is $\alpha^2 x^* = PKE_2(\alpha)$ (Model 2). The predicted daily use is $PDU(\alpha) = \alpha x^*$.*

²⁰When he is not leading the language coordination process, and the interlocutor started the interaction in B , then the weak R -player would know that he is interacting with an R player and would respond in B . If the interlocutor's language is A , then, he will know the interlocutor is not an R player. Consequently, the probability that he is interacting with a monolingual has increased from $(1 - \alpha)$ to $(1 - \alpha)/(1 - \alpha x^*)$ which can be very high, depending on α and x^* . Hence, to avoid the frustration cost c , the weak R -player will play H and answer in language A .

Putting together all the formulas from above, we have

$$\begin{aligned}
PKE_1(\alpha) &= \alpha^2 \left\{ 2 \frac{R(\alpha)(1-\alpha)}{n(\alpha^* - \alpha)} - \left(\frac{R(\alpha)(1-\alpha)}{n(\alpha^* - \alpha)} \right)^2 \right\} \\
PKE_2(\alpha) &= \alpha^2 \frac{R(\alpha)(1-\alpha)}{n(\alpha^* - \alpha)} \\
PDU(\alpha) &= \frac{\alpha R(\alpha)(1-\alpha)}{n(\alpha^* - \alpha)}.
\end{aligned}$$

With the specifications $R(\alpha) = \beta_1 \alpha^{\beta_2}$ for the unknown β_1, β_2 , and $\alpha^* = \alpha^{\beta_3}$ (β_3 also unknown), we get

$$PKE_1(\alpha) = \alpha^2 \left\{ 2 \frac{b_1 \alpha^{\beta_2} (1-\alpha)}{\alpha^{\beta_3} - \alpha} - \left(\frac{b_1 \alpha^{\beta_2} (1-\alpha)}{\alpha^{\beta_3} - \alpha} \right)^2 \right\} \quad (8)$$

$$PKE_2(\alpha) = \alpha^2 \frac{b_1 \alpha^{\beta_2} (1-\alpha)}{(\alpha^{\beta_3} - \alpha)} \quad (9)$$

$$PDU(\alpha) = \frac{\alpha b_1 \alpha^{\beta_2} (1-\alpha)}{(\alpha^{\beta_3} - \alpha)}, \quad (10)$$

where $b_1 = \beta_1/n$. These are the models we will study empirically, based on data for Basque, Irish and Welsh. Note that there is no particular reason why n , the payoff for communicating in A would have seriously changed over the considered time window. Consequently, even though β_1 is not exactly identifiable, you can interpret the development of b_1 over time as the development of β_1 . Moreover, you might be willing to assume that n does not vary (substantially) over the considered regions (Wales, Ireland and Basque Country); in that case you can also interpret the differences in b_1 between regions as the differences in β_1 . In other words, we can identify the differences and changes in the net benefit function $R(\alpha)$ over time and region once we assume n to be constant.

5 Empirical Evidence

For each of the three languages and regions we will present in the following, (a) the parameter estimates for the models (8), (9) and (8), and (b) the resulting functional forms, compared with a nonparametric fit of the observed KE (respectively DU) on α . For the Basque, will further study the distributions of α , KE and EI over the municipals.

In order to estimate the coefficients from the samples $\{KE_{cti}, \alpha_{cti}\}_{i=1}^{n_{ct}}$, $\{DU_{cti}, \alpha_{cti}\}_{i=1}^{n_{ct}}$ for country sample c in year t one might consider either

$$KE_{cti} = PKE_{cti} + \varepsilon_{cti}, \quad (11)$$

(and analogously for DU) or, as a model with multiplicative structure,

$$\log(KE_{cti}) = \log(PKE_{cti}) + \epsilon_{cti} , \quad (12)$$

estimating the parameters $b_{ct,j}$, $j = 1, 2, 3$ by least squares under the constraints that $0 \leq b_1 < 1$ and $\beta_3 < 1$.²¹ Note that models (11) and (12) respectively, require different assumptions on the (random) deviations from the mean. They are, however, not testable, so that we have no particular a priori preference.

While the general findings are quite similar for one or the other estimation strategy, predicting the language use from the logarithmic version (and consequently $\widehat{\log(KE)}$ or $\widehat{\log(DU)}$) is somewhat more complex as one has to correct the - in our case heteroscedastic - error dispersion since $E[\log(KE)|\alpha] < \log E[KE|\alpha]$; the same holds certainly for DU . We therefore concentrate on the presentation of the least square estimates resulting from model (11). All function estimates are given together with nonparametric fits of KE (DU) on α using local quadratic estimators with Epanechnikov kernel. We used local bandwidth such that 25% of all sample points are inside the kernel support.²² For more details see the Appendix.

5.1 Estimation Results

We start with Wales for which we have reliable data only for 2005.²³ In Figure 2 and Table 4 are given the results of estimating the PDU function for Wales. The solid line refers to the parametric model, the dashed one to the nonparametric analogue. The circles indicate the recorded observations.

Table 4: Parameter estimates of equation (8) with additive error for Welsh local authorities; *obs.* indicates the number of observations.

year	2005
b_1	1.105
β_2	0.602
β_3	0.010
obs.	22

Given the small sample size, the estimates may not be very reliable but we see the main characteristics: there is no major difference between the parametric and the

²¹From our discussion above, we see that there is no clear constraint for β_2 , although one would expect $\beta_2 > 0$, as this gives an increasing $R(\alpha)$.

²²More specifically, we used the **R**-procedure `locfit` with $\alpha = 0.25$ and $deg = 2$.

²³The data were actually collected during the period from 2004 to 2006.

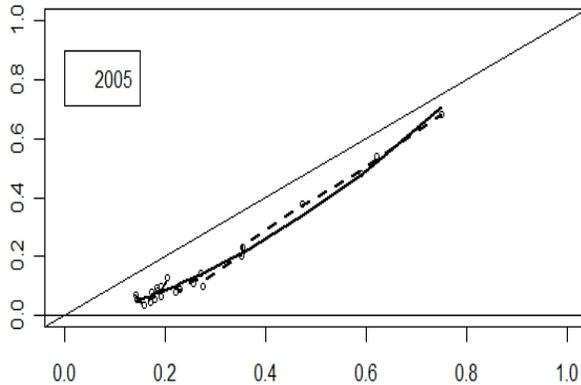


Figure 2: Parametric (solid line) and nonparametric (dashed line) estimates of PDU for Wales, together with angle bisector.

nonparametric curvatures; both are well adapted to the data. That is, our model for x^* seems to fit pretty well what has been observed regarding the daily use of the Welsh language. Given the fact that we have no information about n (payoff of conversations in English), there is no particular interpretation for b_1 , but we know that $\alpha^* = \alpha^{\beta_3}$ while β_2 gives us the speed at which the net benefit function $R(\alpha)$ increases with α . The net benefit increases at a faster rate than $\sqrt{\alpha}$ but with decreasing intensity ($\beta_2 < 1$). The percentage α^* at which Welsh is no longer perceived as a minority language (such that $m = n$) seems to be above 90%.

For Ireland, we have data for different levels of aggregation, namely for about 3400 so called 'electoral divisions', for about 180 'local electoral areas', and for the 34 counties. The last aggregation level is of little help as it exhibits little variation in α . The estimation results for the other samples are given in Table 5 and Figure 3.

Table 5: Parameter estimates of equation (8) with additive error for Irish electoral divisions (first two columns), and Irish local electoral areas (last three columns); *obs.* indicates the number of observations available for that year.

year	2002	2006	2002	2006	2011
b_1	0.017	0.192	0.190	1.660	0.491
β_2	2.307	4.939	2.155	6.724	4.439
β_3	0.978	0.799	0.783	0.715	0.648
obs.	3422	3409	180	180	201

Admittedly, for both Irish and Welsh data there are some uncertainties concerning the DU . As already mentioned for the Irish data, people may use every day language B only at school, but elsewhere they play strategy H . Only since 2006 there is a clear

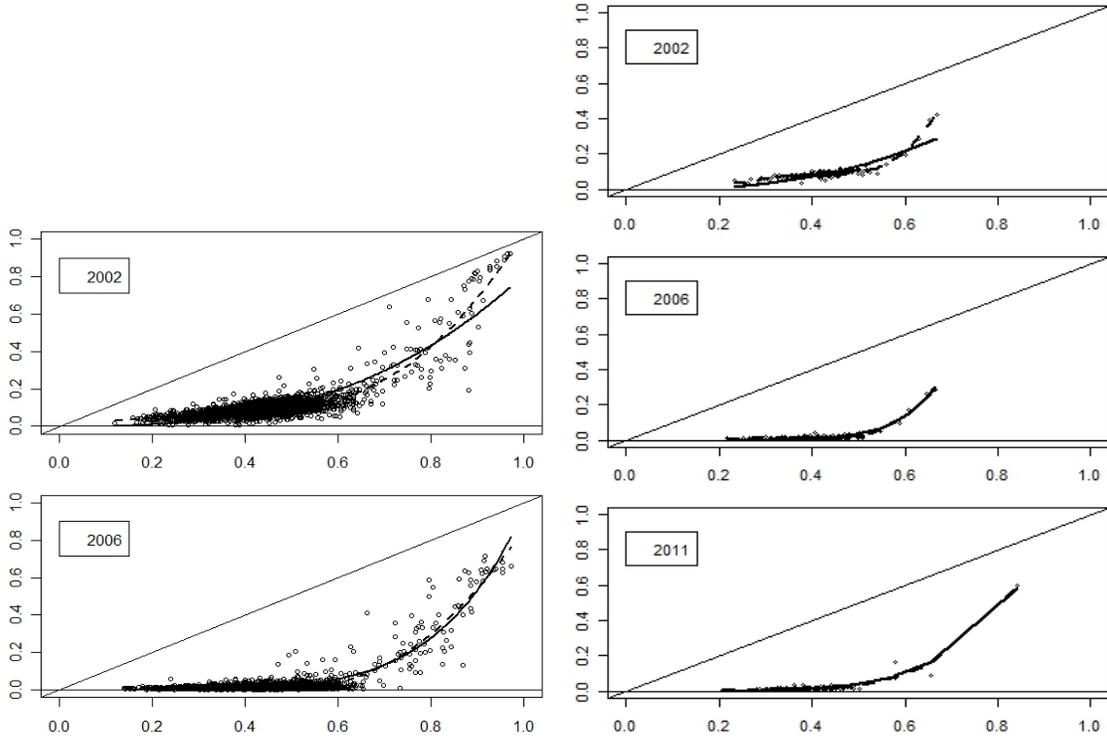


Figure 3: Parametric (solid line) and nonparametric (dashed line) estimates of PDU for Ireland, together with angle bisector: right panel for 'electoral divisions', left panel for 'local electoral areas'.

definition of daily use outside the educational system. Consequently, the difference between the parametric and the nonparametric fit for Ireland in 2002 might be simply to this miss-measurement. We see, however, that for 2006 and 2011 our model fits pretty well the observed data of daily language use, being close to the nonparametric estimate (data fit without a model).

Concerning the parameter estimates we notice that different aggregation levels lead to quite different estimates for b_1 , which is not surprising since these may lead to differently perceived payoffs n . Fortunately, the two parameters with some value of interpretability, β_2 and β_3 , are comparable over the different aggregation levels. We conclude that the net benefit function $R(\alpha)$ has become much steeper with regard to α (from about $\beta_{1,2002}\alpha^2$ in 2002 to about $\beta_{1,2006}\alpha^5$ in 2006), but as $\alpha < 1$, the perceived net benefits have actually diminished dramatically. The α^* , where $m = n$, went down from almost α to just a bit more than $\sqrt{\alpha}$ what is in accordance, i.e. expected from our discussion in Section 4.

For the Basque Country we have data of KE . Moreover, the KE measure did not change so that we can airily compare the five years to study the long-term dynamics (over a period of almost 20 years) later on. The estimates for Model 1, equation

(8) where the R players are of the *strong R-players* type, are shown in Figure 4 and Table 6. The estimates for Model 2 (*weak R-players*), equation (9), are given in Figure 5 and Table 6. In the Figures, the non-parametric estimates are shown as dashed lines whereas the solid lines are the respective PKE model with the parameter estimates as shown in Table 6.

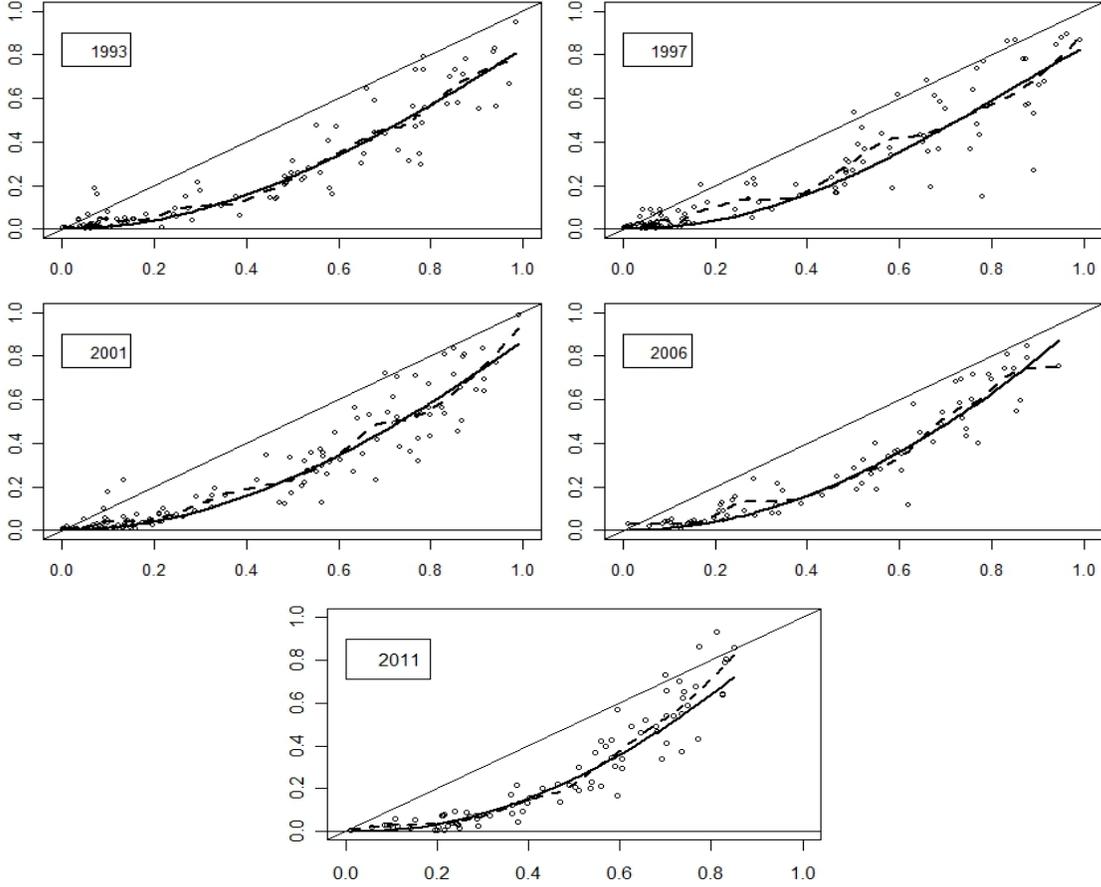


Figure 4: Parametric (solid line) and nonparametric (dashed line) estimates of PKE_1 , **Model 1** for the Basque country, together with angle bisector.

A main difference with respect to the other linguistic zones above, is that the dispersion is much larger here: for α as well as for the α conditioned KE . This might give the misleading impression that our theoretical model would capture less well the reality. As long as we take α as the only varying explanatory variable, this is not true: the nonparametric fit and the theoretical model are very close; in fact the nonparametric one is just a bit more wiggly than the fit with Model 2. This indicates that a better projection of KE on α than the one provide by our theoretical model for x^* is hardly possible. However, the additional, but unexplained variation in KE for given α , may easily be explained by local particularities. To look for those and include them in a regression model might be an interesting exercise, but it is not

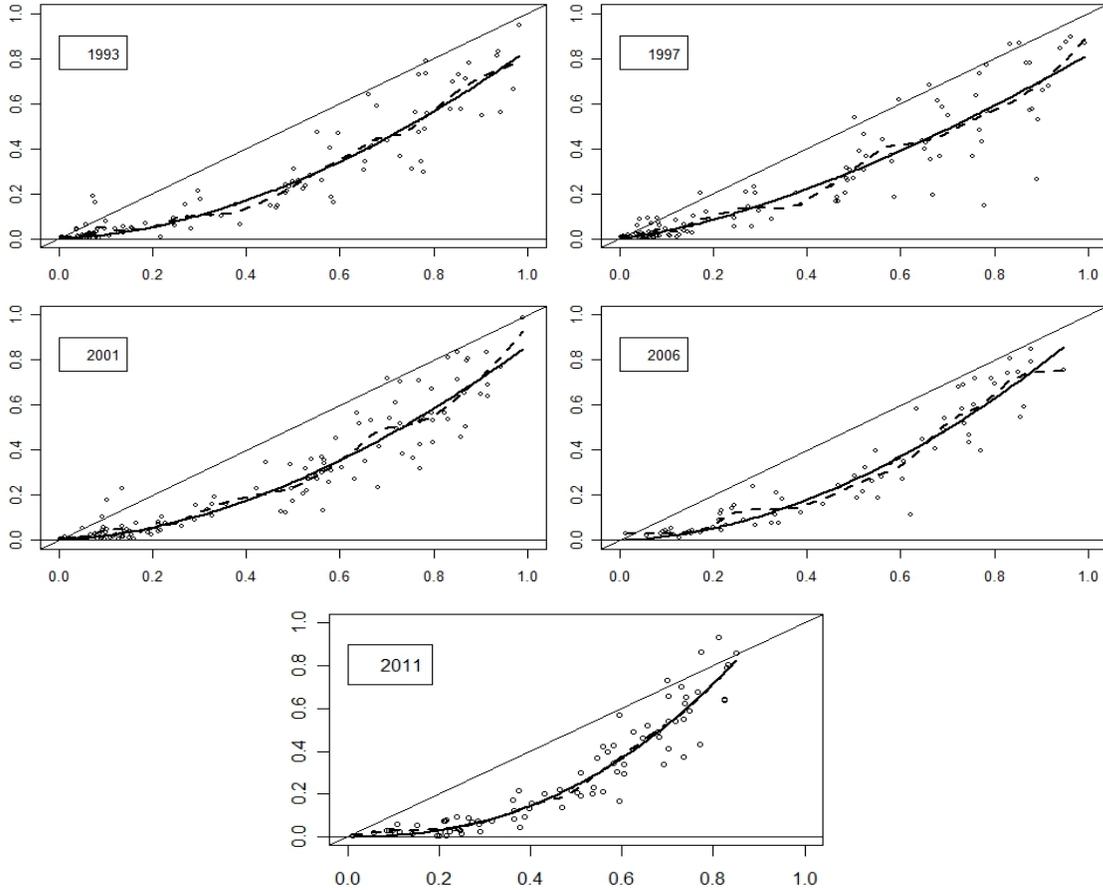


Figure 5: Parametric (solid line) and nonparametric (dashed line) estimates of PKE_2 , **Model 2** for Basque, together with angle bisector.

Table 6: Parameter estimates of equation for the Basque municipals; *obs.* indicates the number of observations available for that year.

year	1993	1997	2001	2006	2011	
Model 1	b_1	0.136	0.116	0.118	0.176	0.685
	β_2	0.192	-0.106	0.204	0.262	0.559
	β_3	0.838	0.858	0.864	0.815	0.434
Model 2	b_1	0.397	0.300	0.963	0.332	0.919
	β_2	0.642	0.824	0.388	0.584	-0.252
	β_3	0.719	0.786	0.292	0.720	0.021
obs.	101	121	134	74	84	

the aim of our language game model.

Instead, it is interesting to compare the data fits provided by Model 1 based on

strong R-players with those provided by Model 2 derived from the behavior of *weak R-players*. By eyeballing we would guess that Model 2 fits better. In fact, for all years the residual variances obtained from Model 1 are much bigger than those from Model 2. Applying a nonparametric specification test, see next subsection, confirms these findings. Thus, it seems that the justifications based essentially on linguistic politeness theory (Brown and Levinson, 1987) to capture the *weak R-players'* behavior, has some positive empirical implications. In the following we therefore only discuss the parameter estimates corresponding to Model 2.

When looking at these estimates, we see a major change only from 2006 to 2011 in b_1 and β_3 . That is, while we hardly see a change in the net benefit function $R(\cdot)$, the α^* went down from about $\alpha^{0.85}$ to $\alpha^{0.43}$. This is a similar development as we observed for Ireland in what concerns α^* , but is contrasted by the stable net benefit $R(\cdot)$, which actually has even increased for Basque due to the in average increasing α . Recall that in Ireland this net benefit went down to almost zero, what strongly fosters the extinction of the Irish language. For Basque we only have a slight, probably insignificant increase in β_2 giving in 2011 a $\beta_2 \approx 0.6$ what results in an increasing $R(\alpha)$ but with decreasing returns to α .

5.2 Testing for significance and Analyzing the Changes over Time

We also tested the functional form of our model nonparametrically for each country and year. This was done along the bootstrap test of Härdle and Mammen (1993), see appendix. One has to know that nonparametric tests conditioned on the design - like almost all nonparametric bootstrap tests are - will always reject once the sample size is large enough compared to the residuals variance. It is then up to the empirical researcher to decide whether the detected statistically significant differences matter for his research question or not. For example, in our context a smooth oscillation around the parametric counterpart like we see it for several years in Figure 5 is of no concern whereas the lack of adaptation of our model to the Irish data in 2002 is (Figure 3).

Recall that for this latter example DU was by far over-reported in 2002, so that our model cannot replicate this (almost constant) shift as our model does not contain an intercept. Consequently, for Irish data of 2002 the test is expected to reject. For 2006, the Irish data exhibit a slightly stronger bend (like of an elbow) than a parametric model can produce, so that given the extremely small residual variance, a residual wild bootstrap test is again expected to reject. For all other years and data we expect to not reject, not even at a 10% level, except maybe Model 1 for Basque in

1997. These are actually the results we obtain from the test, which gives the p-values calculated from 1000 (wild) bootstrap samples as follows: model for Welsh in 2005, $p = 0.54$; model for Irish in 2002 and 2006, $p < 0.01$ but $p = 0.11$ in 2011; Model 1 for Basque in (1993,1997,2001,2006,2011) obtains p-values (0.47, 0.04, 0.22, 0.32, 0.15) whereas Model 2 obtains p-values (0.64, 0.68, 0.43, 0.35, 0.89). Clearly, Model 2 has much higher p -values throughout; so it indeed explains much better the data than Model 1 does.

For Basque we have the most reliable data, which are available for over almost two decades. We can also study the dynamics over time. In Figure 6 we have summarized the changes of the PKE function over time. Though quite stable over the years, we mainly see that the street use of Basque, for given α , seems to steadily increase for municipalities where $\alpha > 0.5$, whereas for those with $\alpha < 0.5$ it is varying over time without a clear tendency. These findings are made independently from looking at Model 2 estimates or at the nonparametric estimates.

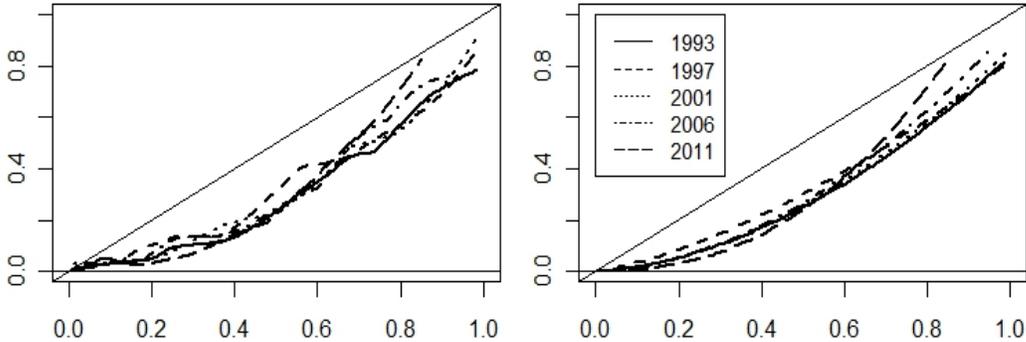


Figure 6: Nonparametric (left panel) and Model 2 based (right panel) estimates of PKE for the Basque country.

Having said that, it would be interesting to contrast it with the development of the α , and also the KE and EI , each separately. The box-plots in Figure 7 illustrate quite well the development of the distributions over the years. First, recall that we are looking at all combinations (α_{ti}, KE_{ti}) (for $t = 1993, 1997, 2001, 2006, 2011$) without weighting them by the population size of municipality i . This explains why it seems that the percentage(s) of bilingual speakers went down though the real total percentage has steadily increased, see Table 1. We see that all years exhibit a huge dispersion for α and KE with no stabilization of any of the distribution of these indices. We observe a shrinking number of municipalities with only small α or/and small values of KE . This might explain why people feel that α has still to increase much more (i.e. b_3 has fallen) to become a normalized language (i.e. $\alpha = \alpha^*$). At the same time, the slightly increased net benefit in 2011 is not clearly reflected in

these box-plots.

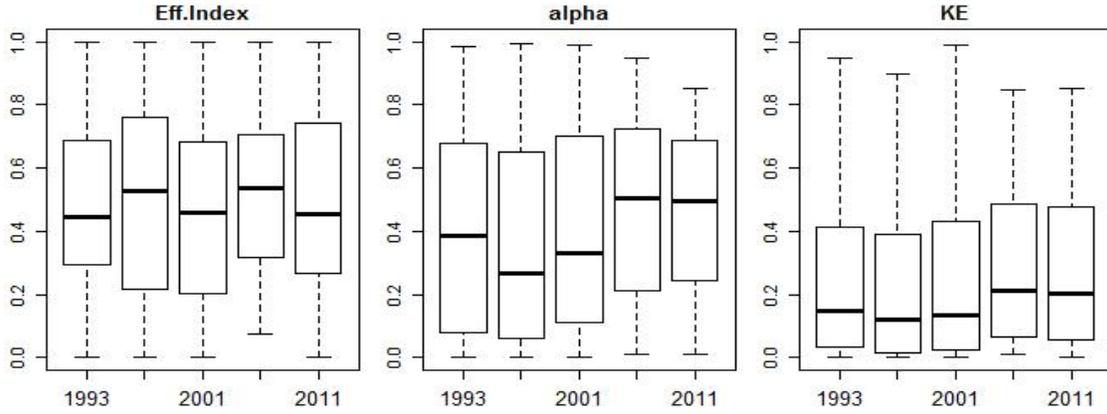


Figure 7: Boxplots of the distributions of the indices α , KE , and $EI = KE/\alpha$ over the regions for each observation year in the Basque Country.

The same analysis for Ireland and Wales show simply in a different way what we already found looking at the tables and figures above, and are therefore skipped.

6 Conclusions and Recommendations

We have built the $x^*(\alpha)$ function that relates each proportion of bilingual speakers, $\alpha \in (0, 1)$, with the corresponding Nash equilibrium proportion of bilingual speakers, x^* , who play the strategy *R - reveal your bilingual type always* - in the Bayesian Language Use Game. We think of this function as a model for the language use measures, KE and DU , of the minority language B . The predicted use of language B , $PKE(\alpha)$ and $PDU(\alpha)$, has two main features: first, it is a strictly increasing convex function on α ; second, for each $\alpha \in (0, 1)$, the equilibrium $x^*(\alpha)$ and, hence the predicted use of B , is evolutionary stable in the associated one-population replicator dynamics.

When we study the data about the actual daily use of Welsh and Irish and the street use of Basque, we observe a relationship between percentage of bilingual speakers α and street use which is as predicted by the theoretical model. Hence, $x^*(\alpha)$ captures the empirical fact that the use of B increases with α . On the other hand, in line with the stability prediction, while the parameters change considerably, when comparing the PKE and PDU forms and locations, they are pretty stable over the years, though not over the countries. However, the latter might be attributed to the different ways of measuring the language use of the minority language in question.

The parametric model has been compared to nonparametric (model-free) fits of the observed language use on their corresponding α . The functions resulting from the theoretical model came astonishingly close to these model-free data fits. This holds also true over time. Since the equilibrium use of language B is evolutionary stable, it could be interpreted as if the bilinguals build linguistic conventions to solve their language coordination problem under imperfect information.

Bilinguals face frequent language choice decisions to coordinate the language with interlocutors of unknown linguistic type. Thus, they are in need of decision procedures to solve fast, without much thinking and no frictions a coordination problem. That kind of procedures takes form when bilinguals reach an evolutionary stable equilibrium or, equivalently, when they build a linguistic convention. Typically, the resulting conventions would be shaped by the language contact situation in which language A is dominant in almost all domains. Empirically, we have shown that Model 2 explains better the data. This finding tells us that the actual street use of Basque is the result of the language behavior of *weak R-players*, who are a type of bilingual speakers who (partly) experiment with strategy H and thus not always signal their true type. The resulting strategy profile (H, H) makes them talk in language A . We do not have data of similar nature as those for Basque, but the daily use of Welsh seems to be the result of similar bilingual speakers' behavior. As for Irish one would say that bilinguals are mainly H -players.

Hence, given the specifications of the models, to the question posed by Fishman (2001) "Why is it so hard to save a threatened language?", we would answer that it is mainly because in a context of language contact and frequent language choices, bilinguals build in the long run linguistic conventions, typically based on the, less effort demanding, strategy H , which is the origin of language coordination failures.

The survival of language B and its related culture depends not on linguistic rights but on the effective use bilinguals make of B . An active language policy should account for this, for example, by promoting the use of these languages in virtually all domains of public life. A main problem is the absence of perfect information mixed with the linguistic politeness (not starting the conversation in A might be considered as impolite or even a source of conflict). This mixture is the main incentive for bilingual speakers to play H (or be a weak R player). Therefore, policy measures might be developed in order to increase bilinguals' use of B by increasing positive signaling where 'positive' means that signals should have a wide social consensus.²⁴ This includes also that policy makers should clarify the implications of language

²⁴In the absence of signals, before starting a conversation, bilinguals should send a message indicating their linguistic type. The content of the message should be well thought so that it should cause no harm or cost to monolinguals in order to respect the politeness convention.

strategies; for example, strategy R has to be related to behavior in favor of *cultural diversity* instead of being a discrimination of monolinguals.

Actually, the linguistic convention introduces a strong stability component into the linguistic behavior of the bilingual population that is hard to break. Roughly speaking, it would be needed political measures to either increase the bilinguals' perceived net benefit of using B or to reduce the imperfect information. A dramatic increase in the proportion α of bilingual speakers is obviously not the key point, as has been proved by comparing Ireland, Wales, and the Basque Country: while the percentage of bilinguals in Ireland doubles the one in the Basque Country, the former is close to extinction while the latter exhibits a pretty stable street use.²⁵ One might speculate that this is because English is the competitor, i.e. the majority language A , is much more dominant (see, for instance, Melitz, 2007, and Ku and Zussman, 2010), what makes it particularly hard for the Irish to survive.²⁶ For this reason we added the Welsh; it has a comparable α like the Basque. Unfortunately, the aggregation level for Wales is too high to draw many conclusions from the model parameter estimates.

References

- Abrams D.M. and Strogatz S.H. (2003) "Modeling the dynamics of language death", *Nature* 424, 900.
- Altuna, O. and Basurto, A. (2013) *A Guide to Language Use Observation. Survey Methods*. Soziolinguistika Klusterra.
- Anderlini, L. and Felli, L. (2004) "Book Reviews: Economics and Language", *Economica* 71, 169-179.
- Blume, A. and Board, O. (2013) "Language Barriers", *Econometrica* 81(2), 781-812.
- Brown, P. and Levinson, S.C. (1987) *Politeness: Some Universals in Language Usage*. (Cambridge: Cambridge University Press).
- Chen, M.K. (2013) "The Effect of Language on Economic Behavior: Evidence from Savings Rates, Health Behaviors, and Retirement Assets", *American Economic Review* 103(2), 690-731.

²⁵Some people might expect it to increase more along α , but recall that the probability of a bilingual random match is α^2 , not just α . In that sense it is correct to say that for a clear increase of KE (or DU) one needs a drastic increase of α .

²⁶Speaking well English would rise the human capital on the international labor market significantly, resulting in higher lifetime income.

- Church, J. and King, I. (1993) "Bilingualism and Network Externalities", *Canadian Journal of Economics* 26(2), 337-345.
- Clinginsmith, D. (2015) "Are the World's Languages Consolidating? The Dynamics and Distribution of Language Populations", *The Economic Journal*, forthcoming.
- Crawford, V. and Sobel, J. (1982) "Strategic information transmission", *Econometrica* 50(6), 1431-1451.
- Demichelis, S. and Weibull, J. (2008) "Language, meaning, and games: a model of communication, coordination, and evolution", *American Economic Review* 98(4), 1292-1311.
- Desmet, K., Ortuño-Ortín, I. and Weber, S. (2009) "Linguistic Diversity and Redistribution", *Journal of the European Economic Association* 7(6), 1291-318.
- Farrel, J. and Rabin, M. (1996) "Cheap talk", *Journal of Economic Perspectives* 10(3), 103-118.
- Fishman, J.A. (1991) *Reversing Language Shift. Theoretical and Empirical Foundations of Assistance to Threatened Languages*. Clevedon, UK: Multilingual Matters.
- Fishman, J.A. (2001) "Why is it so Hard to Save a Threatened Language?" in J.A. Fishman (Ed.) *Can Threatened Languages be Saved?* Clevedon, UK: Multilingual Matters.
- Frank, M.C. and Goodman, N.D. (2012) "Predicting pragmatic reasoning in language games", *Science* 336(6084), 998.
- Gabszewicz, J., Ginsburgh, V. and Weber, S. (2011) "Bilingualism and Communicative Benefits", *Annales d'Economie et Statistique* 101(2), 271-86.
- Ginsburg, V., Ortuño-Ortín, I. and Weber, S. (2007) "Learning Foreign Languages: Theoretical and Empirical Implications of the Selten and Pool Model", *Journal of Economic Behavior and Organization* 64(3-4), 337-347.
- Ginsburgh, V. and Weber, S. (2011) *How Many Languages Do We Need? The Economics of Linguistic Diversity*. Princeton, NJ: Princeton University Press.
- Grin, F., Sfreddo, C. and Vaillancourt, F. (2010) *The Economics of the Multilingual Workplace*. New York, NY: Routledge.

- Härdle, W. and Mammen, E. (1993) "Comparing Nonparametric Versus Parametric Regression Fits", *Annals of Statistics* 21(4), 1926-1947.
- Heller, Y. (2014) "Language, meaning, and games: A model of communication, coordination, and evolution: Comment", *American Economic Review*, 104(6): 1857-63.
- Iriberry, N. and Uriarte, J.R. (2012) "Minority Language and the Stability of Bilingual Equilibria", *Rationality and Society* 2(4), 442-462.
- Isphording, I.E. and Otten, S. (2013) "The costs of Babylon – Linguistic distance in applied economics", *Review of International Economics* 21(2), 354-369.
- Jones, H.M. (2012) *A statistical overview of the Welsh language*. Welsh Language Board, Cardiff.
- Kahneman, D. and Tversky, A. (1979) "Prospect Theory: An Analysis of Decision under Risk", *Econometrica* 47(2), 263-289.
- Ku, H. and Zussman, A. (2010) "Lingua Franca: the Role of English in International Trade", *Journal of Economic Behavior & Organization*, 75(2), 250-260.
- Melitz, J. (2007) "The Impact of English Dominance on Literature and Welfare", *Journal of Economic Behavior & Organization* 64(2), 193-215.
- MERCATOR (2008) *Basque: The Basque language in Education in Spain*. 2nd Edition. www.mercator-research.eu
- MERCATOR (2014) *Welsh: The Welsh language in Education in the UK*. 2nd Edition. www.mercator-research.eu
- Nelde, P.H. (1987) "Language contact means language conflict", *Journal of Multilingual and Multicultural Development* 8(1-2), 33-42.
- Nelde, P.H. (1995) "Language contact and conflict: the Belgian experience and the European Union" In S. Wright and H. Kelly (Eds.) *Languages in Contact and Conflict: Experiences in the Netherlands and Belgium*. Clevedon, UK: Multilingual Matters.
- Patriarca, M. and X. Castelló, Uriarte, J.R., Eguiluz, V.M. and San Miguel, S. (2012) "Modeling Two Language Competition Dynamics", *Advances in Complex Systems*, 15(3-4).
- Rubinstein, A. (2000) *Economics and Language*. Cambridge: CUP.

Selten R. and Pool, J. (1991) "The Distribution of Foreign Language Skills as a Game Equilibrium" in R. Selton ed: *Game Equilibrium Models IV: Social and Political Interaction*, Berlin: Springer-Verlag, 64-87.

Selten, R. and Warglien, M. (2007) "The emergence of simple languages in an experimental coordination game", *Proceedings of the National Academy of Science* 104(18), 7361-7366.

UNESCO (2002). "UNESCO Universal Declaration On Cultural Diversity"

Uriarte, J.R. (2015) "Politeness equilibrium", Mimeo, University of the Basque Country.

Weibull, J.W. (1995) *Evolutionary Game Theory*. Cambridge, Mass.: the MIT Press.

7 Appendix

7.1 Nonparametric Estimation procedure

For the ease of notation we always use KE as observed response variable; for DU the methodology works exactly the same way. Given a sample $\{\alpha_i, KE_i\}_{i=1}^n$ one wants to estimate the conditional expectation $E[KE|\alpha] = m(\alpha)$ under the assumption that $m(\cdot)$ is a smooth function having third order Lipschitz continuous derivatives. The errors $v = KE - m(\alpha)$ have finite variance. One may add some conditions on the distribution of α if one wants to calculate the statistical properties of the now described estimator: For a weight or kernel function $K(\cdot)$ for which we chose the Epanechnikov kernel $K(u) = 0.75 \cdot (1 - u^2)_+$ (the subindex $+$ indicates that the function is set to zero if $1 - u^2$ is negative) and bandwidth h_x we take

$$\widehat{m(x)} = \underset{m, m_1, m_2}{\operatorname{argmin}} \sum_{j=1}^n (KE_j - m - m_1 \cdot (\alpha_j - x) - m_2 \cdot (\alpha_j - x)^2)^2 K\left(\frac{\alpha_j - x}{h_x}\right) \quad (13)$$

as an estimate for $m(x)$. This is the well-known local quadratic kernel estimator. Letting x run over the range of α (here simply over all sample observations α_i) we can draw than the function estimate of $m(\cdot)$ which is compared than with our model for PKE .

7.2 The bootstrap test of Härdle and Mammen (1993)

We want to check the null hypothesis that the parametric model does not significantly deviate from the nonparametric fit which is supposed to reflect the true model

but with a potential smoothing bias. Using the same notation as introduced in the main text, the proposed test statistic is

$$T_{ct} = \frac{1}{n_{ct}} \sum_{i=1}^{n_{ct}} \left(\widetilde{PKE}_{cti} - \widehat{m}(\alpha_{cti}) \right)^2, \quad (14)$$

where $\widehat{m}(\alpha_{cti})$ is the nonparametric data fit of KE on α . Let \widehat{PKE}_{cti} be the parametric prediction along our theoretical model. To avoid potential smoothing bias problems, it is recommended to let it pass through the kernel smoother, too. That is, estimation procedure (13) is applied to $\{\alpha_{cti}, \widehat{PKE}_{cti}\}_{i=1}^{n_{ct}}$, and call the results \widetilde{PKE}_{cti} . To simulate the p-value for test statistic T_{ct} under the null hypothesis one applies wild bootstrap. That is, we keep the α_{cti} but generate new responses by $\widetilde{KE}_{cti} = \widehat{PKE}_{cti} + (KE_{cti} - \widehat{PKE}_{cti}) \cdot N(0, 1)$ (i.e. take the parametrically prediction and add a new normal random term respecting potential heteroscedasticity). Then we calculate the test statistic from this new sample which in fact has been generated under the null hypothesis. This can be done for example a 1000 times. The percentage of these statistics being larger than the original one (14) is a simulated approximate of the p-value of our test.