Working papers series

WP ECON 15.09

Taxation Language learning and communicative benefits

Efthymios Athanasiou (New Economic School, Moscow, Russia)

Juan D. Moreno-Ternero (U. Pablo de Olavide and CORE)

Shlomo Weber (Southern Methodist University, USA and New Economic School, Moscow, Russia)

Keywords: language learning, communicative benefits, language acquisition Nash equilibrium, assignment efficiency, minimal disenfranchisement.

JEL Classification: C72, D62, D63

UNIVERSIDAD PABLO DE OLAVIDE SEVILLA

Department of Economics

Language learning and communicative benefits*

Efthymios Athanasiou† **Juan D. Moreno-Ternero**‡

Shlomo Weber§

June 15, 2015

Abstract

We examine how economic variables impact linguistic diversity and language acquisition. The economic variables in our setting are represented by communicative benefits introduced by Selten and Pool (1991), which subsumes both private monetary rewards and 'pure communicative' benefits of exposure and access to different cultures. The communicative benefits are positively correlated with the number of others with whom individuals can communicate by using one of her spoken languages. Economic examples of communicative benefits are evident in trade, labor market, and migration. We examine the empirical and theoretical literature on language acquisition Nash equilibria and offer an extensive efficiency analysis by using both positive and normative approaches. We also examine various public policies that could enhance the efficiency of selected outcomes.

JEL numbers: *C72, D62, D63.*

Keywords: *language learning, communicative benefits, language acquisition Nash equilibrium, assignment efficiency, minimal disenfranchisement.*

*Thanks are due to Jorge Alcalde-Unzu and Marc Vorsatz for an illuminating conversation on the intricacies of approval voting. Athanasiou and Weber wish to acknowledge the support of the Ministry of Education and Science of the Russian Federation, grant #14.U04.31.0002, administered through the NES Center for Study of Diversity and Social Interactions. Moreno-Ternero acknowledges financial support from the Spanish Ministry of Science and Innovation, grant ECO2011-22919.

†New Economic School, Moscow, Russia.

‡Universidad Pablo de Olavide, Spain and CORE, Université catholique de Louvain, Belgium.

§Southern Methodist University, USA and New Economic School, Moscow, Russia.

1 Introduction

The economic literature has dealt with the way linguistic variables and linguistic diversity impact economic outcomes. The reverse direction that represents how the economic variables impact linguistic diversity, the evolution and the acquisition of languages, is explored to a lesser degree. One can immediately point out that the linguistic diversity and, more specifically, an equal distribution of linguistic skills in multi-lingual societies offers both individual and societal opportunities and challenges. Indeed, there are economic and cultural incentives to acquire languages in addition to one's mother tongue. It may be beneficial for an individual to acquire languages spoken by important trade partners of her own country. If there would be a universal and dominant *lingua franca* spoken by everybody, the incentives of learning foreign languages in our increasingly globalized society would not be that strong. Even though de Swaan (2001) claims that "Globalization proceeds in English", a lot of commercial and trade negotiations in China, India, Latin America and other parts of the world are conducted in languages different from English. China's and India's populations are much larger than that of all English speaking countries combined, and one can wonder why a Chinese businessman would learn English, as he has a direct access to a billion and a half of potential domestic customers. The same applies to individuals in English speaking countries. However, the improving English-language skills in China, and the ever raising demand for studying Chinese in the US and Africa, imply that the language acquisition is enhanced by trade, as Ginsburgh, Melitz and Toubal (2014) demonstrate.

Other examples of economic benefits from learning other languages include direct impact on earnings. Job opportunities are more often open to applicants whose linguistic repertoire includes several languages. In their study of European labor markets, Ginsburgh and Prieto (2011) show that while the foreign language wage premium is almost nonexistent in the UK and quite low in the Netherlands, where the majority of the population is conversant in English, the second language premium is substantial in Austria, Finland, France, Germany, Greece, Italy, Portugal and Spain.[1] Another example is the importance of languages in making immigration decisions. Once in the new country, the

[1] See also Chiswick and Miller (2007), MacManus, Gould and Welsh (1978), Vaillancourt and Lacroix (1985), and Ginsburgh and Weber (2011).

migrant will have to learn, or at least polish the knowledge of the local language to get a job, and thus he faces a learning decision as described above. The importance of linguistic skills for migrants' labor-markets is confirmed by the literature on patterns of language acquisition by immigrants in traditional immigration targets such as Australia, Canada, Germany, Israel, the United Kingdom, and the United States (see Chiswick and Miller, 2014). For the cultural non-economic effects, one may notice that being immersed into a different culture and gaining unfiltered access to its history, arts, and literature in the original language, is worthwhile for some individuals.

From the economic viewpoint, the individuals must evaluate economic benefits of learning other languages and weigh them against the cost of language acquisition, which may include tuition fees charged by schools and payments to private teachers, as well as the opportunity cost of time commited to classes and homework. In evaluating economic benefits of language acquisition, for the large part of this paper we utilize a general concept of *communicative benefits* introduced in the seminal paper by Selten and Pool (1991), which subsumes both private monetary rewards and 'pure communicative' benefits of exposure and access to different cultures. The communicative benefits of every individual are assumed to be positively correlated with the number of others with whom she can communicate by using one of her spoken languages. Naturally, a larger number of people to communicate with raises the attractiveness of the learning of other languages.

The rest of the paper is organized as follows. By relying on the concept of the Selten-Pool communicative benefits we describe and characterize the language acquisition Nash equilibrium that emerges in societies with multiple linguistic groups. We then proceed with the efficiency analysis, which is, in turn, divided into three parts. First, we examine efficiency properties of the equilibrium and discuss public policies, including cost subsidies, for the efficiency enhancement. To complement our efficiency discussion, we extend our search for social beneficial communication patterns within the society and for the socially optimal language acquisition mechanisms. We will then expand our efficiency analysis to include a possibility of selecting official languages for the purpose of efficiency enhancement and increasing communication levels within a society. We conclude with final remarks and the list of references.

2 Communicative benefits

Suppose that the society N, which consists of a finite[2] set of individuals, is partitioned in two linguistic groups, N^E and N^F, the population of which will be denoted by n_E and n_F, respectively. Each citizen speaks the native language of her group, E or F, and, moreover, everybody is unilingual and nobody speaks the language of the other group.

An individual i may choose to learn the foreign language or not. The parameter $a_i \in \{0, 1\}$ captures the individual decision. We write $a_i = 1$ to denote that individual i speaks the foreign language and $a_i = 0$ to denote that she does not. Thus, the vector $a_N \in \{0, 1\}^n$ provides a pattern of language acquisition, or, in simpler terms, determines whether each individual $i \in N$ learns the foreign language or not.

The linguistic composition of the society and patterns of linguistic acquisition determine communicative benefits introduced in Selten and Pool (1991), who cover a wide range of economic, cultural and social advantages gained by learning languages.[3] The societal linguistic data can be conveyed in a succinct manner with the help of a simple drawing. In Figure 1 a set of 5 individuals has been divided in two language groups, $N^E = \{1, 2, 3\}$ and $N^F = \{4, 5\}$. Nodes represent individuals. An arrow stemming from a node and pointing to a set of nodes denotes the fact that the individual that is represented by that node speaks the foreign language. Thus, this simple picture represents both initial conditions and the prevailing pattern of communication. In the situation depicted in Figure 1 only individual 1 learns the foreign language. This means that only 1 among the members of N^E may communicate with each member of N^F and that each member of N^F may communicate with 1, and 1 alone, among the members of N^E.

In broad terms, the communicative benefit assigns a utility value to the partition $\{N^E, N^F\}$ and patterns of language acquisition a_N. In other words, the communicative benefit conveys the individual disposition over situations that can be fully described by pictures that bear the characteristics of the one depicted in figure 1.

[2]The finiteness of the set of individuals is not essential. In fact, some models surveyed in this paper are derived for societies with an infinite number of individuals.

[3]The game-theoretical Selten-Pool model is, in fact, much more general than its variant presented here.

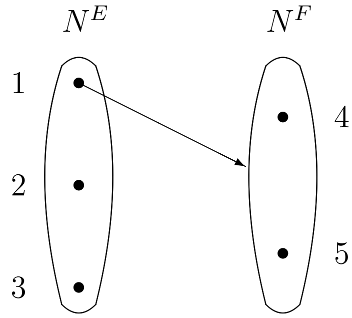

Figure 1

For each $i \in N^E$, let $\theta_i^j \in \mathbb{R}_+$ capture the value that individual i derives form being able to communicate with individual j belonging to the other language group. Let $\theta_i = (\theta_i^1, \ldots, \theta_i^{n_F}) \in \mathbb{R}_+^{n_F}$. In this setting, let $g : \mathbb{R}_+ \to \mathbb{R}_+$ be an increasing function. For each $i \in N^E$ and each $a_N \in \{0,1\}^n$, the communicative benefit of i is given by the expression

$$V_i(a_N) = g\Big(\sum_{j \in N^F} \theta_i^j \min \big\{1, a_i + a_j\big\}\Big).$$

Similarly, for every $j \in N^F$ and each $a_N \in \{0,1\}^n$, the communicative benefit of j is given by the expression

$$V_j(a_N) = g\Big(\sum_{i \in N^E} \theta_j^i \min \big\{1, a_j + a_i\big\}\Big).$$

This formulation accommodates two manifest features of the communicative benefit. Individuals care with whom they communicate. The benefit of communication increases with the number of individuals one is able to communicate with. Each link between two individuals that is generated by the pattern of language acquisition is evaluated separately by the individual to reflect the fact that the possibility of communicating varies across the individuals of the foreign group. However, for each individual, these agent-specific benefits are summed in order to reflect that the incidence of communication has value in itself. Hence, an individual may be driven to learn a foreign language either by some key individual in the foreign group that is of particular significance, or by the mere size of the foreign group. The communicative benefit does not incorporate a utility valuation of the initial conditions, that is the partition of individuals into language groups. Admittedly, this carries value, too. However, the communicative benefit focuses on the evaluation of the pattern of language acquisition and is influenced by the effect of the initial conditions only in so far as they affect the final outcome.

The role of the min operator is to ensure that communication between two individuals via more than one language does not have a bearing on the communicative benefit. Indeed, the value of min $\{1, a_i + a_j\}$ is 1 if at least one of the individuals i and j learns the language of the other group, and zero if both refrain from learning. This formulation treats language as means to an end, rather than an end in itself. Hence, it captures some of the factors (but not all) that drive people to learn foreign languages. The preferences underlying the expression above are effectively defined over the opportunities that language acquisition presents to the individual and their corresponding value. They are not defined over languages themselves.

To complete the specification of the utility function we need to address the issue of the cost language learning entails. For each $\lambda \in \{E, F\}$, let the cost of learning language λ be $c_\lambda \geq 0$. There is no a priori reason for having $c_E = c_F$. A key factor that underlies the cost of learning a foreign language is the linguistic distance between your native and the target language (e.g., Ginsburgh and Weber, 2011). However, this does not necessarily entail that, other things being equal, the cost of learning Spanish if one speaks Portuguese natively is the same as the cost of learning Portuguese for a native Spanish speaker. But individual aptitude has an undisputed influence on the cost. For each $i \in N^E$, let $s_i \in [0, 1]$ be a parameter that captures individual facility with foreign languages. The individual cost of learning language λ can then be captured by the expression $s_i c_\lambda$. Having $s_i = 0$ may reflect a situation where an individual is brought up into a bilingual environment or can learn another language in her sleep. Hence, this modeling choice allows for some form of native bilingualism. Utility is thus given by the expression

$$U_i(a_N) = g\Big(\sum_{j \in N^F} \theta_i^j \min\{1, a_i + a_j\}\Big) - a_i s_i c_F,$$

for each $i \in N^E$. Similarly, for each $j \in N^F$, we have the expression for the net communicative benefit:

$$U_j(a_N) = g\Big(\sum_{j \in N^E} \theta_j^i \min\{1, a_j + a_i\}\Big) - a_j s_j c_E.$$

The existence of linguistic equilibrium under those (and, in fact, more general) conditions has been shown by Selten and Pool (1991). In order to achieve characterization results one needs to impose additional conditions as it has been done in the subsequent literature. Church and King (1993) construct a simplified model, where for all pairs of individuals the value of pairwise communication benefit θ_i^j is independent of i and j and is simply set to one. Moreover, the cost of learning is the same across individuals and language groups. In other words, the utility function for an individual $i \in N^E$ is reduced to:

$$U_i(a_N) = g\Big(\sum_{j \in N^F} \min\big\{1, a_i + a_j\big\}\Big) - a_i c.$$

Under these simplifying assumptions the value of the cost parameter becomes the determining factor in the model and all members of the same group exhibit identical behavior at equilibrium, yielding the existence of corner solutions only. Church and King demonstrate that there are three ranges of the cost parameter c. If this parameter is high then no group learns the language of the other and if the cost is low then everybody learns the language of the other community. The most interesting is the intermediate case when only the members of a smaller community learn the other language, while the members of a larger community refrain from doing so. The explanation is quite obvious; the number of members of the other group justifies the language acquisition for a smaller community but not for a larger one.

The Church and King model has been extended by Ginsburgh, Ortuño-Ortin and Weber (2007) for the case of multiple (more than two) linguistic communities. They show that a unique linguistic equilibrium arises if the function g in the expression for the communicative benefit exhibits super-modularity and individuals share the same preferences for communication. A further extension of this model to the case with heterogeneous learning aptitudes across the country has been extensively studied by Gabszewicz, Ginsburgh and Weber (2011). The assumption that the individuals in the same community face different learning costs, enriches the set of equilibrium outcomes. In fact, they consider the linear functional form of communication benefits for each member $i \in N^E$ and $j \in N^F$, respectively:

$$U_i(a_N) = \sum_{j \in N^F} \theta_i^j \min\big\{1, a_i + a_j\big\} - a_i s_i c_F,$$

$$U_j(a_N) = \sum_{j \in N^E} \theta_j^j \min\big\{1, a_j + a_i\big\} - a_j s_j c_E,$$

where both values s_i and s_j are distributed uniformly over the interval $[0, 1]$.

In this framework, at equilibrium, some (yet not all) individuals in a given language group may opt to learn the foreign language. The size of each language group along with the profile of learning costs fully determine which equilibrium outcome will prevail. Gabszewicz, Ginsburgh and Weber (2011) find that such interior equilibria are often unstable and outline the circumstances under which the stability of equilibrium is guaranteed. They also describe a wide range of comparative statics results that evaluate the dependence of the share of learners with respect to different parameters of the model. In particular, the number of learners declines if the cost of learning the other language raises. Probably, less obvious is the fact that the number of learners in the community is positively correlated with the learning costs of their own language. Indeed, the effect is secondary: the raise of the learning costs of language E reduces the number of learners in F. Thus, the acquisition of F becomes more attractive for members of E, as by learning of F they receive an access to a larger group of unilingual members of F. Also, an increase in population of a group makes its language more attractive to acquire.

This model was taken to data by Ginsburgh, Ortuño-Ortin and Weber (2007), and Ginsburgh, Melitz and Toubal (2014). The latter paper estimates an equation that relates learning decisions of 13 of the most important world languages by citizens who live in some 190 countries and find that learning is influenced positively by trade, which exhibits a special form of communicative benefits.

3. Efficiency

As all papers above recognize, an inherent component of the communicative benefit is that it induces a network externality. Let us assume that individuals only care about the amount of communication they may achieve, although their general disposition towards communication may vary. Let us also assume that all individuals have the same ability to learn the foreign language. For each $i \in N^E$, each $j \in N^F$ and each $a_N \in \{0,1\}^n$, the expression for the individual utility of i thus reduces to

$$U_i(a_N) = \theta_i \sum_{j \in N^F} \min\left\{1, a_i + a_j\right\} - a_i c_F,$$

and

$$U_j(a_N) = \theta_j \sum_{i \in N^E} \min\left\{1, a_j + a_i\right\} - a_j c_E.$$

Consider the following example. Let $N^E = \{1, 2\}$, $N^F = \{3, 4, 5\}$ and $\theta_N = \big[(1,1)(0,0,0)\big]$. For some arbitrarily small $\epsilon > 0$, let $c_E = n_F + \epsilon$ and $c_F = \epsilon$. Each individual's pure strategy is simply a decision on whether to learn the foreign language or not. For each $a_N \in \{0,1\}^n$ and each $i \in N^E$,

$$\text{if } a_i = 1 \text{ then } v_i(a_N) = \theta_i \sum_{j \in N^F} \min\left\{1, 1 + a_j\right\} - c_F = -\epsilon.$$

Therefore, each individual in N^E has a dominant strategy not to learn. The same applies to individuals in the group N^F.

In stark contrast, the linguistic assignment that maximizes the sum of utilities is $a^*_N = \big[(0,0)(1,1,1)\big]$. If this pattern of communication were to prevail, all individuals would be communicating with each other. For efficiency to come about though, a mode of sharing the benefits of communication needs to be established.

Communication externalities may also take the form of mis-coordination. Consider the following example. Let $N^E = \{1, 2\}$, $N^F = \{3, 4\}$ and $\theta_N = \big[(1,1)(1,0)\big]$. For some arbitrarily small $\epsilon > 0$, let $c_E = c_F = 1 + \epsilon$. There are two Nash equilibria: $a^*_N = \big[(1,1)(0,0)\big]$ and $a^{**}_N = \big[(1,0)(1,0)\big]$. Only the first among the two corresponds to the optimal social outcome.

Externalities are inherent in the nature of communication. Consequently, the pattern of communication that emerges in groups of individuals that act on their self-interest fails to achieve the social optimum.

3.1. Efficiency of Equilibria

Let us first examine the case of identical learning cost studied in Church and King (1993). If the social planner cannot discriminate between different communities and individuals, we can still examine whether the equilibrium satisfies *constrained efficiency*.

Note that if there are two Nash equilibria where one of the groups learns the language of the other, only one could be efficient. A smaller community should engage in learning, while the larger one should not do so. Suppose, for the sake of presentation, that the population n_E is larger than n_F. Then, we have three ranges of the cost of learning c:

(I) the costs are high so that no learning occurs in equilibrium, which is also a constrained-efficient outcome;

(II) the costs are low so that the situation where only community F studies the language of E is both an equilibrium and constrained-efficient allocation;

(III) in the intermediate cost range the learning of the language of E by community F is efficient, whereas no learning emerges as a Nash equilibrium. The divergence between the efficient and equilibrium outcomes is rooted in network externalities. When an individual makes a language acquisition choice, she does not take into account that her potential learning of the other language allows others to communicate with her and to raise the total level of societal communicative benefits.

One of the possible remedies for reducing the gap between efficiency and equilibrium is the subsidization of learning costs across the society. By offering tuition reductions (in the case of schools) the government could shrink the range of cost values that yield the "no learning" equilibrium and expand the range where equilibrium and efficiency requirements produce the same outcome where the community F learns the language of E.

The same subsidization policy could be applied in the case of a heterogeneous distribution of learning aptitudes. As Gabszewicz, Ginsburgh and Weber (2011) point out, in the interior equilibrium where only a part of individuals in each community learns the other language, the number of learners is suboptimal and insufficient. The reason is exactly the one mentioned above. In presence of network externalities, the best response considerations do not take into account the impact of learning choices of an individual on members of the other community. The latter group would obviously benefit if the other group learns. In absence of a direct mechanism to enforce the learning, the cost subsidization policies could be a step in the right direction of the societal communicative benefit.

3.2 The Planner's Problem

The purpose of this section is to explore the scope of public intervention in the process of language acquisition. In turn, we discuss the reasons that legitimize public policy, the constraints that incentives place on its design, and the various objectives it should strive to accomplish.

A benevolent Planner aims at bringing about desirable social outcomes. He operates under constraints. Principal among them is the necessity to accommodate individual incentives. The Planner's problem allows for more possibilities if individuals have quasi-linear preferences. We will assume that this is the case. Hence, allowing for the possibility of an individual transfer $t_i \in \mathbb{R}$, the final utility of each agent $i \in N$, at assignment $a_N \in \{0,1\}^N$, becomes

$$u_i(a_N, t_i) = v_i(a_N) + t_i.$$

The problem the Planner faces amounts to assigning a linguistic assignment coupled with a profile of transfers to each possible economy $e = \big((\theta_{N_\alpha}, \theta_{N_\beta}); (c_\alpha, c_\beta)\big)$ in the domain $\{0,1\}^n \times \mathbb{R}^2_+$. A mechanism is a function

$$\varphi : \{0,1\}^n \times \mathbb{R}^2_+ \to \{0,1\}^n \times \mathbb{R}^n_+.$$

A mechanism can be construed as the Planner's prescription. It assigns an outcome to each contingency that may arise. Hence, $\varphi(e) = \big(a_N, (t_1, \ldots, t_n)\big)$, while the individual bundle prescribed to $i \in N$ by φ is denoted by $\varphi_i(e) = (a_i, t_i)$.

3.3. Efficient Assignments

The main reason for studying the Planner's problem stems from the fact that decentralized outcomes are generically inefficient. It is therefore legitimate to wonder whether public intervention may bring about a welfare improvement or even ensure the implementation of efficient outcomes. A linguistic assignment is *efficient* if it maximizes the sum of net communication benefits.

Depending on the particular list of initial parameters at hand, we need to solve the following optimization problem:

$$P(e): \max_{a_N \in \{0,1\}^n} \Big[\sum_{i \in N^E} \Big(\theta_i \big(\sum_{j \in N^F} \min\{1, a_i + a_j\} \big) - a_i c_F \Big) + \sum_{j \in N^F} \Big(\theta_j \big(\sum_{i \in N^E} \min\{1, a_i + a_j\} \big) - a_i c_E \Big) \Big].$$

For each $e \in \{0,1\}^n \times \mathbb{R}^2_+$, let $\Sigma(e)$ be the set of assignments that solve $P(e)$. This is a discrete optimization problem. Athanasiou, Dey and Valletta (2013) provide an algorithm that produces for each $e \in \mathcal{E}$, one $a_N \in \Sigma(e)$. For each $e \in \mathcal{E}$, there are at most $2^{|N|}$ candidate solutions. As the set of candidate solutions is finite, for each $e \in \mathcal{E}$, $\Sigma(e) \neq \emptyset$.

A naive algorithm that solves $P(e)$ enumerates all the candidate solutions. Athanasiou Dey and Valletta propose an algorithm that runs in polynomial time, that is, it enables a computer to solve the problem quickly, far more so compared to the naive algorithm. However, this depends on the assumption of two language groups. For an arbitrary number of language groups their algorithm becomes computationally cumbersome.

Efficient linguistic assignments take various forms depending on the economy. Generically, optimality does not involve all members of a language group to learn. What is more, at an efficient assignment people from either language group may learn. These are not eccentric possibilities. To the contrary, they are prevalent as long as the number of individuals in both groups remains finite. Quite naturally, in large populations, one-sided learning prevails.

These possibilities already allude to the fact that the implementation exercise is involved. A mechanism satisfies efficiency if it always selects an efficient linguistic assignment.

Assignment Efficiency: For each $e \in \mathcal{E}$, if $(a_N, t_N) = \varphi(e)$ then $a_N \in \Sigma(e)$.

3.4 Implementing Efficient Linguistic Outcomes

A natural question to raise is whether a mechanism φ that satisfies *Assignment Efficiency* may also accommodate incentives.

A mechanism can be manipulated in three distinct ways. First, the individual may misreport relevant information she holds private. In particular, she may lie about her disposition towards communication. Second, the individual may choose not to conform to the prescriptions of the mechanism. Third, the individual may refuse to participate. A mechanism is effectively implementable if it is impervious to these forms of manipulation.

First, under all possible circumstances, the mechanism should induce all individuals to reveal their willingness to communicate truthfully independently of whether others choose to do so. This property, called *Strategy-Proofness*, facilitates the implementation exercise in a fundamental way. The Planner does not need to know neither the exact type of each individual, nor the distribution of types across the population.

Strategy − Proofness: For each $e \in \mathcal{E}$, $i \in N$ and $\theta'_i \in \mathbb{R}_+$,

$$u_i(\varphi_i(\theta_N, c_E, c_F); \theta_i) \geq u_i(\varphi_i(\theta'_i, \theta_{N\setminus\{i\}}, c_E, c_F); \theta_i).$$

Second, the mechanism should ensure compliance by all individuals involved. There are two ways in which an individual may exhibit non-compliance. First, he may choose to ignore a prescription to learn. Second, he may learn despite being prescribed not to (he may be 'home-schooled'). Between the two, the latter is significantly more costly to ward against. However, neither of the two possibilities transpire if the mechanism is efficient. If an individual finds it profitable to unilaterally deviate from the linguistic assignment bearing the full cost of her action, then the assignment cannot be optimal. Therefore, *Assignment Efficiency* bears also an incentive justification. It facilitates compliance and reduces the cost of implementation.

Finally, participation to the mechanism needs to be voluntary. The mechanism should refrain from coercion. This is a minimal legitimacy requirement.

Individual Rationality: For each $e \in \mathcal{E}$ and $i \in N$, $u_i(\varphi_i(e); \theta_i) \geq 0$.

Unfortunately, any mechanism that complies with these three properties runs a deficit. This deficit can be construed as a measure of the cost communication externalities place on society. This result can be traced back to Green and Laffont (1979). Athanasiou, Dey and Valletta (2013) provide a proof that is context specific. Moreover, building on results by Krishna and Perry (1997), they show that the phenomenon is quite prevalent by providing a set of sufficient conditions for a deficit to ensue.

Proposition 1. *If a mechanism φ satisfies* Strategy-Proofness*,* Assignment Efficiency *and* Individual Rationality*, then there exists $e \in \{0,1\}^n \times \mathbb{R}^2_+$ such that $\varphi(e) = \big(a_N, (t_1, \ldots, t_n)\big)$ and $\sum_{i \in N} t_i > 0$.*

Thus the cost of implementation takes the form of a loss in welfare that can be readily expressed in pecuniary form. As a phenomenon, the incidence of the deficit constitutes a concrete consequence communication externalities bear on policy. In its intensity it captures the barriers policy faces and may explain the circumstances that impede patterns of language acquisition that enable communication among all.

4. Efficient choice of official languages

Another policy instrument of raising the level of communicative benefit within a multilingual society is the introduction of a relatively small number of official languages,[4] that are used in education, courts, and media (see Pool, 1991). The knowledge of common languages enhances economic, cooperation and cross-regional economic links, enhances economic efficiency and often leads to strengthening of national cohesiveness. However, there is another side to the choice of official languages. A person whose native language is not accepted as an official language can feel *disenfranchised* (Ginsburgh and Weber, 2011). As Pool (1991, p. 495) points out, "[t]hose whose languages are not official spend years learning others' languages and may still communicate with difficulty, compete unequally for employment and participation, and suffer from minority or peripheral status."

Thus, in order to consider the introduction of official languages as as efficiency-enhancing device that raises the level of communicative benefits, one has to carefully examine the selection of official languages. The purpose of this section to place the choice of official language on fair and attractive foundations, by focusing on commonality of agents and mitigating the negative aspects of unavoidable restriction on the number of official languages. We also like to point out the model also applies to establish conventions for the choice of language in a conversation involving a group of multinational agents. In this sense it enriches the setting of Section 2 to account for the existence of a larger number of languages.

Let $M = \{1, 2, \ldots, m\}$ be the set of all languages spoken in the society N. If individual $i \in N$ speaks language $l \in M$, we write $a_{il} = 1$. Otherwise, we write $a_{il} = 0$.[5] As before, we assume that there is no distinction between speaking a language well or not, or between native and non-native languages. Let $A = (a_{il})_{(i,l) \in N \times M}$ denote the resulting $0-1$ matrix that hence summarizes the multilingual reality of the society N. Let $\mathcal{A}$ denote the set of such matrices.

A rule $R : \mathcal{A} \to M$ is a mapping selecting for each matrix A a given language $l \in M$.

[4] Even though there are eleven official languages in South Africa.

[5] Note that in presence of two languages only, as in Section 2, the notation a_i is sufficient to indicate whether individual i speaks another language. In the multi-lingual setting we need to specify whether individual i speaks language l, which necessitates the double index notation.

That is, we assume from the outset that rules only yield *unary* sets. In order to address the subsequent issue of tie breaks, which arises from this assumption, we assume the existence of a given (exogenous) strict ordering $\succ$ among languages. We refer to $\succ$ as the *tie-breaking* rule.

The first obvious rules that come to mind are the so-called dictatorial ones. A *language dictatorship* would always select a given language, whereas an *agent dictatorship* would always select the language spoken by a given agent, provided the tie-breaking rule introduced above.[6] The dictatorial rules would be in conflict with standard formalizations of the principle of impartiality, which refers to the fact that ethically irrelevant information is excluded from the evaluation process. More precisely, agent dictatorships would violate the axiom of *anonymity*, which says that the identity of agents should not matter and the choice of an official language should not depend on who speaks them. Language dictatorships, on the other hand, would violate the natural counterpart axiom of anonymity, *neutrality*, which says that the name of languages should not matter either. Neutrality expresses a symmetric treatment of languages. There are compelling reasons in our setting to endorse a symmetric treatment of agents (as formalized by the axiom of impartiality), but not necessarily of languages. For instance, one might want to provide a slight advantage to the language identified in the status quo. Somewhat related, one might want to break hypothetical ties in a precise way, given our assumption of unary images of rules, i.e., that there exists a unique official language. For these reasons, we shall define *weak neutrality* as the property of neutrality, when restricted to the case in which there are no ties to break. Language dictatorships will also violate this axiom and, hence, we shall discard them from our analysis.

Another axiom with normative appeal in this context is *monotonicity*, which establishes the fact that if a language is selected, it would also be selected after one agent, who did not speak it before, learns it, ceteris paribus.

As stated below, the three previous axioms together characterize a unique rule. This is the so-called *minimal disenfranchisement rule*, which selects a language excluding (disenfranchising) the lowest possible number of agents.[7] Formally,

[6]Note that the tie-breaking rule could be considered in itself as a language dictatorship.

[7]The name is inspired by the notion described by Ginsburgh, Ortuño-Ortin and Weber (2005).

MD: For each $A \in \mathcal{A}$, let

$$A^m = \arg\max \left\{ \sum_{i \in N} a_{il} : l \in M \right\}.$$

Then, $MD(A) = l^\star$, where $l^\star \in A^m$ is such that $l^\star \succ l$ for each $l \in A^m \setminus \{l^\star\}$.

In words, MD is the rule selecting the language with a highest number of speakers (*score*). If several languages have the same score, we resort to the given (exogenous) strict ordering among languages mentioned above, as a tie-breaking norm.[8]

We have the following characterization result:[9]

Proposition 2. *The minimal disenfranchisement rule is the only rule satisfying anonymity, weak neutrality and monotonicity.*

It is worth mentioning that the minimal disenfranchisement rule also selects in this context the language that maximizes communication benefits, as described in the previous section, i.e., the number of pairs of agents that can communicate with a given language.

Our (benchmark) model just described could be considered as a model of collective choice under dichotomous preferences. The contrast between our model and the seminal work on the topic, by Bogomolnaia, Moulin and Stong (2005), is that we consider a specific tie-breaking norm to obtain deterministic rules, whereas they endorse a probabilistic approach and consider lotteries/time sharing.[10] More precisely, the alternative to our MD rule in Bogomolnaia, Moulin and Stong (2005) is the so-called utilitarian mechanism, defined as the uniform lottery over the subset of outcomes liked (languages spoken, in our model) by the largest number of agents. Such a mechanism is anonymous, neutral, strategy-proof and ex-ante efficient (although it is not characterized by such a combination of properties). Somewhat related, our minimal disenfranchisement rule is a specific instance of approval voting (e.g., Brams and Fishburn, 1978), with the proviso that only one alternative is selected (based on our tie-breaking norm). If no specific cardinality is imposed into the definition of a social choice function, then approval voting

[8]Note that $\sum_{i \in N} a_{il}$ yields the number of agents speaking language l, i.e., the 1 entries in the l-th column of A. Obviously, $n - \sum_{i \in N} a_{il}$ yields the number of agents not speaking (disenfranchised by) language l, i.e., the 0 entries in the l-th column of A.

[9]A proof can be found in Moreno-Ternero and Weber (2015).

[10]See also Bogomolnaia and Moulin (2004).

is characterized in this context by anonymity, neutrality, strategy-proofness and strict monotonicity (e.g., Vorsatz, 2007). Massó and Vorsatz (2008) characterize a family of rules generalizing approval voting, upon weakening the neutrality condition and, hence, allowing for heterogeneous importance (modeled as, possibly different, weights) of the alternatives (languages in our case). It is worth mentioning that they work on the general domain of preferences (not necessarily dichotomous). The family of *weighted approval voting* they characterize contains as a specially interesting subclass those lexicographic voting rules that always choose a unique alternative (except when no alternative receives any vote) by applying first approval voting (all weights are the same) and selecting afterwards, among the subset of alternatives with maximal support, the unique alternative that maximizes a given strict order. That is, precisely, our minimal disenfranchisement rule. They characterize the family by means of the axioms of consistency in alternatives (the analogue of Arrow's Choice Axiom; which has a questionable normative appeal in the context we consider here) and voters (which requires that if two disjoint electorates elect a common set out of two feasible alternatives, then exactly this set has to be elected when the two electorates are assembled), anonymity, coherence (which asks that for every alternative, when confronted with another alternative, there must exist a situation, with strictly positive votes for both alternatives, at which the considered alternative is elected, perhaps together with the other one), and the no-support condition (one alternative without any vote is elected, when confronted with another alternative, if and only if the second alternative does not receive any vote either). However, they offer no additional axiom to single-out the above rule that would be a counterpart of our minimal disenfranchisement rule.

The problem of the choice of official languages becomes more complex when one extends the benchmark model presented before in either of several plausible directions. One plausible extension (that we shall not consider here) is that in which a distinction between speaking a language well or not (or between native and non-native languages) is made. Another plausible extension is to analyze the case in which several official languages are allowed. In what follows, we elaborate for the case of two official languages.

Let M^2 be the set of pairs of different languages in M. Now, a rule $R : \mathcal{A} \to M^2$ selects for each matrix A a pair of languages $\{l, q\} \in M$. We shall name those rules as *pair* *rules*, although we might refer to them simply as rules if there is no possible confusion.

A natural starting point is to propose pair rules that extend to this setting the minimal disenfranchisement rule highlighted above. Two plausible options arise. One refers to selecting the pair of languages such that the number of agents speaking none of them is as low as possible. Another refers to selecting the pair of languages with the highest individual scores. Furthermore, the former option suggests a somewhat dual rule, which would select the pair of languages such that the number of agents speaking both of them is as high as possible. As a matter of fact, the three rules just outlined could be seen as members of a general family of *pair-scoring rules*. In order to introduce it formally, we need a piece of additional notation. Let $f : \{0,1\}^2 \to \mathbb{R}$ be a non-decreasing and symmetric function, i.e., $f(0,0) \leq f(1,0) = f(0,1) \leq f(1,1)$. Let $\mathcal{F}$ denote the set of such functions.

D^f: For each $f \in \mathcal{F}$, $A \in \mathcal{A}$, and $\{l,m\} \in M^2$, let

$$\rho_f(l,m) = \sum_{i \in N} f\{a_{il}, a_{im}\},$$

and

$$A^{\rho_f} = \arg\max\left\{\rho_f(l,m) : \{l,m\} \in M^2, l \succ m\right\}.$$

Then, $D^f(A) = \{l^\star, m^\star\}$, where $(l^\star, m^\star) \in A^{\rho_f}$ is such that either $l^\star \succ l$ for each $(l,m) \in A^{\rho_f} \setminus \{(l^\star, m^\star)\}$, or $m^\star \succ m$ for each $(l^\star, m) \in A^{\rho_f} \setminus \{(l^\star, m^\star)\}$.

In words, each rule of the family selects the two languages with a highest score, according to the corresponding function f (with ties broken via $\succ$). Note that if $f(\cdot,\cdot) \equiv \max(\cdot,\cdot)$ then $D^f \equiv D^{\max}$ is the first rule suggested above. Similarly, if $f(\cdot,\cdot) \equiv \sum(\cdot,\cdot)$ then $D^f \equiv D^{sum}$ is the second rule suggested above. Finally, if $f(\cdot,\cdot) \equiv \min(\cdot,\cdot)$ then $D^f \equiv D^{\min}$ is the third rule suggested above.

We now introduce another pair rule, inspired by the notion of communicative benefits.

D^{SP}: For each $A \in \mathcal{A}$, and $\{l,m\} \in M^2$, let

$$\rho_{SP}(l,m) = \left|\left\{(i,j) \in N^2 : \text{ there exists } \lambda \in \{l,m\} \text{ such that } a_{i\lambda} = a_{j\lambda} = 1\right\}\right|,$$

and

$$A^{\rho_{SP}} = \arg\max\left\{\rho_{SP}(l,m) : \{l,m\} \in M^2, l \succ m\right\}.$$

Then, $D^{SP} = \{l^\star, m^\star\}$, where $(l^\star, m^\star) \in A^{\rho_{SP}}$ is such that either $l^\star \succ l$ for each $(l, m) \in A^{\rho_{SP}} \setminus \{(l^\star, m^\star)\}$, or $m^\star \succ m$ for each $(l^\star, m) \in A^{\rho_{SP}} \setminus \{(l^\star, m^\star)\}$.

In words, D^{SP} selects the pair of languages guaranteeing the highest number of pairs of agents that can communicate thanks to a language in the pair.

The following example shows that D^{SP} is not equivalent to any of the three other focal rules presented above.[11]

$$A = \begin{pmatrix} 0 & 1 & 1 \\ 0 & 1 & 1 \\ 1 & 0 & 1 \\ 1 & 0 & 0 \\ 0 & 1 & 1 \end{pmatrix}$$

It is straightforward to see that $D^{SP}(A) = \{1, 3\}$, whereas $D^{sum}(A) = \{2, 3\} = D^{\min}(A)$, and $D^{\max}(A) = \{1, 2\}$, provided $2 \succ 3$.

All the rules presented above satisfy the extensions to this setting of the three axioms characterizing the minimal disenfranchisement rule for the case of the choice of a unique official language. The following axioms allow to distinguish among those rules.

First, we consider the axiom of *consistency*, stating that the (pair) rule always selects as a member of the pair the language selected by MD for the same problem. In other words, the rule consistently extends MD to the setting of pairs of languages.[12] It is straightforward to show that D^{sum} and D^{SP} obey consistency. The following examples show that neither $D^{\max}$, nor $D^{\min}$, do so.

[11] As a matter of fact, D^{SP} is not a pair-scoring rule.

[12] The axiom is inspired by a notion that has played a fundamental role in axiomatic work, and for which normative underpins have been provided (e.g., Thomson, 2012).

$$A^1 = \begin{pmatrix} 1 & 1 & 0 \\ 1 & 1 & 0 \\ 1 & 0 & 1 \\ 1 & 0 & 1 \\ 0 & 0 & 1 \\ 0 & 1 & 0 \end{pmatrix} \qquad A^2 = \begin{pmatrix} 1 & 0 & 0 \\ 1 & 0 & 0 \\ 1 & 0 & 0 \\ 1 & 0 & 0 \\ 0 & 1 & 1 \\ 0 & 1 & 1 \\ 0 & 1 & 1 \end{pmatrix}$$

The next axiom, *no redundancy*, states that if an agent speaks one of the two selected languages, the knowledge of the second does affect the selection of the pair. It is straightforward to show that $D^{\max}$ obeys no redundancy. The following example shows that neither D^{sum}, nor $D^{\min}$, do so.

$$A = \begin{pmatrix} 1 & 1 & 0 \\ 1 & 1 & 0 \\ 1 & 0 & 0 \\ 1 & 0 & 1 \\ 0 & 0 & 1 \\ 0 & 0 & 1 \end{pmatrix}$$

Similarly, the following example shows that D^{SP} also violates no redundancy.

$$A = \begin{pmatrix} 0 & 1 & 1 \\ 0 & 1 & 1 \\ 1 & 0 & 1 \\ 1 & 0 & 0 \\ 0 & 1 & 1 \end{pmatrix}$$

The normative appeal of the no redundancy axiom might be questionable. This motivates an alternative weaker variant, known as *evenness*, which says that if the pair of languages is selected under an uneven distribution of language proficiency, it would also be selected when this distribution is more even. The previous example also shows that $D^{\min}$ violates it.

5. Conclusion

We have focussed in this paper on a relatively unexplored topic dealing with the economic incentives of learning foreign languages. This is in contrast to the main bulk of economic research on languages, which targets (mostly, from an empirical angle) the impact of various facets of linguistic diversity on economic outcomes.

Here we use the Selten-Pool notion of communicative benefits to establish equilibrium patterns of language acquisition, and describe theoretical and empirical finding in this relatively new field of research. It is important to point out the explanation of why equilibrium outcomes often fail the efficiency (and even the constrained efficiency) test. The problem lies in network externality: in making her language acquisition decision, an individual does not take into account how the level of communicative benefits of others is affected by her own decision. We then proceed with the discussion of possible public policies that can reduce the gap between equilibrium and desired efficiency. Our analysis of efficiency leads us identifying the requirements that guarantee socially beneficial communication patterns and optimal language acquisition mechanisms within a society.

We augment our quest for raising the level of communicative benefits (and underlined economic efficiency) by analyzing optimal choices of sets of official languages that facilitate economic cooperation, cultural exchange and enhanced trade relationship with a country and with other trading partners. However, the selection of official languages is a delicate and often daunting task, due to, according to Pool (1991), divisive, symbolic and contentious nature of language conflict, inherent incompatibilities between language communities, reluctance of the majority group to concede linguistic rights to minorities, and the power of civil servants to protect their linguistic privileges. The paper therefore lays down fair and sound principles that help to overcome those difficulties.

We would like to point out that most of the issues described here are relatively new. They therefore require a further research effort, both theoretical and empirical, to address the patterns of language acquisition emerging in our rapidly changing economic and social environment. The tools described here could be also useful for examining economic benefits and costs of linguistic policies in several countries in Europe, South and North America, and Africa.

References

[1] J. Alcalde-Unzu and M. Vorsatz (2009) 'Size approval voting', *Journal of Economic Theory*, 144, 1187-1210.

[2] E. Athanasiou, S. Dey and G. Valletta (2013) 'Groves mechanisms and communication externalities', Mimeo.

[3] A. Bogomolnaia and H. Moulin (2004) 'Random matching under dichotomous preferences', *Econometrica*, 72, 257-279.

[4] A. Bogomolnaia, H. Moulin and R. Stong (2005) 'Collective choice under dichotomous preferences', *Journal of Economic Theory*, 122, 165-184.

[5] S. Brams and P. Fishburn (1978) 'Approval voting', *American Political Science Review*, 72, 831-847.

[6] J. Church and I. King (1993) 'Bilingualism and network externalities', *Canadian Journal of Economics*, 26, 337-345.

[7] A. de Swaan (2001) Words of the World, Polity Press: Cambridge.

[8] J. Gabszewicz, V. Ginsburgh and S. Weber (2011) ' Bilingualism and communicative benefits', *Annals of Economics and Statistics*, 101, 271-286.

[9] B. Chiswick and P. Miller (2007) The Economics of Language, International Analyses, Routledge: London and New York.

[10] B. Chiswick and P. Miller (2014), International migration and economics of language, IZA Discussion paper # 7880.

[11] V. Ginsburgh, I. Ortuño-Ortin and S. Weber (2005) 'Language disenfranchisement in linguistically diverse societies. The case of European Union', *Journal of the European Economic Association*, 3, 946-965.

[12] V. Ginsburgh, I. Ortuño-Ortin and S. Weber (2007) 'Learning foreign languages. Theoretical and empirical implications of the Selten and Pool model' *Journal of Economic Behavior and Organizations*, 64, 337-347.

[13] V. Ginsburgh and J. Prieto (2011) 'Returns to foreign languages of native workers in the European Union', *Industrial and Labor Relations*, 64, 599-618.

[14] . Ginsburgh, J. Melitz and F. Toubal (2014) 'Foreign language learning: An econometric analysis', Manuscript.

[15] V. Ginsburgh and S. Weber (2011) 'How Many Languages Do We Need? The Economics of Linguistic Diversity', Princeton University Press, Princeton.

[16] J. Green and J. Laffont (1979) 'Incentives in public decision making', North-Holland, Amsterdam.

[17] V. Krishna and M. Perry (1997) 'Efficient mechanism design' Unpublished manuscript, Pennsylvania State University

[18] J. Massó and M. Vorsatz (2008) 'Weighted approval voting', *Economic Theory*, 36, 129-146.

[19] W. MacManus, W. Gould, and F. Welsch (1978), Earnings of Hispanic men: The role of English language proficiency, *Journal of Labor Economics*, 1, 101-130.

[20] J. Moreno-Ternero and S. Weber (2015) 'An axiomatic approach to the choice of official languages', Mimeo.

[21] J. Pool (1991) 'The official language problem', *American Political Science Review* 85, 495-514.

[22] R. Selten and J. Pool (1991) 'The distribution of foreign language skills as a game equilibrium' in Selten, R. (Ed.). Game Equilibrium Models, Vol. 4. Berlin: Springer-Verlag, 64-84.

[23] W. Thomson (2012) 'On the axiomatics of resource allocation: interpreting the consistency principle', *Economics and Philosophy*, 28, 385-421.

[24] F. Vaillancourt and R. Lacroix (1985) 'Revenus et langue au Québec', http://www.cslf.gouv.qc.ca/publications/PubD120/D120-6.

[25] M. Vorsatz, (2007) Approval voting on dichotomous preferences, *Social Choice and Welfare*, 28, 127-141.