# LANGUAGE, LEARNING, AND LOCATION

ANDREW JOHN
Department of Economics
University of Virginia
Charlottesville, VA 22903

KEI-MU YI
International Research
Federal Reserve Bank of New York
33 Liberty St.
New York, NY 10045
and
Department of Economics
Rice University
6100 S. Main St.
Houston, TX 77005

October 1995
This version: February 1997

We wish to thank: Paul Beaudry, John Bryant, Russ Cooper, Tina Hickey, Dae Il Kim, Bob King, Ian King, Nobu Kiyotaki, Narayana Kocherlakota, Kevin Lang, Ed Lazear, Peter Mieszkowski, Joyce Sadka and Robert Tamura for useful discussion; seminar participants at Australian National University, Bond University, Georgetown University, University of Hawaii, University of Helsinki, Hong Kong University of Science and Technology, University of Melbourne, University of New South Wales, University of Texas, Virginia Tech University, the 1995 Midwest Macroeconomics Conference, the 1996 NBER Winter Economic Fluctuations Conference, and the 1997 AEA Meetings, for useful comments; and Angela Hung, Fidel Perez, and Dana Rapoport for excellent research assistance. The views expressed here are those of the authors and are not necessarily reflective of the Federal Reserve Bank of New York or the Federal Reserve System. We also wish to thank the Rice University Center for the Study of Institutions and Values and the University of Virginia Sesquicentennial Program for financial support. Errors are our own.

## ABSTRACT

Language is a fundamental tool for communication of ideas between people, and so is an essential input into production and trade. In general, a society will possess more production and consumption opportunities when all its members share a common language. Neighboring societies and communities likewise have a strong incentive to utilize a common language, and indeed there are countless examples of language assimilation, especially in the last one hundred years. Hence, it is puzzling that more assimilation has not occurred. History has recorded numerous examples of communities that coexist with distinct languages and limited economic interaction.

This paper presents a stylized model to reconcile both assimilation and non-assimilation. We abstract from cultural and historical factors, which are of course significant, but are present in both. The model has two languages, two locations, and two time periods. Agents are initially endowed with one or both languages and a location. Agents choose whether or not to learn the other language, and subsequently choose whether or not to move to the other region. Language facilitates production: an agent can produce output only in conjunction with others who share the same location and language. Consequently, there are strong incentives to locate with others who share the same language, and to learn the language that others speak. The cost of learning is endogenous: agents who are learning cannot produce.

Our model delivers a full assimilation equilibrium, as well as geographic and linguistic isolation equilibria. In the latter equilibria, location and language barriers prevent economic interaction from occurring. Increasing returns and strategic complementarities are present, but sometimes they operate locally, in addition to globally.

I. INTRODUCTION

In the last century hundreds of languages have disappeared or fallen into disuse. These include Breton, Welsh, and Irish in Europe, non-Swahili languages in Central sub-Saharan Africa, numerous native North American, South American and Australian languages, most non-English European languages in the U.S. and Canada, and Ainu in Japan. By now, English is the dominant or official language in over 60 countries; in over 75 countries it is " 'routinely in evidence, publicly accessible in varying degrees, and part of the nation's recent or present identity' ". All told about two billion people are now exposed to the English language. The reach of English extends to the Internet, as well, where it has become the dominant language of choice.

At first glance this phenomenon does not appear difficult to explain with economics. Language use exhibits network externalities, as argued by Church and King (1993). For an individual, the returns to speaking a language are increasing in the number of people who speak the same language. Network externalities generate increasing returns that provide a powerful impetus to the use of a single language.

A difficulty with this explanation is that -- despite the disappearance of hundreds of languages -- hundreds of languages and language communities continue to survive and flourish. The number of Spanish speakers in the U.S. has increased at least since 1950 and continues to do so. French speakers in Canada have declined, but only very slowly from 29% to 25% in the 35 years between 1951 and 1986. Catalan, a minority language in Spain, shows no signs of diminished use. More generally, since World War II, the number of countries has tripled with an almost commensurate increase in the number of official languages. A more complete theory of language use must be able to explain why languages persist as well as why they disappear.

Cultural and historical factors undoubtedly play an important role in the persistence of some languages, but languages that die out often also have strong cultural associations. A focus on such factors alone misses an important part of the language community dynamics. The existing economic literature on language use can explain the persistence of a language only by assuming that individuals have a high (exogenous) cost of learning. But the costs of learning a language are surely endogenous: learning takes time, and the opportunity cost of an individual's time depends upon his/her productivity using his/her current language.

A second way to generate language persistence is to assume that speakers of a language are geographically isolated. Even in a world of high-tech telecommunications, most economic activity of any particular individual tends to occur in a relatively localized area. Intuitively, it seems clear that it is easier to maintain a language when its speakers are geographically concentrated, rather than dispersed. Indeed, from the 1990 U.S. Census, we computed the correlation coefficient between the percentage of Asian and Hispanic households - grouped by nationality of origin - where no one over the age of 14 speaks English 'very well', and the percentage of those households living in the center city. The correlation is .64; there is some evidence of geographic concentration of language communities. But, again, the location choices of individuals are surely endogenous and depend upon the languages that are spoken in different regions.

We therefore develop a theory of language use in which both the costs of learning and the choice of location are endogenously determined. We assume that there are two languages and two locations. We think of language as an essential tool of communication and expression, without which production is impossible. We think of geographic proximity as likewise being crucial to production. Consequently, in our model, production occurs in conjunction with other agents who speak the same language and are in the same location. Like Church and King, we assume that there is a network externality: per capita production of an agent is (weakly) increasing in the number who speak the same language. Obviously, each agent wants to locate with others who speak the same language.

In the model, the economy runs for two periods. Initially, each agent is endowed with one or both languages and one location. In the first period, each agent either produces or learns the other language. Between periods, each agent chooses whether or not to move to the other location. In the second period, production occurs again. Production can be thought of as arising from random matching in a search setting or from team production.

There are strategic complementarities in both the decision to learn and the choice of location. The opportunity cost of learning the other community's language depends on the choices of others in one's own community. If one's peers choose to learn, they are unavailable for production, and so the opportunity cost of learning is lower. The incentive to learn is thus higher. Conversely, if others choose not to learn, the opportunity cost is higher, and the incentive to learn is lower. Even if agents speak the same language, however, they may not be able to produce with one another because they are geographically isolated. We model this in the simplest way possible, by assuming that there are two locations (islands) between which no economic interaction is possible.

It is not surprising that an equilibrium of the model is a *full assimilation* equilibrium, in which all agents of one group learn the language of the other group and everyone moves to the same location. Because of the strategic complementarities, however, our model can deliver multiple equilibria. Depending on the initial endowments of location and language there can be two other equilibria in our model: a *geographic isolation* equilibrium, in which agents who share common languages are prevented from interacting because of locational barriers; and a *linguistic isolation* equilibrium, in which agents who live in the same location are prevented from interacting because of language barriers. Both of these latter equilibria can occur with or without some agents learning the other language.

Hence, despite the presence of increasing returns, our model generates equilibria in which assimilation does not occur. It can explain the survival, as well as the disappearance, of languages. It suggests that a few linguae francae will not necessarily replace all other languages. Moreover, our model does not rely on cultural, historical, or other forces -- all of which are relevant, but are not modeled here -- to generate our main implication. What is instead crucial is that location provides a local source of increasing returns working against the global increasing returns that encourage full assimilation. A minority language speaker may be better off not learning, but instead locating and producing with other non-learning speakers.

We make two final observations. First, other researchers have suggested that language can be viewed as a metaphor for more general skills that facilitate economic activity (for example, Lazear (1996) uses the term language to refer broadly to a set of cultural values). We are quite sympathetic to such interpretations, but at the same time, we argue our model does shed light on the sociolinguistics data. That is thus our main focus in the paper. The second related observation is that our analysis abstracts from the social value of language and focuses solely on language as a means of communication. Language is much more than this; it is a repository of cultural and literary values. Our emphasis here on its role in economic activity should not be interpreted as a dismissal of such aspects of language.

The next section presents some stylized facts on language use from the sociolinguistics literature. We present the model in Section III, and discuss the equilibrium allocations in Section IV. Section V discusses equilibria with and without learning. We relate our equilibria to our stylized facts in Section VI, and then briefly examine welfare issues. Section VIII concludes and outlines extensions.

## II. STYLIZED FACTS FROM THE SOCIOLINGUISTICS LITERATURE

Sociolinguistics is concerned with language varieties, functions, and speakers, and their interactions within a speech community. The sociolinguistics literature does not give us a precise definition of bilingualism. It has been defined as narrowly as native-like fluency in two or more languages, and as broadly as the ability to utter meaningful phrases in a second language. Under the broader definitions, it has been estimated that over half the world's population is bilingual. The extent to which a person is bilingual depends on several factors including speaking, listening, and writing proficiency, the frequency with which each language is used, the ability to alternate between languages, and the ability to keep the languages separate.

The literature has identified and discussed the evolution of numerous minority language populations. From an economic perspective, we often observe the following pattern:

1] Two language groups co-exist side-by-side; economic interaction between groups is limited because of language and locational differences.

2] One language group experiences a secular economic boom; industrialization occurs.

3] The second language group gradually assimilates into the economically larger neighbor. Few monolingual speakers of the second language remain.

Minority language populations that have followed this pattern include the Ainu in Japan, the Welsh in Britain, and the Bretons in France. The Ainu led an independent life in Northern Japan, sharing neither language nor location with the rest of Japan until the Meiji restoration of 1868. Subsequently, the Ainu became integrated into the Japanese economy and society. Replacing subsistence living, wage labor emerged; intermarriages occurred at a rapid rate. Between 1927 and four decades later the rate of intermarriage increased from 36 percent to 88 percent. Japanese replaced Ainu as the language of commerce and other everyday communications.

In England and Wales, the demand for labor engendered by the Industrial Revolution and subsequent industrialization integrated Wales into the larger English economy during the first half of the twentieth century. At the beginning of the century, about half of the Welsh population spoke Welsh, a large proportion of them monolingual (Coulmas, p. 179). By 1931 Welsh speakers were only 37 percent of the population, and by 1981, only 19 percent spoke Welsh. It is now estimated that monolingual Welsh speakers constitute only 1 percent of the population of Wales.

A third example of minority language assimilation involves the province of Bretagne in France. Despite earlier government attempts to eliminate the language, 90 percent of Bretons spoke Breton at the turn of the century. Industrialization of the region, according to Coulmas (1992, p. 177) led to a halving of the number of speakers by 1952 and by 1972 only 25 percent of the population, mainly the elderly, used the language in everyday communication.

The United States has a long history of immigrants of different nationalities settling in particular areas and gradually becoming English speakers. The significance of immigration in the United States experience sets it apart from the above-mentioned cases. Yet, the assimilation that ensues once immigration ceases is similar:

1] Immigrant influx leads to communities and regions isolated in terms of both language and location. There is maintenance of the mother tongue language.

2] Immigrant flow ceases and agents slowly assimilate in the larger English-speaking community.

One of the most interesting and well documented immigrant waves is the mid and late 19th century influx of Germans. The German immigration was unique in its size and its concentration (in the Midwest). In 1910, according to Kloss (1966), the German speaking community peaked at 8.8 million first and second generation German-Americans. This was and continues to be an unprecedented share (over 9 percent) of the U.S. population.

The German-American presence in the Midwestern states was large enough for local and state government to pass laws mandating that classes be taught in German. Kloss documents what may have been the first instance of elementary school bilingualism in Ohio in 1839. In the following year it became mandatory for Cincinnati to provide bilingual schools. German language schools also included parochial and private non-sectarian schools. By 1900 over 4,000 schools, with over 500,000 students, taught German and/or taught in German. Up to 10,000 clubs, numerous religious services, and about 500 German language newspapers (with a combined circulation of over 3.4 million), also helped preserve and maintain the language (see Table 1).

Two World Wars, lower German immigration, and greater non-German immigration eventually led to almost complete English language assimilation within the next half century. As of 1965, "at most 50,000 of those under eighteen years of age still speak German natively" (Kloss (1966), p. 248). Interestingly, in Minnesota, the smallest drop off in the use of German in the parishes between 1940 and 1950 was in the Minneapolis-St. Paul area, with presumably the largest number of German-Americans in the state.

Many immigrant groups assimilated more rapidly than the German-Americans. The sheer size of the community was evidently important for German language maintenance. In addition, Kloss (1966, p. 226) argues that an important factor was that "a great many German immigrants lived either in language islands or in monolingual urban sections. They settled in states and territories which were still in a pioneering stage". The relative geographic isolation of the Midwest, according to Kloss, contributed to the relatively slow assimilation of German-Americans.

If current trends continue, the numerical importance and locational concentration of the Spanish speakers in the United States will eclipse that of the German speakers a century earlier (see Table 2). The large immigrant influx and the geographic concentration of residences parallel the German experience. By 1976, those claiming Spanish as a mother tongue was almost three times as large (5.7 million) as the German and Italian mother tongue groups. According to the 1990 census, those claiming Spanish as a mother tongue is greater than 7 percent of the population, (17.3 million); this is more than eight times larger than the second largest language group.

More than half of all Hispanics live in California and Texas, although these states contain less than 20 percent of the total U.S. population. Both foreign born and native born Spanish speakers in Arizona, New Mexico, and Texas have lower rates of English use and higher rates of Spanish monolingualism than their Spanish speaking counterparts in other regions of the U.S. Almost one quarter of Hispanics live in households in which no person over 14 speaks English 'very well' (See Table 3). It seems clear that this pattern is partially explained by the geographic proximity of Mexico to these states.

Finally, Canada is an example of a country where public policy has aided in geographic and linguistic isolation. The percentage of French speakers in Canada has been falling slowly but steadily since 1951. Alarmed by this decline, Québec imposed restrictions on English language use in 1977. In effect, these restrictions impeded both intra-Québec and inter-provincial trade. Ironically, the percentage of French speakers has continued to fall, reaching 24 percent in 1986. But the policy has succeeded in making Canada an increasingly geographically and linguistically isolated country. In 1986, 90 percent of Canadians whose mother tongue was French lived in Québec; 83% of the population of Québec report French as the mother tongue. Both of these proportions have been increasing. Conversely, the proportion of Québecois speaking English at home has declined from 15 percent in 1971 to 12 percent in 1986. While the number of bilinguals in Canada has been increasing, and more than half live in Québec, the province is still only 35 percent bilingual. Most of Québec is monolingual French.

Had German immigrants been dispersed across the United States, assimilation would surely have taken place more quickly. Likewise, if the French-speaking population were distributed more evenly in Canada, English language assimilation would be more likely. The sociolinguistics literature makes clear the significance of geography in determining both the likelihood and the speed of assimilation. We now present a concise model that highlights the role of location in determining language outcomes.

## III. MODEL

We consider a world which lasts for two periods, and has two locations, denoted by {1, 2} and two languages, denoted by {e, s}. For expositional purchases, we refer to the language as English and Spanish, respectively. There is a large, but finite, number of agents, normalized to one. We denote the mass of an individual agent by $\varepsilon$. Agents are initially located in 1 or 2 and they are initially endowed with either English or Spanish, or both (that is, some agents may be initially bilingual, denoted by b). These initial conditions are denoted by a sixtuple: $\{\beta_j^n\}$, n = 1, 2; j = e, s, b. We sometimes use a special case for illustrative purposes, which we call **equal dispersion**: $\beta_j^n = \beta_j \beta^n, n = 1,2; j = e, s, b$. The initial distribution of agents is illustrated in Figure 1. In this case the fraction of agents who speak a given language is the same in the two locations.

In the first period, agents are fixed in their location. They choose either to learn a second language or to engage in production. Between the first and second periods, agents can relocate costlessly, and in the second period, agents again produce. They cannot relocate during a period. Utility of an agent is assumed to be linear in consumption:

$$u(c , c_{t+1}) = c_t + c_{t+1} \quad (1)$$

As noted previously, we think of language and geographical proximity as essential ingredients of production. Consequently, in our model production can only occur when agents are in the same location, and agent i is more productive the more agents there are in i's location that share i's language. We assume that there are no asset markets and that goods are perishable, so i's consumption equals i's production in each period. Specifically, we assume that agent i's consumption in a period is given by:

$$c = h(N_i, N), \text{ if agent i is active (producing)}$$
$$= 0, \quad \text{if agent i is inactive (learning)} \quad (2)$$

where $N_i$ is the number of active agents (including agent i) in the region who share agent i's language, and N is the total number of active agents in the region. We make the following assumptions about this technology.

First, we assume that per capita output is increasing and (weakly) concave in the number of active agents who share agent i's language:

$$h_1(.,.) > 0 \quad (3a)$$
$$h_{11}(.,.) \leq 0, \quad (3b)$$

where $h_1$ denotes the partial derivative of h(.,.) with respect to its first argument and $h_{11}$ denotes the second derivative of h(.,.) with respect to its first argument. Condition (3a) is the network externality in the model. Condition (3b) states that this effect exhibits diminishing marginal returns. Second, we allow for a congestion externality:

$$h_2(.,.) \leq 0 \quad (3c)$$

Third, we assume that there are non-decreasing returns to scale, so that the congestion externality never outweighs the network externality:

$$N_i h_1(.,.) + N h_2(.,.) \geq 0 \quad (3d)$$

When (3d) holds with equality, we have constant returns to scale, and when it is a strict inequality there are increasing returns. One interpretation of (3d) is that agent i is no worse off if one more speaker of i's language enters i's location. Fourth, we assume that even a single agent can produce something; hence, there is always a positive cost to learning the other language. We can think of this assumption as home production. However, this production is relatively inefficient:

$$h(\varepsilon, .) = c << h(x, .), \text{ for all } x > \varepsilon. \quad (3e)$$

The final assumption is a normalization:

$$h(1,1) = 1. \quad (3f)$$

The technology described by (3a) - (3f) is quite general. It encompasses many special cases, such as the technologies in the models of Church and King (1993) and Lazear (1996). These cases can include constant returns to scale:

$$N h_1(.,.) + N h_2(.,.) = 0$$
$$h(N, N) = h(1,1) = 1, \forall N > \varepsilon$$
$$h(N_i, N) = h(N_i/N, 1),$$

as well as increasing returns to scale:

$$N h_1(.,.) + N h_2(.,.) > 0$$

h(N, N) is strictly increasing in N

h(N, N) < 1, $\forall$N < 1.

For our purposes, it is sufficient to specify the technology at this level of generality, without taking a stance on the underlying economic environment and technologies. There are in fact several possible stories that could underlie the production function we have specified. First, we could think of the underlying economic environment as involving search and matching. In the spirit of Diamond (1982), imagine that agents can produce goods costlessly, but can carry only one indivisible unit at a time. Agents randomly meet others in their location, and two agents can swap goods and consume if they can communicate -- that is, if they share a language. A second interpretation of our technology is that production takes place in teams, and team production requires both a shared language and geographical proximity. And, finally, we can think of the technology in (2) as arising in an economy with intermediate goods. For example, imagine that all agents produce one unit of a specialized intermediate input. Suppose further that competitive final goods producers in each location purchase goods from one language group and costlessly assemble goods from these intermediate inputs using a Dixit-Stiglitz technology. Then it is easy to show that we will obtain a technology satisfying (3a) - (3f). In a sense, this story decentralizes the team production structure noted above.

For future reference, we define $\phi(N) = N / h(N,N)$. Here, $\phi()$ captures the degree of returns to scale: under constant returns to scale, $\phi(N) = N$, and under increasing returns to scale, $\phi(N) > N, \forall N < 1$. If h(N,N) = N, $\forall$ N > $\varepsilon$, then per capita output increases linearly with the overall size of the economy. Hence $\phi(N) = 1$, $\forall$ N. We refer to this case as **linear increasing returns to scale (LIRS)** and take it to be the upper bound on the degree of returns to scale. That is, we assume $N \le \phi(N) \le 1$. Finally, if $\phi(N)$ increases for all N, we refer to this as an increase in the degree of returns to scale.

We also define $\sigma(N_i, N) = \dfrac{N \times h(N_i, N)}{N_i \times h(N, N)}$. Here, $\sigma(.)$ measures the concavity of our production function with respect to N . When $\sigma(.) = 1$ $\forall$ Ni, N, then h(.) is **linear** in Ni. Because h(.) is increasing in its first argument, it follows that N/N is the upper bound on $\sigma(.)$. As $\sigma(.) \rightarrow$ N/Ni, then h(Ni,N) $\rightarrow$ h(N,N) for arbitrarily small Ni > $\varepsilon$. We refer to this case as **maximal concavity**. Summarizing, we have $1 \le \sigma(Ni,N) < N/Ni$. Finally, if $\sigma$(Ni,N) increases for all Ni, N, we refer to this as an increase in the concavity of h().

IV. <u>EQUILIBRIUM AND SOLUTION OF MODEL</u>

Our equilibrium concept is pure strategy Nash equilibrium. Recall that the initial conditions of our model are given by the sixtuple $\{\beta_j^n\}$, n = 1, 2; j = e, s, b. An equilibrium of our model is likewise a sixtuple $\{\alpha_j^n\}$, n = 1, 2; j = e, s, b, that satisfies the following conditions:

1] Taking the location choices of others as given, each agent chooses location such that consumption in that location in period 2 is at least as large as in the other location;

2] Taking the learning choices of other agents as given, and correctly anticipating the location choices of all agents, (a) each agent chooses to learn if the gain from learning (extra second period consumption) exceeds the cost of learning (foregone first period consumption); and (b) each agent chooses not to learn if the gain from learning is less than the cost of learning.

*Second Period Location Game*

Because relocation between periods is costless, first-period location does not matter for this decision. We thus need to distinguish between six groups of agents: English monolinguals, Spanish monolinguals, and bilinguals, in locations 1 and 2. This means that there are 26 = 64 arrangements of the agents by location and language. Lemma A.1 in the Appendix establishes that 24 of these are distinct in the sense that the remaining 40 are identical to one of the 24, up to a relabeling of location and/or language.

**Lemma 1:** Any allocation from the location game in which agents of the same type (English-speakers, Spanish-speakers, or bilinguals) are present in both locations cannot be an equilibrium unless the technology is constant returns to scale (CRS) and agents of that type can communicate with all agents in both locations.
**Proof:** See Appendix.

The intuition for Lemma 1 derives from two observations: if an agent moves from location 1 to location 2, then the number of speakers in location 2 increases by one; and, in general, an agent is strictly better off when there is one more speaker with whom she can communicate in her location. The exception arises under constant returns, when any agent who can communicate with all others receives consumption equal to 1, and this is unaffected by the arrival of one more agent.

In the Appendix, lemma A.2 establishes that of the 24 distinct allocations, only five are equilibria of the second period location game:

1] All agents speak English or are bilingual and live in one location.
2] English speakers and bilinguals live in one location, and Spanish speakers live in the other location.
3] English speakers, Spanish speakers, and bilinguals all live in one location.
4] English speakers and some bilinguals live in one location, and Spanish speakers and some bilinguals live in the other location.
5] All agents are bilingual and live in one location.

*First Period Learning Game*

We now examine each of the five equilibria to check whether or not the allocations are consistent with equilibrium learning decisions as well as equilibrium location decisions. Hence, in addition to verifying that the second period location decisions of English monolinguals, Spanish monolinguals, and bilinguals are optimal, we need to verify that the learning decisions of English and Spanish speakers in each location in the first period are also optimal. (Obviously, initial bilinguals do not learn.) Thus there are seven conditions that must be checked for each candidate equilibrium. First, we establish some preliminaries:

**Lemma 2:** Assume that the technology has constant returns to scale. Also, assume that in the first period, agents of a given type can communicate with all others in their region.

Then those agents will not learn.
**Proof:** See Appendix.

**Lemma 3:** In the first period learning game, either all agents of a given type (that is, language and location) learn or none do. Partial learning does not occur.
**Proof:** See Appendix.

**Lemma 4:** The allocation where all agents are bilingual is not an equilibrium of the learning game.
**Proof:** See Appendix..

**Proposition 1**: There exist technologies satisfying (3a) - (3f) and initial conditions such that the allocations listed as [1] - [4] above can be supported as equilibria.
**Proof:** See Appendix.

## V. DISCUSSION

There are strong incentives for agglomeration and assimilation in the world described in this paper. People wish to locate where they can communicate with others, and people have an incentive to learn in order to be able to communicate with more others. Thus our model does indeed contain the ingredients that lead to assimilation. At the same time, it is possible to obtain equilibrium outcomes in which assimilation does not occur, and separate languages instead persist; in some of these cases, localized agglomeration occurs.

The different equilibrium outcomes in the model are driven by the different incentives that agents have to learn, and to relocate. The outcomes also depend, of course, on the initial conditions. For some initial conditions and technologies multiple equilibria are possible. Before discussing the equilibria of the model in detail, therefore, we briefly summarize how the technology and the initial conditions of the model affect the incentives to learn.

**Result 1**: *Greater returns to scale imply more learning* (by 'more learning' we mean that, other things equal, there are greater incentives to learn, and that a larger set of initial endowments of language and location will support learning as an equilibrium outcome). When the returns to scale from the technology are high, there is a significant benefit to being part of a large group. Agents who are members of a small language group therefore possess a large incentive to learn the other language and assimilate into a larger language community.

**Result 2**: *Greater concavity implies less learning.* This result is complementary to Result 1. When the technology exhibits a large degree of concavity, the marginal benefit of interacting with more agents declines rapidly. Therefore, most of the benefits of economic interaction can be achieved with a relatively small group, and the incentive to learn and join a larger language community is correspondingly reduced.

**Result 3**: *More initial bilinguals imply less learning.* Recall that the cost of learning is endogenous in the model: it is the opportunity cost of the economic activity foregone when an agent chooses to learn. If there are more initial bilinguals, then an agent who learns is giving up more first-period consumption.

**Result 4**: G*eographic isolation in the second period implies less learning.* Consider an agent who expects his/her language group to be geographically isolated in the second period. If he/she learns and goes to the other location, he/she must forego interaction with his/her own language group in order to communicate with speakers of the other language. Conversely, if all agents will be in the same location in the second period, then an agent who learns will be able to communicate with all other agents – that is, there is no need to forego interaction with those who do not learn. So the incentive to learn is lower when an agent's peers will be geographically isolated in the second period.

**Result 5:** *The larger the size of an initial group, the less likely learning will occur.* Given that the other members of a group are not learning, the cost of learning for a member is usually higher, the larger is the group. The foregone opportunities for economic interaction are larger.

Armed with these observations, we now consider the four different equilibrium outcomes that can arise in our model. In general, these equilibria can arise either with or without learning. We divide our discussion into equilibria with learning, and equilibria without learning. In much of what follows, we provide intuition only, based on the results just cited. (The detailed derivations are often tedious, and are contained in a technical appendix available from the authors on request.)

For illustrative purposes, we present a number of figures that show the initial conditions (that is, initial distributions of agents) that are consistent with our different equilibria. We present these figures under the assumption that our technology is linear, and we consider both the linear increasing returns case (so that h(N , N) = Ni) and the constant returns case (so that h(Ni, N) = Ni/N). We consider cases with 0, 25%, and 40% initial bilinguals. We also assume in our figures that agents are initially equally dispersed across the two locations. Thus, as in Figure 1b, an initial distribution of agents is represented as a point that shows the relative size of the two regions and the two language groups.

For concreteness, we discuss the equilibria in terms of specific cases. For example, in our discussion of full assimilation, we consider the case where all Spanish agents learn, and all agents locate in region 1. There are of course equivalent cases where all English agents learn, or all agents locate in region 2, or both. Whenever there is learning, we assume the Spanish in 1 learn. Whenever agents are in one location only, we assume they are in location 1. Because we focus on specific cases, the shaded region in our figures shows the set of initial endowments that are consistent with the case in question.

*Equilibria with learning*

Whenever there is learning, there is assimilation. There may be **full assimilation**, meaning that all the members of one language group learn the other language and all agents locate together. Alternatively, there may be partial assimilation, in which case some agents learn and assimilate, but monolinguals of both languages nevertheless remain. There are two subcases here: **partial assimilation with geographic isolation**, in which one monolingual group locates separately from all other agents; and **partial assimilation with linguistic isolation**, where agents locate together, but monolinguals of each language cannot communicate because of the language difference. We consider the three cases in turn.

**Full Assimilation**: {English, Bilinguals}, { $\phi$ }

In the specific example we exposit here, the Spanish speakers in both regions learn English. All speakers (English and the new bilinguals) in region two then migrate to region one. The location decisions of all agents are trivially optimal. Given that the Spanish speakers are learning, the decisions of English monolinguals not to learn are likewise optimal. For the initial Spanish speakers, given that all other Spanish speakers are learning, the cost of learning is the cost of not producing with the initial bilinguals. The gain from learning is the gain to having all the English-speakers as production partners. It is easy to see, then, how learning can be an equilibrium choice for Spanish speakers. Results 1 and 2 imply that the equilibrium is easiest to support in the case of linear LIRS. With maximal concavity by contrast, this allocation cannot be supported as an equilibrium. We illustrate this equilibrium under both constant returns and linear increasing returns in Figure 2. Note that the relative size of the two regions is not relevant under constant returns to scale.

**Partial Assimilation with Geographic Isolation:** {English, Bilinguals}, {Spanish}

In our specific case, Spanish agents in location 1 learn, and Spanish agents in location 2 do not learn. In the equilibrium, the total number of agents in the location where the bilinguals locate must be greater than the number of agents in the other location, since otherwise the bilingual agents would want to move. What forces lead to the Spanish in location 2 not learning? The number of Spanish in 2 must be large enough so that the cost of learning, given that other Spanish speakers in 2 are not learning, is high. On the other hand, if the number of Spanish in location 2 is too large, then the Spanish in 1 will not learn, but will instead move to location 2. There is thus a tension, because the incentives must be such that learning is an equilibrium choice for one group of Spanish speakers, and no learning is an equilibrium choice for the other group. Hence, it is difficult to generalize about the effects of changing concavity or the returns to scale.

Under constant returns, this equilibrium cannot occur. There is no incentive for learning, because what matters under constant returns is the proportion of same-language speakers. In other words, Spanish agents in 1 would have no incentive to learn because they could earn consumption equal to what they get in 1 by locating in 2. Figure 3 shows the set of initial locations and languages that support the equilibrium in both the linear constant returns and linear increasing returns cases.

As the number of bilinguals increases this equilibrium becomes more difficult to support. More bilinguals in location 2 make it more likely that the Spanish agents in location 2 will not learn. More bilinguals in location 1 make it more likely that the Spanish agents in location 1 will not learn, too. Finally, more initial bilinguals make it more likely that the Spanish in 2 will want to move. These last two effects reduce the set of initial distributions for which the equilibrium holds, and are consistent with Result 3, that more bilinguals imply less learning. The more initial bilinguals, the smaller the initial Spanish in region 1 must be, and the smaller region 1 must be to support this equilibrium.

This equilibrium illustrates how the strategic complementarity in learning can lead to multiple equilibria. For a considerable range of initial distributions, it is possible to get both Full Assimilation (all Spanish agents learn), and Partial Assimilation in which Spanish agents in location 1 learn, but those in location 2 do not learn. The complementarities in the second case are a local complementarity similar to those in Benabou (1993).

**Partial Assimilation with Linguistic Isolation:** {English, Spanish, Bilinguals}, { $\phi$ }
The case we exposit here is one in which the Spanish in location 1 choose to learn English, and no other group learns. All agents choose to locate in one region, but cannot all interact, because some agents from each language group have chosen not to learn. This equilibrium trivially satisfies the optimal location conditions. As in the previous case, however, there is a tension in the learning decisions. One group of initial Spanish speakers must be sufficiently small so that learning is an equilibrium, while the other group must be sufficiently large that each agent (given that all other agents are not learning) will choose not to learn. Similarly, the number of initial bilinguals must be sufficiently small so that the Spanish in 1 learn, but sufficiently large so that the Spanish in 2 do not learn. Finally, the number of Spanish who learn and the number of initial bilinguals must be sufficiently large to discourage any English speakers from learning.

Figure 4 shows the initial conditions consistent with equilibrium in the constant returns and linear increasing returns cases. With linear increasing returns and no initial bilinguals, there is no intersection between the region where the English do not learn and the region where the Spanish in 2 do not learn. As we increase the number of initial bilinguals, the region in which the English do not learn becomes bigger (Result 3), so it becomes possible to support this equilibrium. When the number of initial bilinguals gets very large, this makes it less likely that the Spanish in 1 learn; this effect reduces the set of initial endowments consistent with equilibrium.

With linear constant returns, an equilibrium can be supported without initial bilinguals. This is in contrast to the LIRS case and is consistent with Result 1. The more initial bilinguals, the smaller the region of initial distributions consistent with equilibrium.

*Equilibria without learning*

We now turn our attention to equilibria that can be supported without learning. From Results 1-3 we should generally expect these equilibria to be easier to support with high concavity, low returns to scale, and a large number of initial bilinguals. Note that for bilingual agents to be a part of an equilibrium without learning, these agents must have been initially bilingual.

One reason for focusing on these cases separately is that they allow us to discuss steady states. Although our model is essentially static, we can introduce some rudimentary dynamics as follows. Consider a world of successive generations, in which the period 2 outcome from one generation provides the initial conditions for the next generation. A steady state exists when, given an initial allocation $\{\beta_j^n\}$, the final allocation $\{\alpha_j^n\}$, where $\alpha_j^n = \beta_j^n, \forall n, j$, is an equilibrium. A necessary condition for a steady state is that it is optimal for all agents not to learn. All of the following equilibria can be supported as steady states.

**Full Assimilation**: {English, Bilinguals}, { $\phi$ }
For this to be an equilibrium without learning, we must have no initial Spanish speakers. It is then straightforward to see that full assimilation can be an equilibrium. If there are initially only English speakers and bilinguals, then the English speakers will evidently choose not to learn. Thus full assimilation can be a steady state.

**Geographic Isolation**: {English, Bilinguals}, {Spanish}

Geographic isolation can likewise be supported as an equilibrium without learning. In this equilibrium, both groups of English speakers are in one location, and both groups of Spanish speakers are in the other. The (initial) bilinguals will choose to locate in the larger region (inclusive of bilinguals). We discuss the case where they locate with the English monolinguals. Given this decision, we know that the English monolinguals have no incentive to learn. The key to this equilibrium is the conditions that support no learning by both initial groups of Spanish speakers.

We illustrate this equilibrium for the constant returns and the linear increasing returns cases in Figure 5. Under constant returns to scale, the incentives to learn are low (see Result 2); it is relatively easy to support this equilibrium. A high degree of returns to scale tends to encourage assimilation, and so makes it harder to support this equilibrium. Greater concavity makes this equilibrium easier to support because it reduces the benefit from learning.

This equilibrium can arise if there are no initial bilinguals. In this case, the monolinguals sort themselves into separate locations, and there is no interaction at all between speakers of different languages. We think of this as full geographic isolation. For the Spanish not to learn, each group must be large, relative to the English in the same region and the Spanish in the other region. Hence, each Spanish group cannot be too large, because otherwise the smaller group would have an incentive to learn.

Larger numbers of initial bilinguals increase the cost of learning, (Result 3) and make it easier to support this equilibrium. However, more initial bilinguals also means fewer initial Spanish speakers (holding the number of English speakers constant), which makes it more likely that learning will occur (Result 5). If the number of bilinguals is too large, Spanish agents will no longer wish to locate separately, but will instead choose to learn and move to location 1. There is a tension between two forces that influence learning. A larger number of initial bilinguals makes it easier to support this equilibrium, but only up to some threshold. This is illustrated in the linear increasing returns case in Figure 5. From the figure it is clear that distributions where one region is roughly the same size as the other, and where one language group is roughly the same size as the other, are more likely to support this equilibrium.

**Linguistic Isolation**: {English, Bilinguals, Spanish}, { $\phi$ }

In this equilibrium no one learns and everyone chooses to live in location 1. As with geographic isolation, the initial endowments of Spanish speakers must be sufficiently large to make not learning worthwhile, but also must not be so large as to encourage English speakers to learn. The reverse also holds: the initial endowments of English speakers must be sufficiently large to make not learning worthwhile, but must not be so large as to encourage Spanish speakers to learn. This is another way of saying that greater concavity and lower returns to scale make it easier to support this equilibrium. If there are no initial bilinguals, note that the equilibrium implies that there is no communication or interaction across language groups, even though all agents are in the same location. We think of this as full linguistic isolation.

Figure 6 shows the initial distributions consistent with equilibrium in this case. It is hardest to support this equilibrium in the case of linearity and linear increasing returns. Under both constant and increasing returns to scale, the greater the number of bilinguals, the easier it is to support this equilibrium. Moreover, in the constant returns to scale case, the relative size of the two regions does not matter. In the linear increasing returns to scale case, only a large number of initial bilinguals and a roughly equal initial distribution of Spanish and English speakers and of location 1 and location 2 endowments can deliver this equilibrium. The distributions that produce this equilibrium can also support full assimilation. We again see the importance of strategic complementarities in learning.

For the reason noted in Result 4, it is also harder to support this equilibrium than the geographic isolation equilibrium. When agents are geographically isolated, they can only enjoy the benefits of communicating with the other language group if they learn and give up the opportunity to communicate with their own language group. But if agents are in the same location, then an agent does not have to give up interaction with her own language group in order to communicate with the other group. This is clearly illustrated by comparing Figures 5 and 6.

**Bilinguals in both regions**: {English, Bilinguals}, {Spanish, Bilinguals}

This equilibrium is a special case that can only occur under constant returns to scale (with increasing returns, the bilinguals will migrate to one or the other location). Since all agents can communicate with all other agents in their region, all agents receive consumption equal to 1 under constant returns to scale. No agent has an incentive to learn, and no agent has an incentive to relocate.

## VI. EXAMPLES

We now apply our model to some of the examples cited in Sections I and II of the paper.

### A] Language Decline

Consider the phenomenon of language decline -- for example, the decline of Welsh. We could imagine that, prior to industrialization, geographic isolation (with few or even zero bilinguals) was a steady-state equilibrium for English and Welsh. The minority group (Welsh) was sufficiently large that its members had little incentive to learn English. This equilibrium could be maintained as a steady-state as long as the number of Welsh was large enough. (For example, in the linear LIRS case, a necessary condition would be that the number of Welsh exceeds 1/3.)

Now think of the Industrial Revolution as an exogenous shock that increased the number of English agents, measured in efficiency units. (Alternatively, suppose the Industrial Revolution changed the nature of technology from constant returns to increasing returns.) Geographic isolation may cease to be an equilibrium. Welsh agents have an incentive to learn English, become bilingual and then assimilate into the larger English-language community. Full assimilation can obviously be maintained as a steady-state, with no learning and no relocation. The Industrial Revolution changes the steady-state from geographic isolation to full assimilation.

### B] German Immigration and Assimilation

We do not explicitly account for immigration in our framework. But we can think of an initial influx of German-speaking individuals as providing the initial conditions that allow for geographic isolation, as above. Moreover, continued immigration (in the mid and late 19th century) could sustain geographic isolation even in a growing economy, because the relative economic size of the immigrant group remains large. World War I, decreased German immigration, and increased non-German immigration, can all be viewed as shocks that lowered the cost of assimilating, so that full assimilation eventually occurs. We argue that this fits well with Kloss's (1966) description of the German-American experience in the Midwest.

C] Québec

We think of Canada as divided into two regions, Québec and the Rest-of-Canada, in a geographic isolation equilibrium: French monolinguals and bilinguals in Québec; English monolinguals in the Rest of Canada. In this setting, both geographic isolation and full assimilation may be equilibrium outcomes. It can be the case that it would be an equilibrium outcome for French monolinguals to become bilingual, but that this equilibrium is undesirable from the point of view of French speakers (that is, assimilation could be a coordination failure outcome for this language group). The language policies of the provincial government could perhaps even be interpreted as an equilibrium selection mechanism in this case.

D] Spanish-speakers in the United States

In section II we described the geographic concentration of Spanish speakers in the United States. One way to think of this concentration is as an example of a geographic isolation equilibrium: Spanish speakers locate in the Southwest and English speakers everywhere else. Yet this does not seem to capture the extent of Latino integration in many U.S. metropolitan areas, such as New York, Chicago, Los Angeles, Houston and Miami. These and similar metropolitan areas might better be described as a linguistic isolation equilibrium where Spanish speakers, English speakers and bilinguals co-exist and interact. This equilibrium can be supported as a steady-state as long as the number of bilinguals is large enough. In both the linear constant returns and linear, linear increasing returns case, the number of Spanish speakers plus bilinguals must exceed 1/2 and the number of English speakers plus bilinguals must exceed 1/2. Hence, increases in the number of bilinguals make it easier to maintain this equilibrium.

Assuming the number of bilinguals is not too large, Figures 5 and 6 show that for smaller language groups it is easier to support geographic isolation than linguistic isolation. In the U.S., there are fewer Asians than Hispanics; moreover, the Hispanic groups typically speak variations on one language, Spanish, while the Asian groups speak different languages. Hence, the model implies that the Asian households should be relatively more geographically isolated (implying a higher correlation of the percentage of households for each language group that do not speak English very well and the percentage of households that live in the center city) than Hispanic households. We indeed find evidence supporting this implication, as footnote 6 indicates.

## VII. WELFARE

As there are a number of externalities present in the model, it is natural to consider the welfare properties of the different equilibria. It is not in general possible to Pareto-rank the different equilibria, because there are six initial groups of agents who are affected in different ways. For example, those who learn bear the cost of foregone consumption in the first period. There are two externalities from the learning decision. Those who learn bestow a positive externality on monolingual speakers of the other language. But those who learn are unavailable for production in the first period, and so impose a negative externality on their peers.

To focus our discussion, we begin by considering only steady state allocations, thus abstracting from the externalities associated with learning. It is clear that the full assimilation equilibrium has the desirable property that the consumption of all agents is at its maximum value. If we ignore the cost of learning, then full assimilation is desirable. If agents can all communicate with one another, then it is optimal for them to be in the same location.

The more interesting comparisons are between the two cases where there are monolinguals of both languages and bilinguals. As discussed above, there are two possible equilibria: all the agents may be in one location (linguistic isolation), or one group of monolinguals (for concreteness, Spanish) may be in one location, while bilinguals and the other monolinguals are in the other location (geographic isolation). Comparing these two, it is evident that bilinguals are better off in the linguistic isolation case, since they can then communicate with all agents. English monolinguals, however, are worse off under linguistic isolation. The reason is the congestion externality – the difference between the two cases, from the perspective of an English speaker, is simply whether or not there are Spanish monolinguals present. Finally, it is in general ambiguous which allocation would be preferred by Spanish monolinguals. In the linguistic isolation equilibrium they are able to communicate with bilinguals, whereas in the geographic isolation equilibrium they can only communicate among themselves. In the linguistic isolation equilibrium, however, Spanish speakers are hurt by the congestion externality, because there are English monolinguals with whom they cannot communicate. Whenever there is no congestion externality, such as in the linear LIRS case, it follows that linguistic isolation is preferred by all agents to geographic isolation.

Given that full assimilation entails maximum output for all agents, the next natural question to ask is: when would a move to full assimilation -- from either linguistic isolation or geographic isolation -- be Pareto-improving, taking account of the costs of learning?

Assuming that transfers are not possible, it turns out that it is not Pareto-improving to move from linguistic isolation to full assimilation unless the technology is constant returns to scale. The reason is that initial bilinguals are unambiguously harmed by the learning of Spanish-speaking agents in the first period. English-speaking agents, by contrast, are made better off in both periods, since Spanish speakers who learn do not impose the congestion externality in the first period, and are available for production and trade in the second period. Spanish speakers may or may not benefit from learning.

A move from geographic isolation to full assimilation, where the isolated group (Spanish speakers) does the learning, is beneficial to English speakers and bilinguals. It is beneficial to the Spanish speakers if

$$1/2 > h(\beta\ , \beta s).$$

The smaller is the Spanish speaking group, the lower is the cost of learning and the greater is the benefit from learning. The greater the returns to scale, and the lower the concavity, the smaller is $h(\beta s, \beta s)$, in which case it is again more likely that learning is beneficial.

The move from geographic isolation to full assimilation could also be brought about by learning of the English speaking agents. In this case, Spanish speakers are unambiguously better off. English speakers are better off if

$$1/2 > h(\beta b+\beta e, \beta b+\beta e).$$

Bilinguals lose out in the first period because they cannot interact with English speakers, but they gain in the second period by being able to interact with Spanish monolinguals.

They benefit on net if

$$h(\beta b, \beta b) + 1 > 2h(\beta b+\beta e, \beta b+\beta e).$$

We can rewrite this as

$$1 - h(\beta b+\beta e, \beta b+\beta e) > h(\beta b+\beta e, \beta b+\beta e) - h(\beta b, \beta b).$$

The left-hand side is the second-period gain from moving to full assimilation, and the right-hand side is the first period cost.

Even taking account of the cost of learning, and perhaps even if the majority group does the learning, assimilation may be Pareto-improving. As noted previously, however, assimilation – even equilibrium assimilation – need not be beneficial for all groups. And finally, we emphasize again that our discussion ignores any value of language beyond its role as a tool for communication. Languages are, of course, much more than tools of communication; they are repositories of literature, history and culture. Assimilation may thus entail significant social costs associated with the disappearance of minority languages.

## VIII. CONCLUSIONS AND EXTENSIONS

We develop a general model that highlights the role of location in language assimilation. Location matters if there are costs to producing across locations, so agents want to live where other agents live. Language matters if there are costs to producing across languages, so agents want to speak the language that others speak. The initial distribution of languages in the two locations determines the equilibrium degree of assimilation. In our model we can explain both assimilation, and linguistic and geographic isolation, without any underlying heterogeneity among agents other than their initial language and location endowment.

Our model is based on the decision to acquire a second language. One might legitimately ask to what extent language acquisition is actually a conscious decision made by adults. Do adults explicitly consider the costs and benefits of learning another language, or do parents, schools, and public policy dictate the languages a child learns and carries with him or her into adulthood? Perhaps it is more appropriate to think of languages as exogenous from the perspective of the adult entering into the market-oriented years of his or her life. While we acknowledge that most adults do not learn another language, we remind the reader that this is an equilibrium outcome: adults are *choosing* not to learn. Further, in the United States and other countries, immigration as well as national language policies mean that choice of language is a critically important decision for many adults. For people thinking of moving to Québec, the cost of learning French is certainly a part of the calculus. For these and other reasons we (as well as many of the authors cited in our paper) have chosen to model language acquisition explicitly.

Most of our discussion has interpreted language literally, and we believe the model does shed light on the sociolinguistics data. Other research has suggested that language can also be viewed as a metaphor for more general skills that facilitate economic activity. In Lang (1986) language is viewed as just a vehicle for communication between agents, and Lazear (1996) uses culture as his metaphor. We are quite sympathetic to these interpretations, and indeed would offer others. For example, we think language models could be used to shed light on communication across scientific disciplines: interdisciplinary work is costly because it requires investment in the language of the other discipline.

We plan several extensions of our model. Two natural extensions involve allowing agents some possibilities to trade across locations and produce across languages. The former could be accomplished by introducing more than one good and by incorporating 'iceberg' transportation costs of the type that have been used in the international trade literature. Our intuition suggests that introducing multiple goods will increase the likelihood that linguistic and/or geographic isolation is an equilibrium. Producing across languages could involve introducing translators that serve as intermediaries, as in Yavas (1994), or introducing multiple technologies, some of which require a single language, and others which do not.

A third extension involves introducing dynamics. One of the most significant observations that can be drawn from the sociolinguistics literature is the intergenerational assimilation of minority language groups. Veltman (1983, p. 213) states that, except for the Spanish and the Navajo, the pattern in the United States involves two generations: the first generation "undertake the process of learning the English language and making it their own", while the second generation adopts English as their language of everyday use (although they may still learn the minority language as their mother tongue). The rate of assimilation is evidently affected by immigration, immigrant location, and public policy towards immigration and schooling. We are currently pursuing research along these lines using an overlapping generations framework. This framework will allow us to consider other determinants of language acquisition, such as parental choice and schooling.

## APPENDIX

**Lemma 1:** Any allocation from the location game in which agents of the same type (English-speakers, Spanish-speakers, or bilinguals) are present in both locations cannot be an equilibrium unless the technology is constant returns to scale (CRS) and agents of that type can communicate with all agents in both locations.

**Proof:**

Consider an allocation where agent i has optimally chosen to be in location 1. Then:

$$h(N_i^1, N^1) \geq h(N_i^2 + \varepsilon, N^2 + \varepsilon)$$

This implies (from (3d)):

$$h(N_i^1, N^1) > h(N_i^2, N^2)$$

unless h(.,.) is constant returns to scale and $N_i^2 = N^2$. Now assume that an agent of the same type as agent i has optimally chosen to be in location 2. Then it similarly follows that

$$h(N_i^2, N^2) > h(N_i^1, N^1)$$

unless h(.,.) is constant returns to scale and $N_i^1 = N^1$. Thus, we have a contradiction except when h(.,.) is CRS, $N_i^1 = N^1$, and $N_i^2 = N^2$.

**Lemma 2:** Assume that the technology has constant returns to scale. Also, assume that in the first period, agents of a given type can communicate with all others in their region. Then those agents will not learn.

**Proof:**

The gain from learning is always bounded from above by 1-c. Under the above assumptions, each non-learner's first period output = 1. This is the cost of learning.

**Lemma 3:** In the first period learning game, either all agents of a given type (that is,

language and location) learn or none do. Partial learning does not occur.
**Proof:**
Suppose not. That is, suppose we have an equilibrium where some number NL > 0 engage in learning, and some number NN > 0 of agents of the same type do not engage in learning. From Lemma 2, we know that either there is increasing returns to scale, or these agents cannot communicate with all others in their region (or both). This implies that CL > CN , where CL is the cost of learning for an agent who has chosen to learn, and CN is the cost that a non-learner would incur, were she instead to choose to learn. The reason is that, were one of the non-learners to choose to learn, she would be giving up the opportunity to trade with NN - $\varepsilon$ agents, whereas learners give up the opportunity to trade with NN agents. (The argument is similar to that in Lemma 1.) Similarly, the gain to learning for a non-learner is as least as great as the gain to a learner. But then, if it is worthwhile for learning agents to learn, then it must a fortiori be worthwhile for a non-learner to learn. Thus we have a contradiction.

**Lemma 4** The allocation where all agents are bilingual is not an equilibrium of the learning game.
**Proof:**
If all other agents are bilingual, then the gain from learning for an English or Spanish monolingual is zero, and the cost of learning is strictly positive. Thus at least one monolingual would choose not to learn.
**Lemma A.1:** There are 24 distinct ways in which agents can arrange themselves in the two locations.
**Proof:**
The final allocation is described by a sixtuple $\{\alpha_j^n\}$, n = 1,2; j = e, s, b, where $\alpha_j^n$ is the number of agents in location n who speak language j. Each of these elements can be positive or 0 so there are 26 = 64 different configurations. Many of these are identical up to a relabeling of location and/or language, however. When these duplicates are eliminated, we are left with 24 distinct configurations. Please refer to the technical appendix, available from the authors on request, for the list of all 24 distinct allocations.

**Lemma A.2:** Of the 24 distinct allocations, there are 5 distinct equilibria of the second stage location game.
**Proof:**
Three allocations are trivially inadmissible: in one, there are no agents; in the other two there are English-speaking agents only. Thus they violate our assumption that at least one agent can speak Spanish. Seven allocations entail a particular type of agent present in both locations, and either $N_i^1 \neq N^1$ or $N_i^2 \neq N^2$. Hence they can be ruled out by Lemma 1. Five allocations have agents of the same type present in both locations, and have the property that all agents can communicate with all other agents. By Lemma 2, these are equilibria only under CRS, and in this case location is indeterminate and irrelevant. Moreover, these equilibria are equivalent to the equilibrium case (discussed below) where English and bilingual agents are in a single location. Hence they are not distinct equilibria. Two other allocations are special cases of two equilibria with bilinguals; in the special cases the number of bilinguals is just zero. These are not distinct equilibria, either. Finally, two allocations cannot be ruled out on the basis of Lemmas 1 and 2, but

are likewise not equilibria:

$$\{\{\alpha_e^1, \alpha_s^1, \alpha_b^1\}, \{\alpha_b^2\}\}$$
$$\{\{\alpha_e^1, \alpha_s^1\}, \{\alpha_b^2\}\}$$

In the first case, for the English speakers to be locating optimally, we have

$$h(\alpha_e^1, \alpha^1) \geq h(\alpha^2, \alpha^2)$$

For the bilinguals in location 2 to be locating optimally, we have

$$h(\alpha^2, \alpha^2) \geq h(\alpha^1, \alpha^1)$$

But, because h(.,.) is increasing in its first argument,

$$h(\alpha^1, \alpha^1) > h(\alpha_e^1, \alpha^1)$$

so we have a contradiction. A similar argument yields a contradiction in the second case. This leaves just five cases that can be equilibria of the second stage location game. Please see the technical appendix, available from the authors on request, for a complete listing of the allocations that are not equilibria.

**Proposition 1:** There exist technologies satisfying (3a) - (3f) and initial conditions such that the allocations listed as [1] - [4] can be supported as equilibria.
**Proof:**
Allocation [5] cannot be supported as an equilibrium of the learning game, because if all other agents are bilingual, then the gain from learning for an English-speaking or Spanish-speaking agent is zero. The cost of learning, however, is strictly positive.

For details on how allocations [1] - [4] can be supported, we refer the interested reader to the technical appendix.

## REFERENCES


Benabou, R., "Workings of a City: Location, Education, and Production", Quarterly Journal of Economics, 108 (1993) 619-652.

Benabou, R., "Equity and Efficiency in Human Capital Investment: The Local Connection", Review of Economic Studies, (1996)

Breton, A., and P. Mieszkowski, "The Economics of Bilingualism", in W. Oates (ed.) The Political Economy of Fiscal Federalism, Free Press: Lexington, 1977.

Chiswick, B., "Speaking, Reading, and Earnings among Low-skilled Immigrants", Journal of Labor Economics, 9, No. 2 (1991), 149-170.

Chiswick, B., "The Endogeneity Between Language and Earnings: International Analyses", Journal of Labor Economics, 13, No. 2 (1995), 245-287.

Chiswick, B. and P. Miller, "Earnings in Canada: The Roles of Immigrant Generation, French Ethnicity and Language", in T. Paul Schultz, ed., Research in Population Economics, JAI: Greenwich, CT, 1988.

Chiswick, B. and P. Miller, "Language in the Immigrant Labor Market", in B. Chiswick, ed., Immigration, Language and Ethnicity: Canada and the United States, AEI Press: Washington, D.C., 1992.

Church, J. and I. King, "Bilingualism and Network Externalities", Canadian Journal of Economics, 26, 2 (1993), 337-345.

Cooper, R. and A. John, "Coordinating Coordination Failures in Keynesian Models", Quarterly Journal of Economics, 103 (1988): 441-464.

Coulmas, F., Language and Economy, Basil Blackwell: Oxford, U.K., 1992.

Diamond, P., "Aggregate Demand Management in Search Equilibrium", Journal of Political Economy, 90 (1982), 881-894.

Durkin, J., "Immigration, Assimilation, and Growth", manuscript, Wayne State University, 1996a.

Durkin, J., "On the Optimal Location of International Borders", manuscript, Wayne State University, 1996b.

Ethier, W., "National and International Returns to Scale in the Modern Theory of International Trade", American Economic Review, 72 (1982), 389-405.

Fishman, J., ed. Language Loyalty in the United States, Mouton: The Hague, Netherlands, 1966.
Fishman, J., Sociolinguistics: A Brief Introduction, Newbury House: Rowley, MA, 1972.

Fishman, J. and J. Hofman, "Mother Tongue and Nativity in the American Population", in J. Fishman, ed., Language Loyalty in the United States, Mouton: The Hague, Netherlands, 1966.

Fishman, J., R. Hayden, and M. Warshauer, "The Non-English and the Ethnic Group Press, 1910-1960", in J. Fishman, ed., Language Loyalty in the United States, Mouton: The Hague, Netherlands, 1966.

Grenier, G., "An Economic Perspective on Learning a Second Language", Journal of Multilingual and Multicultural Development, 4, No. 6 (1983), 471-483.

Grin, F., "The Economic Approach to Minority Languages", Journal of Multilingual and Multicultural Development, 11, (1990), 153-175.

Grin, F., "The Economics of Language: Survey, Assessment and Prospects", International Journal of the Sociology of Language, 121 (1996), 17-44.

Hofman, J., "Mother Tongue Retentiveness in Ethnic Parishes", in J. Fishman, ed., Language Loyalty in the United States, Mouton: The Hague, Netherlands, 1966.

Katz, K. and Matsui, A. "When Trade Requires Coordination", manuscript, Mount Holyoke College and University of Pennsylvania, 1995.

Katz, M. and C. Shapiro, "Technology Adoption in the Presence of Network Externalities", Journal of Political Economy, 94 No. 4 (1986), 822-841.

Kiyotaki, N. and R. Wright, "On Money as a Medium of Exchange", Journal of Political Economy, 94 No. 4 (1989), 927-954.

Kiyotaki, N. and R. Wright, "A Search-Theoretic Approach to Monetary Economics", American Economic Review, 83 No. 1 (1993), 63-77.

Kloss, H., "German-American Language Maintenance Efforts", in J. Fishman, ed., Language Loyalty in the United States, Mouton: The Hague, Netherlands, 1966.

Kremer, M., "The O-Ring Theory of Economic Development", Quarterly Journal of Economics, 108 (193): 551-576.

Laitin, D. D., "The Tower of Babel as a Coordination Game: Political Linguistics in Ghana". American Political Science Review 88 No. 3 (1994): 622-634.

Lazear, E., "Assimilation: 1900 and Now", manuscript, Hoover Institution, Stanford, CA, 1995.

Lazear, E., "Culture and Language", manuscript, Hoover Institution, Stanford, CA, June 1996.

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

Matsuyama, K., N. Kiyotaki, and A. Matsui, "Towards a Theory of International Currency", Review of Economic Studies, 60 (1993), 283-307.

McManus, W., "Labor Market Costs of Language Disparity: An Interpretation of Hispanic Earnings Differences", American Economic Review, 72 No. 4 (1985), 818-827.

Robinson, C., "Language Choice: The Distribution of Language Skills and Earnings in a Dual- Language Economy", in Research in Labor Economics, vol. 9, JAI Press: Greenwich, CT, 1988.

Romaine, S., Bilingualism, 2nd ed., Basil Blackwell: Oxford, U.K., 1995.

Romer, P., "Endogenous Technological Change". Journal of Political Economy, 98 No. 5 pt. 2 (1990): S71-S102.

Schelling, T., "A Dynamic Model of Segregation", Journal of Mathematical Sociology, (1971).

Statistics Canada, Canada Year Book, 1990, Ottawa, Canada, 1990.

Tamura, R., "An Economics of Translators", manuscript, Clemson University, 1996.

Trejos, A. and R. Wright, "Toward a Theory of International Currency: A Step Further", manuscript, Northwestern University and University of Pennsylvania, 1995.

United States Department of Commerce, Bureau of the Census, Social and Economic Characteristics, United States Summary, 1990.

Veltman, C., Language Shift in the United States, Mouton: Berlin, Germany, 1983.

Williamson, S. and R. Wright, "Barter and Monetary Exchange Under Private Information", American Economic Review, 84 No. 1 (1994), 104-124.

Yavas, A., "Middlemen in Bilateral Search Markets", Journal of Labor Economics, 12 (1994), 406-429.

Yeh, Y., "Searching for Streets Paved with Gold: Immigrants' Residential Choices in Urban America, Theory and Evidence", manuscript, University of Chicago, 1996.