Linguistic Diversity in the Very Long Run

Andrew John*

June 22, 2015

Abstract

A key stylized fact of long-run historical linguistics is that the world experienced a global increase, followed by a decline, in linguistic diversity. This paper develops a model that generates the endogenous creation and destruction of distinct linguistic groups, and that can produce dynamics to match this stylized fact. Specifically, the paper presents an endogenous growth model in which language is an engine of growth: agents who can communicate easily are able to produce and consume more. At the same time, economic interaction among agents affects the evolution of language communities.

*Melbourne Business School, 200 Leicester Street, Carlton, Vic 3053, Australia; andrew.john@mbs.edu.

Etymologists are uncertain of the exact derivation of the term "pidgin." Possibilities include that it is a Chinese corruption of either the English or the Portuguese term for business, that it derives from the Hebrew word for exchange or trade, or that it comes from the Chinese characters "pei, ts'in",meaning "paying money" (Malmkjær, 1991). Etymologists think it very likely, in other words, that there is a link between economic activity and the development of pidgin languages. More generally, since languages evolve and change in part because of interaction among different language groups, and since production and exchange are a major reason for contact among groups, it is natural to believe that economic activity has influenced the development and adaptation of languages throughout history. The relationship between economic and linguistic variables runs in both directions: it is likewise natural to think that language use and language communities affect economic activity.

This paper develops an endogenous growth model that is predicated on such bi-directional links between economic activity and language. The primary goal of the model is to offer an explanation of the emergence and disappearance of languages. Historical linguistics tells us that, far back in history, the number of languages in the world was increasing. More recently, however, the world has seen a large number of languages vanish, and many more are under threat. The major stylized fact that is matched in this paper is this increase and subsequent decline in the number of languages in the world.

The story of the paper is based on research in historical linguistics, which suggests that, in primitive economies where there is little interaction among agents, there was a tendency for languages to diverge; conversely, communication become easier among agents who interact frequently. The economic side of the model is an endogenous growth story. Growth of the economy reduces transportation costs, in turn increasing the ability and incentive of language groups to interact, while the ease with which groups can communicate is itself a factor that can inhibit or encourage economic growth. Economic growth thus reduces the cost of interaction, which in turn leads to more contact between agents. Contact among groups brings languages closer together, so there is an eventual tendency for languages to converge. This mechanism is indeed able to match the stylized fact of an increase followed by a decrease in linguistic diversity.

The most striking feature of the model is that distinct languages are themselves endogenous: the model can generate both the appearance of languages and their subsequent disappearance. The mechanism is as follows: agents *choose* which other agents they will interact with. Other things equal, agents choose to interact more with agents with whom they communicate more easily; likewise they choose to communicate less with those with whom they communicate less easily. Over time, this leads to some groups of agents becoming closer linguistically, while others move further apart. Thus, in equilibrium, we observe distinct sets of agents, demarcated geographically, who are able to communicate easily with each other, but who cannot communicate efficiently with those in other

2

groups.

These language groups are stable in the sense that they can remain in existence for long periods. Eventually, however, it becomes worthwhile for agents to communicate across language groups. When they do so, groups with distinct languages eventually fuse together, so agents who used to be linguistically separated end up speaking a common language. To the best of my knowledge, this is the first model in either the literature on the economics of language or in the linguistics literature that delivers—*as a theoretical result*—the long-run emergence and disappearance of distinct languages.

The model addresses language evolution in the very long run and is therefore by necessity extremely stylized. My use of a highly simplified model to address changes that occur over millennia means that this paper is in the spirit of those by Kremer (1993), Goodfriend and McDermott (1995), and unified growth theory models (see Galor 2005 for a detailed summary and review of this theory). While the model presented here does generate long-run growth, and can also generate protracted periods of very slow growth followed by rapid growth, the paper is intended primarily as a contribution to the literature on economics and language rather than the literature on very-long-run economic growth.

The aim of the model is to highlight the feedback mechanism between language change and economic growth, and to show how this mechanism can explain the historical pattern of linguistic diversity. Numerous forces and factors have influenced both economic growth and the development of languages, most of which are—quite intentionally—not modelled here. Among other things, I neglect the role of population growth, migration, war and conquest both on economic growth and on the spread and demise of languages. While such features have obviously affected the evolution of particular languages and language groups, the paper demonstrates that they need not have been the ultimate driving force behind a global pattern of language growth and decline.

Finally, I emphasise that, in the model, the sole purpose of language is as a means of communication, and thereby an aid to production and exchange. Language is of course much more than this; it is a critical repository of cultural and literary knowledge and values. My focus here on language as a communicative device should not be taken in any sense to imply a dismissal of these other features and values of language. Because I do not include these other attributes of language, it would be misleading to draw welfare implications from the model; I deliberately make no attempt to do so.

1 Economics, Language, and Linguistic Diversity

1.1 Previous Literature

Links from linguistic to economic variables, and *vice versa*, have been a topic of study in economics for the past half century. Marschak (1965) appears to have been the first to introduce economic ideas, such as efficiency and costs, into the study of language. Most of the subsequent economics

literature has investigated how linguistic variables can affect economic decisions and outcomes.

The fields of consumer choice theory and industrial organisation contain work that views language as an attribute of a good. Researchers have considered when to offer goods in multiple language versions (see, for example, Hočevar, 1975, and Caminal, 2010). Grin (2003), meanwhile, discusses the price of and demand for language-specific commodities. [1]

The most studied branch of the literature on economics and language views language skills as a form of human capital. Researchers in this literature have studied the return to learning a second language, with early examples including Grenier and Vaillancourt, 1983, and McManus, Gould and Welch, 1983. There is evidence that language proficiency affects wages (see, for example, Vaillancourt 1980, Carliner, 1980, Grenier, 1984 Garrouste, 2008, and the survey of Zhang and Grenier, 2013), that a shared language affects migration decisions (Beine et al, 2009), and—perhaps not surprisingly given the previous results—there is a large body of evidence showing that immigrants to the United States can obtain a large economic return from learning English (see, for example, Chiswick and Miller, 1995, 1999, 2005, 2012; Yeh, 1996).[2]

Turning to macroeconomic questions, researchers have investigated the influence of linguistic diversity on economic growth and on the patterns of trade. Cross-sectional analyses of economic growth have found that, other things equal, growth is higher in countries where a "major" language is spoken (Hall and Jones, 1999), and that growth is lower in countries with a high degree of linguistic diversity (Tamura, 1997; Easterly and Levine, 1997; Alesina et al, 2003). Gravity models of trade find that countries with a common language generally engage in a significantly greater degree of trade, even after controlling for colonial and historical linkages (Helliwell, 1999; Anderson and van Wincoop, 2004; Rose, 2004; Hutchinson, 2005; Dow and Karunaratna, 2006; Melitz, 2008; Lohmann, 2011).

All the papers mentioned so far focus on the effects of language on economic outcomes. But there is also a small yet significant literature that analyses the economic determinants of language use. This body of work includes papers by Breton and Mieskowski (1977), Grenier-Vaillancourt (1983), Lang (1986), Robinson (1988), Selten and Pool (1991), Church and King (1993), John and Yi (1996), and Lazear (1999). Grin (1996, 2006) and Ginsburgh and Weber (2011) provide excellent surveys.[3] However, the economic literature on language has—with the important exceptions of Michalopoulos (2012) and Clingingsmith (2015), both of which I discuss in the next section—paid little attention

---

[1]There is also a significant related literature in the field of management studies. For example, there is work indicating that the corporate performance of multinationals is influenced by their language structure (Welch, Welch, and Piekkari, 2005; Luo and Shenkar, 2006), while Vaara et al argue that language differences affect the efficacy of mergers.

[2]The results of Beine et al and Grenier are discussed in Ginsburgh and Weber (2011).

[3]There is also research that explores deeper connections between economics and language. For example, Rubinstein (2000) applies economic methods to linguistics in an attempt to understand the construction of natural languages. A rather different example is Chen (2012), who documents an apparent link between the ways in which different languages mark the future tense and the intertemporal decision-making of users of those languages: those who speak languages that clearly demarcate the future tense engage in less forward-looking behavior.

to the historical evolution of languages.

The model considered here is also related to other literature in economics, including research on long-run growth, on local interactions and on conformity. Technically, the paper presents an endogenous growth model in which reductions in transportation and communication costs serve as engines of long-run growth. The model embodies the idea that the inability of agents to communicate outside their language community could act as a impediment to growth; this idea seems plausible, particularly in very primitive societies. That said, the paper's focus is on language outcomes, not the mechanism of growth. As Galor (2005) argues, compelling stories of very-long-run economic growth need to encompass both a Malthusian phase and a transition to modern economic growth. The present model, because it abstracts from fertility and population growth, cannot deliver a true Malthusian outcome. Nevertheless, the mechanisms in the current paper could be embedded in a unified growth model as a source of long-run variation in human capital.

In the model of the paper, agents interact with agents who are relatively nearby, geographically. In this sense, the model has some similarity with work on contagion and local interactions (see, for example, Morris, 2000; Galeotti and Vega-Redondo, 2011; Chen et al, 2013). Unlike most of the work in that literature, however, the optimisation decision of an individual agent in the current model includes a choice about which (and how many) others to interact with. In the equilibrium of the model, agents only choose to travel to those with whom they can communicate most easily. This result carries echoes of Schelling's (1960) analysis of endogenous segregation, as well as more recent work on conformity—see for example Young (2001) and Bisin et al (2006)—and occupational choice as analyzed by Mookherjee et al (2010). In the current model, however, the drive to conformity is driven by production externalities rather than preferences, and—crucially—the similarity of agents (here, their ability to communicate with each other) is endogenous.

1.2   Linguistic Diversity

The origins of language are debated by linguists and cognitive scientists, and there is significant disagreement about when language originally emerged.[4] Some linguists have argued—based on estimates of the time that it takes languages to spread—that languages must have been in existence for at least 100,000 years (see, for example, Nichols, 1992). Malmkjær (2002), however, cites research (drawing on fossil evidence) suggesting that Cro-Magnons still did not possess the vein system required for human speech; this research suggests that the origin of speech may have been as recent as 10,000 years ago. Still, it is widely, even if not universally, accepted that relatively sophisticated language systems must have been in use for at least the last 40,000 years.

The argument that sophisticated language goes back 40 millennia is based on archeological find-

---

[4]The discussion in this section draws on Malmkjær (2002), Crowley and Bowen (2010), Anderson (2002), Pinker (1994), Lyovin (1997), Comrie, Matthews and Polinsky (1996), and Nettle (1999).

ings that humans at that time were engaged in relatively complex social activities. Because of the nature of language evolution, however, it is all but impossible for linguists to infer much about the particulars of language in such distant history. A highly controversial research agenda in historical linguistics has attempted to find the roots of all languages in a single language ("proto-World") that could even be tens of millennia older. Alternatively, it is possible that language developed independently in different geographic locations.

Still controversial, though perhaps marginally less so, is the possibility that there was a relatively small number (perhaps as few as half a dozen) of proto-languages that existed 10,000 - 15,000 years ago, and that provide the basis of today's languages. For example, it has been suggested that the languages of Europe, much of Asia, and North Africa can all be traced to "Nostratic", which—if it existed—was spoken by a people of hunter-gatherers.

The classification of languages is much more reliable for the last few thousand years. There is general agreement among students of historical linguistics that, three to five millennia ago, there was a group of identifiable languages that are the genetic ancestors of modern languages: proto-Indo-European, Uralic, Caucasian; Altaic, Sino-Tibetan, Dravidian, Austro-Asiatic; Afro-Asiatic, Khoisian, Niger-Kordofanian, Nilo-Saharan; Austro-Tai, Indo-Pacific, (aboriginal) Australian; Eskimo-Aleut, and New Guinea.[5] The most disputed regions are Austronesia and the Americas, where some linguists claim to find genetic resemblances among groupings of languages (for example, various Native American languages, or Papuan/New Guinea languages) while others are more circumspect. Still, a large number of the 6,000-7,000 languages in the world today can be traced back to a relatively small group that have been reliably established to have existed a few millennia ago. Of these, probably the most important, and certainly the most studied, is proto-Indo-European, which was spoken over a large stretch of territory including much of Northern Europe and Western Russia, Turkey and Afghanistan, Northern India, and parts of China.

Thus we can trace existing languages to a relatively small number of ancestors. Many of these language groupings are, naturally enough, associated with important migrations. As these languages took hold, other languages were almost certainly displaced and replaced. One therefore cannot conclude directly that there was necessarily once only a very small number of languages in the world. Still, it appears clear that, for much of human history, the number of languages in the world was increasing.

In more recent history, the number of languages spoken in the world has been decreasing.[6] In

---

[5]See Anderson (2002). There are some languages, known as isolates, that do not seem to fit these classifications, such as Basque and possibly also Japanese and Korean. The methods of classification of languages are the subject of much study in historical linguistics; see for example Lyovin (1997), Chapter 1. There is some dispute about some details of these language groupings, particularly in Austronesia and in the Americas, where genetic relationships among languages are least well established. There is also dispute about the languages of the Americas, some of which have been traced to Eskimo-Aleut and the (earlier) language Na-Dene, but others of which are not yet clearly classified.

[6]Matthews (1996) presents a table that places the peak of linguistic diversity at about 17,000 BP, with an estimated $10,000 - 15,000$ languages, but there is no supporting discussion in his chapter.

relatively recent times, the decline in the number of languages has become rapid, and many languages are currently under threat. *Ethnologue* (Lewis, 2009) catalogues 6,909 extant languages, of which 473 are "nearly extinct." Conservative estimates suggest that about 2,500 languages may become extinct over the next century, and some estimates even suggest that up to 90 percent of existing languages will die out (or will become moribund—that is, will have no young speakers) over this period. As linguists often point out, this loss of diversity of languages far exceeds the loss of diversity of species that currently captures so much attention.[7]

Nichols' (1992) summary, while open to debate in the details, is a good statement of the broad stylized fact that I am interested in.[8]

> Three stages can be assumed for the spread of human language. The first begins with the origin of our species, probably in Africa and over 100,000 years ago, and comprises the period when our range was limited to the tropical Old World and perhaps chiefly Africa. This must have been a time of great linguistic diversity. Societies would have been small, simple, and autonomous... There must have been not only proliferation of linguistic diversity, but considerable initial diversity, both genetic and typological. The language of the first modern humans must have inherited a good deal of the substantial diversity that earlier hominid language had developed... And this initial diversity would have multiplied for at least half the lifespan of our species, through the entire first stage.
>
> The second stage can be called the state of expansion. During this time, humans expanded out of the the Old World tropics to colonize Europe, inner Asia, NewGuinea-Australia, and the New World... The stage of expansion can be dated to approximately 60,000 to 30,000 years ago. This stage too must have involved increasing diversity for that part of the linguistic population that participated in the expansion, but it may have ultimately led to extinction of many lineages that did not...
>
> The third stage begins with the end of glaciation. It involves the rise of complex societies and large-scale economies, the spread of languages driven by economic and political prestige, and consequent reduction in linguistic diversity.

1.3   Language Difference and Language Change

A theory of linguistic diversity must include notions of language difference and language change. Linguists have of course investigated both of these topics in detail, and any discussion must begin with the observation is that it is surprisingly hard to define precisely what a language is. Linguists generally accept that there is neither a simple definition of a language, nor a clear distinction between a language and a dialect (although distinctions are sometimes drawn on the basis of shared core vocabulary). Thus the *Ethnologue* catalogue of languages, cited earlier, must be interpreted with care.

---

[7] "We are in danger of losing treasures ranging from Yiddish, with far more words for 'simpleton' than the Eskimos were reputed to have for 'snow,' to Damin, a ceremonial variant of the Australian language Lardil, which has a unique 200-word vocabulary that is learnable in a day but that can express the full range of concepts in everyday speech." (Pinker, 1994 pp. 260-261)

[8] See Nichols (1992), pages 274-275.

Two related branches of linguistics, known as lexicostatistics and glottochronology, calculate measures of language similarity and uses them to estimate when in history two distinct languages might have shared a common ancestry. Measures of language distance are surveyed in Ginsburgh and Weber (2011, chapters 3 and 6); see also Crowley and Bowen (2010) for some discussion. One approach is to look at core vocabulary (which is resistant to change and less likely to be influenced by contact with other languages). For a given pair of languages, linguists then count the percentage of cognates in their core vocabularies and take this as a measure of similarity.[9] Another approach is based on how far apart languages are once they have been classified into language trees (that is, how far back up the tree you have to go to find a common ancestor). Such "cladistic" measures are less precise than lexicostatistic distance measures, but have the advantage that they are not based solely on meaning but also on other aspects of language such as syntax and phonology.

Another approach to linguistic diversity focuses not on the distances between languages, but takes as given a classification of languages and then asks how much linguistic diversity exists within a given geographical area. One measure is the probability that two randomly chosen individuals belong to different linguistic groups; see Greenberg (1956) and the discussion in Ginsburgh and Weber (2011). A more sophisticated measure combines this with measures of distance between languages. Desmet, Ortuño-Ortín and Weber (2009) use such a measure to investigate the link between linguistic diversity and redistribution.

It is well understood that languages change over time and space. Nettle (1998) observes that 'in the absence of positive interaction, people's languages diverge' (p. 359). Nerbonne (2010) investigates the relationship between language and geographic distance, and argues that random social contact, based on geography, can account for about one quarter of linguistic variation. Based on the numbers in Crowley and Bowen (2010), meanwhile, it appears that two isolated languages that are originally identical will lose about 1.5 percent of shared vocabulary every 33 years.[10]

Nettle (1998) documents and analyses the extent of linguistic diversity in the modern world. While Nettle is a linguist rather than an economist, his explanation of linguistic diversity is essentially an economic one, based on the need for risk-sharing in different environments. Specifically, in regions where growing seasons are longer, food supply is less risky, there is less need for risk-sharing, and so smaller language groups are sustainable. He provides empirical evidence that supports this correlation between growing seasons and linguistic diversity.

Michalopoulos (2012) provides a rich empirical analysis relating ethnolinguistic diversity to geographical diversity. Drawing on previous arguments by Nettle and others, he argues that heterogeneity in land quality leads to greater linguistic diversity. His reasoning is that variation in

---

[9]Wichmann et al (2010) compare different automated procedures for evaluating the distance between languages. Ipshording and Otten (2011) use one such procedure to look at trade flows and language acquisition of immigrants.

[10]Crowley and Bowen claim that an individual language retains about 80 percent of its vocabulary after 1000 years. Thus if we take two identical languages, assume that each has a very large number of words, and assume that each changes randomly, then after 1000 years they will retain approximately $0.8^2 = 64$ percent of their words in common. This translates into a decay in joint fluency of about 0.045 percent per year, or 1.5 percent every third of a century.

land endowments leads to human capital that is location-specific. This results in lower mobility, more localised ethnicities, and so greater ethnolinguistic diversity. His paper supports this thesis with careful and creative empirical work, including a finding that the link between geography and ethnolinguistic diversity vanishes in regions characterised by significant recent migrations.

Clingingsmith (2015) looks at the size distribution of languages in the world, and provides evidence that once languages exceed a threshold (35,000 speakers in his analysis), they obey Gibrat's Law: the growth rate in the number of speakers of a language is independent of the current number of speakers. Gibrat's Law implies that the size distribution of languages above the threshold is unchanging. Clingingsmith then uses results from network theory to argue that (given certain assumptions on the underlying graph) the observed size distribution of languages is consistent with the size distribution of locally connected components in a network. The distribution of connected components does not evolve dynamically, however, so his analysis does not consider the case where the set of languages changes over time *across* components.

Both Nettle and Michalopoulos provide plausible stories to motivate their empirical investigations, but neither provides a fully worked-out theory of linguistic diversity. Both researchers are primarily interested in drawing out empirical implications rather than developing a theoretical mechanism to explain the evolution of linguistic diversity. Clingingsmith, meanwhile, prevents a theory of why agents on a component of the network will eventually speak a single language, but does not present a theoretical model of the evolution of language groups.[11] The current paper is thus complementary to the work of these other researchers. In particular, while the endogenous diversity that emerges in the current paper does not require any underlying diversity in the environment, geographic diversity of the kind identified by Nettle and Michalopoulos is compatible with the framework in this paper and would reinforce the mechanisms that drive the current results.

In other work, Nettle (1999) presents a computer simulation of a simple language change, where individuals learn language from those around them in a deterministic manner, and where there is also some random change. He shows that this can generate distinct local dialects. Drawing upon the work of Dixon (1997), Nettle also develops a simple dynamic model of the evolution of language stocks.[12] His story is as follows. Suppose a new people and language arrives in a given unpopulated geographic area (say, a continent). Nettle argues, first, that the production of new stocks will be a decreasing function of time; this is because new stocks will be produced easily as people spread across a virgin continent, but will appear only slowly in an already populated continent. Second, he assumes that a constant fraction of existing stocks will become extinct each period. His model, therefore, is a single equation: $S_{t+1} = 0.95 S_t + A/t$, where A is a parameter. In early periods, the production of stocks exceeds the extinction, and so the number of stocks rises. As time progresses,

---

[11]In his framework, individuals on a network choose which language to use. Language choice is a coordination game: agents' choices depend on their expectations of the language choices of others with whom they are likely to interact. In long-run equilibrium everyone on the same connected component speaks the same language.

[12]Dixon in turn draws upon models of punctuated equilibrium from evolutionary biology.

however, the rate of new language production slows, so the extinction of languages exceeds the production of new languages. The model delivers (for a given geographical area) the phenomenon of an increase, followed by a decrease, in the number of languages.

While Nettle's work perhaps comes closest to my goal in this paper, he is working with—in his own words—a "simple heuristic model." He does not model the appearance and diffusion of languages directly, but simply assumes that the arrival of new stocks declines hyperbolically with time. Likewise, he does not model the forces that cause languages to disappear. More significantly, he does not have a model in which languages evolve based on individual choices or in which linguistic and economic forces interact.[13] Finally, as with the prior economic literature on language, Nettle's equation takes a language to be a primitive concept, whereas languages in the current paper arise endogenously.

The current paper departs from the previous literature on language death in one critical respect. Previous work on language extinction, almost without exception, takes the construct of a language as a primitive concept, and then typically explores decisions about when agents will choose to learn or use languages. [14] In this paper, I adopt a very different perspective: the primitive construct is the communicative ability of pairs of agents. With this approach, it becomes possible for languages themselves to emerge and disappear endogenously.

2   Model

This section sets out the structure of the model. While I do not make extensive use of network *theory*, the model is most compactly described using the terminology of networks. Specifically, the model both a simple geographical network and a language network. I first discuss the underlying assumptions about geography and demography, then explain how language is modelled in the paper. Next I explain how language enters the production/exchange technology and I set out the dynamic linkages in the model. Finally, I define equilibrium.

2.1   Geography and Time

There are M agents, $i = 1...M$, each of whom lives at a specified location. The index i can be used interchangeably to refer to the location or to the agent who inhabits it. The locations represent the nodes of the geographical network. This network takes one of two forms in the paper. First,

---

[13]Abrams and Strogatz (2003) present a simple probabilistic model of language death where one language drives out another; their model consists of a first-order differential equation in which the change in the number of speakers of one language depends on the existing stocks of language speakers and on a parameter reflecting language status. A similar idea, based on a more explicit economic model, can be found in John and Yi (2001).

[14]There is of course a vast literature within linguistics on language *change*: how individual languages evolve, how creoles and pidgin languages are formed, and so on. In some cases (Latin may be a good example), language change is so extreme that the original language effectively dies out. There is also a game-theoretic literature that considers the emergence of language in a highly abstract sense. These literatures, however, are not concerned with the emergence and disappearance of language communities.

I consider the case where agents live on an integer lattice, with the distance between each pair of agents set to unity. Second, I consider the case where agents are evenly distributed on a circle. (In order to develop intuition, I also present some analytical results for an approximate version of the model with an infinite number of agents.)

The population is constant, and while agents can move in a sense to be described below, they always return home: there is no permanent migration in the model. The model comprises sequential generations. At each time period $t = 1, 2, ...$ a new agent is born at each location. Each generation inherits certain characteristics that are determined by equilibrium outcomes in the previous generation, but there are no market or non-market interactions across generations. Time is indexed by t; I suppress this subscript whenever I am simply discussing the features of an individual period.

2.2 Language

As already noted, a distinguishing feature of my analysis, compared to previous work, is that languages, rather than being primitive concepts, emerge endogenously. The primitive concept that I work with instead is the communicative ability of pairs of agents, which I refer to as their *joint fluency*. Just as the locations in the model are the nodes of the geographic network, the individuals living at those locations are the nodes of the linguistic network. Think of each individual as possessing a collection of pieces of linguistic knowledge, such as morphemes, syntactical rules, idioms, phrases, and so forth. For simplicity, call these pieces of knowledge "words". Any two individuals will have a certain number of these words in common. The more that the individuals possess a shared vocabulary (including shared understanding of syntactical rules, and so on), the easier they will find it to communicate. Thus, assume that, associated with any pair of agents $\{i, j\}$, there is an index of their joint fluency $\phi_{ij} = \phi_{ji} \in (0, 1]$, where 0 indicates a complete inability to communicate and 1 indicates perfect mutual intelligibility.[15] The linguistic relationship among all agents can then be represented by the graph (N, Φ), where $N = 1...M$ and $\Phi = \{\phi_{ij}\}$ is the adjacency matrix; this graph represents a weighted undirected network.

What we call a "language" in everyday speech is, roughly speaking, a collection of items of linguistic knowledge that is shared by some subset of agents (the speakers of that language) and not shared by other agents. As Pinker (1994, p. 411) puts it, "'A language' or 'a dialect' refers to the process whereby the different speakers in a community acquire highly similar mental grammars". To make this idea precise in terms of the model, define a joint fluency threshold, $\bar{\phi}$, such that a pair of agents whose joint fluency exceeds that threshold can be said to speak the same language. Define the unweighted network $(N, \tilde{\Phi}; \bar{\phi})$ as the transformation of the original network where $\tilde{\phi}_{ij} = 1 \forall \phi_{ij} \geq \bar{\phi}$

---

[15]There can be considerable redundancy in a world where people speak multiple languages. For example, if two agents both speak both English and Spanish, then they have many more words in common than they need for communication. Two people could have close to perfect joint fluency even if their individual vocabularies are not identical.

and $\tilde{\phi}_{ij} = 0 \forall \phi_{ij} < \bar{\phi}$. Finally, define a *language community* as any complete subnetwork ($L \subseteq N, \tilde{\Phi}' \subseteq \tilde{\Phi}$) that is not a proper subset of some other complete subnetwork. Completeness means that, in a language community, each member can communicate with every other member of that community.[16] The restriction that the set not be a proper subset of another complete network means that each language community is of maximal size: there is no other agent, outside the community, who can also communicate with every member of the community. I make the additional assumption that there is a distinct language associated with each distinct language community.[17] Of course, members of different language communities can also be connected in the larger network.

Now consider three geographically adjacent individuals, $i - 1$, i, $i + 1$. I assume that the joint fluency of agent $i - 1$ and $i + 1$ is equal to the product of the fluency between $i - 1$ and i, and the fluency between i and $i + 1$. That is

$$\phi_{i-1,i+1} = \phi_{i-1,i}\phi_{i,i+1} \ \forall i.$$

Thus, by assumption, individuals who are farther apart geographically also share fewer words.[18] This assumption means that it is possible to construct the entire adjacency matrix Φ from its superdiagonal alone. In other words, given a listing of the joint fluencies of all pairs of *neighbouring* agents, it is possible to completely describe the joint fluency between *any* pair of agents:

$$\phi_{i,i+n} = \prod_{k=i}^{i+n-1} \phi_{k,k+1}.$$

It is then immediate that any language community must comprise a set of agents who form a connected subset on the geographic network.[19]

In the special case where the joint fluency of all neighbouring pairs of agents is identical ($\phi_{i,i+1} =$

---

[16]Equivalently, a language community is a set L defined by

$$L = \{\forall i, j \in L, \phi_{i,j} \geq \bar{\phi} \text{ and } \forall k \notin L, \ \exists j \in L \text{ such that } \phi_{j,k} < \bar{\phi}\}.$$

[17]While it seems intuitively reasonable to define distinct languages in terms of distinct language communities, there need not be—at least in theory—a straightforward mapping from this definition of a language community to the notion of a language as linguists would typically define the term. For example, suppose we have three agents: Agent 1 speaks Estonian and Urdu; Agent 2 speaks Urdu and Xhosa; Agent 3 speaks Xhosa and Estonian. These three agents form a language community according to the definition given. But there is no clear *language* (in the everyday sense) associated with the community. One cannot define the language as the intersection of the agents' shared linguistic knowledge, because that is the null set. But defining language as the union of the shared linguistic knowledge also does not make intuitive sense, as that would define the agents' language as comprising Estonian plus Urdu plus Xhosa. That said, this is probably a theoretical curiosity that is of limited relevance when we think about language communities in the real world. For the purposes of the present discussion, it seems reasonable to posit a one-to-one mapping between languages and language communities.

[18]This is a strong assumption. For example, suppose individuals $i - 1$ and $i + 1$ both refer to a canine as a "dog", while the intervening agent i uses the word "hound". Then imagine that individual i switches to the word "dog" instead. The languages of $i - 1$ and i have now become closer, and the languages of i and $i + 1$ have become closer, but individuals $i - 1$ and $i + 1$ share no more words than before.

[19]Suppose that agents $i - 1$ and $i + 1$ are both members of some language community L'. Then $\forall j \in L$ where j lies to the left of $i - 1$, $\phi_{i,j} \geq \phi_{i+1,j}$. Similarly, $\forall k \in L'$ where k lies to the right of $i + 1$, $\phi_{i,j} \geq \phi_{i-1,k}$. It follows that $i \in L'$.

$\phi, \forall i$), I have

$$\phi_{i,i+n} = \phi^n;$$

that is, an agent's ability to communicate with others decays geometrically with geographical distance. Provided $\phi \geq \bar{\phi}$ (that is, provided each agent can communicate fluently with their immediate neighbour), there will be overlapping language groups. Specifically, if $\bar{n}$ is defined by $\phi^{\bar{n}} \geq \bar{\phi} > \phi^{\bar{n}+1}$, then there will be a distinct language community of size $2\bar{n}+1$ centered on each location.[20] When agents travel too far from their home location, they will find themselves no longer able to communicate with those they meet.

At the other extreme, the subnetworks associated with different language communities may be unconnected. Specifically, suppose that (a) there exists a language community $L' = \{r...v\}$, (b) $\phi_{r-1,r} < \bar{\phi}$, and (c) $\phi_{v,v+1} < \bar{\phi}$. Then this community is not connected to the rest of the network: agents within the community can communicate, but they cannot communicate (at a level above $\bar{\phi}$) with anyone outside the community. If this is true for all language communities, then we can partition the complete set of agents into distinct non-overlapping language communities.

2.3 Production and Exchange

In a primitive subsistence economy, an individual might interact only with those in her household. As trading develops, economic interactions might be almost exclusively with the members of a village. In the modern economy, individuals interact directly with thousands of others, and indirectly with almost everyone on the planet. The production/exchange technology in the model captures this observation that as economies become more complex and more developed, each individual interacts with more and more other agents in the economy. This increased interaction allows specialisation, increased productive complexity, and the exploitation of comparative advantage, all of which lead to increased productivity. At the same time, whether we are talking about trade between two villages millennia in the past, or trade in the modern globalised economy, communication difficulties can impede these productive relationships.

The specification of production and exchange in the paper is deliberately stylised. The basic story can be given a number of interpretations, some of which are discussed in more detail below. The fundamental idea is that individuals face a trade-off between the time they spend producing as individuals and the time they spend interacting with others. To motivate the specification of the technology, consider the following set-up. Each agent in the economy has a single unit of time that she must divide between production and exchange. More specifically, assume that agent i devotes a certain amount of time, s_i^P, to building productive capacity, after which she is able to produce and

---

[20]More precisely, (a) if the agents are located on an integer lattice, then the statement is true for locations at a distance of at least $\bar{n}$ from the end points of the lattice; (b) if the agents are located on a circle, then the statement is true provided $2\bar{n}+1 < M$.

sell (up to) s_i^P units of output at zero marginal cost to every agent that she visits. The remainder of her time $(1 - s_i^P)$ is spent in travelling and exchange. If she devotes s_i^L units of time traveling to her left, and, s_i^R units of time traveling to her right, then her time budget constraint is

$$s_i^P + s_i^L + s_i^R = 1.$$

The number of people the agent can trade with in the leftward and rightward directions depends not only upon how much time she puts into searching for trading partners—that is, traveling to the left and to the right—but also upon the productivity of search. Specifically, the number of agents in each direction that she interacts with are given by

$$Z_i^L = \gamma_i \tau(s_i^L) \tag{1}$$

$$Z_i^R = \gamma_i \tau(s_i^R) \tag{2}$$

where $\tau(s)$ is a (time-invariant) continuous increasing weakly concave function and γ_i is a technology parameter that captures variation in the productivity of search. For example, if it takes a constant time $c/2$ to travel a unit of distance, then the search technology can be written as $\tau(s) = s$ and $\gamma_i = \frac{1}{c}$.[21]

While $\tau()$ is a continuous function, agents will always choose their search intensity to deliver an integer value of trading partners so, without loss of generality, take Z_i^L and Z_i^R to be integers.

Production and Trade with Perfect Fluency. Consider first the case where all agents can communicate perfectly with each other ($\phi_{i,j} = 1 \forall i, j$). Then if agent i sells to agent j, she can produce and sell up to the capacity level s_i^P. The output associated with this transaction is

$$y_{ij} = s_i^P.$$

Assume further that the seller (agent i) is able to capture all the surplus from this transaction. It follows that her objective is to maximise total output, which is obtained by summing the output sold to all her individual trading partners.

$$y_i = \sum_{h=1}^{Z_i^R} y_{i,i+h} + \sum_{h=1}^{Z_i^L} y_{i,i-h}$$
$$= s_i^P \left[Z_i^R + Z_i^L \right].$$

Production with Imperfect Fluency. When agents are not able to communicate perfectly, the exchange process is less efficient. To capture this, assume that agent i's ability to produce and

---

[21] An agent's total travel time includes both an outward and a return journey.

trade with agent j is proportional to their joint fluency. If agent i sells to agent j, then agent i obtains

$$y_{ij} = s_i^P \phi_{i,j}$$

from the transaction. Total output is again obtained by summing over the output produced with all the individual trading partners.

$$
\begin{aligned}
y_i &= \sum_{h=1}^{Z_i^R} y_{i,i+h} + \sum_{h=1}^{Z_i^L} y_{i,i-h} \\
&= \sum_{h=1}^{Z_i^R} s_i^P \phi_{i,i+h} + \sum_{h=1}^{Z_i^L} s_i^P \phi_{i,i-h} \\
&= s_i^P \left[\sum_{h=1}^{Z_i^R} \phi_{i,i+h} + \sum_{h=1}^{Z_i^L} \phi_{i,i-h} \right].
\end{aligned}
$$

And plugging in the expression for joint fluency gives

$$y_i = s_i^P \left[\sum_{h=1}^{Z_i^R} \prod_{k=i}^{i+h-1} \phi_{k,k+1} + \sum_{h=1}^{Z_i^L} \prod_{k=i}^{i+h-1} \phi_{k,k-1} \right]. \tag{3}$$

Production with Constant Fluency. In the case where the joint fluency between all adjacent pairs of agents is identical and equal to ϕ,

$$y_i = s_i^P \left[\sum_{h=1}^{Z_i^R} \phi^h + \sum_{h=1}^{Z_i^L} \phi^h \right].$$

(In the case of perfect fluency ($\phi = 1$) it therefore follows that, as before $y_i = s_i^P \left[Z_i^R + Z_i^L \right]$.)

Production and Exchange: Interpretation.

The specific story of production and exchange in this model is not critical; what matters is the fact that agents choose the extent to which they interact with others, and this decision depends upon their ability to communicate with others. There are several possible interpretations of the production technology derived in the previous subsection.

For example, one more detailed specification of the above model is as follows. Agents do not consume the good that they themselves produce. However, all other goods are perfect substitutes in utility. Agents therefore must exchange in order to consume, and seek to maximise total consumption. Agents produce and sell, as described before, with goods being purchased on credit. After all production and exchange has taken place, each agent owns a bundle of goods, has debts corresponding to all these goods, and has various claims on other agents for the goods that she sold. There is then a centralised secondary market where all debts are cleared, so agents who are net

creditors receive goods from agents who are net debtors. Because goods are perfect substitutes, the relative price of all goods will equal one, and each agent will therefore consume an amount equal to her production.[22]

A different interpretation of the technology is that it captures team production. In this story, think of the capacity decision as being productive effort. Agents form overlapping teams through their travel decisions (think of these as being time spent searching and forming teams); agents travel to work with other agents; and the output of each pair of agents is shared according to the relative effort that each team member applies to their joint production. For example, suppose agent i travels to agent k and agent k also travels to agent i. Then their joint production is $(s_i^P + s_k^P)\phi_{i,k}$, and each earns an amount proportional to effort.[23]

The critical feature of the technology is that agents are more productive when they can interact with a large number of other agents. This increased productivity can be interpreted as a result of specialisation and comparative advantage. Alternatively, it can be thought of as arising from greater complexity in production, in the spirit of the arguments in Kremer (1993a). More complex trade and production relationships require that agents invest resources in trading and producing with others, but such interaction is costly and can be impeded by language barriers.

2.4   The Evolution of Language

Now consider the behaviour of the economy over time. Isolated groups will tend, over time, to develop differences in languages. Sound shifts and other random changes make languages that were originally similar become more different. By contrast, groups of agents who interact a great deal will tend to develop languages that are more alike. This process can arise in a number of ways. Two groups who speak different languages may develop a pidgin language, which in turn may evolve into a creole. The language spoken by one group of agents may disappear, and be completely replaced by another, dominant language. Or languages may simply borrow terms from one another. These changes and adaptations will typically take several generations.

I model such processes as follows. First, I assume that, in the absence of contact among groups, joint fluency has a tendency to decay. Second, I assume that exchange tends to increase joint fluency: when adjacent agents interact frequently, languages tend to become more similar. Specifically, joint fluency evolves according to the following set of equations

$$\phi_{i,i+1;t+1} = \beta\phi_{i,i+1;t} + (1-\beta)f(V_{i,i+1;t}) + \epsilon_{i,i+1;t+1} \tag{4}$$

where $\beta \in (0,1)$ is the rate of decay of joint fluency, $f() : \Re^+ \to [0,1]$ is an increasing function,

---

[22]The secondary market plays a similar role in this story to the market in Lagos and Wright (2005), although the model I describe does not have monetary exchange.

[23]This interpretation still goes through if, say, agent i travels to agent k, but agent k does not travel to agent i. In this case only agent i's effort is productive in the team, and that agent captures all the surplus.

and $V_{i,i+1;t}$ is the number of visitors who travel between agents i and $i+1$. That is, $V()$ represents the total number of agents whose travels take them to both location i and $i+1$. The more such visitors there are in period t, the greater the joint fluency of the next pair of agents to inhabit that location.[24] The final term, $\epsilon_{i,i+1;t+1}$, is a source of random variation in joint fluency.[25]

2.5 The Evolution of Productivity

The final ingredient is the dynamic adjustment of the (agent-specific) productivity parameters, $\gamma_{i;t}$. These variables evolve over time according to

$$\gamma_{i;t+1} = \gamma_{i;t} + max[0, \theta(y_{i;t} - y_{i;t-1})]. \tag{5}$$

The idea is that transportation and communication become more efficient as an external by-product of output growth. Negative changes in output leave productivity unchanged (there is no technological regress), but positive changes in output result in technological improvements that allow agents to search more efficiently. As a result, output growth last period will lead to higher output this period, giving rise to an endogenous growth mechanism in the model.

2.6 Equilibrium

At each date, agents in this model make a single optimization decision along two margins. They choose capacity and search times in each direction (left and right) to maximize their individual output, subject to the time budget constraint. That is, agent i solves

$$\max_{\{s_i^P, s_i^R, s_i^L\}} y_i = s_i^P \left[\sum_{h=1}^{Z_i^R} \prod_{k=i}^{i+h-1} \phi_{k,k+1} + \sum_{h=1}^{Z_i^L} \prod_{k=i}^{i+h-1} \phi_{k,k-1} \right]$$

subject to

$$s_i^P + s_i^R + s_i^L = 1$$
$$Z_i^L = \gamma_i \tau(s_i^L)$$
$$Z_i^R = \gamma_i \tau(s_i^R).$$

The solutions to these problems give the vector $y_{i;t}$, which—given initial values—can then be fed into the expression for the evolution of productivity. The implied solutions for Z_i^L and Z_i^R allow the calculation of the number of visitors to each location, which in turn can be fed into the dynamic equation for the evolution of joint fluencies.

---

[24]The equation is consistent with languages drifting apart in many different ways. The most obvious is lexical change, but there could also be phonological or syntactic changes, for example. Much of Nettle's (1999) book is an exploration of the different kinds of linguistic diversity that exist, and the ways in which they can arise.

[25]In a non-stochastic world ($\epsilon_{i,i+1;t+1} = 0$), this specification is sufficient to ensure $\phi \in (0,1]$. In the stochastic simulations I impose additional restrictions as required to ensure that fluency remains in the required range.

An equilibrium for this model is therefore a sequence $\{y_{i;t}, \phi_{i;t}, \gamma_{i;t}, V_{i;t}, Z_{i;t}^L, Z_{i;t}^R\}$, where, in all periods $t = 1, 2, ...$ (i) agents solve the maximization problem just specified, (ii) team sizes are given by (1) and (2), (iii) outputs for each agent are given by (3), (iv) joint fluencies evolve according to (4), given realisations of $\epsilon_{i,i+1;t+1}$ and given initial values of the state variables $(y_{i;0}, \phi_{i;0}, \gamma_{i;1})$, and (v) search productivities evolve according to (5).

3 Endogenous Growth

Before turning to simulations of the model, I briefly characterise some of the model's properties. The model contains an endogenous growth mechanism, whereby shocks to output propagate through time.

3.1 The Case of a Fixed Language Community

To see how endogenous growth operates in the model, consider first the case of a language community L_K of size $K + 1$ in which all agents can communicate perfectly: $\phi_{i.j} = 1 \forall i, j \in L_K$. Assume also that $\tau(s) = s$, which means that agents are indifferent about their direction of search.[26]

Proposition 1. Assume that $\phi_{i.j} = 1 \forall i, j \in L_K$, $\tau(s) = s$, and that the initial conditions satisfy $4y_0 < \gamma_{i;1} \leq 2K$. Then (i) for $\gamma_{i;t} < 2K$, the output of each agent evolves according to the linear second-order difference equation $\Delta y_{i;t+1} = \frac{\theta}{4} \Delta y_{i;t}$; (ii) the output of each agent initially grows at an increasing rate, reaching a maximum growth rate of $\frac{\theta}{4} - 1$, and thereafter grows at a decreasing rate; (iii) for $\gamma_{i;t} > 2K$, the output of each agent evolves according to the non-linear second-order difference equation $\tilde{y}_{i;t+2} = \frac{K \tilde{y}_{i;t+1}}{1 + \frac{\theta}{K} \tilde{y}_{i;t+1}(\tilde{y}_{i;t} - \tilde{y}_{i;t+1})}$, where $\tilde{y} = K - y$; (iv) as $t \to \infty, y_{i;t} \to K$.

Proof: See Appendix.

The first part of the proposition tells us that, provided agents are not constrained by the size of the language community, output grows without limit, and the second part tells us that, in general, the growth rate of output will initially increase, and then decline. However, when the productivity of search is high enough, the size of the language community becomes a constraint on the agent's decision. In an unconstrained world she would choose to visit a larger number of people, but that option is not available. While specialisation is not explicitly modeled here, the flavour of this result is very familiar: the division of labour is limited by the extent of the market.

In this constrained region, output continues to grow, and hence the productivity of search also continues to grow. The change in output, which was given by

$$\Delta y_{i;t+1} = \frac{\theta}{4} \Delta y_{i;t}$$

---

[26]When $\tau(s)$ is linear in s and $\phi_{i.j} = 1$, agents' marginal efficiency of search is equals and constant in each direction.

in the unconstrained region, now becomes

$$\Delta y_{i;t+1} = \theta K^2 \frac{\Delta y_{i;t}}{\gamma_{i;t+1}\gamma_{i;t}}.$$

(This can be rewritten as the difference equation in $y_{i;t}$ given in part (iii) of the proposition.) It is obvious from this expression that the growth rate of output declines more rapidly than in the unconstrained region, because increases in the productivity of search reduce the growth rate of output. The benefits of increased search productivity decline over time—one way to understand this is that the constraint of optimal search binds more and more tightly. Search productivity grows without limit, so eventually (almost) all time can be spent producing rather than searching for partners, and output tends to a limit of K.

3.2 Trading Across Language Communities

The previous subsection showed that, when the productivity of search is high enough, agents become limited by the extent of the market—that is, the size of the language community in which they live. That analysis assumed there was no possibility of travel to neighbouring language communities. Suppose instead that agent i can trade within her own language community L_K, as before, but that she also has the opportunity of crossing a language boundary and trading in a neighbouring community L_M of size M. Assume that joint fluency at the boundary equals $\hat{\phi}$. The question is: at what point does it become worthwhile for an agent to cross the boundary, and trade with those with whom communication is lower?

Proposition 2. Assume that $\gamma_{i;t} > 2K$, $\phi_{i.j} = 1 \forall i, j \in L_K$, $\phi_{i.j} = 1 \forall i, j \in L_M$, $\tau(s) = s$. Assume that L_K and L_M are adjacent, and that joint fluency at the boundary equals $\hat{\phi} \Rightarrow \phi_{i.j} = \hat{\phi} \forall i \in L_K, j \in L_M$. If $\gamma_{i;t} > \frac{K(1+\hat{\phi})}{\hat{\phi}}$, agent i will trade across the language boundary.

Proof: See Appendix.

Once agents are trading with all of the agents in their own language community, they face a decision about whether to cross a linguistic border and trade with those in a distinct language community. Proposition 1 revealed that, as search technology improves, there are diminishing returns to trading in one's own community. But trading across linguistic borders is less efficient than trading within one's own language community. Proposition 2 shows that, when the productivity of search is high enough, it becomes worthwhile to cross linguistic borders. The condition is intuitive: when the size of the home language community (K) is larger, agents are less likely to cross the language boundary, and when the joint fluency across the two communities ($\hat{\phi}$) is larger, agents are more likely to cross the language boundary.

The forces behind this result play a significant role in the simulations reported below. Language communities form in the model, and for long periods of time agents may optimally choose to trade

only within their own language communities. Overall technological progress, however, is a force that eventually makes it optimal to cross language boundaries and engage in exchange with agents in other language communities.

4   Simulation Results

The specific question of interest is whether distinct linguistic groupings can emerge endogenously in a setting where agents choose their search time in each direction. I answer this question in the affirmative through simulations of the complete model.[27]

Because there is randomness in the evolution of fluency, output has a random component. The following two subsections show two ways in which distinct disjoint language communities can emerge in the model. While the patterns that eventually appear are similar in both cases, the mechanisms are different.

I set the decay of fluency equal to 0.985; this is the number implied by Crowley's (1992) numbers, taking a period as equal to one third of a century. Other than this, however, the simulations are not calibrated; indeed, it is unclear that calibration is particularly meaningful in a model of this kind. Of course, the lack of calibration does mean that I have substantial liberty in choosing functional forms and parameters. Given that my primary aim in these simulations is to demonstrate *possible* outcomes, this freedom should not be viewed as a weakness of the model.[28]

4.1   Endogenous Languages Arising From Boundary Effects

In the first simulation, 100 agents are placed on a finite line segment and the model is run for 1500 periods.[29] Figure 1 shows the outcome of one simulation. The figure contains four panels. The top panels of the figure illustrate the joint fluencies ($\phi_{i,i+1}$) spatially. The horizontal axis represents the 100 agents on a line, while the vertical axis shows joint fluencies between adjacent agents. The three panels are snapshots of the economy: the left panel shows the economy after 500 periods; the centre panel shows the economy after 1000 periods; and the right panel shows the economy after 1500 periods. The colour gradient indicates time: the oldest observations are yellow; more recent observations are orange; the most recent observations are red.

---

[27]The simulations were performed in MATLAB.

[28]The simulations reported here are restricted to simple (bounded) power functions. I have not attempted any systematic analysis of which parameters and functional forms will generate results similar to those reported, but I do note that the qualitative behaviour that I report is easy to generate and appears relatively robust.

[29]The specific equations are

$$\tau(s) = s^{0.5},$$
$$\phi_{i,j;t+1} = max[\phi_{min}, min[1, 0.985\phi_{i,j;t} + 0.0001V_{i,j;t}^{2.7} + \epsilon_{i,j;t+1}]],$$
$$\gamma_{i;t+1} = \gamma_{i;t} + max\{0, \Delta y_{i;t}\}.$$

The standard deviation of the innovations to fluency equals 0.002. Fluency is also bounded away from zero at some small value ϕ_{min} (set equal to 0.005) and constrained not to exceed one; if fluency were ever allowed to reach zero then agents would never find it worthwhile to cross a linguistic boundary.
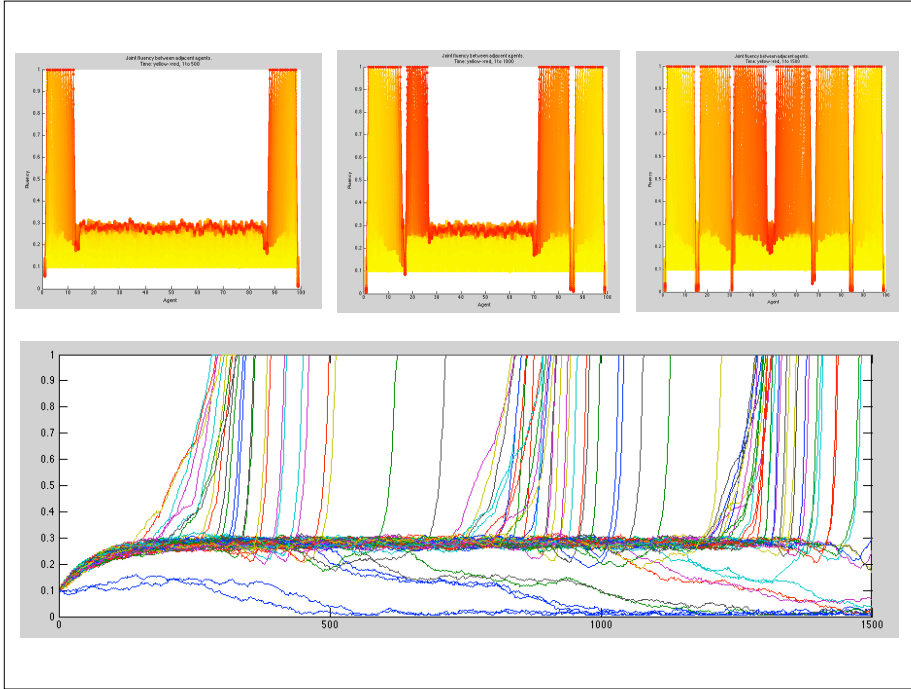
Figure 1: The emergence of distinct language communities

The three panels show the emergence of distinct language communities, starting from the boundaries and moving inwards. Thus, in the first panel, there is a group of about 15 agents at the left boundary who have joint fluency equal to 1. All of those agents can communicate perfectly with each other. There is an analogous group at the right boundary. The remaining agents have joint fluencies of about 0.3, so the ability of those agents to communicate with each other is much lower. By the time 1500 periods have elapsed (the third panel), six distinct language communities have emerged. Agents within these groups can communicate perfectly, but the agents at the edges of adjacent groups are almost unable to communicate with each other; their joint fluency is at the lower bound.

The lower panel shows the evolution of all the individual fluencies over time. The vertical axis again shows joint fluencies, while the horizontal axis shows time. From this panel, it can be seen that joint fluency between most agents moves fairly quickly from the initial value to about 0.3. Thereafter, many joint fluencies rise rapidly to a value of one (perfect fluency), while some others decline.

The emergence of language communities is due to intertemporal strategic complementarities. More specifically, the mechanism is as follows. Agents who live near the boundary are limited in their search choices. For example, the leftmost agent can only search rightwards. Because agents near the boundary are forced to search asymmetrically, there are more visitors to locations near the

boundary. Hence joint fluency increases in this region. After 500 periods, as already noted, this translates into the emergence of linguistic communities near the boundaries.

But the most striking feature of the left panel is what is happening at the *edge* of these linguistic groupings. For concreteness, the following description focuses on the left side of the picture; symmetric behaviour occurs on the right side. Consider the decision of an agent who is living on the right side of the group that has emerged on the left. Their higher joint fluency with agents close the boundary makes it attractive for them to search asymmetrically: specifically, they interact more with people on their left and *less with people on their right*. That is, there is a strategic complementarity operating across time. If agents to the left engaged in more search last period, then there is an increased incentive to travel leftward this period. This behaviour causes a break in fluency to form; the beginnings of this can be seen in the left panel, while in the centre panel the break is well defined. When joint fluency at the break point is low enough (below $\bar{\bar{\phi}}$), the group of agents at the boundary has formed a distinct and unconnected language community.

Once the break in fluencies has emerged, the decision problem of agents to the right of the break is analogous to the decisions of agents at the edge of the lattice. Thus the centre panel shows the process repeating itself, with a new language community forming, and the third panel shows the final two language groups emerging. The model therefore delivers, as an endogenous phenomenon, the arrival of distinct language communities. Moreover, given that these communities are geographically disjoint, it is reasonable to associate a distinct language with each language community. (Even if the languages spoken in different communities initially had some commonalities, the process of language change in isolated regions would lead to greater difference over time.) In other words, the model can generate increased linguistic diversity.

While it is not visible in the figure, the model does also deliver an eventual *decrease* in linguistic diversity. While the languages that emerge in this economy are highly persistent, they do not last indefinitely. As search productivity continues to increase, agents will ultimately find it worthwhile to cross language boundaries. When they do so, joint fluency at the break points increases, and distinct languages fuse together. In the very long run, this economy will have a single common language.

To summarise, this subsection has thus demonstrated that distinct linguistic groupings can emerge endogenously, persist for an extended period, and eventually disappear. Languages emerge because boundary conditions lead to asymmetric search, which then propagates through the economy. Languages disappear because increased productivity of search eventually makes it worthwhile for agents to cross regions of low joint fluency.
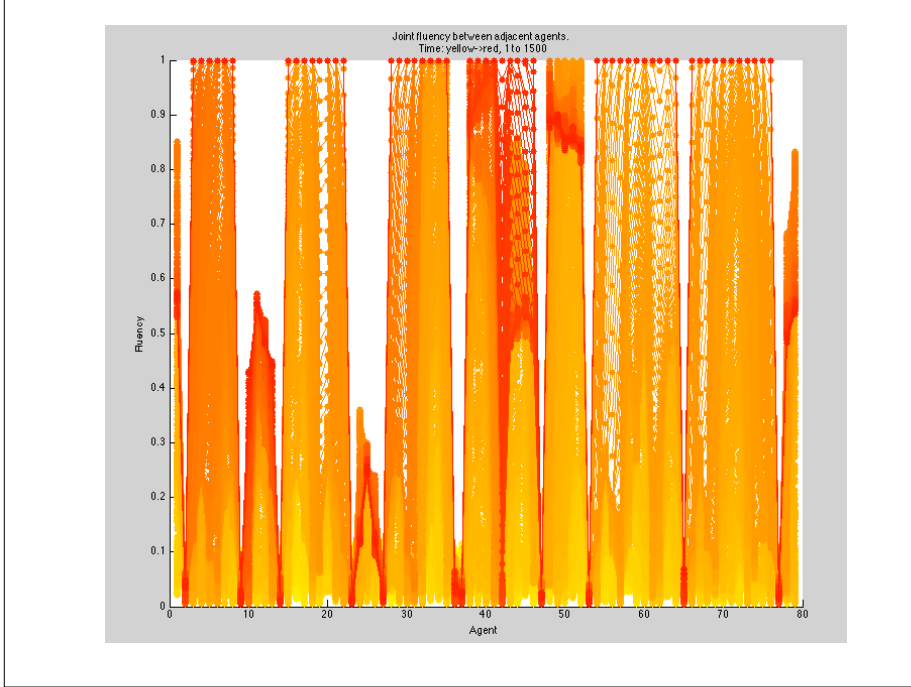
Figure 2: Languages emerging as a result of random variation in fluency

4.2 Endogenous Languages Arising from Randomness

Physical boundaries such as coastlines and mountain ranges do of course exist in the real world. Still, it would be unsatisfactory if boundary effects were the only way of generating distinct languages in the model. This subsection therefore exhibits a simulation where agents instead located on a circle (thus ruling out boundary effects) and shows that distinct languages can still arise. The simulation shown is for 80 agents and is again run for 1500 periods.[30]

The end result of the simulation in Figure 2 broadly resembles that of the previous subsection. As before, distinct language groupings have emerged, separated by boundaries of low joint fluency. Unlike the previous example, though, these language groups are of uneven size, and there are also some regions where joint fluency is neither very high, nor very low. Because the agents in this economy live on a circle, the endpoints have no significance: agent 80 is a neighbour of agent 1.

Figure 3 shows the individual evolution of joint fluencies in the model. For the first 400 periods, joint fluencies are relatively low. They vary randomly, but throughout this time there is little search activity. In fact, agents only ever trade with their immediate neighbours. Still, there is slow

---

[30]The specific equations are

$$\tau(s) = s,$$
$$\phi_{i,j;t+1} = max[\phi_{min}, min[1, 0.985\phi_{i,j;t} + 0.0001V_t^{2.7} + \epsilon_{i,j;t+1}]],$$
$$\gamma_{i;t+1} = \gamma_{i;t} + max\{0, 10\Delta y_{i;t}\}$$

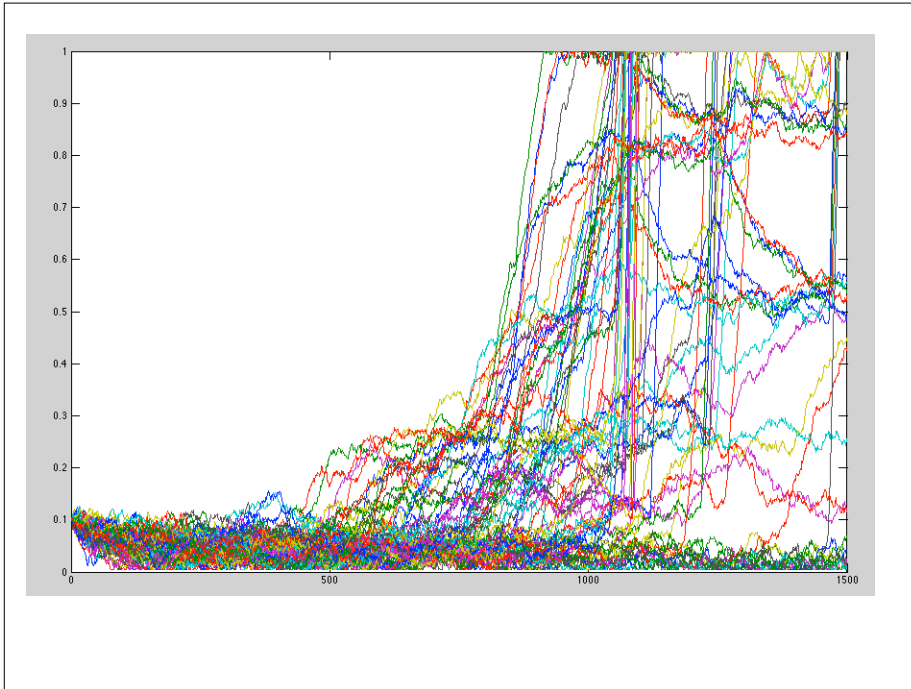The standard deviation of the innovation to fluency is 0.004.

Figure 3: Evolution of joint fluencies over time

technological progress over this period, and eventually some agents' technologies improve enough for them to travel further. Starting around period 300, some locations have three visitors, and by period 500 several regions have four visitors. At this point, as Figure 3 shows, the strategic complementarities start to take effect. As joint fluencies improve, agents find it more worthwhile to search, and as agents search more, joint fluencies improve. At the same time, boundaries of low fluency emerge, as agents find it more and more worthwhile to engage in *asymmetric* search.

5 Discussion

Simulations of this model qualitatively match several of the features of language and economic data. Here I briefly consider the correspondence between the findings of this model and the empirical linkages between language and economics..

5.1 Language Death

I have already discussed the fact that the model generates both the emergence and the demise of languages. In some cases, language death in the model arises because two existing neighbouring language groups fuse together. But, strikingly, Figures 2 and 3 reveal that language death can occur through a more subtle mechanism as well. Look (in Figure 2) at the set of agents in the vicinity of

agent 50. At one point, those agents all had joint fluency equal to 1—they were a distinct language group. But, towards the end of the simulation, the fluency of that group of agents has been declining. This can also be observed in Figure 3, where it is evident that, at some locations, joint fluency has decreased from 1 (perfect joint fluency) to lower values.

Exactly what has happened can be inferred by looking at the agents just to the left—that is, agents in the mid-40s. At one time in the simulation, they were a small distinct language group with an intermediate level of fluency, bounded on the left at about agent 42 and on the right at about agent 47. Eventually, however, some of these agents found it worthwhile to cross these boundaries. Some of those agents traveled to the left to trade with agents in and around number 40; some went to the right to trade with the agents in and around number 50. But then the boundary to the left vanished (visible in Figure 3 as a very steep increase in one joint fluency around period 1450), and suddenly all these agents found it much more productive to travel to the left rather than to the right.

At this point, the relatively small group of agents from about 48 to 53 find themselves with fewer visitors. Moreover, this group has a relatively large language community on each side. These agents therefore are finding it more profitable to deal with the larger language communities that surround them and less profitable to deal with each other. As a consequence, the fluencies in the vicinity of agent 50 are in decline. A similar phenomenon is occurring in the vicinity of agents 1 and 80, though it is less easy to spot in the figure.

Language death in the real world occurs when a language is used less and less, by fewer and fewer people, for communication, production, and exchange. As this happens, its speakers gradually conduct more and more business in other, larger, languages instead. The disappearance of language communities in the simulations follows a process that can be interpreted in exactly this way: agents interact less within their own language community, and more in other language communities instead, with the result that their original language eventually vanishes completely.

5.2   Divergence and Convergence

A major area of study in growth theory is the extent to which economies exhibit convergence or divergence over time. Figure 4 shows the evolution of output in this simulation. A notable feature of the model is that, although the initial conditions are symmetric and the randomness is small, it has propagation mechanisms that result, in the end, in a substantial endogenous dispersion of output. Agents who are lucky enough to end up in relatively large linguistic groups enjoy the benefit of much higher output than agents who are members of smaller linguistic groups. In other words, this is a growth model that generates unconditional divergence of output. This is consistent with the research (discussed earlier in the paper) that has found that countries that speak major languages exhibit higher growth rates.
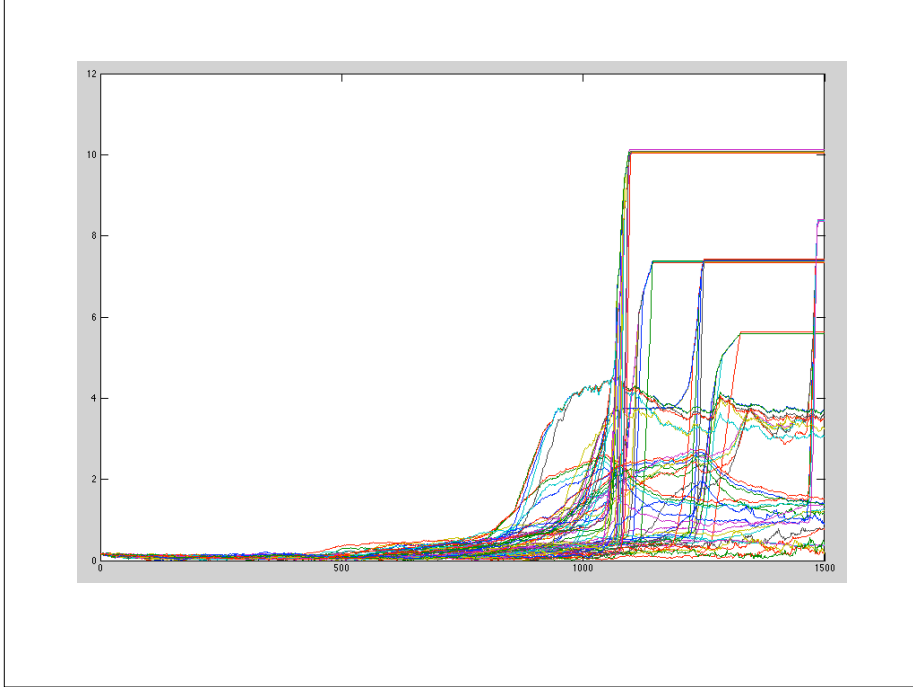
Figure 4: Evolution of output over time

A related feature of the model is that, even though individual search productivities are non-decreasing for all agents, some locations see output decline rather than grow. In the model, regions that are associated with dying languages exhibit lower productivity during the adjustment to the new set of language communities. The causality runs in both directions: areas of declining output receive fewer visitors in the model, contributing to the disappearance of the language.

After we control for the size of the language group, there may still be heterogeneity in output, for two reasons. The first is that there is individual heterogeneity in search costs that may be highly persistent. Second, if $\tau(s)$ is strictly concave, then agents who live near the boundaries of a language community will have lower output than those in the middle, because concavity of $\tau()$ penalises asymmetric search. In this particular simulation the variation in output largely disappears once we control for the size of the language group. In Figure 4, output of all individuals in a language group of a given size is approximately equal ($\tau()$ is linear and the variation due to differences in search productivity is small), which implies that total output will be be approximately equal in all language groups of a given size.

Proposition 1, part (iv), shows that, after conditioning on the size of the language group, the model exhibits convergence: in the limit, as search becomes highly productive, the output of each individual tends to a value equal to the size of the language group. While Proposition 1 considered only the case of linear $\tau()$, this result is evidently more general: as the productivity of search
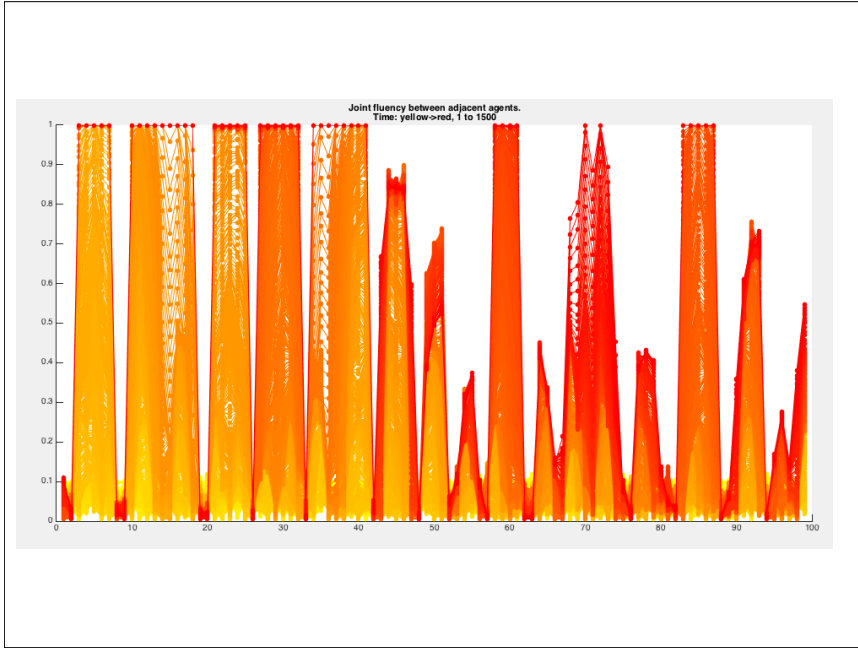
26

Figure 5: Languages emerging with high and low search productivity

increases, the concavity of $\tau()$ has a smaller and smaller effect. Thus the model does (eventually) imply conditional convergence.

5.3   The Relationship Between Mobility and Linguistic Diversity

Michalopolous (2012) argues that linguistic diversity is linked to past differences in specific human capital, which he in turn speculates were driven by geographical diversity. In some locations, according to his argument, human capital will be more tightly linked to the local environment, resulting in lower mobility. There is no direct analogue in my model to the diversity in land quality that, according to Michalopolous, lies behind the ethnolinguistic diversity observed in the modern world. However, variation in search productivity in the model potentially captures a similar notion. Agents who have lower productivity of search will optimally *choose* to be less mobile. Consistent with Michalopolous's hypothesis, this will translate into greater linguistic diversity.

Figure 5 shows the result of a simulation with 100 agents where agents 50-100 (on the right) have initial search productivity that is 40 percent of that of agents 1-50 (on the left).[31] After 1500 periods, these differences in the productivity of search translate into major differences in the language communities that have formed. On the left side of the figure five language groups with perfect joint fluency have formed, while on the right side of the figure the groups are smaller, more numerous, and typically have values of joint fluency that are considerably smaller than one. Thus, just as

---

[31]This simulation was run with the same set of parameters used in the simulations in Section 4.2.
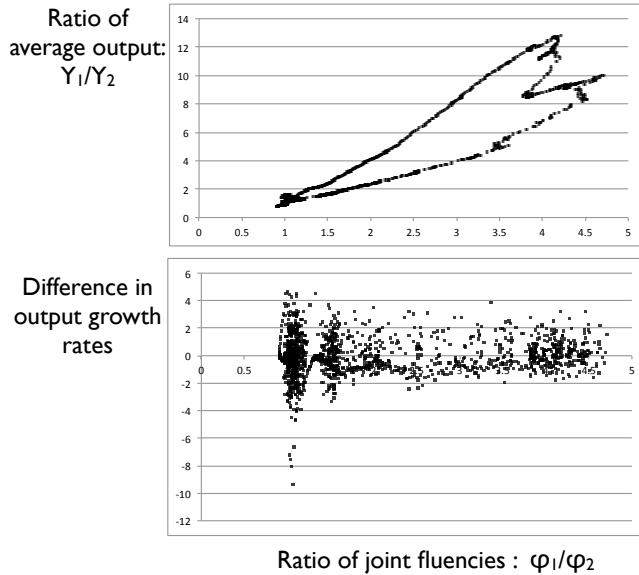
Figure 6: Linguistic diversity, output, and output growth

Michalopolous argues, variations in mobility (which evolve endogenously in the model) translate into variations in linguistic diversity.

5.4 The Relationship Between Linguistic Diversity and Output

As discussed in the introduction, researchers employing cross-country growth regressions have identified a negative relationship between linguistic diversity and output. For the most part, these papers have focused on the relationship between linguistic diversity and the growth rate of output, although Alesina et al (2003) note that linguistic diversity and output per capita are also negatively correlated.

The model in this paper does deliver a negative relationship between linguistic diversity and the *level* of output. Look again at Figure 5. Agents on the right hand side are less able to communicate, so will search less and produce less output. In addition, output is lower because of the lower search productivity. To quantify this, think of two economies: Economy 1 comprises agents 1-50 while Economy 2 comprises agents 51-100. The top panel of Figure 6 shows a scatterplot of the ratio of output in the two economies to the ratio of average joint fluency. When Economy 1's agents have a greater ability to communicate, relative to Economy 2, they also, on average, have much greater output. Note that, initially, the observations are clustered near the bottom left of the graph: the ratio of joint fluency is approximately one, and the ratio of output is about 1.6 due to Economy 1's higher search productivity. But as joint fluencies diverge in the two economies, output becomes
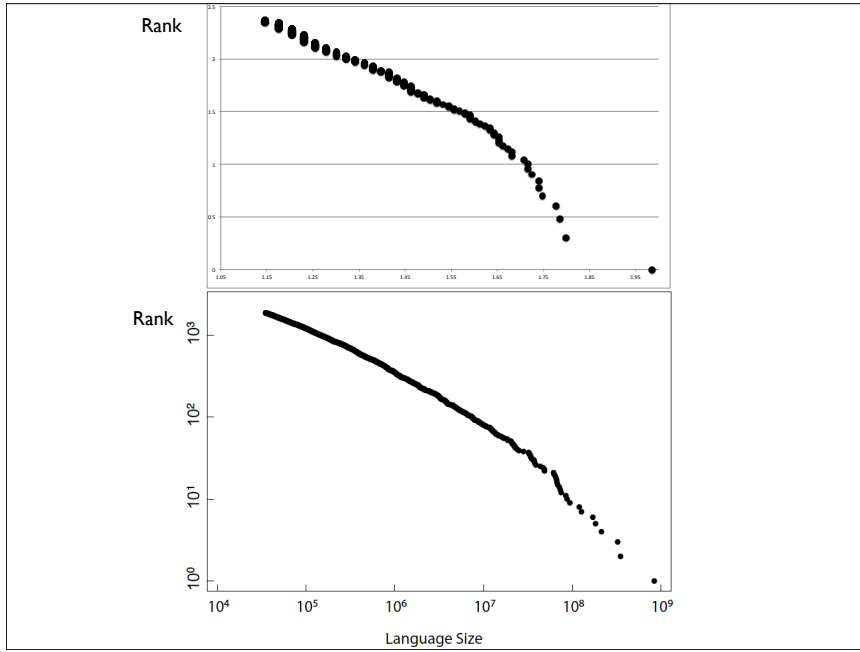
Figure 7: The size distribution of languages (a) as simulated in the current model and (b) as estimated by Clingingsmith

much higher in Economy 1 than in Economy 2 in almost all periods.

The lower panel reveals that this level effect does not translate into a correlation between linguistic diversity and output *growth*, however. It shows a scatterplot of the difference in the growth rates in the two economies to the ratio of average joint fluency. In both economies, output grows most rapidly when groups of agents transition from low joint fluency to high joint fluency—when this occurs, low joint fluency is associated with rapid growth.

5.5 The Distribution of Language Sizes

Clingingsmith's (2015) analysis of the distribution of language sizes found that, once above a threshold minimum size (35,000 speakers), the size distribution can be well-modelled by a double-Pareto distribution. Put differently, log-size is related to log-rank in a piecewise linear fashion. Estimating the size distribution of languages is difficult in the current model because the computational complexity of the simulation increases rapidly as the number of agents increased. I therefore carried out seven separate simulations with identical specifications, each with 1000 agents, looked at the sizes of the language groups that had emerged after 1000 periods, and pooled the results.[32]

I end up with 244 language groups in my simulation, ranging in size from 10 to 97. A small number of very small language groups emerge; I removed language groups smaller than 14 from the sample,

---

[32]I adopted a slightly different specification from the simulations reported earlier in order to get the distribution of languages to converge more quickly. Each run takes approximately two days. I emphasise, however, that I did not search for parameters that would yield the right distribution; I tried one specification only.

bringing the overall number down to 233. (Recall that Clingingsmith's estimation also removes small languages.) Figure 7 compares the results of this simulation (top panel) with the distribution plotted by Clingingsmith (bottom panel). (More precisely the top panel plots $\log_{10}$ of the rank of each language against $\log_{10}$ of the size; I use base 10 logs in order to get the closest comparison to Clingingsmith's picture.) My simulations yield results that resemble those of Clingingsmith quite closely—most noticeably, my graph yields a similar piecewise linear shape. The match is not perfect, however. In particular there are fewer relatively large languages in my simulation, so the plot is relatively steep on the right-hand side of the graph, whereas the difference in slopes is less marked in Clingingsmith's analysis of the data.

6 Concluding Comments

The emergence of language is one of the defining events of the human race. The subsequent development and evolution of languages, over the long sweep of history, reflects all kinds of forces: political, military, cultural, and economic. My focus in this paper is on one of those forces: the interplay between language change and economic change. The model is predicated on feedbacks between economic activity and language development. When individuals share a similar language, they have an economic incentive to produce and trade together; when individuals produce and trade together, their languages become more similar. These strategic complementarities reinforce the emergence of a common language. But complementarities operate negatively as well: when individuals do *not* share a similar language, they have less incentive to produce and exchange together, thus causing languages to drift further apart.

The results in this paper are demonstrations of possibility. It is *possible* for "languages"—that is, demarcated regions of high mutual fluency—to emerge endogenously as a result of the search decisions of individual agents. In constructing the examples in the paper, I have had considerable freedom to choose functional forms and set parameters. With the exception of one parameter, there is nothing so far that can be thought of as calibration. One reason for this is that, when a model is as stylized as the one presented here, it is not clear that calibration is very meaningful. Nevertheless, I plan to investigate whether there are any other functional forms and/or parameters that can be usefully pinned down by existing data.

The geography of this model is also very simple, and it would be interesting to extend the model to a two-dimensional world. Introducing migration—the ability to change location—would also be interesting: the forces that drive economic agglomeration in this model might also lead to spatial agglomeration. Introducing geographic heterogeneity of the kind in Michalopoulos (2012) is likely to be a fruitful extension: different locations could be modelled as having different search costs. Introducing population growth could potentially allow the model to be integrated into the

30

the unified growth literature. Finally, the production structure of the model is simple, and it would potentially be informative to develop a fuller model of production and trade. In terms of language, one possible route would be to combine the economic decision-making that drives this model with simulations of the kind studied in computational linguistics.

7 Appendix: Proofs of Propositions

Proposition 1. Assume that $\phi_{i.j} = 1 \forall i, j \in L_K$, $\tau(s) = s$, and $4y_0 < \gamma_{i;1} \leq 2K$. Then (i) for $\gamma_{i;t} < 2K$, the output of each agent evolves according to the linear second-order difference equation $\Delta y_{i;t+1} = \frac{\theta}{4} \Delta y_{i;t}$; (ii) the output of each agent initially grows at an increasing rate, reaching a maximum growth rate of $\frac{\theta}{4} - 1$, and thereafter grows at a decreasing rate; (iii) for $\gamma_{i;t} > 2K$, the output of each agent evolves according to the non-linear second-order difference equation $\tilde{y}_{i;t+2} = \frac{K \tilde{y}_{i;t+1}}{1 + \frac{\theta}{K} \tilde{y}_{i;t+1}(\tilde{y}_{i;t} - \tilde{y}_{i;t+1})}$, where $\tilde{y} = K - y$; (iv) as $t \to \infty, y_{i;t} \to K$.

Proof: (i) Let s_i be the amount of time spent searching and $1 - s_i$ the amount of time spent producing. (Given that all members of the community can communicate with perfect fluency, and given the linearity of $\tau(s)$, there is no need to distinguish the direction of travel.) The agent encounters $\gamma_i s_i$ trading partners, and produces total output equal to $y_i = (1 - s_i) s_i \gamma_i$, which trivially implies that the optimal search decision is $s_i^* = \frac{1}{2} \Rightarrow y_i = \frac{\gamma_i}{4}$. (By the assumption that $\gamma_t < 2K$, the agent does not wish to visit all the agents in the community, and so her choice is unconstrained.) It follows that

$$
\begin{aligned}
\Delta y_{i;t+1} &= y_{i;t+1} - y_{i;t} \\
&= \frac{1}{4}(\gamma_{i;t+1} - \gamma_{i;t}) \\
&= \frac{\theta}{4} \Delta y_{i;t}
\end{aligned}
$$

(ii) The growth rate of output between t and $t + 1$ equals

$$
\frac{\Delta y_{i;t+1}}{y_{i;t}} = \frac{\theta}{4} \frac{y_{i;t-1}}{y_{i;t}} \frac{\Delta y_{i;t}}{y_{i;t-1}}
$$

Hence the growth rate is increasing over time if $\frac{y_{i;t}}{y_{i;t-1}} < \frac{\theta}{4}$. If the initial conditions satisfy $y_0 < \frac{\gamma_1}{4}$, as assumed, then this condition is satisfied in the first period. Once the growth rate reaches $\frac{\theta}{4} - 1$, the growth rate starts to decline.

(iii) The constant level of search intensity combined with the growth in search productivity implies that, over time, agent i chooses to visit an increasing number of trading partners. Once $\gamma_{i;t} = 2K$, agent i reaches the corner solution where she optimally chooses to visit all the other K agents in the language community. Increases in the productivity of search are still beneficial, because they permit her to visit these agents with the expenditure of less time, thus freeing up more time for production. The amount of search time required to visit all K agents equals $\frac{K}{\gamma_{i;t}}$, leaving $1 - \frac{K}{\gamma_{i;t}}$ available for production. Output in period $t + 2$ is therefore given by $y_{i;t+2} = K(1 - \frac{K}{\gamma_{i;t+2}})$. Using the fact that $\gamma_{i;t+2} = \gamma_{i;t+1} + \theta(y_{i;t+1} - y_{i;t})$, inverting the expression for $y_{i;t+1}$ to obtain $\gamma_{i;t+1} = \frac{K^2}{K - y_{i;t+1}}$, substituting, and rearranging, gives the expression in part (iii) of the proposition.

32

(iv) If search could be undertaken at zero cost, agents would spend their entire time producing, and output would equal K. Thus, K represents the maximum possible level of output for an individual, and the variable $\tilde{y}$ in (iii) measures the gap between this maximum possible level of output and the actual level of output. This variable is declining over time. The dynamic equation $\tilde{y}_{i;t+2} = \frac{K\tilde{y}_{i;t+1}}{1+\frac{\theta}{K}\tilde{y}_{i;t+1}(\tilde{y}_{i;t}-\tilde{y}_{i;t+1})}$ has a unique steady state $\tilde{y} = 0$, which is globally stable so, in the limit, as $t \to \infty, \tilde{y}_{i;t} \to 0$—or, equivalently, $y_{i;t} \to K$.

Proposition 2. Assume that $\gamma_{i;t} > 2K$, $\phi_{i.j} = 1 \forall i, j \in L_K$, $\phi_{i.j} = 1 \forall i, j \in L_M$, $\tau(s) = s$. Assume that L_K and L_M are adjacent, and that joint fluency at the boundary equals $\hat{\phi} \Rightarrow \phi_{i.j} = \hat{\phi} \forall i \in L_K, j \in L_M$. If $\gamma_{i;t} > \frac{K(1+\hat{\phi})}{\hat{\phi}}$, agent i will trade across the language boundary.

Proof: Note first that an will never optimally choose to trade with an agent in L_M, unless she is already trading with all agents in L_K—that is (see Proposition 1) unless $\gamma_{i;t} > 2K$. So, assume that, at date t, agent i spends $s_i = \frac{K}{\gamma_{i;t}}$ units of time in travel to agents within L_K. Let $\hat{s}_i$ be the amount of time that the agent spends travelling to visit agents in L_M. The question is: when will she want to choose a strictly positive value of $\hat{s}_i$?

Thus, choose $\hat{s}_i$ to maximise $y_{i;t} = \gamma_{i;t}(1 - \hat{s}_i - \frac{K}{\gamma_{i;t}})(\hat{\phi}\hat{s}_i + \frac{K}{\gamma_{i;t}})$. The first-order condition from this problem yields

$$\hat{s}_i^* = \frac{\gamma_{i;t} - K}{2\gamma_{i;t}} - \frac{K}{2\hat{\phi}\gamma_{i;t}}$$
$$= \frac{\hat{\phi}\gamma_{i;t} - K(1+\hat{\phi})}{2\gamma_{i;t}\hat{\phi}}$$

(Note that if $\hat{\phi} = 1$, we obtain $s_i + \hat{s}_i = \frac{1}{2}$, as expected.) If the inequality in the statement of the proposition holds, then agent i chooses $\hat{s}_i > 0$.

References

[1] Abrams, D. and S. Strogatz,"Modelling the Dynamics of Language Death," *Nature*, Vol. 424, p. 900, 21 August 2003.

[2] Alesina, A., A. Devleeschauwer, W. Easterly, S. Kurlat and R. Wacziarg (2003), "Fractionalization", *Journal of Economic Growth*, 8,2, (2003)155-194.

[3] Anderson, J. and E. van Wincoop, "Trade Costs." *Journal of Economic Literature*, 42 (3) (2004), 691-751.

[4] Anderson, J. "Historical Linguistics" in K. Malmkjaer (ed.), *The Linguistics Encyclopedia*, 2nd edition, New York: Routledge, 2002, 218-247.

[5] Beine, M., F. Docquier and C. Özden, "Diasporas," Working Paper, Department of Economics, Université Catholique de Louvain, 2009.

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

[7] Beine, M., F. Docquier, and Caglar Özden, "Diasporas," manuscript, Department of Economics, Université Catholique de Louvain, 2009.

[8] Bisin, A., U. Horst, and O. Ozgür, "Rational expectations equilibria of economies with local interactions," *Journal of Economic Theory* 127(1) (2006), 74116.

[9] Carliner, G., "Wage Differentials by Language Group and the Market for Language Skills in Canada", *Journal of Human Resources*, 16 (1980), 384-399.

[10] Chen, H., Y. Chow, and L. Wu, "Imitation, local interaction, and coordination", *International Journal of Game Theory*, 42, 4 (2013), 1041-1057.

[11] Chen, K., "The Effect of Language on Economic Behavior: Evidence from Savings Rates, Health Behaviors, and Retirement Assets," Working Paper, School of Management, Yale University, 2012.

[12] Chiswick, B. and P. Miller, "The Endogeneity between Language and Earnings: International Analyses," *Journal of Labor Economics*, 13: 2 (1995), 246-88.

[13] Chiswick, B. and P. Miller, "English Language Fluency Among Immigrants in the United States," in S. Polachek (ed.) *Research in Labor Economics*, Volume 17, Oxford: JAI Press (1999), 151-200.

[14] Chiswick, B. and P. Miller, "Linguistic Distance: A Quantitative Measure of the Distance Between English and Other Languages," *Journal of Multilingual and Multicultural Develoipment*, 26(1) (2005), 1-11.

[15] Chiswick, B. and P. Miller, "Negative and Positive Assimilation, Skill Transferability, and Linguistic Distance," *Journal of Human Capital*, 6, 1 (2012), 35-55.

[16] Clingingsmith, D.,

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

[18] Comrie, B., S. Matthews and M. Polinsky (eds.). *The Atlas of Languages, Facts on File*: London, Bloomsbury and New York, 1996.

[19] Coulmas, F., *Language and Economy*, Oxford: Basil Blackwell, 1992.

[20] Crowley, T. and C. Bowen, *An Introduction to Historical Linguistics*, 4th edition, New York: Oxford University Press, 2010.

[21] Dixon, R., *The Rise and Fall of Languages*, Cambridge: Cambridge University Press, 1997.

[22] Desmet, K., I. Ortuño-Ortín, and S. Weber, "Linguistic Diversity and Redistribution," *Journal of the European Economic Association*, 7 (2009), 1291-1318.

[23] Dow, D., and A. Karunaratna, "Developing a Multidimensional Instrument to Measure Psychic Distance Stimuli", *Journal of International Business Studies*, 37 (2006), 578602.

[24] Easterly, W., and R. Levine, "Africas Growth Tragedy", *Quarterly Journal of Economics*, 112 (1997) 4, 1203-1250.

[25] Galeotti and Vega-Redondo, "Complex Networks and Local Externalities: A Strategic Approach", *International Journal of Economic Theory*, 7 (2011), 77-92.

[26] Galor, O., "From Stagnation to Growth: Unified Growth Theory," in P. Aghion and S. Durlauf (eds.), *Handbook of Economic Growth*, Elsevier: Volume 1 Part A (2005), 171-293.

[27] Garrouste, C. "Language Skills and Economic Returns", *Policy Futures in Education*, 6, 2 (2008),187-202.

[28] Greenberg, J., "The Measurement of Linguistic Diversity," *Language*, 32 (1956) 109-113

[29] Ginsburgh, V. and S. Weber, *How Many Languages Do We Need? The Economics of Linguistic Diversity*, Princeton: Princeton University Press, 2011.

[30] Goodfriend, M. and J. McDermott, "Early Development," *The American Economic Review*, 85(1) (1995), 116 - 33.

[31] Grenier, G., "The Effects of Language Characteristics on the Wages of Hispanic-American Males", *Journal of Human Resources*, 19 (1985), 35-52.

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

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

[34] Grin, F., "Economic Considerations in Language Policy, " in: Thomas Ricento (ed.), *An Introduction to Language Policy*, Oxford: Basil Blackwell, 2006, 77-94.

[35] Hall, R. and C. Jones, "Why Do Some Countries Produce So Much More Output Per Worker Than Others?," *Quarterly Journal of Economics*, 114:1 (1999), 83-116.

[36] Helliwell, J., "Language and trade', in A. Breton (ed.) *Exploring the Economics of Language*, Ottawa: Canadian Heritage (1999), 5-30.

[37] Hutchinson, W., "'Linguistic Distance' as a Determinant of Bilateral Trade," *Southern Economic Journal*, 72(1) (2005), 1-15.

Isphording, I. and S. Otten, "The Costs of Babylon—Linguistic Distance in Applied Economics," Working Paper, Ruhr-University Bochum, 2011.

[38] John, A., and K. Yi, "Language, Learning, and Location," *Federal Reserve Bank of New York Staff Report Number 26* (1997).

[39] John, A., and K. Yi, "Language and Location," *Working Paper* (2001).

[40] Jones, E., "The Case for a Shared World Language", in M. Casson and A. Godley (eds.), *Cultural Factors in Economic Growth*, Berlin: Springer (2000), 210-235.

[41] Kremer, M. "The O-Ring Theory of Economic Development," *Quarterly Journal of Economics*, 108, August (1993a), 551-575.

[42] Kremer, M. "Population Growth and Technological Change: 1,000,000 B.C. to 1990," *Quarterly Journal of Economics*, 108, August (1993b), 681-716.

[43] Lagos, R., and R. Wright,"A Unified Framework for Monetary Theory and Policy Analysis. *Journal of Political Economy*, 113(2) (2005), 46388.

[44] 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.

[45] Lang, K., "A Language Theory of Discrimination," *Quarterly Journal of Economics*, 101 (1986): 363-382.

[46] Lazear, E., "Culture and Language," *Journal of Political Economy*, 107, December (1999) s95: s126.

[47] Lewis, M. P. (ed.), *Ethnologue: Languages of the World*, Sixteenth edition. Dallas, Tex.: SIL International, 2009. Online version: http://www.ethnologue.com/.

[48] Lohmann, J., "Do Language Barriers Affect Trade," *Economics Letters*, 110(2) (2011), 159-162.

[49] Luo, Y. and O. Shenkar, "The Multinational Corporation as a Multilingual Community: Language and Organization in a Global Context," *Journal of International Business Studies*, 37 (2006), 321-339.

[50] Lyovin, A. V., *An Introduction to the Languages of the World*, New York: Oxford University Press, 1997.

[51] Malmkjær, K. "Creoles and Pidgins," in K. Malmkjær (ed.), *The Linguistics Encyclopedia*, 1st edition, New York: Routledge, 1991, 107-117.

[52] Malmkjær, K. "Origin of Language," in K. Malmkjær (ed.), *The Linguistics Encyclopedia*, 2nd edition, New York: Routledge, 2002, 436-441.

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

[54] Michalopoulos, S. "The Origins of Ethnolinguistic Diversity," *American Economic Review*, 102 (4) (2012), 1508-1539.

[55] Melitz, J., "Language and Foreign Trade," *European Economic Review*, 52 (2008), 667-699.

[56] Mookherjee, D., S. Napel, and D. Ray, "Aspirations, Segregation and Occupational Choice," *Journal of the European Economic Association* 8 (2010), 139-168.

[57] Morris, S., "Contagion", *Review of Economic Studies*, 67 (2000), 57-78.

[58] Nerbonne, J., "Measuring the Diffusion of Linguistic Change," *Philosophical Transactions of the Royal Society B*, 365 (2010), 3821-3828.

[59] Nichols, J., *Linguistic Diversity in Space and Time*, Chicago: University of Chicago Press, 1992.

[60] Nettle, D., *Linguistic Diversity*, New York: Oxford University Press, 1999.

[61] Nettle, D. and S. Romaine, *Vanishing Voices: The Extinction of the World's Languages*, New York: Oxford University Press, 2000.

[62] Pinker, S., *The Language Instinct*, New York: William Morrow and Company, 1994.

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

[64] Romaine, S., *Bilingualism*, 2nd ed., Oxford: Basil Blackwell, U.K., 1995.

[65] Rose, A., "Do We Really Know that the WTO Increases Trade?" *American Economic Review*, 94:1 (2004), 98-114.

[66] Rubinstein, A., *Economics and Language*, Cambridge: Cambridge University Press, 2000.

[67] Schelling, T., *The Strategy of Conflict*, Cambridge: Harvard University Press, 1960.

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

[69] Tamura, R., "Language Homogeneity and Income", Clemson University Working Paper (1997).

[70] Tamura, R., "Translators: Market Makers in Merging Markets," *Journal of Economic Dynamics and Control*, 25 (2001), 1775-1800.

[71] Vaara, E., J. Tienari, R. Piekkari, and R. Sntti, "Language and the Circuits of Power in a Merging Multinational Corporation", *Journal of Management Studies*, 42 (2005), 3595-3623.

[72] Vaillancourt, F., "Difference in Earnings by Language Groups in Québec, 1970: An Economic Analysis", Québec: Centre International de Recherche sur le Bilinguisme, B-90 (1980).

[73] Welch, D. E., Welch, L. S., and Piekkari, R. "Speaking in tongues: The importance of language in international management processes." *International Studies of Management and Organization*, 35(1) (2005), 10-27.

[74] Wichmann, S., E. Holman, D. Bakker and C. Brown, "Evaluating Linguistic Distance Measures," *Physica A*, 389 (2010), 3632-3639.

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

[76] Young, H.P., "The Dynamics of Conformity," in S. Durlauf and P. Young (eds.), *Social Dynamics*, Washington: Brookings Institution Press, 2001, pp. 133-153.

[77] Zhang, W. and G. Grenier, " How Can Language Be Linked to Economics? A Survey of Two Strands of Research, *Language Problems and Language Planning* (2013), 203-226.