Language competition: An economic theory of language learning and production

Harald Wiese

June 2016

Abstract

This paper employs game theory to contribute to sociolinguistics (or the economics of language). From both the synchronic and the diachronic perspective, we are interested in the conditions (of language learning and literary production) that make some languages dominate others. Two results are particularly noteworthy: (i) Translations have an ambiguous effect on domination. (ii) We offer three different explanations of how a past language like Latin or Sanskrit can develop into a standard for literary production.

1. Introduction

This paper discusses language competition. We build simple game-theoretic models that reflect upon English as a global language and Latin or Sanskrit as old languages that have been in use for literary production long after they ceased to be (if they ever were) "mother tongues". Thus, we try to contribute to the economics of language.

A very broad take on language and economy is Coulmas (1992). Grin & Vaillancourt (1997, p. 43) define the economics of language as the study of (i) "the relationships between linguistic and economic variables" and also of (ii) "language-related issues where economic variables have little or no part, but which can nevertheless be examined with the concepts and methods of neo-classical economics." Referring to the surveys by Grin (1996) and Grin & Vaillancourt (1997), our paper is not about socioeconomic variables like consumption, income (inequality), or status, all of which have been examined in relation to languages or language skills, nor about nationalism and language planning. Thus, it does not come under part

(i) of the above definition, but quite clearly under part (ii). More specifically, it can be seen as an application of methods known from an economic subfield called Industrial Organization to language production and learning.

Industrial Organization (often abbreviated by IO) is concerned with theoretical and empirical work on the behavior of monopolistic and oligopolistic firms. These firms set prices, advertise their products, or decide on research and development, etc., while keeping their customers and competitors in mind. Whenever more than two firms are involved, IO uses game theory and employs solution concepts like the Nash equilibrium or backward induction. The standard treatise on IO is Tirole (1988).

Fitting to the IO framework, we model literary producers (the firms) and readers (the customers). We first turn to the readers. They enjoy literature with this enjoyment depending on the language used in that literature. In order to gain access to literature written in foreign languages, readers may decide to learn those languages. These learning decisions depend (i) on the cost of learning and also (ii) on the literature made available to the reader. Thus, language learners are depicted as rational deciders paying attention to the instrumental value of languages, only. As objections against this modelling decision, one may point to symbolic domination (see Bourdieu (1977) and, for a case-based application, Heller (1995)), to the passion fueled by language issues (see Grin (1996, pp. 28-29)) and to motivation and affection (see the recent collection of articles edited by Gabrys-Barker & Bielska (2013)). We do not intent to deny the importance of these aspects. However, for reasons of tractability, we think it a wise decision to disregard them in the models presented here.

The difficulty of language learning is a thorny issue. In our paper, we assume that the readers can weigh their cost of learning against the utility of being able to read a wider range of literary production. We also take into account that people with a given mother tongue may well have different cost of learning a second language (i.e., different talents for foreign languages). In some models, we allow for languages that are "easier" to learn (across all potential learners) than others. Here, we do not elaborate on the findings of contrastive linguistics (see the seminal work by Lado (1957) and the more recent edited volume by González, Mackenzie & Álvarez (2008)) that ease of learning depends on the genetic or typological relationship of the learners' first language and the one they endeavor to learn. In other models, we assume that the languages in question have the same difficulty level.

Contrastive issues are also relevant for another central element in our modelling

that we call accessibility. A reader has full access to literary products employing his mother tongue. He may also profit from literature employing another language when he has learned that language. There may be two different reasons for a reduced accessibility. First, the reader may read the original foreign literature and understand part of it, only. Second, the reader may obtain a translation of the foreign literature. In that case, reduced accessibility stands for the limited amount and/or quality of literary works that are translated into the reader's mother tongue. Or, it may be an expression of the dislike to read translated, rather than original, literature.

The producers in our model do not try to maximize profits as would be the usual assumption in IO models. Instead, they strive to maximize "readership". They do so by choosing a language for producing literature. Naively, literary production can be measured in the number of pages produced or in the number of books (of a certain minimum quality). The assumption of readership maximization seems plausible in different respects. First, a producer may have material reasons for "selling" his books or for winning the patronage of a king (then we are very close to the profit motive of IO). Second, he may want to press a political or religious agenda.

Readership can be operationalized in different manners. It may be based on option demand (see, in the context of prices, Schmalensee (1972)). In that model, readers do not benefit from actual reading, but from the possibility to choose from a wide range of literary products. In other words, the readers care for the option to have many different books at their disposal. An alternative model deals with non-option demand. Here, the readers' benefit revolves around the actually read books. Obviously, both models (option demand and non-option demand) are extreme cases. We consider these extreme cases because we can learn something from each of them, but we certainly need to keep in mind the restrictive assumptions employed.

Readership may also be understood as "current" or as "long-term". Current readership refers to readers that are alive during the producer's lifetime. Alternatively, a literary producer may take a long-term view. For example, the author or the group he represents may want long-term recognition (fame) and influence. This case may be relevant for religious or political groups. Or the author may sell the book rights during his lifetime and therefore has a profit motive for writing for future generations, also.

Our paper is not the first to employ game theory to linguistics. Following the work by Parikh (see, for example, Parikh 2001), game theory has entered

linguistics in efforts to explore pragmatics and, in particular, meaning (see, for example, the introduction to the edited volume by Benz, Ebert, Jäger & Rooij (2011) and also the book by Clark (2011)). Evolutionary game theory has also been used by some linguists (see part II in Benz et al. (2011)). It seems that this part of game theory has, as yet, not been applied to the problem addressed in the current paper: How do the number of speakers, producers, and learners of languages evolve over time?

An early game-theoretic approach to the acquisition of foreign-language skills is Selten & Pool (1991). In that paper, the size of language communities and the learning cost are important factors. The learners strive for communicative benefit and, in contrast to our model, literary production is absent. The same general features are present (and absent) in the model by Church & King (1993). However, since these authors employ more specific assumptions, they can deduce more specific results than Selten & Pool (1991). In particular, they focus on the question of how the cost of learning influences second-language learning.

It seems to us that this paper is the first to employ the IO approach (where both producers and consumers are relevant) to literary production and language learning. In a rough manner, our paper builds on, and reconstructs some stylized facts. We take as given the following facts:

L Language learning:

People tend to learn (L) foreign languages that have many speakers or that exhibit extended literary production, both current and past.

R Language choice by producers:

Authors tend to favor languages that allow for a large readership (R) of their literary products.

Ct Language change, over time:

Languages change over time (Ct) and accessibility is a decreasing function of time gone by.

Cf Language change, forking:

A language 0 may fork into daughter languages 1 and 2. Then, languages 1 and 2 may (or may not) be less accessible from each other than language 0 from language 1 or language 2.

We comment on (Ct) and (Cf). Examples for (Ct) are easy to find. Latin is more accessible from Old French than from New French. Sanskrit is more accessible from Pali than from Hindi.

In contrast, (Cf) is contentious. Forking is to be understood within a tree model of language change. The tree model (also called genetic or cladistic model) has been under attack by linguists (mainly dialectologists) for well over 100 years. François (2014) is a recent summary of alternative models that come under the headings of waves, linkages, or continua[1]. Surely, these alternative models capture language change more accurately than the rather coarse tree model. However, a justification for language trees is provided by Gray & Atkinson (2003, p. 436) who apply some sort of bayesian glottochronology to Indo-European languages and report a "strong tree-like signal". Indeed, comparative linguists seem comfortable using the tree model in their reconstructive work (see, for example, Lehmann (1992, p. 121), Fortson IV (2004, pp. 8-11), or Clackson (2007, pp. 5-15)). Summarizing, despite of the justified critique voiced by many linguists, language trees remain the workhorse model for a lot of relevant work in the diachronic domain. It is in this sense that (Cf) is employed in the current paper.

The accessibility question is also difficult. Assume that some language 0 undergoes two changes, C1 that produces language 1 and C2 that produces language 2. Then, language 1 and language 2 are different from each other by both changes. Instead, language 1 and language 0 differ by one of these changes, only. By this argument, languages 1 and 2 should be less accessible from each other than language 0 from language 1 or language 2. Against this argument, one may point to Spanish and Italian that are more easily accessible from each other than Latin is from Spanish or even Italian. Here, mutual borrowing and the like have lead to a reapproachment.

Using the facts (L) through (Cf), we reconstruct the following stylized facts:

D Dominance:

Both a contemporary language (as English nowadays) or a "dead" language (like Latin or Sanskrit) may prevail over other languages for oral or written purposes.

---

[1]With a view to the continuum perspective, one may question the usual procedure of treating languages as countable objects. Here, again, we like to argue that the discreteness of languages is a simplifying assumption that allows to do comparative work and that allows to do the sort of theoretical exercise we have in mind. It is also a fact that people that are not linguists clearly think and act in terms of discrete languages rather than continua of languages.

N Niche:

> Readers competent in small languages have a smaller choice of literature, but literary production does not totally sideline them.

We begin with synchronic models (next section). In these models, languages coexist (like French and German). We consider two extreme models. The first might be called the option-demand model. In that model, readers do not benefit from actual reading, but from the possibility to choose from a wide range of literary products. In other words, the readers care for the option to have many different books at their disposal. The second model deals with non-option demand. Here, the readers' benefits revolve around the actually read books.

Within the synchronic, option-demand model, we obtain these results. In line with (D), the chances of a language to dominate others are high if the population using that language is large, if the past production is high, and if it is relatively easy to learn. Also, if we have several literary producers, either language may (within a certain range of parameters) become dominant. In contrast, N results from non-option demand.

We then turn to diachronic models (section 3) where languages change over time (as Latin has developed into French). In these models, we investigate the conditions under which production makes use of old languages. We identify two different settings for this to happen. First, producers need to care for future readership, not just for current readership. Second, in the competition between synchronic languages, vernaculars may lose out to their common predecessor language.

2. Synchronic models

2.1. Setup

In our synchronic models, we have two periods $t = p, n$ (see figure 2.1). Here, p stands for "past" and n for "now". We concentrate on period n. Hence, we build a static model. We assume two languages called 1 and 2. In period n, there are q_1 speakers with mother tongue 1 and q_2 speakers with mother tongue 2.

By n_1 (and n_2), we denote language production in the current period that employs language 1 (language 2, respectively). Overall literary production in period n also uses the symbol n and is given by $n = n_1 + n_2$. We assume an "installed base" for language 1 denoted by p_1. This literary production is as accessible as n_1. p_2 is the installed base of literary production in language 2.
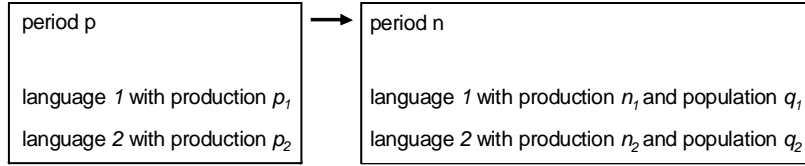
Figure 2.1: The synchronic model

We consider the language learning decisions in the following subsection. We then proceed to the production decision taken by one producer (subsection 2.3) or by several producers (subsection 2.4). In these models, we define readership in an option-demand manner. In subsection 2.5, we assume actual enjoyment, only, and disregard language learning.

2.2. Language learning

The speakers perfectly know their mother tongues, but may choose to learn the foreign language with some cost c. We assume that the language learners know the literary production in languages 0 and 1 and make their learning decisions accordingly. In particular, we work with this payoff (or utility) function for a native speaker of language 1:

$$u_1 = \begin{cases} p_1 + n_1 + (p_2 + n_2) - c, & \text{1 learns language 2} \\ p_1 + n_1 + \alpha_2 (p_2 + n_2), & \text{1 does not learn language 2} \end{cases}$$

Note the different accessibilities for a native speaker of language 1 to the literary products of language 2. It equals 1, if 1 has learned language 2. Thus, we find $1 \cdot (p_2 + n_2) = (p_2 + n_2)$ in the first line of utility function u_1. In the second line, we have $\alpha_2 (p_2 + n_2)$ instead. For example (see also the introduction), a Spanish reader may partly understand Italian literary works (accessibility α between 0.6 and 0.8), while a German reader's accessibility to Italian writings is close to zero. In general, we assume $0 \leq \alpha_1 < 1$ and $0 \leq \alpha_2 < 1$. Thus, a speaker of language 1 has a reduced benefit of the literary production in language 2 if he has not learned language 2. A native speaker of language 1 benefits from learning the other language 2 if

$$c < (p_2 + n_2)(1 - \alpha_2) =: \bar{c}_2 \tag{2.1}$$

holds. Inversely, language learning is not profitable for $c > \bar{c}_2$. Here, as in the rest of the paper, we will not be concerned with the (highly improbable)[2] case of $c = \bar{c}_2$. We assume that the speakers' costs of language learning are uniformly distributed on $[0, C_2]$ with $C_2 > n + p_2$. The latter assumption implies that there are always agents who find language learning too hard. Then,

$$\frac{\bar{c}_2}{C_2} = \frac{(p_2 + n_2)(1 - \alpha_2)}{C_2} < 1 \qquad (2.2)$$

is the proportion of readers of language 1 that learn language 2. This result is in line with (L) in the introduction.

Analogously, we define $\bar{c}_1 := (p_1 + n_1)(1 - \alpha_1)$ and obtain the proportion of readers of language 2 that learn language 1

$$\frac{\bar{c}_1}{C_1} = \frac{(p_1 + n_1)(1 - \alpha_1)}{C_1} < 1, \qquad (2.3)$$

where we assume $p_1 + n < C_1$.

We say that language 1 dominates language 2 with respect to language learning if the percentage of language-1 learners $\frac{\bar{c}_1}{C_1}$ is larger than the percentage of language-2 learners $\frac{\bar{c}_2}{C_2}$. This tends to hold under the following conditions:

- It is relatively easy to learn language 1 ($C_1 < C_2$).

- In the past, literary production in language 1 was relatively large ($p_1 > p_2$).

- The current literary production in language 1 is relatively large ($n_1 > n_2$).

- Accessibility for language-2 speakers to language 1 is relatively small ($\alpha_1 < \alpha_2$).

We comment on the last bullet. In this model, language-learning dominance of language 1 is furthered if there are many and good translations of language-2 literature into language 1. In this case, speakers of language 1 do not have large incentives to learn language 2. From this result, we obtain an immediate policy conclusion. If a country (let us say, France) wishes to make French dominant with respect to language learning, it should translate important works of foreign languages into French (or should subsidize these translations). It should not, however, further

---

[2]Since we are dealing with a continuum of learning costs, the probability of c being equal to a particular value is zero.

translations of French works into foreign languages. Note that the French government sponsored "Centre national du livre" (www.centrenationaldulivre.fr) subsidizes translations in both directions. In contrast, the German Goethe-Institut (www.goethe.de) focuses on translations of German books into foreign languages.

Note, however, that these recommendations follow from the specific model employed. Like all models, some aspects may be left out that possibly militate against these recommendations. For example, one may translate Victor Hugo into other languages and thereby increase the interest for French culture and literature. This increased interest may then feed into more people learning French. This effect is not captured by our model. Another effect comes into play when we consider literary production on top of the learning decision (see the following two subsections).

2.3. Literary production with one producer (option demand)

We first assume that there is one producer, only. The producer chooses n_1 and n_2 in order to maximize his readership (see (R) in the introduction). We assume that n is fixed and that the division of $n = n_1 + n_2$ between languages 1 and 2 is up to the literary producer. That is, the producer has a given capacity for literary production and chooses between languages. Building on the payoff function above, we define the readership for literary production n_1 (and hence $n_2 = n - n_1$) in a very simple and naive manner. We proceed in two steps. First, we define readership for literary production n_1 as n_1 times

$$
1 \cdot q_1 + q_2 \left(1 \cdot \underbrace{\frac{\bar{c}_1}{C_1}}_{\substack{\text{proportion of readers} \\ \text{with mother tongue 2} \\ \text{who have learned 1}}} + \alpha_1 \left[\underbrace{1 - \frac{\bar{c}_1}{C_1}}_{\substack{\text{proportion of readers} \\ \text{with mother tongue 2} \\ \text{who have not learned 1}}} \right] \right)
$$

Readers with mother tongue 1 have full access to n_1 which explains $1 \cdot q_1$. Readers with mother tongue 2 consist of two types. Either they have learned language 1 (proportion $\frac{\bar{c}_1}{C_1}$) and then have full access. Or they have not learned language 2 and hence are characterized by limited access α_1.

Turning to the seccond step, we define (overall) readership for the literary producer by

$$R\left(n_1\right) = n_1\left[q_1 + q_2\left(\frac{\bar{c}_1}{C_1} + \alpha_1\left[1 - \frac{\bar{c}_1}{C_1}\right]\right)\right]$$
$$+ (n - n_1)\left[q_2 + q_1\left(\frac{\bar{c}_2}{C_2} + \alpha_2\left[1 - \frac{\bar{c}_2}{C_2}\right]\right)\right]$$

This definition requires a few comments: First, we have one producer. He can be understood as representative for a diverse set of producers. Alternatively, we can consider him a "benevolent dictator" who strives to maximize readership (if readership is a measure of welfare). Second, the producer knows both languages. Third, readership is of the option-demand type. That is, a reader likes the option to choose from a diverse selection of literature, over and above his enjoyment from actually reading part of that literature. In the subsection after next, we present a non-option model.

Proposition 2.1. *In the synchronic model with 1 (representative) producer (benevolent dictator), language 1 or language 2 become the exclusive literary language for production (compare (D)). In particular, language 1 tends to become the standard language if*

- *the population q_1 of language 1 is large (relative to the population q_2 of language 2),*

- *the cost C_1 of learning language 1 is small (relative to the cost C_2 of learning language 2), or*

- *the literary base p_1 of language 1 is large (relative to the literary base p_2 of language 2).*

The effect of the accessibility parameters on language adoption is ambiguous.

Proof: see appendix.

The effects of the population sizes, the cost of learning, and the literary bases on language adoption are not too surprising. We want to discuss the two interesting aspects of this proposition. First, only one of the languages is used for literary production. This extreme result is due to the option-demand characteristic built into the definition of readership above.

Second, the effect of the accessibility factors is ambiguous. An increase in α_1 has two opposing effects. At first, we disregard language learning, i.e., let (2.3) be constant focussing on the "direct effect" of an increased α_1. The readership function

$$
\begin{align}
R\left(n_1\right) &= n_1\left[q_1 + q_2\left(\frac{\bar{c}_1}{C_1} + \alpha_1\left[1 - \frac{\bar{c}_1}{C_1}\right]\right)\right] + \ldots \tag{2.4}\\
&= n_1\left[q_1 + q_2\left(\alpha_1 + \frac{\bar{c}_1}{C_1}\left(1 - \alpha_1\right)\right)\right] + \ldots \tag{2.5}
\end{align}
$$

reveals that the readership of language 1 increases if accessibility to language 1 increases for non-learners of language 1 (see 2.4). Thus, by this direct effect, the incentives to use language 1 as a medium of production increase with α_1.[3]

However, there is also the "indirect effect" that works through the learning decision of language-2 speakers. They are less enthusiastic about learning language 1 if language 1 is more accessible. This can be seen from the proportion of readers with mother tongue 2 who learn language 1

$$
\frac{\bar{c}_1}{C_1} = \frac{\left(p_1 + n_1\right)\left(1 - \alpha_1\right)}{C_1}
$$

where we have $\frac{d\bar{c}_1(\alpha_1)}{d\alpha_1} = -\left(p_1 + n_1\right) < 0$. Only those readers with mother tongue 2 that have very small learning cost c will endeavor to learn language 1 if α_1 is large. This language-learning effect reduces the producers' incentives to employ language 1 (see 2.5). Thus, we have two opposing effects of accessibility on language adoption.

2.4. Literary production with several producers (option demand)

Let us now turn to two or more producers. We focus on a producer A whose overall production n^A is divided between language 1 $\left(n_1^A\right)$ and language 2 $\left(n_2^A = n^A - n_1^A\right)$.

---

[3]Assume, alternatively, that language learning is very costly and that there is no language learning at all. Then, the producer is the only decision maker. The readership is the one from above with $\bar{c}_1 = \bar{c}_2 = 0$, i.e., it is given by

$$
\begin{align}
R\left(n_1\right) &= n_1\left(q_1 + q_2\alpha_1\right) + \left(n - n_1\right)\left(q_2 + q_1\alpha_2\right)\\
&= n_1\left(q_1 + q_2\alpha_1 - \left[q_2 + q_1\alpha_2\right]\right) + n\left(q_2 + q_1\alpha_2\right).
\end{align}
$$

Thus, language 1 is chosen if $\left(q_1 + q_2\alpha_1 - \left[q_2 + q_1\alpha_2\right]\right)$ is positive or, equivalently, if $\frac{q_1}{q_2} > \frac{1-\alpha_1}{1-\alpha_2}$ holds. Therefore, the effect found in the text above is also obtained if there is no language learning.

The other producers also employ language 1 (n_1^{others}) and language 2 (n_2^{others}). A's readership can then be defined as

$$
\begin{aligned}
& R^A\left(n_1^A\right) \\
= & \ n_1^A\left[q_1 + q_2\left(\frac{\bar{c}_1}{C_1} + \alpha_1\left[1 - \frac{\bar{c}_1}{C_1}\right]\right)\right] \\
& + n_2^A\left[q_2 + q_1\left(\frac{\bar{c}_2}{C_2} + \alpha_2\left[1 - \frac{\bar{c}_2}{C_2}\right]\right)\right]
\end{aligned}
$$

It depends on the production decision by author A, on the populations using languages 1 and 2 as mother tongues (q_1 and q_2, respectively), and on the learning decisions taken by these populations.

As shown in the appendix (proof of proposition 2.2), a producer tends to find production in that language attractive that is used by most other producers. Therefore, we are justified in looking for symmetric equilibria, only. In these equilibria, all producers choose the same language for literary production. Interestingly, there may exist two equilibria:

Proposition 2.2. *In the synchronic model with m symmetric producers, language 1 rather than language 2 becomes the exclusive literary language for production (compare (D)) under the broad conditions stated in proposition 2.1 above. Furthermore, we find:*

(a) *If $\frac{q_1}{q_2}$ is small (large), language 2 (language 1) becomes the exclusive literary language for production.*

(b) *In a medium range of $\frac{q_1}{q_2}$, there exist two symmetric equilibria. In one of them, all producers choose language 1, in the other, all producers choose language 2. The larger the number of producers, the larger the range with two possible equilibria.*

(c) *For two or more producers, the readership maximizing language may not be adopted in equilibrium.*

Thus, the important variable is $\frac{q_1}{q_2}$, the ratio of the population sizes (see also figure 2.2). The exact definition of the thresholds and the proof is given in the appendix. According to (b), there exist parameter constellations such that all producers can coordinate on language 1 or all producers can coordinate on language
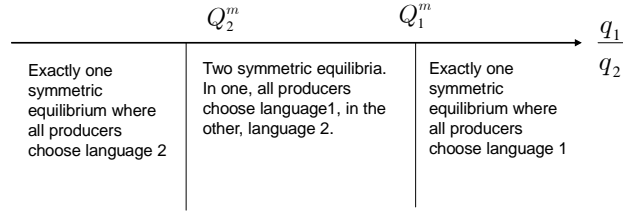
12

Figure 2.2: One or two equilibria for at least two producers

2. The larger the number of producers m, the more extended is the range of $\frac{q_1}{q_2}$ for which either language might be adopted in equilibrium. Although the model is static, a snowball-effect like mechanism might occur: One producer adopts a language and therefore, that language is more attractive to language learners, so that other producers also tend to adopt it. (c) claims that an equilibrium may not be optimal (in the sense of readership maximization).

2.5. Non-option demand

The model in this section differs from the preceeding ones in two respects. First, up to now, readership was defined by some product $q \cdot n$ where q refers to the quantity of potential readers and n to the quantity of literary products. Against this option-demand setup, one might argue that readers have a limited capacity of actual reading. Let us, therefore, understand q as the capacity for reading in the overall population. For example, each reader spends some specific time interval on reading (for example, 10 days a year). Second, in order to keep the model tractable, we disregard learning and assume that readers of language community 1 use language 1 only.

Assume two producers A and B and two languages 1 and 2 with capacities for reading q_1 and q_2, respectively. Literary production of producers A and B employ languages 1 and 2. By n_1^A we denote A's literary production that uses language 1, and so on. We also define

$$
\begin{aligned}
n^A &:= n_1^A + n_2^A, \\
n^B &:= n_1^B + n_2^B.
\end{aligned}
$$

and take n^A and n^B as exogenous.

13

Readerships for producers A and B are defined as

$$R^A\left(n_1^A, n_1^B\right) = q_1 \frac{n_1^A}{n_1^A + n_1^B} + q_2 \frac{n_2^A}{n_2^A + n_2^B},$$

$$R^B\left(n_1^A, n_1^B\right) = q_1 \frac{n_1^B}{n_1^A + n_1^B} + q_2 \frac{n_2^B}{n_2^A + n_2^B}$$

and we find $R^A + R^B = q_1 + q_2$. Again, we assume that producers A and B know how to write in both languages 1 and 2.

We obtain

Proposition 2.3. *In the synchronic reading-capacity (non-option demand) model, the unique Nash equilibrium is given by*

$$\left(n_1^A\right)^* = n^A \frac{q_1}{q_1 + q_2} \text{ and}$$

$$\left(n_1^B\right)^* = n^B \frac{q_1}{q_1 + q_2}.$$

Proof: see appendix.

This proposition shows that languages supported by small communities may subsist (see (N) in the introduction). For example, in many languages (except for the very small ones), journals or magazines are produced. The most notable feature of the equilibrium is proportionality:

- The production using languages 1 and 2 is proportional to the reading capacities:
$$\frac{\left(n_1^A\right)^* + \left(n_1^B\right)^*}{\left(n_2^A\right)^* + \left(n_2^B\right)^*} = \frac{q_1}{q_2}$$

- Each producer uses both languages for production, again proportionally to the reading capacities:
$$\frac{\left(n_1^A\right)^*}{\left(n_2^A\right)^*} = \frac{\left(n_1^B\right)^*}{\left(n_2^B\right)^*} = \frac{q_1}{q_2}$$

If we compare the option demand models in the previous subsection to the non-option demand one in this subsection, we immediately realize an important difference. If producers want to be read (this subsection) and not just to be potentially read (the previous subsections), they will also produce for smaller language communities. This explains why, after 5000 years of literacy, not all literary production is in one language.
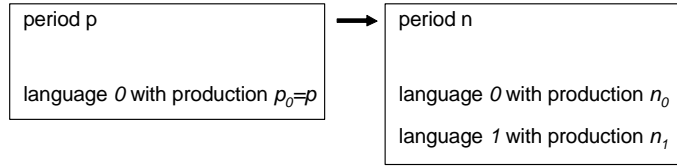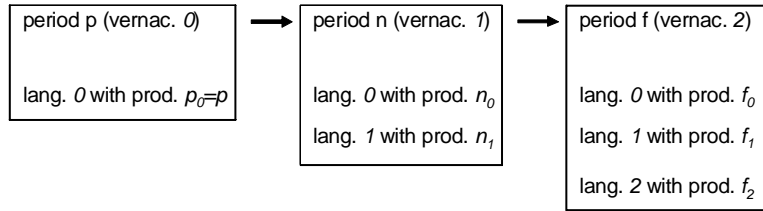
14

Figure 3.1: The diachronic model with two periods

Figure 3.2: The diachronic model with three periods

3. Diachronic models

3.1. Setup

Turning back to option demand and one producer, we now consider one language and its changes over time (Ct). We begin with two periods $t = p, n$ (see figure 3.1). We assume that the vernacular changes over time and call language 0 the vernacular in period p and language 1 the vernacular in period n. The speakers perfectly know their vernaculars. Speakers in period n (who have language 1 as their mother tongue) can learn language 0 at cost c.

The literary productions in the periods p and n are denoted by p and n, respectively. Literary production in period p takes place in the vernacular employed in p, language 0. The quantity of that production is denoted by p or p_0. Literary production in period n can employ language 0 (the language of the past) or language 1, the vernacular of period n. Denoting production in period n by n_0 and n_1, respectively, we then have the literary production

$$n = n_0 + n_1$$

in period n.

After the two-period model (subsection 3.2) we turn to a three-period one (subsection 3.3). The periods are denoted as p ("past"), n ("now") and f which

stands for "future". We also use f for the literary production in period f (see figure 3.2). Finally, we deal with the forking model in subsection 3.4.

3.2. Old-language learning with two periods

We begin with two periods and assume that the language learners take the production in their period into account and adjust their learning decisions accordingly. For the readers of period n, assume this payoff function:

$$u^n = \begin{cases} p_0 + n_1 + n_0 - c, & \text{reader learns language 0} \\ \alpha p_0 + n_1 + \alpha n_0, & \text{reader does not learn language 0} \end{cases}$$

where α is the accessibility of (the past) language 0. In period n, learning language 0 is profitable if

$$c < (1 - \alpha)(p_0 + n_0) =: c_0^n$$

We assume that the speakers' costs of language learning c are uniformly distributed on $[0, C_0^n]$ where $C_0^n > p_0 + n$. Then, a proportion of

$$\frac{c_0^n}{C_0^n} = \frac{(1 - \alpha)(p_0 + n_0)}{C_0^n} < 1$$

speakers alive in period 1 learns language 0.

We now turn to literary production done by one producer, only. We define readership by

$$\begin{aligned} R(n_0) &= q\left[(n - n_0) + n_0\left(\frac{c_0^n}{C_0^n} + \alpha\left[1 - \frac{c_0^n}{C_0^n}\right]\right)\right] \\ &= q\left[n - n_0\left[1 - \frac{c_0^n}{C_0^n}\right](1 - \alpha)\right] \end{aligned}$$

By $\alpha < 1$ and $\frac{c_0^n}{C_0^n} < 1$, readership maximization implies language production with exclusive use of the vernacular in period $n : n_0 = 0$ and $n_1 = n$. We find:

Proposition 3.1. *In the diachronic model with two time periods p (past) and n (present), the producer in period n does not use language 0 for literary production.*

Proof: obvious.

The upshot of this two-period model is this: some readers have an interest in learning language 0, but the producer does not. The reason behind this result

is simple enough. The producer in period n has no reason to produce for those having lived in the past (they cannot read). With respect to those living in the present, he maximizes readership when he uses their vernacular, language 1. After all, only a part of those living now choose to learn language 0. While this result is straightforward, it shows the difficulty of explaining the production in language 0, the vernacular of the past.

3.3. A three-period model of old-language learning

3.3.1. Setup

In order to show how an old language (like Latin) can become a standard language for production, we need to assume three periods (see figure 3.2). The assumptions are close to those of the two-period model. For simplification, we assume that readers learn no or one additional language, only. We need some assumptions on the maximal costs of language learning. It seems plausible to assume $C_0^n < C_0^f$, i.e., that the cost of learning language 0 is smaller for readers in period n than for those in period f. Similarly, one might assume $C_1^f < C_0^f$, i.e., that readers in period f find language 1 (the vernacular of period n) easier to learn than language 0 (the vernacular of period p). However, in order to simplify our formulae, we let $C := C_0^n = C_0^f = C_1^f$. Again, we assume that not all agents find language learning profitable. Formally, the costs of language learning C are sufficiently large (larger than $p + n + f$ is sufficient).

Literary production is described by

$$
\begin{aligned}
p &= p_0 \\
n &= n_0 + n_1 \text{ and} \\
f &= f_0 + f_1 + f_2
\end{aligned}
$$

and readers in periods n and f have the payoff functions

$$
u^n = \begin{cases} p_0 + n_1 + n_0 - c, & \text{reader learns language 0} \\ \alpha p_0 + n_1 + \alpha n_0, & \text{reader does not learn language 0} \end{cases}
$$

$$
u^f = \begin{cases} p_0 + n_0 + \alpha n_1 + f_2 + f_0 + \alpha f_1 - c, & \text{reader learns language 0} \\ \alpha p_0 + \alpha n_0 + n_1 + f_2 + \alpha f_0 + f_1 - c, & \text{reader learns language 1} \\ \alpha^2 p_0 + \alpha^2 n_0 + \alpha n_1 + f_2 + \alpha^2 f_0 + \alpha f_1, & \text{reader does not learn any language} \end{cases}
$$

respectively. Consider, for example, the second line of u^f where we find

$$\alpha p_0 + \alpha n_0 + n_1 + f_2 + \alpha f_0 + f_1 - c = \alpha (p_0 + n_0 + f_0) + n_1 + f_2 + f_1 - c.$$

The reader in period f can perfectly understand his own vernacular (f_2) and also literary production that uses language 1 ($n_1 + f_1$). Language 1 is only one step away from language 0. Therefore, readers from period f who know language 1 have a better understanding (indicated by factor α) of language-0 production than those who have not learned language 1 (see factor α^2 in the third line).

3.3.2. Current readership

Consider the literary producers of the middle period n. We develop two submodels. In the first (this section), these producers want to maximize their current readership, only. In the following section, they try to maximize their long-term readership. See the introduction for a short discussion. We will see below that long-term readership maximization can, and current-readership maximization cannot, explain the use of language 0 by the producers of the middle period n.

Current readership is defined by

$$R = q_n \left[n_1 + n_0 \left(r_0^n + \alpha \left[1 - r_0^n \right] \right) \right] \text{ (people alive in period } n)$$

where r_0^n is the percentage of readers in period n that have learned language 0.

We obtain

Proposition 3.2. *In the diachronic model with three time periods p (past), n (present), and f (future), if producers aim to maximize current readership, only, the producers in periods n and f use only their respective vernaculars for literary production.*

Proof: see appendix.

An extension of this model to any number of periods yields the same result: Current readership maximization implies that producers solely use their vernaculars for literary production. The reason is always the same. Past readers are dead, producers do not care for future readers (by definition), and therefore, they maximize readership within period n by using the language that people in this period know best (the current vernacular).

3.3.3. Long-term readership

We now turn to maximization of long-term readership. The readership of the middle-period producers is now enlarged and is composed of period-n and period-

f readers:

$$R = q_n \left[n_1 + n_0 \left(r_0^n + \alpha \left[1 - r_0^n \right] \right) \right] \text{ (people alive in period } n)$$
$$+ q_f \left[n_1 \left(r_1^f + \alpha \left[r_0^f + r_{no}^f \right] \right) + n_0 \left(r_0^f + \alpha r_1^f + \alpha^2 r_{no}^f \right) \right] \text{ (people alive in period } f)$$

where r_0^n, r_0^f, and r_1^f denote the percentage of readers in periods n or f that have learned languages 0 or 1, respectively. r_{no}^f is the percentage of period-f readers that have not learned any second language.

Proposition 3.3. *In the diachronic model with three time periods p (past), n (present), and f (future), if producers aim to maximize long-term readership, the producer in period f uses his vernacular, only. Assume $q_n = q_f$. We find:*

- *The producer in period n employs language 0 if C is sufficiently small, and language 1, otherwise.*

- *The chances for language 0 being used are smaller with increasing literary production ($p_0 < n$) than with decreasing literary production ($p_0 > n$).*

Proof: see appendix.

The producer in period f is in a situation comparable to that of a producer in period n in the two-period model. He maximizes his readership (there is no difference between the current and the long-term one) by using the language known perfectly by everybody alive.

The producer in period n has the choice between the vernacular language 1 and language 0, the past language. He may consider using the past language if the cost of language learning is relatively small. Then, a substantial portion of people alive in periods n and f will consider learning language 0. Of course, he can reach the people in period n better with language 1. Thus, it is for the sake of period-f readers that he considers using language 0. Here, we have our first explanation of how a past language can develop into a standard for literary production. Turning to the second bullet, the willingness to learn language 0 depends on the production in that language. It might be produced in period n (the proposition is about the conditions for this to happen) and it might have been produced in period p. If that production is large in comparison to the production in period n ($p_0 > n$), language 0 stands a higher chance of becoming the standard language for production.
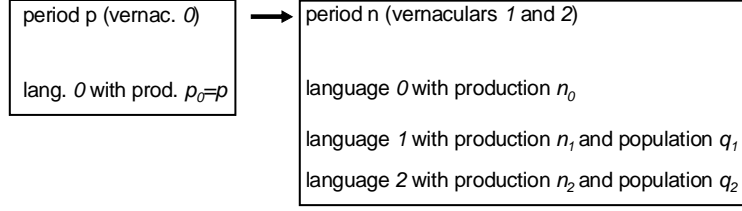
Figure 3.3: The forking model with two periods

3.4. A forking two-period model (Old French and Old Spanish)

We now turn to the forking model (see (Cf) in the introduction). There exists one language 0 in period p that develops into two different languages 1 and 2 that are vernaculars in period n (see figure 3.3). We focus on language learning and production in period n. Overall production in period n is fixed and given by

$$n = n_0 + n_1 + n_2.$$

Speakers in period n (who have language 1 or language 2 as their mother tongue) can learn language 0 or the other vernacular. We denote the accessibility of language 1 from language 2 by α (and also the other way around) and the accessibility of language 0 from any of the daughter languages by α_0. The reader in period n whose mother tongue is 1 has the payoff function

$$u_1 = \begin{cases} p_0 + n_0 + n_1 + \alpha n_2 - c, & \text{reader learns language 0} \\ \alpha_0 \left(p_0 + n_0\right) + n_1 + n_2 - c, & \text{reader learns language 2} \\ \alpha_0 \left(p_0 + n_0\right) + n_1 + \alpha n_2, & \text{reader does not learn any language} \end{cases}$$

We define readership by

$$\begin{aligned} R & = q_1 \left[n_0 \left(r_0^1 + \alpha_0 \left(1 - r_0^1\right)\right) + n_1 + \alpha n_2 \right] \\ & + q_2 \left[n_0 \left(r_0^2 + \alpha_0 \left(1 - r_0^2\right)\right) + \alpha n_1 + n_2 \right] \end{aligned}$$

where r_0^1 (r_0^2) is the percentage of readers with mother tongue 1 (with mother tongue 2) that have learned language 0.

We now present two propositions. Both assume (without loss of generality) $q_2 > q_1$. Then, if one of the two period-n languages is chosen, language 2 is the dominant one. The propositions differ with respect to the accessibiliy assumption. The first is based on $\alpha < \alpha_0$. Thus, the reader with mother tongue 1 (Old French)

who does not learn any language finds the literature of language 0 (Latin) more easily accessible than the literature of language 2 (Old Spanish). Here, we assume that language learning is very difficult. Taking this assumption to the extreme, language learning is excluded. The second proposition is based on the inverse inequality $\alpha > \alpha_0$. Both propositions identify conditions under which Latin (or another old language) may become the standard for literary production.

Proposition 3.4. *In the diachronic forking model with two time periods p (past) and n (present) and no language learning, assume $\alpha < \alpha_0$ (Old French closer to Latin than to Old Spanish) and $q_2 > q_1$.*

- *The producer in period n employs language 0 if*

    - *α_0 is relatively large,*
    - *α is small in comparison with α_0 (Old French much closer to Latin than to Old Spanish), and*
    - *the population sizes do not differ a lot.*

- *Otherwise, he employs language 2.*

Proof: see appendix.

According to the proposition, Latin may dominate Old French and Old Spanish if the cost of language learning is very large so that (practically) no language learning occurs. Also, α_0 has to be large, much larger than α. Thus, a producer is enticed to use Latin for his literary production because this literature can also be enjoyed by speakers of Old French and Old Spanish. Finally, the population sizes must not differ too much. Otherwise, the larger population of one language would lead to the domination of this language.

Proposition 3.5. *In the diachronic forking model with two time periods p (past) and n (present), assume $\alpha > \alpha_0$ (Old French closer to Old Spanish than to Latin) and $q_2 > q_1$.*

- *The producer in period n employs language 0 if*

    - *C is sufficiently small*
    - *α is relatively large and α_0 is relatively small (Old French much closer to Old Spanish than to Latin), and*

21

– *literary production is decreasing ($p_0 > n$) or only mildly increasing ($\frac{1-\alpha}{1-\alpha_0} n < p_0 < n$).*

- *Otherwise, he employs language 2.*

Proof: see appendix.

Here, Latin may become the standard language if the cost of language learning is low or, for medium cost of language learning, if literary production in period p is larger than (or only minimally smaller than) literary production in period n. The reason behind this result is the following: The speakers of Old French can enjoy Old Spanish due to the similarity between these two languages (α is relatively large). Therefore, if p_0 is large in comparison to n, they find learning Latin attractive. This attractiveness of learning Latin carries over to the literary producer in period n who also finds Latin an attractive medium.

4. Conclusion

In order to understand the interconnections between language learning and literary production, we employ standard methods from microeconomics, in particular from Game Theory and Industrial Organization. Among our findings, we want to highlight the following ones:

1. Translations (modeled by accessibility) have contradicting effects on helping a language gain readers and/or literary producers (see subsection 2.2).

   - For given literary productions, language learning is made unattractive if good translations from foreign languages exist.[4]

   - For given learning decisions, the readership of a language increases if good translations into foreign languages are available. Then, that language becomes more attractive from the producers' point of view.

2. There can be a theoretical ambiguity as to which of several languages is chosen by producers and hence preferred by language learners (see subsection 2.4). This ambiguity is typical for network effects and the resulting path dependence is well-known from the literature on competing technologies (see Arthur (1989) and below).

---

[4]Colomer (1996) deals with the related question of how to compare the social benefits and costs of language learning versus translations.

3. Option demand models tend to reduce the number of surviving languages drastically, while non-option demand (see subsection 2.5) explains the persistence of small languages (possibly above a certain threshold, see Grin (1992) and sections 2 and 3 in Fishman (1989)).

4. We offer three explanations of how a past language like Latin or Sanskrit can develop into a standard for literary production.

   - First, a producer may use the old language in order to reach out to future readers (see subsection 3.3.3).

   - Second, in a forking situation, a language like Latin may dominate its daughter languages Old French and Old Spanish for a while (see subsection 3.4). This result might be obtained for large learning cost and pronounced dissimilarities between daughter languages (see proposition 3.4).

   - Third, Latin may also dominate in a forking situation where we have small learning cost together with similar daughter languages (see proposition 3.5).

In a survey article, Grin (1996, pp. 28-30) rightly warns against the facile use of analogy in the overlapping fields of linguistics and economics. The current author likes to side with Church & King (1993) who point out the similarities between language and technology adoption. In fact, we encounter so-called network effects in both fields. For product markets, network effects are said to be present when demand depends positively on past and expected sales. Examples of network-effect goods are communication systems, television standards, and video systems. Network effects provide large-scale producers with advantages over small-scale producers. A comprehensive survey is Farrell & Klemperer (2007). Similarly, languages spoken by many speakers (see the effects of q_1 and q_2 in our models) are attractive for language learners and literary producers, still increasing the number of speakers. English is the obvious case in point.

Technologies and languages can be more or less "compatible" with each other. Thus, in order to enable communication, telephones have to adhere to specific technical specifications which they may fulfill, completely or partly. Languages can also be more or less similar to each other (see our remarks on contrastive linguistics in the introduction). In our model, we have a parameter for compatibility that we call accessibility.

Another similarity between technologies and languages is "sponsorship". A technology may be "sponsored", i.e., a firm possessing (or using) a technology may be interested in ensuring that the technology becomes standard (see Katz & Shapiro (1986)). Similarly, languages may be sponsored by national or supranational governments. In contrast to technologies, governments may sponsor several and even minority languages (see the Swiss or the Irish examples in Coulmas (1991)). We are then entering the extensive literature on language planning and language policy. See Wardhaugh (2006, chapter 15) for a survey, Ricento (2009) for a collection of articles on language policy, or Pool (1996) for an optimal set of official languages in the EU. Our paper also contributes to that literature, if only by exploring the effect of translations on the competitive success of languages (see 1. above).

Our model might be extended in different ways. For example, one might include the communicative benefit of knowing a language many other people know (and thus link up to previous models based on that network effect). Also, the fact of knowing two or more languages does not settle the question of which language is used (most of the time or for specific purposes). A decision-theoretic analysis in that direction is presented by Grin (1990).

Future research might also turn to specific language histories. There, one might engage in finding out whether these histories bear out our results or, if they do not, in identifying the specific cultural and political conditions that are responsible. In the Indian case, for example, a thorough treatment has been given by Pollock (2006). For Latin and its successor languages, the reader might profitably turn to Auerbach (1965), Hammond (1976), Irvine (1994), or Herman (2000).

5. Appendix

A. Proof of proposition 2.1

The second derivative of the readership R with respect to n_1 is positive so that we have extremal solutions. By $n + p_2 \leq C_2$, we obtain

$$
\begin{aligned}
R\left(n_1 = n\right) \quad &> \quad R\left(n_1 = 0\right) \Leftrightarrow \\
\frac{q_1}{q_2} \quad &> \quad \frac{1 - \left(\frac{(p_1+n)(1-\alpha_1)^2}{C_1} + \alpha_1\right)}{1 - \left(\frac{(p_2+n)(1-\alpha_2)^2}{C_2} + \alpha_2\right)} =: Q^1
\end{aligned}
$$

Thus, if $\frac{q_1}{q_2} > Q^1$ holds, the representative producer employs language 1, only. If $\frac{q_1}{q_2} < Q^1$ holds, the representative producer employs language 2, only.

B. Proof of proposition 2.2

Assume m literary producers, among them a producer called A. Focus on producer A who is faced with the literary productions of the other authors n_1^{others} (language 1), n_2^{others} (language 2), and $n^{\text{others}} = n_1^{\text{others}} + n_2^{\text{others}} < n$. Indicating the production of producer A with A, we define

$$
\begin{aligned}
n^A &: = n - n^{\text{others}} \\
n_1^A &: = n_1 - n_1^{\text{others}} \\
n_2^A &: = n_2 - n_2^{\text{others}}
\end{aligned}
$$

and obtain $n_2^A = n^A - n_1^A$. The second partial derivative of $R^A\left(n_1^A, n_1^{\text{others}}, n_2^{\text{others}}\right)$ (as given in the main text) with respect to n_1^A is positive. We find

$$
R^A\left(n_1^A = n^A, n_1^{\text{others}}, n_2^{\text{others}}\right) > R^A\left(n_1^A = 0, n_1^{\text{others}}, n_2^{\text{others}}\right) \Leftrightarrow
$$

$$
\frac{q_1}{q_2} > \frac{1 - \left(\frac{\left(p_1 + n_1^{\text{others}} + n^A\right)(1-\alpha_1)^2}{C_1} + \alpha_1\right)}{1 - \left(\frac{\left(p_2 + n_2^{\text{others}} + n^A\right)(1-\alpha_2)^2}{C_2} + \alpha_2\right)} =: Q^{\text{crit}}
$$

where the last inequality uses the assumption made in the main text that C_2 is sufficiently large. Assume that the others (predominantly) choose language 1 so that n_1^{others} is large and n_2^{others} small. Then, Q^{crit} is small and producer A also uses language 1. Therefore, we are justified in looking for symmetric equilibria, only. A separating equilibrium might exist, but it would be unstable. Thus, all producers choose language 1 if

$$
\frac{q_1}{q_2} > \frac{1 - \left(\frac{\left(p_1 + n_1^{\text{others}} + n^A\right)(1-\alpha_1)^2}{C_1} + \alpha_1\right)}{1 - \left(\frac{\left(p_2 + n_2^{\text{others}} + n^A\right)(1-\alpha_2)^2}{C_2} + \alpha_2\right)} = \frac{1 - \left(\frac{(p_1 + n)(1-\alpha_1)^2}{C_1} + \alpha_1\right)}{1 - \left(\frac{\left(p_2 + \frac{n}{m}\right)(1-\alpha_2)^2}{C_2} + \alpha_2\right)} =: Q_2^m
$$

$$
\text{with } Q_2^\infty := \lim_{m \to \infty} Q_2^m = \frac{1 - \left(\frac{(p_1 + n)(1-\alpha_1)^2}{C_1} + \alpha_1\right)}{1 - \left(\frac{p_2(1-\alpha_2)^2}{C_2} + \alpha_2\right)}
$$

In contrast, all producers choose language 2 if

$$\frac{q_1}{q_2} < \frac{1 - \left(\frac{\left(p_1 + n_1^{\text{others}} + n^A \right)(1-\alpha_1)^2}{C_1} + \alpha_1 \right)}{1 - \left(\frac{\left(p_2 + n_2^{\text{others}} + n^A \right)(1-\alpha_2)^2}{C_2} + \alpha_2 \right)} = \frac{1 - \left(\frac{\left(p_1 + \frac{n}{m} \right)(1-\alpha_1)^2}{C_1} + \alpha_1 \right)}{1 - \left(\frac{(p_2+n)(1-\alpha_2)^2}{C_2} + \alpha_2 \right)} =: Q_1^m$$

$$\text{with } Q_1^{\infty} := \lim_{m \to \infty} Q_1^m = \frac{1 - \left(\frac{p_1(1-\alpha_1)^2}{C_1} + \alpha_1 \right)}{1 - \left(\frac{(p_2+n)(1-\alpha_2)^2}{C_2} + \alpha_2 \right)}$$

Now, from

$$\begin{aligned} Q^1 &= Q_2^m = Q_1^m \text{ for } m = 1 \text{ and} \\ Q_2^{\infty} &< Q_2^{m+1} < Q_2^m < Q^1 < Q_1^m < Q_1^{m+1} < Q_1^{\infty} \text{ for } m \geq 2 \end{aligned}$$

we obtain the results (a) though (c) of the proposition. In particular, in case of

$$Q_2^m < \frac{q_1}{q_2} < Q^1 < Q_1^m$$

language 2 should be adopted according to proposition 2.1, but language 1 is adopted in one of the two equilibria according to case (b) of the current proposition.

C. Proof of proposition 2.3

Forming the derivative of R^A with respect to n_1^A yields

$$\frac{\partial R^A}{\partial n_1^A} = q_1 \frac{n_1^B}{\left(n_1^A + n_1^B \right)^2} - q_2 \frac{n_2^B}{\left(n^A - n_1^A + n_2^B \right)^2}$$

with negative second derivative. Thus,

$$q_1 \frac{n_1^B}{\left(n_1^A + n_1^B \right)^2} = q_2 \frac{n^B - n_1^B}{\left(n^A - n_1^A + n^B - n_1^B \right)^2} \tag{C.1}$$

is the first-order condition for readership maximization on the part of A. Similarly,

$$q_1 \frac{n_1^A}{\left(n_1^A + n_1^B \right)^2} = q_2 \frac{n_2^A}{\left(n^B - n_1^B + n_2^A \right)^2} \tag{C.2}$$

is the corresponding condition for B. The Nash equilibrium is the tuple of strategies $\left(n_1^A, n_1^B \right)$ that fulfills both equalities.

D. Proof of proposition 3.2

We now solve the three-period model by backward induction. Since no decisions are made in period p, we have two stages. We begin with language learning in period f.

- Learning language 1 is better than learning no language if

$$c \leq (1 - \alpha)(f_1 + n_1) + (\alpha - \alpha^2)(f_0 + n_0 + p_0) =: c_1^f$$

- Learning language 0 is better than learning no language if

$$c \leq (1 - \alpha^2)(f_0 + n_0 + p_0) =: c_0^f$$

- Learning language 0 is better than learning language 1 if

$$f_0 + n_0 + p_0 > f_1 + n_1$$

According to this last inequality, we distinguish two cases.

- Large language-0 base: $f_0 + n_0 + p_0 > f_1 + n_1$

  - In period f, language 0 is learned by the readership proportion $r_0^{f,\text{la } 0} := \frac{c_0^f}{C} = \frac{(1-\alpha^2)(f_0+n_0+p_0)}{C}$
  - In period f, no language is learned by the readership proportion $r_{\text{no}}^{f,\text{la } 0} := 1 - \frac{c_0^f}{C}$

- Small language-0 base: $f_0 + n_0 + p_0 < f_1 + n_1$

  - In period f, language 1 is learned by the readership proportion $r_1^{f,\text{sm } 0} := \frac{c_1^f}{C} = \frac{(1-\alpha)(f_1+n_1)+(\alpha-\alpha^2)(f_0+n_0+p_0)}{C}$
  - In period f, no language is learned by the readership proportion $r_{\text{no}}^{f,\text{sm } 0} := 1 - \frac{c_1^f}{C}$

Disregarding q_n (which does not make any substantial difference), the readership of period-f literary products is given by

$$\begin{cases} f_2 + f_0 \left(r_0^{f,\text{la }0} + \alpha^2 r_{\text{no}}^{f,\text{la }0} \right) + f_1 \alpha, & \text{large language-0 base} \\ f_2 + f_0 \left(\alpha r_1^{f,\text{sm }0} + \alpha^2 r_{\text{no}}^{f,\text{sm }0} \right) + f_1 \left(r_1^{f,\text{sm }0} + \alpha r_{\text{no}}^{f,\text{sm }0} \right), & \text{small language-0 base} \end{cases}$$

In order to maximize period-2 readership, the producers choose the vernacular, only:

$$\begin{aligned} f_2 &= f \\ f_0 &= 0 \\ f_1 &= 0 \end{aligned}$$

Then, we obtain:

- Large language-0 base: $n_0 + p_0 > n_1$ or $n_0 > \frac{n - p_0}{2}$

    - In period f, language 0 is learned by the readership proportion $r_0^{f,\text{la }0} :=$ $\frac{c_0^f}{C} = \frac{\left(1 - \alpha^2\right)(p_0 + n_0)}{C}$
    - In period f, no language is learend by the readership proportion $r_{\text{no}}^{f,\text{la }0} :=$ $1 - \frac{c_0^f}{C}$

- Small language-0 base: $n_0 + p_0 < n_1$ or $n_0 < \frac{n - p_0}{2}$

    - In period f, language 1 is learned by the readership proportion $r_1^{f,\text{sm }0} :=$ $\frac{c_1^f}{C} = \frac{(1-\alpha)n_1 + \left(\alpha - \alpha^2\right)(p_0 + n_0)}{C}$
    - In period f, no language is learned by the readership proportion $r_{\text{no}}^{f,\text{sm }0} :=$ $1 - \frac{c_1^f}{C}$

The readership of period-2 literary products is f.

We now turn to period n. Period-n learners are in the very same position as in the two-period model. Thus, a proportion of

$$r_0^n := \frac{(1 - \alpha)(p_0 + n_0)}{C}$$

speakers alive in period n learn language 0 while the proportion $r^n_{\text{not } 0} := 1 - r^n_0$ does not learn language 0.

Since the producers of period n maximize current readership, only, they are in the same position as in the two-period model, i.e., we obtain

$$n_0 = 0 \text{ and } n_1 = n.$$

E. Proof of proposition 3.3

The proposition builds on the results obtained for current readership (see the previous appendix). In particular, we need to distinguish between large and small language-0 bases as defined above. We can safely disregard $q_n = q_f$. For a large language-0 base ($n_0 + p_0 > n_1$ or $n_0 > \frac{n - p_0}{2}$), we obtain the period-n producer's readership

$$
\begin{aligned}
& R^{n,\text{la } 0}\left(n_0\right) \\
= \ & n_1 + n_0\left(r^n_0 + \alpha r^n_{\text{not } 0}\right) \text{ (people alive in period } n\text{)} \\
& + \alpha n_1 + n_0\left(r^{f,\text{la } 0}_0 + \alpha^2 r^{f,\text{la } 0}_{\text{no}}\right) \text{ (people alive in period } f\text{)} \\
= \ & n\left(1 + \alpha\right) + n_0\left(-1 + \alpha^2 + \frac{\left(1 - \alpha\right)^2\left(p_0 + n_0\right)}{C} + \frac{\left(1 - \alpha^2\right)^2\left(p_0 + n_0\right)}{C}\right)
\end{aligned}
$$

The second derivative with respect to n_0 is positive.

Similarly, for a small language-0 base ($n_0 + p_0 < n_1$ or $n_0 < \frac{n - p_0}{2}$), we obtain the period-n producer's readership

$$
\begin{aligned}
& R^{n,\text{sm } 0}\left(n_0\right) \\
= \ & n_1 + n_0\left(r^n_0 + \alpha r^n_{\text{not } 0}\right) \text{ (people alive in period } n\text{)} \\
& + n_1\left(r^{f,\text{sm } 0}_1 + \alpha r^{f,\text{sm } 0}_{\text{no}}\right) + n_0\left(\alpha r^{f,\text{sm } 0}_1 + \alpha^2 r^{f,\text{sm } 0}_{\text{no}}\right) \text{ (people alive in period } f\text{)} \\
= \ & \left(n - n_0\right)\left(1 + \alpha\right) + n_0\left(\alpha + \alpha^2\right) + n_0\left(1 - \alpha\right)\frac{\left(1 - \alpha\right)\left(p_0 + n_0\right)}{C} \\
& + \left[\left(n - n_0\right)\left(1 - \alpha\right) + n_0\left(\alpha - \alpha^2\right)\right]\frac{\left(1 - \alpha\right)\left(n - n_0\right) + \left(\alpha - \alpha^2\right)\left(p_0 + n_0\right)}{C}
\end{aligned}
$$

again with positive second derivative.

Besides the 0-base (which can be large or small), we need the distinction between

- increasing literary production which is given by $p_0 < n$ and

- decreasing literary production which is given by $p_0 > n$.

We then obtain the following matrix:

	$p_0 < n$	$p_0 > n$
$n_0 > \frac{n-p_0}{2}$	Case "incr. prod. + large 0-base"	Case "decr. prod. + large 0-base"
$n_0 < \frac{n-p_0}{2}$	Case "incr. prod. + small 0-base"	excluded (see below)

Assume, first, $p_0 < n$. We find:

- for the case "incr. prod. + large 0-base"

$$R^{n,\text{la } 0}(n) \;>\; R^{n,\text{la } 0}\left(\frac{n-p_0}{2}\right) \Leftrightarrow$$

$$C \;<\; \frac{1}{2}\frac{1-\alpha}{1+\alpha}\left(2\alpha + \alpha^2 + 2\right)\left(3n + p_0\right) := C^{\text{la } 0, p_0 < n, n \succ \frac{n-p_0}{2}}$$

- for the case "incr. prod. + small 0-base"

$$R^{n,\text{sm } 0}\left(\frac{n-p_0}{2}\right) \;>\; R^{n,\text{sm } 0}(0) \Leftrightarrow$$

$$C \;<\; \frac{1}{2}\frac{1-\alpha}{1+\alpha}\left(n\left(2\alpha + \alpha^2 - 2\right) + \alpha^2 p_0\right) := C^{\text{sm } 0, p_0 < n, \frac{n-p_0}{2} \succ 0}$$

$\frac{n-p_0}{2}$ can never be optimal because of these two implications:

$$R^{n,\text{la } 0}\left(\frac{n-p_0}{2}\right) \;>\; R^{n,\text{la } 0}(n) \Rightarrow R^{n,\text{sm } 0}(0) > R^{n,\text{la } 0}\left(\frac{n-p_0}{2}\right) \text{ and}$$

$$R^{n,\text{sm } 0}\left(\frac{n-p_0}{2}\right) \;>\; R^{n,\text{sm } 0}(0) \Rightarrow R^{n,\text{la } 0}(n) > R^{n,\text{sm } 0}\left(\frac{n-p_0}{2}\right)$$

Thus, the important comparison concerns $n_0 = n$ versus $n_0 = 0$. We obtain

$$R^{n,\text{la } 0}(n) \;>\; R^{n,\text{sm } 0}(0) \Leftrightarrow$$

$$C \;<\; \frac{1-\alpha}{1+\alpha}\left(n\left(\alpha + 1\right)^2 + p_0\left(\alpha + \alpha^2 + 2\right)\right) =: C^{p_0 < n}$$

We now turn to decreasing literary production $(p_0 > n)$ which excludes $n_0 < \frac{n-p_0}{2}$. Thus, we have the case "decr. prod. + large 0-base". Here, we find

$$R^{n,\text{la }0}(n) > R^{n,\text{la }0}(0) \Leftrightarrow$$

$$C < \frac{(2\alpha + \alpha^2 + 2)(1-\alpha)}{(1+\alpha)}(n+p_0) =: C^{p_0>n}$$

The preceeding calculations imply the proposition:

- Language 0 is employed if C is sufficiently small. In particular:

  - For $p_0 > n$ (decreasing literary production), the producer employs language 0, only, in case of

    $$C < \frac{(2\alpha + \alpha^2 + 2)(1-\alpha)}{1+\alpha}(n+p_0) =: C^{p_0>n}$$

    and language 1, otherwise.

  - For $p_0 < n$ (increasing literary production), the producer employs language 0, only, in case of

    $$C < \frac{1-\alpha}{1+\alpha}\left(n(\alpha+1)^2 + p_0(\alpha + \alpha^2 + 2)\right) =: C^{p_0<n}$$

    and language 1, otherwise.

- $C^{p_0>n} > C^{p_0<n}$ is easily confirmed. Thus, the chances for language 0 are smaller with increasing literary production than with decreasing literary production.

F. Proof of proposition 3.4

Assume $\alpha < \alpha_0$, $q_1 < q_2$ (without loss of generality), and $C > p_0 + n$. We solve a one-stage model (no language learning). Readership is then defined by

$$
\begin{aligned}
R &= q_1[n_0\alpha_0 + n_1 + \alpha n_2] + q_2[n_0\alpha_0 + \alpha n_1 + n_2] \\
&= n_1[q_1(1-\alpha_0) - q_2(\alpha_0 - \alpha)] \\
&\quad + n_2[q_2(1-\alpha_0) - q_1(\alpha_0 - \alpha)] + n\alpha_0(q_1 + q_2).
\end{aligned}
$$

By the above assumptions, we have these three cases, only:

- $0 < q_1 (1 - \alpha_0) - q_2 (\alpha_0 - \alpha) < q_2 (1 - \alpha_0) - q_1 (\alpha_0 - \alpha)$ or, equivalently, $\frac{q_1}{q_2} < \frac{\alpha_0 - \alpha}{1 - \alpha_0}$

  Readership is maximized for $n_0 = 0, n_1 = 0, n_2 = n$.

- $q_1 (1 - \alpha_0) - q_2 (\alpha_0 - \alpha) < 0 < q_2 (1 - \alpha_0) - q_1 (\alpha_0 - \alpha)$ or, equivalently, $\frac{1 - \alpha_0}{\alpha_0 - \alpha} > \frac{q_1}{q_2} > \frac{\alpha_0 - \alpha}{1 - \alpha_0}$

  Readership is maximized for $n_0 = 0, n_1 = 0, n_2 = n$.

- $q_1 (1 - \alpha_0) - q_2 (\alpha_0 - \alpha) < q_2 (1 - \alpha_0) - q_1 (\alpha_0 - \alpha) < 0$ or, equivalently, $\frac{1 - \alpha_0}{\alpha_0 - \alpha} < \frac{q_1}{q_2}$

  Readership is maximized for $n_0 = n, n_1 = 0, n_2 = 0$.

Together with $q_1 < q_2$, the last inequality implies $\frac{1 - \alpha_0}{\alpha_0 - \alpha} q_2 < q_1 < q_2$ and $\alpha_0 > \frac{1 + \alpha}{2}$.

G. Proof of proposition 3.5

Assume $\alpha > \alpha_0$, $q_1 < q_2$ (without loss of generality), and $C > p_0 + n$. We solve the two-stage model by backward induction. We begin with language learning in period n. For a reader whose mother tongue is language 1, learning language 0 is better than learning language 2 if

$$(1 - \alpha_0)(n_0 + p_0) > (1 - \alpha) n_2 \text{ or}$$
$$n_0 + p_0 > \frac{1 - \alpha}{1 - \alpha_0} n_2$$

We denote this case by "la 1" and the opposite case by "sm 1". We now turn to the question of whether learning a language is better than learning no language. For readers with mother tongue 1, we have:

- In case of "la 1", learning language 0 is better than learning no language if

$$c \leq (1 - \alpha_0)(n_0 + p_0) =: c_0^1$$

  holds (readership proportion $r_0^{1, \text{la } 1} := \frac{c_0^1}{C} = \frac{(1 - \alpha_0)(n_0 + p_0)}{C}$),

- In case of "sm 1", learning language 2 is better than learning no language if

$$c \leq (1 - \alpha) n_2 =: c_2^1$$

  holds (readership proportion $r_2^{1, \text{sm } 1} := \frac{c_2^1}{C} = \frac{(1 - \alpha) n_2}{C}$).

32

Interchanging the indices for 1 and 2 yields corresponding results for readers with mother tongue 2. In particular, we obtain $r_0^{2,\text{la }2}$ and $r_1^{2,\text{sm }2}$. In the "la 1"-"la 2" case $(n_0 + p_0 > \frac{1-\alpha}{1-\alpha_0} n_2$ and $n_0 + p_0 > \frac{1-\alpha}{1-\alpha_0} n_1)$, we obtain the readership

$$
\begin{aligned}
& R^{\text{la 1, la 2}}\left(n_0, n_1, n_2\right) \\
= \ & q_1 \left[n_0 \left(r_0^{1,\text{la }1} + \alpha_0 \left(1 - r_0^{1,\text{la }1}\right)\right) + n_1 + \alpha n_2\right] \\
& + q_2 \left[n_0 \left(r_0^{2,\text{la }2} + \alpha_0 \left(1 - r_0^{2,\text{la }2}\right)\right) + \alpha n_1 + n_2\right] \\
= \ & q_1 \left[(n - n_1 - n_2)\left(\alpha_0 + \frac{(1-\alpha_0)^2\left((n - n_1 - n_2) + p_0\right)}{C}\right) + n_1 + \alpha n_2\right] \\
& + q_2 \left[(n - n_1 - n_2)\left(\alpha_0 + \frac{(1-\alpha_0)^2\left((n - n_1 - n_2) + p_0\right)}{C}\right) + \alpha n_1 + n_2\right].
\end{aligned}
$$

The "la 1"-"sm 2" readership (case $n_0 + p_0 > \frac{1-\alpha}{1-\alpha_0} n_2$ and $n_0 + p_0 < \frac{1-\alpha}{1-\alpha_0} n_1$) is

$$
\begin{aligned}
& R^{\text{la 1, sm 2}}\left(n_0, n_1, n_2\right) \\
= \ & q_1 \left[n_0 \left(r_0^{1,\text{la }1} + \alpha_0 \left(1 - r_0^{1,\text{la }1}\right)\right) + n_1 + \alpha n_2\right] \\
& + q_2 \left[\alpha_0 n_0 + n_1 \left(r_1^{2,\text{sm }2} + \alpha \left(1 - r_1^{2,\text{sm }2}\right)\right) + n_2\right] \\
= \ & q_1 \left[(n - n_1 - n_2)\left(\alpha_0 + \frac{(1-\alpha_0)^2\left((n - n_1 - n_2) + p_0\right)}{C}\right) + n_1 + \alpha n_2\right] \\
& + q_2 \left[\alpha_0 (n - n_1 - n_2) + n_1 \left(\alpha + \frac{(1-\alpha)^2 n_1}{C}\right) + n_2\right]
\end{aligned}
$$

and the "sm 1"-"la 2" readership is (reversing the roles)

$$
\begin{aligned}
& R^{\text{sm 1, la 2}}\left(n_0, n_1, n_2\right) \\
= \ & q_1 \left[\alpha_0 (n - n_1 - n_2) + n_1 + n_2 \left(\alpha + \frac{(1-\alpha)^2 n_2}{C}\right)\right] \\
& + q_2 \left[(n - n_1 - n_2)\left(\alpha_0 + \frac{(1-\alpha_0)^2\left((n - n_1 - n_2) + p_0\right)}{C}\right) + \alpha n_1 + n_2\right]
\end{aligned}
$$

The readership in the "sm 1"-"sm 2" case is

$$
\begin{aligned}
R^{\text{sm 1, sm 2}}(n_0, n_1, n_2) &= q_1 \left[\alpha_0 n_0 + n_1 + n_2 \left(r_2^{1,\text{sm 1}} + \alpha \left(1 - r_2^{1,\text{sm 1}} \right) \right) \right] \\
&\quad + q_2 \left[\alpha_0 n_0 + n_1 \left(r_1^{2,\text{sm 2}} + \alpha \left(1 - r_1^{2,\text{sm 2}} \right) \right) + n_2 \right] \\
&= q_1 \left[\alpha_0 (n - n_1 - n_2) + n_1 + n_2 \left(\alpha + \frac{(1-\alpha)^2 n_2}{C} \right) \right] \\
&\quad + q_2 \left[\alpha_0 (n - n_1 - n_2) + n_1 \left(\alpha + \frac{(1-\alpha)^2 n_1}{C} \right) + n_2 \right]
\end{aligned}
$$

In all these four cases, the Hessian with respect to n_1 and n_2 is positive definite. Thus, the readership functions are strictly convex and at least one of the three variables n_0, n_1, or n_2 is zero. Letting one of these variables be zero, the restricted readership functions are strictly convex, too. Therefore, it is sufficient to focus on the three extreme cases:

- $n_0 = n, n_1 = 0, n_2 = 0$

  Here, we have $n + p_0 > \frac{1-\alpha}{1-\alpha_0} \cdot 0$ (where the 0 may stand for n_2 or n_1), i.e., the "la" cases, and we find the readership

$$
\begin{aligned}
&R^{\text{la 1, la 2}}(n_0 = n) \\
&= q_1 n \left(\alpha_0 + \frac{(1-\alpha_0)^2 (n + p_0)}{C} \right) + q_2 n \left(\alpha_0 + \frac{(1-\alpha_0)^2 (n + p_0)}{C} \right)
\end{aligned}
$$

- $n_0 = 0, n_1 = n, n_2 = 0$

  For speakers of language 1, we have the "la" case ($n + p_0 > \frac{1-\alpha}{1-\alpha_0} \cdot 0$), while speakers of language 2 may obey $p_0 > \frac{1-\alpha}{1-\alpha_0} n$ or not. If $p_0 > \frac{1-\alpha}{1-\alpha_0} n$ holds, we have the readership

$$
R^{\text{la 1, la 2}}(n_1 = n) = n (q_1 + \alpha q_2)
$$

  If, however, $p_0 < \frac{1-\alpha}{1-\alpha_0} n$, we have the "la 1"-"sm 2" readership

$$
R^{\text{la 1, sm 2}}(n_1 = n) = q_1 n + q_2 n \left(\alpha + \frac{(1-\alpha)^2 n}{C} \right)
$$

34

- $n_0 = 0, n_1 = 0, n_2 = n$

  If $p_0 > \frac{1-\alpha}{1-\alpha_0} n$, we have the readership

  $$R^{\text{la 1, la 2}} (n_2 = n) = n (q_2 + \alpha q_1)$$

  If, however, $p_0 < \frac{1-\alpha}{1-\alpha_0} n$, we obtain the readership

  $$R^{\text{sm 1, la 2}} (n_2 = n) = q_1 n \left(\alpha + \frac{(1-\alpha)^2 n}{C} \right) + q_2 n$$

So far, we have not made use of $\alpha > \alpha_0$ and of $q_1 < q_2$. From now on, we call $p_0 > \frac{1-\alpha}{1-\alpha_0} n$ "p_0 la" and $p_0 < \frac{1-\alpha}{1-\alpha_0} n$ "p_0 sm". Starting with the latter, we find

$$R^{\text{la 1, la 2}} (n_0 = n) > R^{\text{sm 1, la 2}} (n_2 = n) \Leftrightarrow$$

$$C < \frac{(1-\alpha_0)^2 (q_1 + q_2) (n + p_0) - q_1 (1-\alpha)^2 n}{q_1 (\alpha - \alpha_0) + q_2 (1-\alpha_0)} =: C^{p_0 \text{ sm},0 \succ 2}$$

and, analogously,

$$R^{\text{la 1, la 2}} (n_0 = n) > R^{\text{la 1, sm 2}} (n_1 = n) \Leftrightarrow$$

$$C < \frac{(1-\alpha_0)^2 (q_1 + q_2) (n + p_0) - q_2 (1-\alpha)^2 n}{q_2 (\alpha - \alpha_0) + q_1 (1-\alpha_0)} =: C^{p_0 \text{ sm},0 \succ 1}$$

By $q_2 > q_1$ and $\alpha > \alpha_0$, we have

$$C^{p_0 \text{ sm},0 \succ 2} < C^{p_0 \text{ sm},0 \succ 1}$$

$q_2 > q_1$ implies

$$R^{\text{sm 1, la 2}} (n_2 = n) > R^{\text{la 1, sm 2}} (n_1 = n).$$

Thus, for "p_0 sm" ($p_0 < \frac{1-\alpha}{1-\alpha_0} n$), we obtain:

- for $C < C^{p_0 \text{ sm},0 \succ 2}$, all production takes place in language 0,

- for $C > C^{p_0 \text{ sm},0 \succ 2}$, all production takes place in language 2.

We now turn to case "p_0 la" and observe

$$R\left(n_0 = n\right) > R^{p_0 \text{ la}}\left(n_1 = n\right) = n\left(q_1 + \alpha q_2\right) \Leftrightarrow$$

$$C \quad < \quad \frac{n\left(1 - \alpha_0\right)^2\left(q_1 + q_2\right)\left(n + p_0\right)}{q_1 n\left(1 - \alpha_0\right) + q_2 n\left(\alpha - \alpha_0\right)} =: C^{p_0 \text{ la},0\succ1}$$

and analogously

$$R\left(n_0 = n\right) > R^{p_0 \text{ la}}\left(n_2 = n\right) \Leftrightarrow$$

$$C \quad < \quad \frac{n\left(1 - \alpha_0\right)^2\left(q_1 + q_2\right)\left(n + p_0\right)}{q_2 n\left(1 - \alpha_0\right) + q_1 n\left(\alpha - \alpha_0\right)} =: C^{p_0 \text{ la},0\succ2}$$

We have

$$C^{p_0 \text{ la},0\succ2} < C^{p_0 \text{ la},0\succ1}$$

and

$$R^{p_0 \text{ la}}\left(n_1 = n\right) < R^{p_0 \text{ la}}\left(n_2 = n\right)$$

Thus, for "p_0 la" ($p_0 > \frac{1-\alpha}{1-\alpha_0}n$),

- for $C < C^{p_0 \text{ la},0\succ2}$, all production takes place in language 0,

- for $C > C^{p_0 \text{ la},0\succ2}$, all production takes place in language 2.

By $\alpha > \alpha_0$, we find

$$C^{p_0 \text{ sm},0\succ2} < C^{p_0 \text{ la},0\succ2}$$

and hence, the producer in period n decides in the following manner:

- For $C < C^{p_0 \text{ } sm,0\succ2}$ (very small cost of language learning), he employs language 0.

- For $C > C^{p_0 \text{ la},0\succ2}$ (very large cost of language learning), he employs language 2.

- For $C^{p_0 \text{ sm},0\succ2} < C < C^{p_0 \text{ la},0\succ2}$ (medium cost of language learning),

  - he employs language 0 if $p_0 > \frac{1-\alpha}{1-\alpha_0}n$ (not strongly increasing literary production) holds, but

  - he employs language 2, otherwise.

36

References

Arthur, W. B. (1989). Competing technologies, increasing returns, and lock-in by historical events, *The Economic Journal* **99**: 116–131.

Auerbach, E. (1965). *Literary Language and Its Public in Late Latin Antiquity and in the Middle Ages*, Princeton University Press. Translated from the German original (Literatursprache und Publikum in der lateinischen Spätantike und im Mittelalter, Francke Verlag, 1958) by Ralph Manheim.

Benz, A., Ebert, C., Jäger, G. & Rooij, R. (2011). *Language, Games, and Evolution*, Springer.

Bourdieu, P. (1977). L'economie des echanges linguistiques, *Language francaise* **34**: 17–34.

Church, J. & King, I. (1993). Bilingualism and network externalities, *Canadian Journal of Economics* **26**: 337–345.

Clackson, J. (2007). *Indo-European Linguistics*, Cambridge University Press, Cambridge et al.

Clark, R. L. (2011). *Meaningful games: Exploring language with game theory*, MIT Press.

Colomer, J. M. (1996). To translate or to learn languages? an evaluation of social efficiency, *International Journal of the Sociology of Language* **121**: 181–197.

Coulmas, F. (1992). *Language and Economy*, Blackwell.

Coulmas, F. (ed.) (1991). *A Language Policy for Europe*, Mouton de Gruyter.

Farrell, J. & Klemperer, P. (2007). Coordination and lock-in: Competition with switching costs and network effects, *in* C. M. Armstrong & R. H. Porter (eds), *Handbook of Industrial Organization, Vol. 3*, North-Holland, Amsterdam et al., chapter 5.

Fishman, J. A. (1989). *Language and ethnicity in minority sociolinguistic perspective*, Multilingual Matters.

Fortson IV, B. W. (2004). *Indo-European Language and Culture*, Blackwell, Oxford.

François, A. (2014). Trees, waves and linkages: Models of language diversification, *in* C. Bowern & B. Evans (eds), *The Routledge Handbook of Historical Linguistics*, Routledge, pp. 161–189.

Gabrys-Barker, D. & Bielska, J. (eds) (2013). *The Affective Dimension in Second Language Acquisition*, Multilingual matters.

González, M. d. l. Á. G., Mackenzie, J. L. & Álvarez, E. M. G. (2008). *Current trends in contrastive linguistics: functional and cognitive perspectives*, Vol. 60, John Benjamins Publishing.

Gray, R. D. & Atkinson, Q. D. (2003). Language-tree divergence times support the Anatolian theory of Indo-European origin, *Nature* **426**: 435–439.

Grin, F. (1990). The economic approach to minority languages, *Journal of Multilingual & Multicultural Development* **11**(1-2): 153–173.

Grin, F. (1992). Towards a threshold theory of minority language survival, *Kyklos* **45**(1): 69–97.

Grin, F. (1996). The economics of language: survey, assessment, and prospects, *International Journal of the Sociology of Language* **121**: 17–44.

Grin, F. & Vaillancourt, F. (1997). The economics of multilingualism: overview and analytical framework, *Annual Review of Applied Linguistics* **17**: 43–65.

Hammond, M. (1976). *Latin: A Historical and Linguistic Handbook*, Harvard University Press.

Heller, M. (1995). Language choice, social institutions, and symbolic domination, *Language in Society* **24**: 373–405.

Herman, J. (2000). *Vulgar Latin*, Pennsylvania State University Press.

Irvine, M. (1994). *The Making of Textual Culture. 'Grammatica' and Literary Theory, 350-1100*, Cambridge University Press.

Katz, M. L. & Shapiro, C. (1986). Technology adoption in the presence of network externalities, *Journal of Political Economy* **94**: 822–841.

Lado, R. (1957). *Linguistics across cultures: Applied linguistics for language teachers.*, University of Michigan Press.

Lehmann, W. P. (1992). *Historical Linguistics*, 3 edn, Routledge, London/New York.

Parikh, P. (2001). *The Use of Language*, CSLI Publications.

Pollock, S. (2006). *The Language of the Gods in the World of Men*, University of California Press.

Pool, J. (1996). Optimal language regimes for the european union, *International Journal of the Sociology of Language* **121**: 159–179.

Ricento, T. (2009). *An introduction to language policy: Theory and method*, John Wiley & Sons.

Schmalensee, R. (1972). Option demand and consumer's surplus: Valuing price changes under uncertainty, *The American Economic Review* pp. 813–824.

Selten, R. & Pool, J. (1991). The distribution of foreign language skills as a game equilibrium, *in* R. Selten (ed.), *Game Equilibrium Models IV*, Springer.

Tirole, J. (1988). *The Theory of Industrial Organization*, MIT Press, Cambridge (MA)/London.

Wardhaugh, R. (2006). *An introduction to sociolinguistics*, Blackwell Publishing.