9 Language spread, shift and maintenance: how groups choose their language

> Languages, like organic species can be classified into groups and subgroups … Dominant languages and dialects spread and lead to the gradual extinction of other tongues.
>
> Charles Darwin, *The Descent of Man and Selection in Relation to Sex*

> The lure of English has not left us. And until it goes, our own languages will remain paupers.
>
> Mohandas Gandhi (1965)

Outline of the chapter

Some societies are characterized by relatively stable language arrangements; others are more volatile. This chapter addresses the questions of how and why language choices by individuals and groups bring about incremental change of sociolinguistic arrangements in language-contact settings. By way of conceptualizing the inequality of the world's languages, it provides a brief review of their distribution and offers a five-tiered scheme as a general orientation. It then goes on to consider language-demographic statistics, explaining the difficulties of obtaining reliable data. The concepts of language loyalty, ethnolinguistic vitality, territories and domains, and utility are introduced as the most promising theoretical tools for analysing unstable language arrangements. By way of illustration, reference is made to the spread of languages on the Internet and to the ascent of English to the status of global language.

Key terms: migration, language spread, language decline, language shift, language maintenance, language loyalty

Stable and unstable language arrangements

One of the hallmarks of diglossia is its stability over long periods of time, a feature it shares with grassroots bilingualism. In both cases the population is territorially stationary, exhibiting no abrupt changes in residence patterns or language choice. If we think, for example, of German-speaking Switzerland, the Basque country and Jammu, north-west India, the

language scenarios of these regions have lasted for many generations. No uncertainties or challenges undermined the alternating language-choice conventions for Swiss German/High German, Basque/Spanish and Pashto/Urdu. Stability is a rather vague notion. If, for present purposes, it is defined to mean that nobody alive remembers different language-choice conventions, it implies that basic patterns have been in place for at least four generations. In this sense the three cases, different though they are in terms of sociogenesis and functional allocation of codes, exemplify stable language arrangements.

However, the stability of multilingual arrangements is by no means the rule. Around the globe speech communities live in close proximity, influencing each other. From a language-centred point of view as suggested by Darwin's above-quoted analogy, languages influence each other, expand, contract or die. The impact of one language on the lexicon, phonology and syntax of another has long been considered as vitally important for the understanding of language change. Contemporary scholarship recognizes that language contact involves speakers and, therefore, has to draw on psychological and sociological as well as linguistic approaches. Historical linguistics asked questions such as 'What happened to language X under the influence of language Y?' and 'Are there any features of X that can be traced to Y?' Contact linguistics,[1] by contrast, is more interested in the communicative aspects and processes of intergroup dynamics, varying communication norms and shifting patterns of language choice. From a speaker-centred point of view, the question to ask is how individuals and groups of speakers react when they get in touch with other groups and their languages. It is very rare that contact is between equals and more or less symmetric. More commonly the relationship between the communities involved is characterized by differences in terms of size, power, wealth, prestige and vitality. Variable combinations of these factors often make for inherently unstable language arrangements where speakers adjust their language-choice patterns during their lifetime and/or from one generation to the next. A major cause of such changes is migration.

Mass migration

Mass migration is not an isolated phenomenon. It involves an area and a community of origin, a destination and, perhaps, transitory stations along the way. On the individual level migration is usually a once-in-a-lifetime event, but seen in a larger historical context it is an extremely long-term process. For example, the Indo-European expansion has been going on for some 6,000 years (Gunaratne 2003), bringing in its wake the dislocation of many languages, the formation of new ones through pidginization and creolization, and the extinction of an indeterminable number of others which

have disappeared without a trace. Nowadays, in the age of globalization, this process continues, interacting with others that engender new language-contact situations. At 216 million (in 2010), the global migrant population would equal the fifth most populous country of the world, after China, India, USA and Indonesia.[2] Continuing migration flows of such an order of magnitude have given the questions of what these populations do with their languages and how they communicate with speakers of other languages renewed importance.

Immigrant groups find themselves in a new social and linguistic environment to which they adjust in many ways and which adjusts to them. Sociolinguistic evidence reveals a wide range of patterns of mutual adjustment with host communities insisting on assimilation to various degrees and immigrant communities showing variable inclination to yield to assimilation pressure. The resulting differences in language-choice patterns constitute a major subject of macro-sociolinguistic studies.

Language spread

Dutch enjoys official status and is widely spoken in Suriname, English in New Zealand, French in Madagascar, Portuguese in East Timor, Spanish in Cuba, Chinese in Singapore, Arabic in the Comoros, Turkish in Cyprus and Hindi in Mauritius. These are just some examples of languages that have been carried from their homeland to other parts of the world where they coexist with other languages which had been there before or were brought in later to add another layer to the mix.

The above-mentioned languages have spread, meaning that the populations speaking these languages grew more than the primary speech community in their homeland. For example, the estimated 130 million speakers of French[3] outnumber the population of France by a rate of 2:1. With a demographic strength of some 420 million, Spanish has almost ten times as many speakers as Spain has inhabitants. The ratio of Arabic speakers and Saudi Arabian nationals is more than 10:1, and that of Portuguese speakers and Portuguese nationals exceeds 20:1. Notice that, with the exception of Arabic, all of these languages belong to the Indo-European family of languages, which in a long-term historical perspective has spread more extensively than any other language family, picking up speakers around the globe. And we have not even mentioned the language whose growth over the past 300 years and especially during the twentieth century eclipsed that of any other language, English, which has an estimated 1,000 million speakers in some sixty countries. While the languages of Europe account for just 3 per cent of the world's 6,800 languages, about half of the world population of 7 billion speak European languages.[4] Clearly, a few select European languages have expanded more vigorously than all other languages on this globe.

Table 9.1 *Chinese demographic development*

China's population (estimates)
1751 ± 250 million
1851 ± 400 million
2012 ± 1340 million

Chinese, too, has a huge speech community exceeding 1,300 million speakers if all varieties are included. In a period of two and a half centuries, it has grown by a factor of 5.3 (see Table 9.1). However, the increase of the speech community parallels that of the Chinese population.

To be sure, as a result of Chinese emigration the Chinese language nowadays enjoys some currency in Chinatowns around the world from Amsterdam to New York City and Sydney, but it has not spread outside these settlements. It was never backed up by military might, colonialism and imperialism as was English, promoted by the British Empire first and then the US army, commerce, proselytism and popular culture. The Chinese written language was adopted by literary elites in neighbouring countries,[5] but vernacular Chinese never displaced the indigenous languages. Similarly, Hindi which follows Chinese as the second most populous language in terms of native speakers, has a very compact speech community with few overseas extensions.

Arabic, on the other hand, was disseminated both as a holy written language and a spoken vernacular from the Arabian Peninsula to other lands in the wake of the Islamic conquest. Today it is widely spoken in some twenty-three countries, ranking as one of the world's top ten languages in terms of demographic strength, although no reliable statistics about number of speakers are available (see below).

As these examples show, language spread is not a natural development. Rather, it is a complex process driven by many interacting variables such as the expansion of groups of people, migration, trade, slavery, conquest and the subjugation of others, domination, the establishment of institutions that promote and protect certain languages, such as churches, schools and armed forces, and the undermining of cultures, but also demand for some languages which primary speakers of other languages perceive as more useful or prestigious. Language spread is not just a result of these factors but has an effect on them, too. A language with a wide reach differs from one confined to a small speech community not just in size. The functional range of 'big' languages is wider and more differentiated than that of small ones. For example, the majority of the world's languages are not used for the functions most relevant for modern life. Only a small number of languages have spread to the Internet. In 2007, 17.5 per cent of the world population were connected to the Internet (Pimienta, Prado and Blanco 2009). At the time,

about 2 per cent of all languages accounted for close to 100 per cent of online communication, with huge disproportions between them. The only non-European languages of the top-ranking twelve are Chinese, Japanese and Korean. Then follow Arabic, Turkish and Thai. However, the linguistic diversity of the Internet keeps evolving. Various different measures are used to assess the presence of languages in computer-mediated communication and the penetration of the Internet in countries and speech communities. According to a calculation by O'Neill, Lavoie and Bennett (2003), English accounted for 72 per cent of random samples of webpages and Chinese for 2 per cent. In the meantime, Chinese has caught up by leaps and bounds. Internet World Stats (2010) counts 536.6 million English language internet users compared with 444.9 million Chinese users. Spanish has climbed to third rank, Japanese dropped to fourth. According to the same source, English is the language of 26.8 per cent of all internet users, compared to 24.2 per cent who use Chinese. In the first decade of this century, the English language internet population tripled, whereas the Chinese language internet population grew by 1500 per cent, starting from a low baseline. In the eight years since the publication of the first edition of this textbook, the online population has tripled. Figure 9.1 then cited 649 million as of March 2003, compared with 1,966 million in June 2010. This is just one indication of how much the situation is still in flux.

The 27 per cent of the online population who use English are a smaller proportion of the total than English accounted for in the early days of the Internet; however, even this reduced share is indicative of the role of English as the international lingua franca, which at the same time is the language of the hegemonic power of our days. As the only language with truly global reach, it constitutes the topmost layer of the multitiered system of the world's languages today. Supranational languages come next, notably French, Spanish, Portuguese, German, Russian, Arabic and Chinese, followed in turn by a number of languages whose speech communities are still large, but more geographically concentrated and unchallenged in their proper territories, such as Hindi–Urdu, Bengali, Tamil, Malay–Indonesian and Kiswahili. Highly developed but demographically small languages such as Dutch, Danish, Czech and Greek, and all other languages regularly used in writing, form the next two tiers. And finally there are a large number of unwritten languages spoken by small groups of speakers. These five tiers make up the world system of languages which is characterized by an inverse relation between number of languages and number of speakers, as schematically depicted in Figure 9.2. Ten major languages, each spoken by more than 100 million people, account for almost half of the world's population, whereas 52 per cent of all languages are spoken by less than 10,000 people, hundreds of them having fewer than 100 speakers.[6] Approximately 96 per cent of the world's population is being educated in one of the top twenty languages in terms of size. These

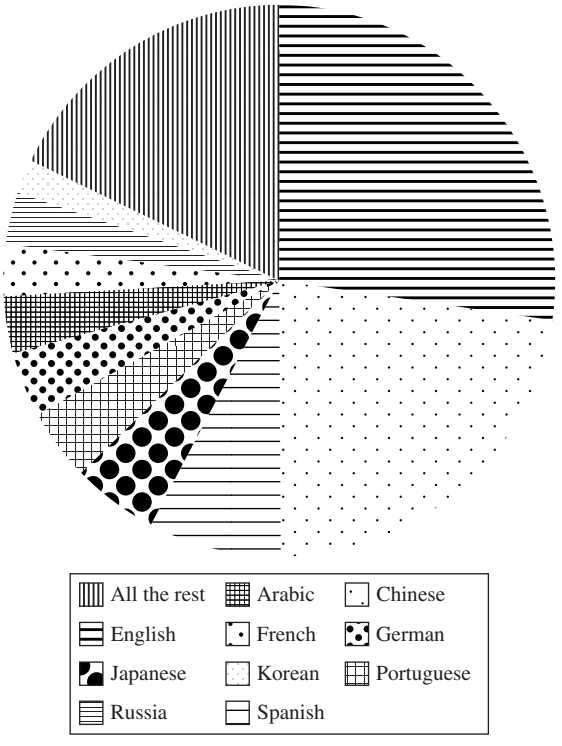

Figure 9.1 *The ten most frequently used languages on the internet. Online language populations, total: 1,966 million (June 2010). Source: Internet World Stats: www.internetworldstats.com/stats7.htm*

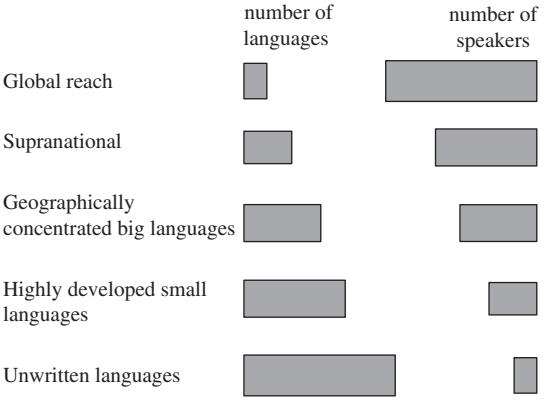

Figure 9.2 *Few languages with many speakers, many languages with few speakers: the linguistic diversity of the world*

are Chinese, Hindi–Urdu, English, Spanish, Arabic, Bengali, Portuguese, Russian, Malay–Indonesian, Japanese, German, French, Bihari,[7] Korean, Vietnamese, Italian, Telugu, Tamil, Turkish and Javanese.

Being equipped with the necessary means to function as the medium of instruction and being formally recognized as such gives these twenty languages an enormous competitive advantage in the endless push and pull and vying for domains and territories between the world's languages.[8] How did this system come into existence, and how does it evolve? As the comparison between Chinese and English shows, demographic strength is just one of several variables involved. It is, however, very important and must be kept in mind because present-day disparities imply that the potential to spread is very uneven for the languages of the world, as are the chances of survival. So great is the imbalance of power between 'big' languages and the 'little' ones, that many scholars predict a dramatic reduction of global linguistic diversity in the near future. The *UNESCO Atlas of the World's Languages in Danger*, available in its 2010 edition in English, Spanish and French, keeps track of some 2,500 languages that are thought to be at risk of extinction.[9]

Language replacement is a process that has been going on for thousands of years, but in the age of globalization it has accelerated dramatically, bringing about a global situation which is quite unstable. The world system of languages is the result of long-term historical processes. How this system changes as we live with it must be investigated on a smaller scale where issues of language shift and maintenance are addressed.

Data

A note of caution is in order here. The above numbers seem neat and precise; in truth they are quite problematic, constituting approximations at best. In addition to the usual hitches with demographic data, a number of problems mentioned earlier, such as language names, the vague and often arbitrary difference between language and dialect, and inconsistencies in self-declared language affiliations[10] make exact assessments difficult. Are Flemish and Dutch one language or two? Maltese used to be considered an Arabic dialect, but in 1990 the government of Malta decided to recognize it as a proper language (Hull 1994: 344). What does diglossia imply for counting languages? Who counts as a speaker of a language: natives only; bilinguals; speakers with reduced proficiency for lack of use? These and similar problems must be dealt with, if only arbitrarily, if data are to be classified and processed.

Language spread can be assessed only on the basis of representative data. The very first step in the study of this process makes it necessary to ascertain who speaks the language under investigation in a given community; what functional roles it fulfils in that community; and whether there is inter-generational change. It should be pointed out that these and other related

questions concern not just L1 speakers of that language but L2 speakers, too, for L2 speakers are crucially important for language expansion. Gathering data of this sort is a highly complicated matter.

For illustration, consider what was said about Arabic above. No reliable statistics about number of speakers are available. There are several reasons for this. First, we need a clear-cut definition of 'Arabic'. Suppose we settle on Modern Standard Arabic, which is quite clearly defined, then the question arises as to how to determine the number of speakers of this variety. Only a research instrument that is uniformly designed and applied in all Arabic-speaking countries and in other countries with sizeable Arabic-speaking minorities could yield reasonably accurate results. Now, recall for a moment the relationships between the languages and dialects spoken in Morocco as depicted in Figure 8.1 (p. 151). It is clear from this figure that it contains just approximate percentages, one of the reasons being that Moroccan Arabic, Middle Moroccan Arabic and Literary Arabic are not distinct entities but should be seen, rather, as overlapping sets. Similar sets constitute the language situations in other Arabic-speaking countries which, however, differ from Morocco in various ways. A large part of the Moroccan population are L1 Berber speakers. Many of them are bilingual to various degrees with proficiency in Arabic ranging from totally fluent to severely restricted. How are they to be counted? Whatever the answer, it would be difficult to ensure comparability with data gathered in other Arabic-speaking countries.

Alternatively, both a definition of Arabic and a standard of proficiency could be dispensed with, every variety and every level of aptitude being admitted. The returns of such a survey would be accordingly vague. Since testing every respondent's aptitude level is impracticable in representative surveys, language demographic data are usually based on self-report. Self-report data are notoriously unreliable, but useful information can be extracted from them nevertheless if they are broken down for social variables such as sex, class and generation. In any event, language-demographic data are hard to come by. National census questionnaires rarely include language-related items. Questions about language are sometimes deliberately excluded as potentially troublesome. Where language-related questions are included in a national census, they are often inadequate. For example, respondents may be requested to disclose their mother tongue in such a way that only a single answer is permissible. Language-contact situations call for fine-grained research instruments capable of recording complex patterns of language affiliation, language use, language choice and language preference. As an example of what this involves, consider a selection of some questions about language included in the Ethnic Diversity Survey of the Government of Canada,[11] a country that has a history of immigration and concern with ethnolinguistic diversity.

Language-related questions of the Ethnic Diversity Survey of the Canadian government

(1) What was the language that you first learned at home in childhood?

INTERVIEWER: Specify up to 3 responses. Accept multiple responses only if languages were learned at the same time.

If the respondent answers 'Chinese', ask 'Would that be Cantonese, Mandarin or another Chinese language?'

English
French
Cantonese
Mandarin
Italian
German
Punjabi
Spanish
Polish
Portuguese
Arabic
Tagalog (Pilipino)
Other – Search
Refused
Don't know

(2) Can you still understand ^ FirstLang?

Yes
No
Refused
Don't know

(3) Using a scale of 1 to 5, where 1 is not well at all and 5 is very well, how well can you understand ^ FirstLang now?

(1) – not well at all
(2)
(3)
(4)
(5) – very well

Refused
Don't know

(4) Can you still speak ^ FirstLang?

Yes
No
Refused
Don't know

(5) **Using a scale of 1 to 5, where 1 is not well at all and 5 is very well, how well can you speak ^ FirstLang now?**
(1) – not well
(2)
(3)
(4)
(5) – very well

Refused
Don't know

(6) **Besides the language of interview and your first language, are there any other languages that you speak well enough to conduct a conversation?**
INTERVIEWER: Report only those languages in which the respondent can carry on a conversation of some length on various topics.

Yes
No
Refused
Don't know

(7) **What languages are these?**
INTERVIEWER: Specify up to 6 responses.
Report only those languages in which the respondent can carry on a conversation of some length on various topics. If the respondent answers 'Chinese', ask 'Would that be Cantonese, Mandarin or another Chinese language?'

English
French
Cantonese
Mandarin
Italian
German
Punjabi
Spanish
Polish
Portuguese
Arabic
Tagalog (Pilipino)
Other – Search

(8) **Using a scale of 1 to 5, where 1 is not well at all and 5 is very well, how well can you speak ^ SpokenLang?**

(1) – not well
(2)
(3)

(4)
(5) – very well

Refused
Don't know

(9) What language do you speak most often at home?

INTERVIEWER: Mark up to 3 responses. Accept multiple responses only if languages are spoken equally often.
For a person who lives alone, report the language in which the respondent feels most comfortable (this can be the language the respondent would use for talking on the telephone, visiting at home with friends, etc.).

Do not include languages used exclusively in a homebased business.

Language of Interview
Language 1 derived
Language 2 derived
Language 3 derived
Language 4 derived
Language 5 derived
Language 6 derived
Language 7 derived
Language 8 derived
Language 9 derived
Refused
Don't know

(10) Are there any other languages you speak on a regular basis at home?

Yes
No
Refused
Don't know

(11) What languages are these?

INTERVIEWER: Mark up to 3 responses.

Do not include languages used exclusively in a homebased business.

Language of Interview
Language 1 derived
Language 2 derived
Language 3 derived
Language 4 derived
Language 5 derived
Language 6 derived
Language 7 derived

Language 8 derived
Language 9 derived

(12) **What language do you speak most often with your friends?**
INTERVIEWER: Mark up to 3 responses. Accept multiple responses only if languages are spoken equally often.
If the respondent uses different languages with different groups of friends, ask them to answer in a general way (for example, report the language they used most often, thinking of contact with all their friends together).

Language of Interview
Language 1 derived
Language 2 derived
Language 3 derived
Language 4 derived
Language 5 derived
Language 6 derived
Language 7 derived
Language 8 derived
Language 9 derived
Don't have any friends
Refused
Don't know

(13) **Now, I have two questions about your use of languages when you were growing up**.
Up until you were age 15, what language did you and your parents use most of the time when speaking to each other?
INTERVIEWER: Mark up to 3 responses. Accept multiple responses only if (1) languages were used equally often, or (2) different languages were spoken with each parent, or (3) respondent used one language when speaking to their parents and their parents used another language when speaking to them.

(14) **Up until you were age 15, what language did you and your brothers, sisters, and any other children in your household, use most of the time when speaking to each other?**
INTERVIEWER: Mark up to 3 responses. Accept multiple responses only if languages were used equally often or if different languages were spoken with different siblings.

If a survey using an instrument of this sort is conducted at regular intervals, it can reveal changes in a community's sociolinguistic arrangement, and even a single survey produces information that is indicative of the status and function of a language and tendencies of change. For example, significant differences in the responses to questions 1 and 9 between generations would

suggest intergenerational change, whereas similar results across all age groups are indicative of the stability of a language in a given society. Research along these lines, on the basis of both census data and more specialized research instruments, has been carried out first and foremost in 'classic' immigrant countries, especially the United States (Fishman et al. 1966), Australia (Clyne 1991), but also Western Europe (Verdoodt 1989; European Commission 2006). Various approaches have been developed to explain how languages fare when carried by their speakers to novel environments.

Language shift and maintenance

The key notions here are 'language shift' and 'language maintenance'. Language shift, defined by Weinreich (1953: 68) as the gradual replacement of one language by another, is often the outcome of language contact in an immigration situation. Language maintenance, on the other hand, refers to a situation where a speech community, under circumstances that would seem to favour language shift, holds on to its language. For instance, the transmission of Korean to the next generation of speakers in South Korea is not the result of language maintenance, but in Japan, which has a Korean minority of some 600,000, it is (Maher and Kawanishi 1995). There is no language contact in South Korea which could induce language shift, but in Japan Koreans are in direct contact with Japanese, virtually all of them being bilinguals. Fishman, who laid the groundwork for the scientific investigation of language shift, states the general direction of research:

> The study of language maintenance and language shift is concerned with the relationship between change and stability in habitual language use, on the one hand, and ongoing psychological, social and cultural processes, on the other, when populations differing in language are in contact with each other. (Fishman 1964: 32)

Language loyalty

It is not easy to ascertain that language shift has occurred or is occurring, some of the reasons for which were given above. Why it occurs is just as difficult to find out. This question can refer to the language choice of an individual, a family or a whole community. Generally, the same variables apply as in language change, that is, sex, age and class (cf. Chapter 5), intercommunity relations – between immigrant and host community or minority and majority – constituting an additional dimension of variation. Focusing on the community-building and culture-sustaining functions of language, Fishman (1972: 123f.) has used the notion of language loyalty, that is, the variable function of language for group maintenance, as an explanatory concept. In immigrant situations, the language loyalties of

immigrant and host community meet, producing expectations on both sides as to the immigrants' acquisition of the majority language and maintenance of their language of ancestry. The extent to which a society uses language as an index of acculturation, Fishman maintains, is culturally variable, putting linguistic minorities under more pressure to adapt in some societies than in others. Research has shown that different immigrant communities behave differently in the same host community which would seem to indicate that language attitudes on both sides, the host community and the immigrant community, interact to produce differing language shift and maintenance patterns.

Since both language loyalty and tolerance for linguistic pluralism vary across communities, it is hard to predict the fate of community languages. Migration often leads to language shift within three generations. For the monolingual speakers of the first generation, L1 is the default choice. L2, the dominant language of the target society, will remain a foreign language for most of them. Their children will grow up bilinguals, speaking L1 at home and L2 in most other domains. They have a real L1/L2 choice, but unless they continue to use L1 domestically, the third generation is likely to be L2 dominant if not monolingual. This pattern has been observed in several immigrant communities in the United States, but rather than representing a universal rule it reflects the assimilation pressure at work in the host society. In other environments migration does not necessarily bring language shift in its wake. For instance, Laitin (1993) reports that migrant communities throughout India retain the languages of their areas of origin. Since this is true of various different communities, it is not attributable to greater language loyalty on the part of the migrants, but reflects relatively higher tolerance for linguistic pluralism and cross-linguistic communication practices in India than in the United States.

At the same time, language loyalty is not a constant but a variable part of linguistic culture. In Australia, different ethnolinguistic groups were found to exhibit very different language maintenance and shift patterns (Clyne 1991). For example, after three generations, Danish immigrants had shifted to English almost completely, just 0.6 per cent maintaining active competence and use of Danish. In sharp contrast, after the same length of time Turkish immigrants had a maintenance rate of 83.6 per cent. A comparison reveals a pronounced distinction between, on the one hand, groups with a very high language maintenance rate, such as Turks, Greeks, Lebanese and Vietnamese, and, on the other, groups with low language maintenance rates and rapid language shift, most of them of northern European origin (Table 9.2).

These statistics testify to differing roles of language for individual and group identity in different societies. This is how the fact has been interpreted that the speakers of some immigrant languages living in a host-language-dominant environment abandon their language to replace it with the

Table 9.2 *Immigrant language maintenance in Australia by generation (Clyne 1991: 66f.)*

	1st gen.	2nd gen.	2nd + half gen.	3rd + gen.
French	66.1	39.8	6.5	2.0
French mix	22.6	13.4	4.8	0.8
Mauritian	68.3	31.9	6.5	23.4
German	61.1	27.3	3.7	1.3
German mix	14.5	9.6	2.5	0.4
Austrian	52.4	24.0	3.7	6.2
Dutch	48.9	14.7	2.5	10.3
Dutch mix	13.8	3.3	0.9	0.6
Danish	53.2	19.3	1.9	0.6
Norwegian–Swedish	45.6	21.1	1.2	0.8
Finnish	75.1	59.5	13.3	7.1
Russian	70.0	49.5	14.3	13.4
Russian mix	20.4	11.9	2.4	1.8
Ukrainian	72.1	51.7	43.3	42.1
Polish	73.3	39.3	12.5	13.6
Polish mix	13.6	7.5	1.5	0.8
Czech	64.1	33.4	8.0	18.9
Slovenian	64.4	55.4	9.2	21.1
Hungarian	70.6	49.4	12.8	36.0
Croatian	94.8	92.7	54.2	62.5
Serbian	91.9	91.3	60.0	67.2
Macedonian	96.2	91.2	64.5	61.7
'Yugoslav'	79.5	66.7	26.0	35.8
Greek	92.2	88.3	56.6	48.5
Greek mix	31.4	34.9	11.4	3.6
Maltese	70.5	38.6	12.5	25.3
Italian	88.0	70.0	31.7	27.6
Italian mix	26.4	25.4	6.7	2.4
Spanish	84.2	76.3	6.6	8.1
Spanish mix	41.3	28.3	4.5	0.8
Latin American	89.3	84.2	7.5	42.2
Portuguese	83.0	78.2	15.0	18.2
Turkish	93.6	93.4	67.6	83.6
Lebanese	92.5	82.0	27.8	35.8
Arabic	87.8	87.3	24.3	32.8
Sinhalese	33.8	6.5	2.6	16.2
Vietnamese	94.9	89.4	41.4	79.9
Lao	94.4	93.7	–	76.5
Khmer	91.2	79.1	–	71.6
'Filipino'	75.2	46.0	25.0	50.9
Chinese	81.9	66.4	22.7	16.6
Chinese mix	30.6	13.1	3.5	0.6
Irish	2.2	0.4	0.3	0.2

dominant language faster than speakers of other immigrant languages living in the same environment. How social class interacts with the host country's linguistic culture and with the immigrant group's language loyalty is another dimension of the problem which has to be studied on a case-by-case basis. Social advancement is often contingent on a good command of the dominant language, and the opportunity to acquire the dominant language depends on the employment situation. Immigrants who have to hold down two or three jobs to make ends meet are not in a position to give high priority to language study, which may translate into language maintenance. On the other hand, climbing up the social ladder is often considered more important than language loyalty. How these two tendencies are balanced out depends on culturally transmitted attitudes and educational support. For instance, the maintenance of Turkish in Australia has been on the decline for some time, leading Beykont (2010: 104) to conclude 'that additional policy and program measures are necessary to sustain Turkish bilingualism beyond the second generation'.

Ethnolinguistic vitality

Another concept that has been developed to predict language behaviour in language-contact situations is ethnolinguistic vitality (EV). Giles, Bourhis and Taylor (1977) have defined EV as an aggregate of sociocultural factors that determine a group's ability to function as a distinct collective entity. The major factors involved in a group's EV have already been mentioned in passing. They are *demography*, *institutional support* and *status*. Demography refers to the absolute size of the group and its relative strength in the total population as well as residence patterns – concentrated or dispersed – birth rate, endogamy and continuing migration. Institutional support concerns the presence of the group's language in the institutions of various social arenas such as education, government, media and religion. And status refers to the group's position in a social prestige hierarchy which is itself a composite factor involving the group's immigration history as well as social, economic, cultural and linguistic aspects.

EV theory has been criticized for methodological shortcomings, especially the difficulty of controlling the many variables involved and the resultant problems of comparing different cases. Yet, it has fostered many studies on language shift and maintenance (Allard and Landry 1992; Landry and Allard 1994).

An important distinction in EV studies is between objective EV and subjective EV, the latter being the *perceived* rather than the actual demographic strength, media presence and prestige of the language. Subjective EV is widely considered a better predictor of changing language-choice and language-use patterns than objective EV because people act on the basis of perceptions, assumptions and assessments rather than facts, most of which,

such as the actual size of their ethnic group, are unknown to them. It has been a default assumption that minority languages with low subjective EV are more likely to be replaced by the dominant language than languages with high-perceived EV. Bilingual speakers will be influenced in their language choice, so the argument goes, by their subjective assessment of their own ethnic group's vitality. Speakers with low vitality will tend to increase their choice of the dominant language, leading to linguistic assimilation, that is, language shift. Many case studies corroborate this correlation between subjective EV and language choice. More generally, high- and low-perceived EV predict, respectively, language maintenance and language shift. However, contrasting patterns have also been found (Harwood, Giles and Bourhis 1994), which show the complicated interaction between ethnolinguistic vitality perceptions, intercommunity relations and language-choice patterns in contact situations. Subjective EV is a relative concept which concerns a community not in isolation but in its relation to another, usually dominant, community and, as the case may be, further immigrant/minority communities. Immigrant communities hailing from the same country have different vitality ratings in different immigration contexts and, accordingly, develop different language-choice patterns and preferences.[12]

Territories and domains

Among the factors determining language shift and maintenance two have attracted special attention: (1) the micro-social arena of the family as the agent of spontaneous intergeneration and language transmission; and (2) the macro-social arena of group settlement in a 'territory'. The absolute demographic strength of a group means little if its members are widely dispersed, providing few social settings for using its language outside the family. And if a language ceases to be transmitted domestically the bedrock of its continuing tradition is undermined. Micro- and macro-social factors interact in that families are influenced in their language use by the surrounding community. Family and concentrated minority-residence areas are domains and territories of language. Language-contact situations differ in terms of the separation, upholding and invasion of domains.

Nations, ethnic communities, tribes as well as many animal species are examples of actors which operate mainly within certain territories. These can be understood literally in terms of physical space (Johnstone 2010) or in terms of abstract domain characteristics (Dailey-O'Cain and Liebscher 2011). In this sense, languages and speech varieties are tied to territories and domains. A language can be said to be robust if no other language can invade its territory. What are the conditions for robustness to obtain and, conversely, for a language to invade another's territory? Under normal circumstances a single newcomer is unlikely to make natives convert to

his or her language because the natives can communicate more easily among themselves using their native language. Newcomers, therefore, have a strong tendency if not to assimilate at least to adopt the host community's language. However, in social settings characterized by an extreme power differential, such as is characteristic of colonial situations, a small minority may be able to invade a language territory because certain members of the host community – i.e. co-opted colonial elites – defer to the invader by using his language. The time-tested principle *cuius regio, eius lingua* 'whose realm, whose tongue' comes to bear here.

As another example of successful invasion, consider the setting of a scholarly conference in a small country such as Holland. In such a setting a single non-Dutch-speaking participant will in all likelihood function as a carrier of English which thus invades the territory of Dutch. Since Dutch academics all speak English and since, in a context of scholarly exchange, they will give priority to the cooperative principle (see above, pp. 12f.) over other considerations they may entertain, such as attaching importance to the use of Dutch as a language of science, they will accommodate the invader by choosing English rather than Dutch for the discussion.[13] Communication accommodation theory accounts for ways in which individuals modify their speech behaviour in relation to an interlocutor (see above, pp. 37f.). The Dutch conference participants' choice of English can be explained in terms of opting for the lingua franca of the international domain of science by way of accommodation. However, at the same time it is an example of domain invasion if the reference domain is Dutch academia. In this particular case, domain invasion is facilitated by the fact that scientific discourse is one of the domains most susceptible to the forces of globalization.

Eventually domain invasion may lead to colonization, a process that occurs whenever the location/domain of a less successful species is taken over by a more successful neighbour or invader. This is what has happened in Dutch academia over the past two decades or so, though it was not an invasion by brute force. It must be acknowledged that Dutch academics, rather than being passive victims of the process, have an active part in it. Having a choice between Dutch and English, they allowed English into the academic domain, as a supplementary language in exceptional cases at first, whose employment in scientific discourse then gradually expanded. If we want to explain how English penetrated the educational domain in the Netherlands, the role played by those who chose to deliver their lectures, engage in discussions and publish their papers in English must not be overlooked. From a language-centred point of view, English can be said to invade Dutch academia, extending its territory by adding yet another domain of use. A speaker-centred approach focuses on individual acts of choice emphasizing pull- rather than push-factors.

In the maintenance of language, the role of time perspectives is critical. When a sufficient number of speakers anticipate future encounters and care

enough about 'their future together', the conditions are good for a linguistic tradition to be continued. Clearly, Dutch academics do not see their future together endangered by the domain invasion of English, because they consider English an addition to, not a replacement of, Dutch. Academia is just one of several communication domains of Dutch society, and in others Dutch is unchallenged. Functional domain allocation and division of labour of languages and language varieties is common in all societies. It should be noted, however, that for a language to be driven out of a domain means a reduction of its functional range which, in turn, will gradually lead to an erosion of the register appropriate to that domain, making a reversal of the process in future ever more unlikely. Partial language shift thus occurs which may spill over into other domains. A rearrangement of functional domain allocations with no consequences for the vitality of the language that forfeits a domain is unlikely. Rather, domain invasion must be considered the most obvious indication of language shift (Fasold 1984: 213; Appel and Muysken 1987: 39f.). Observers have diagnosed the domain invasion of English in the Dutch language area as an indication of on-going functional adjustment:

> Dutch may not be threatened with extinction in the short or medium term, but it is in danger of losing domains. It could eventually become just a colloquial language, a language you use at home … but not the one you use for the serious things in life: work, money, science, technology. (Van Hoorde 1998: 6)

In many small speech communities, language shift proceeds domain by domain. The domestic domain, researchers agree, is the last bastion, but a language that is reduced to the domestic domain is increasingly unlikely to be maintained because each domain loss lessens its utility (see below). The general trend is for 'bigger' languages to invade domains of 'smaller' ones. English invades domains of Dutch, Dutch invades domains of Frisian, German invades domains of Rhaeto-Romance, French domains of Breton, etc., and not vice versa. However, in complex multilingual situations it is not necessarily the biggest languages that are the most vigorous invaders. In many African contexts where a three-tiered arrangement of local vernaculars, regional lingua francas and former colonial languages is typical, shift from a local vernacular to a widely spoken African language such as Swahili, Hausa, Manding and Songhai is more common than to English, Portuguese or French. In the event, language shift is not the result of social engineering but happens spontaneously as communities adapt their means of communication to changing communication needs. Matthias Brenzinger, an expert on language shift in Africa, gives the following explanation: 'Assimilation by choice will be the main cause for the worldwide decline of minority languages in the future' (Brenzinger 1996: 282).

As we have seen, the reasons for language shift and the worldwide reduction of linguistic diversity it brings about are many. One sticks out,

however, which warrants further discussion. Language shift is always in the direction of the language of greater utility.

Utility

A case of language shift in South Africa well illustrates the complexity of the process, involving as it does an immigrant community's shift to a language that is not native to South Africa but has successfully invaded the South African language territory. In the event, a shift is taking place from Telugu to English. Prabhakaran (1998) describes the adoption of English instead of Telugu as the domestic language in the South African Indian community as a conscious choice that parents made for their children. In their decision they were influenced by several external factors, one of them being the South African government's language policy which assigns no Indian language official status. Yet, they could have chosen to transmit their ancestral language to their offspring, as many immigrants do. No threat of reprisals deterred them from so doing. The principal reason behind the South African Indians' decision not to raise their children as Telugu speakers is that English has greater utility and, therefore, commands higher prestige in South Africa than Telugu.

Utility is an economic notion which has been invoked in the analysis of language shift and other macro-sociolinguistic processes (Coulmas 1992a; Grin 1996). It is assumed that language itself has a utility utility-value which is variable, and that the actual and the perceived utility-value of languages in contact situations are valid predictors of language shift and maintenance.

In the most general sense, economic theory is concerned with the optimization of means–ends relations. Language is a tool, a means designed to carry out certain tasks and to achieve certain ends. If people act rationally, which most theoretical models of economics assume, they will, as a general principle, minimize their expenditure of time and effort to achieve their ends and, therefore, choose the instrument optimally suited for a certain job, that is, the instrument with the highest utility-value. In a number of language-choice situations, it is easy to see the general validity of this principle, especially where the deliberate acquisition of additional languages is concerned. Discussing the reasons why learners find Esperanto unattractive compared to English and why governments fail to add Esperanto to school curricula, Li Wei (2003:44) explains:

> In terms of language choice, with the exception of situations where a predominant language is imposed (as is typical of colonial societies) the choice of one second or foreign language over another hinges essentially on the perception of how useful the target language is.

The foreign language market works much like other markets, foreign languages functioning as marketable goods whose exchange value depends on

supply and demand which in turn reflect their utility-value. The vast majority of all languages are not marketed at all or in very specialized niche markets only because they have no utility for non-native speakers, and there is accordingly no demand for them. For the rest, the foreign language market is heavily dominated by a few European languages, English occupying the unrivalled position of greatest demand and greatest supply. As a commodity, English supports a bigger industry, generates a higher turnover and more revenues than any other language (McCallen 1989). And, most importantly, in most contexts the expenditure for its acquisition as a foreign or second language promises higher returns than that of other languages.

The lure of English

Arguably, the dynamics at work in the foreign language market are not the same as those in the world language system at large. When it comes to choosing their mother tongue or their children's mother tongue, people do not submit to market forces as readily as with regard to foreign language acquisition. This is true, because a language has not just instrumental value (utility), but symbolic value, too, the topic of the next chapter. Also, it must be noted that, just as ethnolinguistic vitality has an objective and a subjective dimension, the objective utility-value of a language may differ from its subjective perception. However, in multilingual settings the foreign language market and decisions concerning foreign language learning have implications that are fundamentally different from those in environments operating under monolingual assumptions. Kamwangamalu (2003) discusses the choice of English-medium schools in South Africa's black urban communities as a harbinger of language shift, since it is spreading rapidly from the school to the family domain. The main reason is, as Grin (2001: 73) puts it, that even at low levels of competence English is always associated with higher earnings.

Trivial and, as many who reflect on human nature would think, depressing as it may seem, the promise of higher earnings alone goes a long way towards explaining the phenomenal spread of English[14] which furnishes the goal of language shift in numerous speech communities on all continents. There can be no doubt that relative utility-value differential is the most consistent predictor of language spread and language shift. The utility-value of a language is an aggregate encompassing the following factors:

- the communicative range of a language measured as the demographic strength of the community using it as a first and a second/foreign language;
- the investment made in a language in terms of lexical recording, dictionaries, translations, electronic processibility, etc.;
- demand for a language in the international language market;
- a language's level of development as a societal means of production.

These factors are of a social and economic nature, which should, however, not obscure the fact that linguistic factors, too, feed into the utility-value of languages. In language, as in the case of other instruments, structure is a function of use, which is another way of saying that every language is the product of collective labour, shaped by millions of speech acts to suit the needs of its users. If this is so, it follows that language spread has cumulative effects because it means that a language which spreads is adjusted to ever new functions and domains. In our day, English is the most formidable example. Though looking back on only a relatively short literary history (compared, for example, with Chinese or Greek), it has been acculturated and nativized on all continents and put to use to a greater variety of functions in more diverse cultural contexts than any language in history. Its multiple uses have made it a most powerful and versatile instrument which increases its utility in a snowball-fashion by the day as people of different social and cultural backgrounds around the globe see in it a means of social advance and its acquisition as a profitable investment.

English is unrivalled in terms of the information accessible in it; it is flexible and open to innovation, absorbing elements of other languages uninhibitedly, and as a result commands a breadth and depth of lexical differentiation unmatched by other languages. Some increasingly important domains such as science and technology are dominated by English, which is also the language of choice by those critical of these developments. Hegemony expresses itself in English, and its discontents, too. That is, those who lament the spread of English at the expense of other languages contribute to its growing expressive power. In the passage quoted at the outset of this chapter, Mohandas Gandhi, champion of India's liberation from colonial rule, refers to the 'lure of English' which turns Indian languages into 'paupers'. By choosing English rather than Indian languages for purposes of higher communication – education, science, government, law – the Indian elites failed to develop their languages and adapt them to the requirements of modern life.

What Gandhi observed is quite common. Minor languages are not adapted to modern life which to many of their speakers seems incompatible with the traditions and community values they associate with their language. Such attitudes find expression in the belief that a language itself is a territory that must be protected against invasion by other languages in the form of borrowing, code-switching and other ways of language mixing which are seen as a first step towards impending language shift. Ironically, attempts at protecting the integrity of a language are often counterproductive, accelerating its stylistic deterioration and eventual demise instead of safeguarding it for future generations. As Hill (1993: 84) observes, language shift is driven in some cases by an ideology of linguistic purism promoted by self-styled language masters who denounce the speech of others for use of loanwords and grammatical interference. An unintended side effect of such protectionism is that

castigated speakers are discouraged from using the language in public, further pushing it back to the domestic domain. The noble choice of the right and the pure may thus turn out to be the deathblow for a small language under pressure. Purism, and more generally language ideology, is yet another variable that affects language shift.

English is the very opposite of a protected market, functioning as the biggest donor and borrower language at the centre of worldwide linguistic exchange. The resultant lexical and structural properties reflect the multifarious uses to which English has been put and are at the same time conducive to its application to ever new tasks and domains. Manifesting the same self-reinforcing mechanism of mutual dependency in reverse, languages that lose domains undergo lexical and structural impoverishment, continuously weakening the functional incentives of maintaining them.

Although language is not just an instrument, its instrumental functions are vital. A language with limited or no instrumental utility is more likely to be abandoned by its speakers than one that is serviceable in a wide range of domains. The lesson to be learnt from purism and the role it plays in language shift and maintenance is that utility can be ignored only at a cost. At the heart of the study of language shift must be the question to what extent utility as a determinant of collective language choice can be mitigated by cultural and sociopsychological factors. Language shift from smaller to bigger languages locally and the worldwide spread of English testify to the inclination of individuals and communities to act according to the principle of getting the maximum benefits for the minimum of effort. Against this background, the student of language shift and maintenance asks to what extent communities in contact situations rely on cultural and sociopsychological values to withstand the forces of the language market. These forces exert a strong influence on people's choices which, however, does not mean that they have no choice. Language shift occurs as the result of choices made by individuals, most importantly in the domestic domain, in accordance with their own motivations, expectations and goals which they may or may not share with other members of their community. Taken together, individual choices form collective choices that impact on the future of a speech community and its language. It lies in the nature of collective choice that the individuals involved in it do not foresee all consequences of their acts, for example, that by choosing – in the best interest of their children's social advance – the majority language of the larger environment as their domestic language, they may contribute to the recession and demise of their ancestral language. That, by insisting on a native expression rather than a loanword, a purist language guardian may contribute to the same result is likewise a choice with unforeseen consequences. Both in immigrant communities and in small autochthonous communities coming under pressure in their native environment, the continuation of a linguistic tradition depends, ultimately, on individual acts of choice made in a social environment

incorporating a vast number of variables whose interaction is hard to predict for social actors as well as for social scientists.

Conclusion

The key insight from the voluminous literature on language spread, language shift and language maintenance is that while speakers have a quasi-natural attachment to their native language, there is great variability in regards to the propensity of individuals and groups to maintain their language under adverse circumstances or, conversely, to abandon it in favour of another. The factors that have a bearing on language maintenance and shift are many and complex. Linguistic diversity/homogeneity, culture, demographics and socioeconomic conditions interact as determinants of individual speech behaviour and language choice. Mass migration and imperial expansion have transplanted many languages from their native lands into disparate environments causing changes in the linguistic ecology. The quest to understand how immigrant languages fare in their new environments and how language arrangements are affected by the penetration of an imperial language or language of wider communication has engendered a number of theoretical notions discussed in this chapter, notably language loyalty, ethnolinguistic vitality, territory, functional domain and utility. They all highlight separate aspects of unstable language arrangements and factors that have a bearing on the language choices that ultimately shape them. For sociolinguists it is important to adopt a speaker-centred perspective: it is not a language that thrives or withers, but speakers who continue or cease to use it. Without speakers no linguistic environment exists. Yet, the extant linguistic environment influences speakers' behaviour and choices, although they do not necessarily realize how. This is at the bottom of the complexity of explaining unstable social language arrangements. It is appropriate, therefore, to conclude this chapter with Georges Lüdi's (1986: 220) pointed remark that 'in language choice, as in other domains of linguistic performance, speakers are not always aware of what they are doing'.

Questions for discussion

(1) What is the role of demographic strength in language spread?
(2) Are language shift and language maintenance mirror-image processes or can they be understood each in their own right only?
(3) What is ethnolinguistic vitality, and how does it bear on language spread and language shift?
(4) Think of a setting where (the speakers of) one language successfully invaded the territory of another prompting language

shift. Is the concept of domain invasion applicable to all language-shift situations or is it more suitable for some than for others?

(5) Since language is an instrument, a utility differential in language-contact situations is a predictor of language shift, although utility is not the only determinant of language choice. What other factors modulate its effectiveness?

Notes

1. The range of contact linguistics as an interdisciplinary field relying on methods and theories developed in psychology and sociology, as well as auxiliary disciplines such as demographics and migration studies, is documented in Nelde (1995).
2. UN statistics: http://unstats.un.org/unsd/demographic/sconcerns/migration/
3. Figures in this chapter are taken from *Ethnologue*, 16th edition, released 2009 (www.ethnologue.com/) and the United Nations Population Division (www.un.org/esa/population/).
4. Data from *Ethnologue*, 16th edition, 2009.
5. Cf. Coulmas (1999) for an overview.
6. Figures are again from *Ethnologue* and Internet World Stats (www.internetworldstats.com/). Notice that statistics of this sort are rather short-lived, yet they are indicative of the general situation.
7. Bihari is a cover term for Bhojpuri, Magahi and Maithili.
8. For a system-theoretic analysis of the hierarchical ordering of languages in the world system, see Swaan (2001).
9. UNESCO 2010: www.unesco.org/new/en/culture/themes/endangered-languages/; cf. also Crystal (2000).
10. Khubchadani (1983:103–7) deals at length with the problem of oscillation of census returns regarding mother tongue in India, especially north-central India which he calls a 'fluid zone'. He explains: 'The composite characteristics of communication patterns in the entire Hindi–Urdu–Panjabi region are at variance with the concerns for identity expressed through the claims of mother tongues, which have a marked tendency to shift in every decennial census depending on the prevailing sociocultural climate.' Kamusella (2012) examines extant standards of language classification, demonstrating that many listings by missionary-linguists of the Summer Institute of Linguistics and the *Ethnologue* are arbitrary and ad hoc.
11. Language statistics of Canada's 2006 census can be inspected at www12.statcan.gc.ca/census-recensement/2006/rt-td/lng-eng.cfm.
12. See, for example, Yagmur (2004) who compares Turkish immigrant communities in Australia and Western Europe. While Turkish has generally high vitality scores as compared to other immigrant languages, differences between countries point to complex interrelations between immigrant and host communities and immigrant and dominant language. For instance, in the Netherlands, Turkish has the highest vitality score of all immigrant languages correlating with high language maintenance rates. In France, on the other hand, second-generation Turkish immigrants exhibit a marked tendency to shift to French.
13. The dominant role that in the past half-century English has come to play in science has become a field of sociolinguistic research in its own right with its own scholarly journals and societies and an incalculable number of publications. Cf. Ammon (2002), Sandelin and Sarafoglou (2004).
14. Current explanations for the spread of English fall into two camps. Phillipson (1992) leads the faction of those who explain the spread of English as the result of deliberate language policies by the British and US governments associating it with the evils of global capitalism let loose. By contrast, Crystal ([1997] 2003) takes the view that the

spread of English is a grassroots development largely unconnected to the processes of enrichment and impoverishment accompanying globalization. See also Pennycook (1994), Fishman, Conrad and Rubal-Lopez (1996) and Chapter 13 below.

Useful online resources

The Language Observatory Project (LOP) constitutes a worldwide consortium of partners: www.language-observatory.org/

World Network for Linguistic Diversity: http://maaya.org

Index of Linguistic Diversity: www.terralingua.org/linguisticdiversity/

UNESCO interactive atlas of endangered languages: www.unesco.org/new/en/culture/themes/endangered-languages/news/dynamic-content-single-view-news/news/the_interactive_atlas_of_endangered_languages_updates/

Migration Policy Institute: www.migrationinformation.org/datahub/europe.cfm

Further reading

Brenzinger, Matthias. 1992. Patterns of language shift in East Africa. In Robert Herbert (ed.), *Language and Society in Africa*. Cape Town: Witwatersrand University Press, 287–303.

Coulmas, Florian. 2010. The ethics of language choice in immigration. *Policy Innovations*. www.policyinnovations.org/ideas/commentary/data/000162

Clyne, Michael. 1991. *Community Languages: The Australian Experience*. Cambridge University Press.

Crystal, David. 2000. *Language Death*. Cambridge University Press.

Extra, Guus. 2010. Mapping linguistic diversity in multicultural contexts: demolinguistics perspectives. In J. A. Fishman and O. Garcia (eds.), *Handbook of Language and Ethnic Identity*. 2nd edn. New York: Oxford University Press, 107–22.

Fishman, Joshua A. 1991. *Reversing Language Shift: Theoretical and Empirical Foundations of Assistance to Threatened Languages*. Clevedon: Multilingual Matters.

Hyltenstam, Kenneth and Åke Viberg (eds.) 1993. *Progression and Regression in Language*. Cambridge University Press.