



Words  
of the  
World

ABRAM DE SWAAN



# Words of the World



# Words of the World

The Global Language System

*Abram de Swaan*

Polity

Copyright © Abram de Swaan 2001

The right of Abram de Swaan to be identified as author of this work has been asserted in accordance with the Copyright, Designs and Patents Act 1988.

First published in 2001 by Polity Press in association with Blackwell Publishers Ltd

*Editorial office:*

Polity Press  
65 Bridge Street  
Cambridge CB2 1UR, UK

*Marketing and production:*

Blackwell Publishers Ltd  
108 Cowley Road  
Oxford OX4 1JF, UK

*Published in the USA by*

Blackwell Publishers Inc.  
350 Main Street  
Malden, MA 02148, USA

All rights reserved. Except for the quotation of short passages for the purposes of criticism and review, no part of this publication may be reproduced, stored in a retrieval system, or transmitted, in any form or by any means, electronic, mechanical, photocopying, recording or otherwise, without the prior permission of the publisher.

Except in the United States of America, this book is sold subject to the condition that it shall not, by way of trade or otherwise, be lent, re-sold, hired out, or otherwise circulated without the publisher's prior consent in any form of binding or cover other than that in which it is published and without a similar condition including this condition being imposed on the subsequent purchaser.

A catalogue record for this book is available from the British Library.

**Library of Congress Cataloging-in-Publication Data**

Swaan, A. de.

Words of the world : the global language system / Abram de Swaan.

p. cm.

Includes bibliographical references and index.

ISBN 0-7456-2747-1—ISBN 0-7456-2748-X (pbk.)

1. Language and languages—Philosophy. 2. Language policy. 3. English language—Foreign countries. I. Title.

P107 .S93 2001

401—dc21

2001036802

Typeset in 10  $\frac{1}{2}$  on 12 pt Sabon

by Kolam Information Services Pvt. Ltd., Pondicherry, India

Printed in Great Britain by TJJ International, Padstow, Cornwall

This book is printed on acid-free paper.

# Contents

*Preface* ix

<b>1</b>	<b>Introduction: the global language system</b>	<b>1</b>
1.1	The global language system: a galaxy of languages	4
1.2	A historical atlas of the world as a language system	6
1.3	Supercentral constellations in the present language system	11
1.4	Scope and approach of this book	17
1.5	Plan of the book	20
<b>2</b>	<b>The political economy of language constellations</b>	<b>25</b>
2.1	Languages as 'hypercollective goods'	27
2.2	The communication potential of a language: the Q-value	33

<b>3</b>	<b>Language, culture and the unequal exchange of texts</b>	<b>41</b>
3.1	Texts as commodities in international exchange	42
3.2	Protectionism and free trade in cultural exchange	47
3.3	Monoglossia, polyglossia and heteroglossia	53
3.4	Discussion	57
<b>4</b>	<b>India: the rivalry between Hindi and English</b>	<b>60</b>
4.1	Characterization of the Indian constellation	61
4.2	State formation, nation-building and language unification	63
4.3	The vicissitudes of language policy in India	69
4.4	Discussion	73
<b>5</b>	<b>The triumph of <i>bahasa Indonesia</i></b>	<b>81</b>
5.1	Gandhi's dream	81
5.2	The rise of Malay	83
5.3	The demise of Dutch	86
5.4	The rejection of Javanese	90
5.5	Discussion	93
<b>6</b>	<b>Africa: the persistence of the colonial languages</b>	<b>96</b>
6.1	A two-by-three comparison	99
6.2	Three francophone constellations south of the Sahara	102
6.3	Three English-centred constellations south of the Sahara	116
<b>7</b>	<b>South Africa: the survival of the old language regime</b>	<b>127</b>
7.1	The language regime under Apartheid	128
7.2	Language policy after Apartheid	132
7.3	The dynamics of the constellation	136



7.4 Discussion	140
----------------	-----

<b>8 The European Union: the more languages, the more English</b>	<b>144</b>
---	------------

8.1 Civil Europe (1): language unification in national constellations	146
8.2 Civil Europe (2): Q-values in the European Union	151
8.3 Institutional Europe (1): the public level	166
8.4 Institutional Europe (2): the bureaucracy	171
8.5 Discussion	173

<b>9 Conclusions and considerations</b>	<b>176</b>
---	------------

9.1 Conclusions	177
9.2 Considerations	187

<i>Notes</i>	194
<i>References</i>	225
<i>Index</i>	244



# Preface

This book is much more a product of the global language system than its author ever intended it to be. I began work on it in 1993, in the hospitable, anglophone environment of Cornell University, NY, where I held the Luigi Einaudi chair of international studies. Once I returned to the familiar and increasingly bilingual setting of the University of Amsterdam, I continued to write in English, using Netherlandish mostly for essays and columns in Dutch newspapers, much as one might take a plane to remote destinations and ride one's bike on nearby trips. In the autumn of 1995, I found myself abroad once again, this time in Budapest, where I occupied the European Union chair of social policy at the Eötvös Loránd University. This time I was surrounded by Hungarians who had such facility in English (and German, and French and Russian) that I never had a chance to pick up more Magyar than the most indispensable: *nagyon szépen köszönöm*, many thanks to my colleagues there, and most of all to Professor Zsuzsa Ferge, my thoughtful hostess and longtime friend. The decisive spurt in writing this book came in the academic year 1997/8, when I was elected to the *chaire européenne* of the Collège de France at the initiative of Pierre Bourdieu, who thus contributed much to the completion of this project. The professors of the Collège are expected to present thirteen *conférences* each year, and write them up, in French, of course. I made the best of my predicament and told the students that my halting French was an apt demonstration of the realities of the world language system. They were duly impressed,

by the realities. Indeed, I did write the first complete version of this book entirely in French: ‘Oops, wrong language!’ I rewrote the text, once again, in English: the version that lies before you. Finally, Leonoor Broeder translated it most competently into Dutch. Reading one’s own writings in translation feels like being impersonated by an actor. But seeing the translation into my own native tongue, I felt as though I were being confronted with a *Doppelgänger*, who to everyone else appears perfectly indistinguishable, but to me can never be the same.

In this edition, too, there is a slight sense of alienation: English is not the language of my childhood years (although it was my parents’ secret language which therefore had to be urgently deciphered).

Native speakers of English have the great fortune that their language has become *the* lingua franca of the world, which saves them much effort and yields them great opportunities. On the other hand, many hundreds of millions have appropriated this global vehicular as a foreign language, with considerable exertion, and much to their advantage.

A variety of Englishes has emerged in different parts of the world, but so far they have all remained mutually intelligible. The more peripheral and the more recent varieties command less prestige and attention than the established usage of educated native speakers in the UK and the US. Moreover, for a long time to come, American and British media will remain in control of the worldwide distribution of texts and performances in English. For all those who learned English after their childhood years, I hope that the privileged position of its native users will gradually erode. But that is not likely to happen soon. In the meantime, even the foreign speakers, underprivileged as they are, still share in the advantages of the one global language.

Among the pleasures of working on this book was the collaboration with a sequence of research assistants: Jannes Hartkamp in Budapest, Jeanne Kouta and Florence Colas in Paris, Jeroen Starrenburg and Jacek Magala in Amsterdam. For some reason they all turned out to be both gifted and devoted, a surprising combination that put me much into their debt.

The Royal Netherlands Academy of Science, the European Cultural Foundation and the Brouwers Stichting provided me with grants. Numerous colleagues from different disciplines and countries helped me with facts, references and critical comments. My intellectual debt to them should be apparent from the notes and citations in the book. One among them is mentioned here: David Laitin, who was an early companion in the political sociology of language.

I took the world as my research area, compared five different language constellations, and combined notions from economics, linguistics, history, political science and sociology in a synthetic perspective. A huge topic is not necessarily more difficult than a small one, nor is a multidisciplinary approach more learned than a specialist view, but they can be a bit more intimidating. I was encouraged throughout by my friends Johan Goudsblom and Kitty Roukens, who read each chapter as a draft and gave me their views with insight, generosity and tact.

This is in English, not in French, or Dutch, or Swahili (and the preceding phrase is absolutely untranslatable). Why it should be in this language rather than another can be explained from the dynamics of the global language constellation, and that is what this book is about.

Abram de Swaan



# Introduction: the global language system

The human species is divided into more than five thousand groups each of which speaks a different language and does not understand any of the others. With this multitude of languages, humankind has brought upon itself a great confusion of tongues. But nevertheless, the entire human species remains connected: the division is overcome by people who speak more than one language and thus ensure communication between different groups. It is multilingualism that has kept humanity, separated by so many languages, together. The multilingual connections between language groups do not occur haphazardly, but, on the contrary, they constitute a surprisingly strong and efficient network that ties together – directly or indirectly – the six billion inhabitants of the earth. It is this ingenious pattern of connections between language groups that constitutes the global language system. That is the subject of this book.

This worldwide constellation of languages is an integral part of the ‘world system’. The population of the earth is organized into almost two hundred states and a network of international organizations – the political dimension of the world system; it is coordinated through a concatenation of markets and corporations – the economic dimension; it is linked by electronic media in an encompassing, global culture; and, in its ‘metabolism with nature’, it also constitutes an ecological system. The idea of a global human society which indeed constitutes a system on a world scale has regained much attention in recent years. However, the fact that humanity, divided by a multitude

## 2 Introduction: the global language system

of languages, but connected by a lattice of multilingual speakers, also constitutes a coherent language constellation, as one more dimension of the world system, has so far remained unnoticed. Yet, as soon as it has been pointed out, the observation seems obvious.<sup>1</sup>

The global language constellation will be discussed in this book as an integral part of the world system. This implies that language constellations are considered as a – very special – social phenomenon, which can be understood in terms of social science theories. This, too, is new, albeit not entirely without precedent.<sup>2</sup> Rivalry and accommodation between language groups will be explained with the aid of the political sociology of language and the political economy of language. The former focuses on the structure of the language system and its subsystems, and looks at ‘language jealousies’ between groups, at elite monopolization of the official language, at the exclusion of the unschooled, and at the uses of language to achieve upward mobility; the latter approach analyses how people try to maximize their opportunities for communication, how this confronts them with dilemmas of collective action that may even provoke stampedes towards another language and the abandonment of their native tongue, and what occurs in the unequal relations of exchange between small and large language groups. Many of these notions from sociology and economics have never before been applied to languages or language groups.<sup>3</sup> Together they constitute a coherent theoretical framework that can explain events in such disparate language constellations as India and Indonesia, Sub-Saharan Africa and South Africa, or the European Union.

That language has emerged at all is a cause for marvel; its evolution into innumerable, mutually unintelligible languages is an equally amazing testimony to human ingenuity. As languages grew apart in the course of collective transmission and transformation, new forms of pronunciation must have emerged, thousands of new words appeared and hundreds of grammatical and syntactical rules (and as many exceptions) evolved. All of this was the result of human action and almost none of it was the outcome of human intention.

It seems increasingly likely that all languages that are currently spoken on earth are related and have developed from a common predecessor, roughly following the evolutionary path of present human beings from a common genetic stock in the course of some hundred and twenty thousand years. Evolutionary genetics, comparative linguistics and archaeology are now producing a quickly growing body of evidence for this shared origin.<sup>4</sup> But even if it turns out to be the case that the human species and its languages come from several,



diverse origins, there is no doubt that at present all human groups constitute a single interdependent whole, and that their languages together form a global constellation that represents one dimension of the modern world system.

Five or six thousand languages are spoken on earth. The number cannot be specified more exactly, because languages are not always countable. In this respect they resemble clouds: it is hard to tell where one begins and the other ends, and yet most clouds and languages are obviously distinct, with a clear expanse separating them.

In their inexhaustible variety and almost impenetrable complexity, languages are best compared to that other most complex and variegated phenomenon, life itself. Much as a biological species is defined by the capacity of pairs of male and female members to reproduce, a language may be defined by the capacity of any two speakers to understand one another. Two languages are considered distinct if the speakers of one and the other are mutually unintelligible. Just as species are subdivided into many varieties that can indeed interbreed, within languages various mutually intelligible dialects are discerned. Biological varieties of one species shade into one another as do dialects of the same languages, and that is why, in both fields, classification is so often controversial.<sup>5</sup> Indeed, cognate languages can be very hard to tell apart. Where in biology the proof is in the mating, in linguistics it is in the understanding. But mutual intelligibility is not simply a characteristic of the two languages involved; not entirely unlike interbreeding, it also depends on the individuals involved. They may have greater or lesser skills in understanding strangers, they may be more or less eager to communicate with one another, and the context of their encounter may be so structured as to facilitate mutual comprehension or hinder it.

There is no doubt that Chinese and Dutch are two entirely different languages, but it is a matter of controversy whether German and Dutch are indeed distinct languages,<sup>6</sup> while almost everybody would agree that Flemish and Dutch are two varieties of the same language (since their respective speakers would have no trouble at all explaining to one another, each in their own idiom, how insuperable the differences between the two are). Granting the cloudy nature of languages, nevertheless most of the time they are discussed here as if they were distinct entities, separated by barriers of incomprehensibility.

## 4 Introduction: the global language system

### 1.1 The global language system: a galaxy of languages

Mutually unintelligible languages are connected by multilingual speakers, but not at all in random fashion. In fact, the scheme of all the world's languages and of the multilinguals that connect them displays a strongly ordered, hierarchical pattern, quite similar to those reversed tree-structures that the French call 'organigrammes', charts used to depict the organization of armies or large bureaucracies.

The vast majority of the languages in the world of today, some 98 per cent of them, are situated in the lower part of this chart: these are the 'peripheral languages' and although there are thousands of them, all together they are used by less than 10 per cent of humankind. Very little of what has been said in all these languages has ever been recorded, be it on clay, stone, papyrus, paper, tape or disk. They are the languages of memory, and whatever was uttered in these languages could only endure because it was heard and remembered, repeated, understood and memorized again.<sup>7</sup> Rather than being defined by what they are not, as 'unwritten' or 'scriptless' languages, these languages deserve to be identified by what constitutes their strength: they are the languages of conversation and narration rather than reading and writing, of memory and remembrance rather than record.

Any two peripheral groups are mutually connected through members that speak the languages of both. But on the whole such ties tend to be scarce. Or rather, they are becoming scarcer since communication between the inhabitants of adjacent villages has become less important, as they increasingly come to deal with traders and administrators in the district capital. As a result, members of the various peripheral groups are more likely to acquire one and the same second language, one that is therefore 'central' to these groups. All or most communication between the peripheral groups occurs through this central language. The peripheral languages, grouped around the central language, may be compared to moons circling a planet. There may be about one hundred languages that occupy a central or 'planetary' position in the global language system.<sup>8</sup> Together they are used by some 95 per cent of humankind. The central languages are used in elementary education and usually also at the level of secondary and higher education. They appear in print, in newspapers, in textbooks and in fiction, they are spoken on radio, on cassettes and increasingly

on television. Most of them are used in politics, in the bureaucracy and in the courts. They are usually 'national' languages, and quite often the official languages of the state that rules the area. These are the languages of record: much of what has been said and written in those languages is saved in newspaper reports, minutes and proceedings, stored in archives, included in history books, collections of the 'classics', of folk tales and folkways, increasingly recorded on electronic media, and thus conserved for posterity.

Many of the speakers of a central language are multilingual: first of all, there are those whose native speech is one of the satellite, peripheral languages, and who have later acquired the central language. In fact, everywhere in the world the number of this type of bilinguals is on the increase because of the spread of elementary education and the printed word, and through the impact of radio broadcasting. The second type, on the other hand, that of the native speakers of the central language who have learned one of the peripheral languages, is much less common. Apparently, language learning occurs mostly upward, in a 'centripetal' mode: people usually prefer to learn a language that is at a higher level in the hierarchy. This again reinforces the hierarchical nature of the world language system.

If the mother-tongue speakers of a central language acquire another language, it is usually one that is more widely spread and higher up in the hierarchy. At this next level, a number of central languages are connected through their multilingual speakers to one very large language group that occupies a 'supercentral' position within the system. It serves purposes of long-distance and international communication. Quite often this is a language that was once imposed by a colonial power and after independence continued to be used in politics, administration, law, big business, technology and higher education. There are about a dozen of these supercentral languages. Their position in the global language system resembles that of so many suns surrounded by their planets, the central languages, which, in turn, are encircled by their respective satellites, the peripheral languages. The supercentral languages are Arabic, Chinese, English, French, German, Hindi, Japanese, Malay, Portuguese, Russian, Spanish and Swahili. All these languages, except Swahili, have more than one hundred million speakers and each serves to connect the speakers of a series of central languages. In subsequent chapters a number of regional constellations will be discussed, each centring on one or more of these supercentral languages, such as the Indian constellation around Hindi and English; the Indonesian constellation around Malay (*bahasa Indonesia*); the French-centred constellation of 'francophone' West Africa and the East African constellation that hinges upon English;

## 6 Introduction: the global language system

the South African constellation, where English and Afrikaans compete; and, finally, the constellation of the European Union, where a dozen national languages are increasingly linked by English, less and less by French and hardly any more by German.

If an Arab and a Chinese, a Russian and a Spaniard, or a Japanese and a German meet, they will almost certainly make themselves understood in one and the same language, one that connects the supercentral languages with one another and that therefore constitutes the pivot of the world language system. This ‘hypercentral’ language that holds the entire constellation together is, of course, English.

In the present world, English is the language of global communication. It is so to speak at the centre of the twelve solar language systems, at the hub of the linguistic galaxy.<sup>9</sup> English has not always held that position. On the contrary, it has now done so for only half a century or so and one day it may lose its hypercentral functions again, but in the next decades it is only likely to reinforce its position even further.

If the origins of language correspond closely to the origins of the human species, the spread of languages across the globe is intimately connected with the history of humanity. For scores of millennia, languages spread with demographic expansion and migration. In historical times, they followed in the wake of conquest, commerce and conversion. It is only since a century ago at most that languages spread more frequently through formal schooling than in any other way. But the educational system certainly does not operate independently of the political, economic and cultural context, which continues to shape the patterns of language acquisition.

### 1.2 A historical atlas of the world as a language system

The best way to visualize the evolving global language constellation is through a series of maps of the world.<sup>10</sup> Quite probably, in prehistoric times, as the human species scattered across the continents, small bands must time and again have left their main group, crossed mountains and seas, to settle in areas that were quite distant from the next human population. There, in isolation, and in the absence of any written texts, their languages may have changed rather quickly, reach-

ing unintelligibility with respect to the original language in the span of a few dozen generations.<sup>11</sup> Encounters with other human groups and the ensuing language contacts produced new amalgamations. Thus an imaginary map of the prehistoric distribution of languages would render language areas as fairly small circles, extending and elongating as language groups spread and trekked across new territory, stretching to the breaking point, when a separate circle would indicate the emergence of a 'new' language in that location.

Thus, the hypothesis of 'monogenesis', the evolution of all languages from a single predecessor, does not at all contradict the existence of a great many, mutually unintelligible, languages, once the human species had scattered across the continents. The early distribution of human languages was much more fragmented than the present world language system. Yet it is quite likely that bands in adjacent territories traded and intermarried and some people learned the language of the next group. The circles, no matter how small, may have shown some overlap in the more densely populated areas. As people settled and began to work the land, they must have developed a language for communication between adjoining villages: an early *lingua franca*, which appears on the map as a dotted line, enclosing the entire area where the linking language is used. That is where the pattern of language distribution is regaining some coherence.

The early 'military-agrarian' regimes, based on military conquest of agrarian communities, demanded the payment of tribute for protection (against other warriors and themselves).<sup>12</sup> With their dominion, they also usually imposed their religion, and their language. Thus the first 'central' languages emerged, linking the peripheral languages of the conquered communities through bilingual speakers to the language of the victors: the language of conquest, conversion and commerce. On the language map, the territory of such regimes would be rendered in a solid but rather pale colour, indicating its wide extension and relatively low density. The circles of the peripheral languages would still clearly show through in their respective areas.

The next stage of integration of the language system occurred with the formation of empires. Marching armies laid one people after another under tribute, maintained roads and harbours, and protected trading routes across the territory. The map of the language constellation in the year 1 shows several such 'world empires'. Not much is known about the western hemisphere or Africa, in that period, but on the Eurasian continent plenty of written records have survived. At least three languages had already spread along very long, but very thin lines. First of all, Latin, emanating from Rome all along the Mediterranean coast, stretched across the southern half of Europe and, more

## 8 Introduction: the global language system

sparsely, further to the north, into the Germanic and Celtic lands. Latin was a spoken and a written language; it served to administer the conquered areas, to carry out diplomatic missions and trading ventures, and to spread new knowledge and technology. Soon, moreover, it was to be the vehicle of Christian expansion. After the fall of the Roman Empire, Latin served for another fifteen hundred years as a major European linking language. But in all the many language groups of Christendom there were only a few individuals, clergy usually, who had learned the language of the church and hence could communicate with their peers all over the continent. They served as translators and mediators to connect their communities with the continental network. Until the Renaissance, Latin hardly had competitors as the language of learning and long-distance communication. The connecting web may have been extremely tenuous, the Latin speakers very few in number, but in the domains of scholarship, law and religion it held together until the nineteenth century. Thus, Europe, with Latin as its supercentral language, already constituted a coherent, if precarious, language system more than two thousand years ago. The language map of the era would have displayed the supercentral presence of Latin by a pattern of rays in a single colour, extending from Rome in ever thinner lines across the continent and overlaying the solid patches of central languages with the circles of the peripheral languages still visible underneath.

The second imperial language of that era was of course Chinese. In the core area of contemporary China, a 'pre-classical' version of Han Chinese already functioned as the language of long-distance communication, spoken and written by clerics and scholars, and used in court as the language of rule and administration. In South Asia, learned and religious men used Sanskrit for the same purposes and an equally fine and extended grid overlaid the language map of that subcontinent. If they had strictly limited their encounters to their peers, clergy and courtiers, at the time, might have travelled all across the Eurasian landmass, using only Chinese, Sanskrit and Latin (and maybe some Persian or Greek). But with the common people, the innkeepers and the traders, let alone peasants and soldiers, these languages would have been quite useless.

One thousand years later, the great classical languages had spawned vernacular versions all over their respective regions. Yet they continued to serve for long-distance communication in the fields of administration, diplomacy, religion, science, literature and trade over an area that had grown even larger in the meantime. Right at the centre of the Afro-Eurasian land mass, a fourth language had been spreading for some time: Arabic, originating in the Arabic peninsula

and extending its lines across northern Africa to the southern tip of Spain, along the East African coast and deep into Central Asia.

Clearly, the regions of the classical languages more or less coincided with the areas of the great religions, Islam, Christendom, Hinduism, Confucianism and Buddhism (the last of which overlaps with the preceding two). The supercentral networks were vast, but still very thin, as so few people could understand, and even fewer could read or write, the corresponding languages.

Another five hundred years later, at the 'dawn of the Modern Era', around 1500, the pattern of long-distance communication had begun to change perceptibly. First of all, the vernaculars that had sprouted from the great classical languages were coming into their own: they were crafted into standard versions by poets, writers and scholars, increasingly used in trade, science, law, administration and, in the West, also in religion and at princely courts.<sup>13</sup> In Europe, among the many popular languages that derived from Latin, Italian became a literary language early in the fourteenth century. It soon developed into a language of scholars, courtiers, politicians, artists, scientists and the military.<sup>14</sup> From the flourishing Italian city states it spread over southwestern Europe as a vehicular language of diplomacy and learning.

The other vernaculars that derived from Latin each also spread over broad territories of their own, helped by a vast increase in the circulation of their written versions through the new printing presses. Increasingly, they were used at the royal court and in the courts of law, in parliament, in the schools and academies. As will appear later, they succeeded in driving out the peripheral languages (or the marginalized, formerly central languages of conquered territories), in the process each becoming the hegemonic language of its realm.<sup>15</sup> A similar development rendered Russian and German hegemonic in their respective territories. The language maps of the epoch for Europe increasingly show closed, single-coloured areas of a more and more intense hue, while the patches of the peripheral languages slowly fade away. These processes of national language unification that occurred throughout Europe represent another stage in the integration process, this time on a smaller scale but with much greater density than in the preceding empires.

The new European vernaculars travelled overseas with the explorers to Africa, Asia and America, where they initially found a tiny niche near estuaries or on islands near the coast. Thus began their long career on distant continents as the languages of rule, trade and conversion.

Around this time, Arabic reached its zenith as a world language. But the language of the Koran was to be conserved in its unadulter-

## 10 Introduction: the global language system

ated form; any divergence could only spell degeneration. Hence, the vernaculars it engendered never developed into distinct, acknowledged languages as did the descendants of Han, Sanskrit and Latin.

The overseas expansion of Chinese was brought to a halt, once the Ming rulers suspended maritime trade and exploration in the early sixteenth century. As a result, henceforth their language could only spread over land, albeit across a vast expanse. In India, in the course of the sixteenth and seventeenth centuries, the Mughal kings extended their rule ever southward, at a time when the vernacular languages (the *prakrits* in the north, Dravidian languages in the south) had established themselves, each in its own area. One of those, the version of Hindi that was current in the Delhi region, finally became the vernacular and the vehicular of the Mughal realm.

The Russians had conquered a good part of what is known today as 'Russia' in the seventeenth century and were laying the groundwork for an expansion towards the east that would continue for centuries, until all of Siberia, and most of central Asia, had been conquered. Throughout this vast area, Russian functioned as the supercentral language, taught increasingly in the schools as the first 'foreign' language.

The Modern Era was very much the period of the expansion and imposition of European vernaculars across the globe. Portuguese, Spanish and English between them almost entirely covered the western hemisphere; English became the dominant language of the Australian continent; French prevailed side by side with Arabic in northern Africa. Russian came to dominate all of northern Asia. All these new territories were settled in large numbers by colonists from the European 'mother' country.

In Sub-Saharan Africa and in most of South and South-East Asia, English, French and Portuguese had spread with colonial conquest, and functioned increasingly as media of administration, trade, higher education and long-distance communication, but they never eliminated the indigenous languages. One reason was that Europeans migrated to those lands in much smaller numbers. But in almost all the former colonies, the European language continued to serve key functions, even after the departure of the colonizers, and still do after half a century of independence. The end of this worldwide presence of European languages is not yet in sight. And one of these vernaculars, English, is still increasing its hypercentral prominence almost everywhere on the globe.

A map of the present global language system looks quite like a composite of political maps from the eighteenth, nineteenth and twentieth centuries. It shows how much language constellations are



determined by political events, but also how they often survive long after this political base has disappeared. Thus, Spanish and Portuguese came to the southern part of the western hemisphere as colonial languages, and so did English and French in North America. And although almost the entire continent became independent of the European mother countries between the end of the eighteenth and the middle of the nineteenth century, the languages of the former colonizers, English, Spanish and Portuguese, still prevail there.

Equally, by the end of the nineteenth century almost all of Africa had been divided between the West European powers. After World War I, Germany was divested of its African possessions. Today, three or four decades after independence, the former colonial languages, English, French and Portuguese, still function throughout Africa; the linguistic map does not look very different from the political map of, say, 1920.

Nor has the map of European languages changed much when compared to the political map of, say, a century ago. The central languages of many European countries coincide with the state borders (although a more detailed map would reveal incongruities in almost every country). But this apparent stability hides great upheavals that occurred in the twentieth century.<sup>16</sup> German spread with the Nazi conquerors, and receded as soon as they were defeated. Russian was imposed on the countries of Central and Eastern Europe and swiftly abolished after the Transition. Likewise, Japan expanded in the past century from Manchuria to New Guinea, and Japanese followed the paths of conquest, only to disappear almost entirely after the defeat of Japan.

In Asia also, the military conquests of preceding centuries very much determine the present distribution of languages. The most notable exception is Indonesia, where after Independence Dutch disappeared completely, while Malay spread all over the archipelago. But the language constellations of China and India, the Philippines, Thailand and Malaysia still coincide rather closely with the political patterns of a century ago. In Vietnam, Laos and Cambodia, French had to make room for English in the wake of more recent wars.

### **1.3 Supercentral constellations in the present language system**

The supercentral languages mostly spread in two ways: over land and over sea. German, Russian, Arabic, Hindi, Chinese and Japanese each cover a large contiguous area, more or less coterminous with the

## 12 Introduction: the global language system

territory once or still controlled by a major empire: these are the 'land-bound' languages that spread with marching empires.<sup>17</sup> English, French, Portuguese and Spanish, on the contrary, spread with the conquest of territories overseas. Swahili and Malay initially functioned as regional vehicular *lingua francas* and became national languages after Independence.

The Chinese language constellation covers mainland China and Taiwan. The vast majority of continental Chinese are by now able to read the most current Chinese characters and speak or at least understand *putonghua*, a standardized version of the Mandarin variety from Beijing that has been taught in the schools all over China since 1948. Mandarin differs considerably from the other varieties of Han Chinese that are spoken mostly south of the Yangtse. Of the 1.2 billion mainland Chinese, 96 per cent use one of the varieties of Han, 93 per cent are native Han speakers and 63 per cent are mother-tongue speakers of its most important variety, Mandarin.<sup>18</sup> The largest non-Han language spoken in China is Zhuang, with 13 million speakers.<sup>19</sup> Abroad, varieties of Han Chinese continue to be used by many millions of emigrants, mostly at home and on social occasions, and in newspapers for the local immigrant community. The supercentral position of Mandarin *putonghua* in mainland China is illustrated by the fact that the lion's share of 'minority' members who did acquire a second language (18 million) learned a Han variety (almost all of them Mandarin) and only a tiny fraction (0.9 million) learned other 'minority languages', whereas very few Han speakers (1.2 million) learned one of the peripheral languages. Yet, for its huge numbers and ancient tradition, Han Chinese in its several versions plays a rather minor role in communication beyond China's borders, except for the many millions of emigrants.

Unlike China, India was occupied by a Western colonial power for almost a century and a half. Once again, the language map reproduces features of earlier political maps: English, more than half a century after independence, is still very much present as a second language, in stiff competition with Hindi. The presence of the Dravidian languages in the south very much complicates the Indian constellation. This is the subject of a separate chapter (chapter 4).

German, Russian and Japanese nowadays are barely supercentral languages, confined as they are to the remaining state territories. In the wake of centuries of conquest and rule each of these languages became established as the official language, the vernacular and the predominant mother tongue in a vast and contiguous area. In the course of the twentieth century, as a result of military expansion all three of them spread far beyond their former limits, and receded again once the defeated conquerors had to surrender their territorial gains.

Since the early 1990s, the Russian language constellation has been rapidly coming apart. The Central and East European satellite states have regained their full autonomy and quickly did away with Russian as the first foreign language, turning instead to English or German. The former autonomous republics of the Soviet Union became 'independent states' and likewise discarded obligatory Russian, reverting to their regional languages, from Latvian, Lithuanian and Estonian on the Baltic Sea, to Armenian and Georgian on the Kaspic shores, or Kazakh, Kirghiz, Turkmen and Uzbek, and Tadjik in Central Asia. Siberia retains Russian, although indigenous peripheral languages remain current locally.

But even in the core area of the Soviet empire, the speakers of Belorussian and Ukrainian increasingly came to see their languages as essentially different from Russian and abandoned the idea of complete mutual intelligibility that had prevailed without much controversy during the Soviet era.<sup>20</sup> As a result, Russian began to lose many of its supercentral functions in the former Soviet empire. English took over these linking tasks almost everywhere, Turkish grew in importance in the Central Asian republics and German plays a modest linking role in Central and Eastern Europe.

Just as Russian was discarded by its European vassal states immediately after the transition, some forty years earlier German had been uprooted and abolished in the same region, while in East and South East Asia, Japanese was 'forgotten' just as quickly.<sup>21</sup> On the other hand, in the respective core territories, where political rule by the central state had already been consolidated during the nineteenth century, these languages have conserved all their functions and are still spoken by practically all citizens, almost without competition from smaller indigenous languages. Apparently, in the twentieth century imperial conquest did not pay in terms of enduring language expansion. Maybe national languages had already taken hold too deeply in the newly conquered territories to be eradicated definitively; possibly also the foreign occupation did not last long enough to establish the conqueror's language for good.

In Europe, German has the most numerous native speakers covering an area that comprises present-day Germany, Austria, the northern part of Switzerland (where it coexists with the *Schweizerdeutsch* variety) and Luxembourg (where the local variety is *Letzeburgesch*). Moreover, as Alsatian it is spoken by most inhabitants of Alsace-Lorraine, an area that became German in 1871, only to revert to France in 1919. By the end of the nineteenth century, German had become one of the most important commercial and scientific languages, almost on a par with English and French.<sup>22</sup> Had Germany