Chapter 6
The Organization of Urban Systems

Anne Bretagnolle, Denise Pumain and Céline Vacchiani-Marcuzzo

6.1 Introduction

Cities[1] are a major form of the material, social, and symbolic organization of societies. They are persistent and adaptive structures that fulfill a variety of social functionalities: habitat, production, services, political control over people and territories, as well as technical and symbolic mediation between nature and culture, groups and individuals. For a long time, and especially since the first industrial revolution, they have been both the spaces where most innovations occur and spaces for which most innovations are designed (Chapter 8). Location in space and time has to be considered as a very important feature for these systems, since societal evolution is much more rapid than its biological counterpart is. Being adaptive, cities alter through a variety of intentional actions that may produce non-intentional persistent features: urban structures are produced partly through design and planning, partly through self-organization. This chapter is an analysis of a conception of urban systems as complex systems. We develop a theory of the organization of cities and systems of cities as multilevel networks, including two main observable levels, where specific emergent properties can be observed. Compared to the earlier seminal conception by Berry (1964) of "cities as systems within systems of cities," we emphasize a multilevel organization perspective and specific dynamic features that produce this structure. We also add an evolutionary perspective, which is based on observations about the way cities co-evolve within urban systems, through a variety of social interactions. Empirical evidence of the corresponding patterns and processes is provided for three main styles of urbanization, exemplified in four different parts of the world.

A. Bretagnolle (✉)
Université Paris I Panthéon Sorbonne, UFR de Géographie, Paris, France

[1] The word "city" is used here to designate generic urban entity defined as a geographically and functionally consistent urbanized area, whatever the administrative or political boundaries.

D. Lane et al. (eds.), *Complexity Perspectives in Innovation and Social Change*, Methodos Series 7, DOI 10.1007/978-1-4020-9663-1_7,

6.2 Cities and Systems of Cities as Different Levels of Social Organization

The organization of urban systems can be described on three main levels (Pumain, 2006): the micro level represents elementary units (individual people, firms, institutions) that are living together in a city, the meso level corresponds to the city itself (defined as a consistent geographical entity), and the macro level is the system of cities, made up of a large number of towns and cities which interact under unified control (in a national political territory or a global economic network). This organization is shaped by interactions operating on different spatial and temporal scales of observation.

6.2.1 The City: A Collective Evolving Entity

In a very general way, a city can be seen as a collective entity whose specific properties, although mostly produced by intentional agents at the individual level, cannot be simply explained or predicted from these intentions, nor derived by summing the characteristics of its inhabitants. Concepts such as urban function, centrality, or morphology were invented by urban scientists to understand cities' emergent properties. Produced through reiterated interactions between individuals, these are only defined at the aggregate level of the city and cannot be estimated from the mere summing of the attributes of the individuals that compose the urban entity. The collective character of a city also emerges over historical time because the evolution of cities is extremely coherent. Each of them has a specific history that defines its identity. Even if they participate in the same historical events and trends, cities each follow an original trajectory, which is influenced mainly by some of their political options or by their successes and failures in the process of their socio-economic co-evolution. The concept of a city as an urban entity is a relatively autonomous and persistent system of locally dense and frequent daily interactions. It is rooted (as a material and a symbolic entity) in the collective process of specification and identification. This is constrained by two distinct trends, bottom up by way of internal interactions (among urban citizens, firms, and institutions) and top down by way of external interactions (among competing cities), that form interlocking networks.

6.2.1.1 The "One-Hour" Traveling Time Constrains the City's Development

For a relevant analysis of the dynamics of this complex system, we have to translate this concept into proper measurements of city size that are comparable over space and through historical times. Measures of city size are difficult because cities are expanding not only inside the boundaries of a fixed territory, but over these limits. We have then to consider an entity according to a common reference in time-space. In this framework, we can visualize a city as the envelope for the daily activities

of its inhabitants and the buildings hosting them. The concept of a city that we use is not restricted to elementary political entities governed by a municipality or an administrative circumscription with fixed boundaries. Indeed, in the course of time, it is frequently observed that urban growth crosses these boundaries and spills over into neighboring circumscriptions, sometimes including other nearby cities and towns in "conurbations," as noted by Patrick Geddes as early as 1905. In order to correctly identify a city as a consistent geographical entity, and considering that its spatial expansion as well as its *in situ* development are part of its growth, we define it as a "daily urban system" (this concept was the basis for the definition of Standard Metropolitan Areas in the US in the fifties, and its name has been chosen because it allows for frequent social interactions across the boundary, which usually take place within one day). Recent surveys in European, American, or African cities demonstrate that during a working day, an urban dweller typically needs to connect to three or four different places of activity, and devotes on average about one hour each day to traveling for these purposes (Crozet & Joly, 2004). It seems that this travel time, representing about 15% of a 24 hour working day, has remained rather stable over centuries (in the literature on transport, this is known as "Zahavi's law").

While the maximum spatial development of cities seems always to have been constrained by this typical one-hour time-budget, improvements in the speed of the means of transportation available have enabled the average commuting distance between places of work or urban services and places of residence, to be multiplied by ten in the last two centuries. In London, for instance, the city maps of the 17th century (Hollar, 1667) represent an urban agglomeration whose radius is 3 to 4 km. This radius has evolved (6 to 8 km in 1830, 20 km in 1900) according to the innovations in transportation technologies, for which London was a pioneer in Europe (1829, horse omnibus; 1836, steam railway; 1863, steam metropolitan; 1905, electric metropolitan). Today, a remote sensing image of London representing built-up areas (*CORINE Landcover*, EEA, 2000) shows a first perimeter of compact constructions over a circle of 20–30 km radius, with expanding branching developments reaching out to zones located nearly 70–80 km from the centre (Fig. 6.1). The historical adaptation of an urban entity, whose expansion was driven by economic success and social pressure towards the maximum possible spatial extension of the time (and such despite policies aiming at "the containment of urban England" – Hall, Gracey, & Drewey, 1973), reveals the general limitation of a one-hour traveling time, which constrained its development. A behavioral parameter defined for spatial interactions at the individual level is reflected in the organization of the urban entity at a higher level.

6.2.2 The System of Co-Evolving Cities

Cities never actually develop as isolated entities; they are always engaged in many types of interactions with other cities. They are linked through a variety of social networks, among which there are not only visible infrastructures, such as roads,

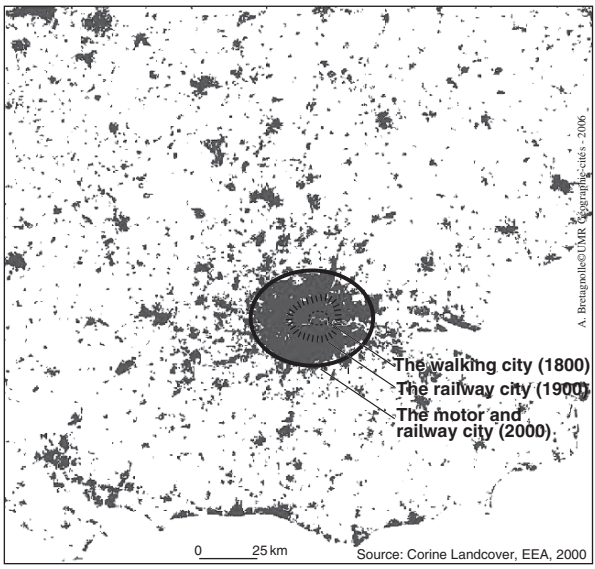


Fig. 6.1 The agglomeration of London, an evolving spatial entity delimited by a "one-hour" constraint of traveling time

railways or airlines, and the kind of exchanges usually accounted for, such as population migrations or trade of goods, but also more "invisible" networks: capital and information flows, for example. (The latter are more involved particularly in the process of innovation diffusion.) The resulting *systems of cities*, although mostly self-organized, ensure the same social functionality: to control territories or networks. The nature of the control has evolved through historical times, mainly from political to economic. The political control involved, which at first may have coincided with the scope of the market area of one city, spread to kingdoms or empires and then to national territories associating several cities. Economic control, expressed first by local entrepreneurs, broadened to national then multinational firms. Individual cities, which in the past were the main actors in their own development – seeking to overcome the limitations of their immediate environment by building networks to exploit distant resources – have become instruments of control of wider territories and networks, a control that is now assumed by national or supranational actors using the networks that cities have built through their interactions. *What we call systems of cities are evolutionary objects that may include subsets of cities connected by long-distance networks or cities belonging to unified political territories*. The general trend is a historical increase in the number of cities that are integrated, through ever more intense and more frequent interactions, although political events or economic crises may reduce their number and impact locally for some time. The precise identification of systems of cities is very difficult, due to the changing nature of the interactions that need to be considered, and the fluctuations in their spatial extension.

6.2.2.1 An Integrated System of Co-Evolving Cities Through Multilevel Spatial Interactions

The direct interactions, which could be named *first order interactions*, produce strong interdependencies in the evolution of cities and give rise to a macro-organizational level: the system of cities. This organization of "city systems" that develop specific emergent properties was noticed a long time ago (the first mention of a "system of cities" can be found in the writings of a French Saint-Simonian engineer, Jean Reynaud, as early as 1841) and now is part of a geographical theory on urban systems (Berry, 1964; Pred, 1977; Pumain, 1997). Systems of cities always show a differentiation of city sizes according to several orders of magnitude (today, from a few thousands to tens of millions of inhabitants), which follow a very regular statistical distribution, which is lognormal or of the Pareto-Zipf type. This hierarchy of sizes also corresponds to a hierarchy of urban functions and to more or less regular settlement patterns. These regularities were summarized, for a while, by central place theory (Christaller, 1933; Berry, 1967), which became embedded into a more general evolutionary theory of urban systems. Systems of cities also are characterized by their functional diversity and the co-evolution of their socio-economic profiles (see Chapter 8), as well as by distributed growth that can be summarized in first approximation by a Gibrat model (1931).

These first order exchanges, although very often reciprocal, are not fully symmetric. Asymmetries in the interaction flows between cities, when reiterated, produce a diversity of quantitative and qualitative differences in terms of city size, economic specialization, social composition, cultural features and urban landscapes. These differences lead to *second order interactions*, which are effects of *selection*, constraining a city's development according to its rank in the hierarchy of city sizes, its economic specialization, and its "image" in the individual and collective representations of agents. Second order interactions can be observed indirectly by comparing the evolution of qualitative urban features and measuring the evolution of relative sizes in large sets of interacting cities (Pumain, 2006). More specifically, some types of networking are, at least momentarily, restricted to small samples of very large or very specialized cities.

In order to delineate a system of cities geographically, according to a "classic" definition of a "system," we have to identify a set of interacting urban entities, including cities that have more intense or more frequent interactions with cities inside the system than with cities outside. This is often implicit when systems of cities are considered within the boundaries of national territories. National boundaries delimit a community of political and social rules as well as cultural features that reduce external interactions. However, even in the frame of this static view, cities can exceed these limitations, according to their size and function. In particular, national capitals or cities that are specialized in international activities have a broader range of exchanges than smaller or less specific towns. This wider opening of some cities can be observed in their first order interactions, but, most of the time, it is better illustrated by the interdependencies in their evolution that are produced by second order interactions. Therefore, the envelope of systems of cities should not

be seen as a strictly delimited boundary, but rather as a membrane that is more or less permeable according to the size and function of the cities inside the system. Because the range of interactions usually is strongly influenced by city size, the ability to develop outside interactions and to co-evolve with other urban systems is closely related to the hierarchical differentiation inside the urban system. But there often are, in a system, a few smaller cities that specialize in international activities, for example finance or tourism.

6.2.2.2 Global Cities Since the Middle Ages

At all times, a few outstanding cities have dominated the exchanges in the interconnected "world" formed by systems of cities, crossing the political boundaries. Considering an integrated "system of cities," defined as a coherent set of cities already engaged in commercial and political competition, it is possible to follow the evolution of the weight of a given city within the system over several centuries and interpret its trajectory in relation to the greater or lesser success of the city in having a share in successive innovations, whether or not these lead to specialization. A few cities, for long periods of time, exhibit coefficients of allometric growth that remain systematically greater than one, when compared to the evolution of the other cities. This dynamic behavior is explained by the fact that these cities have much larger interaction networks than those of the other cities. Figure 6.2 presents three examples of such "global cities" that bypass the European urban field at three historical periods. We have computed the weights of Venice, Lisbon and London within the European total urban population at each date and represented the trajectory of this relative weight in phase space (ordinate: the weight of the city in the system for a given year; abscissa the weight[2] for the city the preceding year). The trajectories are illustrative of the successive success of Venice, in Mediterranean and Hanseatic trade during Middle Ages, then Lisboa in Atlantic maritime trade with the New World in the 15th century, and then London as the center of a colonial Empire and the industrial revolution at the end of the 18th century.

6.2.3 A Crucial Point in Urban Ontology: The Consistent Delimitation of Urban and Territorial Entities

The dynamics of urban systems have been interpreted in many different or even contradictory ways because the underlying observations of the differential growth of cities are based on different measures from one author to another, varying with the chosen definition of the city, the delimitation of the territory involved, and the samples of cities selected within these territories. Our theoretical conception of systems of cities makes it possible to gain a better understanding of the dynamics

[2] Europe can be considered approximately as forming a system of cities since the reopening of long distance trade in 12th or 13th century (cf. Pirenne 1927)

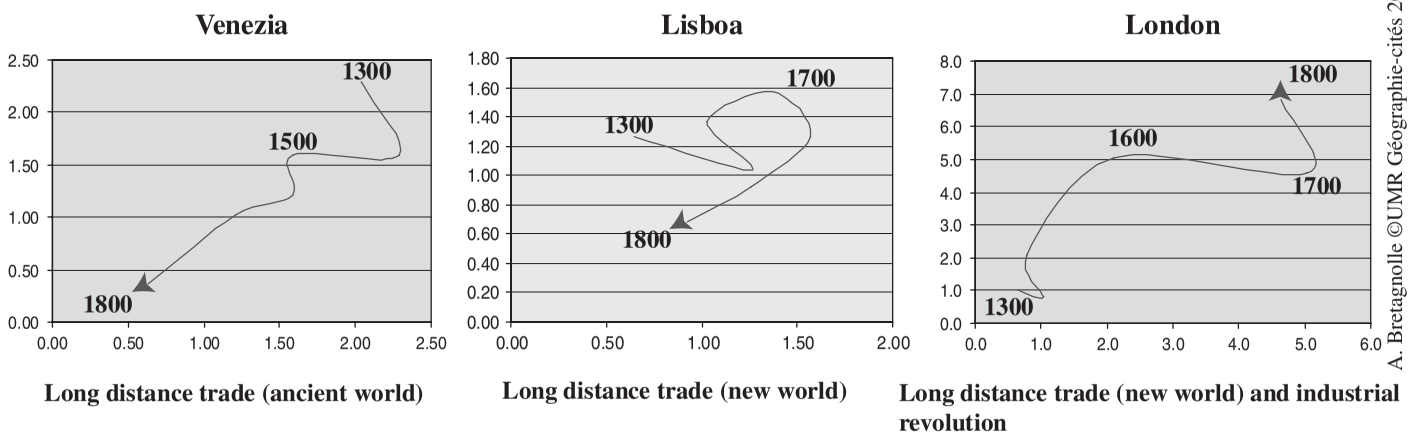


Fig. 6.2 Trajectories of individual cities in the European urban system. The ordinate is the weight of the city in the system for a given year; the abscissa is the weight for the city the preceding year

involved by relating them to their social functionalities and to the artifacts that make these functionalities feasible. As we have seen, at the scale of the city, the main functionality is the coordination of day-to-day activities, while at the scale of a city system the issue is the control of a territory or multiple networks, mainly by means of political power at the outset, and economic power subsequently. We have defined comparable databases within the frame of national or continental boundaries and we shall return to the consequences of this choice (Chapter 12).

Comparing city systems over time and in different areas of the world requires careful preparation of databases. These are never directly exploitable as they are collected and produced by statistical institutes. Official databases have their limitations, first, in terms of the indicators available. Since successive industrial revolutions have widened the gaps in standard of living, the best measure of the success of a city might be its economic power. Most of the time, we must be content with data on numbers of inhabitants of cities, the only indicator that can be mobilized on scales of time and space of this magnitude. However, even regarding population, harmonization is required so that entities termed "cities" are comparable from one country to another and through history (Pumain, Saint-Julien, Cattan, & Rozenblat, 1991; Bretagnolle et al., 2007).

Today census bodies use two main approaches to define the city. The first outlines the *urban agglomeration*, which is formed by the continuity of built-up area and by minimum population or density threshold values; the second outlines the *urban area*, which is much wider than the agglomeration since it also includes peri-urban rings that send part of their working population on a daily basis to the functional pole of the agglomeration. To make urban populations comparable from one country to another and at different times does not necessarily mean adopting identical criteria. For instance, it is considered today that in Europe the minimum functions that can be associated with a city characterize aggregates of more than 10,000 inhabitants, while in South Africa, more recently urbanized and involving a smaller surface area, it is more reasonable to lower this threshold to 5,000 inhabitants. Likewise, a strong feature of the USA is found in the extremely wide spatial encroachment of its cities, arising from the tendency of the population to undertake long-range daily commuting.[3]

When defined for highly integrated territories with intense, frequent interactions (for instance, with exceptions, most national states today), city systems present a certain number of characteristic emergent properties. Among the main "evolutionary laws," we insist on four major trends: historical path dependence, competitive expansion and distributed growth, and reinforced urban hierarchy.

---

[3] The different urban definitions that we have adopted in our comparative study are as follows: in Europe and India, agglomerations of more than 10,000 inhabitants; in South Africa, functional agglomerations of more than 5,000 inhabitants (including the white city and the non-white townships that are economically linked to the city by home-to-work commuting); in the USA, the populations of cities and towns up to 1940, then the *Standard Metropolitan Areas* (known today as the *Metropolitan Statistical Areas*) and the *Micropolitan Statistical Areas* up to 2000 (for a more detailed description of these databases, see Bretagnolle Pumain, & Vacchiani-Marcuzzo , 2007; Bretagnolle, Giraud, & Mathian, 2007).

6.3 Historical Path Dependence

Despite the multiplicity and apparent diversity of national urban systems in the world, three major styles are recognizable in their hierarchical and spatial organization. These styles are differentiated because of different historical trajectories (Table 6.1). Their properties vary according to their period of emergence (technological conditions during the urban transition determine space-filling parameters) and according to any major exogenous impacts (such as colonization).

Table 6.1 Three major styles of urban systems in the world

Urban System	Typical world region	Selected example
Long-standing urbanization, slow and regular evolution	European urban system National urban systems in Europe	European urban system
Long-standing urbanization, major exogenous impact (colonialism)	Asia Black Africa	India
New countries and waves of urban creations	U.S.A. Australia South Africa	U.S.A, South Africa

6.3.1 Diversity in Morphogenesis

In countries with long-standing and continuous settlement processes, cities emerge more or less simultaneously all over the territory, and the city systems are characterized at once by the long-standing nature of their urbanization and by the regularity of their development over time. European countries are a good illustration of this. From Antiquity, long-range inter-city exchange networks became established, and several metropolises had more than 100,000 inhabitants (Ancient Athens) or even one million (Ancient Rome). After periods of relative stagnancy in the Early and Late Middle Ages, European cities were once again undergoing vigorous growth and the reactivation of exchange networks in the Early Modern period, following a "national" or "international" logic (at least for the largest or most specialized). Figure 6.3 shows a marked persistence of the distribution over geographical space of towns and cities since AD 1200 in the western part of Europe and since AD 1500 in eastern Germany and central Europe. The urban growth that is characteristic of the industrial revolution (the number of towns and cities doubled) did not significantly alter spatial distribution, even if there was a strong densification effect along the coasts and in the major mining and metalworking basins. A similar stability also characterizes the top of the urban hierarchy. The majority of the large European metropolises already dominated national networks since the end of the Middle Ages, despite the fact that the major economic centers holding sway over the Mediterranean in the Middle Ages gradually gave way to the trading cities on the

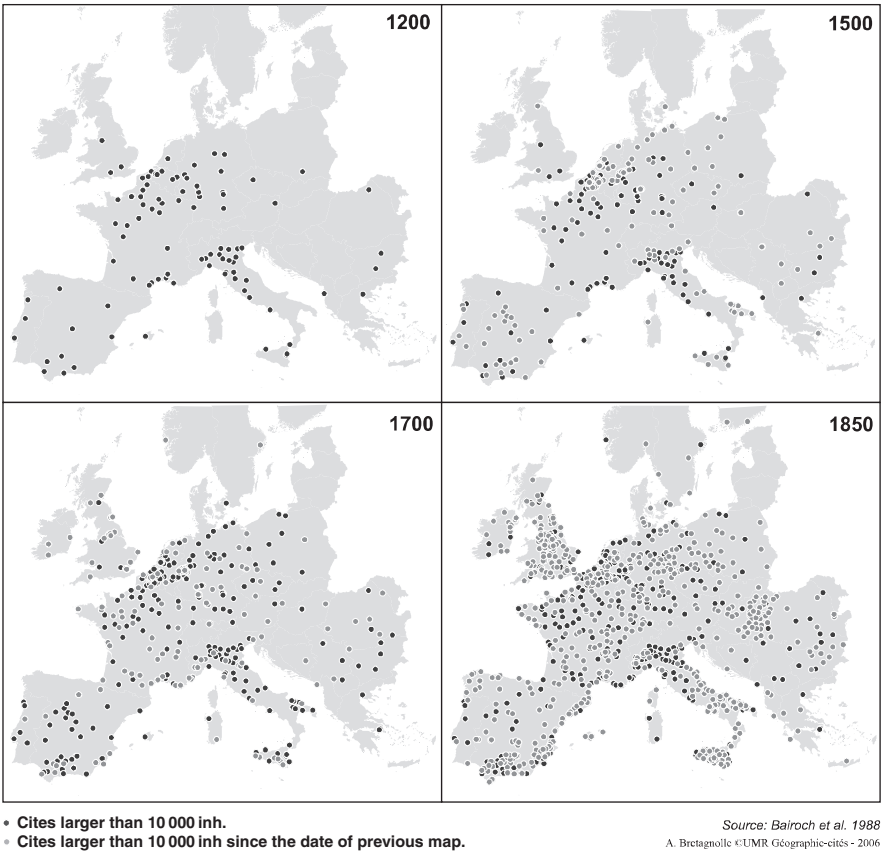


Fig. 6.3 First urban system style – continuity (Europe)

Atlantic coastline, until the industrial revolution sent Britain to the front (Braudel, 1967; de Vries, 1984).

A second style of urban system is represented by countries where urbanization is long-standing but where disruptions have been experienced, like Asia or Black Africa. The example of India is illustrated in Fig. 6.4. This country underwent a major phase of urbanization in Antiquity and a second phase in the Middle Ages with Muslim expansion. At these times, the country was divided into some twenty kingdoms, and the main cities in existence (Agra, Delhi) were located inland rather than on the coasts, because the main activities were related to domestic trade and territorial control. While the first Portuguese, Dutch, French, and later British trading posts were set up from the 16th century (Bombay in 1532, Madras in 1639, Calcutta in 1690), they did not have any decisive impact on the urban patterns in India before the 19th century. However, when India officially came under the British crown in 1847, the administrative and economic orientations changed radically and altered the distribution of Indian cities in a durable manner. From then on, the main

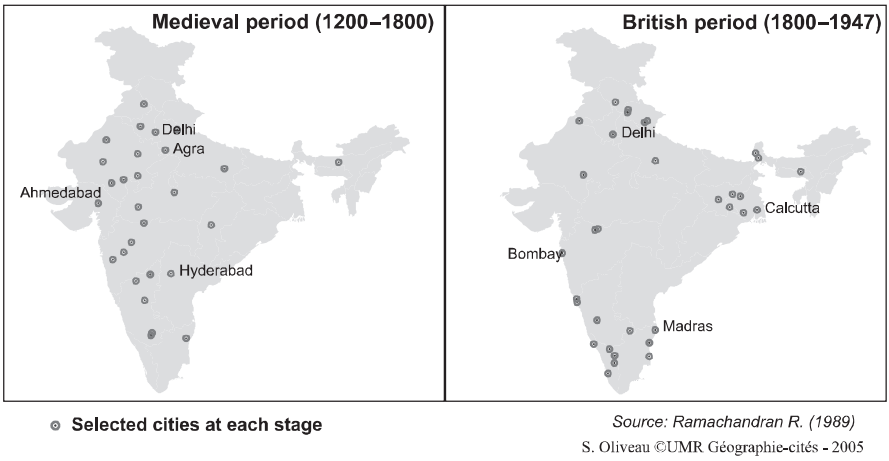


Fig. 6.4 Second urban system style – disruptions (India)

cities were created or developed along the coasts and the main rivers to facilitate exchanges with Britain.

The third type of urban system is that of "New World" countries, where towns and cities were imported by settlers and spread in successive waves of penetration, either driven or accompanied by canals and railways. In the USA, the occupation of space by cities took place in relation to waves of settlement moving in from the coastlines. A first frontier moved west, reaching the Mississippi in the 1850s, the Rockies in the 1870s, and the western coastline in the 1890s (Fig. 6.5). This frontier started from the first communities established in the 17th century along the east coast (Manhattan in 1614, Philadelphia in 1654) which developed slowly up to the start of the 19th century. In the 1790 census, only five cities were larger than 10,000 inhabitants, one of which was New York with only 33,000 inhabitants at this date (while London had reached 948,000 inhabitants). The exponential urban development that followed independence is out of all proportion with the growth that characterized Europe over the same period. In less than one century, New York became the second largest city in the world, with more than three million inhabitants in 1900. A second settlement front opened up in the 18th century on the Pacific coastline when the Spanish moved up from Mexico. The first towns of more than 5,000 inhabitants were San Francisco and Sacramento in 1850.

In South Africa, the process is slightly different: between 1652, when Cape Town was founded, and the start of the 20th century, several waves of Dutch and later British settlements spread along the coast towards the east, within a mainly agro-pastoral economy. The decisive impetus to the creation of urban settlement was the discovery of diamond resources (Kimberly in 1867) and gold (Johannesburg in 1883) in the central province of Gauteng, leading to a major shift of the center of gravity from the coastline towards the interior. From this period onwards, the cities grew more through internal migrations than through international immigration. The formation of the South African state in 1910 and the development of exchanges

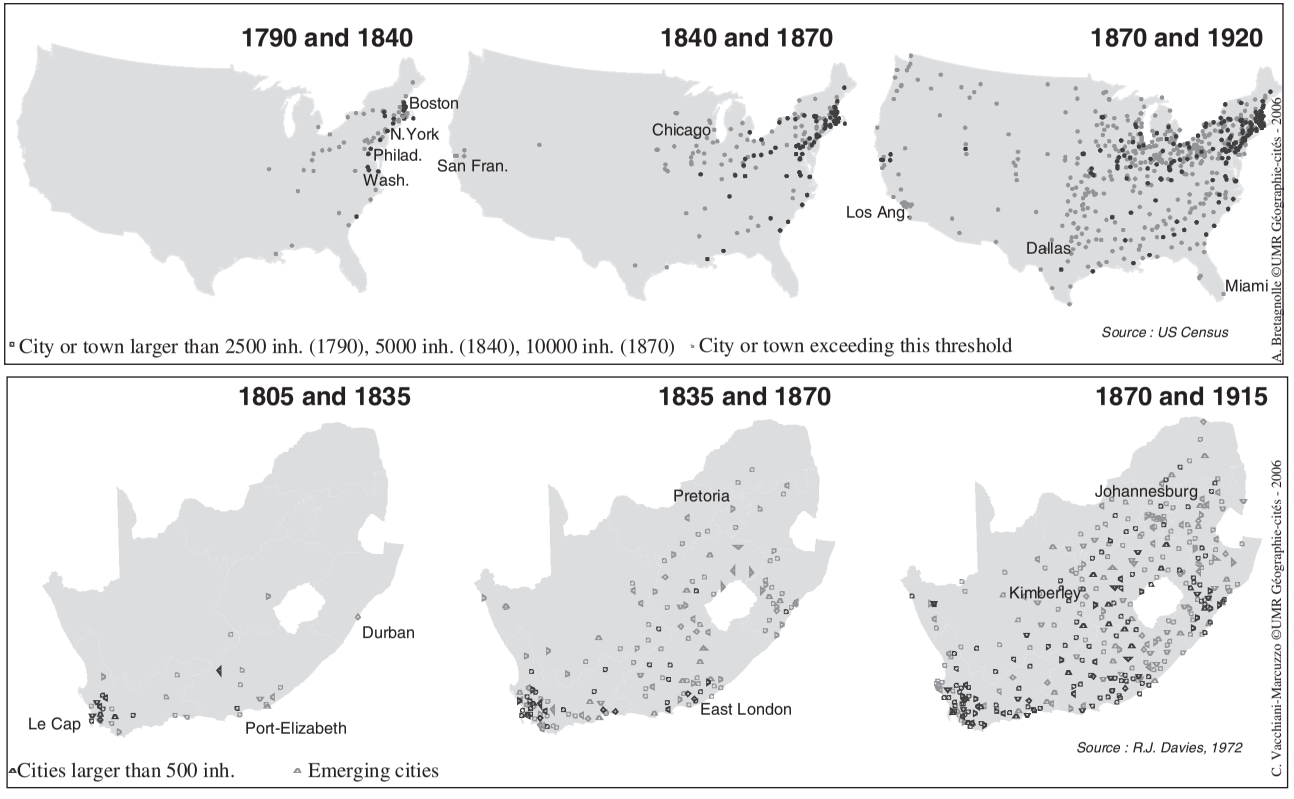


Fig. 6.5 Third urban system style – waves of urban creations (U.S.A., South Africa)

within an economy that had become industrial favored the emergence of a genuine city system in the 1950s (Fig. 6.5).

These three major types of morphogenesis do of course considerably oversimplify the diversity of situations observed worldwide. It would be interesting, in particular, to analyze in a more precise manner the particular place of the city systems in Latin America, North Africa, and the Middle East, as each of these areas had its own specific involvement in colonial processes. However, these three main types do make it possible to identify forms that still today most strongly differentiate hierarchical and spatial configurations in city systems across the world.

6.3.2 Typical Emergent Properties of the Three Major Styles of Urban System

The three types of system described above are not differentiated by their histories alone, but also by certain features of their hierarchical and spatial configurations that are still perceptible today.

A first feature that differentiates the three styles of systems is their morphology, i.e. the pattern of their occupation of space as measured by the density of the cities and the degree of hierarchization (Table 6.2). When cities were established at a time when the means of transportation were very slow, via the spontaneous emergence of agricultural markets or by the establishment of relay posts on the main communication routes, they were very close to one another. For example, in Europe and India alike, the average distance between two towns or cities is around 15 km. The newer countries show a wider spacing between towns and cities, marked concentration in the largest cities, and an urban hierarchy that is characterized by greater variation in size. As a convenient measure of the degree of hierarchical inequality, the slope of the straight line adjusting the rank-size distribution of cities and towns can be used. This parameter value is under one for Europe and India, while it is clearly above

Table 6.2 Typical parameters of the three major styles of urban systems

	Average distance to nearest neighbor in km (2000)	Inequalities in city sizes (2000)	Average annual growth rate (%) during the urban transition (19th or 20th century)	Macrocephaly index (2000)
Long standing urbanization: Europe	15	0.94	1–2, distributed	2 to 3
Long standing and external shock: India	16	0.99	2–3,dual	4 to 8 (regional indices)
Recent systems:				
(a) U.S.A	36	1.20	3–4, higher on frontier	1.5
(b) South Africa	32	1.19	3–4, higher on frontier	2

NB: the degree of inequality is measured by the absolute value of the adjustment slope in the rank-size graphs.

one for the USA and South Africa. Indeed, in the latter two, countries, towns and cities developed in a pioneer logic, i.e. aiming to occupy the widest space possible, even if this was in an extensive manner, and they developed with faster and more efficient means of transportation (especially the railways). This results in systems where towns and cities are less numerous, less dependent upon the initial agricultural settlement, more widely spaced, sizes more contrasted, and where the largest cities can reach sizes greater than those observed in the Old World. The automobile later enabled a small number of urban centers to have an influence over very distant outer rings: for instance in 2000, the mean radius of functional areas is 46 km for the American metropolitan statistical areas (MSAs), as compared with 13 km for French urban areas.[4]

A second feature that differentiates the three styles of urban systems concerns the regimes of city growth. By analogy with demographic transition, what is known as urban transition (Zelinski, 1971) is a period of massive urban growth in the course of which, settlements, thus far made up of villages and fairly homogeneous and scattered, became much more heterogeneous by concentration around urban centers. This transition, which began at the time of the great industrial revolution, is complete in countries where it took place early (Europe, the USA), but it is still underway in countries in Asia (China, India) and is just starting in certain African countries. While the main influx of new urban settlers was above all from rural areas in the 19th century urban transition, it also was fed by demographic growth specific to the towns themselves in the 20th century transition. This increase in population (Table 6.2) is spread across towns and cities according to three main patterns:

1. In countries where urbanization is long-standing, and where it developed in the continuous mode (European type), urban growth was *distributed*, i.e. spread across all parts of the territory in a manner that is proportional to the size of the towns and cities, even if a slightly higher relative growth rate is seen in the large cities. Average growth rates were low throughout the 19th century, at around 1 to 2% per year (1.7 for London and 1.3 for Paris), except in certain localities in industrial basins (Bradford 4%, Valenciennes 3%).
2. In new countries, urban growth followed a "*wave*" settlement pattern, with markedly higher growth rates in the newly created towns and cities. In the USA, for instance, Chicago saw an annual growth rate of 12% per year between 1850 and 1870, and urban growth was more intense on average during the urban transition than in older countries (around 3 to 4% a year, mainly because of immigration from abroad).

[4] Caution is required in comparing these figures: indeed the commuting range threshold taken into account to define functional areas is not the same in the USA (15%) and France (40%), and it refers to very different dimensional grids (a ratio of 100:1 on average between the surface areas of *counties* and that of the French *communes)*. But the choices made by the census bodies involved is an actual reflection of the specific nature of urban settlement patterns in each country.

3. In systems that were reorganized following the impact of colonialism (some Asian and African countries), urban growth was *dual* i.e. fairly small and late endogenous growth of markets, administrative centers, and local artisan activity is superimposed on very marked concentration in the large cities, often the capital cities, or acting as an interface with the capital. This macrocephaly (highest ratio between the populations of two cities of consecutive rank) is typical: while the ratio is 1:2 or 1:3 on average in most European countries, it is 1:6 in Ivory Coast and 1:7 in Mali. In India, several regional capitals were driven by colonialism (Bombay, Calcutta, Delhi, with 16 million, 13 million and 12.5 million inhabitants respectively in 2001), and each is far ahead of the other cities in their respective regions (the populations of these cities today are four to eight times that of the largest city in their hinterland). In these developing countries, the mean growth rates during urban transition are high, often more than twice those observed in "older" countries, especially because that can coincide with demographic transition and its reduced urban mortality (Table 6.2).

6.4 Competitive Expansion and Distributed Growth

It could be thought that such marked historical differences would result in different patterns of urban dynamics, and that each main type of city system thus identified would evolve in its own specific way. In fact, once the city system is established and integrated into a political territory, the resemblances in the way they evolve are striking. Paradoxically, it can even be said that it is because all city systems, once formed, evolve in the same manner that they continue to carry the marks of their histories. Indeed, such marks are not the indication of the "inertia" of geographical structures, but conversely of their extraordinary ability to adapt.

6.4.1 Increase in Size and Number of Cities

Whatever the level of development, and however long-standing the urbanization, city systems always have been characterized by a tendency to grow both in terms of their urban population (the maximum city size increased from one million in 1800 to three million in 1900 and 30 million in 2000) and in terms of their number (Table 6.3).

For a total surface area of around 4.5 million km^2, Europe today has more than 5,000 cities, but the period of greatest expansion was between 1700 and 1800 (where the number of cities increased threefold) and then between 1800 and 1900 (where the number increased 2.5 times, while it barely doubled in the following century). India, where the urban transition began much later, made up the difference in a spectacularly short period between 1950 and 2000, since for a total surface area of 3.2 million km^2 the total number of cities is 3300 (the same density as Europe). In the USA, the progression of the number of cities is regular through the 19th and

Table 6.3 Evolution of the number of cities from 1900 to 2000 (Europe, India, USA, South Africa)

	1900	1950	2000	Surface (million km^2)
Europe	2532	3702	5123	4.8
India	580	1095	3285	3.2
United States	382	717	934	7.8
South Africa	14	41	307	1.2

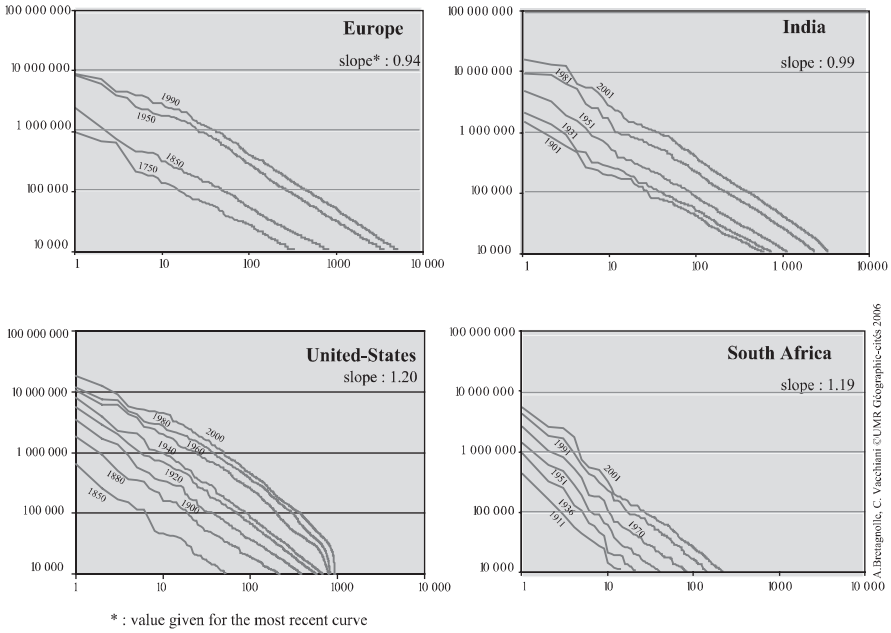
Sources and data bases: Europe: Bairoch, Batou, and Chèvre (1988), Pinol (2003); India: Census, Oliveau (2005); United States: Census of the U.S., Bretagnolle, Giraud (2006); South Africa: Davies (1972), Vacchiani-Marcuzzo (2005).

20th centuries, the decrease in the rate of growth that can be seen from 1950 being due to grouping of contiguous urban units into the SMAs. South Africa is a case apart because of its very recent urbanization, and a city system that was not mature before 1950.

The important fact is that, whatever the country, over recent centuries, urban populations do grow faster on average than the number of cities, which results in greater concentration in urban settlement. The number of cities is still increasing in developing countries, but in highly urbanized countries, the appearance of new cities with populations above a certain threshold is slowing down, and the number of urban units can even decrease on account of the fusion to form larger units as a result of urban sprawl. It is likely that in a few decades, the stabilization of the overall world population will be reflected in the stabilization of the growth of urban populations, which will not necessarily prevent their continued economic growth. It is here that the lack of reliable indicators of economic growth in cities makes itself felt. Nevertheless, recent and present-day dynamics of city systems are characterized by a historical tendency to vigorous growth. How does this growth spread over the different cities?

6.4.2 Hierarchical Differentiation and Distributed Urban Growth

The size of a city is the product of a long-term process of local accumulation. There were many fluctuations in city sizes over historical times, but once systems of cities are established in a given territory, the rapid upsurge of new cities as well as the sudden collapse of some of them become very rare, and less and less probable. As already mentioned, there are very large differences in city sizes, when measured by the total population they concentrate, their surface areas, or their economic gross product. The number of cities is in inverse geometrical proportion to the number of their inhabitants, as summarized by Zipf in his famous "rank size rule" (1941, 1949). This hierarchical differentiation within systems of cities is an emergent property that characterizes the organization of consolidated and integrated urban systems (Pumain, 2006). In Fig. 6.6, the remarkable stability of this hierarchical structure in the long term can be seen for each of the four city systems under observation, whatever the historical "style" to which they are assigned.



* : value given for the most recent curve

Fig. 6.6 Pareto-Zipf distribution of city size in the long-term period for the four categories of city systems

Not only does the shape of the distribution of city sizes remain very similar over decades, the hierarchical and spatial order given by the population sizes also evolves very little. The maps in Fig. 6.7 show this evolution (the surface area of the circles increases proportionately to the population of the cities between 1850 or 1900 and 2000 for the four cases under study). The similarities in evolution as represented in this manner are striking: on average, cities grow in a manner that is proportional to their size. This explains how initial differences tend to persist over very long periods. The reference model is that developed by Gibrat, explaining the non-symmetrical distribution of city sizes (a lognormal distribution that differs slightly from Zipf's law) by way of a random growth process in which all cities, at each time interval, possess the same probability for growth rate. This model, which has been termed "distributed growth" (Pumain, 1997) because all parts of the system grow at more or less the same rate, gives a good account of growth processes in established city systems, although with certain systematic divergences that will be examined below. This statistical model merely informs us that the causes of variation in city populations are so numerous and diverse that it is sufficient to use a straightforward random process to represent this variation, and, thereby, to explain why all systems have the same general distribution pattern for city sizes. But to go further in interpretation, and in particular to understand phenomena of divergence from this model, we also need to understand what properties of city systems explain this distributed growth.

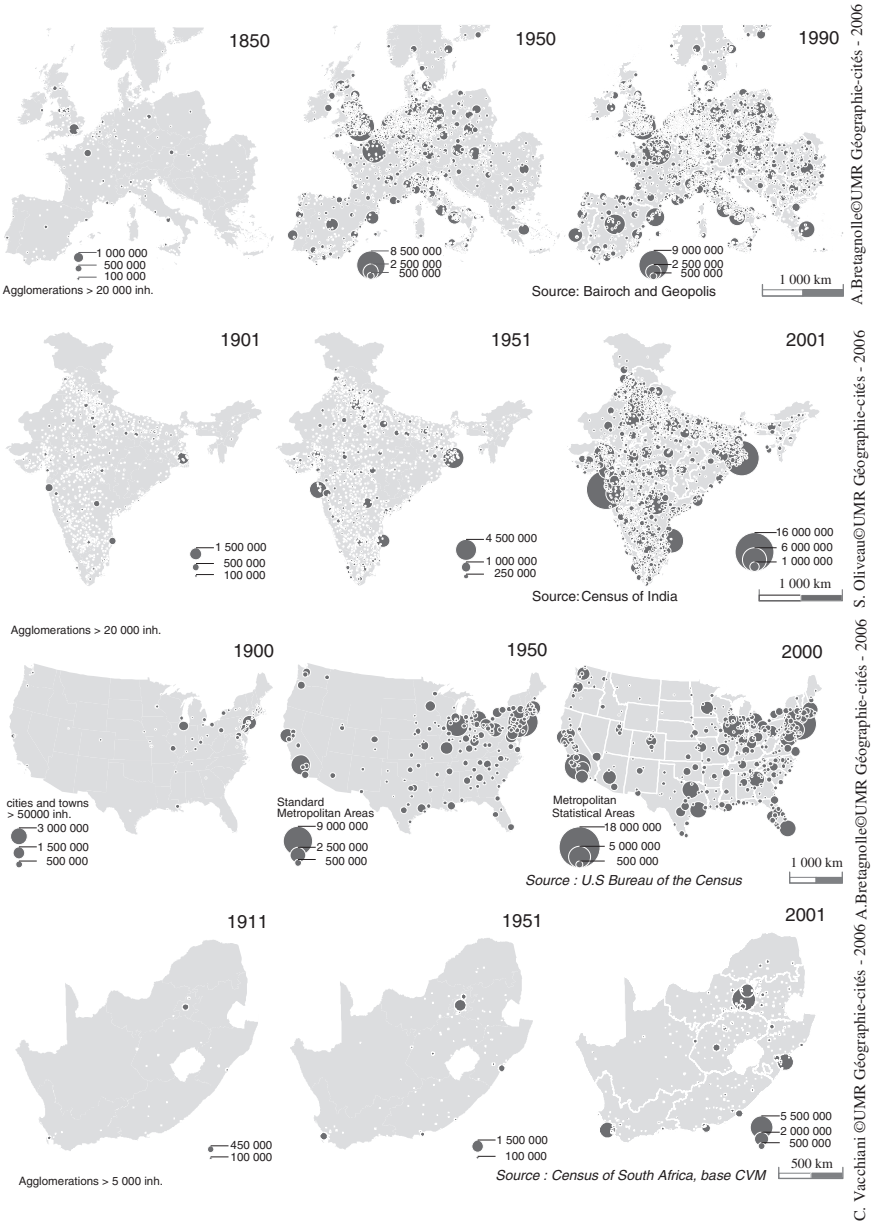


Fig. 6.7 Evolution of urban patterns in integrated city systems

Going back at least to the enlightened work by Botero in the 16th century (Pumain & Gaudin, 2002), there has been an awareness that cities are constantly *in competition* to capture resources and innovation so they can continue to make good use of what they have already acquired and maintain or increase their influence within the city systems with which they entertain relationships. This competition explains why innovation spreads quickly from one city to another and why the resulting qualitative and quantitative changes are of more or less the same magnitude, over short periods, in integrated systems. The consequence of this mode of growth is that city size hierarchy and the inequalities that arise following various types of political accidents (war, choice of capital city) or economic accidents (functional specializations, as in the industrial revolution), are maintained over periods that last a lot longer than the events that caused them.

6.5 Reinforcement of Urban Hierarchy

In the course of time (at least since the first industrial revolution), inequalities in city sizes have shown a tendency to grow. This growing inequality can be explained both by the arrival of new smaller towns in the system and by the fast growth of the largest cities. A first divergence from Gibrat's model that is often noted is that the variations in growth are not totally independent of city size: over long periods of time, the large cities grow a little faster and the smaller towns and cities a little more slowly than the average, and inequalities in size reinforce more markedly than what is expected by the model (Bretagnolle, 1999).

6.5.1 Concentration of Urban Population in the Integrated Systems

Whatever the system considered, the degree of hierarchization increases over time in integrated systems (see Table 6.4; an exception is noted for South Africa, which can be explained by the late period when the system became established).

The observation of adjustment slopes in the rank-size graphs for the last three decades could suggest that this process of historical concentration is terminating

Table 6.4 Increase in inequalities among city sizes

	1900	1950	2000
Europe	0.74 (1850)	0.91	0.94
India	0.76	0.86	0.99
United States	0.97	1.15	1.20
South Africa	1.39	1.16*	1.19

*This slope is for 1960, because from this date on the entries of small towns and creations became less frequent.
The degree of inequality is measured by the absolute value of the adjustment slope in rank-size graphs.

in countries with long-standing industrialization. However, observations once again show that the choice of delimitation criteria for urban areas has a strong influence on the results of measures of growth and concentration. Using France as an example, if evolving delimitations are adopted that take the increase in transportation speeds into account, it can be seen that concentration, far from falling off, actually increases (Bretagnolle, Paulus, & Pumain, 2002). In the USA, the results are less clear-cut: whatever the delimitations used (official SMA/MSA definitions or the harmonized database that we have developed), concentrations do indeed decrease between 1970 and 1980, and fluctuate after that time.

6.5.2 Selection Process and Hierarchical Diffusion of Innovations

The reinforcement of urban hierarchy can be explained by the process of hierarchical and selective diffusion of innovation, which combines two effects: (1) a growth advantage for the largest cities in a system by way of early adoption of innovations and (2) a tendency of relative decline for smaller towns and cities short circuited by these innovations. This process will be examined in more detail in Chapter 8. The capture of innovations by the large cities is explained by the complexity and diversity of their functions and infrastructures (the result of a long history of successive adaptations), which provide better access to information and a greater capacity to carry the high cost and risk associated with innovation. Over the long term, the most marked growth trends that arise from the initial advantage benefit the larger cities and result in "top-down" hierarchization within city systems.

Conversely, smaller towns and cities gain access to innovation at a later date, or not at all, which results in a "bottom up" hierarchization effect. The diffusion of innovations in transportation illustrates this selection process. Whether we consider the pre-industrial era with its mail-coach network or the later railway network, motorways, or airline networks, the large cities and the smaller towns possessing an attractive specialization first profited from such services. From as early as 1820, the quest for increased speed, which is not solely a modern concern, led to a reduction in the number of intermediate stops throughout the 19th and 20th centuries. Figure 6.8 shows how, over and above specific savings brought about by any given technical innovation (for instance the braking power of locomotives or the electrification of the railways), there is a constant and regular increase in the average speed from Paris to the main cities in the French urban system.

For the towns or cities that lost their "nodal" position of access to the fastest networks of their time, effects are visible in the long term, even if demographic variables alone are taken into consideration (see Bretagnolle, 2003). Thus, in disagreement with certain widespread assumptions about the growing universal availability of information, in history the reverse has occurred. Innovation in transport and communication networks does indeed spread among towns and cities overall but with time lapses in relation to city size, which leads to greater concentration of the channels for the circulation of information that is essential at a given time.

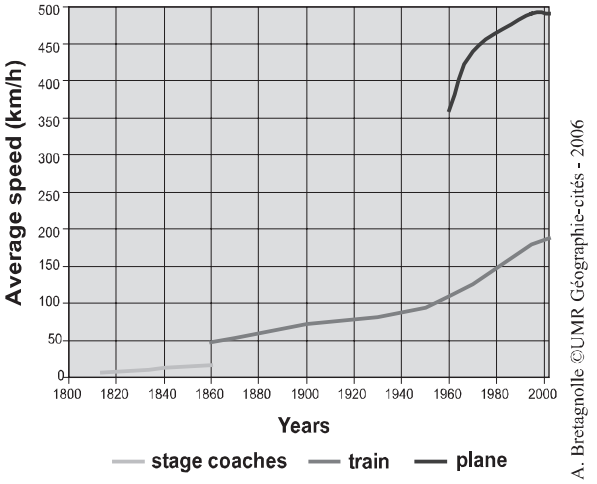


Fig. 6.8 Increasing speed of transportation, between Paris and main French metropolises (1800–2000)

As an example of the selective diffusion process, we have mapped the successive extensions of the postal road network in France from 1584 to 1833 (Bretagnolle & Verdier, 2005, 2007). As an innovation, these postal roads connected first the cities that were already in the upper part of the urban hierarchy, following a dual process of hierarchical diffusion and space-filling (Fig. 6.9).

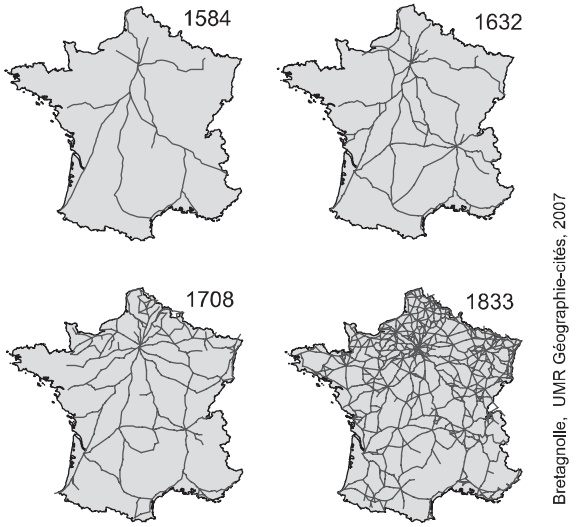


Fig. 6.9 The evolution of the postal roads between 1584 and 1833 (current boundaries)

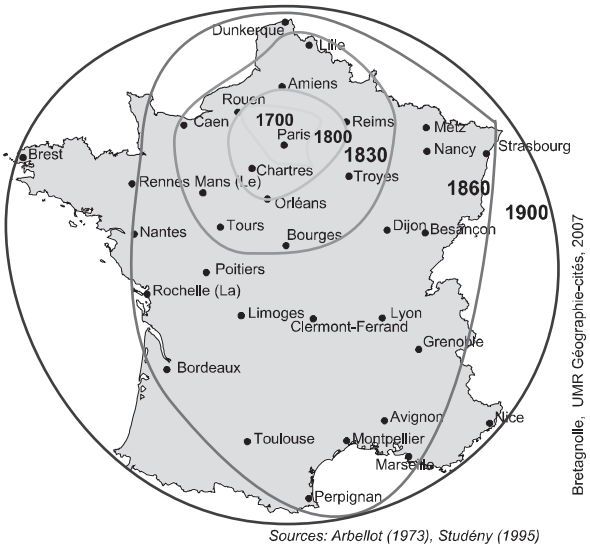


Fig. 6.10 Isochronic map: cities that are located at a one day distance from Paris (1700–1900)

As an example of the acceleration in the evolution of social interaction space from postal network to railway network, we have mapped the French cities that are located at a one-day distance from Paris between 1700 and 1900. These were the places that had the benefit of rapid interaction and that were not accessible to non-connected cities (Fig. 6.10). The consequence of this "space-time contraction" (Janelle, 1969) is that the smaller towns and cities grow on average at lower rates than the largest (Table 6.5).

Table 6.5 Average annual growth rates (%) of cities according to their size

Size class	Europe (since 1850)	India (since 1901)	USA* (since 1940)	South Africa (since 1951)
>100 000	1.38	2.45	1.79	3.21
50–100 000	1.13	1.95	0.92	3.11
< 50 000	0.99	1.71	*	3.10

*Values are computed for SMAs, i.e. since 1940 and with a minimal size of 50,000 inh.

6.6 Conclusion: Toward an Evolutionary Theory of Urban Systems

Through observation of progress through history, and comparing different regions in the world, we have identified several patterns of city system dynamics. We have underlined the importance of identifying urban and territorial entities that are geographically relevant to comparisons in time and space. We went on to distinguish phases in the establishment of these systems, or in their de-structuring

and re-structuring, and phases of "normal" evolution in consolidated systems, when interactions among the towns and cities are controlled and regulated in a more homogeneous manner, for instance inside national frontiers. We have shown that the structural properties of these integrated urban systems are partly similar, insofar as they result from these dynamics, and partly dependent on the history of the political territories to which they belong. The main difference observed between the "older" city systems and those in "new" countries can thus be explained by the evolution of the material conditions in which individuals, goods and information were circulating at the time when they became established. The second difference, distinguishing more "monogenetic" systems from "dual" systems, can be explained by the interference between dynamics belonging inside the system and the dynamics of relationships outside their territory. Thus, there is indeed a correlation between the structure of city systems and the terms of exchanges between cities, over time.

These exchanges and interactions among cities, on all scales, are of the center-periphery type, which generates asymmetry, and, which in turn, enables the accumulation of population and activity in certain places and some redistribution via diffusion. Exchanges among cities are multiform, and their range is highly varied; they produce complex networks where patterns of hierarchical structuring can be detected, but it is very difficult to observe such interaction in a direct manner. Just as we had to infer the dynamics of city size from the observation of the evolution of population sizes, in Chapter 8 we will attempt to deduce the logic of the ability of cities to generate and adopt innovation from observations of changes in their economic activities and in their social composition. Understanding these processes is essential to analyze how urbanization actually contributes to innovation and to globalization, so as to go on to identify levers for action.

References

Bairoch, P., Batou, J., & Chèvre, P. (1988). *La Population des Villes Européennes de 800 à 1950, Banque de Données et Analyse Sommaire des Résultats* (Vol. 2). Publications du Centre d'Histoire Economique Internationale de l'Université de Genève. Geneva, Switzerland: Librarie Droz.

Berry, B. J. L. (1964). Cities as systems within systems of cities. *Papers of the Regional Science Association, 13*, 147–163.

Berry, B. J. L. (1967). *Geography of market centers and retail distribution*. Englewood Cliffs, NJ: Prentice Hall.

Braudel, F. (1967). *Civilisation matérielle et Capitalisme*. Paris, France: Colin.

Bretagnolle, A. (1999). *Les systèmes de villes dans l'espace-temps: effets de l'accroissement de la vitesse des déplacements sur la taille et l'espacement des villes*. Thèse de doctorat, Paris, France:Université Paris I.

Bretagnolle, A. (2003). Vitesse et processus de sélection hiérarchique dans le système des villes françaises. In D. Pumain, M.-F. Mattéi (Eds.), *Données Urbaines* (Vol. 4). Paris, France: Anthropos.

Bretagnolle, A., & Verdier, N. (2005). Images d'un réseau en évolution: les routes de poste dans la France pré-industrielle (XVII^{ème}- début XIX^{ème}s.). In *Mappemonde*, No. 79 (2005-3) http://mappemonde.mgm.fr.

Bretagnolle, A., & Verdier, N. (2007). Expanding the network of postal routes in France (1708–1833). In Le Roux (Ed.), *Post offices of Europe 18th–21th century. A comparative history* (pp. 155–173), Paris, France: Comité pour l'Histoire de la Poste. http://halshs.archives-ouvertes.fr/halshs-00144669/en/. Accessed 5 November 2007.

Bretagnolle, A., Paulus, F., & Pumain, D. (2002). Time and space scales for measuring urban growth. *Cybergeo, 219*.

Bretagnolle, A., Pumain, D., & Vacchiani-Marcuzzo, C. (2007). Les formes des systèmes de villes dans le monde. In D. Pumain, & M.-F. Mattéi (Eds.), *Données Urbaines*, 5, Paris, France: Anthropos.

Bretagnolle, A., Giraud, T., & Mathian, H. (2007). L'urbanisation des Etats-Unis, des premiers comptoirs coloniaux aux Metropolitan Standard Areas, *Cybergeo*.

Christaller, W. (1933). *Die Zentralen Orte in Süddeutschland*. Jena, Germany: Fisher.

Crozet, Y., & Joly, I. (2004). Budget-temps de transport: les sociétés tertiaires confrontées à la gestion paradoxale du bien le plus rare. *Cahiers Scientifiques du Transport, 45*, 27–48.

Davies, R. J. (1972). *The urban geography of South Africa*. Institute for Social Research, Durban, South Africa: University of Natal.

de Vries, J. (1984). *European urbanisation: 1500–1800*, London, UK: Methuen.

Gibrat, R. (1931). *Les Inégalités Economiques*. Paris, Sirey.

Hall, P., Gracey, P., & Drewey, R. (1973). *The containment of urban England*. London, UK: Sage Publications.

Hollar, W. (1667). *Hollar's Exact Surveigh of the City of London*. Reproduction, Edward Stanford Ltd., N.D. 83 × 54 cm. De la Feuille James (1690), Londini angliae regni metropolis novissima *and* accuratissima, http://www.british-history.ac.uk.

Janelle, D.G. (1969). Spatial reorganisation: a model and concept. *Annals of the Association of American Geographers, 59*, 348–364.

Pinol, J. L. (2003). *Histoire de l'Europe Urbaine. 1 – De l'Antiquité au XVIII^{e} siècle*. Paris, France: Le Seuil.

Pirenne, H. (1927). *Les Villes au Moyen Age: Essai d'Histoire Economique et Sociale*. Brussels, Belgium: Lamertin.

Pred, A. (1977). *City systems in advanced societies*. London, UK: Hutchison.

Pumain, D. (1997). Vers une théorie évolutive des villes. *L'Espace Géographique, 2*, 119–134.

Pumain, D. (Ed.). (2006). *Hierarchy in natural and social sciences*. Methodos Series Vol. 3, Dordrecht, The Netherlands: Springer.

Pumain, D., & Gaudin, J.-P. (2002). Systèmes de villes et pouvoir. L'analyse de Giovanni Botero à l'époque de la Renaissance. *Cybergeo, 227*.

Pumain, D., Saint-Julien, T., Cattan, N., & Rozenblat, C. (1991). *The statistical concept of city in Europe*. Eurostat.

Reynaud, J. (1841) Villes. *Encyclopédie Nouvelle* (Vol. 8, pp. 670–687). Paris, France: Gosselin.

Vacchiani-Marcuzzo, C. (2005). *Mondialisation et système de villes: les entreprises étrangères et l'évolution des agglomérations sud-africaines*. Thèse de Doctorat, Université Paris 1, tel.archives-ouvertes.fr/tel-00011351/en/.

Zelinski, W. (1971). The hypothesis of the mobility transition. *Geographical Review, 61*, 219–249.

Zipf, G. K. (1941). *National unity and disunity*. Bloomington, IN: Principia Press.

Zipf, G. K. (1949). *Human behaviour and the principle of least effort*. Cambridge, MA: Addison-Wesley Press.