Proceedings of the 14th International Space Syntax Symposium

25

The natural scale of urban networks

Intelligibility is not scale-free

Sean Hanna; Space Syntax Laboratory, Bartlett School of Architecture, UCL; s.hanna@ucl.ac.uk

---

ABSTRACT

Angular street segment measures are used to analyse the degree to which network intelligibility of cities becomes evident at different scales. It is shown that the phenomenon is not scale-invariant, but has a distinct peak at which correlations are highest, providing a recognisable scale of any given city map that results from its broad network structure rather than individual features such as mean street length. For actual street networks it is found that this scale corresponds with the radius at which Syntax measures such as choice best predict human movement within the city.

These results may explain the relationship between individual navigation strategies and urban form, and why Space Syntax techniques as different as centrality measures and agent simulation should both be effective in predicting movement. Syntactic intelligibility appears to operate at a particular radius around 2-4km; at this scale, both a perfectly informed strategy of route minimisation (network centrality measures) and a locally opportunistic navigation (stochastic agent simulations) yield similar patterns. Empirical evidence indicates this is the scale at which both correlate with observed movement and of an average journey length. The networks examined can thus be considered to be optimised to convey the most relevant information on likely journeys to navigators throughout the network, and to do it at the most useful scale.

KEYWORDS

Scale-free networks, power law, urban networks, intelligibility, space syntax analysis

1 INTRODUCTION

Cities are sometimes described as self-similar across a wide range of scales, and yet, conversely, as particularly dependent on specific scales. The ‘scale-free’ nature of cities has been claimed both in their overall distribution (Batty 2006, Zipf 1949) and internal morphology (Carvalho and Penn 2004; Jiang 2009). A consistent power-law distribution observed in languages, structures like the internet and gene regulatory networks, explained by processes of utility maximation (Zipf 1949) or preferential attachment (Barabasi and Albert 1999), has also been attributed to urban form, resulting in the assertion that its basic structure is fractal (Batty and Longley 1994), with street connectivity and length defined by a power-law (Jiang 2009) irrespective of scale. Yet basic features such as block size have limits: they are not so small to be measured in millimetres or so large as kilometres, and in reading a map it is often easy to estimate their approximate scale. It appears also that more complex network properties have a scale: Space Syntax analyses have found certain scales or radii (e.g. axial radius 3) more useful in predicting natural movement or revealing particular urban features or functions (Krenz 2017). Particular patterns appear in analyses at different scales: angular choice at a low radius highlights a series of discrete 'hot spots' while high radius separates a 'foreground' of linear routes from a 'background' network.

Where such scales have found expression in practical rules of thumb for conducting analyses, they are used because they have proven expedient in revealing other phenomena in question, such as commercial land use or pedestrian traffic routes. Scale itself is less often the subject of analysis or explanation. What is it about these scales that makes them particularly relevant? And do they differ from one city to the next?

Urban scale is examined here through the Space Syntax concept of *intelligibility*, which refers not to a cognitive understanding of space but to the degree to which information about long range structure is made evident at any given point. Intelligibility is quantified as a correlation between a local measure (traditionally axial connectivity) and a long-range measure (traditionally axial integration, radius 3 or n); high values are normally expected in city networks, with low values thought to indicate pathological structure (Hillier et al 1987; Hillier 1996; Penn 2003). When a city is intelligible, a visitor unfamiliar with the map will be able to infer the centrality of their location and decisions on optimal routes based only on local properties of their current location, and thus intelligibility is thought to support rational and natural movement (Hillier et al. 1993). It is also a property that is not reducible to any particular parts of the city in isolation, but of the overall network. If intelligibility is manifest at a particular scale, it would be crucial both to our navigation and our understanding of the city.

The hypothesis of this paper is that there is an optimal scale at which the city is most intelligible. It tests this using a method that differs from the most prevalent measure of axial intelligibility but is more directly tied to actual cognitive experience and more precisely quantifiable in terms of distance. It then explores the relationship of such a scale to the prediction of movement and investigates the extent to which it is consistent across cities. To the extent such a scale exists, it is an important feature of a city's morphology with respect to movement. To the extent it is universal from city to city, it suggests a means by which our cognitive and navigational abilities may operate consistently regardless of the city we are in, and by which we intuit the scale of any city by looking at its map.

2 THEORY

2.1 Scale-free or scale-dependent: power laws and fractals in cities

The description of cities as scale-free is based on observations that certain properties are fractal (Batty and Longley 1994) or self-similar at a wide range of different scales. Fractal dimension has been proposed as a distinguishing feature, and a power law has long been observed to describe the size of urban areas or populations, such that the size of the city is highly correlated with the reciprocal of its rank (Zipf 1949, Batty 2006) In Zipf's (1949) original ranking of US metropolitan districts, for example, the second city was approximately half the population of the first, the third 1/3 the population, the forth 1/4, and so on.

The same power law, or rank-size rule, has been found to describe the distribution of streets within a city, e.g. as quantified by the length or connectivity of axial lines (Carvalho and Penn 2004; Jiang 2009; Zhang et al. 2013), however in these cases the fit has not been as close. In particular, the power law tends to break down toward the extreme ends of the scale, for the highest and/or lowest ranked samples. This is to be expected at the largest end of the data set due to limited sample sizes in finite sets as wholes (there are very few *largest* streets in a city), and also at the smallest end because there is a limitation to the size of the individual elements (axial lines cannot be smaller than a single segment, i.e. length of a block).

Where the power law is claimed to hold for city morphology, then, is in the middle range, not at the extremes, and it may be bounded at both ends by features that have definite scale, namely segment length and overall city size. It is this intermediate range of scales over which it may appear to be scale-free. This paper will test this middle range, to determine whether even here there is a recognisable scale of the network as a whole that is independent of either street length or overall size.

2.2 Long-range and local structure: intelligibility

The properties of the city most often considered in Space Syntax are non-local network measures of network structure, such as centrality measures integration and choice, which are observed to relate to observed movement and socioeconomic behaviour. Relationships have also been found between both long-range and local properties, such as the association of land-prices with both high radius network integration (Chiaradia at al. 2009) and local street-view photographs (Law et al. 2019). The most prevalent relationship involving both local and long-range structure of the network is *intelligibility*, proposed as a fundamental property in early Space Syntax theory describing the degree to which "the whole can be read from the parts"

(Hillier et al 1987). Despite the apparent reference to cognition of the name, intelligibility is a purely morphological property of the street network, quantifying how well long-range structural properties can be predicted from the local.

Intelligibility is typically calculated as the correlation between the local connectivity (number of connections, or degree) of an axial line and its integration within the network at a larger scale, e.g. radius 3. This measure is not directly compatible with more recent methods like angular segment analysis (Hillier and Iida 2005) that have been effective in revealing morphological properties or predicting behaviour (Turner 2009), or with maps based on road centre lines rather than axial. Other measures have been proposed which do use these representations, their primary principle being a correlation between a value representing a long-range network property relevant to movement and a value representing a property that is directly visible locally (Hanna 2022).

Segment angular choice (or betweenness centrality), is the long-range Space Syntax measure that is theoretically representative of through-movement (Hillier et al. 1987; Hillier and Iida 2005), empirically best able to predict observed movement (Hillier and Iida 2005; Hanna 2021) and has been found to reveal the “foreground” network of major streets and through-routes that appears an important feature of large-scale morphology.

The most local property of a street segment within the graph is the angle at which it joins those street segments immediately adjacent to its current location. A random walk can be simulated in which agents make a choice of direction based only on these angles. Such a random walk is analogous to those taken by EVAS agents in DepthmapX (Turner and Penn 2002), but when implemented within the constraints of a segment graph, can be calculated efficiently and exactly as a function of probability without resorting to stochastic approximation (Hanna 2021). The distribution of agents on each segment converges over time such that some streets have a greater density of traffic (fig. 1); these resemble visually the pattern of segment angular choice.

The two measures are idealisations of two different ways in which one might navigate through a city. Choice is a measure of all the *optimal routes* (i.e. minimising angle) through each point in the graph, and so represents the action of navigators with a complete and accurate knowledge of the whole network, on which they base decisions. A simulated random walk is composed of agents’ decisions based only on *immediately visible features*, without broader knowledge or memory from one step to the next. As such, they comprise the long-range and local properties of intelligibility. They also can be quantified at a precise range of scales (in metres): measuring choice by setting a limiting radius of analysis, and random walks by taking the distance walked.

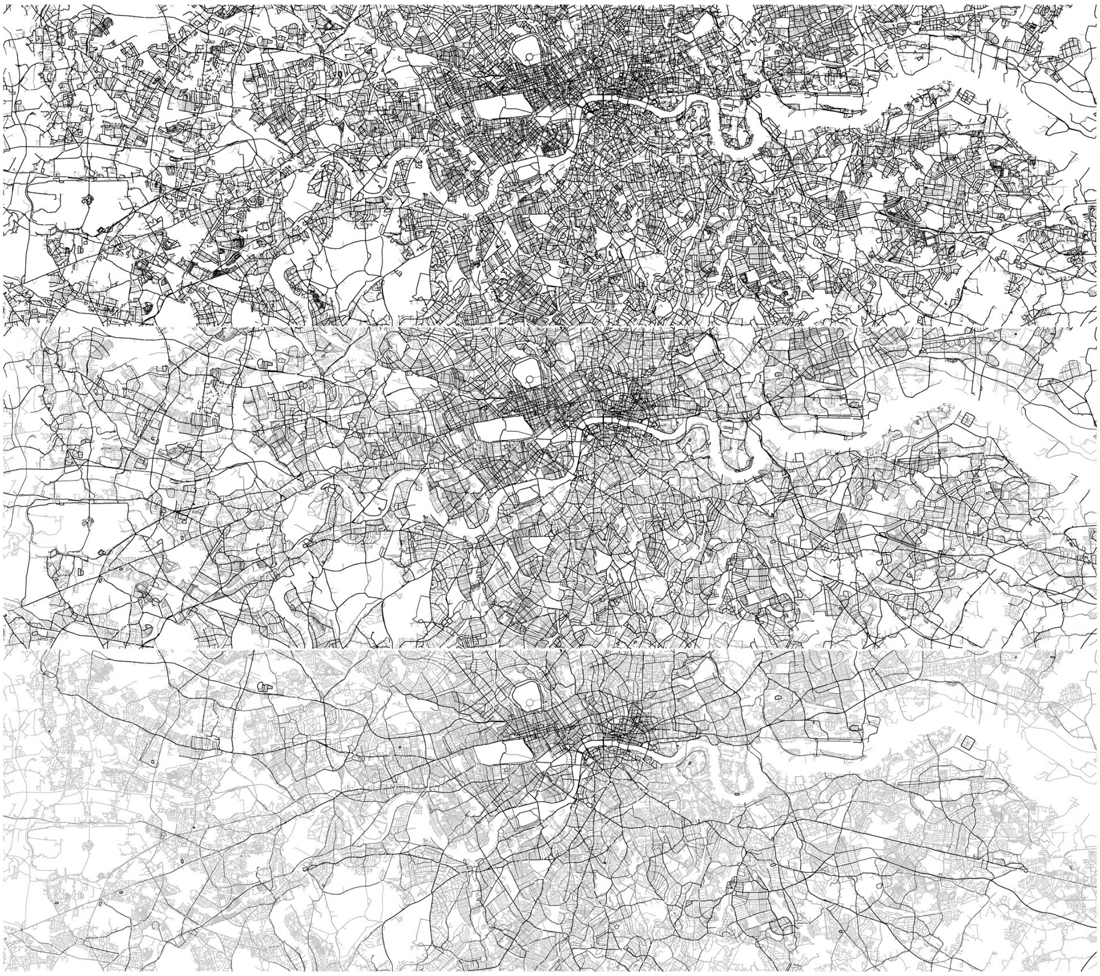

Figure 1: Weighted random walks in London: distribution of simulated traffic at steps 5, 26 and 100. As the walk progresses, values converge on a pattern approximating medium to high radius of choice (betweenness centrality).

2.3 The scale of prediction and intelligibility

Intelligibility is connected to movement, as has been demonstrated empirically. Intelligibility has been observed to be strongly related to the degree to which movement in a given area can be predicted by spatial integration or choice (Hillier et al 1987; Penn 2003, Zhang et al 2013). Theoretically, intelligibility is related to the theory of natural movement (Hillier et al 1993) in which aggregate pedestrian or vehicle traffic is determined, or predicted, by the street network; two potential mechanisms have been implied. The first is a cognitive understanding of the network as a whole, in which intelligibility allows us to see “a picture of the whole urban system […] built up from its parts” (Hillier 1996, p.94); Penn (2003) describes human cognition as a “correlation detector” with which one learns to infer global properties from what is visible locally. The second is immediate action based only on local affordances, or “exosomatic” (Turner 2007) navigation, as has been used in agent models (Turner and Penn 2002) and observed in some navigation tasks (McElhinney et al. 2022).

These individually related properties together have a triangular relationship, which diagrams (fig. 2, after Hanna 2022) the mutual dependence of intelligibility and predictability. Global measures, at the top of the diagram, may be at various scales, and it has been observed that some of these predict movement better than others: urban traffic prediction based on angular choice has been observed most effective around 4 km (Serra and Hillier 2019, Hanna 2022). In theory, this suggests that the same scale might apply to intelligibility. This refines the basic hypothesis of the paper, which posits the existence of an optimal scale of intelligibility, to suggest it might also be found around 4 km.

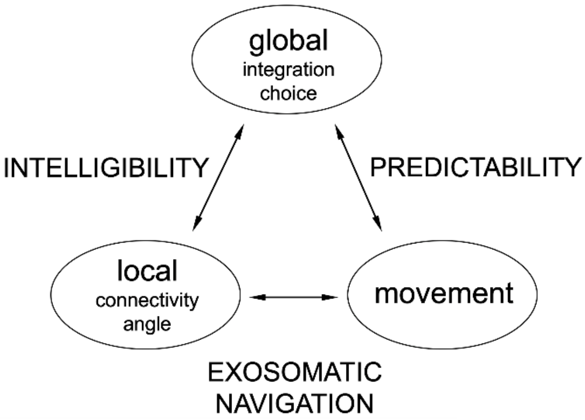

Figure 2: The mutual relationship between intelligibility, predictability and exosomatic movement. (Redrawn from Hanna 2022).

3 DATASETS AND METHODS

A set of 21 cities was selected to represent a diversity of potentially relevant variables: overall size[1] (16,914 to 976,445 street segments), mean street segment length (20.4 m to 81 m) and geographical location (North America: 5; South America: 3; Europe: 10; Asia: 3). See figure 3, and table 1, for the full list. To maintain some consistency with previous studies of intelligibility and due to very rare cases in which road centre lines can fail to capture lines of sight relevant to a random walk (see Hanna 2021), 18 of these were selected from axial map representations used in previous Space Syntax studies. Road centre line representations were used for three large cities (Atlanta, London, and New York City), because axial maps were unavailable for the full extents used. In all cases, analyses were performed on line segments, so any continuous (i.e. axial) lines were divided into separate units at the intersections with other streets.

---

[1] For those concerned with power laws in sets of cities, it might be noted that the data set conforms nicely: size (in segments) and rank^{-1} are very strongly correlated, with r=0.985. This was not an aim in choosing the cities, but simply a result of selecting a representative set.

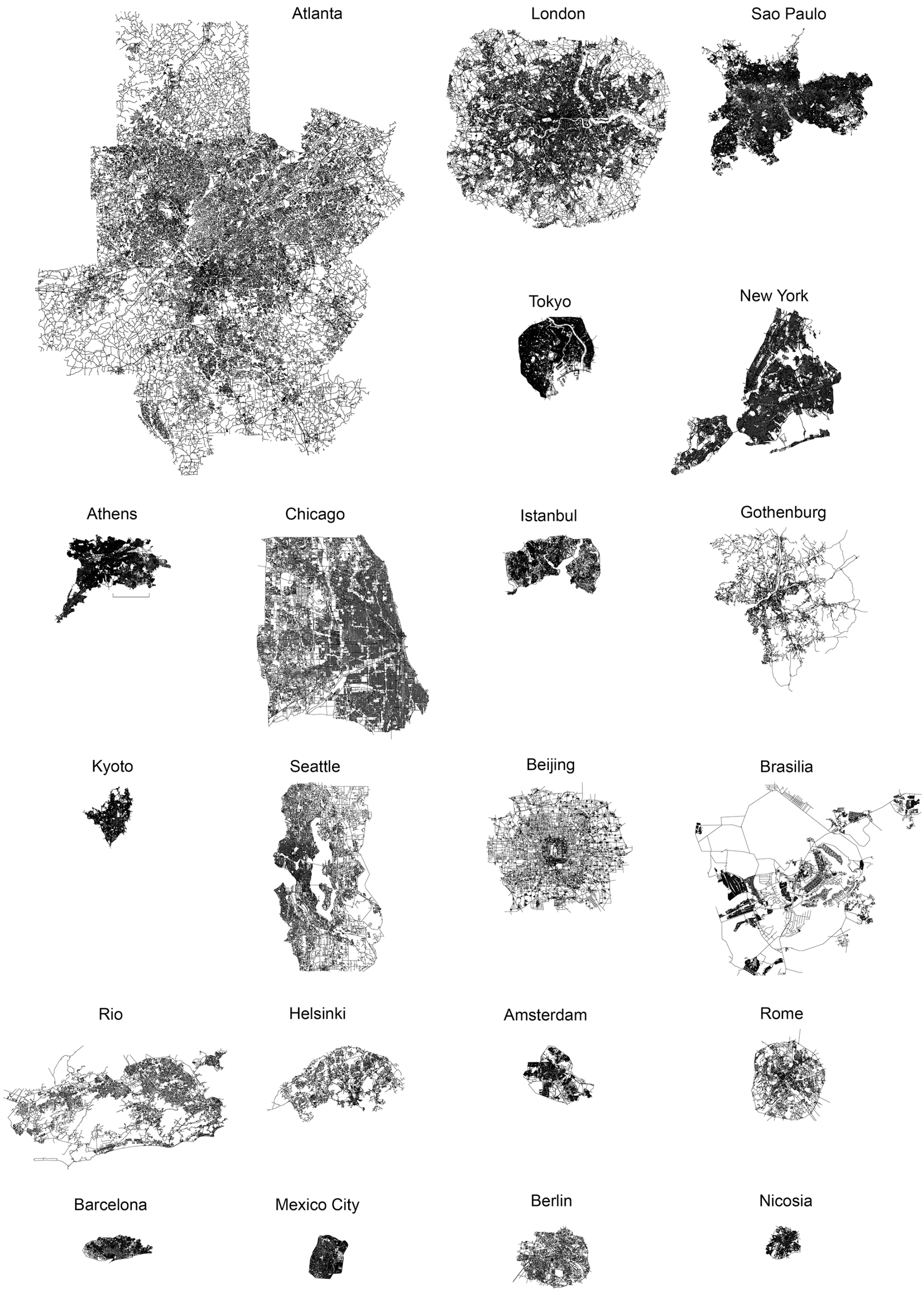

Figure 3: All 21 cities, arranged in order of the greatest to fewest number of street segments.

3.1 Long-range optimal routes

To assess globally or long-range optimal routes, centrality analyses were performed on all segment maps using DepthmapX (2017), version 0.8.0. Angular choice (a form of betweenness centrality) measures were taken at standard radii raging approximately from street segment length to city radius, in approximately doubling increments: 50 m, 100 m, 200 m, 400 m, 1 km, 2 km, 4 km, 10 km and 20 km. The result of each analysis is a value for each street segment quantifying its overall centrality at every radius, with higher choice indicating more central streets and an expectation of greater traffic under the assumption that all travellers take optimal routes for journeys of that radius.

3.2 Local exosomatic navigation

To assess exosomatic navigation based only on local properties, simulated random walks were performed on each city using a non-stochastic calculation of the distribution at each time step, using a method found to predict observed movement in London and the surrounding region (see Hanna 2021, for full details). The probability of selecting a route from the choices available at each intersection is given by a function of angle, such that straight routes are more probable than sharp turns. Hanna (2021) uses weights each by a non-linear function of angle raised to an exponent, namely $angle^4$ or $angle^5$, as these best approximate the probabilities of choice made by the random walking EVA agents (Turner and Penn 2002) typically used in Space Syntax and implemented in DepthmapX (2017). The same function has also been found to result in random walks that correlate well with observed human movement (Hanna 2021, 2022). For the present work, a range of exponents from 1 to 10 was tested for peek correlations between the resultant random walk and choice values, and $angle^5$ was again found to correlate best with a mean of r=0.58 (Pearson) across the set of cities used.

The random walk results in a density of traffic at each street segment, which changes in time from one iteration to the next as agents step between their current position and adjacent segments. Beginning in the first iteration with a uniform distribution of identical traffic density at all segments, this approaches a steady state distribution over time, with the greatest changes occurring in the initial time steps and gradually converging. Walks are run for 300 iterations, equivalent to a range of 45 m to 13.8 km for the average city.

3.3 Intelligibility

In analyses based on axial line representations, intelligibility has been quantified by the correlation between long range integration and local connectivity (Hillier et al. 1987; Hillier 1996; Penn 2003). Here it is assessed by the correlation between the two measures above: choice as a long-range measure of optimal routes, and the density of random walkers as a measure of local decisions. For a given radius (choice) and a given step distance (random walk), each segment has a unique value for both of these measures, so the correlation between both is calculated across the full range of scales noted above (radii 50 m to 20 km, and random walks approximately 45 m to 13.8 km). Because these are sampled at discrete intervals, and they are seen to form continuous curves, correlations for intermediate scales are found by cubic interpolation to determine peaks (figures 4, 5).

This value, representative of intelligibility, is lower than expected based on previous published correlations between random walks and choice (and with traffic). London had been in the range of 0.7 or 0.8 (Hanna 2021, 2022), with Hanna (2022) reporting high values also for New York City, using Spearman rank correlation. Two reasons exist, the first being that rank correlations tend to be higher: Spearman rank correlation values were tested for the present set and found to have mean of 0.68 for a random walk using $angle^5$, and peaking at 0.69 with a slightly lower exponent of $angle^3$. The second reason for the lower mean correlations is that London appears to be particularly intelligible: it was found that London has the highest correlation of all cities in the sample (Pearson linear r=0.74 for $angle^5$; Spearman rank rho=0.775 for $angle^3$). The variance in overall intelligibility among cities (e.g. Athens at the low extreme of r=0.46, rho=0.63) is not unimportant, but is independent of the question of the scale at which intelligibility is greatest for any given city, which is the concern of this paper.

The use of rank correlation may be useful in some cases as the data are non-uniformly distributed toward the lower values of both choice values and walk distribution, and the linear Pearson correlations may be thought to miss important variation in this range while being swayed unduly by comparatively few data points of high choice and walk traffic. However, given that these high values represent the distinction of the ‘foreground’ from the ‘background’ network of the city, it is the linear Pearson correlation and the random walk function of $angle^5$ that are normally used in this paper.

4 RESULTS

For a given choice radius, correlations between a random walk and choice are plotted over time; figure 4 shows the first 100 steps in London. A common pattern is found in all cities as the walk progresses, in which this value begins at zero (for the uniform distribution of random walk traffic) and typically increases over the initial iterations, before either plateauing or decreasing again slightly after a number of steps. There is likewise a common pattern in the sequence of curves overall, in which the lowest radii of choice have poor correlation values overall, which increase for larger radii, and then fall again after a certain point. The peak correlation values in each direction therefore correspond to two scales in which the city is maximally intelligible: the first indicating the length of the random walk journey and the second the radius of choice.

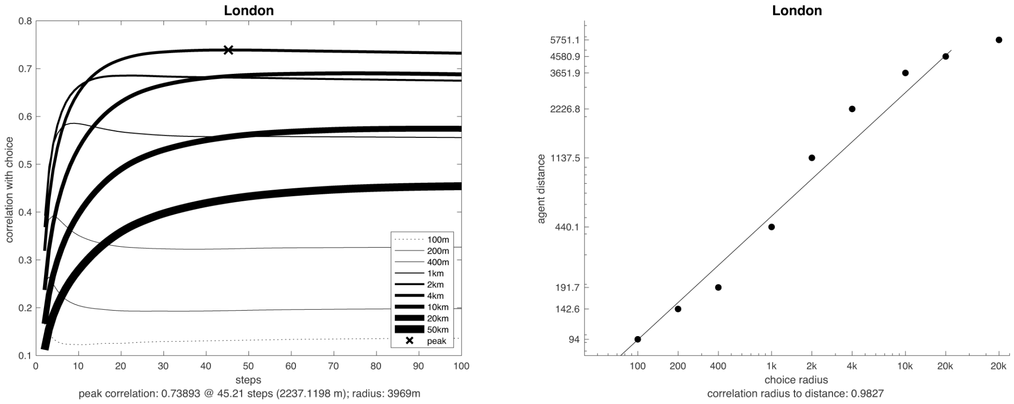

Figure 4: Correlation between choice and random walk distributions from step 1 to 100 of the random walk in London; successive curves represent increasing choice radii (left). As radii increase, the mean distance at which the random walk best predicts choice increases linearly.

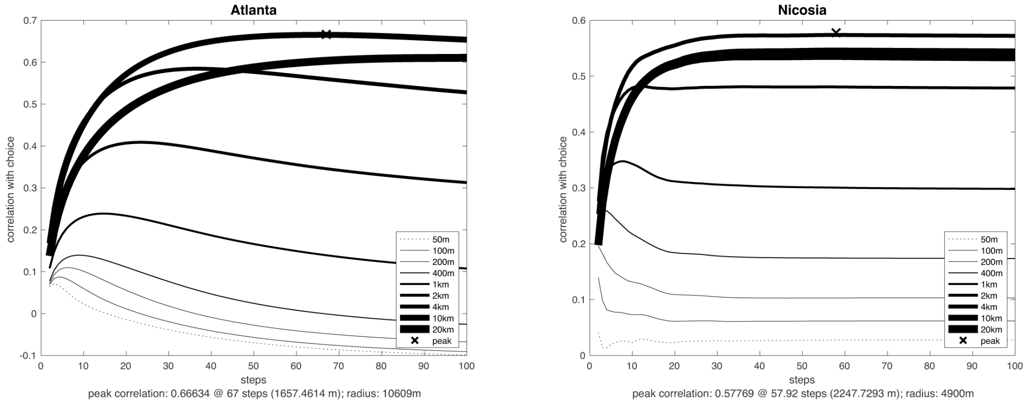

Figure 5: Choice and random walk correlations follow a similar pattern for all cities: Atlanta and Nicosia, the largest and smallest cities in the dataset, for comparison with London. The step at which maximum correlation occurs is marked with an X.

The number of steps at which correlations peak increases with increasing radius, suggesting a relationship between the distance in which agents walk and the radii of most intelligible paths. The number of steps indicates only an approximate distance travelled by the agents; however, the actual distance, as an average of all agents, can be calculated precisely as the cumulative sum at each step of the graph segment lengths, where each segment is weighted by the number of agents present on it at each time step. Thus, if agents converge more to the longest segments over time, for instance, their distance travelled will increase by more.

Figure 4 (right) plots the peak correlation values for each choice radius (horizontal axis) against the mean agent distance at which that peak occurs, also for London. The distance at which agents correspond with radius 100m routes is 94m; 2km radius is best approximated by agents at 1.1 km; 20km radius is approximated at 4.6km. For all cities, there is an almost perfect linear relationship between the two (r=0.98 for London); the slope of this function varies somewhat between cities, but in all cases the distance travelled by agents is a smaller value than the corresponding choice radius.

This linear relationship between peak choice radius and walk distance is so closely correlated in every city as to suggest a general law of correspondence between the two, and yet the values for agent distance are not the same as choice radius but uniformly lower. Two observations may be made about why this should be the case. First, as can be seen from the 4km radius curve in London (fig. 4, left), while correlations peak at 2.2 km, they remain high over a considerable plateau: the curve reaches 99% of its maximal value at a distance of 1,288 m, and retains this until a distance of 5,118 m. Thus, at journeys of 4 km, agent paths still correlate extremely well with 4 km choice. Second, while the agent distances are mean journey lengths, choice corresponds to maximum journeys: radius 4 km determines a measure of shortest paths within a maximum of 4 km of a given node. The actual distances are dependent on the configuration of the network and are not provided in the measure of choice, but it may be that the average length of these paths may be much less, and closer to that of the agents.

4.1 Cities compared: the general scale of intelligibility

The qualitative pattern of correlations seen in London is evident in all cities studied, with similar sets of curves seen in figure 5 for the largest and smallest cities in the set, Atlanta (976,445 segments) and Nicosia (16,914 segments).

We can determine a single point at which the overall maximum correlation occurs for any city, corresponding both to the peak radius of choice and the random walk point at which these routes are most intelligible. This is found by interpolating between each of the correlation curves (e.g. in figs. 4 and 5) and finding the maximum on the corresponding surface. This provides a maximal value of intelligibility for the city (the vertical axis, figs. 4, left, and 5) and distances both for the agent journey (the horizontal axis) and choice radius (axis not shown). It is marked with an X in figures 4 and 5. The random walk distance in metres is calculated as a cumulative sum of agent's segment lengths, as above.

The random walk distance and choice radius of these peak correlations are listed for all 21 cities in Table 1, which lists cities in order of size, from the greatest number of segments to the least. The mean values for the set as a whole give a choice radius of 4,700m and a slightly lower walk distance of 3,283m, suggesting that the cities as a group are most intelligible at a scale of around 3-5 km, similar in magnitude to the scale of expected actual journeys and that at which observed movement tends to be best predicted by choice. As with London, the agent distance is lower than the choice radius.

Table 1: Correlations, choice radii and random walk distances corresponding with maximal intelligibility for all cities, ranked from largest to smallest.

city	peak R	rand walk steps	walk dist (m) (est.)	choice radius (m)	mean segment length (m)	num segments
Atlanta	0.67	67	1657	10609	20.9	976445
London	0.74	45	2237	3969	46.8	312037
Sao Paulo	0.54	14	532	2209	36.7	263201
Tokyo	0.58	14	301	1369	20.4	262394
NYC	0.52	36	1620	3481	47.0	163084
Athens	0.46	186	6262	2500	32.0	145708
Chicago	0.71	96	6824	8100	65.7	136960
Istanbul	0.55	203	7749	2401	37.9	73306
Gothenburg	0.57	31	1437	3600	36.7	67350
Kyoto	0.66	28	622	3249	20.4	64963
Seattle	0.64	212	13615	3721	63.6	63873
Beijing	0.67	101	8368	14641	60.9	61530
Brasilia	0.53	31	2254	6724	64.6	54470
Rio	0.56	14	1208	4225	81.0	46600
Helsinki	0.54	19	956	3481	46.0	36508
Amsterdam	0.49	11	455	1936	39.0	35977
Rome	0.54	27	1694	5625	60.4	27339
Barcelona	0.56	20	856	3481	38.0	27309
Mexico City	0.53	14	594	3721	40.9	27166
Berlin	0.54	101	7464	4761	68.7	22241
Nicosia	0.58	58	2248	4900	35.4	16914
mean	***0.6***	***63***	***3283***	***4700***	***45.9***	***137399***
median	***0.6***	***31***	***1657***	***3721***	***40.9***	***63873***
RandTri	0.66	3	114	540	45.9	52580
RandSts	0.37	4	171	892	45.9	32953

Neither scale is constant; there is some variation across the cities in the set, but the majority of them are within a consistent range. Figure 6 plots both peak choice radii (heavy line) and peak walk distance (lighter line) for the cities in groups classified geographically by continent, in which it can be seen that the majority of cities, including all cities in Europe, have greatest intelligibility at choice radii between 2-6 km. A few exceed this scale, namely Brasilia, Atlanta, Chicago, and Beijing; it may be relevant that these cities were either begun (Brasilia) or greatly expanded during the latter half of the 20th century in a period privileging automobile traffic, and are so not particularly ‘walkable’ over much of their network.

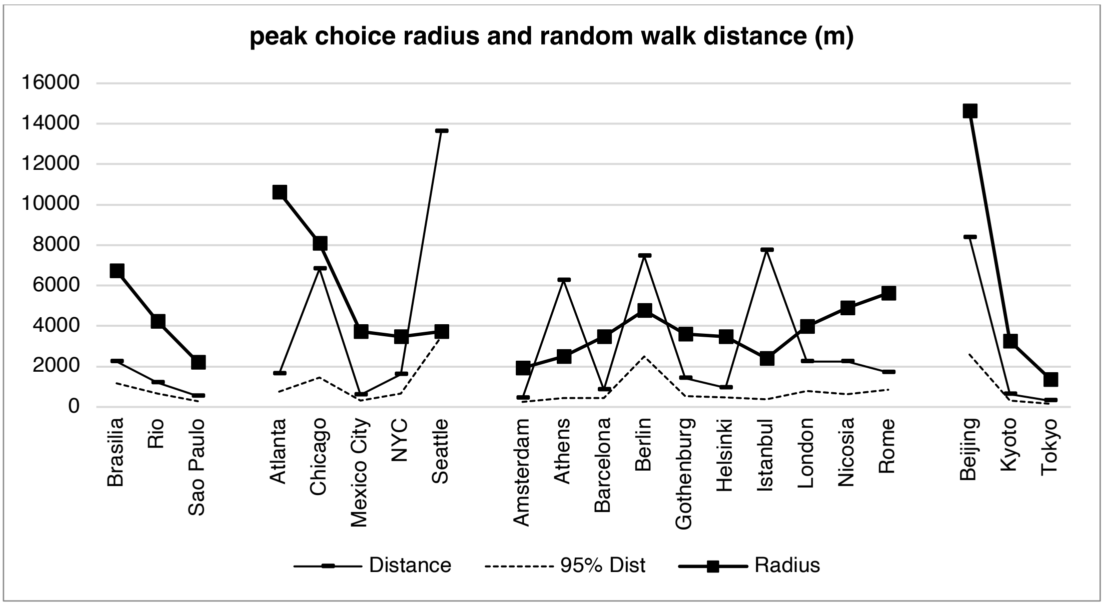

Figure 6: Choice radius (mean 4700m) compared with random walk distance (mean 3283m) of maximum intelligibility. Cities are grouped geographically, by continent. Most cities fall within an approximate range of 2-6km radius; random walks are more varied.

The random walk distance at which intelligibility is maximised is considerably more varied across the cities, and quite independent of the scale of choice (correlation between the two is not strong: r=0.29) as indicated by the plot (thin line) in figure 6. While the mean value is 3.3 km, several cities otherwise in the 'walkable' set of 2-6 km peak choice radii have much greater peak walk distances: e.g. Berlin (7.5 km), Istanbul (7.7 km) and Seattle (13.6 km). However, on closer inspection of these cases it should be noted that the peak value, like the example of London (section 4.0) is on a very flat plateau in which there is very little change in correlation over a long period of steps, and so the confidence in their precise values is low. If the point in the random walk at which the intelligibility first reaches 99% of its maximum is used instead (dotted line, fig. 6) these distances drop to values comparable to the other cities, and below their corresponding choice radii.

Several cities do appear to be outliers: Beijing and Chicago have intelligibility that peaks both for choice of high radii (14.6 and 8.1 km) and for long walk distances (8.4 and 6.8 km) more than one standard deviation from the mean. But the majority of cities have maximal intelligibility for optimal journeys of a (choice) radius close the mean of 4.7 km, correlating best with random walks of a distance close to the mean of 3.3 km.

4.2 Cities and non-cities

The data above suggest that actual city networks are intelligible at a particular scale, but it is not obvious whether this is a unique property of cities or of a broader class of two-dimensional networks. Unintelligible networks exist, and yet high intelligibility has also been found in surprisingly different sorts of geometries, both random (Hanna 2022) and labyrinthine (Zhang et al. 2013). Are cities particularly optimised to be intelligible at a unique scale? If so, can the features that accomplish this be identified? To understand the difference between cities as a class and other potential networks, they are compared here with two randomly generated graphs of different descriptions.

The first is a connected set of randomly placed street intersection points. Coordinates for 20,000 points are generated at random, and each of these then connected by street segments to its nearest neighbours by Delaunay triangulation to form a fully triangulated grid. Segments that are the longest of two adjacent triangles are removed, resulting in a network of segments describing both triangular and quadrilateral blocks. The second is a set of randomly placed and sized lines. 9,000 start points are generated at random, along with corresponding random end points within a region roughly 1/50 the size of the whole network area. Resulting lines typically intersect several others; these are divided into segments at the intersections and short, dead end, stubs are removed. Both randomly generated networks are within the range of city sizes by number of segments, 52,580 and 32,953 respectively, and their size in metres was scaled such that the mean segment length of each is exactly the mean segment length of the set of cities: 45.9 m. Both underwent the same analysis of correlations across choice radii and random walk distances as the cities, and results are shown in the lower rows of Table 1.

The two graphs differ in their maximum intelligibility: the network of randomly placed lines (*RandSts*) has a much lower correlation value than all cities (r=0.37) but the triangulated intersections (*RandTri*) appear quite intelligible with a value above the mean (r=0.66). Where both networks are similar, and both quite distinct from the group of real cities, is the scale at which this intelligibility peaks. Both the peak choice radii (540 m for *RandTri*; 892 m for *RandSts*) and the peak walk distances (114 m and 171 m) are well below even the lowest values for any city and an order of magnitude smaller than the averages. Given that the scale of gross features of total size and mean segment length were set to be equivalent, this difference must be due to the geometry of the networks.

What sort of street network would result if we were to assume the scale of peak intelligibility was that of cities? We can scale both networks such that the choice radius at which the correlation peaks is the mean of the city set, 4,700 m, and find that the *RandTri* network has a mean street segment length of 399 m and *RandSts* has segments of mean 242 m. If instead we increase the scale to match that of the expected random walk distance of the cities, 3,283 m, we have a *RandTri* network with segments averaging 1,320 m and *RandSts* of 881 m. Where a typical city has street segments around 45 m, or about 30 seconds walking, these scaled random grids may have the nearest intersection more than 1 km, or 10 minutes walk away.

Such streets are clearly unlike any real city, but a recognisable similarity is evident if we look at the scaled *RandTri* map next to a real city, as in figure 7, where both London and *RandTri* are shown. While the mean length of all street segments is roughly 13 times that of London, it very closely approximates London's 'foreground' network both in scale and geometry. The greyscale weight indicates choice values at radius 4 km for both networks, resulting in a typical 'foreground' (bold) and 'background' distinction between streets in London, whereas *RandTri* is very uniform throughout, with most segments lying on high choice routes.

This result suggests a qualitative equivalence between the full *RandTri* network and the foreground network of a real city, with the implication both that this foreground network is what drives intelligibility and that it does so at a particular scale of around 4 km. While many other potential graphs exist, it is notable that the two randomly generated ones used here are most intelligible at this scale.

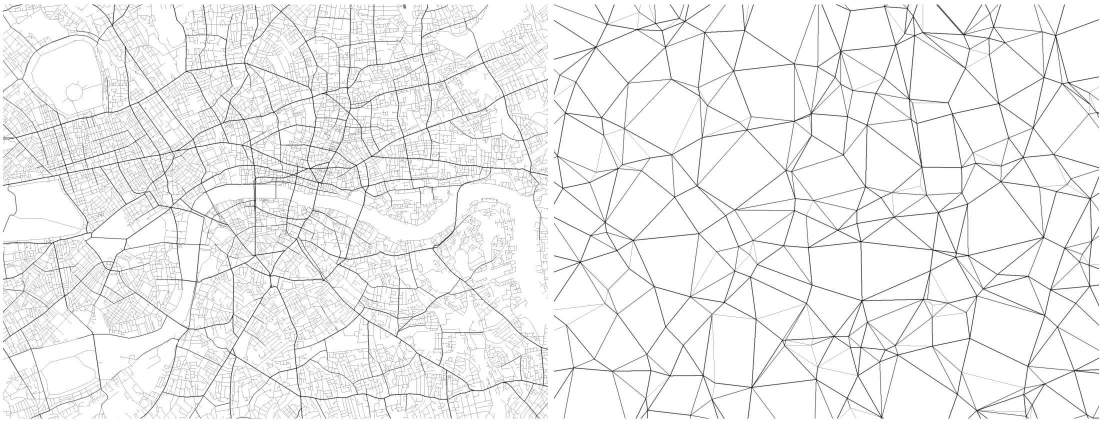

Figure 7: An area of London (left) and the *RandTri* network (right), with the latter scaled to be intelligible at the same scale as cities. The *RandTri* map closely approximates the scale and geometry of London's foreground network (bold lines).

5 CONCLUSIONS

Each of the 21 cities sampled has a unique scale at which intelligibility is maximal, as measured by the correlation between the long-range measure of angular choice and the distribution of locally-based random walks. This is not identical in every city, but is typically within an expected range of approximately 4 km radius of choice (mean 4.7 km, median 3.7 km) and slightly less in walk distance (mean 3.3 km, median 1.7 km). This scale corresponds to the radius at which Syntax measures such as choice best predict human movement within the city.

Several cities are genuinely outliers in that both measures are above this range. Beijing and Chicago have a peak intelligibility more than one standard deviation from the mean both for choice of high radii (14.6 and 8.1 km) and for long walk distances (8.4 and 6.8 km). These, along with others with higher scale, coincide with the middle-late 20th century prevalence of vehicle traffic over pedestrian, and further research may determine whether and how the scale of a given city may be related to its relative accommodation to walking or to driving.

The scale of intelligibility is a property of the network, rather than of individual features or overall size. A relationship might be expected between random walk distance and segment length, both because of existing expectations that movement is based more on topological than metric distance (Hillier and Iida 2005; Turner 2009) and because the latter is used directly in calculating the former: some correlation (r=0.42) is observed between the two. Likewise, there is a small correlation between the overall city size in segments and the peak choice radius (r=0.31), although this is negligible with random walk distance (r=-0.15). But the scale of intelligibility does appear to be substantially independent and consistent across a sample set that spans considerable range, from Nicosia (16,914 segments) to greater Atlanta (976,445 segments), more than 50 times in size.

As such, it is possible to determine the approximate scale of any real city without knowledge of the actual size either of its total footprint or any individual street segment; real cities are not ‘scale-free’, but deeply scale-dependent. The same can be done for artificial street networks, but when scaled based on intelligibility *RandTri* was found to have improbably long and uniform street segments, resembling only a foreground network of main roads, and thereby an extreme example of a single scale. The axial lines of such a grid, for example, approximate a normal distribution rather than a power law. Real cities, more complex, have also their background network of smaller streets which spans a range of scales and can appear to have a fractal nature for certain properties. But network intelligibility has a particular scale.

This scale of intelligibility appears to be related to what we consider to be the foreground network, as evidenced by its similarity to the rescaled *RandTri* (fig. 7). While the foreground has been observed at multiple scales (across a range of higher radii) this suggests that scale may be important. While the foreground has been treated as a morphological feature, this suggests a plausible link between this morphology and cognitive processes of navigation.

These results may explain the relationship between individual navigation strategies and urban form, and why Space Syntax techniques as different as centrality measures and agent simulation should both be effective in predicting movement. Syntactic intelligibility appears to operate at a particular radius around 2-4 km; at this scale, both a perfectly informed strategy of route minimisation (network centrality measures) and a locally opportunistic navigation (stochastic agent simulations) yield similar patterns. Empirical evidence indicates this is the scale at which both correlate with observed movement and average journey length. The networks examined can thus be considered to be optimised to convey the most relevant information on likely journeys to navigators throughout the network, and to do it at the most useful scale.

ACKNOWLEDGEMENTS

Thanks to Tasos Varoudis for the centrality (integration and choice) analyses of greater Atlanta, the largest of the maps, in DepthmapX.

REFERENCES

Barabasi AL and Albert R (1999) Emergence of scaling in random networks. *Science*, 286, 509–512.

Batty M (2006) Rank clocks. In *Nature*, 444, 592–596. doi: 10.1038/nature05302

Batty M and Longley P (1994) *Fractal Cities: A Geometry of Form and Function.* Academic Press, San Diego, CA and London.

Carvalho R and Penn A (2004) Scaling and universality in the micro-structure of urban space. *Physica A*, 332, pp. 539–547.

Chiaradia A, Hillier B, Barnes Y, Schwander C (2009) Residential property value patterns in London: space syntax spatial analysis. In: Koch D, Marcus L and Steen J (eds.) *(Proceedings) 7th International Space Syntax Symposium*. (pp. 15-) Royal Institute of Technology (KTH), Stockholm, Sweden.

DepthmapX development team (2017) *depthmapX* (Version 0.6.0) [Computer software]. Available at: https://github.com/SpaceGroupUCL/depthmapX/

Hanna S (2021) Random walks in urban graphs: A minimal model of movement, *Environment and Planning B: Urban Analytics and City Science*, 48(6), pp. 1697–1711. doi: 10.1177/2399808320946766.

Hanna S (2022) Route Choice from Local Information: Comparing Theories of Movement and Intelligibility. In: Van Nes A and de Koning RE (eds.) *Proceedings of the 13th Space Syntax Symposium*, Bergen, Norway, pp. 377:1–22.

Hillier B (1996) *Space is the Machine*. Cambridge.

Hillier B, Burdett R, Peponis J and Penn A (1987) Creating life, or, does architecture determine anything? *Architecture and Behavior/Architecture et Comportement*, 3, pp. 233–250.

Hillier B and Iida S (2005) Network and psychological effects in urban movement. In: Cohn AG and Mark DM (eds) *Proceedings of spatial information theory*, Ellicottville, NY, USA, 14–18 September 2005, pp.475–490. Berlin: Springer.

Hillier B, Penn A, Hanson J, Grajewski T, and Wu J (1993) Natural Movement: or, Configuration and Attraction in Urban Pedestian Movement. *Environment and Planning B*, vol (20), 29-66.

Jiang B (2009) Ranking spaces for predicting human movement in an urban environment, *International Journal of Geographical Information Science*, 23(7), 823–837. doi: 10.1080/13658810802022822.

Krenz K (2017) Regional Morphology: The Emergence of Spatial Scales in Urban Regions. In *Proceedings of the 11th Space Syntax Symposium*, Lisbon.

Law S, Paige B and Russell C (2019) Take a look around: using street view and satellite images to estimate house prices, *ACM Transactions on Intelligent Systems and Technology*, 10(5) pp. 1–19.

McElhinney S, Zisch F, Hornsberger M, Coutrot A, Spiers H and Hanna S (2022) Exosomatic Route Choice in Navigation: Evidence from video game player data. *Proceedings of the 13th Space Syntax Symposium*.

Penn A (2003) Space syntax and spatial cognition or why the axial line? *Environment and Behavior*, 35(1), 30–65.

Serra M and Hillier B (2019) Angular and metric distance in road network analysis: A nationwide correlation study. Computers, Environment and Urban Systems 74: 194–207.

Turner A (2007) The ingredients of an exosomatic cognitive map: isovists, agents and axial lines? In: Hölscher, C. and Conroy Dalton, R. and Turner, A., (eds.) *Space Syntax and Spatial Cognition: Proceedings of the Workshop held in Bremen, 24th September 2006.* pp. 163-180. Universität Bremen: Bremen, Germany.

Turner A (2009) The role of angularity in route choice: An analysis of motorcycle courier GPS traces. In: *COSIT'09 Proceedings of the 9th international conference on spatial information theory*, Aber Wrac'h, France, pp.489–504. Berlin: Springer.

Turner A and Penn A (2002) Encoding natural movement as an agent-based system: An investigation into human pedestrian behaviour in the built environment. Environment and Planning B: Planning and Design 29(4), pp. 473–490.

Zhang L, Chiradia A and Zhuang Y (2013) In The Intelligibility Maze of Space Syntax: a Space Syntax Analysis of Toy Models, Mazes and Labyrinths. In: Kim YO, Park HT and Seo KW (eds.) *Proceedings of the Ninth International Space Syntax Symposium*, Seoul.

Zipf GK (1949) *Human Behaviour and the Principle of Least Effort*, Cambridge, MA: Addison Wesley.