

中国人口流动网络中的语言纽带效应

赵梓渝¹, 王士君², 李卓伟², 韩钟辉³

(1. 青岛大学旅游与地理科学学院, 青岛 266071; 2. 东北师范大学地理科学学院, 长春 130024;
3. 中国海洋大学管理学院, 青岛 266100)

摘要: 本文利用2019年腾讯人口迁徙与中国语言分布数据, 揭示方言相似性对于中国人口流动影响的语言纽带效应, 并进一步探究其时空异质性、交通方式异质性、空间距离对其影响的调节作用。研究发现: ① 方言相似性能够促进中国人口城际流动网络的路径形成概率与关系强度, 形成了人口流动网络中的语言纽带效应。地理距离和语言距离分别可视为人口流动的显性和隐性空间成本; ② 语言纽带效应存在时空异质性和交通方式异质性。地区间的方言相似性每增加一个单位, 在“春运”和日常阶段下, 中国人口流动网络的路径形成概率分别增加47.7%、49.5%, 流动强度分别增加17.1%、9.9%。语言纽带效应对于不同汉语方言区、不同交通方式的人口流动的影响存在正负与程度差异, 展示出其对于中国人口流动影响的空间差异与地理含义; ③ 语言纽带效应受到空间衰减规律的制约, 并且在不同交通方式和时间段下呈现出差异性和复杂性。本文扩展了当前以经济、社会、人口学特征为主要视角的中国人口流动影响因素研究, 对于揭示中国人口流动驱动机制具有科学意义, 对于多民族融合的中国社会治理和文化多样性保护具有实践价值。

关键词: 人口流动; “春运”; 汉语方言; 语言纽带; 文化纽带; 中国

DOI: 10.11821/dlxb202404004

1 引言

人口流动是中国改革开放以来规模最大、意义最为深远的地理过程之一^[1]。社会与经济的区域发展差异对于人口流动影响的重要性与作用机制已被广泛论证^[2-4], 作为一个典型的“人情社会”和“熟人社会”^[5], 文化的空间相似性与差异性是如何影响中国的人口流动? 语言从古至今就是地区文化最重要的外显性符号, 相同或相近的语言是维系国家、民族、族群凝聚力的纽带。本文将基于方言相似性视角, 探究中国人口流动网络中的语言纽带效应。

文化是人类历史与社会的凝结和沉淀。特定文化代表了该区域群体的意识形态、价值观念、传统习俗、语言文字^[6]。现实世界中的文化分布具有空间异质性, 形成了有别于行政、制度等有形边界的无形边界, 同时不同地区的相同或相近文化又呈现关联性^[7]。文化相似性促进了人类的流动与迁移^[8-10], 形成了影响人口流动的文化纽带效应。

收稿日期: 2022-11-07; 修订日期: 2023-09-24

基金项目: 国家自然科学基金(42171198, 42001176); 教育部人文社会科学研究青年基金(20YJCZH241) [Foundation: National Natural Science Foundation of China, No.42171198, No.42001176; Youth Foundation of Humanities and Social Sciences of Ministry of Education, No.20YJCZH241]

作者简介: 赵梓渝(1986-), 男, 吉林长春人, 教授, 硕士生导师, 中国地理学会会员(S110017165M), 研究方向为城市网络与人口流动。E-mail: zhaozy@qdu.edu.cn

通讯作者: 王士君(1963-), 男, 黑龙江延寿人, 教授, 博士生导师, 主要从事经济地理学和城市地理学研究。
E-mail: wangsj@nenu.edu.cn

尽管文化是一个难以量化的概念,然而语言为揭示人口流动网络中的文化纽带作用提供了契机。语言是人类表达思想意识、描述客观事物的工具,也是人们在社会活动中身份认同的基石。作为具有 5000 余年文化渊源的古老国家,中国幅员辽阔、多民族融合,不同地区方言差异显著。方言多样性促进了社会多样性,同时深刻影响着中国人口流动的能动性与时空模式。

本文将从 3 个方面揭示中国人口流动网络中的语言纽带效应及其复杂性。① 人口流动网络由节点(城市)和路径(城市间人口流动关系)2 个要素构成。本文将基于 2019 年腾讯人口迁徙与中国语言分布数据,通过指数随机图模型(Exponential Random Graph Model, ERGM)、普通最小二乘法(Ordinary Least Squares, OLS)线性回归模型,分别检验方言相似性对于中国人口流动网络的路径形成概率与关系强度的影响;② 本文将通过中国春节运输(简称“春运”)人口迁徙和日常商旅通勤两种类型的人口流动、以方言区为研究单元的人口流动和区分航空、铁路、公路 3 种交通方式的人口流动,探究语言纽带效应对人口流动影响的流动类型差异、空间差异和交通方式差异;③ 本文将通过测度出行距离对语言纽带效应的调节作用,探究空间衰减规律对于语言纽带效应的影响。本文以方言相似性为代表的语言纽带效应及其复杂性,对于深入理解中国人口流动地理过程的驱动机制具有科学意义,对于理解人类流动性具有学术价值^[10],同时对于中国多民族融合的社会治理、保护文化多样性和制定区域协调发展策略具有实践价值。

2 文献综述

人口流动和迁移是影响中国当前与未来数十年经济社会发展的重大问题。城镇化转型中后期,中国人口流动进入总量稳定下的空间调整新阶段^[11]。已有研究一方面关注中国人口流迁的空间格局及其演化特征^[12-13];另一方面,基于户籍制度改革、区域发展差异与城乡差距、劳动报酬与经济收益、公共服务设施、空间距离、环境污染^[2-4, 11, 14]等方面探究中国人口流迁的影响因素与驱动机制。与此同时,随着信息化技术的发展,以百度、腾讯、高德为代表的关系型、高时空精度的人口流动大数据,在揭示中国人口流动的动态过程、空间格局及其演化、组织结构特征、交通方式与空间尺度异质性等方面展示出了重要的研究价值^[15-21]。相较于人口普查等统计型数据,关系型数据不仅能够揭示转型期中国人口流动的时空动态性与结构复杂性^[15],也促进了网络研究方法在人口流动研究中的应用^[15-17]。城市在网络中主要承担着枢纽作用,为人口流动网络研究提供了载体依托^[20-21]。

已有研究中,相对于经济、社会、人口学特征等视角,文化对于中国人口流动影响的程度与作用机制尚未得到充分的揭示^[8]。语言是文化的重要表征和影响人口流动的重要因素^[8-9]。在该领域语言经济学基于成本收益角度展开了论证。经济、非经济的收益与成本是人口流动的重要驱动力。一方面,语言作为一种人力资本^[22],相近的语言有助于流动者增强社会认同、获得信息优势、减少就业歧视^[23],从而显著提高他们的收入水平^[24];而语言差异会增加沟通成本并成为劳动力跨地区流动的非正式壁垒^[25]。在欧洲是否熟练掌握迁入地语言将影响流动人口获得平等工作和教育的机会^[26],这种效应即便在一个国家推广标准话的情况下仍然产生着长远影响^[27]。

另一方面,流动者面临着以精神成本为代表的非经济成本,其源自远离家乡、家庭和亲友的隐性成本,能够体现人口流动到文化陌生环境的心理代价与适应性。相近的语言意味着相近的文化背景、行为规范与思维方式,是促进外来者融入社会并实现有效互动的核心元素^[9],有助于流动者社会关系网络的重新构建和社会资本的积累^[28],进而增加

流动幸福感^[29]。正是因为这些因素的影响, 人口倾向于向语言相近、文化背景相似的地方流动迁移^[9, 30]。

结合已有研究, 本文将通过方言相似性探究语言纽带效应对于中国人口流动的影响, 并突出如下3个问题: ① 不同时间段的人口流动行为存在着性质上的差异^[12, 15, 21], 同时不同区域人口流动的空间特征与影响因素也存在着显著差异, 然而已有研究较少涉及方言或文化对于中国人口流动影响的流动类型差异性与时空差异性。② 随着近年来中国区域交通基础设施的建设与完善, 人口出行的交通方式选择日益多样化。不同交通方式下人口流动的组织结构与空间范围差异显著^[17], 显然也隐含了其背后驱动机制的差异。然而少有研究关注方言或文化对于不同交通方式下人口流动影响的差异性。③ 已有研究基于人口流动迁出地与迁入地的地区属性构建计量模型, 其结果丰富了经典推—拉理论的内涵, 然而人口流动作为一个具有组织复杂性的网络结构, 二元关系变量对网络结构特征的影响机制尚未得到充分揭示。综上所述, 本文将区分“春运”和日常两个研究时间段, 航空、铁路、公路3种不同交通方式, 基于网络研究视角探究中国人口流动网络中的语言纽带效应及其复杂性问题。

3 研究数据与方法

3.1 研究数据

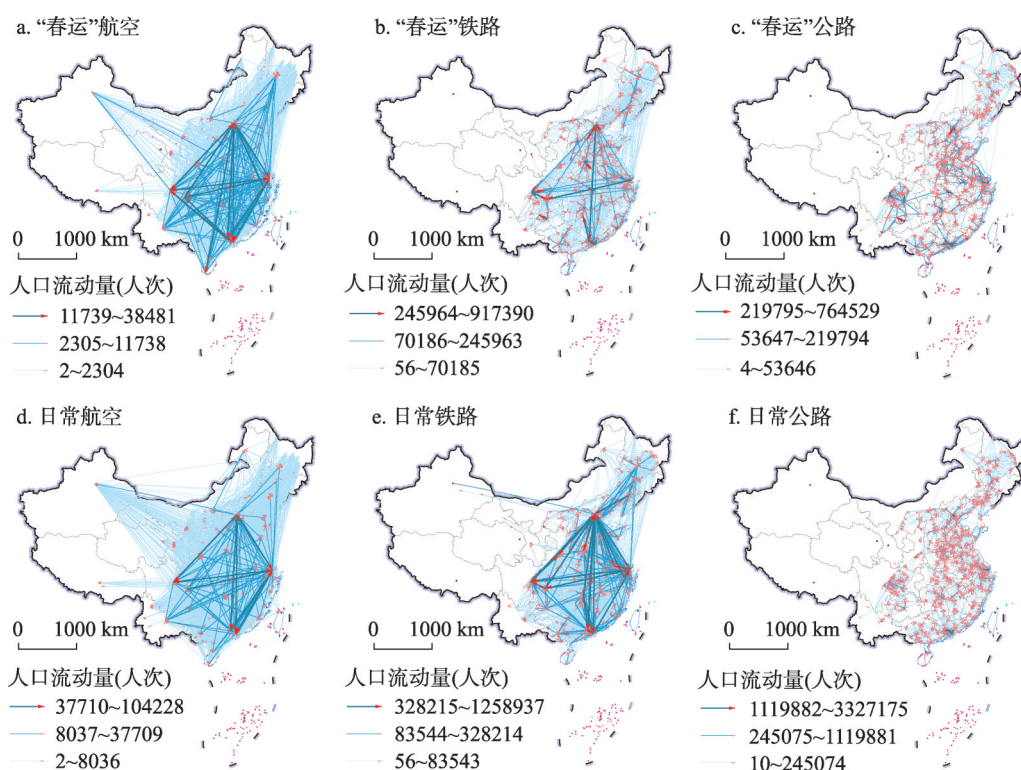
3.1.1 人口流动 人口流动数据来源于腾讯位置大数据 (<https://heat.qq.com/qianxi.php>), 其通过手机用户的定位信息映射人口流动轨迹, 记录了人口流动过程中的起止城市、流动强度、时间和交通方式(航空、铁路、公路), 数据所蕴含的丰富的人口流动信息, 对于中国人口流动的空间格局及其动态变化研究具有重要的学术价值^[16-18]。

腾讯人口迁徙以移动定位设备跨市辖区作为判定标准, 因此人口流动的空间统计单元为直辖市和地级行政区划。研究区域包括中国4个直辖市、香港特别行政区、澳门特别行政区及349个地级行政区(其中台湾省仅包括台北市), 共计355个地级及以上行政区划。鉴于2020年及以后COVID-19感染对人口流动的空间结构产生显著影响^[31], 本文选择2019年人口流动数据。考虑到一年中不同时段发生的人口流动存在类型差异, 选取“春运”和日常两个研究时间段, 分别代表“春运”人口迁徙和日常商旅通勤。其中, 通过“春运”返乡流(2019年1月21日—2月4日)研究“春运”时间段; 通过2019年6月人口流动研究日常时间段。腾讯人口迁徙展示了每个城市与之发生人口流动强度前10位城市的网络路径。研究数据共包含21778条、11.680亿人次人口流动(图1)。

3.1.2 方言及其相似性 语言类型及分区来源于《中国语言地图集》第2版^[32]。《中国语言地图集》由中国社会科学院语言研究所、中国社会科学院民族学与人类学研究所、香港城市大学语言资讯科学研究中心在1987年第1版的基础上编著, 全面表现中国汉语言和少数民族语言的分类与空间分布。

根据语言的语音、词汇和语法的地域差异, 中国可分为不同的方言区。《中国语言地图集》第2版中汉语方言划分为东北官话、北京官话、晋语、吴语等22个方言分区(表1), 每一个方言分区又可细划为不同分区小片, 以更精确化语言差异。同时, 使用《中国语言地图集》中的少数民族语言数据, 对文中所涉及的非汉语言地区进行信息补充, 以达到研究区域的全覆盖。

依据方言相似性建立语言纽带网络。本文建立5层语言关系矩阵, 按照联系强度依次为: 同一方言分区并处于同一小片的城市, 联系=5; 处于同一方言分区不同小片, 联



注：基于自然资源部标准地图服务网站GS(2020)4619号的标准地图制作，底图边界无修改。

图1 2019年“春运”节前、日常6月份中国人口流动空间分布

Fig. 1 Spatial distribution of China's population flows in the first 15 days of the Spring Festival travel rush and in June of 2019 (person-time)

系=4，尽管属于不同的语言小片，但同一方言区内部是大体上能够直接语言沟通的，因为他们同属于一种方言基础；方言官话之间，联系=3，官话具有大范围的适用性，其间存在一定的可交流程度；官话与非官话之间，联系=2；剩余城市之间，联系=1。相同的城市对则取高值。最后，将城市间的语言关系转为无向加权网络。

3.2 研究方法

3.2.1 网络构建 根据一天内各城市间人口流动数据构建关系矩阵 $T=(T_{ij})$ ， T_{ij} 为城市 i 到城市 j 的人口流权重。

$$T = \begin{matrix} & \begin{matrix} j_1 & j_2 & \cdots & j_{n-1} & j_n \end{matrix} \\ \begin{matrix} i_1 \\ i_2 \\ \vdots \\ i_{n-1} \\ i_n \end{matrix} & \begin{bmatrix} 0 & T_{12} & \cdots & T_{1(n-1)} & T_{1n} \\ T_{21} & 0 & \cdots & T_{2(n-1)} & T_{2n} \\ \vdots & \vdots & \ddots & \vdots & \vdots \\ T_{(n-1)1} & T_{(n-1)2} & \cdots & 0 & T_{(n-1)n} \\ T_{n1} & T_{n2} & \cdots & T_{n(n-1)} & 0 \end{bmatrix} \end{matrix} \quad (1)$$

3.2.2 子群结构 使用子群结构对人口流动网络的拓扑关系进行分析。子群结构可反映网络中个体局部性特征及其相互间的联系，并分析与预测整个网络各元素间的交互关系^[33]。同一子群的城市间具有密切的人口流动联系，而不同子群之间的联系相对较小。模块化是计算网络子群的常用算法，其公式如下：

$$Q = \sum_{s=1}^m \left[\frac{l_s}{L} - \left(\frac{d_s}{2L} \right)^2 \right] \quad (2)$$

表1 语言名称及其所辖区域
Tab. 1 Language name and its jurisdiction

序号	方言分区	分区小片	所辖行政区划
1	东北官话	哈阜片、黑松片、吉沈片	辽宁、吉林、黑龙江、青海、内蒙古
2	北京官话	京承片、朝峰片	河北、北京、辽宁、内蒙古、天津
3	冀鲁官话	保唐片、沧惠片、石济片	河北、北京、山东、山西、天津
4	胶辽官话	不分片	辽宁、黑龙江、山东
5	中原官话A	洛嵩片、漯项片、南鲁片等8个	河南、安徽、河北、山东、江苏、陕西
6	中原官话B	汾河片、关中片、河州片等5个	四川、陕西、甘肃、宁夏、青海、山西、河南
7	中原官话C	南疆片	新疆、甘肃
8	兰银官话	北疆片、河西片、金城片、银吴片	新疆、甘肃、宁夏、辽宁、湖南
9	江淮官话	洪巢片、黄孝片、泰如片	安徽、江苏、湖北、浙江、江西
10	西南官话A	桂柳片、湖广片	湖北、湖南
11	西南官话B	川黔片、川西片、湖广片、西蜀片	四川、陕西、重庆
12	西南官话C	川黔片、桂柳片、湖广片等5个	贵州、广西、云南
13	晋语	并州片、大包片、邯新片等8个	河南、内蒙古、山西、黑龙江、河北、陕西
14	吴语	金衢片、瓯江片、上丽片等5个	江苏、安徽、浙江、福建、上海、江西
15	闽语A	宣州片、闽北片、闽东片等7个	广东、福建、台湾、浙江、江苏、江西
16	闽语B	雷州片、琼文片	海南、广东
17	客家话	海陆片、宁龙片、汀州片等8个	湖南、广东、江西、安徽、福建、台湾、江苏、贵州
18	粤语	高阳片、勾漏片、广府片等7个	澳门、广西、广东、贵州、香港
19	湘语	辰溆片、衡州片、娄邵片等5个	湖南、广西、四川
20	赣语	昌都片、大通片、洞绥片等9个	安徽、湖南、江西、湖北、福建
21	徽语	绩歙片、旌占片、祁婺片等5个	安徽、浙江、江西
22	平话和土话	桂北片、桂南片、湘南片、粤北片	广西、湖南、广东

式中： Q 为模块度函数； m 代表计算出的子群数量； L 为网络中人口流动总量； l_s 为第 s 个子群的人口流动强度； d_s 为子群 s 中与各节点城市相关联的人口流量之和。

3.2.3 影响因素分析 本文通过ERGM探究方言相似性对于人口流动网络路径形成概率的影响。ERGM基于指数形式来研究网络结构，能够对影响网络结构和生长发育的各种因素进行综合分析和定量检验^[34]。

ERGM通过网络内生结构、节点属性和网络协变量的统计量来表示实际网络在所有随机网络中出现的概率。以本文为例，在ERGM中，人口流动网络被视为宏观网络作为因变量，局部网络结构和各种其他节点属性变量、外生网络变量作为自变量，通过概率函数推导某种节点属性或边连接效应，进而定量解释节点属性或外生网络对人口流动网络的影响效应。

ERGM通过R语言进行构造和测度。宏观城市网络中由部分城市组成的合集 $G=(V, E)$ ， $V=\{1, 2, \cdots, n\}$ 表示网络中城市合集（节点）， E 表示城市之间的连接关系（边）；令 Y 代表与 G 相关的邻接矩阵， Y_{ij} 代表 G 中存在一条由节点 i 指向节点 j 的边 $e \in E$ ， $y=[y_{ij}]$ 是 Y 的特定实现，若节点 i 与节点 j 发生人口流动，则 $y_{ij}=1$ ，否则 $y_{ij}=0$ 。

ERGM的一般设定形式如下：

$$Pr(Y=y|\theta)=P_{\theta}(y)=\frac{1}{k}\exp\left[\sum_H\theta_Hg_H(y)\right] \tag{3}$$

式中： $k=k(q)$ 是确保网络结构概率取值之和在0~1之间的参数常量； H 代表影响网络结构的各类影响因素，分别为内生结构因素、节点属性因素和外生网络因素； $g_H(y)$ 为影响网络结构因素 H 的统计量； θ_H 是相应的估计系数。

本文选取的解释变量具体如表2所示。① 内生结构变量。内生结构变量是用于检验网络自我组织特征的变量，为ERGM自带变量。边数是ERGM中最主要的内生结构变量，可以用来反映网络中关系形成的随机性。② 外生网络变量。选取方言网络（*dia*）和距离网络（*dis*）作为协变量，分别检验语言关系与空间距离对于人口流动网络路径形成概率的影响。人口流动遵从地理要素流动的空间衰减效应^[35]，文中的公路、铁路距离使用2019年中国基础公路交通网和铁路交通网数据进行最优路线求解，航空距离使用直线距离进行测度。③ 节点属性变量。选取4个指标为节点属性进行分析。第一，基于劳动报酬与经济收益视角，流动人口通过空间流动行为以谋求更高的收入，因此工资水平影响其流动目的地的选择，通过在岗职工平均工资（*wage*）来解释^[1]；第二，城市间社会经济发展水平差异也是人口流动的主要原因^[36]，理性的个体试图从发展中地区流向发达地区，通过人均地区生产总值（*PGDP*）来解释^[37]；第三，1980年以后新经济迁移理论强调家庭作为人口迁移决策主体的重要性，社会公共资源衡量了城市对于流动人口的接纳与安置能力^[35]。其中，以教育设施、教育投入为代表的公共服务水平的高低对流动人口的长期居留意愿具有显著影响^[1]，用普通高等学校数和中等职业教育学校数相加（*sch*）来表达这一变量；第四，建成区面积（*area*）作为城市规模的衡量指标，为流动主体提供了一种综合的预期效应。这种综合预期效应尽管不能直接体现工资水平与就业机会，但却构成了隐性的非现金收入和成本价值^[36]，因此用其衡量当前中国城市发展的规模效应及其的显著不平衡对于人口流动的影响。

表 2 解释变量及其描述
Tab. 2 Description of explanatory variables

机制类型	变量名称	变量内容	检验
内生结构变量	<i>edges</i>	边数	基础效应
外生网络变量	<i>dia</i>	城市间的方言相似性程度	关系网络对人口流动网络结构影响
	<i>dis</i>	人口出行距离(km)	
节点属性变量	<i>area</i>	行政区域建成区面积(km ²)	节点属性对于网络结构影响
	<i>wage</i>	在岗职工平均工资(元)	
	<i>PGDP</i>	人均地区生产总值(元)	
	<i>sch</i>	普通高校与中等职业学校之和(个)	

模型中城市节点属性的数据来源于《中国城市统计年鉴》和香港、澳门、台湾省统计年鉴。各年鉴部分计量单位和计量方法有所不同，均以《中国城市统计年鉴》为参考进行转化。部分缺失数据在各市（州、盟）统计年鉴补充。考虑到社会经济发展因素对于人口流动意愿影响的时间滞后性，因此使用2018年数据。

需要对变量选取进行补充说明：已有研究基于多角度揭示了中国人口流动影响因素，但本文的核心变量是方言相似性，同时考虑到是利用腾讯人口迁徙构建人口流动网络、选择“春运”和日常两种类型的人口流动、3种出行交通方式，并对各方言区人口流动进行逐一回归，因此控制变量选取所遵循的原则是，在已有研究中被广泛验证了对中国人口流动产生重要影响，而且这些变量应具有在不同数据、模型、时间或空间下中国人口流动影响研究中应用的广泛性与普遍解释力，而非特定时间、区域或特定人口流动类

型。同时，囿于《中国城市统计年鉴》与港澳台年鉴部分指标的统计差异、腾讯人口迁徙包括355个地级及以上行政区划，而部分自治州、县级市的统计数据只能通过政府公开数据进行查询，也在一定程度上限制了本文模型的变量选取。

综上所述，具体模型如下：

$$Pr(Y=y|\theta)=\frac{1}{k(\theta)}\exp\left(\frac{\theta_1 edges+\theta_2 dia+\theta_3 \ln dis+\theta_4 \ln area+\theta_5 \ln wage+\theta_6 \ln PGDP+\theta_7 \ln sch}{\theta_5 \ln wage+\theta_6 \ln PGDP+\theta_7 \ln sch}\right)$$

(4)

本文主要是探究网络节点属性变量和网络协变量对人口流动网络的影响，因此对模型结果采用极大伪似然估计法（Maximum Pseudo Likelihood Estimation, MPLE）进行估计检验。其原理是用给定网络其余部分的每个关系的条件概率的乘积代替ERGM所要估计的极大似然函数^[38]。在结果评价方面，采用 z 统计值检验参数的显著性。本文采用R语言中Statnet程序进行模型处理。

通过OLS线性回归模型检验方言相似性对于路径关系强度（*flow*）的影响，同时进一步区分城市节点作为人口流入地和流出地时的差异。节点变量与ERGM所采用的一致（表2）。因为对数变换不会影响原始变量之间的变化态势，且对数变化可消除异方差现象^[39]，所以除变量方言相似性（*dia*）外，对因变量和各解释变量均进行了对数变换。模型表示如下：

$$\ln flow = \beta_0 + \beta_1 dia + \beta_2 \ln dis + \beta_3 \ln O_area + \beta_4 \ln O_wage + \beta_5 \ln O_PGDP + \beta_6 \ln O_sch + \beta_7 \ln D_area + \beta_8 \ln D_wage + \beta_9 \ln D_PGDP + \beta_{10} \ln D_sch + \varepsilon$$

(5)

式中： β_0 为常数项； β_i 为各变量的回归系数； ε 为随机误差项。

4 结果分析

4.1 时空异质性

ERGM检验了自变量对于人口流动网络路径形成概率的影响，即中国城际人口流动是否发生在经济社会发展水平高、空间距离相近、方言相似性高的城市之间。模型中（表3），边数（*edges*）系数结果显著，说明网络并非是完全随机形成，探究人口流动网络影响因素是有意义的。方言相似性（*dia*）所表征的语言纽带效应对于中国人口流动网络的组织结构产生显著正向影响，即人口流动偏向于发生在方言相似性高的城市之间。城际方言相似性每增加一个单位，在“春运”和日常阶段下，人口流动发生概率分别增加47.7% $[\exp(0.390)-1 \approx 0.477]$ 和49.5% $[\exp(0.402)-1 \approx 0.495]$ 。

表3 ERGM回归结果
Tab. 3 Regression results of ERGM

变量	“春运”阶段			日常阶段		
	估计系数	标准误	z 值	估计系数	标准误	z 值
<i>edges</i>	-35.960***	1.387	-25.917	-34.106***	1.324	-25.750
<i>dia</i>	0.390***	0.021	18.420	0.402***	0.020	20.295
<i>ln dis</i>	-0.462***	0.028	-16.386	-0.538***	0.027	-20.250
<i>ln PGDP</i>	-0.044	0.038	-1.157	0.132***	0.035	3.786
<i>ln wage</i>	1.272***	0.076	16.691	1.039***	0.072	14.452
<i>ln sch</i>	0.381***	0.023	16.599	0.388***	0.021	18.168
<i>ln area</i>	0.633***	0.027	23.605	0.674***	0.025	26.820

注：***表示 $p < 0.01$ ，**表示 $p < 0.05$ ，*表示 $p < 0.10$ 。

OLS 检验了自变量对于人口流动网络路径关系强度的影响。模型中（表4），调整 R^2 分别为0.609、0.570， F 变化量分别为790.664、1109.852，可见回归模型具有较高的合理性；各变量 VIF 均小于10，即不存在明显的共线性问题。结果表明，首先，方言相似性对网络路径关系强度产生显著正效应影响，城际方言相似性每增加一个单位，在“春运”、日常阶段下，人口流动的路径关系强度分别增加17.1%、9.9%；其次，标准化系数显示，方言相似性变量的作用强度高于本文所选取的部分社会经济指标。例如“春运”期间人口迁出地的普通高校与中等职业学校之和（ $\ln O_sch$ ），人口迁入地的行政区域建成区面积（ $\ln D_area$ ）。这表明在后续相关研究中，应更加重视以语言为代表的文化因素的影响；最后，对比表3、表4可知，语言纽带的作用程度存在时间差异性，就网络路径形成概率而言，在日常阶段高于“春运”阶段，而在路径关系强度方面则相反。

进一步，通过对各方言区人口流动网络分别进行 OLS 回归（表5）。结果表明，语言纽带效应对于不同地理位置分布的方言区人口流动的影响呈现出差异性。以返乡流为代

表 4 OLS 回归结果
Tab. 4 Regression results of OLS

变量	“春运”阶段				日常阶段			
	B	标准误	Beta	<i>t</i>	B	标准误	Beta	<i>t</i>
常数项	-13.034***	1.648		-7.907	-19.686***	1.423		-13.836
<i>dia</i>	0.171***	0.021	0.110	8.186	0.099***	0.017	0.061	5.893
<i>Indis</i>	-1.647***	0.030	-0.782	-54.488	-1.898***	0.026	-0.802	-72.379
<i>lnO_area</i>	0.363***	0.044	0.134	8.294	0.261***	0.036	0.090	7.165
<i>lnO_wage</i>	1.588***	0.150	0.141	10.571	1.104***	0.122	0.092	9.086
<i>lnO_PGDP</i>	-0.006	0.062	-0.001	-0.101	0.007	0.051	0.001	0.134
<i>lnO_sch</i>	0.268***	0.037	0.092	7.140	0.479***	0.031	0.154	15.207
<i>lnD_area</i>	0.164***	0.047	0.083	3.509	0.591***	0.038	0.273	15.456
<i>lnD_wage</i>	1.393***	0.127	0.192	10.986	2.425***	0.109	0.298	22.325
<i>lnD_PGDP</i>	-0.641***	0.073	-0.162	-8.771	-0.537***	0.060	-0.127	-9.003
<i>lnD_sch</i>	0.259***	0.040	0.110	6.460	0.034	0.033	0.013	1.036

注：***表示 $p < 0.01$ ，**表示 $p < 0.05$ ，*表示 $p < 0.10$ 。

表 5 不同方言区回归标准化系数(Beta)
Tab. 5 Standardized coefficient of OLS regression (Beta) in different linguistic regions

序号	方言分区	“春运”阶段	日常阶段	序号	方言分区	“春运”阶段	日常阶段
1	东北官话	0.079*** (3.455)	0.115*** (6.366)	12	晋语	0.238*** (10.724)	0.172*** (9.485)
2	北京官话	0.130*** (4.691)	0.082*** (3.676)	13	吴语	0.113*** (6.486)	0.002 (0.170)
3	冀鲁官话	0.135*** (6.649)	0.102*** (6.419)	14	闽语 A	0.109*** (5.113)	0.020 (1.100)
4	胶辽官话	0.143*** (5.082)	0.147*** (7.555)	15	闽语 B	0.323*** (8.483)	0.303*** (9.104)
5	中原官话 A	0.095*** (5.01)	-0.023 (-1.428)	16	客家话	0.039** (2.373)	-0.049*** (-3.376)
6	中原官话 B	0.128*** (4.045)	0.038 (1.377)	17	粤语	0.160*** (7.945)	0.075*** (4.779)
7	兰银官话	0.133*** (3.747)	0.069** (2.386)	18	湘语	0.143*** (5.567)	-0.013 (-0.577)
8	江淮官话	0.102*** (6.984)	0.028* (2.157)	19	赣语	0.088*** (4.366)	-0.013 (-0.730)
9	西南官话 A	0.118*** (4.666)	0.033 (1.491)	20	徽语	0.242*** (8.149)	0.109*** (4.034)
10	西南官话 B	0.084*** (4.073)	0.003 (0.161)	21	平话和土话	0.156*** (4.622)	0.155*** (7.018)
11	西南官话 C	0.218*** (9.271)	0.100*** (5.293)				

注：***表示 $p < 0.01$ ，**表示 $p < 0.05$ ，*表示 $p < 0.10$ ；括号内为 t 值；受人口流量小因素影响，中原官话 C 模型无法计算。

表的“春运”阶段, 方言相似性对方言区人口流动强度均产生显著的正效应影响。在日常阶段, 中原A、中原B、西南A、西南B、吴语、闽语A、湘语、赣语的方言相似性影响结果不显著, 这些方言主要分布在安徽、江苏、陕西、四川、湖南、福建、江西等省份; 在显著的方言区中, 除了客家话以外均为正效应。客家话主要分布于粤东、粤北、赣南、闽西和珠江口东岸等地区, 客家话结果为显著的负效应, 即人口更趋向于前往方言差异性较大区域, 这与客家人拼搏和开放的精神有着密切关系。这一结果反映出中国人口流动行为的地域差异和文化背景差异, 揭示了语言纽带效应在影响中国人口流动的过程中, 所呈现出的空间差异与地理含义。

以上结果揭示了中国人口流动网络中的语言纽带效应的时空异质性。中国人口在一年中不同时间段的出行存在类型差异, 例如“春运”流动人口迁徙、日常商旅通勤、以旅游为主的节日出行(劳动节与国庆节), 人口流动的类型差异影响其目的地选择, 进一步导致空间结构特征的差异。就“春运”时段而言, 1978年改革开放以来, 中国区域社会经济差异的巨大差异导致大规模劳动力采取用脚投票的方式长距离流动迁移, 劳动力在不同区域和城市间的集聚与扩散是对地方经济地理功能的回应, 尤其是对以制造业为代表的劳动密集型产业集聚产生的劳动力市场需求的回应^[20], 而“春运”这一中国特有的重大社会现象为流动人口迁徙研究提供了契机。“春运”期间的人口流动往往发生在故乡和工作地之间, 出行目的是探亲 and 返回工作地。上述研究结果表明, “春运”阶段, 无论是全部人口流还是分方言区人口流, 由方言相似性代表的语言纽带效应, 对于中国人口流动网络结构均产生了显著的正效应影响。一方面, 中国流动人口在外务工迁徙的空间选择上, 更偏向于方言相似性高的地区。这不仅仅意味着社会交往、日常工作中沟通成本的降低。同时, 相近的语言和文化也意味着相近的行为规范与思维方式, 这将使流动者在迁入地更容易获得归属感与认同感^[29], 有助于其社会关系网络的重建和社会资本的积累; 另一方面, 先迁移的流动人口会带动故乡的亲友迁移到相同目的地, 这种链式迁移可以减低迁移过程中的风险和成本, 即人口在流动过程中受到相近文化影响的依赖性迁移现象^[40]。

本文选择了2019年6月份作为日常时段, 该期间的人口流动以商旅出行、工作通勤、物流运输为主。人口流动是生产要素的空间再配置过程, 在日常时间段, 各种企业及组织间以价值创造为导向的横向、纵向的正式与非正式的联系构成了人口流动网络存在的基础^[20]。全部日常人口流模型结果表明, 方言相似性将促进城际人口流动路径形成概率的发生和流动强度的增加。进一步, 在考虑地理差异时, 21个方言区的语言纽带效应呈现出作用的空间异质性。方言距离不同于空间距离。相较于空间距离的作用普遍呈幂律形式^[35], 方言距离具有复杂性。中国当前的方言分布格局, 是地理环境与历史变迁的综合作用结果。例如, 五岭这一地理屏障分隔了湘语和粤语, 体现了地理因素对方言分布的影响^[41]。以东北官话、胶辽官话、西南官话C、晋语、闽语B、徽语、平话和土话等为代表的方言区, OLS回归模型中的较高标准化系数结果表明, 以上区域人口流动过程中, 方言相似性已经成为重要的影响因素。

4.2 交通方式异质性

表6分别从中国人口流动网络的路径形成概率与关系强度两个角度, 验证了以方言相似性为代表的语言纽带效应的交通方式异质性问题。结果表明, 方言相似性对航空流网络产生负效应影响, 对铁路、公路流网络均产生正效应影响。该效应在“春运”与日常阶段的网络路径形成概率与关系强度中均得到了一致性的检验。具体而言, 城市间方言相似性每增加一个单位, 在“春运”和日常时间段, 航空流网络的发生概率分别减少

表6 不同交通方式下 ERGM 和 OLS 回归结果
Tab. 6 Regression results of ERGM and OLS under different travel modes

模型	变量	“春运”阶段			日常阶段		
		航空	铁路	公路	航空	铁路	公路
ERGM	<i>edges</i>	-48.878*** (1.640)	-33.285*** (2.281)	10.111** (3.357)	-45.868*** (1.541)	-32.223*** (1.741)	10.173** (3.357)
	<i>dia</i>	-0.178*** (0.028)	0.454*** (0.035)	0.609*** (0.046)	-0.080** (0.026)	0.238*** (0.026)	0.607*** (0.046)
	<i>Indis</i>	1.135*** (0.039)	-2.044*** (0.056)	-3.812*** (0.103)	1.162*** (0.036)	-1.760*** (0.041)	-3.814*** (0.103)
	节点变量	Yes	Yes	Yes	Yes	Yes	Yes
OLS	<i>dia</i>	-0.107*** (-5.367)	0.112*** (4.569)	0.238*** (8.182)	-0.049*** (-2.986)	0.062*** (2.929)	0.239*** (9.733)
	<i>Indis</i>	-0.03 (-0.632)	-1.013*** (-29.063)	-1.668*** (-36.115)	-0.049 (-1.313)	-0.845*** (-24.710)	-1.857*** (-43.454)
	节点变量	Yes	Yes	Yes	Yes	Yes	Yes
	常数项	-26.563*** (-15.502)	-2.71 (-1.473)	-6.273*** (-3.187)	-32.359*** (-21.947)	-13.438*** (-7.405)	-5.873*** (-3.299)

注：***表示 $p < 0.01$ ，**表示 $p < 0.05$ ，*表示 $p < 0.10$ ；ERGM 括号内为标准误，OLS 模型括号内为 t 值。

16.3%、7.7%，人口流动强度分别减少 10.7%、4.9%；铁路流网络的发生概率分别增加 57.5%、26.9%，流动强度分别增加 11.2%、6.2%；公路流网络的发生概率分别增加 83.9%、83.5%，流动强度分别增加 23.8%、23.9%。方言相似性对于人口流动网络的影响程度总体上呈现公路>铁路>航空的特征。

这一结果与不同交通方式下的人口出行距离、出行目的以及客流规模的差异有着密切关系。对此，本文测算了航空、铁路、公路人口流动网络的模块化指数，“春运”阶段分别为 0.003、0.548、0.768，日常阶段分别为 0.008、0.284、0.740。结合图 2 中子群结构，航空交通主要承担远距离出行，最具跨区域交织能力^[17]，跨区域的语言与文化差异可能会带来一定的知识和信息交换，促进了以商务合作、创新创业和信息交流为目的的城际出行规模的增大^[42]，这导致方言相似性因素呈负效应影响且作用程度低；而铁路、公路通常承担中、短距离的人口出行，尤其是公路流呈现显著的空间衰减效应和边界约束，局部地方效应最为显著^[17]，形成了较为明显的子群结构（图 2c、2f），更贴近汉语方言分区所覆盖的空间范围。例如，粤港澳大湾区、长三角、成渝都市圈、东北地区的人口流动网络所形成的紧密联系“小世界”，分别对应粤语、吴语、西南官话 B、东北官话的大体空间覆盖范围，导致方言相似性因素呈正效应影响且作用程度高。

4.3 调节效应

为探究方言相似性因素是否受到空间衰减规律的制约，本文通过交互项以捕捉地理距离对于语言纽带效应的调节作用。在网络路径形成概率方面，本文通过在 1000 次模拟网络中随机选取 1 个人口流动网络进行分析，以验证语言纽带效应的稳定性。针对所选取的“春运”、日常阶段下全部和不同交通方式的 8 个人口流动网络，计算各观测网络（初始网络）变量的平均值与模拟网络中对应变量平均值的比值。结果表明，8 个模型的观测值/模拟值介于 0.992~1.005，表明模拟网络具有较高的准确性。在此基础上，对选取的模拟网络中加入 $dia \times Indis$ 变量进行 ERGM 回归。

回归结果如表 7 所示。在全部人口流的两个研究时间段下，方言相似性与空间距离的交互项（ $dia \times Indis$ ）系数在 1% 的水平上显著为负，即空间距离削弱了方言相似性对于

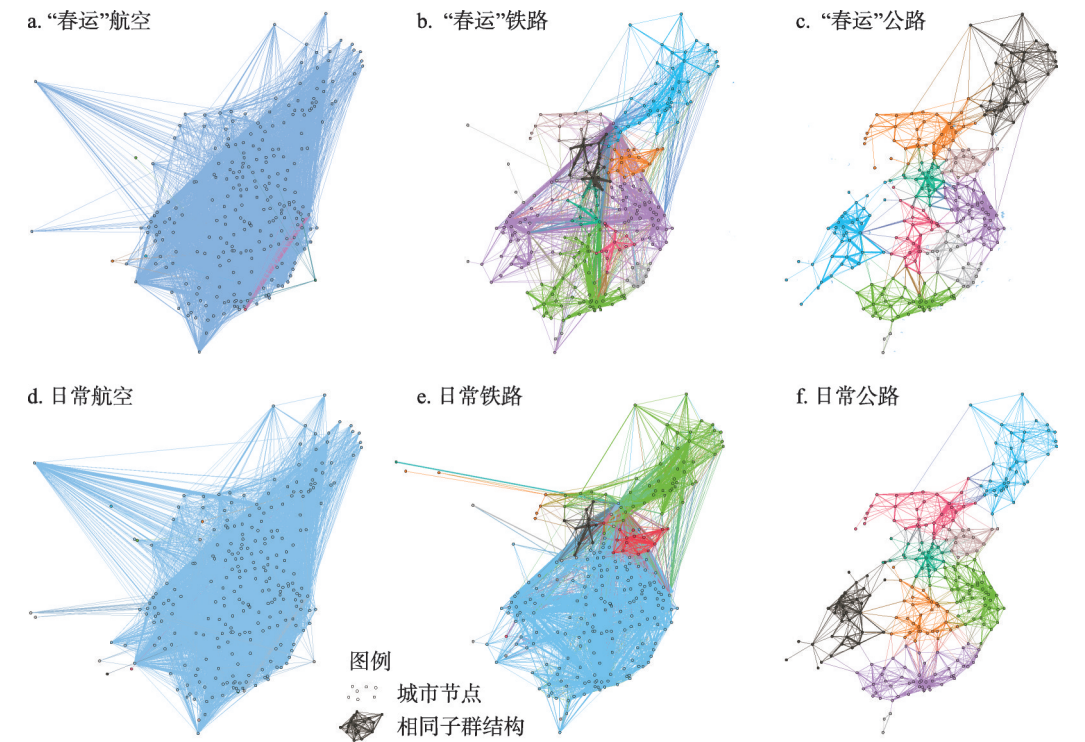


图2 不同交通方式人口流动网络的子群结构

Fig. 2 The subgroup structure of population flow network under different travel modes

表7 网络路径形成影响因素ERGM交互结果

Tab. 7 ERGM interaction results of the influencing factors of network path formation

变量	全部人口流		航空		铁路		公路	
	“春运”阶段	日常阶段	“春运”阶段	日常阶段	“春运”阶段	日常阶段	“春运”阶段	日常阶段
<i>edges</i>	-32.890*** (1.367)	-31.978*** (1.312)	-38.476*** (1.576)	-35.177*** (1.481)	-38.847*** (2.407)	-40.176*** (1.764)	-9.628*** (3.370)	-21.602*** (3.172)
<i>dia</i>	0.214*** (0.024)	0.261*** (0.023)	-0.174*** (0.028)	-0.121*** (0.026)	0.420*** (0.051)	0.184*** (0.035)	1.004*** (0.120)	0.861*** (0.109)
<i>lndis</i>	-0.275*** (0.032)	-0.361*** (0.030)	0.987*** (0.040)	1.088*** (0.037)	-2.036*** (0.069)	-1.717*** (0.046)	-4.210*** (0.161)	-4.328*** (0.144)
<i>dia×lndis</i>	-0.287*** (0.022)	-0.269*** (0.022)	0.130*** (0.037)	0.108*** (0.033)	-0.096*** (0.042)	-0.083*** (0.031)	0.317*** (0.094)	0.254*** (0.089)
节点变量	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
<i>AIC</i>	23121	25767	17030	19791	7565	13278	4083	4533
<i>BIC</i>	23188	25834	17097	19858	7632	13345	4150	4600

注：***表示 $p < 0.01$ ，**表示 $p < 0.05$ ，*表示 $p < 0.10$ ；括号内为标准误。为避免多重共线性问题，对交互项中的各变量进行了中心化处理。

人口流动网络路径形成概率的正向作用。进一步，就不同交通方式而言，空间距离对方言相似性的调节效应依然存在。空间距离削弱了方言相似性对于航空流网络路径形成概率的负向作用、对于铁路流网络路径形成概率的正向影响。在公路流网络中，方言相似性与空间距离交互项的系数显著为正，空间距离强化了方言相似性对于人口流动网络路径形成概率的正向作用。

在人口流动网络的路径关系强度方面（表8），在全部人口流的两个研究时间段下，空间距离削弱了方言相似性对于人口流动网络路径关系强度的正向作用，在“春运”和日常阶段，当空间距离分别超过1764.52 km和4488.85 km后，方言相似性对于人口流动强度的总效应变为负向作用。就不同交通方式而言，空间距离加强了两个研究时间段下方言相似性对于航空流网络路径关系强度的负向影响；空间距离削弱了“春运”阶段下方言相似性对铁路流网络路径关系强度的正向影响，当空间距离超过3370.02 km后，方言相似性对于人口流动强度的总效应变为负向作用；空间距离加强了日常阶段下方言相似性对于铁路流和公路流网络路径关系强度的正向影响。

表 8 路径关系强度影响因素 OLS 交互结果

Tab. 8 OLS interaction results of the influencing factors of the strength of path relationships

变量	全部人口流动		航空		铁路		公路	
	“春运”阶段	日常阶段	“春运”阶段	日常阶段	“春运”阶段	日常阶段	“春运”阶段	日常阶段
<i>dia</i>	0.205*** (9.744)	0.100*** (5.965)	-0.144*** (-6.988)	-0.083*** (-4.767)	0.138*** (5.293)	0.051** (2.368)	0.231*** (7.374)	0.201*** (7.318)
<i>Indis</i>	-1.605*** (-53.044)	-1.869*** (-67.726)	0.056 (1.142)	0.024 (0.629)	-1.027*** (-29.224)	-0.867*** (-24.763)	-1.669*** (-36.108)	-1.875*** (-43.542)
<i>dia</i> × <i>Indis</i>	-0.173*** (-9.796)	-0.053*** (-3.345)	-0.249*** (-6.273)	-0.166*** (-5.733)	-0.058*** (-2.904)	0.058*** (2.880)	0.025 (0.554)	0.136*** (3.144)
节点变量	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
常数项	-21.542*** (-13.205)	-31.269*** (-22.063)	-26.548*** (-16.302)	-32.129*** (-22.678)	-7.378*** (-3.972)	-19.278*** (-10.498)	-14.225*** (-7.126)	-15.478*** (-8.610)
调整 <i>R</i> ²	0.616	0.571	0.605	0.548	0.566	0.414	0.509	0.565

注：***表示 $p < 0.01$ ，**表示 $p < 0.05$ ，*表示 $p < 0.10$ ；括号内为*t*值。为避免多重共线性问题，对交互项中的各变量进行了中心化处理。

以方言相似性为代表的语言纽带效应影响了中国人口流动网络的路径形成概率与路径关系强度，叠加空间距离的作用后，该效应产生了复杂作用（图3）。语言纽带效应对于人口流动网络的路径形成概率与路径关系强度均产生了正向的扩张作用，即网络组织的繁盛发展；而地理距离对于语言纽带效应产生了空间制约影响，即随着空间距离的增加，语言纽带效应影响下的人口流动网络组织呈现由繁盛向衰退的结构变化。

就不同交通方式而言，在航空流网络中，文化纽带对其路径形成概率与关系强度均呈现负向的收缩效应，而空间距离分别削弱、加强了这一负效应，即随着空间距离的增加，语言纽带效应影响下的人口流动网络组织呈现由衰退向分散的结构变化。在公

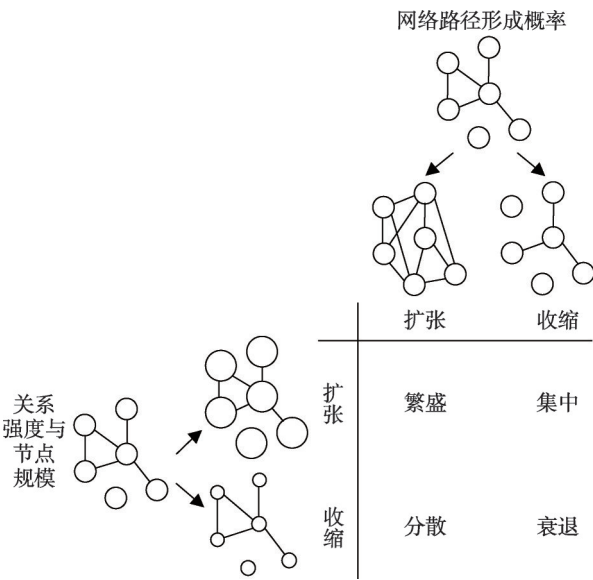


图 3 网络结构变化分解

Fig. 3 Decomposition of network structure changes

路流网络中,文化纽带对其网络结构呈现正向的扩张效应,而空间距离强化了该效应,即网络组织结构向繁盛发展。铁路流网络中,在“春运”阶段,文化纽带对其网络结构呈现正向的扩张效应,而空间距离削弱了这一效应,即铁路流网络组织由繁盛向衰退变化;在日常阶段,空间距离削弱了文化纽带对于网络路径形成概率的正效应,同时强化了对于关系强度的正效应,即网络组织结构由繁盛向集中变化。以上分析表明,一方面,地理学为深入探究中国人口流动行为、过程、现象背后的文化因素的复杂性与多样性提供了学科视角与基础支撑;另一方面,基于网络维度的分析,能够加深对于中国人口流动地理过程的理解,有助于认知其背后的行为机制与组织逻辑。

5 结论与讨论

5.1 结论

以语言为代表的文化的空间关联与差异塑造着人类流动。汉语方言相似性所产生的语言纽带效应增加了中国人口城际流动网络的路径形成概率与关系强度,因此语言距离与地理距离共同构成了中国人口城际流动的空间成本。尽管20世纪50年代中国政府开始推广普通话,然而方言对于人口流动影响的纽带效应至今仍产生着长远的影响。方言相似性因素对于人口流动的影响程度高于部分本文所选择的社会经济指标,展示出以语言为代表的文化因素对中国人口流动产生了重要的影响。这一结果拓展了对于中国人口迁移流动空间格局及其影响因素的规律性认识。

由方言相似性所表征的语言纽带效应存在显著的时空异质性和交通方式异质性。城市间方言相似性每增加一个单位,“春运”和日常阶段下,人口城际流动的概率分别增加47.7%、49.5%,人口流动强度分别增加17.1%、9.9%;方言相似性对于不同地理位置分布的方言区人口流动的影响存在多样性与差异性,反映了文化在塑造中国人口流动过程中所呈现的空间差异与地理含义;就交通方式异质性而言,方言相似性对航空流产生负效应,对铁路流、公路流产生正效应影响,这与不同交通方式的出行距离、出行目的及其流动网络结构密切相关。

地理距离总体上制约了语言对于中国人口流动影响的纽带效应。空间距离均削弱了方言相似性对于人口流动网络的路径形成概率与路径关系强度的正向作用,在“春运”和日常阶段,当距离分别超过1764.52 km和4488.85 km后,方言相似性对于人口流动强度的总效应变为负向作用。同时随着空间距离的增加,语言纽带效应影响下的人口流动网络组织呈现出多样性的结构变化,这一发现有助于揭示中国人口流动网络中语言纽带效应的复杂性。

5.2 讨论

理解文化对于人口流动的影响具有现实意义。语言趋同性对于人类社会的发展是一把双刃剑。一方面,以方言为代表的区域文化差异和沟通障碍抑制了人口流动迁移,阻碍了人力资本在空间上的有效配置^[9]。已有研究指出,方言多样性会强化市场分割^[43],使流动人口在流入地劳动力市场的融入过程中遭遇摩擦,降低外来劳动力收入和人均产出^[24]。因此,促进语言的趋同性对于增强人口要素流动、优化人力资本的空间配置、促进区域一体化发展和构建全国统一市场具有重要意义。另一方面,增加语言趋同性也将会导致文化多样性的丧失。不同地区人口的流入,增加了流入地的语言多样性与文化多样性。随着流动人口对于地方语言的学习与掌握,尽管这有利于流入人口更好地适应迁入地的生活环境并改善生活际遇,但自然也会导致以语言多样性为代表的文化多样性日益消失。

因此,探究文化在人口流动过程中的作用,对于多民族融合的中国社会治理具有重要意义。语言纽带效应的复杂性及其背后所反映的人口流动问题,提醒城市管理者与政策制定者,谨慎部署人口流动管理的相关措施,通过鼓励跨方言区域人口流动和尊重文化多样性等公共政策,在保障人力资本有效配置的同时,促进文化多样性的和谐发展。

最后需要对研究数据进行补充说明。人口流动是一种大规模的群体性时空行为,受制于统计时间、结构化内容限制,人口普查数据难以揭示人口流动空间性、社会性的复杂特性和时空动态性,以及中国快速城镇化背景下城市间日益复杂的交互关系;而社会调查数据的样本数量较难达到大规模的量级。应用个体粒度的人口流动海量时空轨迹数据,得以对人类时空行为进行综合、高精度、连续性的观测,为揭示人口流动的空间性与社会性提供了基础数据支撑。尽管人类的社会活动存在显著的不确定性,但这种随机性和不可预见性的范围是有限的。当累积足量的个体空间行为轨迹,数据中的群体特征即显现出来^[31]。

地理时空大数据一方面让我们能够在更精细的粒度下对人口流动这一行为进行近乎全覆盖的统计推断,在总体层面上评估文化多样性在人类流动模式形成中的潜在作用;另一方面,以腾讯人口迁徙为代表的人口流动大数据同样存在一定的研究限制与不足。首先,数据只包含城际人口流动总量,不带有流动者的人口学特征、社会属性、流动目的,无法识别流动性质或目的,“春运”和日常是基于中国发展背景下,依托特定时段和事件下识别人口流动的宏观问题;其次,由于移动数据只记录了特定人群的时空轨迹(如腾讯大数据迁徙只采集了使用腾讯定位服务手机用户的数据),使其识别的人口流动强度将低于实际情况。尽管该数据不能涵盖所有流动人口,但其占比远超过常规流动人口抽样调查的比例^[44];最后,移动定位数据的采集方式可能会导致部分中转等原因用户的行程被拆解^[45],一定程度上增加了人口流动强度、扩大了交通枢纽城市的人口流动规模。

参考文献(References)

- [1] Zhu Yu, Lin Liyue. Studies on the temporal processes of migration and their spatial effects in China: Progress and prospect. *Scientia Geographica Sinica*, 2016, 36(6): 820-828. [朱宇, 林李月. 中国人口迁移流动的时间过程及其空间效应研究: 回顾与展望. *地理科学*, 2016, 36(6): 820-828.]
- [2] Shen J F. Changing patterns and determinants of interprovincial migration in China 1985-2000. *Population, Space and Place*, 2012, 18(3): 384-402.
- [3] Zhang W L, Chong Z H, Li X J, et al. Spatial patterns and determinant factors of population flow networks in China: Analysis on Tencent Location Big Data. *Cities*, 2020, 99: 102640. DOI: 10.1016/j.cities.2020.102640.
- [4] Gu Hengyu, Shen Tiyan, Liu Ziliang, et al. Driving mechanism of interprovincial population migration flows in China based on spatial filtering. *Acta Geographica Sinica*, 2019, 74(2): 222-237. [古恒宇, 沈体雁, 刘子亮, 等. 基于空间滤波方法的中国省际人口迁移驱动因素. *地理学报*, 2019, 74(2): 222-237.]
- [5] Fei Xiaotong. *From the Soil*. Beijing: People's Publishing House, 2008. [费孝通. 乡土中国. 北京: 人民出版社, 2008.]
- [6] Lambert B, Kontonatsios G, Mauch M, et al. The pace of modern culture. *Nature Human Behaviour*, 2020, 4: 352-360.
- [7] Wang H R, Zhang H P, Tang G A, et al. Inter-city association pattern recognition by constructing cultural semantic similarity network. *Transactions in GIS*, 2022, 26(5): 2225-2243.
- [8] Wu W J, Wang J H, Dai T S. The geography of cultural ties and human mobility: Big data in urban contexts. *Annals of the American Association of Geographers*, 2016, 106(3): 612-630.
- [9] Huang Zongye, Yang Jing. The impact of dialect on inter-provincial migration in China. *Population Research*, 2020, 44(4): 89-101. [黄宗晖, 杨静. 方言对省际人口迁移的影响. *人口研究*, 2020, 44(4): 89-101.]
- [10] Lee S H, Ffranco R, Abrams D M, et al. Matchmaker, matchmaker, make me a match: Migration of populations via marriages in the past. *Physical Review X*, 2014, 4(4): 041009. DOI: 10.1103/PhysRevX.4.041009.

- [11] Liu Tao, Zhuo Yunxia, Wang Jiejing. How multi-proximity affects destination choice in onward migration: A nested logit model. *Acta Geographica Sinica*, 2020, 75(12): 2716-2729. [刘涛, 卓云霞, 王洁晶. 邻近性对人口再流动目的地选择的影响. *地理学报*, 2020, 75(12): 2716-2729.]
- [12] Pan Jinghu, Lai Jianbo. Research on spatial pattern of population mobility among cities: A case study of "Tencent Migration" big data in "National Day-Mid-Autumn Festival" vacation. *Geographical Research*, 2019, 38(7): 1678-1693. [潘竞虎, 赖建波. 中国城市间人口流动空间格局的网络分析: 以国庆—中秋长假和腾讯迁徙数据为例. *地理研究*, 2019, 38(7): 1678-1693.]
- [13] Ke Wenqian, Zhu Yu, Chen Chen, et al. Spatio-temporal evolution of migration in China from 1995 to 2015. *Acta Geographica Sinica*, 2022, 77(2): 411-425. [柯文前, 朱宇, 陈晨, 等. 1995-2015年中国人口迁移的时空变化特征. *地理学报*, 2022, 77(2): 411-425.]
- [14] Cao Guangzhong, Liu Jiajie, Liu Tao. Examining the role of air quality in shaping the landscape of China's internal migration: Phase characteristics, push and pull effects. *Geographical Research*, 2021, 40(1): 199-212. [曹广忠, 刘嘉杰, 刘涛. 空气质量对中国人口迁移的影响. *地理研究*, 2021, 40(1): 199-212.]
- [15] Zhao Ziyu, Wang Shijun. A spatial-temporal study of inter-provincial migration pattern during Chinese Spring Festival travel rush. *Population Research*, 2017, 41(3): 101-112. [赵梓渝, 王士君. 2015年我国春运人口省际流动的时空格局. *人口研究*, 2017, 41(3): 101-112.]
- [16] Wang Lucang, Liu Haiyang, Liu Qing. China's city network based on Tencent's migration big data. *Acta Geographica Sinica*, 2021, 76(4): 853-869. [王录仓, 刘海洋, 刘清. 基于腾讯迁徙大数据的中国城市网络研究. *地理学报*, 2021, 76(4): 853-869.]
- [17] Pan Jinghu, Wei Shimei, Zhang Rong, et al. Spatial structure characteristics of intercity travel network of Chinese residents: Based on Tencent migration data. *Acta Geographica Sinica*, 2022, 77(10): 2494-2513. [潘竞虎, 魏石梅, 张蓉, 等. 中国居民城际出行网络的空间结构特征. *地理学报*, 2022, 77(10): 2494-2513.]
- [18] Zhang Rong, Pan Jinghu, Lai Jianbo. Characteristics of intercity trip network structure of residents under different traffic modes: A case study of Spring Festival travel rush. *Progress in Geography*, 2021, 40(5): 759-773. [张蓉, 潘竞虎, 赖建波. 不同交通方式下居民城际出行网络结构特征: 以“春运”为例. *地理科学进展*, 2021, 40(5): 759-773.]
- [19] Liu Wangbao, Shi Enming. Spatial pattern of population daily flow among cities based on ICT: A case study of "Baidu Migration". *Acta Geographica Sinica*, 2016, 71(10): 1667-1679. [刘望保, 石恩名. 基于ICT的中国城市间人口日常流动空间格局: 以百度迁徙为例. *地理学报*, 2016, 71(10): 1667-1679.]
- [20] Zhao Ziyu, Wei Ye, Pang Ruiqiu, et al. Alter-based centrality and power of Chinese city network using inter-provincial population flow. *Acta Geographica Sinica*, 2017, 72(6): 1032-1048. [赵梓渝, 魏冶, 庞瑞秋, 等. 基于人口省际流动的中国城市网络转变中心性与控制力研究: 兼论递归理论用于城市网络研究的条件性. *地理学报*, 2017, 72(6): 1032-1048.]
- [21] Shi Xiang, Wang Shijun, Wang Dongyan, et al. Characteristics and influencing factors of daily population flow among cities in China. *Scientia Geographica Sinica*, 2022, 42(11): 1889-1899. [施响, 王士君, 王冬艳, 等. 中国市域间日常人口流动特征及影响因素. *地理科学*, 2022, 42(11): 1889-1899.]
- [22] Chen Z, Lu M, Xu L. Returns to dialect: Identity exposure through language in the Chinese labor market. *China Economic Review*, 2014, 30: 27-43.
- [23] Chiswick B R, Miller P W. Occupational language requirements and the value of English in the US labor market. *Journal of Population Economics*, 2010, 23(1): 353-372.
- [24] Ma Shuang, Zhao Wenbo. Linguistic diversity and labor income of the migrants: An empirical research using CHFS. *China Economic Quarterly*, 2019, 18(1): 393-414. [马双, 赵文博. 方言多样性与流动人口收入: 基于CHFS的实证研究. *经济学(季刊)*, 2019, 18(1): 393-414.]
- [25] Guiso L, Sapienza P, Zingales L. Cultural biases in economic exchange? *The Quarterly Journal of Economics*, 2009, 124(3): 1095-1131.
- [26] Beier H, Kroneberg C. Language boundaries and the subjective well-being of immigrants in Europe. *Journal of Ethnic and Migration Studies*, 2013, 39(10): 1535-1553.
- [27] Falck O, Heblich S, Lameli A, et al. Dialects, cultural identity, and economic exchange. *Journal of Urban Economics*, 2012, 72(2-3): 225-239.
- [28] Chu Rongwei, Xiong Yihan, Zou Yi. The determinant factors of migrant workers' social identity: An empirical study in Shanghai. *Chinese Journal of Sociology*, 2014, 34(4): 25-48. [褚荣伟, 熊易寒, 邹怡. 农民工社会认同的决定因素研究: 基于上海的实证分析. *社会*, 2014, 34(4): 25-48.]

- [29] Kang Huilin. Can language competence affect the subjective well-being of migrants: Analysis based on China urbanization and labor migration survey data. *Journal of Social Development*, 2020, 7(4): 201-218, 242. [康慧琳. 语言能力会影响流动人口的主观幸福感吗? 基于中国城镇化与劳动移民调查数据的分析. *社会发展研究*, 2020, 7(4): 201-218, 242.]
- [30] Lu Yonggang, Zhang Kai. Geographic distance, dialect difference, and spatial labor mobility. *Statistical Research*, 2019, 36(3): 88-99. [鲁永刚, 张凯. 地理距离、方言文化与劳动力空间流动. *统计研究*, 2019, 36(3): 88-99.]
- [31] Zhao Ziyu, Han Zhonghui, Wei Ye, et al. Effect of population flow control in restraining COVID-19 in China. *Acta Geographica Sinica*, 2022, 77(2): 426-442. [赵梓渝, 韩钟辉, 魏冶, 等. 中国人口流动管控应对COVID-19疫情效应评估. *地理学报*, 2022, 77(2): 426-442.]
- [32] Institute of Linguistics, CASS, Institute of Ethnology and Anthropology, CASS, Language Information Science Research Centre, City University of Hong Kong. *Language Atlas of China*. 2nd ed. Beijing: The Commercial Press, 2012. [中国社会科学院语言研究所, 中国社会科学院民族学与人类学研究所, 香港城市大学语言资讯科学研究中心. *中国语言地图集*. 2版. 北京: 商务印书馆, 2012.]
- [33] Zhao Ziyu, Wang Shijun, Chen Xiaofei, et al. Cross-regional relatedness and influencing factors of China's automobile production network from the perspective of "global-local": A case study of FAW-Volkswagen. *Progress in Geography*, 2022, 41(5): 741-754. [赵梓渝, 王士君, 陈肖飞, 等. “全球—地方”视角下中国汽车生产网络的跨域关联与影响因素: 一汽—大众案例. *地理科学进展*, 2022, 41(5): 741-754.]
- [34] Sheng Kerong, Zhang Hongxia, Zhao Chaoyue. Determinants of the urban spatial network in China: An analysis through the lens of corporate networks within electronic information industry. *Geographical Research*, 2019, 38(5): 1030-1044. [盛科荣, 张红霞, 赵超越. 中国城市网络关联格局的影响因素分析: 基于电子信息企业网络的视角. *地理研究*, 2019, 38(5): 1030-1044.]
- [35] Zhao Ziyu, Wei Ye, Yang Ran, et al. Gravity model coefficient calibration and error estimation: Based on Chinese interprovincial population flow. *Acta Geographica Sinica*, 2019, 74(2): 203-221. [赵梓渝, 魏冶, 杨冉, 等. 中国人口省际流动重力模型的参数标定与误差估算. *地理学报*, 2019, 74(2): 203-221.]
- [36] Tong Yufen, Wang Yingying. The choice of floating migrants in China: Why megacities are always preferred? A cost-benefit analysis. *Population Research*, 2015, 39(4): 49-56. [童玉芬, 王莹莹. 中国流动人口的选择: 为何北上广如此受青睐?: 基于个体成本收益分析. *人口研究*, 2015, 39(4): 49-56.]
- [37] Wang Y X, Wang F H, Zhang Y, et al. Delineating urbanization "source-sink" regions in China: Evidence from mobile app data. *Cities*, 2019, 86: 167-177.
- [38] Tang Xiaobin, Cui Maosheng. Research on the dynamic change of goods trade network structure and its impact mechanism of countries along the Belt and Road. *Journal of Finance and Economics*, 2020, 46(7): 138-153. [唐晓彬, 崔茂生. “一带一路”货物贸易网络结构动态变化及其影响机制. *财经研究*, 2020, 46(7): 138-153.]
- [39] Duan Dezhong, Du Debin, Chen Ying, et al. Technology transfer in China's city system: Process, pattern and influencing factors. *Acta Geographica Sinica*, 2018, 73(4): 738-754. [段德忠, 杜德斌, 谌颖, 等. 中国城市创新技术转移格局与影响因素. *地理学报*, 2018, 73(4): 738-754.]
- [40] Lv Chen, Sun Wei. Impacting factors of population agglomeration areas on migration: A case study in Dongguan City. *Progress in Geography*, 2014, 33(5): 593-604. [吕晨, 孙威. 人口集聚区吸纳人口迁入的影响因素: 以东莞市为例. *地理科学进展*, 2014, 33(5): 593-604.]
- [41] Huang Jinghu. *Chinese Dialectology*. Xiamen: Xiamen University Press, 1987. [黄景湖. *汉语方言学*. 厦门: 厦门大学出版社, 1987.]
- [42] Tang Jinyue, Zhang Weiyang, Wang Yifei. The pattern and influencing factors of daily population movement network in the Yangtze River Delta. *Geographical Research*, 2020, 39(5): 1166-1181. [唐锦玥, 张维阳, 王逸飞. 长三角城际日常人口移动网络的格局与影响机制. *地理研究*, 2020, 39(5): 1166-1181.]
- [43] Ding Congming, Ji Zhenlin, Lei Yu, et al. Dialect and market segmentation: Evidence from the synthetic metropolitan area. *Economic Research Journal*, 2018, 53(11): 148-164. [丁从明, 吉振霖, 雷雨, 等. 方言多样性与市场一体化: 基于城市圈的视角. *经济研究*, 2018, 53(11): 148-164.]
- [44] Liu Zhang, Qian Jiale, Du Yunyan, et al. Multi-level spatial distribution estimation model of the inter-regional migrant population using multi-source spatio-temporal big data: A case study of migrants from Wuhan during the spread of COVID-19. *Journal of Geo-information Science*, 2020, 22(2): 147-160. [刘张, 千家乐, 杜云艳, 等. 基于多源时空大数据的区际迁徙人群多层次空间分布估算模型: 以COVID-19疫情期间自武汉迁出人群为例. *地球信息科学学报*, 2020, 22(2): 147-160.]

- [45] Li Tao, Wang Jiaoe, Gao Xingchuan. Comparison of inter-city travel network during weekdays and holiday in China. *Acta Geographica Sinica*, 2020, 75(4): 833-848. [李涛, 王姣娥, 高兴川. 中国居民工作日与节假日的城际出行网络异同性研究. *地理学报*, 2020, 75(4): 833-848.]

A study on the effect of language ties in China's population flow network

ZHAO Ziyu¹, WANG Shijun², LI Zhuowei², HAN Zhonghui³

(1. School of Tourism and Geography Science, Qingdao University, Qingdao 266071, Shandong, China;

2. School of Geographical Sciences, Northeast Normal University, Changchun 130024, China;

3. Management College, Ocean University of China, Qingdao 266100, Shandong, China)

Abstract: Using Tencent population migration data in 2019 and Chinese language distribution data, the paper reveals the language ties effect of dialect similarity on Chinese population flow, and further explores its spatiotemporal heterogeneity, traffic mode heterogeneity, and the moderated role of spatial distance on its impact. The results show that: (1) Dialect similarity increases the path formation probability and the relationship strength of the inter-city population flow network in China, and forms the language ties effect in the population flow network. Geographic distance and language distance can be regarded as the explicit and implicit spatial costs of population flow, respectively. (2) There are spatiotemporal heterogeneity and traffic mode heterogeneity in the language ties effect. With each unit increase in dialect similarity, the path formation probability of China's population flow network will increase by 47.7% and 49.5%, respectively, and the flow intensity will increase by 17.1% and 9.9%, respectively, in the Spring Festival travel rush and the daily periods. At the same time, there are positive, negative, and degree differences in the influence of the language ties effect on population flow in different language regions and different modes of transportation, revealing the spatial differences and geographical implications of the language ties effect in the process of affecting China's population flow. (3) The effect of language ties is restricted by the law of spatial decay, and shows differences and complexity in different modes of transportation and periods. The paper expands the current research on the influencing factors of China's population flow from the perspective of economic, social and demographic characteristics, helps to understand the driving mechanism of China's population flow, and has practical significance for social governance and cultural diversity protection in multi-ethnic integrated China.

Keywords: population flow; Spring Festival travel rush; Chinese dialects; language ties; cultural ties; China