論文の内容の要旨

A mathematical model of cultural evolution on a population network and its application to the analysis of spatial distribution of dialects

(集団ネットワーク上の文化進化の数理モデルと、方言の空間的分布の解析への応用)

高橋　拓也

In analogy with genetic traits, cultural traits diffuse from one population to another through interaction and learning process of individuals. To deduce how newly invented traits spread from their original populations and to analyze how observed geographic pattern of cultural traits has evolved, network is a convenient mathematical model. This thesis, composed of three substantial chapters enumerated Chapter 2, 3, and 4, aims to shed light to the cultural evolution among multiple populations with the aid of a network whose nodes represent populations and edges represent transmission of cultural traits. As an application of this model, I specifically treat the spatial pattern of dialects, or lexical variants.

In Chapter 2, I discuss a peculiar spatial distribution of dialects; some dialect words are shared among geographically distant groups of people without close interaction. Such a pattern may indicate the current or past presence of a cultural center exerting a strong influence on peripheries. For example, concentric distributions of dialect variants in Japan may be explicable by repeated inventions of new variants at Kyoto, the ancient capital, with subsequent outward diffusion. In Chapter 2, I develop a model of linguistic diffusion within a population network to quantify the distribution of variants created at the central population. Equilibrium distributions of word ages are obtained for idealized networks and for a realistic network of Japanese prefectures. My model successfully

replicates the observed pattern, supporting the notion that a center-periphery social structure underlies the emergence of concentric patterns. Unlike what has previously been claimed, my model indicates that a novelty bias in linguistic transmission is not always necessary to account for the concentric pattern, whereas some bias in the direction of transmission between populations is needed to be consistent with the observed absence of old words near the central population. My analysis on the realistic network also suggests that the process of linguistic transmission is not much affected by between-prefecture differences in population size.

In Chapter 3, I further generalize the model used in Chapter 2 and consider the dynamics of cultural macroevolution, which concerns a long-term evolutionary process involving transmission of non-genetic or cultural traits between populations as well as birth and death of populations. To understand the spatial dynamics of macroevolution, I present a network model of cultural transmission in which any population may innovate a novel trait. Borrowing the method of ancestral backward process from population genetics, my model explores the genealogy of a cultural variant sampled in the present generation. Mathematical analysis of the model enables us to predict the distribution of cultural age in each population of the network, investigate the frequencies of cultural variants originating from given populations, discuss the time it takes for a cultural variant to diffuse between a given pair of populations, and calculate the frequency of each variant in each population under a finite-variant assumption. I also perform numerical analysis on random networks of populations to investigate the effect of network topology and innovation rate on the age and origin of cultural variants in each population. The results suggest that a cultural variant tends to be maintained at a higher frequency if its original population is characterized by more innovations, more influence on other populations, and/or less influence from other populations. My finding also includes that a cultural variant invented in a given population tends to spread more rapidly if the population is more influential to and/or more influenced by other populations.

In Chapter 4, I investigate the spatial distribution of lexical variants of Japanese language based on both empirical data and the mathematical model presented in Chapter 3, focusing mainly on the relationship between linguistic distance (LD) and geographic distance (GD). For empirical data, I calculate the edit (Levenshtein) distance among lexical variants recorded at 2400 localities surveyed in the Linguistic Atlas of Japan (LAJ) project, showing that linguistic distance between localities is strongly correlated with geographic distance. Since a pair of words derived from the same source usually bears a marked similarity, simulating when and where lexical variants were invented is an essential way to examine the linguistic distance between localities. For this purpose, I

develop a network of the surveyed localities of LAJ, based on their geographic positions and population sizes, and apply the model of Chapter 3 to quantify the origin and expected age of variants. I show that variants are likely to be originated from localities with large population sizes, whereas variants originating from an arbitrary population occupy every population with a relatively similar probability, unless cultural transmission occurs exclusively between proximate populations. In addition, simulating the linguistic distance for locality pairs on the network, I show that linguistic distance correlates strongly with the logarithm of geographic distance (Séguy's curve) if variants transmit mostly between closely positioned localities. Conversely, if variants can also transmit between remote localities, linguistic distance correlates more strongly with geographic distance itself, resulting in a linear relationship between LD and GD. Moreover, my simulation reveals that accumulation of linguistic distance is not spatially isotropic when lexical variants can only transmit between proximate localities. Within the framework of approximate Bayesian computation with Markov Chain Monte Carlo (MCMC) sampling, we infer parameter values of our model based on empirical data in LAJ.