COEVOLUTION OF LANGUAGES AND SOCIAL NETWORKS

STATEMENT OF RESEARCH PROJECT

Natural languages around the world exhibit tremendous diversity with different levels of structural complexities. For example, languages vary greatly in the size of their phonological inventories, and in the degree to which they make use of inflectional morphology. Two related foundational questions in linguistics and cognitive science at present are:

a) How did languages grow more complex in the first place, and how do they maintain or lose complexity? and
b) What is the source of variation in linguistic complexity across the world?

One appealing hypothesis is that the structural properties of human languages may be influenced, in part, by the community structure of the linguistic populations in which they thrive. As such, languages are seen as evolving organisms that adapt to fit a social niche, similar to living organisms adapting to an ecological niche. For example, small, isolated linguistic communities oftentimes develop languages with great structural complexity, elaborate and opaque morphology, rich patterns of agreement, and many irregularities. By contrast, languages with large communities of speakers, such as Mandarin or English, appear to be morphologically much simpler. There is a historical tendency toward structural simplification as languages gain an ever-larger community of speakers. Some researchers have proposed mechanistic explanations for why languages should become structurally simpler with larger communities. However, to our knowledge no one has offered a quantifiable mechanistic account of how languages become more complex in the first place. Current accounts are descriptive in nature at best but not explicitly explanatory. Thus, the proposal requires further investigation, and, most importantly, the mechanism whereby the social context of learning and use affects the grammatical evolution of a language needs elucidation. In addition, community size and degree of isolation are only a few among several other potential social network measures (see below) whose contribution to language complexity has not been investigated yet.

This project will explore the impact of topological properties of social and communication networks in which speakers reside on the evolution of natural languages while providing the first computationally explicitly causal account for the emergence of structural complexity of human languages. We do so by combining three novel methodologies:

1) Agent-based computer models of language evolution in social networks of simulated agents will provide a proof-of-concept for the hypotheses at hand.
2) Human agent-based simulations of communication between actual speakers engaged in language evolution games will provide empirical complementary evidence that the mechanisms explored with simulated agents are also in place in real-time human communication.
3) Brain imaging techniques combined with Network Analytic methods will correlate individuals' brain responses to linguistic changes with topological properties of social networks (e.g., modularity and spectral analysis). This last approach can measure the degree of language change happening in the brains of individual speakers of a given community, as well as between communities, thus providing a proxy measure of the rate at which the linguistic complexity of a given community changes. It will provide a novel proxy measure for language attrition for the Mother Tongues speakers in Singapore.

SCOPE OF WORK FOR SELECTED PHD STUDENT

The aim of this project is to train the candidate PhD student on a set of theories and methodologies that combined together can offer 'converging multi-method evidence' for a given hypothesis. The project can be divided in three components:

Component 1 - Computer simulations of language games in social networks

The student will be guided (mainly by the Co-Supervisor Oh in WKWSCI) to design computational algorithms, code them using programming languages, such as R and Python, and simulate language games in social networks, with the expectation that the student will learn core concepts in network science, complexity science, and information science, as well as practical skills necessary for computational methods, social network analysis, and other advanced statistical analysis. The high performance computing centre (HPCC) in NTU provides the required computational resources. The Supervisor (Onnis in LMS) has expertise in computer modeling of language, and will also lend his expertise with the conceptualisation and hypothesis link to the linguistic literature, notably evolutionary linguistics.

Component 2 - Human simulations of social learning

The student will create an online computer platform to implement interactive language games. The Supervisor has expertise in this methodology and will guide the student, while the Co-Supervisor will advise on the programming and analysis components of the task. The purpose is to test the predictions generated by the simulations in Component 1, and verify which ones hold in the face of real human communication. To our knowledge, most of the psycholinguistic literature has focused on individual speakers or dyadic relations at best, and thus this line of research would be a trailblazer in a nascent field of network psycholinguistics.

Component 3 - Brain effects of social structure

Language maintenance and attrition are important issues to Singapore, who currently faces a surge of English at the expense of the other official languages. Do the three ethnic communities (Chinese, Malay, Indian) undergo language attrition at different rates, and if so why? The student will examine the topological properties of social networks of participants and relate it to psycholinguistic proxy measures of language attrition. The latter are provided by EEG/NIRS brain responses during psycholinguistic tasks. The goal is to correlate the magnitude of these brain responses (and thus the degree of attrition in the Mother Tongue) to the type of language network the individual is exposed to, using aggregated Relational Data (ARD [5]) as a proxy for each speakers' networks. The Supervisor has published work using EEG methods and psycholinguistics, and is conducting a study with NIRS. The Co-Supervisor will advise on methods for obtaining Aggregated Relational Data (ARD) from participants' social networks.