Tanmoy Bhattacharya tanmoy@santafe.edu

Project 1: Words in languages often display polysemy, that is, they have more than one meaning. It is of inherent interest whether these diverse meanings form patterns across languages. Furthermore, word polysemy is generally believed to be the first step in semantic change in words, and so understanding polysemy is also of interest to historical linguists. Analysis of empirical data on this topic, however, is difficult to collect and needs to control for various systematics. Members of our group have developed three separate projects with different aims that have collected data on occurrences of polysemy across languages. What now needs to be done, and can be done in the scope of an REU project, is to unify the results of the three projects and compare them, seeing which common generalizations emerge, and exploring if differences are due to differences in aims, methodology and scope.

Project 2: Languages change over time at various levels. For example, words change in pronunciation and meaning. It is commonly held that the pronunciation changes are extremely regular, and follow their own rules, without regard to the meaning of the word. Over time, however, phonetic patterns like stress are lost, and borrowing across languages or dialects obscures these regular patterns. Mere sound similarity, however, occurs too often at random to be easily usable to discover relationship between languages. A systematic study of the amount of relationship expected by chance has, however, not been performed. The Evolution of Human Languages project has a database of words from many language families, and we have developed various similarity scores to compare the phonetic similarity. The REU project will involve quantifying the long distance similarities and figuring out the comparative merits of the scores and representations for discovering long distance relationships.