



CALIFORNIA SCIENCE & ENGINEERING FAIR 2018 PROJECT SUMMARY

Name(s) Sophia L. Balkovski	Project Number 38686
Project Title Geographic or Linguistic Distance: Defining African Color Terms	
Objectives/Goals Basic color terms can be defined as words representing a wide set of hues, like "red" or "blue". Color terms differ from language to language. The goal of this project was to find out whether linguistic and/or geographic distances among African languages affected the differences among their color terms. Abstract Methods/Materials Data from the World Color Survey, which collected information on color term usage in 110 world languages, were analyzed using the Mathematica language. A quantitative methodology was used to compare color charts created from the data. Three types of distance were found between any two languages: geographic, linguistic (distance on a language tree) and chart (percent similarity of color terms). Linear regressions were made for each pair of the distances. The languages were then separated into two groups based the number of color terms they had, and the mean inter- and intra-group distances were compared for all distances. Both the chart and linguistic distances were used to reconstruct the geographical distance using multidimensional scaling. Results It was ultimately found that both linguistic and geographic distances have an effect on the way that color terms are used. The linear regressions for all distance pairs found them to be correlated with high statistical significance. Furthermore, for language number groups, the average intra-group distances were smaller than the average inter-group distances, meaning that languages with similar numbers of color terms were close both linguistically and geographically. Surprisingly, it was possible to roughly predict the location of the languages based on color chip names, using multidimensional scaling. Conclusions/Discussion This project has been a study of African languages, but it can be applied in a much wider context. More languages and other linguistic features (e.g. syntactic structures) can be studied. This project is representative of using STEM methodology (mathematics and statistics) to study problems in human behavior. It is also a step closer to being able to mathematically represent the spread of culture between languages. With the expansion of globalization, it is becoming increasingly important to preserve Native American, African and Aboriginal cultures. The long-term goal is to create a model that could help design the best strategies to protect the diversity of these indigenous cultures.	
Summary Statement I found a way to mathematically model the spread of culture, by looking at color terms in African Languages	
Help Received I discussed my ideas with a professor, and he suggested I use multidimensional scaling, which helped me further prove how correlated linguistic, geographic and chart distances were.	