This project explores language ideologies implicated in the online translation service offered by Google Translate Hindi and Google Translate Urdu (https://translate.google.com/). It emerges out of the assumption that spoken Hindi and Urdu (often described as Hindi-Urdu or Hindustani) have a shared literary and cultural heritage despite nationalist attempts in India and Pakistan to establish separate and linguistically distinct genealogies of Hindi and Urdu, which are also the official languages of the Indian and Pakistani nation-states, respectively. This shared language is also one long used in the commercial film industry and in the ever so popular “Hindi” film songs. Since the mid-nineteenth century there have been efforts to “cleanse” or “purify” the Hindi and Urdu languages by replacing common (Hindi-Urdu) words from a shared linguistic lexicon with more Arabicized and Persianized (in the case of Urdu) as well as more Sanskritized (in the case of Hindi) vocabulary.

The goal of the project is to uncover such processes as they are also taking place on Google Translate, which over the past seven years that I have been observing the site has shifted from a dictionary of Hindi-Urdu to one representing the official languages of two nation states, i.e. Hindi and Urdu. The findings of this project will also be useful for students in introductory, intermediate and heritage-language Hindi-Urdu classrooms: dictionaries are not “neutral”, but based on decisions influenced by language ideologies and language politics.

Together with the faculty member, the undergraduate student will prepare a representative sample of Hindi, Urdu and Hindi-Urdu words as they are commonly promoted on both Google Translate Hindi-English and Google Translate Urdu-English websites. Once this part of the project is completed, the student will move on to an even more comparative analysis of Hindi-Urdu language use in South Asian popular movies and songs.

Ideally, the undergraduate student is a Hindi-Urdu heritage speaker. In any case, knowledge of the Hindi (Nagari) and Urdu (Nastaliq) scripts is required for this project.