

Abstract

Language change refers to the phenomenon of alterations in the linguistic structures of the language used by a community over time, which may subsequently give rise to one or more languages or dialects. Phonological change is a special type of language change that involves alteration in the pronunciation patterns. The aim of this thesis is to develop computational models for cases of phonological changes that have been observed for real languages and validate the same against real data.

Several theories have been proposed to explain the cause and nature of phonological change, which, nevertheless, cannot be validated and established firmly due to the lack of sufficient linguistic data and scope of experimentation. Computational models not only provide a formal description of the linguistic theories, but also serve as excellent virtual laboratories. The inherent complexities of our cognitive and social structures, however, make it hard or impossible to replicate reality within a computational framework. Consequently, most of the present computational models are developed for toy languages and are rarely compared with real linguistic data – an obvious weakness that has attracted criticisms from linguists and computer scientists alike.

In this thesis, we present computational models pertaining to the different explanations of phonological change for two cases of real world language change: *schwa deletion in Hindi* (SDH) and *change affecting the verb inflections of the dialects of Bengali*. Hindi and Bengali are two major languages of the Indo-Aryan family that are spoken in the Indian sub-continent. The models are developed in an incremental manner, where the findings of a more detailed, but computationally intensive model are used to abstract out the details in the

subsequent phases of the modeling. This, in turn, leads to extensive models that are not as detailed as their precursors, but computationally more tractable and linguistically equally plausible. This, which we call the *hierarchical abstraction methodology*, is the prime contribution of the thesis. Besides this, the novel formalisms, analysis techniques and the applications that have been developed in the context of the two specific problems tackled are also salient contributions of this work.

The models proposed in this work include: a constrained-optimization model for SDH that has been solved analytically to show that the pattern of SDH is optimal under certain assumptions; two multi-agent simulation models for SDH, where it has been shown that under suitable circumstances the pattern of SDH automatically emerges in these systems without any explicit optimization or global control; a multi-objective genetic algorithm based optimization model to explain the morpho-phonological change of Bengali verb inflections and the dialect diversity arising as a consequence of the same. While the optimization models put forward functional explanations of phonological change, the multi-agent simulation models provide an emergent explanation grounded in the principles of *phonetically-based phonology* and *evolutionary phonology*.

To summarize, this thesis shows that, through the concept of hierarchical abstraction methodology, it is indeed possible to build computational models of real world language change. Thus, we believe that in the future computational techniques will serve as an important and inevitable tool for research in diachronic linguistics.