Book Review

The Computational Nature of Language Learning and Evolution

Partha Niyogi
(University of Chicago)

The MIT Press, 2006, xviii+482 pp; hardbound, ISBN 0-262-14094-2, $42.00, £27.95

Reviewed by
Tony Belpaeme
University of Plymouth

Darwin already remarked that evolutionary thinking also applies to the study of language. Language is heritable, in the sense that the language of offspring will likely resemble that of the parents, and during language learning variation is inevitably introduced. If on top of this a selection mechanism is operating that allows individuals using a particular language to have more descendants, some languages are more likely to spread through the population than others. The evolutionary nature of language change has been extensively studied in diachronic or historical linguistics, but in *The Computational Nature of Language Learning and Evolution* Niyogi takes a fresh approach by providing a formal study of evolutionary language change. In this he focuses on the population instead of on the individual language users, and provides a thorough analysis of the population dynamics resulting from individuals learning and using a language.

Language evolution has only recently been subjected to the rigor of mathematical analysis. The expertise built up in theoretical biology, game theory, information theory, statistical physics, and complex systems research seems to be particularly well-suited to study and report on language dynamics at a macroscopic level. In addition, these disciplines rely on a set of trusted analytical tools that can be employed to study dynamical aspects of language learning and evolution.

Niyogi sees language acquisition as a mapping of example sentences onto a private grammar. For this the language learner uses a particular learning mechanism, and the first part of the book is concerned with the "logical problem of language acquisition" (Gold 1967), and a number of formal learning mechanisms are presented that will be used in later chapters. Niyogi makes no commitment to the representation of grammars. Grammars can be generative grammars, but can equally be phonological rules, probabilistic grammars, sentence–meaning pairs, or any other combinatorial and compositional structure. However, he does assume that the learner has some sort of bias in its learning mechanism, a universal grammar if you will, that constrains the grammar acquisition process so that a stable language is maintained in a community.

In the first chapters, a number of learning mechanisms are presented and explored, including the memory-less learner, which bases its next hypothesis only on its current hypothesis and the current sentence, and the batch learner, which waits until all example sentences have been received and then chooses the most likely hypothesis. If the learner is exposed to more example sentences, the hypothesis of the learner will home in on the target grammar of the teacher. Only if an infinite number of sentences are presented will the learner be able to acquire the exact same grammar as its teacher. In reality, no child ever hears an infinite number of sentences, so it is bound to learn a grammar that varies slightly from that of its caretakers. Exactly this variation drives

diachronic change in languages. From this very simple formal framework, it already follows that language-learning in a population exhibits complex dynamics. These might result in chaos, but more often than not in the case of language, individuals chaotic regimes typically steer clear of. If bifurcations do arise from the non-linear dynamics of language acquisition, these might help explain major transitions in language, such as rather abrupt changes in word-order structure.

The next part of the book focuses on language change, or its counterpart: language stability. If language learners form a distributed system without coordination, how can coherence in the language arise and how can it be maintained? Niyogi shows that coherence does not imply that all learners need to have the same grammar; they only need a grammar that produces sentences that are correctly interpreted by others. Language change, as with most changes in a large population, has often been believed to follow a logistic S-shaped curve. However, the nature of the change depends, among other things, on the learning algorithm and distributional assumptions of sentences offered to learners. Niyogi shows how under certain circumstances language change diverges from the typical logistic curve, and populations end up in alternative absorbing states of the dynamical system. He explores a wide range of parameter settings of his system, changing the number of languages in the population, the structure of the population, and introducing maturation in the population. Language change is highly non-linear, and a slight variation in a parameter setting can result in very different evolutionary trajectories. In Chapter 7 this framework is applied to the historical change of the placement of clitics in Portuguese, and Chapter 8 does the same for phonological change in the Wu dialects in China. Interestingly, Niyogi fails to explain how the phonological change occurred, but at same time—through presenting possible evolutionary scenarios—manages to discount a number of alternative hypotheses for the change.

In the first eight chapters, only vertical language evolution is considered, in which language is transmitted from generation to generation. Chapter 9 extends the formal models to handle horizontal, or cultural, evolution. These are built upon the theories of Cavalli-Sforza and Feldman (1981), in which a set of cultural traits with a limited number of parameter settings is passed on between individuals. Niyogi explores the competition between linguistic traits and illustrates this with a study of the evolution of English syntax from Old over Middle to Modern English. Chapter 10 explores some variations on the model. Until now Niyogi considered only infinite populations, making the models' predictions insensitive to interaction dynamics. Having finite populations and a spatial organization between individuals introduces stochasticity in the interaction dynamics, which—although more realistic—makes it more difficult to make generalizing conclusions.

The last part of the book is concerned with the origin of language, and in particular with the conditions under which communication systems are selected for. Niyogi defines a measure for the fitness of an individual, the communicative efficiency, which measures how mutually intelligible a language is. Interestingly, in a case study using phonemic contrast in English, where it is shown that there is quite some perceptual confusion in English, Niyogi warns that communicative efficiency might not have been very important in the evolution and origins of language. Chapter 12 is an extended version of Nowak, Komarova, and Niyogi (2001), studying populations of interacting linguistic agents where communicative efficiency forms a selective Darwinian pressure. It is formally shown that the learning fidelity in the model, which is the probability of a child learning the parent's language, should be sufficiently high for coherence to emerge in a population. If the fidelity is not high enough, the population will not end up speaking one majority language and linguistic communication will not take off.

Interestingly, Niyogi remarks that this predicts that having only one or a few languages is optimal, but given the many thousands of languages spoken today, this result is rather counterintuitive. In Chapter 13, Niyogi looks at circumstances under which linguistic coherence might emerge without natural selection on linguistic performance. Instead, social learning is used, and individuals now learn from the entire population, as opposed to learning from a single parent as in previous models. It is shown that under a large range of conditions, the evolutionary trajectory bifurcates and the population ends up speaking one language. An interesting insight is that if a learner only takes input from a single parent, natural selection is needed on communicative fitness to allow a shared linguistic system to emerge (this might, for example, be the case for bird song). However, if the learner receives input from a community, natural selection is not a necessary requirement for linguistic coherence to emerge.

Chapter 14 concludes and again underlines the importance of bifurcations: moments in history where linguistic evolution suddenly takes a sharp turn. These bifurcations result from the nonlinear interactions between language users, and could explain major transitions in linguistic history. Niyogi also points to future directions for this type of work, mostly involving more complicated versions of the current models.

In addition to mathematical analysis, in recent years computational modeling of language evolution has also received equal if not more attention. Whereas formal analysis allows insights in the macroscopic regime of an evolutionary system—through, for example, state-space analysis—computational modeling allows for more realistic models (albeit still rather simple compared to the actual phenomena that are being modeled). Niyogi for the most part prefers formal analysis, and although this requires abstraction to the extreme, the book clearly shows how this can still provide intriguing insights. The main reason for this is that formal analysis (and computational modeling) brings issues into focus in a way that verbal arguments cannot. Niyogi uses the language of statistical physics to give an evolutionary account of language change and language origins, which for some readers might be a little daunting. However, an effort has been made to reach out to different disciplines and together with a peppering of practical examples, *The Computational Nature of Language Learning and Evolution* will not only be of interest to researchers modeling the evolution of language, but also deserves attention from adventurous historical linguists.

References

Cavalli-Sforza, Luigi and Marcus Feldman. 1981. *Cultural Transmission and Evolution: A Quantitative Approach.* Princeton University Press, Princeton, NJ.

Gold, Mark E. 1967. Language identification in the limit. *Information and Control*, 10(5):447–474.

Nowak, Martin, Natalia Komarova, and Partha Niyogi. 2001. Evolution of universal grammar. *Science*, 291:114–118.

Tony Belpaeme is a Lecturer in Intelligent Systems at the University of Plymouth (UK). He works on computational modeling of concept acquisition under the influence of language; he then uses insights gained from this to construct cognitive systems for AI and robotic applications. Belpaeme's address is University of Plymouth, School of Computing, Communications and Electronics, A318 Portand Square, PL4 8AA Plymouth, UK; e-mail: tony.belpaeme@plymouth.ac.uk.