CALL & Complexity

Stephen Shucart

INTRODUCTION

Ultimately all teaching, including CALL, is based on the teacher's model of language learning. Such models may be explicit or implicit, but they are the foundation on which the teacher builds classroom practices and the standard for judging their effectiveness. While many models of language learning exist, none have been able to link what is happening at the neural level with the student's overt performance, nor have they been able to effectively integrate all of the various factors which so obviously affect language learning. This paper presents a new model of language learning that overcomes these limitations. Below, some ideas are introduced and then the model is described. Finally the model's application to CALL is briefly outlined.

Language as a Hierarchy

We will begin with the concept of language as a hierarchy. No matter which language feature we select, there are prerequisites that must be learned before the student can master that feature. For example, before a vocabulary item can be mastered orally, the student must have control of the pronunciation, a knowledge of the parts of speech, and the ability to link the verbal representation to meaning, to list only a few necessary elements.

Although it would take a huge effort, it is possible to specify the prerequisites for each and every language feature. These features could then be arranged in a tree with the highest level processes as the leaves and the underlying physical processes as the roots. Once the tree is defined, the prerequisites for any language feature would be those nodes that form a path leading to that feature. With a fully specified tree of features, it is possible to define different levels using the number of nodes between the feature in question and the base roots of the tree.

Language as Usage

ELA: Emergence, Learning, and Acquisition

Gaining a language can be considered from three different perspectives. One is *learning*, in which specific elements of language are explicitly taught and practiced. A second perspective is *acquisition*, in which a person gains the language through exposure to meaningful input. Finally, in a complexity-based model, language is said to *emerge* from the interaction between the variables involved.

In keeping with general practice, the term *learning* will be used when precision is not necessary. The context will clarify any contrastive use of this term.

Classes of ELA Variables

Obviously, factors other than the language itself will affect the efficiency of usage. We have all experienced those ups and downs that seem to make the language either the easiest thing in the world or virtually impossible. Factors ranging from our mood to the temperature of the room to the method used by the teacher have wide-ranging effects on how well we are using the language. The specific factors involved in learning are almost infinite in number, but they can be summarized by seven classes of variables.

Student

The factors in this class include but are not limited to such things as the student's mood, linguistic ability, prior learning and experiences, motivation, and expectations. One aspect of the student's experience is the amount of exposure and practice that the student has had with the language.

Class

These factors concern the effects of the psychological interrelationships between the various members in a conversation, or a class. For example, when the students become a single supportive group, learning is enhanced, but if the group contains cliques and factions, learning can be severely inhibited.

Teacher

This class of variables relates to the skills, knowledge, training, and experience of the teacher as well as such factors as personality and current mood.

Method

These variables concern the specific method, or lack of method, that the teacher is using.

Materials

The physical materials introduced into the classroom as well as their linguistic and thematic content also have an effect, positive or negative, on the student's learning.

Environment

The physical location where the students are working and its characteristics affect usage. These factors include such things as temperature, comfort of the chairs, color of the room, and amount of free space.

Administration

Administration refers to factors such as scheduling, curriculum, availability of equipment, and social expectations, factors related to the school and society but which are beyond the control of the student or the classroom teacher.

Fitness Landscapes

A fitness landscape is a multidimensional graph in which the dependent variable is the effectiveness or efficiency of a process and the independent variables are those factors that influence that effectiveness. A single point on the landscape can be represented by a large binary number where each bit represents the presence or absence of a single feature or the minimum

measure of a variable, digital or analog, that is represented by a series of bits, each representing the minimum distinction that we can make within that variable.

Characteristics of a fitness landscape

Each combination of variable values will have a fitness associated with it and a plot of these values will form a surface, called a landscape because it has flat areas, valleys, and peaks. With a small number of variables the landscape will have a single peak, but as the number of variables increases the number of peaks will increase until the entire landscape is covered with steep-sided peaks and valleys. Stuart Kauffman of the Santa Fe Institute has done extensive research on landscapes. He found that among other things landscapes where the variables interact with each other are extremely rugged and very small changes in the variable values may cause large changes in the fitness (Kauffman, 1993, 1995).

The landscape for language use

The landscapes for language use are extremely rugged, since most of the variables are interconnected. For example, if the student is in a bad mood, this may affect the teacher's mood, causing both of their responses to the room temperature to be different than it would be if they were alone. This means that theoretically we can expect small changes in the ELA variables will elicit large changes in the student's ability to use a language, explaining the non-reproducibility of many applied linguistics experiments.

A Fitness Landscape for a Specific Linguistic Feature

First, each of the prerequisite features must be in place before the fitness landscape usage arises. The independent variables consist of those variables from the ELA which have an effect on the use of the present feature. The fitness will be a measure of the likelihood that the feature will be appropriately and accurately used.

The Terraced Labyrinth Model

The original Terraced Labyrinth model (Crutchfield & van Nimwegen, 1999), a tree-based model of punctuated evolution, includes the following basic components.

Fitness

The fitness is a measure of the probability of an individual successfully reproducing.

Genotypes

A genotype is a very large but finite binary number that represents the DNA of an individual. The actual instantiation of a genotype is called a phenotype. Genotypes and fitness landscapes represent different aspects of the model, genotypes applying to the tree and fitness landscapes to the subbasins.

Subbasins

A subbasin is a multi-dimensional volume, a fitness landscape, that describes the particular species at that level of evolution.

Portals

A portal is a subset of genotype bits, called a constellation, that when set to a specific value allow movement to the next higher subbasin. This movement represents the emergence of a new species. A description of the new species can not be predicted before its emergence and can only be found by observation.

Learning a language

This section puts the above pieces together to form a comprehensive model of the emergence, acquisition, and learning of languages.

A Terraced Labyrinth Model for Language Learning

This model for language learning is based on individuals rather than populations as in the original Terraced Labyrinth. Also it is limited to only those features directly related to language.

New definitions of fitness

There are three landscapes associated with this model, each with its own fitness. The first landscape is a sequencing variable that specifies the level of the subbasin. The second describes the probability that a particular genotype bit will change. The independent variables are the states of the other bits and the environment. The third is the probability with which an emergent feature will appear appropriately in the student's language. These independent variables are contained in the ELA sets.

Specification of genotype

As in the general model, the genotype is a large, but finite, binary number representing the presence or absence of particular features, physical or linguistic in the case of language.

Subbasins as individual aspects of language

Subbasins now represent the individual features of language, rather than species as in the evolutionary model. A subbasin, however, represents different things at different levels of the tree. At lower levels, a subbasin would represent something as specific as having and being able to control a specific muscle that is used to make a sound. At a higher level, another subbasin might represent the ability to incorporate the surrounding sounds in the decision to actually produce a particular sound in an utterance. At still higher levels, subbasins would represent larger and more complex features of the language, including grammar, vocabulary, function, and all the other features that appear.

Portals between linguistic subbasins

A portal is a constellation that represents the presence of a prerequisite set of linguistic features that allow the emergence of a new feature.

Effects of input

The genotype settings change according to the linguistic input received and the current settings of the other genotype bits. Future research may show that the ELA variables also have an effect, but until such evidence is available, Occam's Razor will be applied and the ELA variables will be considered to influence only the usage of a feature. In either case, the final result is the same.
At present we can not determine where the students are in the language tree, so we do not know what input they need next. Also different students in a class will be at different levels. The best solution seems to be to provide the students with a large amount of comprehensible inputAs the details of the tree develop, however, it should be possible to tailor the input more closely to the needs of the next subbasins.

Differences between Learning and Acquisition

The inclusion of input in the model naturally focuses attention on the differences between learning and acquisition. The Terraced Labyrinth Model helps clarify the differences, but there is still room for further research in this area. It may well be that the acquisition/learning dichotomy may need modification.

Language Acquisition

Acquisition within this model is the adjustment of the genotype bit settings through exposure to comprehensible target-language input. The values are changed in a semi-random way that depends on the statistical features of the language input and its comprehensibility. Thus, the individual's path up the language tree will be idiosyncratic, depending on that student's current location in the language tree, the input, and the ELA variable states.

Language Learning

Learning, or being taught a specific feature, is an attempt to ignore the tree. A feature is presented to the students whether or not the prerequisites are present and the students are expected to learn the feature, or fail with all its negative consequences. The students will attempt to place the feature into their general tree. Some will find that the prerequisites are available and the feature will become part of the language tree. Others will find a place for the feature in another part of their general tree. These students will consciously understand the feature but not be able to consistently employ it as language. The remainder will find they are missing prerequisites and will neither understand nor be able to use the feature.

Possibilities for Verifying the Model

Although it will require much work, the model can be specified in detail and verified. The

first step would be to arbitrarily select language features and determine the necessary prerequisites. Then the prerequisites would be analyzed for their prerequisites. This process would continue until a level was reached where we could assume that the students would bring the features with them. At the same time, the initial features could be used to help determine the prerequisites for the upper level features. Eventually the entire language tree for a specific language could be specified in detail. Throughout the process, concurrent research using MRIs, cognitive psychology, traditional linguistics, and other methods could be used to verify the subbasins at each level.

Implications for CALL

The complete language tree will contain a vast amount of data and the only practical way to organize and maintain the data will be in a computer database. As the details become available computerized tests mat be developed to determine what prerequisites a student is missing for those language features that are the goal of the present learning session. The CALL program will be able to administer a series of these short tests, possibly disguised as something else for psychological reasons, to determine what prerequisites are yet to be acquired and then teach these before going on to the points in question. By allowing the teaching to be focus exactly on what is needed the efficiency will increase substantially. The CALL program itself will be able to tailor the content of the instruction to the exact needs of the individual student.

Conclusion

Models generalize, delete, and distort and, therefore, refer only to certain aspects of the actual phenomenon. The Terraced Labyrinth is no exception, which means that it should be thought of as supplementing rather than replacing other models of language and learning. The particular value of this model is its potential applications to CALL and its ability to allow the user to micro-teach, that is, to focus as closely as possible on what is needed by the student to take the step to the next subbasins in his or her linear trip up the language tree.

References

Crutchfield, J. P., & van Nimwegen, E. (1999, February) . The Evolutionary Unfolding of Complexity. Santa Fe Institute Working Paper 99-02-015.

Kauffman, S. (1993) . The Origins of Order: Self-organization and selection in evolution. Oxford: Oxford University Press.

Kauffman, S. (1995) . At Home In The Universe: The search for the laws of self-organization and complexity. Oxford: Oxford University Press.