Schulze, M. (2008). Modeling SLA processes using NLP. In C. A. Chapelle, Y.-R. Chung, & J. Xu (Eds.), *Towards adaptive CALL: Natural language processing for diagnostic language assessment* (pp. 149- 166). Ames, IA: Iowa State University.

Modeling SLA Processes using NLP

Mathias Schulze
Department of Germanic and Slavic Studies / Centre of Arts and Technology
University of Waterloo, Ontario, Canada

Attempting to understand and to capture the complex and dynamic nature of language learning processes is a non-trivial task for researchers in Second Language Acquisition (SLA) and Computer-Assisted Language Learning (CALL). After sketching major developments in SLA and student modeling for Intelligent CALL—the intersection of Artficial Intelligence and CALL—this paper proposes a conceptual framework of the dynamic language proficiency of students using the example of the Mocha project. The main project goals are outlined. Dynamic Systems Theory and Construction Grammar are motivated as the theoretical foundation of our thinking about SLA and student modeling.

Individualization has been praised as one of the major strengths and advantages of computer-assisted language learning (CALL). Advocates of CALL frequently mention the value of students working at their own pace and receiving feedback immediately and—sometimes—based on an individual context. However, individualization from this perspective often means having the student work individually using a tutorial CALL package or simply lack of instructor control. Individualization in this sense does not imply that individual characteristics of the students are considered and that the computational learning environment is tailored accordingly. Often it does not even mean that the prior learning path—learning events with their contents and the students' achievements—is recorded and has an influence on such decisions as which learning objects are presented or what kind of feedback is provided. I will argue in the course of this paper that a non-mechanical and more humanistic approach to individualization in CALL can only be achieved when the CALL system comprises a student model, a model which is informed through the analysis of learner texts.

Let us begin with a cursory look at trends in second language acquisition (SLA) research from which one might seek guidance for development of student models and note the assessment problem that arises in this context. The subsequent discussion of student models will show how different modeling techniques correspond to different approaches in SLA and will highlight their strengths and limitations that pertain to individualization. The main part of this paper will sketch a new approach to modeling language learning, an approach which relies on aspects of Dynamic Systems Theory (DST).[i] The consequences such a modeling approach has for the underlying natural language processing (NLP) techniques will be illustrated through the description of a project currently underway which aims to conceptualize and construct a student model for early and intermediate learners of German at a Canadian university.

APPROACHES TO SECOND LANGUAGE ACQUISITION

Our current understanding of second language acquisition processes can be said to have started with Lado (1957) and his 'contrastive hypothesis.' The stronger version of this hypothesis claims that areas of learning difficulty can be predicted on the basis of typological analyses of the language acquired first (L1) and the language learnt or acquired later (L2). It then goes on to claim, for example, that if structural features of L2 differ greatly from comparable features in L1, learners of L2 will make errors when using linguistic structures with these features. This hypothesis had—and still has—some intuitive appeal. However, its predictive power is limited: learners make errors and have some difficulty with language phenomena that are very similar in their L1 and L2 (e.g., the learning and un-learning of cognates and false friends) and, vice versa, few or hardly any transfer-induced errors occur in some areas which are different in L1 and L2 (e.g., German learners of English usually have little trouble spelling English words in spite of the fact that the orthographic systems of the two languages are radically different). This shows that typological differences between L1 and L2 certainly influence the nature of the language learning process, but this variable is not an exclusive and sufficient predictive factor of language learning behavior and success—not even a dominant one—because the contrastive hypothesis in its reliance on comparing two language systems—two langues—fails to consider the language learner.

SLA researchers made an attempt to address these limitations by focusing on the learner language through error analysis (Corder, 1974, 1981). This approach did indeed improve our understanding of language learning processes by identifying areas of genuine difficulty for certain groups of learners. However, the exclusive focus on errors—the negative results of an individual learner's efforts—led to the exclusion of error avoidance strategies (Schachter, 1974) learners employ because these could not be detected on the surface in the analysis of an individual text. The focus on errors also meant that learner language was described as flawed, shortcomings were emphasized, and learner success and interesting, creative, and meaningful aspects of learner utterances and texts were ignored. Thus, the learner's identity construction in text—their image as depicted in the what and how of the text—was reduced to the classificatory description of selected negative features. Social and individual characteristics of the learner as text producer did not feature in error analysis.

The sole concentration on learner errors was overcome in interlanguage analysis (Selinker, 1974, 1992). All linguistic structures in texts by individual learners over time were considered. This interlanguage continuum—both across groups of learners and for each learner over time—was conceptualized as a variety space (Klein, 1986; Klein & Dittmar, 1979; Klein & Perdue, 1992). Each variety was assumed to have its own system of rules. Thus interlanguage was and still is described as systematic, a conceptualization which provides a basis for a computational implementation of the interlanguage grammar. These language systems consist of rules which are identical to rules in L1 or other previously acquired languages (language transfer) or in L2 (acquired rules) as well as rules which are specific to this particular variety (e.g., through overgeneralization and simplification). The often purely structural focus of interlanguage studies has been extended more recently to include the pragmatics of interlanguage (e.g., Kasper & Rose, 2002). Despite the theorized systematicity of individual interlanguages, considerable variability is also observed, and as a consequence, modeling interlanguages has proven more difficult than one might initially expect.

Perhaps even more problematic for conceptualizing student models in CALL, interlanguage studies have largely ignored individual learner differences (Dörnyei, 2005) and concentrated on groups of learners. Even when individual differences are studied in SLA research, researchers often looked at one variable after attempting to statistically or experimentally eliminate all others. This contributed to—what was perceived by some as—a dominance of quantitative approaches in SLA (Firth & Wagner, 1997). The concentration on individual variables yielded interesting insights into important aspects of language learning processes and language learners and instructors. However, the exclusion of other variables—often minor ones—resulted in these (quantitative) studies presenting contradictory results (Larsen-Freeman, 1997).

It is likely that contextual differences in these studies have a disproportionate effect on their findings. In other words, quantitative studies which should be generalizable, may be impossible to replicate because of seemingly innocuous contextual variables. In contrast, qualitative (case) studies made every attempt to consider all variables which might possibly influence the learner, the learning process, and the learning outcomes as well as the instructor. Such studies posed challenging new questions and highlighted interesting factors which influence language learning, but they are difficult to generalize and even more difficult to apply and implement in a computational context such as student modeling. Comprehensive, generalizable, and robust findings in SLA can provide a solid basis for student models in CALL.

Addressing this dilemma of quantitative vs. qualitative research methods, integrative approaches to SLA have been suggested more recently (de Bot, Lowie, & Verspoor, 2005, 2007; Ellis & Larsen-Freeman, 2006; Larsen-Freeman, 1997, 2000, 2003). These approaches, which can be referred to under the umbrella of Dynamic Systems Theory (DST) have in common that they view language learning as a dynamic system: a system which changes over time and is changed by individual speakers in communicative events. The DST perspective conceptualizes second language acquisition as:

- nonlinear, i.e., it incorporates spurts, progression, plateaus, fossilization, and retrogression, …;

- nonperiodic, i.e., segments of the language learning process will not be repeated and will not recur, but smaller segments of the learning curve might be similar to larger segments—a self-similarity of the plot of learning events which has often been described as fractality;

- nonmonotonic, i.e., the speed of L2 acquisition and / or attrition varies over time;
- complex, i.e., a large number of variables will have to be considered at any given point in time, their correlations have to be taken into account as well as the changes they undergo through the interaction with other variables and over time.

From a DST perspective, initial conditions—and even tiny differences in these initial conditions—often result in large differences in the end state of the system due to the complex nature of the dynamic system. Based on observation of this complex system, it appears to be impossible to predict the quality of the end state of the system, e.g., it is impossible to predict after observing an early language learner whether she will ever reach near-native proficiency and what this proficiency will look like for her. However, given meaningful data points over time, it should be possible to predict the nature of the next state as a (mathematical) function of the path through prior states. To use an example from another dynamic system, it is possible today to make relatively reliable predictions about the weather of the next twelve to twenty-four hours, based on weather data of the last hundred years or so, but it is impossible to predict the moment when the weather system is coming to a rest (when the weather will not change any longer) and the quality of that end state nor is it possible to predict the exact nature of the weather at some point in time in the more remote future (Lorenz, 1993). Similar claims can be made about language learning: It is our hypothesis that it is possible to make valid predictions of the next state of a particular learner's language learning system, but it will be impossible to make predictions about the state of the language learning process in the more remote future.

The DST approach to SLA is relatively new with a growing number of researchers entering into the discussion. There are promising results however in other social sciences: cognitive psychology and first language acquisition (Hollich, 2000; Hollich et al., 2000; van Geert, 1994, 1997, 1999, 2000; van Geert, Verhoeven, & van Balkom, 2004), pedagogy (Haggis, 2005), and bilingualism (Herdina & Jessner, 2002). The application of the underlying philosophical approach in DST to SLA—the consideration of a multitude of variables[ii] (or at least their concatenation) in context—promises a framework for a more integrative conceptualization of language learning. Moreover, DST has a mathematical basis, which may provide a basis for its computational implementation and therefore an impetus for student modeling in CALL.

BRIEF EXCURSION INTO LANGUAGE ASSESSMENT

A DST conceptualization of SLA may provide a promising basis for student models, but on the surface it presents a challenge for testing procedures. If we want to measure complex variables such as proficiency or even communicative competence then we need to gain an understanding of the complexity of the learner and of the language they use. Classical testing theory relies on calibrating the difficulty of decontextualized (often sentential) test items and to measure students' ability as a reverse proportional function of the difficulty of the most difficult items correctly answered. However, resulting scores provide only a holistic estimate of proficiency or only one aspect of proficiency (see Jang, this volume). In order to develop assessments that can provide more fine-tuned diagnostic information that can play a role in individualized CALL, we need to be able to model learners in their complexity. For example, if the student model can predict the nature of the next language learning system state the learner is likely to be in, we can present the learner with a test item which corresponds to that state.

Given our notion of language learning as a dynamic system, we are able to view a language test as a complex language learning event which 'interacts' with prior language learning events and whose test items interact with each other. A detailed linguistic analysis—in computer-assisted language testing preferably a computational analysis—is a necessary prerequisite for the analysis of test results as a whole. And again, modeling the language learning of individual students in context is a key to successful adaptive (diagnostic) testing in an integrative, communicative approach to SLA. What could such models look like?

APPROACHES TO STUDENT MODELING

Student models gather and structure information about the student's knowledge. In computational terms, a student model can be defined as a data structure that contains information about the student. "But we cannot directly observe what a student does and does not *know*; this we must infer, imperfectly, from what a student does and does not *do*" (Mislevy & Gitomer, 1996, p. 253). Thus, a distinction has to be made between a surface level student model which "represents the scheduled problem solving plans and applied procedural knowledge" and a deeper level student model which "must infer and model the student's belief by interpreting the surface level student model" (Villano, 1992, p. 469).

The simplest way of maintaining such a data structure is by recording performance data in the form of scores. These scores, however, do not contain any information about the kind of knowledge that has been acquired; they only reflect how much knowledge has been gained (Gisolfi, Dattolo, & Balzano, 1992, p. 329). For example, looking at the score of a simple grammar test conducted in the foreign language classroom, we could note that student X answered 80% of the questions correctly. This result might be about 20% higher than that of a comparable previous test (probable knowledge gain). On the other hand, if we looked at this score later, we would be unable to ascertain what exactly the knowledge items were the student had acquired or not. Most student models capture the knowledge state(s) of the student relative to the domain of learning. Far fewer incorporate individual characteristics of the learner (Milne, Shiu, & Cook, 1996) because individual characteristics are much more difficult to obtain (Mabbott & Bull, 2004).

Student models can be used in a number of ways (Elsom-Cook, 1993): corrective (providing tailored feedback), elaborative (extending the knowledge of the student), strategic (guiding decisions on teaching interventions), diagnostic (to determine the knowledge state of the student), predictive (anticipate future student behavior) and evaluative (assessing the level of student achievement). The diagnostic and evaluative uses can be incorporated into CALL programs or can function in language assessments.

A number of different approaches of creating and maintaining a student model facilitate the gathering of detailed information about the language learner and her learning process and avoid the exclusive reliance on score-keeping. Here, I will briefly discuss the bug library technique, the model tracing technique, and constraint-based student modeling.[iii]

The Bug Library Technique

The bug library technique comprises error descriptions of student errors and their explanations. Murray (1999, p. 99) distinguishes two different ways of recording student bugs: runnable models (student knowledge as subset of expert system rules plus some buggy rules) and overlay models (assign competency or probability to different rules according to inferences the system has made). A version of the bug catalogue is the so-called perturbation approach. Reyes (1998, p. 330) suggests this approach to student modeling for the domain of Pascal programming. This technique does not rely on a set of buggy rules, that is, anticipated student errors. Instead, it uses transformations of rules that the expert system possesses. She applies perturbation, that is, a set of meta-rules which modify operators, delete sub-expressions, exchange operands and alter variables (Reyes, 1998, p. 330). Perturbation or other modeling techniques discussed later are applied because the error descriptions are very expensive to create since they are built on empirical analyses of errors previously encountered. The error libraries are also restricted in that they are not transferable from one student population to another. Errors often occur because the student applied a similar rule, schema, or pattern to a new problem,[iv] or because the learner employed an existing correct rule which is not appropriate for the problem or the context of the problem at hand (Burton & Brown, 1982). How problematic this is in language learning and teaching becomes apparent if one considers that the number of utterances in any language is infinite and because each of these utterances could at least contain one error, the number of erroneous utterances is also infinite. Programs with a built-in expert system—in Intelligent CALL (ICALL) systems, with their application of artificial intelligence techniques to CALL, this is usually the case with an in-built natural language parser—have been based all too often on the assumption that the student's knowledge is simply a subset of the knowledge of the expert system (e.g., Cerri, Cheli, & McIntyre, 1992). Accordingly, the main function of the system is to impart the complementary subset of facts and rules onto the student. This is, of course, a fallacy and simplifies the teaching and learning process (Burton & Brown, 1982, p. 51). If one compares facets of the bug library approach, it is apparent that there are interesting overlaps with the contrastive and error analyses approaches in SLA. Thus, it is difficult to apply this student modeling technique in the context of current discourses in SLA.

The Model Tracing Technique

The model tracing technique monitors each step the student takes in the problem solving process instead of attempting to infer from final answers. "The student is modeled as the set of rules which matched his or her steps in the traced performance" (Ohlsson, 1992, p. 433). The technique thus depends on a set of correct and incorrect rules and has to rely on anticipating the rules that might get violated. Tasso, Fum and Giangrandi (1992) developed a version of this technique, that is, of backward model tracing. It was implemented for a later version of ET (English Tutor), which concentrates on verb conjugation in English. Backward model tracing utilizes all techniques of model tracing, but does not rely "on an a priori established catalogue of correct and incorrect productions but is able to dynamically generate mal-rules necessary to explain the student performance" (Tasso et al., 1992, p. 154). The student input is compared with a version of the same utterance which is generated by not just relying on the expert system, but also on the information already contained in the student model. Accordingly, the actual student performance is compared to the expected student behavior as predicted by the learner model. Tasso et al. (1992) state:

> If the two answers are equal, the Modeler assumes that the student has applied the same knowledge utilized in the simulation process and this constitutes a useful piece of information for discriminating among possible hypotheses still active from preceding cycles. On the other hand, if the two answers are different, the Modeler executes the two analyses of commission [application of an inappropriate rule] and omission rules [ignoring of a necessary rule], which will eventually produce new hypotheses about the student knowledge. (Tasso et al., 1992, p. 158)

It is not surprising that such an application of model tracing was to language learning in Krashen's (1982) narrow sense of learning (vs. acquisition), e.g., to the learning of individual grammatical rules. Only such learning procedures can be described algorithmically, whereas the production of chunks of discourses in meaningful communication cannot be broken down into a neat sequence of steps the learner or text producer has to follow. In other words, if the system is intended to model the different steps students take when learning grammatical knowledge items, model tracing appears to be a viable option, whereas if the CALL system is intended as a tool or tutor (Levy 1997) in communicative language teaching, it is not necessarily possible to establish and trace an ordered sequence of deterministic steps the learner took or will have to take. This limitation led ICALL researchers to the exploration of constraint-based approaches, which came from formal linguistics and robust parsing, in order to decrease the reliance on (error) anticipation.

Constraint-Based Modeling

Constraint-based modeling was originally proposed by Ohlsson (see e.g. Mitrovic, 1998, p. 415). This approach concentrates on learner errors and attempts to correct them. It presupposes that diagnostic information is not attained from the (intractable) learning problem solving process the student undergoes. Instead it is obtained from the problem state at which the student arrives or the final results. The constraint-based approach has similarities with the bug library technique described above. Both start with a knowledge base in the expert system. The bug library technique generates possible deviant solutions with meta-rules that transform rules and facts from the expert system. The constraint-based approach relaxes these constraints during the analysis in order to determine the rule(s) that might have been violated. Each constraint consists of a relevance condition that, in turn, determines when to apply the satisfaction condition (Ohlsson, 1992, p. 437). For example, if the parser finds what could be a direct object, then the phrase in question needs to be marked for accusative case. If this constraint is relaxed, however, the parser would successfully parse the sentence containing the direct object even though the object might be marked with the dative case. The parser would record that a phrase that should have been accusative-marked was actually dative-marked to later on have, for example, a basis to provide feedback to the learner about the deviant case-marking. The constraint-based modeling approach is very efficient because it does not rely on the anticipation of student errors as an intractable problem (Heift, 1998; Heift & Nicholson, 2001; Heinecke, Kunze, Menzel, & Schröder, 1998; Menzel, 1988, 1992a, 1992b; Menzel & Schröder, 1998; Schröder, 2002; Schulze, 1998, 1999, 2001; Vandeventer, 2001). However, it poses a set of different problems because it has to work with an immensely large search space. For example, parse forests, that is, all syntactic trees representing one and the same sentence or text fragment, get increasingly larger depending on the number and kind of constraints that can be violated. This potentially prolongs the analysis of student input. But, more importantly, it requires a selection of the most appropriate analysis of this input for feedback generation and for the maintenance of the student model. This, in itself, is a very complicated and time-consuming task.

This modeling approach bears some similarity with earlier conceptualizations of interlanguage (Selinker, 1974, 1992) in that it starts off with the assumption that the ultimate goal of language learning—the target variety—is the standard variety used by L2 native speakers and that surface structures of this variety are the main ‘content’ for acquisition. This modeling approach pays close attention to interlanguage processes such as transfer, overgeneralization, and simplification, but would have significant problems with what Corder described as errors of appropriateness (Corder, 1974), i.e., how communicative intentions were met successfully or otherwise.

We are not aware of any other modeling approaches that have been applied to CALL successfully (Heift & Schulze, 2007). We could assume that it would be necessary to rely on machine learning approaches to capture the dynamics and the complexity of the second language acquisition processes, but this remains a question which will have to be answered by future research although some problems of modeling learning in such domains have already been addressed in student modeling research.

General Problems with Student Modeling

McCalla et al. (2000) believe that learner modeling "should be easier than in the past given the vast amount of information that will be available about learner interaction in the emerging information technology intensive world" (p. 61). Accordingly, this conceptualization of a student model views modeling as a computation, that is, a continuous, fragmented process rather than a data structure. As a result, the immense development costs of learner models, which others have tried to achieve with generic student models, are reduced. The challenge of this approach to student modeling, however, does not primarily lie in the information collected about the learner, but in the fact that the large amount of information available needs to result in a coherent student model in spite of its many facets and traits. Student models are more complex than other user models because misconceptions and inconsistencies in the student's knowledge have to be considered. Mitrovic et al. (1996) identify four different sources of this noise: Inconsistent student behavior, dynamic and nonmonotonic nature of human learning, ambiguity of possible explanations for the observed behavior and indeterminacy of student answers. Certain aspects of observable student behavior can only be described as stochastic. As discussed in the previous section, we are addressing this challenge of complexity and nonlinearity through our reliance on DST approaches.

The following are a few possible reasons for students' inconsistent beliefs:

- In language learning, students are often testing hypotheses (Output Hypothesis - Swain, 1985) and they expect feedback on these attempts. There are instances when students deliberately employ a wide variety of surface structures which—according to their hypothesis—convey the same meaning to learn which of these will be accepted.

- Students are often unable to consider all relevant issues simultaneously when solving a problem because nobody has full access to one's full body of knowledge at all times. Considering issues such as conflicting communicative (sub)goals, meaning and form of lexical items, word order rules, subject-verb agreement, subcategorization of verb arguments, case-marking of noun phrases, … all at the same time results in nonlinear, inconsistent behavior.

- Students react adversely to external pressures. Such pressures often stem from testing situations but also factors such as fatigue, lack of motivation can play a role. It is also possible that one source of inconsistent student behavior can be attributed to the fact that students are using computers for language learning. Low levels of computer literacy and lack of typing skills can adversely affect language learning behavior. On the other hand, a preference for working with computers, a positive attitude to computer-mediated communication, for example, can also be conducive to language learning.

These are some of the factors which will have to be considered when conceptualizing a student model for language learning.

Conceptualization of the Mocha Student Model

We are basing our approach to student modeling in the Mocha project[v] on an integrative and balanced approach to SLA (Larsen-Freeman, 2000) by relying heavily on DST

(Lorenz, 1993; van Geert, 1994; Williams, 1997) to explain individual language learning processes in context. The learnt language in our case is German, but we are hoping that our findings will be applicable to the learning and teaching of other languages. The goal of the Mocha project is to build a student model which is informed by current thinking in SLA. The dynamic complexities of second language learning processes for individual learners in the context of computational modeling in CALL necessitates theoretical, analytical framework, which on the one hand is capable of considering the evolving complexity of variables that influence this learning process and on the other is sufficiently formal—in a mathematical sense—to provide the basis for a computational implementation. This is the reason why we are looking to DST to inform our analysis and modeling of individual language learning processes over time.

As has been outlined above, this makes it impossible to employ modeling techniques which rely on the anticipation of errors (bug library and model tracing). The problem with the large search space when using relaxed constraints makes the adoption of this modeling technique very hard. We are currently experimenting with an approach which borrows from machine learning, i.e., the system will be capable to 'learn' new linguistic information and new modeling rules and then be able to handle unanticipated student output. The linguistic analysis is informed by a formal variant of construction grammar[vi] (Kay, 2002). It is our hypothesis that employing this grammatical formalism—with its exclusive reliance on the construction as the unit of analysis and unification to explain the relation of more or less grammatically and lexically specified constructions—will help us to avoid limitations of modeling methods which either rely on error anticipation or on constraint relaxation by identifying well-formed and ill-formed constructions without attempting to anticipate either and by unifying learner constructions which are already known to the learner model. This approach to construction grammar-based parsing is currently only at the stage of conceptualization since the necessary formal approaches to construction grammar are relatively recent and hardly any implementations have been documented in the literature to date.

A formal approach to grammar is, of course, a necessary prerequisite for its computational implementation. However, it is the fact that construction grammar is usage-based and therefore also considers pragmatic features of constructions in addition to morpho-syntactic and semantic features which led to its successful application in first language acquisition research (e.g., Tomasello, 2003). Attempts have also been made to employ it for linguistic analyses in SLA (e.g., Haberzettl, 2007). Tomasello (2003), for example, was able to show that young children acquire their first language by repeating holophrases first—short constructions which consist of fixed lexical material and have to be used with the same meaning and context. Later they manipulate acquired constructions and produce item-based constructions, in which clearly defined parts of the construction are substituted with other suitable items, to then arrive at the level of abstract construction: "a form-meaning pair (F, M) where F is a set of conditions on syntactic and phonological form and M is a set of conditions on meaning and use" (Lakoff 1987, p. 467; quoted in Fischer & Stefanowitsch, 2007, p. 5). Constructions are not just manipulated over time, but they can also be merged or blended. It is our hypothesis that this acquisition order from more concrete to more abstract constructions can also be applied to SLA and that the usage-based construction grammar is a useful tool for textual analysis under the DST approach. But how can complexity of the language learning system be captured for a student model?

Such complex, dynamic systems, then, can be described by separating data points—the value of variable at a selected system state—and plotting these values in a time series graph: $Y = \{Y_t : t \in T\}$. If the lag time between neighboring data points is identical, then the time series can easily be converted to phase space. Here each data point is seen as a function of a previous data point: $f(x_{t+1}) = a \cdot x_t$. In other words, instead of plotting data points over time, they are plotted 'against each other': $\{(x_t, x_{t+1}), (x_{t+1}, x_{t+2}), (x_{t+2}, x_{t+3}), (x_{t+3}, x_{t+4}), ...\}$. This basically means in our context that the language learning process is plotted in such a way that a language learning event in the past is seen as the starting point or basis for the current language learning event or—to illustrate the predictive power of the student model—the current or past language learning events are understood to be the basis for a language learning event in the immediate future. So, for example, if we have plotted all constructions—by labeling each construction with a numeric identifier first—that a student in an elementary language class has used in the first four weeks, we would find a number of constructions with noun phrases in the nominative or accusative case. We could then make predictions about how to facilitate the learning of indirect objects and their dative-marking, considering this student's understanding of case-marking in the context of her understanding of grammatical phenomena such as word order, verb conjugation as well as based on the evidence the model has collected from the student's text on individual characteristics such as willingness to communicate (based on the frequency and length of her textual contributions relative to her peers) and her willingness to take risks with structural text elements (diversity of lexical and grammatical material identified thus far).

The question which arises is what variable(s) should be measured for plotting at each data point. Our goal of an integrative and balanced view of SLA prevents us from selecting one quantifiable variable and from ignoring the context by trying to eliminate all other variables and their interaction with each other over time. On the other hand, we cannot afford to plot a multitude of variables over time. Our phase space would have as many dimensions as we have variables. The mathematics of multidimensional spaces is complicated and certainly beyond the grasps of this author; some of it is still not known or has not been proven yet. We therefore decided to measure one variable which shows traces of all other linguistic, contextual and individual variables which might play a role: text. We view text as a product (the text we analyze) and as a window onto text as process (the production of text which also reflects the learning process to a large extent). Text is, of course, a very complex variable and it shows traits of a multitude of text-external variables such as individual differences and instructional variables. In other words, by capturing and plotting a student's language use over time in form and their use of constructions in text, we are using the complex variable text as a window on the complex and dynamic learning process.

We are currently investigating the possibility of measuring the discourse complexity of learner utterances and plotting these over time to get a good approximation of the individual learning process. We selected utterances because they are the smallest linguistic sign with a communicative meaning and because we assume that they incorporate the influence of the dynamic variables that are relevant to this language learning event. In order to measure the complexity, we conduct a construction grammar analysis of each utterance and estimate the level of abstraction for each construction by calculating the frequency of the variants of that construction which the student produced before. Basically, we are attempting to get an approximation of the entropy in the text—how little or often constructions get repeated in text over time and how significantly one construction differs from another—with respect to the construction in question because we can then assume that text entropy and text complexity are proportional. Different instantiations of abstract constructions clearly result in a higher text entropy and are assumed to be an indication of a higher level of the complexity of the learner text. At the other end of the continuum, holophrase constructions which are identical to input material the student was likely to have seen are assumed to have a very low level of discourse complexity. Traditionally, this kind of complexity would have been described informally as the range of vocabulary and the range of grammatical constructions. The different complexity levels are then plotted in a phase space. This graph will give us some indication of what constructions learners used at different times in their language learning process and how they varied in discourse complexity relative to one another.

We have started analyzing written utterances students produced at various stages of a one-semester online course. We are developing our analytical tools—computational German construction grammar—concurrently with our data sets of language learning processes. Using this approach, we intend to model individual cases of complexity and use of learner language first. Also, we intend to examine later whether these individual differences have places of overlap and, if they exist, how they can be used to improve the predictive power of the model.

CONCLUSION

Information from such a student model can be used in diagnostic testing. It will provide a more holistic, balanced picture of the complexity of utterances the learner is able to produce in L2. Having some information about the complexity level an individual learner is at will also enable the system to provide better corrective, error-contingent feedback because the probability of mapping a well-formed and intended construction onto the learner's construction or utterance is much higher. Similarly, the probability of the system being able to suggest a suitable learning object based on a good evaluation of the learner is much higher if the system has information on the prior language learning path and an identification of strengths, weaknesses, and learning preferences.

As research at the intersection of DST, Construction Grammar, SLA, and ICALL is very new, many unanswered questions remain, and many claims await empirical testing. However, the philosophical and methodological slant of DST as well as Construction Grammar, their mathematical foundations, and their integrative nature hold great promise for further progress in student modeling in ICALL and, probably more importantly, for an improved, more comprehensive understanding of SLA processes.

REFERENCES

Antos, G. (1982). *Grundlagen einer Theorie des Formulierens. Textherstellung in geschriebener und gesprochener Sprache*. Tübingen: Niemeyer.

Burton, R. R., & Brown, J. S. (1982). An investigation of computer coaching for informal learning activities. In D. Sleeman & J. S. Brown (Eds.), *Intelligent tutoring systems* (pp. 79-88). London: Academic Press.

Cerri, S. A., Cheli, E., & McIntyre, A. (1992). Nobile: Object-based user model acquisition for second language learning. In M. L. Swartz & M. Yazdani (Eds.), *Intelligent tutoring systems for foreign language learning: The bridge to international communication* (pp. 171-190). Berlin: Springer Verlag.

Corder, S. P. (1974). The significance of learners' errors. In J. C. Richards (Ed.), *Error Analysis: Perspectives on second language acquisition* (pp. 19-27). London: Longman.

Corder, S. P. (1981). *Error Analysis and Interlanguage*. Oxford: Oxford University Press.

de Bot, K., Lowie, W., & Verspoor, M. (2005). Dynamic Systems Theory and Applied Linguistics: The Ultimate "so what"? *International Journal of Applied Linguistics*, *15*(1), 116-118.

de Bot, K., Lowie, W., & Verspoor, M. (2007). A Dynamic Systems Theory approach to second language acquisition. *Bilingualism: Language and Cognition*, *10*(1), 7-21.

Dörnyei, Z. (2005). *The psychology of the language learner: Individual differences in second language acquisition*. Mahwah, NJ; London: Lawrence Erlbaum Associates.

Ellis, N. C., & Larsen-Freeman, D. (Eds.). (2006). *Language emergence - Implications for Applied Linguistics. Special Issue of Applied Linguistics (27/4)*. Oxford: Oxford University Press.

Elsom-Cook, M. (1993). Student modeling in intelligent tutoring systems. *Artificial Intelligence Review*, *7*(3-4), 227-240.

Firth, A., & Wagner, J. (1997). On discourse, communication, and (some) fundamental concepts in SLA research. *The Modern Language Journal*, *81*(3), 285-300.

Fischer, K., & Stefanowitsch, A. (2007). Konstruktionsgrammatik: Ein Überblick. In K. Fischer & A. Stefanowitsch (Eds.), *Konstruktionsgrammatik. Von der Anwendung zur Theorie* (pp. 3-17). Tübingen: Stauffenberg Verlag.

Gisolfi, A., Dattolo, A., & Balzano, W. (1992). A Fuzzy Approach to student modeling. *Computers and Education*, *19*(4), 329-334.

Gleick, J. (1987). *Chaos: Making a new science*. New York, N.Y.: Viking.

Haberzettl, S. (2007). Konstruktionen im Zweitsprachenerwerb. In K. Fischer & A. Stefanowitsch (Eds.), *Konstruktionsgrammatik. Von der Anwendung zur Theorie* (pp. 55-77). Tübingen: Stauffenberg Verlag.

Haggis, T. (2005). *'Knowledge Must Be Contextual': Some possible implications of Complexity and Dynamic Systems Theories for educational research.* Paper presented at the Complexity, Science and Society Conference, Liverpool.

Heift, T. (1998). Designed Intelligence: A Language Teacher Model. Unpublished PhD Thesis, Simon Fraser University, Burnaby.

Heift, T., & Nicholson, D. (2001). Web delivery of adaptive and interactive language tutoring. *International Journal of Artificial Intelligence in Education*, *12*(4), 310-325.

Heift, T., & Schulze, M. (2007). *Intelligence and errors in CALL. Parsers and pedagogues*. New York: Routledge.

Heinecke, J., Kunze, J., Menzel, W., & Schröder, I. (1998). Eliminative parsing with graded constraints. In *Proceedings of Coling-ACL'98* (pp. 526-530).

Herdina, P., & Jessner, U. (2002). *A dynamic model of multilingualism: Perspectives of change in psycholinguistics*. Clevedon; Buffalo; Toronto: Multilingual Matters.

Hollich, G. J. (2000). Mechanisms of word learning: A computational model. Unpublished Dissertation/Thesis, http://www.il.proquest.com/umi, Univ Microfilms International, US.

Hollich, G. J., Hirsh-Pasek, K., Golinkoff, R. M., Brand, R. J., Brown, E., Chung, H. L., et al. (2000). Breaking the language barrier: An emergentist coalition model for the origins of word learning. *Monographs of the Society for Research in Child Development*, *65*(3), v-123.

Kasper, G., & Rose, K. R. (2002). *Pragmatic development in a second language*. Oxford: Blackwell.

Kay, P. (2002). An informal sketch of a formal architecture for construction grammar.

Grammars, 5(1), 1-19.

Klein, W. (1986). *Second language acquisition*. Cambridge: CUP.

Klein, W., & Dittmar, N. (1979). *Developing grammars. The acquisition of German syntax by foreign workers*. Berlin: Springer Verlag.

Klein, W., & Perdue, C. (1992). *Utterance structure (Developing grammars again).* Amsterdam/Philadelphia: Benjamins.

Krashen, S. D. (1982). *Principles and practice in second language acquisition*. Oxford: Pergamon.

Lado, R. (1957). *Linguistics across cultures: Applied Linguistics for language teachers*. Ann Arbor: The University of Michigan Press.

Lakoff, G. (1987). *Women, Fire, and Dangerous Things. What Categories Reveal about the Mind.* Chicago: University of Chicago Press.

Larsen-Freeman, D. (1997). Chaos/Complexity Science and second language acquisition. *Applied Linguistics, 18*(2), 141-165.

Larsen-Freeman, D. (2000). Second language acquisition and Applied Linguistics. *Annual Review of Applied Linguistics, 20*, 165-181.

Larsen-Freeman, D. (2003). *Teaching language: From grammar to grammaring*. Southbank, Victoria: Thomson/Heinle.

Levy, M. (1997). *Computer-assisted language learning: Context and conceptualisation.* Oxford: Oxford University Press.

Lorenz, E. N. (1993). *The Essence of chaos*. Seattle: University of Washington Press.

Mabbott, A., & Bull, S. (2004). Alternative views on knowledge: Presentation of open learner models. In J. C. Lester, R. M. Vicari & F. Paraguacu (Eds.), *Intelligent Tutoring Systems: 7th International Conference* (pp. 689-698). Berlin: Springer-Verlag.

MacWhinney, B. (1998). Models of the emergence of language. *Annual Review of Psychology, 49*, 199-227.

McCalla, G. I., Vassileva, J., Greer, J. E., & Bull, S. (2000). Active learner modelling. In G. Gauthier, C. Frasson & K. VanLehn (Eds.), *Intelligent Tutoring Systems. 5th International Conference, ITS 2000, Montréal, Canada, June 2000, Proceedings* (pp. 53-62). Berlin: Springer Verlag.

Menzel, W. (1988). Error diagnosing and selection in a training system for second language learning. In *Proceedings of the Twelfth International Conference on*

Computational Linguistics (pp. 414-419).

Menzel, W. (1992a). Constraint-based diagnosis of grammatical faults. In J. Thompson & C. Zähner (Eds.), *Proceedings of the ICALL Workshop, UMIST, September 1991* (pp. 89-101). Hull: University of Hull, CTI Centre for Modern Languages.

Menzel, W. (1992b). *Modellbasierte Fehlerdiagnose in Sprachlernsystemen*. Tübingen: Niemeyer.

Menzel, W., & Schröder, I. (1998). Constraint-based diagnosis for intelligent language tutoring systems. In *Proceedings of the IT&KNOWS Conference at IFIP'98 Congress* (pp. 484-497). Wien/Budapest.

Milne, S., Shiu, E., & Cook, J. (1996). Development of a model of user attributes and its implementation within an adaptive tutoring system. *User Modeling and User-Adapted Interaction, 6*, 303-335.

Mislevy, R. J., & Gitomer, D. H. (1996). The Role of probability-based inference in an intelligent tutoring system. *User Modeling and User-Adapted Interaction, 5*, 253-282.

Mitrovic, A. (1998). Experiences in implementing constraint-based modeling in *SQL-Tutor*. In B. P. Goettl, H. M. Halff, C. L. Redfield & V. J. Shute (Eds.), *Intelligent Tutoring Systems. 4th International Conference, ITS 1998, San Antonio, Texas, USA, August 1998, Proceedings* (pp. 414-423).

Mitrovic, A., Djordjevic-Kajan, S., & Stoimenov, L. (1996). INSTRUCT: Modeling students by asking questions. *User Modeling and User-Adapted Interaction, 6*, 273-302.

Murray, T. (1999). Authoring intelligent tutoring systems: An analysis of the state of the art. *International Journal of Artificial Intelligence in Education, 10*, 98-129.

Ohlsson, S. (1992). Constraint-based student modeling. *Journal of Artificial Intelligence in Education, 3*(4), 429-447.

Reyes, R. L. (1998). A domain theory extension of a student modeling system for Pascal programming. In B. P. Goettl, H. M. Halff, C. L. Redfield & V. J. Shute (Eds.), *Intelligent Tutoring Systems. 4th International Conference, ITS 1998, San Antonio, Texas, USA, August 1998, Proceedings* (pp. 324-333).

Schachter, J. (1974). An error in Error Analysis. *Language Learning, 27*, 205-214.

Schönefeld, D. (2006). Constructions [Electronic Version]. *Constructions, SV 1*, from www.constructions-online.de

Schröder, I. (2002). Natural language parsing with graded constraints. Unpublished PhD

Thesis, University of Hamburg, Hamburg.

Schulze, M. (1998). Teaching grammar - learning grammar: Aspects of second language acquisition. *CALL, 11*(2), 215-228.

Schulze, M. (1999). From the developer to the learner: Computing grammar - learning grammar. *ReCall, 11*(1), 117-124.

Schulze, M. (2001). Textana - Grammar and grammar checking in parser-based CALL. Unpublished PhD Thesis, UMIST, Manchester.

Selinker, L. (1974). Interlanguage. In J. C. Richards (Ed.), *Error Analysis: Perspectives on second language acquisition* (pp. 31-54). London: Longman.

Selinker, L. (1992). *Rediscovering Interlanguage*. London: Longman.

Swain, M. (1985). Communicative competence: Some roles of Comprehensible Input and Comprehensible Output in its development. In S. M. Gass & C. G. Madden (Eds.), *Input in second language acquisition* (pp. 235-253). Rowley: Newbury House.

Tasso, C., Fum, D., & Giangrandi, P. (1992). The Use of explanation-based learning for modelling student behavior in foreign language tutoring. In M. L. Swartz & M. Yazdani (Eds.), *Intelligent tutoring systems for foreign language learning: The bridge to international communication* (pp. 151-170). Berlin: Springer Verlag.

Tomasello, M. (2003). *Constructing a language: A usage-based theory of language acquisition*. Cambridge, Mass. ; London: Harvard University Press.

van Geert, P. (1994). *Dynamic Systems of Development: Change between Complexity and Chaos*. Harvester Wheatsheaf, Hertfordshire, HP2 7EZ: England.

van Geert, P. (1997). Nonlinear Dynamics and the Explanation of Mental and Behavioral Development. *Journal of Mind and Behavior, 18*(2-3), 269-290.

van Geert, P. (1999). *Vygotsky's Dynamic Systems*. Taylor & Frances/Routledge, Florence, KY, US, [URL:http: //www.routledge.com].

van Geert, P. (2000). The Dynamics of General Developmental Mechanisms: From Piaget and Vygotsky to Dynamic Systems Models. *Current Directions in Psychological Science*, *9*(2), 64-68.

van Geert, P., Verhoeven, L., & van Balkom, H. (2004). A dynamic systems approach to diagnostic measurement of SLI. In *Classification of developmental language disorders: Theoretical issues and clinical implications.* (pp. 327-348). Lawrence Erlbaum Associates, Publishers, Mahwah, NJ: US.

Vandeventer, A. (2001). Creating a Grammar Checker for CALL by Constraint

Relaxation: A Feasibility Study. *ReCall, 13*(1), 110-120.

Villano, M. (1992). Probabilistic Student Models: Bayesian Belief Networks and Knowledge Space Theory. In C. Frasson, G. Gauthier & G. I. McCalla (Eds.), *Intelligent Tutoring Systems. Second International Conference, ITS'96, Montréal, Canada, June 1992, Proceedings* (pp. 491-498). Berlin: Springer Verlag.

Williams, G. P. (1997). *Chaos Theory Tamed*. Washington, D.C.: Joseph Henry Press.

---

[i] Different theoretical approaches are subsumed here (for reasons of stylistic convenience rather than because they are identical): Dynamic Systems Theory, Chaos Theory, Complexity Theory, Emergentism (see e.g., Gleick, 1987; Lorenz, 1993; van Geert, 1994; Williams, 1997).

[ii] A multitude of variables influence the path of this process and potentially 'decide' about success or failure. It is the interaction of these variables which contribute to the "emergence of language" (MacWhinney, 1998).

[iii] For a more detailed discussion of different modeling techniques and student models in CALL see Heift and Schulze (2007).

[iv] Here I follow Antos (1982) who argued that although text production might not be a problem solving process, it is fruitful in Applied Linguistics research to depict it as such.

[v] Principal investigator: Mathias Schulze; Co-investigator: Trude Heift; The research is supported by the Social Sciences and Humanities Research Council (SSHRC) of Canada; grant number 410-2007-2549

[vi] Construction grammar is an umbrella term for a number of more or less formal approaches which all have in common that they view constructions as the central syntactic unit. For an overview, see (Fischer & Stefanowitsch, 2007; Schönefeld, 2006).