The Dynamic Systems Approach in the Study of L1 and L2 Acquisition: An Introduction

PAUL VAN GEERT
University of Groningen
The Heymans Institute
Grote Kruisstraat 2/1
9712 TS Groningen
The Netherlands
Email: p.l.c.van.geert@rug.nl

The basic properties and concepts of dynamic systems theory are introduced by means of an imaginary, literary example, namely, Alice (from Wonderland) walking to the Queen in *Through the Looking Glass* (Carroll, 1871). The discussion encompasses notions such as time evolution, evolution term, attractor, self-organization, and timescales. These abstract notions are then applied to the study of first language (L1) and second language (L2) acquisition. The steps that should be taken toward the construction of dynamic systems models of L2 learning are described, including a discussion of the difficulties and the theoretical opportunities that accompany this model building. Finally, a concrete example of model building and of the study of variability as an indicator of developmental transitions in L1 acquisition is given.

ABOUT 20 YEARS AGO, THE TERM *DYNAMIC systems* began to appear in the titles of articles in developmental psychology. Since then, the appearance of the phrase "dynamic systems" has been steadily but rather slowly increasing in the field, in terms of the number of publications and areas of application. Nevertheless, the dissemination of information regarding dynamic systems has not been very widespread. In my view, this is disappointing because I believe knowledge of dynamic systems can make a significant contribution to understanding development. The process of development cannot be truly understood if we confine ourselves to looking at it in the usual way, which is through investigating associations between variables across populations (for a discussion of this viewpoint, see van Geert, 1998a; van Geert & Steenbeek, 2005). Part of the explanation of why dynamic systems has not really taken off lies in the confusion that exists in developmental psychology about what dynamic systems are and what dynamic systems theory means.

The Modern Language Journal, 92, ii, (2008)
0026-7902/08/179–199 $1.50/0

I will first present a description of what I think dynamic systems means and show how it differs from the descriptive, explanatory, and methodological practices that prevail in the scientific disciplines that address psychological development, learning, and acquisition, including language acquisition. I will then explain how dynamic systems approaches the processes of change in general. I will then proceed with a discussion of complex systems and present some issues in the acquisition of first (L1) and second language (L2) that the theory of complex dynamic systems might fruitfully address.

THROUGH THE LOOKING GLASS: BASIC PROPERTIES OF DYNAMIC SYSTEMS

In Lewis Carroll's (1871) *Through the Looking Glass*, the following intriguing dialogue appears:

> 'Well, in our country,' said Alice, still panting a little, 'you'd generally get to somewhere else—if you run very fast for a long time, as we've been doing.' 'A slow sort of country!' said the Queen. 'Now, here, you see, it takes all the running you can do, to keep in the same place. If you want to get somewhere else, you must run at least twice as fast as that!' (p. 25, 2005 ed.)

To provide some context, Alice and the Queen move about in the Queen's "most curious country," with ". . . a number of tiny little brooks running straight across it from side to side, and the ground between . . . divided up into squares by a number of little green hedges . . . ." (Carroll, p. 22, 2005 ed.) The country resembles a giant chessboard, formed by the squares of land divided by the brooks. The story, which is a sequel to *Alice in Wonderland*, is about Alice, a girl who passes through a mirror to find herself in a country where everything seems backward. Now let us try to look at this little scene from the viewpoint of dynamic systems.

The State and State Space of a System

To begin with, we must identify the main variable of interest here, which is the "place" or "position" in the *Looking Glass* world that is occupied by either Alice or the Queen. The places or positions form a *system*. A system can be defined as any collection of identifiable elements—abstract or concrete—that are somehow related to one another in a way that is relevant to the dynamics we wish to describe. Trivially, the possible places or positions in the chessboard world are related through distances and specified by at least two measurement dimensions, longitude and latitude (or the geographical x and y axes). At any point in time, the presence of Alice, or of the Queen for that matter, will define a particular place, which is the position she occupies at that time, for instance, under the tree or on a particular square surrounded by green hedges. The *state* of system, then, is defined by Alice's position (or the Queen's, or both, depending on what one wishes the system to contain) on a particular date at a particular time, for instance, September 22, 2006, at 12:09:01 p.m. The chessboard landscape is the *state space*, in other words, the space of possible states or positions of Alice. Alice's—or the Queen's—position can be specified by two numbers, one for the longitude, one for the latitude. That is, the state space is a two-dimensional manifold (a two-fold, as Alice would probably call it).

The Time Evolution of a System and the Evolution Rule

What matters to Alice in the first place is how to get from one place to another. That is, if Alice is in one place, what does she do to get to another place, and how does she get there? Alice provides the answer: run. To run means to take steps, one at a time and in succession, in a particular direction and with a particular speed (in Alice's "slow" country, walking is a form of running slowly, and standing still is a form of running VERY slowly). If you run in Alice's country, your position on the map will change over time. The change of positions over time is called the *time evolution* of Alice's position (including standing still, if you run VERY slowly) and the taking of one particular step at a time is the *evolution rule* or *evolution term*. That is, if you apply the evolution rule—taking a step—to any particular place in the chessboard land, it will bring you to another place in the chessboard land (and since this is a part of the *Looking Glass* world, some steps will make you stay in the same place). To be of interest, the evolution rule must be applied iteratively. For instance, as Alice enters into the *Looking Glass* world (the first position or first state of the system), she must try to get to the Queen, where she will have the discussion that we read in the first quote. Alice must take one step to get to her next position, then from that position take another step to get to the next position, and so forth. Why describe Alice's path from the entrance to the Queen in the form of such small steps? The reason is obvious: The way Alice gets to her destination is by taking one step at a time. If you want to understand Alice's path, you must understand it as an iterative sequence of steps.

Constraints and Parameters

To understand why she makes the detour, we must know that in order to take a step you must have a solid, supportive ground, and water does not have those qualities. (If Alice were a bird, the evolution rule of her position would follow very different constraints and thus lead to very different time evolutions, such as crossing the brook without a bridge.) The necessity of solid ground is an example of a *constraint* that governs the application of the evolution rule. Many other such constraints can be specified; a particularly interesting one could be that it requires less effort to take a step down a slope than up a slope. If the chessboard landscape is very hilly, we should in fact expand our manifold to three dimensions, including height, in order to be able to explain why Alice prefers a particular path (for instance, avoiding steep slopes). But if the slopes are weak, we can just as well stick to the two original dimensions, since adding height will not add to our understanding of Alice's path.

The evolution rule—the taking of a step—requires three *parameters* for its specification. One is the length of the step, another is the frequency (how many steps per time unit), and the third is the direction of the step. Some parameters can be *constants*, such as the length of the step, which, for the sake of simplicity, is assumed to depend solely on the length of Alice's legs. Other parameters are variable, in that they can be made to vary, for example, by taking more steps at a time (running). These are called *control parameters* because they control, that is, determine, the form of the path that Alice will take. For instance, if each step (of similar length, as we agreed on) changes direction by 90 degrees in comparison to the preceding step, Alice will return to her point of departure after four steps.

In the chessboard landscape, the direction of Alice's steps is determined by the position of the Queen because that is where Alice wants to go to. Now we are getting closer to complex systems, as the direction of Alice's steps seems to be determined by other elements in the system, which is the position of the Queen relative to Alice's own position, and Alice's wish or intention to approach the Queen, which we have so far taken for granted.

Complex systems are systems with many components that interact, meaning that they codetermine each other's time evolution. Adding the Queen makes the Alice "system" just sufficiently more complex to understand the notion of a *control system*. As Alice wants to get to the Queen, the direction of her steps is determined by Alice's perception of where she is relative to the Queen. A good rule of thumb (for the real world) is to try to take steps in such a direction that the view of your target position remains as close as possible to the middle of your visual field as you look in the direction of your next step. Interestingly, in the looking-glass world, Alice must walk in the opposite direction to arrive at the Queen.

Attractors and Perturbations

Based on this more complex evolution rule—taking one step at a time in a direction controlled in a particular way by the view of the Queen—Alice will arrive at the Queen's position after some time, irrespective of where she starts her journey, if there are no obstacles in the landscape. This point, the Queen's position, is the *attractor state* of the system, given the system's particular evolution rule. Since we are in the looking-glass world, we can take the liberty to do as Lewis Carroll did and change the rules to serve our explanatory needs. Let us assume that Alice can indeed move toward the Queen, while trying to keep the Queen's position at a 30-degree angle from Alice's central field of vision when looking ahead. With such an evolution rule, Alice can start from any place in the chessboard land and arrive at a place where her path will automatically begin to circle around the Queen because that circle maximally obeys the 30-degree rule. In this strange case, the system's attractor is a *cyclical attractor*; that is, given the particular evolution rule, Alice will end up in a circle around the Queen, regardless of where she starts.

Note that since we describe the chessboard landscape in the form of a two-dimensional manifold (longitude and latitude), our system of places does not include information about brooks or hedges. It is as if we observe an Alice who is equipped with a global positioning system transmitting her coordinates to us every second, thus reducing the actual landscape to something that appears as a black box to us, the observers. However, if we know the evolution rule that governs Alice's path and we notice that at some places she diverges from the path that the evolution rule predicts, we can be pretty sure that this is because there are *obstacles* in the landscape, such as hedges, brooks, bushes, poison ivy, and so forth. (For simplicity, I will continue to assume that it is still Alice's intention to get to the Queen.) Assuming that we have no direct observational access to these obstacles, we can infer their existence on the basis of the divergence between the real path and the path that should be followed if the evolution rule could be applied without disturbance. When Alice gets stuck somewhere, for instance, because she sees her path blocked by brooks and other obstacles, she will most likely tend to move away from them in an attempt to find a path that leads to her goal. Such places are the opposite of attractors; they are *repellors*, or places that push Alice away from them. However, their function as repellors depends on Alice's evolution rule, the rules that govern her steps.

Another term for an external disturbance of a path or time evolution of a system is *perturbation*. For instance, we can block a frequented path, but given Alice's evolution rule, it will be fairly simple for her to bypass the obstacle and reach her goal. We also can build a wall around Alice and observe that she is no longer able to reach the Queen. The effect of perturbations on the course of the trajectory can thus be very informative of the nature of the rules that govern the dynamics.

Self-Organization and Emergence

Technically speaking, the chessboard landscape is a multidimensional or even infinite-dimensional manifold, consisting of dimensions describing each place in the landscape as being occupied by pavement, hedges, poison ivy, water, and so forth. We can specify all the places in the landscape in terms of their likelihood of being visited by Alice. Some places, like where the Queen stands, or some dead-ends in the landscape where Alice will get stuck (e.g., between two brooks), will stand out as unique features, that is, salient attractors or repellors, in the landscape. Likewise, some paths will occur more often, for instance, because they are easier to walk on. As time goes by and Alice tends to visit the Queen more often, some of the paths will become more and more salient, as Alice's steps create clearly visible paths through the grass. Thus, before Alice ever tries to reach the Queen, the chessboard landscape contains literally an infinite number of possible paths to the Queen. However, given Alice's principle of walking, the infinite number of possibilities will soon be reduced to a small number of actual trajectories or paths. That is, the landscape gets organized into a limited number of endpoints and a limited number of paths. Thus, after a while, the landscape becomes more orderly, more structured or organized, in terms of how it relates to Alice's walking. This organization, that is, the reduction of the number of most likely paths, is a direct consequence of the (mostly hidden or unknown) properties of the chessboard landscape and of Alice's dynamics. For this reason—the organization being a direct consequence of the dynamics itself—we might call it *self-organization,* which amounts to the spontaneous formation of *patterns*.

Self-organization is a characteristic property of (complex) dynamic systems. Self-organization has interesting properties: It is flexible and it is adaptive. Thus, if the Queen's pawns were to change the position of the brooks and hedges overnight (a nice example of a perturbation), Alice would soon create alternative paths, hence the flexibility of this self-organization. The number of preferred paths is likely to be small in any case, but there are many possibilities for establishing such paths. The adaptive nature of the self-organization implies that the paths are optimally adapted to the nature of the evolution rules, Alice's walking, but this adaptation arises from the fact that the means (the preferred paths) are to a great extent created by the function (the walking to the Queen). Self-organization, the spontaneous occurrence of patterns due to the dynamics itself, is a particular form of *emergence* (Crutchfield, 1994; Pessa, 2004).

Emergence is the spontaneous occurrence of something new as a result of the dynamics of the system. To give a different sort of example, as long as Alice walks below a certain speed, her pattern of locomotion will be walking. However, as Alice increases the speed with which she moves one leg in front of the other, her walking will suddenly shift toward a new pattern of locomotion, running. The walking and running patterns are not built into Alice's legs or brain; they emerge under the influence of the control parameter, which is, in this case, speed. The speed of putting one leg in front of another is the control parameter here. A continuous increase of the parameter, such as slowly driving the speed up, results in a sudden shift to a new pattern of locomotion (running), and the speed at which this occurs is the *critical* parameter value. The pattern of locomotion (either the typical walking or the typical running pattern) also can be called a *phase*, and the sudden shift from walking to running or from running back to walking is a *phase transition*. If we ask the walking Alice to increase her speed, she will make the phase shift from walking to running at, say, seven Wonderland miles per hour (the critical value). Now we ask Alice to slowly reduce her speed, and we observe that she will shift from running to walking at the speed of five Wonderland miles per hour. That is, the critical value for walking to running is different from running to walking. This asymmetry is called *hysteresis*. It implies that transitions (switching from one state to another) depend not only on the controlling variable (such as Alice's walking speed) but also on the system's history (whether she goes from running to walking or from walking to running).

Nonlinearity and Stability

So far, I have not said anything about a peculiar property of the chessboard landscape described earlier. From the story, it is clear that Alice can reach the Queen from her starting point, but once she is near the Queen she undergoes a strange influence of the royal highness on the outcome of her taking steps: "Now, here, you see, it takes all the running you can do, to keep in the same place," says the Queen, and indeed, whereas a step far away from the Queen brings Alice a bit closer to the Queen, a step in the vicinity of the Queen leads to remaining in the same place. We can say that the effect of taking a step on the crossing of a certain distance is a *nonlinear* property of the distance from the Queen, with the effect decreasing to zero once Alice is near the Queen. Simply said (but not entirely faithful to the original mathematical meaning), nonlinearity means that the effect of the step is not proportional to the distance that needs to be crossed. Proportionality would mean that Alice makes greater strides the farther away she is from the Queen, thus implying that the length of the step becomes zero if the distance from the Queen is zero. But this is not a linear system: We have seen that as Alice walks away from the Queen every (same) step brings her closer to the Queen, and as she is near the Queen, every (same) step leaves her where she is, at the same place. This way of remaining in the same place is a beautiful illustration of *dynamic stability*, or stability in a dynamic framework. In the linear case, stability is literally static: If you do not take a step, you do not move. That is, in the linear case, stability is the absence of dynamics. But in the nonlinear Wonderland, in the vicinity of the Queen, stability, or remaining in the same place, is the effect of taking steps (and doing so with great speed, for that matter). Hence, a dynamically stable state of the system is defined as a position in space where a step from a particular place moves you to that same place. Although this seems like a very weird and unnatural property, worthy of only a truly Lewis Carollian universe, it is in fact one of the most central features of the living world as we know it, where standing still (stability) is just as dynamic as moving. Another example of nonlinearity is the hysteresis effect described in the preceding section: The critical value depends on the direction from which you reach it.

Timescales

Although the issue of timescales does not explicitly feature in *The Looking Glass*, it is not difficult to imagine a situation in which the distinction becomes apparent. Imagine Alice saying that she walks from Point A to Point B, and some *Looking Glass* character—for instance, the *Looking Glass* tortoise—then objects that she does not walk, but that what she does is raise one foot, bring that foot to another place, shift her balance, raise another foot, and bring that foot to yet another place, and that all this raising and placing is all there is. Alice and the tortoise are clearly referring to different *timescales*. Alice's is the long-term timescale of locomotion, going from place A to B, and the other is the short-term timescale of taking a step and iteratively combining these steps into the act of locomotion. Thus, the dynamic constituents of the long term are single steps, and the dynamic constituents of single steps are movements of the limbs and body. To a certain extent, it makes no difference to the long-term timescale how, exactly, Alice takes the steps. If Alice had one leg tied, she could hop from A to B, and if she had hurt her foot and were walking with crutches, she would sort of limp from A to B. (Note that these patterns, hopping for instance, are typical examples of self-organizing patterns, given particular constraints, such as the tied leg.) In each case, her path would not lead across brooks or ponds or over walls (which would be the case if the dynamics of her locomotion constituted a form of flying, for instance). However, the short-term dynamics of her locomotion will affect the long-term dynamics in subtle but interesting ways. With the crutches or tied leg, for instance, Alice would automatically tend to avoid rough terrain that she could easily navigate if both legs were unhampered. Thus, the form of the long-term path will reflect the constraints and possibilities of the short-term dynamics of her walking. In fact, if we for some reason are unable to directly observe the short-term dynamics, some of the observable properties of the long-term dynamics (the sort of path Alice takes) can suggest the nature of the underlying short-term dynamics.

Summary: Through the Looking Glass to Development

Of course, the *Looking Glass* world is entirely fictional and, at times, physically impossible. Its properties nevertheless illustrate some of the basic principles of dynamic systems. I have used this fictional world to demonstrate that dynamic systems is not a specific theory but that it is a general view on change, change in complex systems, in particular, or, systems consisting of many interacting components, the properties of which can change over the course of time. The developing person, embedded in his or her environmental niche, is an example of such a complex system, which, for simplicity's sake, I will call the *developmental system* (Ford & Lerner, 1992). The properties that we saw in the *Looking Glass* world also will very likely apply to the developmental system: It will follow a particular time evolution depending on its evolution rules; its time evolution will depend on particular constraints and parameters; it will move toward particular attractor states and be affected by perturbations; it will be characterized by self-organization and emergence of a characteristic structure; its change will depend, in a nonlinear fashion, on particular driving forces; its stability, in the form of its attractors, will be as dynamic as its actual change; and finally, it will be characterized by different, interacting timescales.

In the next major section, I will discuss some possible applications of dynamic systems to the study of L2 acquisition. Where examples are not available, I will resort to illustrations from the L1.

DYNAMIC SYSTEMS AND L2

Resolving Potential Confusion

Dynamic systems already has a long history (see, e.g., van Geert, 2003, for an overview), and there is comparatively little disagreement in mathematics, physics, or biology as to what "dynamic systems theory" means. However, the situation is different for the developmental sciences. If one searches the developmental literature on dynamic systems—which is not extensive, but steadily increasing in terms of publications and fields of application—one is likely to first come across publications that define dynamic systems theory as a theory of embodied and embedded action (especially in the publications of Thelen and Smith; see, e.g., Smith, Thelen, Titzer, & McLine, 1999; Thelen & Smith, 1994). In essence, cognition, thinking, and action are explained as dynamic patterns unfolding from the continuous, "here-and-now" interaction between the person and the immediate environment. A particularly clear description, in the context of cognition and intelligence, comes from Smith (2005): "The embodiment hypothesis is the idea that intelligence emerges in the interaction of an organism with an environment and as a result of sensory-motor activity. The continual coupling of cognition to the world through the body both adapts cognition to the idiosyncrasies of the here and now, makes it relevant, and provides the mechanism for developmental change" (p. 205). The dynamic system at issue is the continuous coupling between the organism and its environment, showing a time evolution that takes the form of intelligent action.

A second line of thought emphasizes the major qualitative properties of complex dynamic systems. The developmental system is viewed as a self-organizing system, showing attractor states, nonlinearity in its behavior, emergence, and so forth. The inspiration for this view comes from the study of dynamic systems in other fields, which has demonstrated that such systems show the properties in question. An eloquent defender of this view is Marc Lewis, who has primarily focused on social interaction, emotions, and personality, and, recently, has shifted his attention to brain development (Lewis, 1995; Lewis, 2000; Lewis, 2005; Lewis, Lamey, & Douglas, 1999).

A third approach is the one that I have defended for about 15 years now, which is that dynamic systems is basically a very general approach to describing and explaining change, focusing on the time evolution of some phenomenon of interest, including the principles or "rules" that describe this time evolution (for general overviews, see van Geert, 1994, 2003; van Geert & Steenbeek, 2005).[1]

Fortunately, the differences in viewpoints among the embedded–embodied, the qualitative, and the general approach, if I may call them by these names, are smaller than they appear to be at first sight. (Although some points of disagreement remain, especially with regard to the way in which cognition, or language for that matter, should be studied in a dynamic systems framework.) If we apply the general approach to systems that are complex and sufficiently permanent—such as human beings, societies, and also more transient phenomena such as an action or a conversation—we soon will find out that they display a host of interesting qualitative properties, namely, emergence, self-organization, attractor states, nonlinearity, and so forth. Thus, the application of the general approach to systems that are of interest, say, to students of L2 acquisition, leads to the properties focused on by the qualitative approach.

The relationship with the embedded–embodied approach is a little more complicated and requires some understanding of the concept of dynamics on various timescales, as introduced in the preceding section. When a person says something in response to what another person has just said before, an observer can confine oneself to describing the utterance as a sentence, comprising a particular syntactic structure and categories. There is no doubt that this is an adequate (though certainly not exhaustive) description of what the observed speaker does at the time of uttering. The question is, of course: How does this sentence emerge, in the real context and timeframe of its occurrence? This is the sort of question that arose with Alice claiming to be walking (a shorthand description of a macroscopic pattern, so to say) and the imaginary tortoise claiming Alice was not walking but in fact moving one leg (or part of a leg) at a time. The tortoise was thus referring to the microscopic, short-term appearance of the macroscopic property called "walking." The short-term dynamics of language use and understanding is an example of the short-term dynamics of human action and communication, and its understanding requires an understanding of intentionality (goal-directedness, the interest-driven nature of action), and embodiment and embeddedness in the context (see, e.g., Clark, 1997, 2006; Clark & Chalmers, 1998; Harris, 2004; Love, 2004).

At present, we are only beginning to understand some of the short-term dynamics of how actions, including language, emerge through the confluence of internal abilities and external, contextual affordances. However, this relative lack of understanding at the "microscopic" level should not prevent us from trying to understand the dynamics of a particular phenomenon, such as L2 learning, at the macroscopic level of longer timeframes (e.g., months or years during which the L2 is learned). However, the short-term dynamics that explains the actual production and understanding of sentences on the spot is characterized by a number of interesting properties, one of which is termed "fluctuation and variability" (e.g., Thelen & Smith, 1994). These properties must be taken into account by any model focusing on longer term timescales and not be disregarded as, for instance, measurement error. I will revisit this issue later.

In summary, the first approach to dynamic systems discussed in this section refers to a particular, in general, short-term, timescale of phenomena that also can be dealt with on the long-term timescale of macroscopic change. However, the basic principles of dynamic systems theory in the general sense must apply to all timescales. (This is the reason connectionist network models, which operate on the timescale of input–output relations in the brain, are entirely compatible with dynamic systems principles and are thus examples of dynamic systems [Pessa, 2004; Thelen & Bates, 2003].)

Steps in the Construction of a Dynamic System for L2

A State Space Description of Developmental Trajectories. A characteristic feature of a dynamic systems approach is that it is geometric in nature. In the *Looking Glass* example, the system described Alice's changing position in a world that, basically, was described by means of two coordinates, latitude and longitude. Other dimensions can be added to the system as needed to specify any additional variable by which the system is characterized. We thus arrive at a mathematical object that consists of an abstract space of n dimensions (with n = any arbitrary number). The state of the system at any point in time is represented by the system's coordinates on all its dimensions as a point in the n-dimensional space, or *state space*. The bad thing about this is that it is a pretty abstract way of describing concrete things; the good thing is that we do not need to know or measure each of the dimensions to be able to use this geometric concept of the system. This notion of state space will allow us to describe and understand certain general properties of the system.

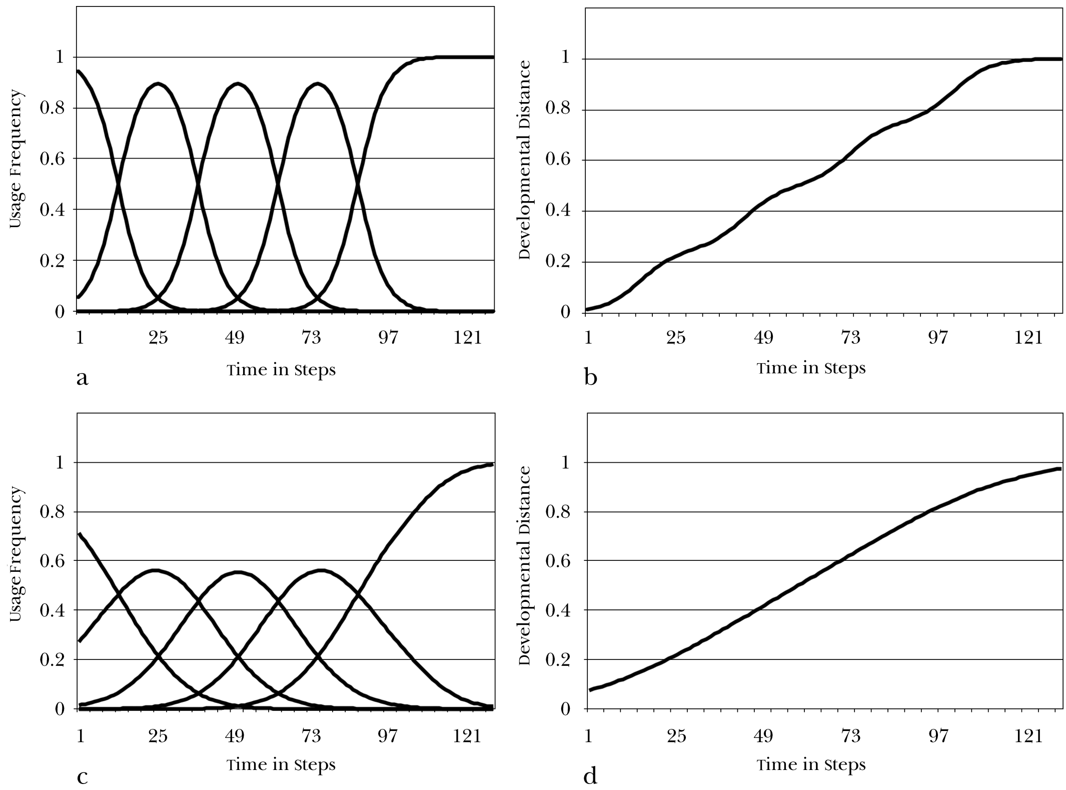
Imagine a state space containing all the dimensions (variables, properties, etc.) a researcher needs to more or less exhaustively specify the state of any conceivable L1 and L2 user and/or learner, and let us, for simplicity, confine ourselves to L1 speakers who wish to learn an L2. Somewhere in this state space there is a region where we can find the L1 speakers who have not yet begun to learn the L2 in question, and somewhere we can locate a region where we will find those L1 speakers who, according to whatever criteria we wish to use, have mastered the L2. In order to study complex dynamics—of which L2 learning is most likely an example—it is strategically wise to conceptually reduce the complex system to a single dimension or very simple state space, in which the qualitative properties of the system dynamics can be observed. In the given L1–L2 state space, we easily can imagine a single dimension that describes the distance between the "starting point" region and the "end point" region. This simplification does not entail an ontological claim, in the sense that it does not imply that "in reality" L2 learning is a one-dimensional process, which clearly would be absurd. Instead, the simplification is a mathematical projection (in the sense of projective geometry), projecting a complex pattern onto a highly reduced state space, which, if done properly, opens the complex pattern for study and description.

Empirically, this one-dimensional distance can be viewed as "L2 proficiency," which, for now, remains a somewhat vague term. The important thing is that the steps an individual takes on his or her way to L2 acquisition can be rank-ordered on this dimension (ties not excluded). What a dynamic systems–oriented researcher would wish to know, in this case, is the following: Let $\Delta d/\Delta t$ be the distance toward L2 covered over some time interval Δt (a day, a week, a month, etc.). How does $\Delta d/\Delta t$ relate to time, that is, the time passing from some arbitrarily determined starting point of the person's L2 learning to some arbitrarily determined end point of L2 learning? On a short-term timescale (e.g., days), for instance, $\Delta d/\Delta t$ is likely to map out a nonsmooth, "jerky" function, which represents short-term fluctuation and is characteristic of the short-term dynamics of L2 use and learning. On a longer term timescale (e.g., months), $\Delta d/\Delta t$ is likely to follow a path of accelerations and decelerations, which is characteristic of a stage-wise pattern of acquisition.

I shall discuss an example based on a free (and idealizing) interpretation of a model advanced by Pienemann (1998). The model describes the L2 acquisition of German word order in the form of five stages that represent a particular grammatical structure. Let t_0 be the time at which a person starts to learn German, and t_f a later point in time at which that person uses correct word order. The stages imply that the frequencies of use of each of the word order types change. In the beginning, a person will almost uniquely use the stage 0 structure, and as time passes, the proportional frequency of more advanced structures will increase. Finally, the proportion of correct word order structures will approach 1 (i.e., 100%). It is likely that the proportions of the structure types wax and wane, until only the correct structures remain. I have simplified and idealized this situation by assuming that the structures form overlapping waves of usage frequencies (see Siegler, 2005, for a general overlapping waves model), as represented in Figures 1a–1d.

The state space describing the dynamics of German word order acquisition thus requires five dimensions, each representing the frequency of use of a particular word order structure. The developmental distance vector runs from 100% use of the most primitive structure to 100% use of the most complex structure. If, for simplicity, we put the stages on equal distances from each other in terms of increasing complexity, we can weigh the stages quantitatively by 0, 1, 2, 3, and 4, respectively, 4 indicating the most complex or correct structure. At any point between t_0 and t_f, the learner's position on the distance dimension is the weighted sum of $(0*p_0+1*p_1+2*p_2+3*p_3+4*p_4)/4$, for p_0 to p_4 equals the proportions of usage of each of the structures at a particular point in time. When we apply this equation to the overlapping waves from Figure 1a, we obtain a stepwise growth function for the one-dimensional state space of word order acquisition (Figure 1b). Note that if the frequencies overlap differently, as in Figure 1c, the resulting one-dimensional curve

FIGURES 1a–1d
Imaginary Successions of Frequency of Use of Five Developmentally Ordered Types of Word Order Structures (Left) and the Corresponding One-Dimensional Developmental Distance Representation (Right)



approaches a continuous S-shaped curve (see Figure 1d).

What this example illustrates is that it is possible to transform a qualitative model of structural stages into a geometric model of quantitative change over time. As with any model, the quantitative model is as good as the assumptions on which it has been based, and in the current example these assumptions are highly simplified and idealized. The assumptions have nevertheless helped us link qualitative and quantitative representations of the same phenomenon.

The choice of the state space dimension(s) reflects the organization of the system under observation. A complex system (i.e., the sort of system we are interested in) shows a high degree of organization—a high amount of coordination of its many components. In the case of complex, developing, and psychological systems, such as an L2 learner, this organization or coordination is expressed in the form of a few dominant, or salient, variables. An example of such a salient variable is the word order from the preceding illustration; another example is the one-dimensional notion of proficiency that characterizes a particular person's mastery of an L2. These salient variables are called the *order parameters* of the phenomenon in question, for instance, L2 learning of German or English.

Distinguishing a one-dimensional, general variable such as L2 proficiency or mastery is primarily a matter of descriptive choice, but this choice must be descriptively adequate; in other words, it must reflect something that is real or functional. A complex system such as the L2 learning system requires various perspectives, various order parameters and state space descriptions, which, together, capture the L2 learning phenomenon in all its richness.

Order parameters can have a categorical format, such as the word order structures distinguished by Pienemann (1998), or they can have a continuous, dimensional format, such as the proficiency or mastery level parameter. Continuous dimensions come close to the notion of a *metric*, or ruler, along which the states of the system can be measured in the classical sense of the word (see, e.g., Fischer & Bidell, 2006, and Fischer & Rose, 1999, for an application in developmental psychology in general). In linguistics and psychology, continuous order parameters often are represented in the form of tests providing a continuous or quotient score (an intelligence test is a prime example). However, the test and the state space dimension are not the same thing: They refer to each other, but it is not necessary to have a test (e.g., of L2 proficiency) in order to fruitfully apply a particular state space dimension or order parameter to a system (e.g., L2 proficiency as a convenient descriptive construct). In some cases, the dimension or order parameter can be empirically mapped onto a series of variables for which several tests are available, without implying that the dynamics must be described by means of a state space with as many dimensions as there are tests (or as many dimensions as there are factors in a factor analysis or principal component analysis).

The Short-Term Dynamics of Learning and the Role of Intentionality. In order to obtain a good understanding of the long-term dynamics of developmental and learning processes, it is important to specify the short-term dynamics of action. In L2 learning, actions involve communicating with other people, such as native speakers, listening to conversations, studying grammar in the textbook, making assignments, and so forth. These actions contribute to learning and development. However, learning and development determine the nature of the actions carried out (compare listening to a conversation by a novice and listening to that same conversation by a more advanced L2 learner, in terms of the information they can retrieve, or what they actually learn, from that conversation).

In a recent series of papers (Steenbeek & van Geert, 2005, 2007, 2008; van Geert & Steenbeek, 2005), a dynamic model of action in the context of development has been presented. Its major components are the person's concerns (goals, intentions, etc.), the actions available to realize those concerns, the person's subjective evaluation of the realization of those concerns through emotions and emotional expressions, and, finally, a direct social component that accounts for the "contagiousness" of the actions of others and their impact on the actions of the person through imitation. A person's goal relates to—but is not the same as—an attractor in the state space as described earlier. For instance, a person wishes to learn a foreign language and wants to pass an exam with excellent marks. These goals co-determine the person's actions, but the actions also are co-determined by a host of other factors. An important additional factor is the person's preferred balance between the goal at issue (e.g., pass the language exam) and other goals (e.g., the person's wish to go out or relax, or the person's job and family interests). The attractor state—the person's having learned the foreign language, the marks obtained on the exam—in fact self-organizes as a result of the dynamic relationships among all the factors involved. (For an application in the context of children's play, see the publications mentioned earlier in this paragraph.) An important factor of the short-term dynamics of actions related to learning and development is the real-time emotional evaluation of a person's actions in light of his goals or concerns (Posner, Russell, & Peterson, 2005). Positive emotions (both self-conscious and other-related) are of primary concern (see, e.g., Dewaele, 2005, and Immordino-Yang & Damasio, 2007, in the context of L2 learning).

Tschacher and Haken (2007) recently have presented an interesting general explanation of intentional action, which can help explain the ordered and self-organizational nature of actions and the resulting developmental and learning trajectories. First, developmental state spaces are characterized by what Tschacher and Haken call a *gradient*. On the long-term timescale, the gradient is the difference between lower and higher levels of development (learning, mastery, etc.). On the short-term timescale, the gradient represents the distance between the intended state, or goal state, on the one hand and the current realization of that state on the other hand. Second, in complex state spaces consisting of many descriptive dimensions, such as the L2 developmental state space or the state space of L2 learning activities, a system (e.g., an L2 learning individual) theoretically can take any route leading from a lesser to a more developed state. That is, theoretically, any trajectory in the state space has the same probability of being taken by the acting system. In reality, however, we see that the number of trajectories is limited (there are likely to be more than one, but in general not more than a few; Shore, 1995). The limitation of the trajectories to only a few is an example of self-organization. Tschacher and Haken (2007) claim that complex systems must resolve gradients by means of optimally dynamically efficient routes, which are in fact like minimal-energy solutions to crossing the distance from, in this particular case, a lower to a higher level of L2 proficiency. Tschacher and Haken's model thus attempts to explain why the number of possible developmental trajectories is limited, and why the number of possible action trajectories leading from a goal to its perceived, satisfactory realization also is limited. These trajectories are the result of self-organizing processes, which include but are not limited to explicit intentions and action plans. In this sense, both action and long-term learning and development are examples of "organized self-organization" (van Geert, 1989).

Static Versus Dynamic Accounts of L2. In psychology, or any other science that deals with human behavior, including language, it is customary to indirectly approach the dynamics of a phenomenon by finding out how the distribution of a particular phenomenon in a group (population or sample) is related to distributions of other phenomena in the group. For instance, one might wish to know the association between a student's oral performance on the one hand and the student's task goal orientation on the other hand. One can approach this question by taking a large sample of students, measuring their oral performance and goal orientation, and calculating the correlation or some other indicator of association between the two measures, which, as we will find out, takes a positive value. This association indicator is a so-called *static* measure (there is absolutely no pejorative or evaluative meaning associated to this term), meaning that it reflects the position of students in a group relative to each other. (For a recent example relating to adaptive language learning, see Woodrow, 2006. Note that I am using this relationship as an example about which I will speculate freely in order to try to illuminate the difference between a static and a dynamic model. I am in no way claiming to discuss or criticize Woodrow's findings here.) We can assume that the positive correlation means that we will tend to find higher or better task orientation with students who have better oral performance and, by definition, also the other way around. Technically, we can specify this relationship as follows:

$$y_i = f(x_i)$$

with y representing oral performance and x task orientation (or vice versa). The f stands for *function* and means a positive association here, which allows us to predict a value for x, given a value of y, within certain limits (the error term, or percentage of explained variance).

However—and this is a very important point—the fact that an association holds over a population or sample does not necessarily imply that such an association holds for a time evolution or trajectory, that is, a dynamic system. Thus, if we follow the time evolution of oral performance and the time evolution of task orientation in a particular subject—which is where the evolution takes place—we do not necessarily find that higher task orientation is related to higher performance.

This possible absence of positive relationship has nothing to do with eventual random fluctuation or error. What I mean to say is that it is very well thinkable that a static, positive relationship over a population (including all kinds of error fluctuation) might correspond with a negative, dynamic relation (or anything else) in any or eventually all of the individual members of the population. Think, for instance—and this is purely speculative but not impossible in my view—that a person's goal orientation, including the strength of his intention to perform well and invest more effort in accomplishing the task, increases as he or she perceives the task as more difficult and demanding, as in reading and understanding texts of high lexical and syntactic complexity, for example. It is not unlikely that such demanding tasks require the student to invest more effort in aspects other than oral performance, and thus that his or her oral performance in such tasks is on average *lower* than in tasks that the student considers easy and not requiring a stronger task orientation than he or she is used to. Thus, seen in a dynamical perspective against the student's own standards of performance and perception of competence, the relationship between task orientation and oral performance might be a negative one, whereas that same relationship is positive if the topic of analysis is the sample of students. The crucial thing here is that a dynamic relationship represents a process mechanism, whereas a static relationship over a sample represents an association between the distributions of two variables over the sample.

The finding that a static and a dynamic relationship are not the same thing is an example of the so-called *ergodicity problem* (Hamaker, Dolan, & Molenaar, 2005; Molenaar, 2004; Molenaar, Huizenga, & Nesselroade, 2003). The problem refers to the possible similarity of a space average (e.g., an average or association) calculated for an ensemble of participants, and a time average (i.e., an average calculated over a sequence of steps in an individual). The speculative example of the oral performance and task orientation relationship is an illustration of a situation in which this similarity does not necessarily apply. Although this example is speculative, there exists research in a different field that shows this inequality empirically. A study by Musher-Eizenman, Nesselroade, and Schmitz (2002) showed that the relation between perceived control and academic performance found in a cross-sectional or classical longitudinal design (few repeated measurements over a comparatively long time) was different from the relation found when relatively short-term within-person change patterns were studied. Thus, the short-term dynamics of perceived control and academic performance in individuals is different from the long-term or cross-sectional relationship. If this reasoning is applied to the oral performance versus task orientation example, it would hold that whereas in the short run the relationship between the two is negative, in the long run the relationship becomes positive again, in that a long-term increase in goal orientation in an individual can be associated with a long-term increase in that person's oral performance. (For an example of a distinction between relationships holding over a sample and holding between time points in a dynamics of language learning and instruction, see van Geert & Steenbeek, 2005.)

In order to avoid possible confusion arising out of the ergodicity discussion, I must add that I do not mean to say that development is some sort of internal, solipsist thing. Dynamic rules or relationships pertain to particular dynamics, and in the case of L2 learning, this is the dynamics of an individual learning the language in a continuous interaction with other persons who use the L1 and L2 or teach the L2. There is no doubt that L2 learning is a highly social process. In addition, the ergodicity discussion does not imply that individual cases, or even exceptions, are more important than what is common among individuals or groups. In a dynamic systems approach, a sample is the collection of dynamic trajectories, some of which are more common than others. A valid dynamic model must thus be able to explain the common pattern as well as the uncommon patterns and exceptions.

Finding the Evolution Term and Explaining the Long-Term Dynamics. In order to try to obtain a better understanding of this ergodicity problem and of the principles of dynamics in general, consider the simplest possible format of a dynamic system, which is:

$$y_{t+1} = f(y_t) \text{ and } x_{t+1} = g(x_t)$$

The equation says that the next state (at time $t+1$) of a variable y or x is a function (f or g) of the variables' preceding states (at time t). If this is so, then any next state is a function of its preceding state, and the equation produces a picture of the *time evolution* of y and x:

$$y_t \rightarrow y_{t+1} \rightarrow y_{t+2} \rightarrow y_{t+3} \rightarrow y_{t+4} \rightarrow \ldots \text{ and}$$
$$x_t \rightarrow x_{t+1} \rightarrow x_{t+2} \rightarrow x_{t+3} \rightarrow x_{t+4} \rightarrow \ldots$$

This is a simple application of the iterativeness (or recursiveness) explained earlier. Thus, the first question to answer is: What is the content of, or the mechanism behind, the enigmatic f, assuming that y represents L2 proficiency, the variable quality or level of L2 performance of a person? The simplest possible answer to this question is that, if L2 performance is part of a learning–teaching process, the change of L2 performance likely is to be some form of *accretion*, some form of adding some performance quality to the preceding state of L2 performance quality. An accretion model of lexical learning implies that a person learns a fixed number of words a day (plus or minus some random variation due to purely external causes). From a learning-theoretical[2] point of view, the accretion model is very unlikely because it views the learning, at any point in time, as completely independent of what the person already knows and of what the person does not yet know.

Studies of learning and development converge on the finding that the increase of performance, skill, or knowledge over time is proportional to the performance level already attained, proportional to the difference between what is already learned and what has yet to be learned, and proportional to the available resources (e.g., the quality of the teaching, the student's motivation, etc.). This proportionality leads to various models of long-term change, the most likely or applicable of which is the so-called *logistic growth model*, the model of growth under limited resources. This dynamic growth model has been thoroughly explained elsewhere (de Bot, Lowie, & Verspoor, 2007; van Geert, 1991, 1994, 2003). Under ideal conditions, it leads to a smooth, S-shaped pattern of increase of performance, which stabilizes at a level it is able to maintain given the available resources. For an L2, this might include such resources as the frequency with which high-quality (native) L2 is heard by the L2 learner, the frequency of use, and the speaker's linguistic talent (whatever that may mean). However, as stated before, the dynamic growth model must reckon with the properties of the short-term dynamics of actual L2 production and L2 reception, which require explanation at a different timescale. As stated before, a prime example of such a property is the natural fluctuation of performance over time.

In the preceding section, I discussed the example of oral performance and task orientation, which demonstrated the importance of a particular coupling of variables. Thus, task orientation is likely to affect oral performance over time, and we have seen that this relationship need not be the same as the relationship found over a sample of participants. The relationship between variables relates to a particular property of dynamic systems, namely, the possibility to dynamically couple the variables in various ways. The formal format of a coupled dynamic system is as follows:

$$y_{t+1} = f(y_t, x_t) \text{ and } x_{t+1} = g(x_t, y_t)$$

meaning that the next step of y depends on the preceding value of itself and of another variable, namely, x, whereas the next step of x depends on its preceding value and the preceding value of y. The variables y and x can refer to processes within a person and also to processes between persons. For instance, y can represent a level of proficiency in a particular L2 learner, such as the type of word order frequencies in the learner's L2 production. x, however, can represent corrections or demonstrations made by an L2 speaker or L2 teacher, following the L2 learner's apparent mistakes. If the learner is sensitive to these corrections and assimilates them, his or her L2 level will change and induce other types of corrections, if any are needed, in the L2 teacher.

In the simplest possible sort of model, the relationship between one variable and another can be positive, as is the case if increased task orientation results in better oral performance. Or this relationship can be negative, as is the case if an increase in task orientation is due to perceived task difficulty eventually resulting in lower performance than if the task is easy or less demanding. These relationships are unidirectional; that is, they go from one variable to another but not the other way around.

More interesting dynamics result from mutual relationships between variables. Take, for instance, the following speculative example. It is likely that in the early stages of L2 learning, L2 use requires considerable effort, in that it is not yet a more or less automatic task and requires monitoring, explicit attention, and so forth. More effort will, in principle, result in better performance (all other things, such as prior L1 and L2 knowledge, support given, etc., being equal). Good performance, if it is recognized as such by a more competent L2 conversation partner, will in principle help to maintain the speaker's motivation and thus the investment of effort. However, high levels of effort lead to fatigue, and fatigue is likely to affect performance negatively. Hence, the relationship between performance and fatigue, via effort, is mutual but asymmetrical. It is positive from high performance (via high effort) to fatigue in that high effort makes fatigue increase. It is negative from fatigue to performance, in that higher fatigue results in lower performance. This type of asymmetrical relationship, which can be compared to a predator–prey relationship, is likely to produce oscillating patterns, waves of high performance and high fatigue. (This is, of course, an idealized example). However, the effort component depends clearly on the level of mastery of the L2. The more L2 use becomes an automatic task performance, the less it will appeal to the effort component, and the less it will be a cause of fatigue. But fatigue from a different source (e.g., jet lag at the beginning of a conference) will still affect L2 performance, even in an experienced L2 speaker. Thus, the nature of the dynamic relationships changes over time and is also context and origin dependent.

To return to the ergodicity problem and the issue of static versus dynamic models, assume that a researcher has a sample of participants at different levels of performance and at different levels of effort required to achieve that performance. The relationship between performance and fatigue over this sample will likely be anomalous; that is, the correlation may be small and not statistically significant. However, there is no direct link between this null relationship over the sample and the fatigue–performance relationships that govern the dynamics of L2 acquisition and performance in all the independent individuals composing the sample. In individuals, the link is still present and influencing the individual dynamics to a considerable extent. However, the nature of the link itself is subject to developmental change and thus not uniform over participants.

Short-Term Dynamics and Fluctuation. The short-term dynamics of language refers to the mechanisms operating on the actual production and understanding of language, for instance, in conversations or oral or written monologues, such as the homework assignments of an L2 learner. Fluctuation, or intraperson variability, is a characteristic feature of the short-term dynamics of language use (de Bot, Lowie, & Verspoor, 2007; Ellis, 1985, 1994; Pienemann, 2007). Fluctuations originate, for instance, from the availability of alternative forms or strategies (e.g., various patterns of word order) and from contextual variations (e.g., monologues versus conversations with native or nonnative L2 speakers). Fluctuations thus partly originate from the long-term dynamics of L2 learning, which makes certain forms, strategies, or rules available or not.

Short-term dynamics also are likely to contribute to the long-term dynamics of the L2 learning process. Variability in linguistic productions of L2 learners among different action contexts provides possibilities for spontaneous explorations of possible forms or rules of the L2 that is being learned, for instance, because this variability in production allows for corrections or rephrasing by the L2 teacher or native speaker. Since the fluctuation is a product of the underlying short-term dynamics (e.g., the dynamics of sentence formation or the dynamics of a conversation), the fluctuation must not be treated as noise, that is, as variability due to some independent, external source of disturbance. Rather, the fluctuation, if properly described and understood, may contain information about the underlying short-term dynamics and its effect on the long-term dynamics of increasing L2 performance. An example from L1 learning will be discussed in the next section (Bassano & van Geert, 2007).

Another example, also in early L1 acquisition, comes from a study with Marijn van Dijk on variability peaks in the use of spatial prepositions, marking the transition to rule-governed use of those preposition types in the child's language (van Dijk & van Geert, 2007). Other examples of this phenomenon are presented in the article by Verspoor, Lowie, and van Dijk (this issue). In short, the theoretical specification of the evolution term explaining the time evolution of L2 performance on the long term will have to reckon with the consequences of the short-term dynamics, such as variability. These short-term dynamics are the subject of study of disciplines such as cognitive psychology (Costa, La Heij, & Navarrete, 2006) or neurocognition, or dynamic systems as defined by Thelen and Smith (1994; see also Colunga & Smith, 2005, and Jones & Smith, 2005, for an application of this theoretical approach to word learning and use).

Transitions in Early Language Development: An Example of Dynamic Systems Model Building

A Hypothetical Model of Three Stages in Early Language Acquisition. In a study I carried out in collaboration with Dominique Bassano, an attempt was made to build a dynamic model of long-term changes in language production, focusing on sentence length as an indicator of underlying principles of L1 production (for details see Bassano & van Geert, 2007). The model that we wished to check goes back to a hypothesis about L1 production that has been around for almost 30 years now. It assumes that, in the construction of a genuine syntactic language, children begin with a stage in which one word expresses a complex referential meaning ("holophrase"). The hypothesized generating mechanism of language at this stage is therefore called the "holophrastic generator." In a second stage, the child is assumed to simply combine these complex holophrases into utterances consisting of a few (two or three) such units, and thus generate language by means of combination, hence the stage of the "combinatorial generator." It is assumed that, while combining words, children become sensitive to the way the language spoken by the linguistic environment solves the combinatorial problem, which involves the use of syntax (order rules, word correspondence, etc.). Thus, the third and final stage, according to this hypothesis, is the stage at which language is supposed to be based on a "syntactic generator." These stages are assumed to occur in the form of overlapping waves. The alternative hypothesis is that language development is a continuous process (whatever the exact nature of the underlying continuous processes).

Data and Method. The main set of data used in this study came from the longitudinal corpus of one French girl, Pauline, who was studied from ages 1;2 to 3;0. Additional data are from another French child, Benjamin, who was studied from age 2;0 to 3;0. (For analyses of the children's language development, see Bassano, 1996.) Data were obtained using a free speech sampling method. For each child, frequencies of one-word, two-word, three-word utterances, and so on (W1, W2, W3, etc.) were calculated for each monthly sample (120 utterances) and for the subsamples (60 utterances and 30 utterances, respectively). Figure 2 shows the smoothed curves of the raw data, based on a Loess smoothing procedure (locally weighted, least-squares smoothing), which provides a representation of the changes in the sentence frequencies and is able to capture eventual temporal regressions and accelerations in the growth rate.

A Dynamic Model of Stepwise Grammatical Development. The aim of this model is to explain the quantitative evolution of three types of utterances or phrases, namely, one-word, two- and three-word, and four-word-and-more phrases, called W1, W23, and W4+, respectively. These types of phrases are supposed to be linked with the linguistic generators described earlier. Before proceeding with the model, two remarks should be made. The first is that the relationship between a particular type of phrase, for instance, W1, and its assumed underlying generator, is in itself a transient relationship in that it changes over developmental time. Thus, an early W1 is considered a product of the holophrastic generator, but a late W1, occurring in the midst of W23 and W4+ phrases, is most likely an expression of the assumed syntactic generator. This changing relationship must be taken into account when interpreting the meaning of the linguistic data. Second, the model we wish to build does not explain the emergence of a new generator out of an old one. In order to do so, a different type of model would be required, showing how a procedure of word combination, for instance, grows out of a procedure of holophrase use, based on various kinds of linguistic inputs, due to increasing participation in communicative events with other, and in principle, (more) mature speakers. Moreover, the model—which is a model of long-term change—does not include the short-term dynamics of phrase production; for instance, it does not provide an explanation of how and why children, in their actual speech producing processes, make a choice between a holophrastic and combinatorial generator at developmental stages when both are available.

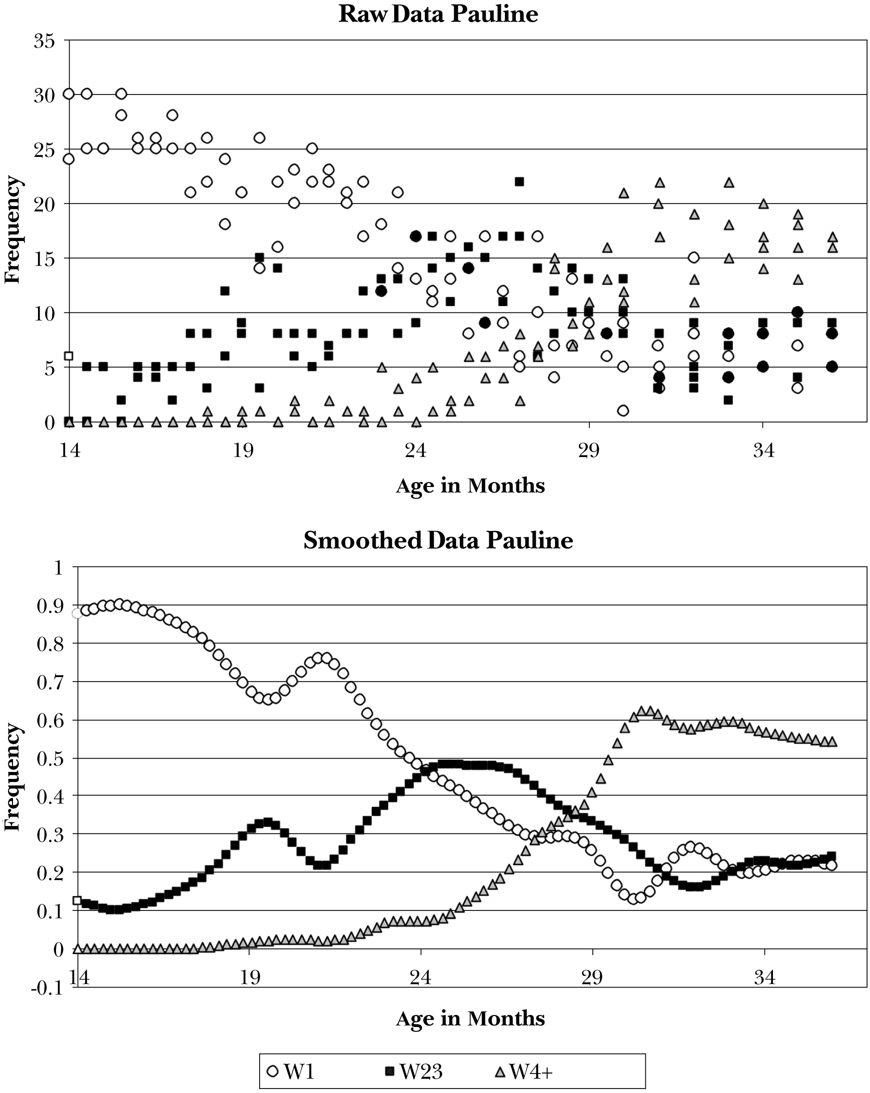
The term *explain* might give rise to confusion. Of course, the dynamic model might be said to "explain" this choice by invoking the notion of a random selection between two available generators, given the expected frequencies at some particular moment in time. However, this model-theoretic random selection is not intended as a literal description of the actual short-term process of phrase production and does not in itself imply that a language-producing child literally makes a stochastic response choice. In fact, for reasons of simplicity, the long-term model treats the short-term process as if it were a stochastic choice as convenient for the time being and leaves the description of how the process actually occurs in the short term of phrase production to a theory and design that studies this short-term timescale directly.

In line with comparable growth models of long-term change, we postulated that the hypothesized generators—which are hierarchically ordered in terms of assumed developmental complexity and order—were characterized by bidirectional asymmetric relationships (see Figure 3).

Relationships run from the less to the more advanced generator and from the more to the less advanced.

First, the relationships from the less advanced to the more advanced generator are *supportive* and *conditional*. For instance, the holophrastic generator explains the increase in the production of one-word phrases. It is assumed that a minimum level of one-word phrase production is necessary to make the transition to a combinatorial grammar, which generates two- and three-word phrases. This minimum required level is needed for the combinatorial generator to emerge, hence the notion of a conditional relationship from holophrastic to combinatorial generator.[3] The supportive relationship between holophrastic and

FIGURE 2
Raw Data (Top) and Smoothed Curves (Bottom) of W1-, W23-, and W4+ Sentences of Pauline Between 14 and 36 Months of Age

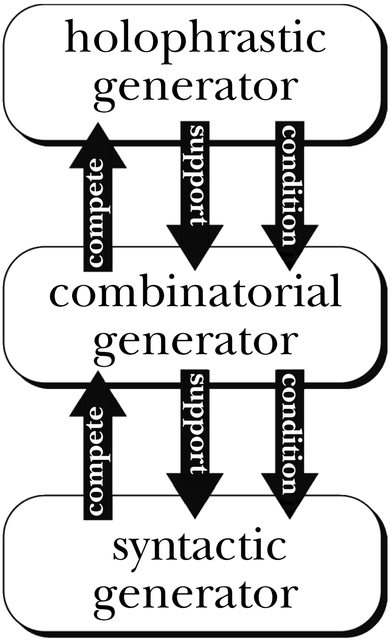


Note. W1 = one-word utterance; W23 = two- and three-word utterance; W4+ = four-word-and-more phrase.

combinatorial generator implies that there is a linear, positive relationship between the number of W1 phrases actually produced and the increase in the number of W23 phrases. That is, "simpler" formats, such as W1 phrases produced by a holophrastic generator, are assumed to stimulate the production of more complex formats, such as W23 phrases produced by a combinatorial generator. I see this as a very simple, basic developmental assumption, which can be explained on the basis of classical developmental and learning-theoretical rules or principles (see van Geert, 1998b).

Second, the relationships from the more advanced to the less advanced generators are *competitive*. Such a relationship implies that an increase in the use of the more advanced generator is related to a decrease in the use of the less advanced generator, which will contribute to the latter's decline (see MacWhinney, 1998 and van Geert, 1991 for a more thorough discussion of the origins of the competitive relationship).

FIGURE 3
Competitive and Supportive Relationships Among Hypothetical Syntactic Generators in L1 Acquisition



The competitive relationship is, in all likelihood, a transitive relationship. That is, it holds for all generators that are less complex or less mature. Thus, a competitive relationship from the syntactic to the combinatorial generator also implies a competitive relationship from the syntactic to the still older holophrastic generator.

The next step is to specify a mathematical model that provides a formal representation of the set of relationships. For instance, the supportive relationship entails a positive effect of the productivity of a more mature generator on a less mature one. Thus, if C_t represents the frequency of W23 at time t, and S_t represents the frequency of W4+ phrases at time t, the increase in S_t, which we can write as $\Delta S/\Delta t$, is proportional to C_t. Further, if the parameter d stands for that proportion, $\Delta S/\Delta t$ depends on $d * C_t$. In one of the preceding sections, we have seen that the effect of a growth factor or learning factor on a variable is likely to be proportional to the current level of that variable. That is, the magnitude of the effect of C on S is proportional to the current level of S. If e is a parameter that stands for that particular proportion, we can now combine the two effects in a simple equation that specifies the effect of C of S, as follows:

$$\Delta S/\Delta t = d * C_t * e * S_t$$

If, for simplicity, we set $d * e$ equal to a value c, we find that:

$$\Delta S/\Delta t = c * C_t * S_t$$

which is the expression for a supportive relationship. If c takes a negative value, the equation expresses a competitive relationship.

It can be shown that growth and learning processes are processes of limited increase (see van Geert, 1991). A trivial example of a limitation is the amount of time a person can spend producing phrases (which is certainly limited by a person's life span). A less trivial example, relating to growth of the lexicon, is that such growth is limited by the number of words available in the language. This reasoning leads to the following model, which is the coupled logistic growth model. For simplicity, it first will be applied to the increase, over time, of the number of W1 phrases uttered by a child, for instance, on an average day. Given that a child must build up a lexicon from an arbitrarily small number of first words, and thus from an arbitrarily small number of possible one-word phrases (H, for holophrases), the number of one-word sentences thus will grow over time. This growth is, as we have seen, most likely proportional to the level of H already attained. Thus, the increase of H is expressed as, $\Delta H/\Delta t = b * H_t$, with b as a proportion that specifies the rate of increase. However, as I have just stated, that increase is limited. The simplest way of specifying this limitation is that b, the rate of increase, will decrease as H gets bigger and thus approaches its limitations, and this decrease is proportional to H (let us say the symbol a represents this proportion). This relationship can be expressed mathematically in the following way:

$$\Delta H/\Delta t = (b - a * H_t) * H_t$$

We also have stated that H "suffers" from the increase of more complex phrases, based on the combinatorial generator. Thus, the increase in H over some time t can be expressed as follows:

$$\Delta H/\Delta t = (b - a * H_t) * H_t - g * C_t * H_t$$

Applying a comparable logic to the growth of W23 (the C-generator utterances) and W4+ (the S-generator utterances), we arrive at the following coupled dynamic system:

$$\Delta H/\Delta t = (b - a * H_t) * H_t - g * C_t * H_t - h * S_t * H_t$$

$$\Delta C/\Delta t = (i - j * C_t) * C_t + k * H_t * C_t - l * S_t * C_t$$

$$\Delta S/\Delta t = (m - n * S_t) * S_t + s * C_t * S_t$$

which describes the time evolution of W1, W23, and W4+ phrases. (Remember that the connection between a type of phrase and a generator, for example, W1 and the H-generator, changes over time, which makes the connection between the model and the empirical data more complicated.) This is a dynamic system described by a three-dimensional state space (the H, C, and S dimensions). For simplicity, it can be represented by three separate one-dimensional time evolutions (the time evolution of H, C, and S separately).

Mathematically, this model is equivalent to the following notation (for reasons of simplicity and without loss of generality, we can confine ourselves to the *H* variable):

$$H_{t+1} = H_t + b * (1 - H_t/K_H) - g * C_t * H_t - h * S_t * H_t$$

with K_H the limit level of H (the highest possible level H can attain under the present limiting circumstances). This K parameter has an interesting property: Since the limit level, specified by K, is a function of all the available resources, whether they be internal or external, K amounts to a single-value representation of all the available resources. The fact that it is at all possible to represent the very complex structure of resources (memory, effort, help given, motivation, talent, available books, etc.) by a single value is a direct consequence of the fact that the resources are represented in function of the dynamic effect they have on the growth and stabilization of the variable at issue (H, in this case).

Finally, the coupled-dynamics model specifies only the supportive and competitive relationships, not the conditional relationships, which have been left out from the equations to avoid further complication. They can be added to the equations in the following way (the example is limited to the C-component):

$$\Delta C/\Delta t = 0 \text{ if } H_t < H_P$$

$$\Delta C/\Delta t = (i - j * C_t) * C_t + k * H_t * C_t - l * S_t * C_t \text{ if } H_t \geq H_P$$

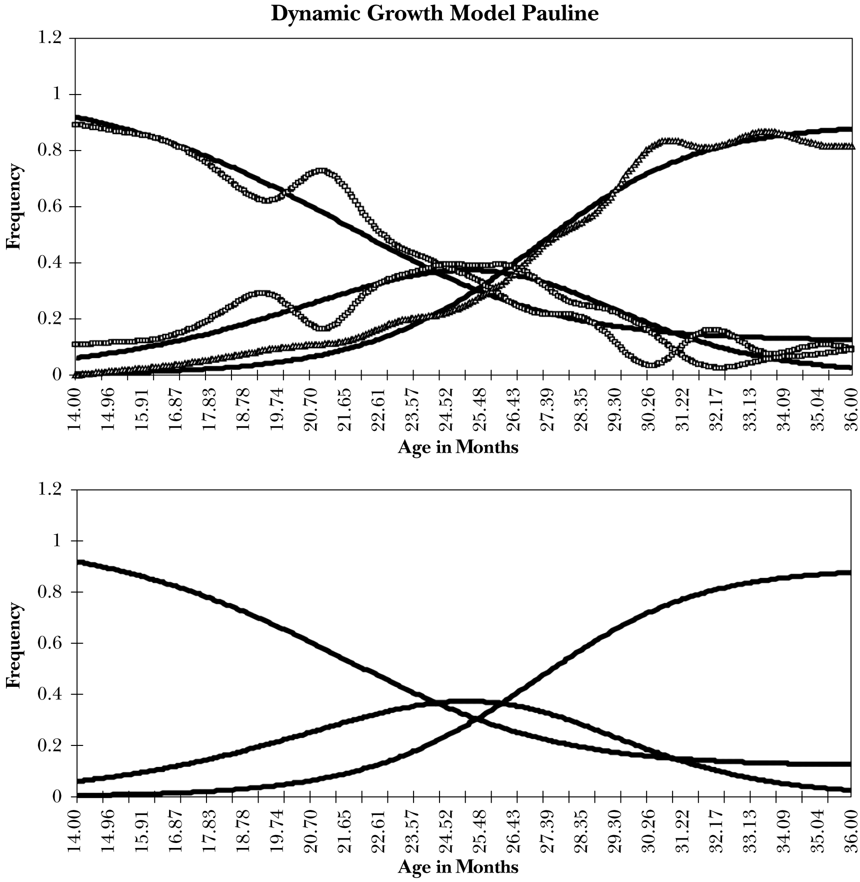
with H_P the conditional value H must have for C to grow (e.g., if H_P is set to 100, it is assumed that a child must have at least 100 different words before he can start making two-word combinations in a productive way, and thus for the combinatorial grammar to increase in terms of productivity).

The next question is, of course, what the value of these *a*, *b* parameters should be. They can be estimated by mere guessing and then by trying out the model with these guessed values to see whether the model based on these parameter values fits the data. However, they also can be estimated by means of parameter optimization programs that calculate the best possible set of parameters, the set for which the model produces the best possible fit with the data. Since the W23 and W1 phrases occurring later in developmental time are "absorbed" by the emerging syntactic generator (meaning that a late W1 is produced by the S-generator, and no longer by the original H-generator, for instance), the data must be rescaled in order to account for this absorption process. This rescaling was done by normalizing the data to values between 0 and 1 and by subtracting a baseline value. Figure 4 shows the result of this optimization procedure, demonstrating that the model provides a good fit with the normalized and rescaled data.

Evidence of Stage Transitions? The dynamic model predicts that the probabilities of W1, W23, and W4+ sentences increase and decrease smoothly. For instance, every day the expected frequency of W1 phrases decreases a little bit, whereas the expected frequency of the W23 gradually increases. This means that for any episode of language production long enough to arrive at reliable frequency counts—say, an episode of a couple of hours or days—the expected frequencies will change smoothly. Since the *observed* frequencies depend on the *expected* frequencies, which are in fact probabilities of production, the observed frequencies will, of course, fluctuate over days or hours. These fluctuations must be within the bounds predicted by the probabilities of production, represented by the smoothed frequencies that the dynamic model so nicely simulates. So far, all this is an expression of an orderly world. However, we began this article by visiting Alice's Wonderland, and so we will end. It is time to imagine some strange events.

Just assume that around the age of 22 months, Pauline produces only W1 phrases on even calendar days and only W23 phrases on uneven calendar days. Thus, in a truly Alicean way, it is either Holophrase or Combinatorics for Pauline (around the age of 22 months, that is; things might change afterward). In other words, the two

FIGURE 4
Smoothed Data Curves Compared With Curves Generated by the Dynamic Growth Model



generators are in harsh competition, and there is an age (i.e., 22 months) where both have the same chance to win the contest and thus to determine the sort of phrases that will be uttered during a particular day. If this were so, we would observe a maximal fluctuation over days between 100% W1 and 100% W23 for some time. This day-to-day fluctuation will thus be considerably greater than the fluctuation we expect to find if the generators wax and wane in the smooth way of the dynamic model, which so nicely fits the smoothed data. However, this condition of increased fluctuation is less *Looking Glass*–like than one might expect. In fact, increased fluctuation was exactly what we found in Pauline's and Benjamin's data. We found two periods of fluctuation where the variability was significantly greater than should be expected on statistical grounds, given the observed probabilities of the three phrase types (details can be found in Bassano & van Geert, 2007). What this finding means is that reality lies somewhere between the continuous dynamic model and the discontinuous world inspired by the *Looking Glass* magic. That is, as a new generator emerges, there is a short period during which the child feels (moderately) "torn" between two generators, and this is the kind of thing one expects to find during true developmental transitions (see, e.g., Ruhland & van Geert, 1998 for an application of catastrophe theory to language development). To put it differently, the study of fluctuations and intra-individual variability may add to our understanding of the underlying developmental processes. It also helps us understand how, in development, continuity and gradualness on the one hand and discontinuity and transition on the other hand go together (see van Dijk & van Geert, 2007). To avoid any misunderstanding, the specific dynamical phenomena discussed in the present chapter pertain to specific phenomena in

L1 development and do not necessarily generalize to L2 development. However, the principles of dynamic systems—types of relationships between components, occurrence of changing levels of variability, and so on—are assumed to apply to any developmental process (see de Bot, Lowie, & Verspoor, 2007, for an application of these principles to L2 learning).

CONCLUSION

There are various ways in which dynamic systems theory can be applied to language learning and acquisition, and there are different timescales and phenomena to be explained. Researchers can resort either to making mathematical models or to detailed analyses of fluctuations in the data. These possibilities still do not exhaust the richness of this approach. It must be emphasized that dynamic systems is not an approach that allows for simple general answers, and there is so much that has yet to be explored and tested. It is not an approach that aims to abandon the current, established way of studying language growth, learning, and change. However, an understanding of dynamic systems is crucial if we want to go beyond the static or structural relationships between properties or variables and wish to understand the mechanism of development and learning as it applies to individuals. The road toward an understanding of the dynamics of L1 and L2 acquisition and learning is steep and long and paved with difficulties, but it is a road well worth taking.

NOTES

[1] This entire overview of names and publications is highly selective in that it does not do justice to many others who have made contributions to the field.

[2] The term *learning-theoretical* refers to the viewpoint of classical or general learning theory (see, e.g., Mazur, 2005).

[3] The minimal level hypothesis is related to the critical mass hypothesis in lexical development (see, e.g., Marchman & Bates, 1994; Bates & Goodman, 1997, 2001).

REFERENCES

Bassano, D. (1996). Functional and formal constraints on the emergence of epistemic modality: A longitudinal study on French. *First Language, 16,* 77–113.

Bassano, D., & van Geert, P. (2007). Modeling continuity and discontinuity in utterance length: A quantitative approach to changes, transitions and intra-individual variability in early grammatical development. *Developmental Science, 10,* 588–612.

Bates, E., & Goodman, J. C. (1997). On the inseparability of grammar and the lexicon: Evidence from acquisition, aphasia, and real-time processing. *Language and Cognitive Processes, 12,* 507–584.

Bates, E., & Goodman, J. C. (2001). *On the inseparability of grammar and the lexicon: Evidence from acquisition.* Malden, MA: Blackwell.

Carroll, L. (1871). *Through the looking glass and what Alice found there.* London: MacMillan. (References in the text from the digital facsimile version obtainable from Google Books and published by Digital Scanning, Inc., 2007, based on Carroll, L. (2005). *Through the looking-glass.* Fairfield, IA: 1st World Library Literary Society.)

Clark, A. (1997). Being there: Putting brain, body and world together again. Cambridge, MA: MIT Press.

Clark, A. (2006). Language, embodiment, and the cognitive niche. *Trends in Cognitive Sciences, 10,* 370–374.

Clark, A., & Chalmers, D. (1998). The extended mind. *Analysis, 58,* 7–19.

Colunga, E., & Smith, L. B. (2005). From the lexicon to expectations about kinds: A role for associative learning. *Psychological Review, 112,* 347–382.

Costa, A., La Heij, W., & Navarrete, E. (2006). The dynamics of bilingual lexical access. *Bilingualism: Language and Cognition, 9,* 137–151.

Crutchfield, J. P. (1994). The calculi of emergence: Computation, dynamics and induction. *Physica D, 75,* 11–54.

de Bot, K., Lowie, W., & Verspoor, M. (2007). A dynamic systems theory approach to second language acquisition. *Bilingualism: Language and Cognition, 10,* 7–21.

Dewaele, J. M. (2005). Investigating the psychological and emotional dimensions in instructed language learning: Obstacles and possibilities. *Modern Language Journal, 89,* 367–380.

Ellis, R. (1985). Sources of variability in interlanguage. *Applied Linguistics, 6,* 118–131.

Ellis, R. (1994). *The study of second language acquisition.* Oxford: Oxford University Press.

Fischer, K. W., & Bidell, T. R. (2006). Dynamic development of action, thought, and emotion. In R. M. Lerner & W. Damon (Eds.), *Handbook of child psychology: Vol. 1. Theoretical models of human development* (6th ed., pp. 313–399). New York: Wiley.

Fischer, K. W., & Rose, S. P. (1999). Rulers, models, and nonlinear dynamics: Measurement and method in developmental research. In G. Savelsbergh, H. van der Maas, & P. van Geert (Eds.), *Nonlinear developmental processes* (pp. 197–212). Amsterdam: Royal Netherlands Academy of Arts and Sciences.

Ford, D. H., & Lerner, R. M. (1992). *Developmental systems theory: An integrative approach.* Thousand Oaks, CA: Sage Publications.

Hamaker, E. L., Dolan, C. V., & Molenaar, P. C. M. (2005). Statistical modeling of the individual:

Rationale and application of multivariate stationary time series analysis. *Multivariate Behavioral Research, 40,* 207–233.
Harris, R. (2004). Integrationism, language, mind and world. *Language Sciences, 26,* 727–739.
Immordino-Yang, M. H., & Damasio, A. (2007). We feel, therefore we learn: The relevance of affective and social neuroscience to education. *Mind, Brain, and Education, 1,* 3–10.
Jones, S. S., & Smith, L. B. (2005). Object name learning and object perception: A deficit in late talkers. *Journal of Child Language, 32,* 223–240.
Lewis, M. D. (1995). Cognition-emotion feedback and the self-organization of developmental paths. *Human Development, 38,* 71–102.
Lewis, M. D. (2000). The promise of dynamic systems approaches for an integrated account of human development. *Child Development, 71,* 36–43.
Lewis, M. D. (2005). Self-organizing individual differences in brain development. *Developmental Review, 25,* 252–277.
Lewis, M. D., Lamey, A. V., & Douglas, L. (1999). A new dynamic systems method for the analysis of early socioemotional development. *Developmental Science, 2,* 457–475.
Love, N. (2004). Cognition and the language myth. *Language Sciences, 26,* 525–544.
MacWhinney, B. (1998). Models of the emergence of language. *Annual Review of Psychology, 49,* 199–227.
Marchman, V. A., & Bates, E. (1994). Continuity in lexical and morphological development: A test of the critical mass hypothesis. *Journal of Child Language, 21,* 339–366.
Mazur, J. E. (2005). *Learning and behavior* (6th ed.). Englewood Cliffs, NJ: Prentice Hall.
Molenaar, P. C. M. (2004). A manifesto on psychology as idiographic science: Bringing the person back into scientific psychology, this time forever. *Measurement: Interdisciplinary Research & Perspectives, 2,* 201–218.
Molenaar, P. C. M., Huizenga, H. M., & Nesselroade, J. R. (2003). *The relationship between the structure of interindividual and intraindividual variability: A theoretical and empirical vindication of developmental systems theory*. Dordrecht, The Netherlands: Kluwer Academic.
Musher-Eizenman, D. R., Nesselroade, J. R., & Schmitz, B. (2002). Perceived control and academic performance: A comparison of high- and low-performing children on within-person change-patterns. *International Journal of Behavioral Development, 26,* 540–547.
Pessa, E. (2004). Quantum connectionism and the emergence of cognition. In G. G. Globus, K. H. Pribram, & G. Vitiello (Eds.), *Brain and being: At the boundary between science, philosophy, language and arts* (pp. 127–145). Amsterdam: John Benjamins.
Pienemann, M. (1998). Developmental dynamics in L1 and L2 acquisition: Processability Theory and generative entrenchment. *Bilingualism: Language and Cognition, 1,* 1–20.
Pienemann, M. (2007). Variation and dynamic systems in SLA. *Bilingualism: Language and Cognition, 10,* 43–45.
Posner, J., Russell, J. A., & Peterson, B. S. (2005). The circumplex model of affect: An integrative approach to affective neuroscience, cognitive development, and psychopathology. *Development and Psychopathology, 17,* 715–734.
Ruhland, R., & van Geert, P. (1998). Jumping into syntax: Transitions in the development of closed class words. *British Journal of Developmental Psychology, 16,* 65–95.
Shore, C. M. (1995). *Individual differences in language development*. Thousand Oaks, CA: Sage Publications.
Siegler, R. S. (2005). Children's learning. *American Psychologist, 60,* 769–778.
Smith, L. B. (2005). Cognition as a dynamic system: Principles from embodiment. *Developmental Review, 25,* 278–298.
Smith, L. B., Thelen, E., Titzer, R., & McLin, D. (1999). Knowing in the context of acting: The task dynamics of the A-not-B error. *Psychological Review, 106,* 235–260.
Steenbeek, H., & van Geert, P. (2005). A dynamic systems model of dyadic interaction during play of two children. *European Journal of Developmental Psychology, 2,* 105–145.
Steenbeek, H., & van Geert, P. (2007). A theory and dynamic model of dyadic interaction: Concerns, appraisals, and contagiousness in a developmental context. *Developmental Review, 27,* 1–40.
Steenbeek, H., & van Geert, P. (2008). The empirical validation of a dynamic systems model of interaction: Do children of different sociometric statuses differ in their dyadic play interactions? *Developmental Science, 11.*
Thelen, E., & Bates, E. (2003). Connectionism and dynamic systems: Are they really different? *Developmental Science, 6,* 378–391.
Thelen, E., & Smith, L. B. (1994). *A dynamic systems approach to the development of cognition and action.* Cambridge, MA: Bradford Books/MIT Press.
Tschacher, W., & Haken, H. (2007). Intentionality in non-equilibrium systems? The functional aspects of self-organized pattern formation. *New Ideas in Psychology, 25,* 1–15.
van Dijk, M., & van Geert, P. (2007). Wobbles, humps and sudden jumps: A case study of continuity, discontinuity and variability in early language development. *Infant and Child Development, 16,* 7–33.
van Geert, P. (1989). Cognitive development: A process of organized self-organization. In G. J. Dalenoort (Ed.), *Paradigm of self-organization: Current trends in self-organization* (pp. 146–166). Amsterdam: Gordon and Breach.
van Geert, P. (1991). A dynamic systems model of cognitive and language growth. *Psychological Review, 98,* 3–53.

van Geert, P. (1994). *Dynamic systems of development: Change between complexity and chaos*. Hertfordshire, UK: Harvester Wheatsheaf.

van Geert, P. (1998a). We almost had a great future behind us: The contribution of non-linear dynamics to developmental-science-in-the-making. *Developmental Science, 1,* 143–159.

van Geert, P. (1998b). A dynamic systems model of basic developmental mechanisms: Piaget, Vygotsky and beyond. *Psychological Review, 105,* 634–677.

van Geert, P. (2003). Dynamic systems approaches and modeling of developmental processes. In J. Valsiner & K. J. Conolly (Eds.), *Handbook of developmental psychology* (pp. 640–672). London: Sage.

van Geert, P., & Steenbeek, H. (2005). Explaining after by before: Basic aspects of a dynamic systems approach to the study of development. *Developmental Review, 25,* 408–442.

Verspoor, M., Lowie, W., & van Dijk, M. (2008). Variability in L2 development from a dynamic systems perspective. *Modern Language Journal, 92,* 214–231.

Woodrow, L. J. (2006). A model of adaptive language learning. *Modern Language Journal, 90,* 297–319.