This is a Wordfile version of the article that appeared in Journal of Cognitive Education and Psychology

Unfolding potential as dynamic emergence

A view from the theory of complex, nonlinear dynamic systems

The notions of *unfolding* and *potential*

The question of unfolding potential is of key importance for the sciences of education and developmental psychology. It relates to two important, but hard to define concepts, namely that of unfolding and that of potential. A good way to refresh one's view on a particular concept is to go back to its origins of meaning. Etymology, is a branch of linguistics that can help us to so. An etymological analysis of the notion of *unfolding* reveals nothing new: we un-fold something that has been folded up such that we can see its real and final form. Let us compare this to the etymology of another important concept, namely the concept of *development*. De-velopment literally means to un-wrap. That is the same when we speak about the development of something, we imply that there is something there that needs to be unwrapped, that is already there but still invisible (for an etymological analysis of development and related terms, see Van Geert 1996). In this sense, the notion of unfolding is very closely related to that of development, if the latter is interpreted in its original meaning. Development as well as unfolding imply the notion of *potentiality*, i.e. that there is something there that needs to be unwrapped of unfolded to show its real form.

The concept of *potential*, as in learning potential or developmental potential has its historical roots in Aristotelian philosophy. These roots have very important consequences for modern theorizing. The

Aristotelian notion of potentiality implies a conceptual contrast between potential and actual. Potentiality implies a set of possibilities of the thing, given the way this thing works. It implies a possible actuality in the future. Actuality then implies the meaning of being-at-work-here-and-now. Actuality also implies a certain directedness which is inherent in the way this particular thing is "at work", how it functions. One may of course question the meaningfulness of this semantic analysis of unfolding potential and categorize it as a form of antique mumbo-jumbo, with little consequences for modern empirical science. However, in these antique notions reside meanings that are still of fundamental interest to a science of education and development, and that find an interesting scientific reinterpretation in the theory of complex, nonlinear dynamic systems.

The theory of complex nonlinear dynamic systems shown a rapid development during the last couple of decades, and its theoretical and mathematical foundations are now very well-developed. Nevertheless, its effect on the behavioral sciences, including education and developmental psychology, has so far been relatively small. Is this because the theory of complex dynamic systems is in essence not really applicable to psychological and educational phenomena? Or is this is because the application of the theory requires a level of empirical sophistication and data that the psychological and educational sciences cannot at present provide?

In order to answer these questions, we need to go little deeper into the notions of *complex*, *dynamic*, *nonlinear* and *system*.

Complex, nonlinear dynamic systems

The architecture of a complex system

A complex system is a collection of several components or elements, "several" ranging from a relatively small number to many hundreds, thousands or more (for more details, see for instance Van Geert 2008). These components are usually defined on a particular level of aggregation. A school class for instance can be defined on the level of its composing persons, and then consists of about 25 elements, one being the teacher and all the others being the students. The same class can also be defined on the level of educational roles, in which case it consists of two components, one the teaching person, the other the learning persons. The brain for instance, if defined on the level of neurons consists of billions of brain cells. If defined on the level of functional areas, the so-called Brodmann areas, it consists of about 52 components, each characterized by their own cytoarchitectonics.

In a complex system, the components are causally connected, and these connections can be direct or indirect (i.e. running via other components). A causal connection means that properties or events of one component affect, i.e. influence, events or properties of another component. For instance, a teacher's activity of giving a particular assignment is causally connected to the students making that assignment, or to some students actively resisting making that assignment. The firing of a group of neurons may stimulate other neurons in the brain to become activated too, or it may inhibit the activity of other neurons in the brain. In a complex system, the connections may be reciprocal, and reciprocity may be direct or indirect. For instance, if a teacher gives a particular assignment, and the students react with clear signs of resistance or discomfort, the teacher may decide to replace that assignment by one that is more adapted to the current level of the students. Hence, there is a causal connection from teacher to students and from students to teacher. In the brain, activity in one region may produce activation in other regions, which on their turn affect still other regions, which on their turn affect the region that started the activation wave. In a complex system, be it the brain, or a psycho-educational social system such as a classroom, the connections between the components may become quite complicated. Neuroscientists, for instance currently try to unravel the way the functional regions of the human brain are connected, and they have called this pattern of connections the *connectome* (see for instance Sporns, 2011). In spite of the incredible variety of possible connections between components of systems that consist of even relatively few components (e.g. between 10 and 50), modern network theory has succeeded in finding certain qualitative properties that distinguish families of connection patterns (Newman, 2003, 2010). These qualitative properties are related to the ways the complex systems in question are functioning. The pattern of connections between functional regions in the brain, for instance, appears to share typical qualitative characteristics with the pattern of connections we can find among the computer servers that together form the World Wide Web (Sporns, 2011; Bullmore & Bassett, 2011; Rubinov & Sporns, 2010). This characteristic, which is the so-called "small world" pattern, implies that every component is connected with relatively few other components, in such a way that most components can be reached from any other component in a relatively small number of steps (i.e., via a small number of components that are connected to each other). This type of connectivity leads to systems that are relatively robust, i.e. that are relatively insensitive to changes from the outside.

Properties of complex systems

The fact that events in one component causally influence particular events in another component (or components), and that such causal connections may be directly or indirectly reciprocal, leads to several interesting properties of such systems. The first property is that the system is characterized by a particular temporal flow of events. What happens in the classroom for instance is a particular temporal flow of actions of the teacher and of the students, in which one action triggers another one, which on its turn triggers yet another one and so forth, with many such actions occurring simultaneously. Another property is that, because of this flow of events, the components in the system tend to change, that is to say that in the long run they can change the way they react to events in other components. For instance, a teacher who tends to overestimate the cognitive levels of their students and who consistently provides assignments that are too complicated for them, may cause his students to adopt an overly resistant attitude towards his assignments, which might then lead the teacher to lower the demands put on the students, or, alternatively, lead the teacher to persist even further and to consolidate his tendency to give overly complicated assignments. As this example illustrates, the changes in the components are not deterministic, and, in fact, depend on the way the components interact and on the components' history. The third property is in fact the most basic property of a complex dynamic system: as a result of the temporal flow in the system, the components including their connections undergo changes, and these changes result in creating a particular order, structure or organization of the system. If everything, in some way, affects everything else, one might expect that the flow of events in a system might become utterly chaotic and random, also because the system is always affected by random events from outside (e.g. events from outside the classroom, or from outside the brain). What happens, however, is exactly the opposite. Systems consisting of components that react to events in other components on the basis of a limited set of reaction principles, tend to undergo a spontaneous organization, leading to self-sustaining patterns or structures that resist influences from outside, meanwhile remaining flexible enough to show long-term adaptations. This originating of self-sustaining patterns as a consequence of the interactions between the components is what is called self-organization, which is a defining feature of complex dynamic systems (for general discussions, see Van Geert 2003, Van Geert and Fischer, 2009; Van Geert and Steenbeek, 2005). Complex systems, such as the weather, or the economy, are characterized by typical forms of organization that result from the way in which their components interact. In the economic system, the components are participants such as consumers, producers, financers, politicians and so forth, and their interactions produce dynamical states that we characterize as recessions, growth, booms, cycles, sudden crisis and so forth. For relatively simple systems, researchers have been able to unravel the principles of self-organization, but for systems of increasing complexity, self-organization can be observed but not necessarily explained in all its details. What is important, though, is that the self-organization is like a default option for the world in which we live. That is to say, the patterns we observe in the world - from the formation of flocks of birds to the complicated cycles of human economies or the growth of a mammalian fetus - are most likely the result of self-organization. That is, based on what we know about physical and in particular living systems, we may safely assume that such structures neither result from design rules originating from some master organizer, nor result from a simple addition of all the influences that affect them.

Self-organization

The notion of self-organization refers to the origin of a particular structure, organization or pattern. For instance, when children develop language, the use of particular kinds of words and particular word orders becomes organized into a pattern that we can describe as the grammar of a particular language, or the person's language competence, i.e. it way of organizing sounds (or gestures) and meanings. When children develop a strategy or multiple strategies for solving cognitive problems, they develop particular ways of organizing information retrieval, material changes, memories etc. that lead to the solution of particular problems (correct or incorrect). In psychology, the question as to where this organization comes from has traditionally been answered by referring to an external organizing source. For instance, children learn the pattern of language because they imitate this pattern from their parents. Children develop problem-solving strategies because these strategies are taught by their teachers. That is to say, there exists some form of organization — for instance in a parent's language — and this organization is then transmitted from one source (a parent) to another (a child). Beginning with Chomsky's 1959 critique on Skinner's book on verbal behavior (Chomsky, 1959), it has become clear that the mechanism of transmission does not work for complex abilities such as language, and that it is highly likely that it does not work for abilities involving complex cognition either (Van Geert, 2009). The classical answer to the failure-of-transmission argument was to revert to another source of structure, namely the genome. If the complex grammar of natural languages or a complex form of cognition, cannot be transmitted, they must be present in the form of genetic instructions. That is, the particular organization that unfolds during development is in a sense copying the organization already present in the genes. However, it is highly unlikely that genes can be conceived of as complete instruction lists (Kauffman, 1993). The importance of the complex, nonlinear dynamics paradigm is that it has shown that organization at the level of a whole (e.g. language as a whole, theory of mind as a whole, problem-solving strategies as a whole etc.) can spontaneously result from the interactions of the components or elements that constitute the whole, i.e. that self-organization is more or less the default option in natural systems. What is needed is that a component reacts to another component in some systematic way. These reaction rules are mostly very simple, such as the rule that one component increases in magnitude if another increases in magnitude, or if yet another one decreases in magnitude. A typical biological illustration is the emergence of movement patterns of groups of animals, such as a school of fish or the V-pattern of migrating geese. These patterns are typical of the whole, e.g. of the ensemble of fish or the ensemble of geese and they are of course not a property of a single fish or a single goose. Likewise, the pattern we call "theory of mind" or grammar is not in the individual components, such as separate judgments of other people’s intentions, or separate words or sentences, but in the totality of these components, given the ways they interact. That self-organization is in fact the most likely thing to occur in systems of interacting components, is the major discovery of the theory of complex, nonlinear dynamic systems.

In fact, the theory of complex, nonlinear dynamic systems forms a new scientific paradigm, in the sense that it offers a starting point — a default theory — for the phenomena under study. In the educational and developmental sciences, the often tacit, underlying paradigm is quite the opposite: phenomena such as a child’s language ability or its use of particular problem-solving strategies are conceived of as the sum of independent variables that affect them, such as intelligence, motivation, quality and duration of instruction, and so forth. The default option is that if we find some sort of structure or pattern arising in an individual's behavior, such as language or problem-solving strategies, that structure or pattern must have been transmitted from an external source, such as a teacher or a set of genetic instructions (note that "external" means external to the variable of interest; for instance an external influence on a child's language could be the child's cognition, or the instruction the child receives). The main problem with his general linear paradigm is that it leads to an infinite regression, and secondly to an obviously wrong prediction about the course of development on various time scales. First, if structures arise because they are borrowed from another source, then where do the structures in that other source come from, if not from yet another source? Second, if structures arise because they are transmitted from one source to another, then over the long run, structure must inevitably diminish in terms of complexity and coherence (the justification of this argument exceeds the scope of the present article, but see Van Geert, 2009 on transmission theory). If it is indeed so that, in nature, structure emerges out of the interactions between the components, the problem of infinite regression is immediately solved. In addition, if structure is emergent on the interactions between the components, the new structures can take the role of new components, implying that over the course of time (e.g. over successive human generations) structures such as human cognition or technology become more instead of less complex. In summary, the theory of complex nonlinear dynamic systems should be treated as a new paradigm, leading to a new "default" view on the nature of reality, including educational and psychological phenomena, as self-organizational processes by default. Simple linear additions of effects and forms of transmission can be treated as special, limiting cases of self-organizing processes.

It is interesting to note that the classical European theories of cognitive development, such as the theory of Piaget or Vygotsky, contain all the elements of what is currently conceived of as the theory of complex, nonlinear dynamic systems (Van Geert, 1998). The North American, learning theoretical tradition on the other hand, strongly proposed a linear into transmission view on educational and psychological processes. In short, the position taken by the present author and other advocates of the dynamic systems approach in the life sciences is that we should treat educational and psychological phenomena as complex dynamic, self-organizing systems, unless proven otherwise.

Development, emergence and novelty

The theory of complex, nonlinear dynamic systems also provides a natural way for describing and explaining the problem of novelty, which is a major theoretical and empirical problem in the field of educational and developmental sciences. When children acquire language — by whatever means they do so — they have acquired something that is entirely new for them, in comparison with the abilities and cognitive contents that they already possessed. When children develop a concept of systems of relationships (Fischer and Bidell, 2006), they have acquired something that is entirely new in comparison with the knowledge that preceded the concept of systems of relationships. As we have seen, a traditional way of solving this kind of problem is by referring to the process of transmission: new structures are transmitted from other sources, e.g. teachers. However, transmission does not solve the problem: how can a structure characterized by a particular level of complexity (e.g. cognitive complexity at the level of single relationships between concepts) accommodate for a structure that requires an entirely different level of complexity (e.g. systems of relationships), if that level of complexity is not already there, waiting to receive a content of a particular level of complexity. This reasoning is known as *Fodor's argument for the impossibility of learning* (Fodor, 1980, Amini, 2011). Fodor's argument rests upon two assumptions, namely the assumption that knowledge is defined by having representations (and that thinking amounts to processes carried out on such representations) and the assumption that (cognitive) development occurs through transmission of cognitive contents. These two assumptions are also deeply rooted in the standard linear approach to psychological and educational phenomena. Fodor's argument refutes the transmission assumption, and replaces it by the innateness assumption. However, according to the theory of complex, nonlinear dynamic systems, knowledge is not a matter of representations, but rather a matter of embodied processes, consisting of perception-action loops involving a person's memory and affordances of the material context. In addition, the theory of complex, nonlinear dynamic systems proposes self-organization as the prime source of structure and organization in complex phenomena and circumvents the position that knowledge structures are either transmitted or innate.

The notion of self-organization and the problem of how novel, more complex forms can arise during development — in individuals or in the form of cultural evolution — come together in an important notion characteristic of complex dynamic systems, namely the notion of the *emergence*. A dynamic system is a set of interacting components, that is, the components are bound together through their interactions, thus forming a whole. For instance, relationships between words belonging to different categories such as prepositions or nouns constitute a whole, which is a language characterized by a coherent grammar. Once a particular coherent structure, such as a grammar has originated, it becomes a source of causal influences in its own, affecting the behavior of its components (Haken, 2006), thus contributing to its self-sustenance. The process of creating properties of the whole as a consequence of the dynamic relationships between components is called emergence, and is considered to be a very important concept in the theory of complex, nonlinear dynamic systems. We may thus assume, that most, and certainly the important features of the human cognitive system that originates during development are the result of emergence. The concept of emergence is perfectly compatible with notions such as guidance and teaching. In fact, for typical human cognitive abilities and skills to emerge, the process of emergence must be scaffolded by processes of social mediation, support and instruction mostly by more competent others, i.e. teachers (Van Geert, 1998). Put differently, education is primarily a process supporting emergence in individuals, and not in the first place a process of transmission and shaping by external sources. There is no doubt, that many researchers will share these general ideas on the nature of education and development. The problem is however that they require another underlying paradigm than the linear paradigm they are used to. Unfortunately, the standard paradigm is constantly verified by the way we do our research, which is basically to aggregate data over individuals and to study associations between variables over individuals instead of studying such associations within individuals, i.e. in the form of individual processes.

(Non-)Ergodicity in educational and psychological processes

This point brings us to another, primarily statistical point, which in spite of its importance, is still hardly known by the educational in developmental research community, namely the point of non-ergodicity of the phenomena under study.

Let us assume that a clinical psychologist must decide about some form of therapy or intervention for a child with problems regarding attention. At school, the child shifts his attention too rapidly from his math assignment to chatting with this peers, the child is overly sensitive to immediate rewards and so forth. If the psychologist consults the literature, he will find models of the associations between a whole range of important variables, such as working memory, various executive functions, motivation, intelligence and so forth. This theory about the associations between all the variables is typically based on research designs that compare individuals. For instance, with regard to working memory and say, analogical reasoning, researchers will commonly select a representative sample of children , and for each of these children the level of working memory and analogical reasoning is determined by means of standardized tests, which in the end results in a specification of the association between working memory and analogical reasoning. It is important that the sample is representative of the population, in order for the results to be generalizable, which means generalizable to the population. The idea is that a *general* theory is by definition applicable to *specific* cases. That is to say it is assumed that generalizable findings about the association between working memory and analogical reasoning, or between working memory and executive functions are applicable to the child for whom we wish to design a particular treatment, or whose problems with attention we seek to understand.

The reasoning that a theory holding for a representative set of all possible observations of a particular phenomenon, also holds for a specific subset of those observations, for instance in the form of repeated observations in a single person over a sufficient amount of time, appears so obvious that it is almost not worth discussing. However, this reasoning relies upon an unproven and most likely also false assumption that the statistical structure of relationships between variables holding between subjects - which is what the observation of all possible observations of that relationship actually entails -, is statistically similar to the structure of relationships between those variables holding within subjects. There exists a particular term referring to the condition in which the statistical structure of some property of interest, derived from the ensemble of subjects, is similar to the statistical structure of this property of interest, derivable from any single individual from this ensemble, which is ergodicity. Suppose a researcher wishes to know how inhabitants of Amsterdam like the parks in their city. According to Tarko (2013)

> One idea is to take a momentary snapshot: to see how many people are this moment in park A, how many are in park B and so on. Another idea is to look at one individual (or few of them) and to follow him for a certain period of time, e.g. a year. Then, you observe how often the individual is going to park A, how often he is going to park B and so on. Thus, you obtain two different results: one statistical analysis over the entire ensemble of people at a certain moment in time, and one statistical analysis for one person over a certain period of time. … The idea is that an ensemble is ergodic if the two types of statistics give the same result.
> (<http://news.softpedia.com/news/What-is-ergodicity-15686.shtml>)

Chances are that the profile of park preferences defined by the design that sampled over individuals will apply to relatively few people, each of whom is likely to have its own particular profile. The difference between the methods is that they provide a different kind of information. The first method tells you something about parks in Amsterdam, namely how popular they are with people, with “people” being the “unidentified” ensemble of all people in the city of Amsterdam. The second method tells you something about individual people, namely how fond they are of parks and how they differ in their preferences (e.g. some people don’t like parks at all, others like very particular parks, and still others like parks that attract many people).

If we apply this example to the method of finding relations between variables by sampling across the ensemble of all possible persons, we can conclude that the first statistical method - sampling across individuals - tells us something about *variables*, namely how they are distributed across people. The method of finding relations between variables by sampling across consecutive events in a particular person tells us something about *people*, namely how particular variables behave during the course of their activities, in the short or in the long run. Moreover, this method tells us something about *specific* people, namely the people in whom we have sampled the successive observations. That is, the observations lead to *idiosyncratic* models, i.e. models that apply to special cases (Molenaar, 2004; Molenaar & Newell, 2010; Molenaar & Campbell; 2009; Molenaar, Sinclair, Rovine, Ram & Corneal, 2009). The information provided by these two methods serves different purposes, in terms of generalization or specification. Given the first type of information, based on samples across subjects, you can formulate particular expectations about some new or unknown sample of subjects (including a sample consisting of just one subject). Given the second type of information, based on samples within a subject, i.e. time samples, you can formulate particular expectations about some new time sample for this subject, for instance the time sample in the near future. The idiosyncratic theory upon which counseling and treatment of a particular child must be based is very likely to be different from the theory based upon observations *across* subjects (Fisher, Newman and Mo lenaar, 2011; Hatfi, McCullough, Frantz & Krieger, 2010; Hayes, Laurenceau, Feldman, Strauss & Cardaciotto, 2007; Hoenders, Bos, de Jong & de Jonge, 2012; Laurenceau, Hayes & Feldman, 2008; Polman, Bouman, van Geert, de Jong & den Boer, 2011;Rosmalen, Wenting, Roest, de JOnge & Bos, 2012; Schiepek, 2009).

A related question deals with the validity of idiosyncratic theory: is it just an illustration, with some individual particularities, of the general principles found by comparing many subjects? Shouldn't a scientific theory be a general theory, one that applies to all specific cases? Although idiosyncratic models may differ for each subject they describe, in the limit implying that there will be as many idiosyncratic models such model should be based on a general theory of the underlying processes. The theory of complex, nonlinear dynamic systems is ideally suited for constructing general theories of the underlying dynamics, that can be applied to individual cases, resulting in idiosyncratic models based on general theory. Examples of such general principles of change might be that each idiosyncratic pathway is the result of dynamic interactions in a network of connected variables, based on specific connection principles. An example of such specific connection principles is that the network may be connected in the form of a so-called small world network (most of the connections occur between relatively similar variables, with very few connections occurring between variables that are quite different from one another). It has been shown for instance that this principle which provides an explanation of the structure of the world wide web, also applies to the way functional brain regions are connected (Barabasi, 2009; Bullmore & Sporns, 2009; Bullmore & Bassett, 2011; Rubinov & Sporns, 2010). Knowing that the brain or any other system for that matter is connected in the form of a small world network, allows one to explain why the system is relatively stable on the one hand, but why it can sometimes show quite massive and relatively rapid reorganizations.

The methodological consequences of these principles are that research should focus on individual time series, because it is the individual time series that provides information about the working of a particular system. The very important question of whether a particular type of system is indeed ergodic or not (sufficiently close to ergodicity to warrant research based on groups of individuals) can only be answered by means of detailed studies of individual trajectories of learning, clinical intervention and development. The term individual should be taken quite broadly however, and used in a generic sense, not necessarily meaning a single person. For instance, an individual time series may apply to the short-term and long-term interaction between a specific teacher and a specific child, or a specific teacher and a specific school class, a specific family and so forth. Research should become idiosyncratic, in the sense of individual-centered, and process-oriented, i.e. consisting of time series collected from individual systems. In the context of developmental psychopathology, individual-centered and process-oriented approaches to developmental psychopathology could offer a way out the over-diagnosis problem, that is to say that too many children are diagnosed as having ADHD, for instance, because for some time in their lives they showed behavioral properties that approach the standard diagnostic criteria (Batstra & Frances 2012abc; Batstra, Hadders-Algra, Nieweg, Van Tol, Pijl & Frances, 2012). If a condition, such as ADHD, is considered as a stable trait property of an individual, a once-only diagnosis tends to become a lifetime assignment. An individual-centered and process-oriented approach might provide information about the extent to which developmental psychopathological trajectories might temporarily converge but later also diverge from a particular standard pattern.

The common objection against individual-centered and process-oriented research, which in view of its labor-intensive nature is often published in the form of single case studies or very-small-n studies, is that its results are not generalizable. However, it is important to note that there are at least two different notions of generalizability that should not be confused (Lee & Baskerville, 2003). One, which is the more standard notion, refers to the relationship between a statistical property of a sample (e.g. correlation between two variables) and the corresponding property of the population from which the sample is taken. The other refers to the relationship between a theory and all possible individual cases. That is to say, a general theory of a particular educational or developmental process should be applicable to any individual process case. The theory of complex, nonlinear dynamic systems claims that it can offer such general process theories of educational and developmental processes, although we must acknowledge, however, that such an approach to education and development is still in its infancy, and still needs to prove its viability.

From theory to specific models

Modeling dynamic systems

Technically, a dynamic system can be described as a means of describing how one state of the system develops into another state over the course of time (Weisstein, 1999). The state of the system can be described in terms of the components, or in terms of some emergent property of the system. Let us give an example of a relatively simple system, consisting of a mother and a child, each of which is defined by mean length of utterance (MLU) in spontaneous speech (van Dijk et al., 2013). Because it is a mother-child system, we describe the system by the mother's mean length of utterance (MLU) in her child-directed speech, and the child by her mean length of utterance in mother-directed speech. As this example very clearly shows, we have reduced the real mother-child system to a single variable of interest, namely their mean length of utterance. We have done so, because we have strong reasons to believe that, up to a certain age, MLU is an important indicator of a speaker’s mature linguistic competence. We could have used other indicators, such as preposition use, or the mean number of different nouns in an utterance etc. (Bassano et al., 2011). Each indicator can tell us something about the child's linguistic capacities at a particular moment in time. More importantly, since we are interested in the question whether the mother adapts her language production to that of the child, and the child learns from the mother, many other indicators would have been equally feasible. If we confine ourselves to the MLU example, the state of the mother-child system at a particular moment of time - for instance the first of a series of observations of spontaneous language use, when the child is 14 months old - is the MLU used in the mother's child directed speech, and the MLU used by the child (which at this particular age is probably equal to one). During the next observation, one week later, we might observe that the mother is using less or more words per utterance in her child-directed speech than a week before, whereas the child is still in its one-word stage. The third observation showed yet another combination of MLU's in mother and child, and so forth until the end of the sequence of 36 observations. In fact, the state of the system is described in a two-dimensional space, one dimension being the MLU of the mother, and the other dimension the MLU of the child (see figure 1). With two dimensions, the state space is similar to the familiar X- Y graphs that are often used to chart the associations between variables.

Insert figure 1 about here

It is easy to see that at each moment in time, e.g. defined by a particular observation session, the mother-child system is characterized by a single point in this two-dimensional space (see the figure). During successive observations sessions, i.e. successive moments in time, the points characterizing the mother-child system will shift position, simply because the MLU of the mother and of the child change across time. If we connect the consecutive points in the two-dimensional space by a line, we obtain a drawing of the trajectory that the mother-child system follows in the two-dimensional MLU space over the course of time, e.g. during one and half year, specified by weekly observation sessions. Technically, this two-dimensional descriptive space is called the system's state space, because it is the space in which the states of the system, as defined by its values on the two dimensions, can be specified. The state of the system can also be specified by the region in the state space within which the observations very during a particular period of time (see figure).

The main question for a dynamic systems model builder is to provide a formal specification, for instance in the form of a mathematical equation, of how a particular state of the system develops into another state, which is the next state in the systems trajectory. More simply stated, what the dynamic model builder says is: give me a description of a particular state of the system, and I will give you the description of the state that follows it. Or, to put it differently, the next state is some sort of function (f) of the present state. Therefore, in its crudest possible form, a dynamic model always looks like this:

$X_{t+1} = f(x_t)$

and a dynamic model for the mother-child system characterized by MLU, in which the MLU of the mother is affected by the MLU of the child, and the other way around, would in its crudest possible form, look like this

$MLU\_M_{t+1} = f(MLU\_M, MLU\_C_t)$

$MLU\_C_{t+1} = g(MLU\_C_t, MLU\_M_{t \text{ him and}})$

Hence, a dynamic model shows how the current state of the system is some sort of transformation of the preceding state of the system, and how the next state of the system is a transformation of the current state, and so forth. This sequence of transformations in time — for the time steps t, t+1, t+2, t+3, t+4 … t+n r and so forth — produces a trajectory in the state space. If the dynamic model provides an adequate description of the underlying system dynamics, the trajectory that results from the dynamic model will be similar to the observed trajectory.

In principle, the next state of the system differs from the preceding state (but note that the difference may be zero, we shall see later why this is important). This difference is the amount of change that the system undergoes over a particular time duration (e.g. the time between one observation session and the next one). Hence, another way of expressing a dynamic system is to specify how the *amount of change* of the properties that define the system's current state, depend on those properties. Such specification takes the form of a differential equation (or a difference equation, such as the ones presented above, but those are technical details). Differential equations abound in physics, but they are very hard to find in the behavioral sciences, including the sciences of education and development. This is quite surprising, because the latter disciplines are focusing primarily on change. A typical differential equation might look like this:

$\Delta MLU\_C/\Delta t = g(MLU\_C, MLU\_M)$

which should be read as follows: the *change* in the child's MLU over a particular time change, is a function, expressed as *g*, of the child's MLU and the mother's MLU.

Suppose now we have some starting point, for instance a parent using the habitual mean length of utterance in her standard adult-directed speech, and a 14 month old child who has just entered her single word phase, i.e. whose MLU is 1. Suppose also we have a dynamic model in the form of differential equations describing the change in these two variables over the course of time. We apply the equations to the starting point, obtain the next point, apply it to the next point and obtain a successive point and so on. What we will see in typical dynamic models of developmental and educational change is that the amount of change in the variables gradually diminishes, up to a point where the application of the differential equations to the state of the system produces a state that is equal to the preceding state. That is to say, the dynamics produce zero change. Zero change, which is a consequence of the dynamic principles operating on the system, means that the system has now moved toward stability, which means that its principles of change now *create and maintain* stability. This stable level is the system's *attractor state* under the given dynamics. We call it an attractor state, because the principles of change automatically lead the system towards this particular state, and to different, somewhat metaphorical way of saying this is that the system is attracted towards this state. For the mother-child's MLU system, the stable state is one in which the child's mean length of utterance is (approximately) similar to the mother's mean length of utterance in child directed speech, and where the mother's mean length of utterance in child directed speech is (approximately) similar to her mean length of utterance in her common, adult-directed speech.

The graphical structure of the model is represented in figure 2

Insert figure 2 about here

The dynamic model can be represented as a set of three coupled differential equations, in which I stands for linguistic “input”, i.e. the mother’s child-directed speech, capital L stands for the child’s level, in this case the child’s MLU, and A stands for the mother’s tendency to adapt their language to that of the child (for more details, see Van Dijk et al., 2013).

First, the mother’s actual adaptation to the language of the child is expressed as follows:

$I_t = I_H + s\, A_t\, (L_t - I_H)$

(I_H is the parent’s habitual level in adult directed speech, L_t is the child's current level of the linguistic variable at issue, A_t is the parent’s current tendency to adapt her language to that of the child, and *s* is a damping parameter).

It was assumed that the parent's tendency to adapt her child-directed speech, represented by the A_t parameter, changes in function of the actual acquisition rates of the child. If a child makes only little progress in the learning of a particular linguistic variable, in spite of the parent's adaptation, the parent is likely to increase her tendency to adapt. That is, she will tend to bring the level of her language more closely to the level of the child's language (which is a generalization of dynamic scaffolding theory, Van Geert and Steenbeek, 2006). The equation expressing the eventual change in the parent's tendency to adapt is as follows:

$$\Delta A_t = a\ A_t\ (1 - \Delta L_t/S)$$

(a is a damping parameter, and S is a scaling parameter; and and ΔL_t is the current amount of change in the child's linguistic variable, which in this present case is MLU).

Finally, the child's change in the preferred level of MLU, represented by ΔL_t , is modeled by a logistic growth equation (for an explanation of the logic underlying these equation see Van Geert, 1991).

$$\Delta L_t = r\ L_t^p\ (1 - L_t/(q\ I_t))$$

(p is an exponent modifying the effect of the level of L on its own growth, and q is a proportion specifying the extent to which the child can currently approach the parent's actual input level, I_t).

Figure 3 represents smoothed data curves and fits of the dynamic model (for the concept of smoothing, see further)

Insert figure 3 about here

Given these equations, the mother-child MLU system typically moves towards a single point in the state space, i.e. towards a point attractor. In practice, behavioral systems will never move towards literal point attractors. Rather, they will move towards a particular, confined region of the state space. The actual states of the system will vary within this region. Other systems might move towards more complicated attractors, for instance cycles. Sleep-wake cycles, and more generally the circadian rhythm, are typical examples of a cyclical attractor. They are regulated by neurophysiological processes in the suprachiasmatic nucleus, which is part of the hypothalamus, producing a cyclical behavior which is the result of the couplings of about 20,000 neurons (which in itself is a nice example of a complex, nonlinear dynamic system). The process can adapt itself to changes in daylight and intentional sleep-wake cycles, as anyone knows who has ever experienced a jet lag. In a jet lag, one can experience how a cyclical attractor pattern gradually changes into a new, stable cycle over the course of days. Another example from the behavioral sciences, which might be a bit more speculative, are the cycles of aggressive behavior in students with behavioral disorders and behavioral corrections by educators (Van Geert & Steenbeek, 2003; Abraham, 2014). Attractors may change in various ways. For instance, the MLU used by a child is a typical range attractor if observed over the short-term time course of conversations between a particular mother and child (Van Geert & van Dijk, 2002). However over the longer-term time course of months, it tends to change in a relatively gradual way. As with many comparable variables, the long-term changes take the form of accelerated and decelerated growth, resulting in typical S-shaped patterns (the growth of prepositions follows a similar pattern for instance, Bassano et al., 2011). Attractors may also change in discontinuous ways. That is to say, a child may function either in an attractor state A, or in an attractor state B, and never occupy a position in between. This pattern of discontinuous change has been extensively studied in areas of cognitive development, such as conservation or the use of cognitive rules in problems such as the balance scale (van der Maas & Molenaar, 1992 additional references;).

Studying educational and developmental phenomena as dynamic systems

In order to empirically investigate a particular educational or developmental phenomenon, such as a child's learning to read or the development of conservation in a child, the focus of study must be on the temporal flow in the system. Note that we defined a dynamic system as a state space, with each dimension of the space representing a component of the system, which is usually a variable characterizing subjects (e.g. as in our mother-child preposition system) or a variable characterizing components of system (e.g. a system consisting of the use of prepositions and the mean length of utterance of a particular child, Van Dijk & Van Geert, 2011). The flow is the movement of the system's state through the state space, as explained in the preceding section. Maybe the notion of "flow" is not entirely adequate here, because it assumes a kind of smoothness in the change that the real behavioral data seldom show. Real behavior data change in a relatively erratic way, in that they jump from one point in the state space to another one (see the graph of the MLU data), although the jumps usually take place within a confined region of the state space, which changes over the long-term. Technically, we can call such jumping behavior a "map" in that one state of the system, as defined by the system’s state space, is mapped onto another state of the system, namely the system's next state. What should we do to collect data on a system's dynamic flow or dynamic map?

Step one: selecting a system of interest

To begin with, we must first select a system of interest. An example of such a system is a particular child and her mother, which is defined by the child's and the mothers MLU. Another example is a particular child, who is defined by her use of prepositions and her mean length of utterance, or by the frequency of use of various strategies for solving a particular type of problem. Note that we explicitly refer to specific individuals, specific dyads (e.g. a specific mother-child dyad) or a specific group (e.g. a specific school class). A sample of children, or a sample of mother-child dyads is not a system: the sample does indeed consist of components (e.g. the individual children) but they are clearly not causally interacting with one another (the nature of the statistical sample is that the cases or "components" are strictly independent, whereas a system is defined by components that are co-dependent in some way). Something that can be conceived of as a connected system and that can be characterized by some sort of state space, will show a particular trajectory over time in its state space, and it is this trajectory that we will have to empirically investigate.

Step two: defining the structure of the state space

The system is defined by particular variables, that is to say, by properties that can change. If the variables are of the ratio type, such as the frequency data of the use of prepositions, we can use a state space that consists of numerical variables (which, in the case of frequency data are non-negative integer variables). In the educational or developmental sciences, many variables are of a categorical type. For instance, emotional interactions between a mother and a child may be scored as being positive, neutral or negative. In this case, we have a categorical state space, which looks a bit like a checkerboard. In the example of three emotional categories for each participant, the total number of states or cells in the categorical state space is equal to nine (any combination of an emotion of the mother with an emotion of the child). In the state space, the system may jump from one cell to another, and do so in a characteristic way. Figure 4 provides an example of a system consisting of a mother presenting food to an infant in various ways (she can scoop the food, wait, knock on the child’s mouth, put the spoon in the child’s mouth and retract the spoon) and an infant accepting the food in various ways (e.g. by swallowing it, refusing it, spitting it out and so forth; Van Dijk, Hunnius & Van Geert, 2012).

Insert figure 4 about here

The study of dynamic maps (temporal patterns of change) in categorical state spaces is now becoming increasingly popular in the educational and developmental sciences. The success is in part due to the fact that many empirical investigations in these disciplines use categorical rather than ratio variables, and in part due to the fact that there is a versatile piece of freeware available (the so-called Gridware; Hollenstein, 2013). Gridware allows researchers to easily map data from individual patterns of change, and to directly compute the statistics that are typically related to such temporal mappings (we shall discuss some later). In fact, the success of these categorical state space methods might lead to the erroneous impression that the application of dynamic systems thinking in the educational and developmental sciences basically amounts to applying state space grids.

Step three: defining the sampling frequency

A system changes over time: during each conversation, mother and child each produce a particular number words per utterance, and this number differs from conversation to conversation. Note that this particular case the typical time step corresponding with the way we have defined our variable (MLU during a conversation) corresponds with a mother-child conversation session. Hence, in order to sample the system, the researcher should sample every conversation the mother and child are holding. This is of course not practically feasible, and most researchers will try to sample enough data to obtain a reliable impression of the way the system changes over time. For instance, in our own studies of math lessons in special education, we focused on a particular student and teacher and videotaped a teacher-student math lesson once a week, for a total duration of two years (Steenbeek, Jansen & Van Geert, 2012; see figure 5).

Insert figure 5 about here

The best sampling frequency is a compromise between practical possibilities of measurements or observations on the one hand, and the nature of the dynamic pattern or trajectory in the state space. In the very unlikely case that a system follows a strictly linear increase through the state space, we can suffice with two measurements, and the length of the interval between these two measurements will, theoretically, not matter (a straight line is always defined by two points, irrespective of the place of those points on the line). The problem is of course, that for most developmental and educational process phenomena we might be interested in, we have only very little preceding knowledge of their natural temporal flows or mappings, which implies that the choice for a sampling frequency is often done on the basis of an best educated guessing. A solution might be the adoption of a mixed frequency design, that is a combination of bouts of very frequent observations or measurements interspersed by longer intervals (Van Dijk & Van Geert, 2002; Lichtwarck-Aschoff et al., 2008).

While defining the sampling frequency, it is very important to take into consideration that short-term fluctuations (e.g. short-term fluctuations in the frequency of MLU or preposition use, or short-term fluctuations in the score on a conservation test which is administered repeatedly) provide important information about the underlying dynamic system. Short-term fluctuations should not be treated as measurement error (Van Dijk & Van Geert, 2002). Such fluctuations, i.e. intra-individual variability, characterize specific patterns of change. For instance if the underlying system undergoes a so-called phase transition (i.e. the system changes its underlying organization) intraindividual variability may temporarily increase (Bassano & Van Geert, 2007; Van der Maas & Molenaar, 1992). As a particular system, for instance the cognitive system underlying the reading process, increases its internal coordination and quality of functioning, fluctuations may change from so-called white noise (indistinguishable from sequences of independent random events) to so-called pink noise, in which the random variations are correlated over longer time frames (Wijnants et al., 2009, 2012). Many psychological phenomena are considered to be scale free in their variability, which implies that variability sampled over short intervals is statistically indistinguishable from variability sampled over longer intervals (Kello et al., 2010). In order to discover whether a process is scale free - a property that depends on the underlying process dynamics - one must sample over short as well as long intervals.

If distinct events are the subject of study, the sampling problem is easier to solve. For instance, in our studies of scientific reasoning - children's explaining and predicting of phenomena in the context of simple science experiments - we videotaped a relevant classroom interaction and then scored the cognitive level of each act of explanation or prediction by means of hierarchical skill theory (van der Steen et al., 2014; Meindertsma et al., 2013). By doing so we can sample all instances of reasoning during a particular educational event, namely a particular science class. However, as soon as we wish to study the long-term changes in scientific reasoning of a particular child, or of a particular school class, we have to decide on how many such science class events we shall have to sample: all of them, one class every month,...? The answer to such questions can only be given by following various schedules of observation and comparing the evidence resulting from it.

Step four: describing and statistically processing the data

Many real time series look rather erratic, and this becomes increasingly problematic with multivariate timeseries. For instance, figure 6 shows the cognitive complexity levels of questions asked by an adult and answers given by a boy of four years and eight months old (from van der Steen et al., 2014).

Insert figure 6 about here

We expect to find a relationship between the level of the questions and the levels of the answers, but this relationship is hard to detect in the raw data. In fact, a simple correlation of the cognitive levels of the adult's questions and the cognitive levels of the boys answers shows a value of -0.13, which might lead us to conclude that the association between questions and answers is negative but very weak (about 1% of common variance). However, in educational conversations, associations between questions and answers may be considerably more complex than the relationship that can be expressed in the form of a correlation, and correlations may actually conceal what is really going on. For instance, there might be delays in the child's picking up the cognitive challenges in the questions of the adult, and the adult might adapt the level of her questioning only after repeated observations of unsuccessful hints. There exist many techniques for finding relationships in time serial data, but in the current article I will first focus on techniques that are primarily visual, i.e. that show the complexity of the eventual interrelationships, instead of reducing them to a single numerical value that, in many instances, tells us little to nothing about the actual process.

Nonlinear smoothing techniques

A standard technique to extract information from erratically varying data is to calculate some sort of central tendency. For instance, data from a sample of individuals may be reduced to an average value. An association between two variables - such as cognitive levels and age - may be reduced to a central tendency in the form of a linear regression line. From a dynamical point of view, these reductions to central tendencies are ways of summarizing or describing the data, and should not be considered methods to find the "real" curve underlying the ever-laden measurements. For instance, the fact that we can draw a linear regression line through a timeseries does not imply that the straight line is the real trajectory, and that the observations are the result of measurement error imposed on the real value. In fact, the time series should be described in such a way that the more interesting details are conserved and the less interesting details are temporarily put aside, leading for further analysis, if needed. One way to reduce excessive information from time serial data is to use *nonlinear smoothing techniques*. A simple form is the moving average: the researcher determines a certain window size, e.g. five consecutive observations, and calculates the average value for these five observations, then shifts the window to the right with one step, i.e. one observation, calculates the average again and so forth. More sophisticated techniques apply weighted regression models over moving windows of observations. One such technique which is widely applied, is LOESS smoothing, which stands for locally weighted regression modeling.

Figure 7 is an example of the change in mean length of utterance in a parent and the child over 38 consecutive observations sessions between the ages of 14 and 32 months, which we discussed earlier. The raw data show considerable fluctuations, especially in the mother (see section xx).

Insert figure 7 about here

The figure shows the LOESS smoothed data in addition to the raw data. In addition to the local variations in the central tendencies (let's say, the current local centers of variation), the figure also shows a smoothed representation of the bandwidth of the data. The bandwidth is the local range of variation in the data. In the child, the bandwidth is gradually increasing, but in the mother the bandwidth is quite variable, with a maximum around the onset of the 2-3 word sentences in the child.

In order to better see the eventual coherence between mean length of utterance in the mother and in the child, the smoothed curves can be *normalized*, implying that both then vary between zero and one (technically speaking, they are made dimensionless). Figure 8 shows an example of normalized and smoothed frequency curves of preposition use in two mothers and their children.

Insert figure 8 about here

In particular the curves representing preposition use in Pauline and her mother, show a clear pattern of mutual adaptation. Before about the 20th observation, the mother has adapted her mean length of utterance to the mean length of utterance of the child (which is basically the period of one word sentences), shows a rapid support adaptation, followed by a downward adaptation, and, after observation 20, and adaptation that closely parallels the increase of MLU in the child. The data show a pattern of mutual fine-tuning that is typically adaptive in the way complex systems co-adapt, namely with temporary divergences that are later corrected. The data of Jan show a pattern that is considerably less adaptive.

The language data are typical of a "long-term" dynamics, occurring over a period of about 18 months, which is indeed long given the typical rate of early language development.

Figure 9 represents data on the levels of cognitive complexity in an adult's questions and the child's answers in the context of simple science experiments (Van der Steen et al., 2014).

Insert figure 9 about here

The data are typically microgenetic, spanning about 500 seconds, i.e. 8 to 9 minutes of adult-child cognitive interaction. By smoothing the data and normalizing them, we obtain a qualitative representation of the temporal unfolding of cognitive complexity during the course of a single adult child interaction session. The figure also covers a macrogenetic time scale by showing three consecutive sessions, with an interval of 3 months each. The short-term pattern of complexity follows a typical oscillating trajectory. Sometimes, the increase and decrease of complexity in the adult and the child are quite closely coordinated, and sometimes they diverge quite considerably. As to the long-term, a comparison of the three sessions shows that whereas during the first session, the adult is clearly leading the increase in cognitive complexity, the pattern is reversed in the third section, in which the child's changes in cognitive complexity are typically followed by corresponding changes in the questions of the adult. In short, what these data show is how cognitive complexity literally unfolds over the course of a single adult-child conversation, and how the dynamic relationship between adult and child is changing over the course of consecutive sessions.

Categorical methods and state space grid methodology

In the preceding examples, event data relating to the production of linguistic categories or levels of cognitive complexity, were transformed into ratio variables by smoothing the raw data. The resulting timeseries showed how the variables unfolded over time, and how the association between the timeseries changed during the interaction and across successive interaction sessions. However, it is also possible to describe the temporal dynamics of such a system by means of categorical methods. One method makes use of the state space grids already discussed in the preceding section. The following example (figure 10) is based on a study of music lessons, in which a child learns to play the violin from a teacher who follows the Suzuki method (Kupers et al., 2013). The state space grid representation of a single lesson shows the dynamic relationship between the students actions, which are either correct, partially correct, or incorrect, given the teacher's questions or assignments, and the teacher's reaction on the student's actions in the form of either giving the student a more complex assignments (forward), a less complex assignment (backward), an assignment on the same level of difficulty (repeat) or an assignment that is and related to the previous assignment (new task).

Insert figure 10 about here

This type of state space allows us to study processes of *scaffolding*. Adequate scaffolding implies that if a student makes an error, the teacher gives the student of assignment that is a bit less complicated than the current one, in order to bring the child back on the level of correct performance. If the student's action is partially correct, the teacher should repeat the current assignment, and if the student's performance is correct, the teacher should give the student a somewhat more complex assignment or proceed to a new task. The state space grid from figure 10 shows that most of the time this particular teacher-student system moves within the cells within or close to the range of adequate scaffolding.

An interesting feature of state space grid representations of a process, in this case a learning-teaching process unfolding during a music lesson, is that several *statistical indicators of dynamic properties* can be calculated. For instance, it is possible to calculate a measure of *dispersion* which corresponds with the distribution of the observed states over smaller or greater portions of the state space. Another measure refers to the *amount of variability*, namely the probability that the system stays within a particular cell of the state space or moves to another cell. The more cells are visited over the course of time, the less stability occurs in the process. Transitions occurring within the same cell or between adjacent cells (adjacent in terms of some underlying order principle, such as the principle of adequate scaffolding) provide information about the extent to which a particular system is captured in a particular attractor state, moves between attractor states, or is erratically dispersed across the whole range of possible states.

The values of these dynamic properties may vary across events, for instance over consecutive music lessons, and be used as indicators of long-term dynamics. For instance, in a study on the development of prepositional phrases, we analyzed the spontaneous utterances of children with regard to the occurrence of correct and incorrect prepositional constructions and the length (number of words) of the utterance in which a particular prepositional constructions occurred (Van Dijk & Van Geert, 2009). Figure 11 gives an example of the resulting state space grids for one of the children, Heleen (from a total of 25 observation sessions we selected six sessions equally distributed across the entire observation period.)

Insert figure 11 about here

Visual inspection of the grids shows that the dispersion of the blue points, corresponding with the system states, over the state space is first increasing and then decreasing, corresponding with a transition from an initial attractor of isolated prepositions in one word sentences to an attractor characterized by syntactically correct prepositional constructions in 4+ word sentences. Over six consecutive observations sessions, corresponding with a period from one year six months to two years six months, the level of the dispersion criterion shows a typical inverted U-shaped pattern, which was quite similar for the two children in this particular study (see figure 12).

Insert figure 12 about here

What we see here is the long-term evolution of a dynamic parameter of short-term processes, namely conversations between a child and the caretaker. The dynamic parameter refers to the amount of variability, i.e. the "chaoticity", in the linguistic patterns produced by a child across a particular conversation. That is, a particular linguistic features such as the production of prepositional constructions, may unfold over time either in a highly ordered fashion or quite erratically. The latter, erratic type of dynamic flow is typical of intermediary developmental states.

Categorical methods and state transition diagrams

In the state space grid, the transitions from one state to another are not always easy to follow. For this reason, researchers might use the format of state transition diagrams to represent the set of possible transitions from one state to another. The state transition diagram is a visualization of a matrix of conditional probabilities, which represents a so-called Markov chain. An example will show that this idea is actually very intuitive. Ensing et al. (2014) recently proposed a dynamic systems oriented definition of learning potential, as "the way a child is able to profit from the scaffolding situation that emerges within the dynamics between a particular teacher and this child at a particular moment" (page xx). To empirically describe scaffolding situations, the authors observed interactions between a child and a teacher in the context of the child's performing a particular task assigned by the teacher. A simple coding system was used, including every event in the teacher-child interaction during the task performance as either help eliciting behavior, help given by the teacher, correct performance by the child and feedback given by the teacher. By doing so, the interaction dynamics is reduced to four possible states and to the transition probabilities from one state to another, including the transition from a state to itself (e.g. if help eliciting behavior is followed by another instance of help eliciting behavior). Circles represent states, and arrows represent transitions from one state to another or from one state to itself. The size of the circles and arrows is (approximately) related to their frequencies. In principle, only the most frequent transitions are represented. Figure 13 provides an example of a state transition diagram of one child copying a star by means of triangles of various sizes that have to be glued on the right places.

Insert the figure 13 about here

The state transition diagram of dish child can then be compared with that of another child, doing the same task and with the same teacher, and statistically significant and educationally relevant differences between the transition diagrams can be calculated. The state transition diagrams represent important properties of the learning potential of the children in the context of particular tasks in particular teachers. For instance, in some children, help eliciting behavior is more often followed by help given by the teacher, but teachers can have a strategy of avoiding to respond to the child's help eliciting behavior, for instance if they think that the child asks for help too often and should be obliged to be more autonomous in his task performance activities.

More generally speaking, a state transition diagram — or a state space grid for that matter – is a directed graph. Directed graphs occur in many mathematical descriptions of processes, and can be conceived of as definitions of dynamic systems (see Abraham, 2014). They can be used to represent any network that produces a particular dynamics. For instance, the mother-child system concerning MLU can be represented by three nodes, namely the child's MLU in conversations with his mother, the mother’s MLU in her child-directed speech, and the mother’s MLU in her habitual, adult-directed speech. These tree nodes are connected by means of arrows, each of which represents how exactly a node affects the properties of another node. These relationships are expressed in the form of the mathematical equations that we presented before and that form the heart of the mother-child dynamic system as related to the use and development of prepositions.

Statistical methods for time series

Recently, a number of authors have made significant progress in finding techniques for the estimation of dynamic parameters from time series. One such technique is state space embedding or time delay embedding, which implies that parameters are estimated across multiple time scales (Ferrer & Zhang, 2009; Ferrer et al., 2007; Von Oertzen, T., & Boker, S., 2010; Deboeck, P., Montpetit, M., Bergeman, C., & Boker, S. , 2009).

A technique that is of particular relevance to dynamic model fitting is nonstationary time series modeling, which has been applied to the analysis of individual developmental processes by Molenaar et al. (2009) (see also Molenaar's chapter in the present volume). An important tool is the Kalman filter, which is used to estimate the value of a changing, i.e. dynamic, latent variable on the basis of multivariate, time varying observations , based on a particular dynamic rule or evolution term (See Chow, Ferrer and Nesselroade, 2007; Molenaar, Sinclair, Rovine, Ram, & Corneal, 2009). An example of application of this approach is the Molenaar et al. (2009) study on changes in the structure of emotional experiences of adolescents as they communicated with their (step)fathers.

A technique that is likely to become increasingly important is cross-recurrence quantification. In a highly variable timeseries — and most developmental and educational timeseries are indeed highly variable — many of the observed values, such as a specific MLU occurring in during a particular observation session, will tend to come back at some later observation. That is, at some later observation we will measure of value that is approximately similar to value that we have measured before. Put differently, values, but also combinations of values, will tend to recur in the timeseries. There is empirical and mathematical proof for the fact that the pattern of recurrence can tell us something about the nature of the underlying dynamic system, for instance the extent to which it is rigid or flexible, the extent to which it is generated by a system with more or less internal coordination and so forth. So far, the technique has been applied to a relatively limited set of developmental data, such as language development (Cox and Van Dijk, 2013; Dale and Spivey, 200x); motor development (reference), parent-child interactions in aggressive children (Lichtwarck Aschoff et al., 2012), mother infant interaction and infant sleep (de Graag et al., 2012). The aim of the technique is to look for evidence of internal structure and complexity of the system that generates the behavior. The assumption is that the system generating human behavior are so complex that it is virtually impossible to estimate the influence of a single parameter on the behavior of the system, simply because all parameters are interconnected. The determination of the effect of single variables on the basis of associations occurring in samples consisting of many individuals tells us little to nothing about the way such variables are intertwined individual dynamic systems (which is a direct consequence of the principle of non-entropy in individual developmental and educational processes).

In all these cases we assume that there is some explicit dynamic model that we wish to fit to the data, the parameters of which we wish to estimate. These statistical timeseries techniques are of particular importance in those cases where the parameters of the model change over developmental time, or where the parameters depend on intervening contextual factors, such as in dyadic communication.

Further details concerning methods of model fitting can be found in recently published specialized handbooks, such as Boker and Wenger's "*Data analytic techniques for dynamical systems*" (2007), Molenaar and Newell's "*Individual pathways of change: Statistical models for analyzing learning and development"* (2010), and Valsiner, Molenaar, Lyra and Chaudhary's *Dynamic process methodology in the social and developmental sciences" (2009).*

Conclusion

The theory of complex, nonlinear dynamical systems provides a modern, theoretical and empirical framework for the notion of development and education as forms of unfolding potential. The notion of unfolding directly refers to the processes of change explained by the underlying mechanisms of change, operating on the products of change, namely the actual states of a particular educational or developmental system. A dynamic system explains how every next step in the process is a direct result of the preceding step or steps, and by doing so provides a formal description of what "unfolding" actually entails. In the classical developmental metaphor, the content to be unfolded was already there in the beginning, it was something that had to be unwrapped. The theory of complex, nonlinear dynamic systems, however provides a general framework for understanding and explaining of what is in fact the opposite of unwrapping, or literal unfolding, namely emergence. The notion of emergence allows us to get a better grip on the notion of potential. A system's potential is the set of possible developmental trajectories and possible processes of emergence that the system might show as it follows its inherent principles or mechanisms of change.

In addition to providing a theoretical ground for the notions of unfolding and potential, the theory of complex, nonlinear dynamical systems also provides new possibilities for model building and for empirical research. Model building focuses on attempts to explain the principles of change by means of mathematical models in the form of differential and the difference equations. Empirical research focuses on the study of processes in systems, and on how and where such processes actually occur, which is in individual systems, such as individual children, mother-infant dyads, classes with their teacher and so forth. Generalization is aimed at finding general principles of change in general dynamic models, and that can be easily individualized by individualizing the parameter values, while keeping the general underlying change model intact.

The theory of complex, nonlinear dynamic systems is not so much a new theory, as a new paradigm, in that it starts from a different view on the nature of reality, and in particular to reality of educational and developmental processes. These new starting point is the assumption that the systems that the life sciences are studying are, by default, nonlinear in their behavior, are showing self-organization and emergence and are based on models of mutual, coordinated causal interactions between the many components that constitute any living system or subsystem. This new paradigm is indeed still in its infancy as far as the educational and developmental sciences are concerned. However, an increasing number of researchers is adopting this new approach, and new and exciting methods, models and findings are merging that have the potential to change our view on how education and development should be studied, described and explained.

References (incomplete)

Amini, M. (**2011**) **Fodor** and the Impossibility of Learning, in Just the Arguments:
100 of the Most Important Arguments in Western Philosophy (eds M. Bruce and S. Barbone),
Wiley-Blackwell, Oxford, UK

Amini, M. Fodor and the Impossibility of Learning. *Just the Arguments: 100 of the Most Important Arguments in Western Philosophy*, 359-361.

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*(5), 588-612.

Bassano, D., Korecky-Kröll, K., Maillochon, I., van Dijk, M., Laaha, S., van Geert, P., & Dressler, W. U. (2013). Prosody and animacy in the development of noun determiner use: A cross-linguistic approach. *First Language*, *33*(5), 476-503.

Bassano, D., Maillochon, I., Korecky-Kröll, K., Van Dijk, M., Laaha, S., Dressler, W. U., & Van Geert, P. (2011). A comparative and dynamic approach to the development of determiner use in three children acquiring different languages. *First Language*, *31*(3), 253-279.

Bassett, D. S., & Bullmore, E. D. (2006). Small-world brain networks. *The neuroscientist*, *12*(6), 512-523.

Bullmore, E. T., & Bassett, D. S. (2011). Brain graphs: graphical models of the human brain connectome. *Annual review of clinical psychology*, *7*, 113-140.

Bullmore, E., & Sporns, O. (2009). Complex brain networks: graph theoretical analysis of structural and functional systems. *Nature Reviews Neuroscience*,*10*(3), 186-198.

Chomsky, N. (1959). A review of BF Skinner's Verbal Behavior. *Language*,*35*(1), 26-58.

de Graag, J. A., Cox, R. F., Hasselman, F., Jansen, J., & de Weerth, C. (2012). Functioning within a relationship: Mother–infant synchrony and infant sleep. *Infant Behavior and Development*, *35*(2), 252-263.

Fischer, K. W., & Bidell, T. R. (2006). Dynamic development of action and thought. *Handbook of child psychology*./

Fodor , Jerry . “ On the Impossibility of Acquiring ‘ More Powerful ’ Structures , ” in Language and Learning: The Debate between Jean Piaget and Noam Chomsky , edited by Massimo Piattelli - Palmarini, 142 – 62 . London : Routledge & Kegan Paul , 1980

Haken, H. (2006). Synergetics of brain function. *International journal of psychophysiology: official journal of the International Organization of Psychophysiology*, *60*(2), 110–24.

Kauffman, S. A., 1993. *The origins of order*. Oxford University Press.

Kauffman, S. A., 1995. *At home in the universe*. Oxford University Press.

Lee, A. S., & Baskerville, R. L. (2003). Generalizing Generalizability in Information Systems Research. INFORMATION SYSTEMS RESEARCH, 14(3), 221–243. Retrieved from generalizing

Lichtwarck-Aschoff, A., Hasselman, F., Cox, R., Pepler, D., & Granic, I. (2012). A characteristic destabilization profile in parent-child interactions associated with treatment efficacy for aggressive children. *Nonlinear Dynamics-Psychology and Life Sciences*, *16*(3), 353.

Newman, M. E. J. (2003). The Structure and Function of Complex Networks. *SIAM Review*, *45*(2), 167. doi:10.1137/S003614450342480

Newman, M. E. J. (2003). The structure and function of complex networks. *SIAM Review, 45,* 167-256.

Newman, M. E. J. (2010). *Networks: an introduction*. Oxford: Oxford University Press.

Rubinov, M., & Sporns, O. (2010). Complex network measures of brain connectivity: uses and interpretations. *Neuroimage*, *52*(3), 1059-1069.

Sporns, O. (2011). The human connectome: a complex network. *Annals of the New York Academy of Sciences*, *1224*(1), 109-125.

Van der Steen, S., Steenbeek, H., & Van Geert, P. (2012). Using the dynamics of a person context system to describe children's understanding of air pressure.*Current Topics in Children's Learning and Cognition*, 21-44. /Kloos, H., Morris, B. J., & Amaral, J. L. (2012). Current Topics in Children's Learning and Cognition.

Van Dijk, M., Van Geert, P., Korecky-Kroell, K., Maillochon, I., Laaha, S., Dressler, W. U., & Bassano, D. (2013). Adaptation between child-directed speech and early child language: A developmental perspective. *Language Learning*, *63*, 214-270.

van Geert, P. (1998). A dynamic systems model of basic developmental mechanisms: Piaget, Vygotsky, and beyond. *Psychological review*, *105*(4), 634.

Van Geert, P. (2003). Dynamic systems approaches and modeling of developmental processes. In J. Valsiner and K. J. Conolly (Eds.), *Handbook of developmental Psychology*. London: Sage. Pp. 640-672.

van Geert, P. (2009). Transmission, self-organization, and the emergence of language: A dynamic systems point of view. *Cultural transmission: Psychological, developmental, social, and methodological aspects*, 48-69./ Schönpflug, U. E. (2009). *Cultural transmission: Psychological, developmental, social, and methodological aspects*. Cambridge University Press.

Van Geert, P., & Fischer, K. W. (2009). Dynamic systems and the quest for individual-based models of change and development. *Toward a unified theory of development: Connectionism and dynamic systems theory re-considered*, 313-336.

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*(3), 408-442.

Figures

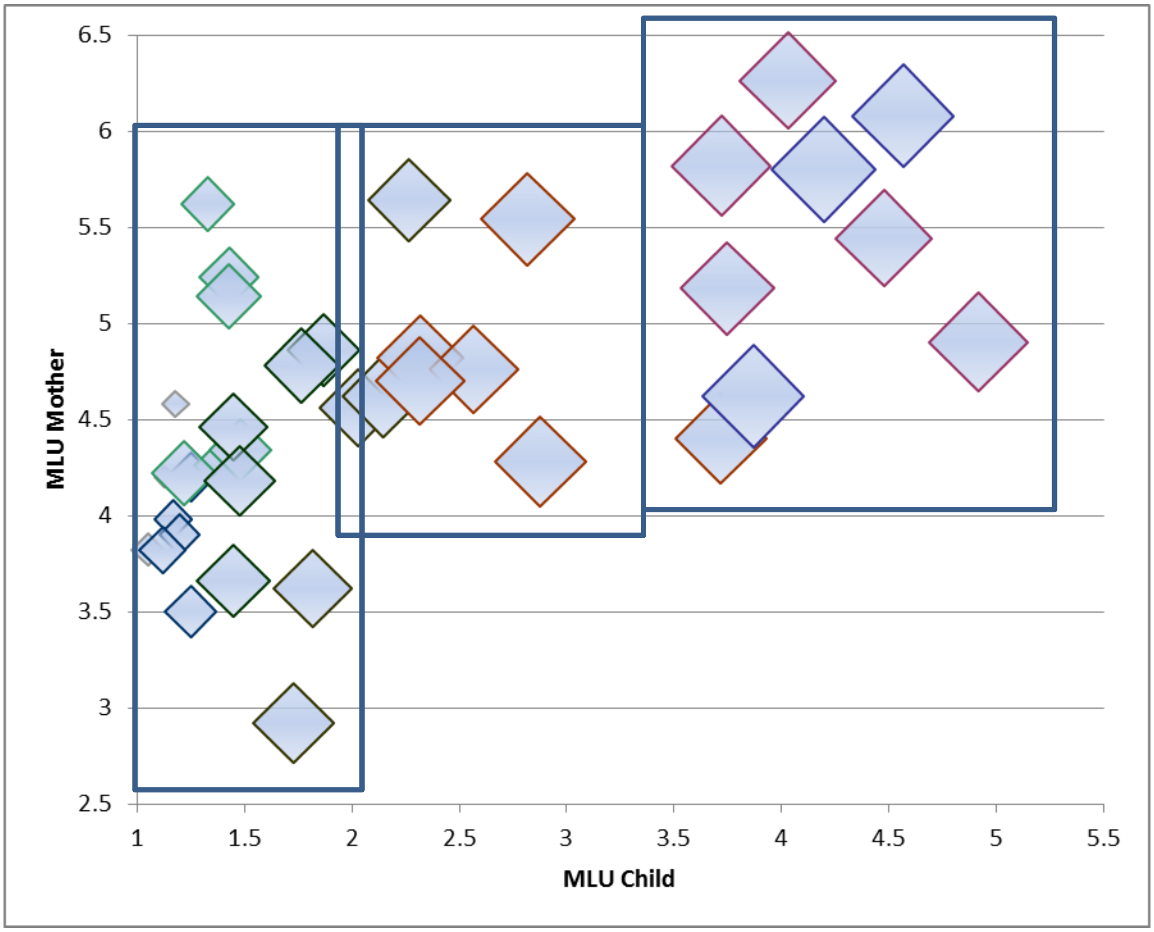

Figure 1 State space representation of MLU of mother and child; size of markers is approximately representative of time of occurrence (greater markers refer to later observations); the pattern can be roughly divided into three clusters or “stages”

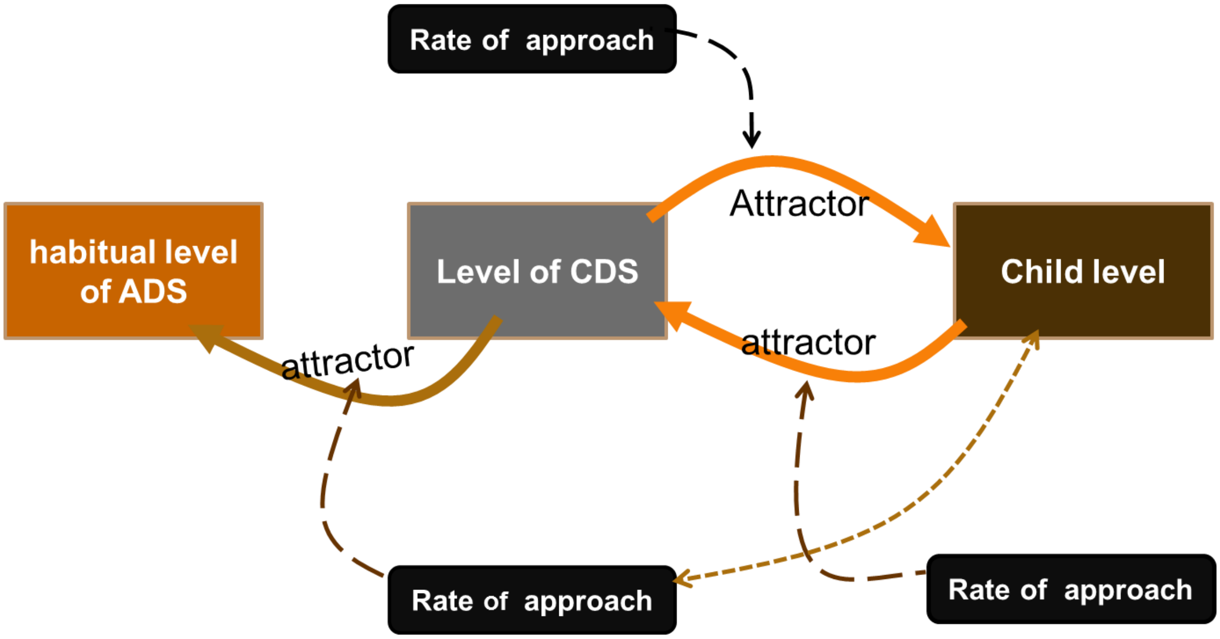

Figure 2: Graphical structure of the dynamic model of mother-child language interaction

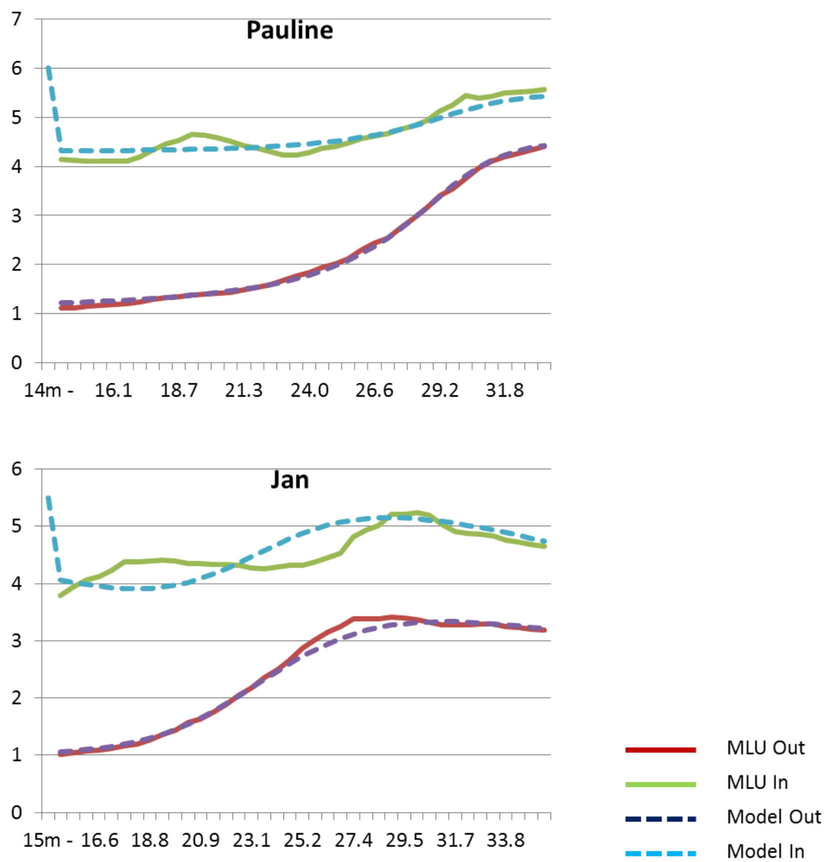

Figure 3: Smoothed data curves and curves generated by the dynamic model based on three coupled difference equations.

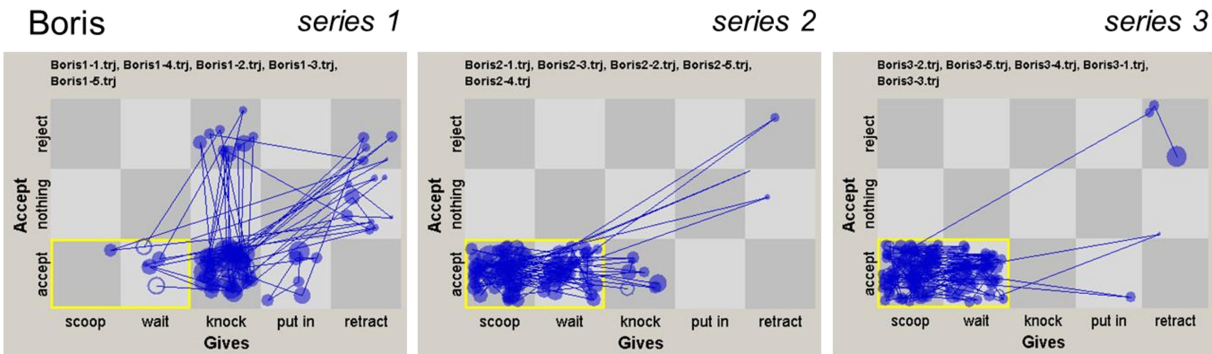

Figure 4: State space representations of mother-child spoon feeding sessions. The three sessions, taken from different stages in the development of spoon feeding and eating, show a pattern of increasing order and emergence of an attractor state (scoop-wait-accept).

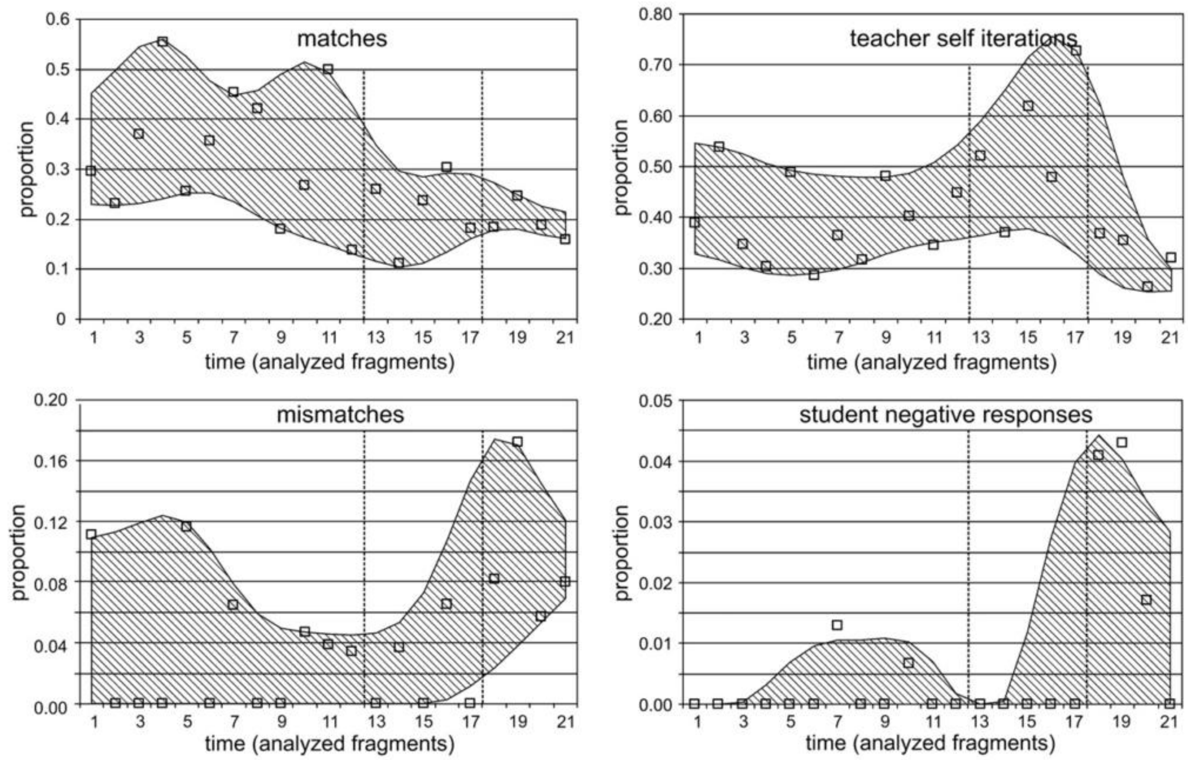

Figure 5: nonlinear changes in child and teacher variables in individual math lessons, over a period of two years. Dotted, vertical lines represent two change points, where the pattern of the variables change into new, temporary states.

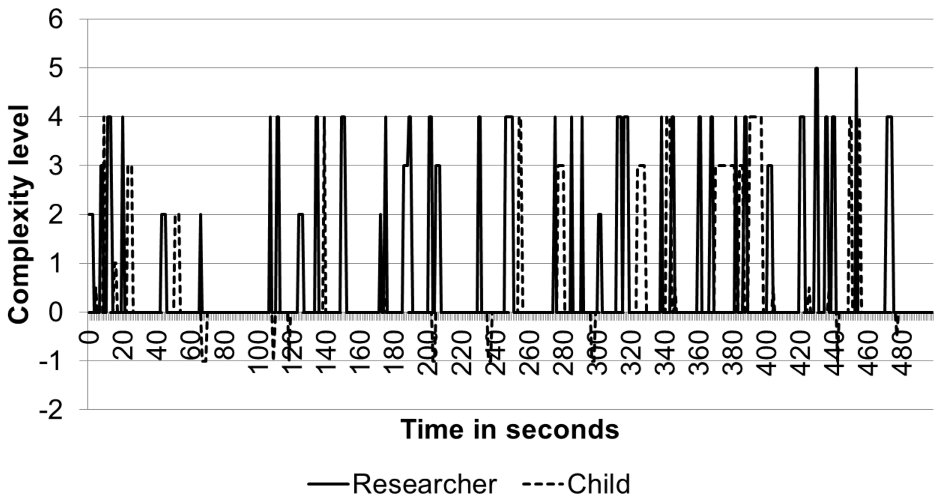

Figure 6: Time-serial illustration of the complexity levels measured in a researcher's questions and a boy's answers during session 1. Utterances classified as incorrect are depicted as −1, and right answers to close-ended questions are marked as 0.5.

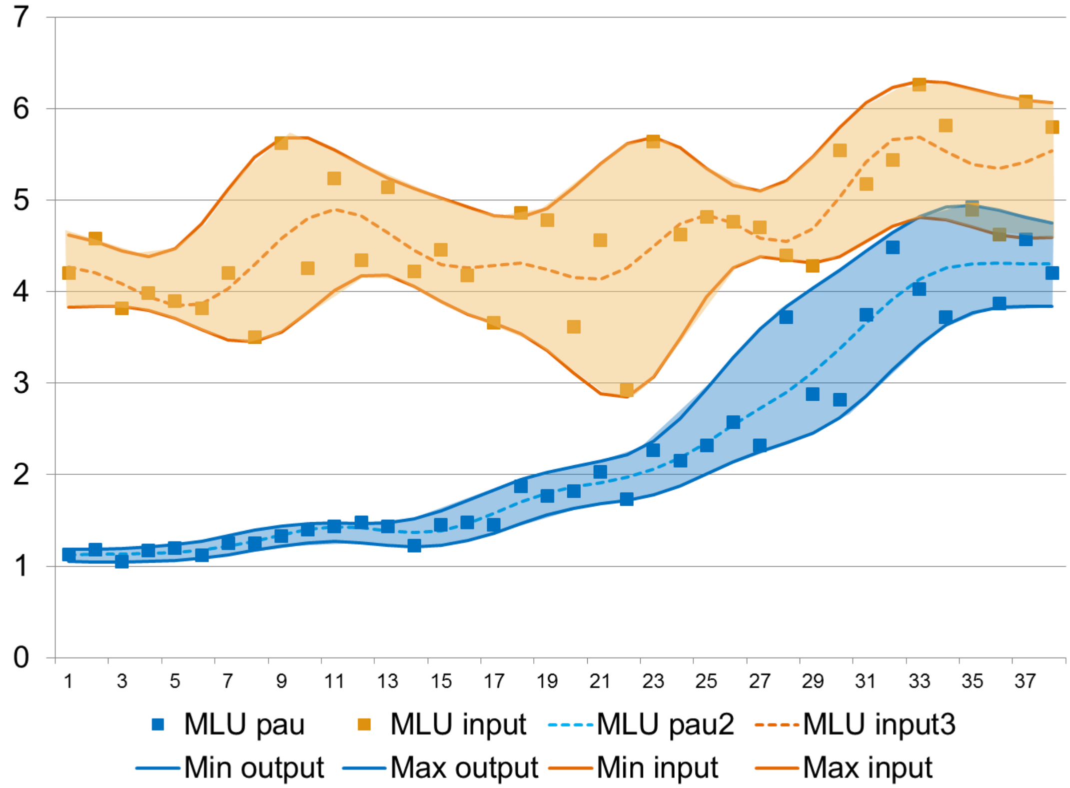

Figure 7: Raw data of MLU in a mother and her child, with smoothed curves (central lines) and smoothed curves representing the bandwidth of variability.

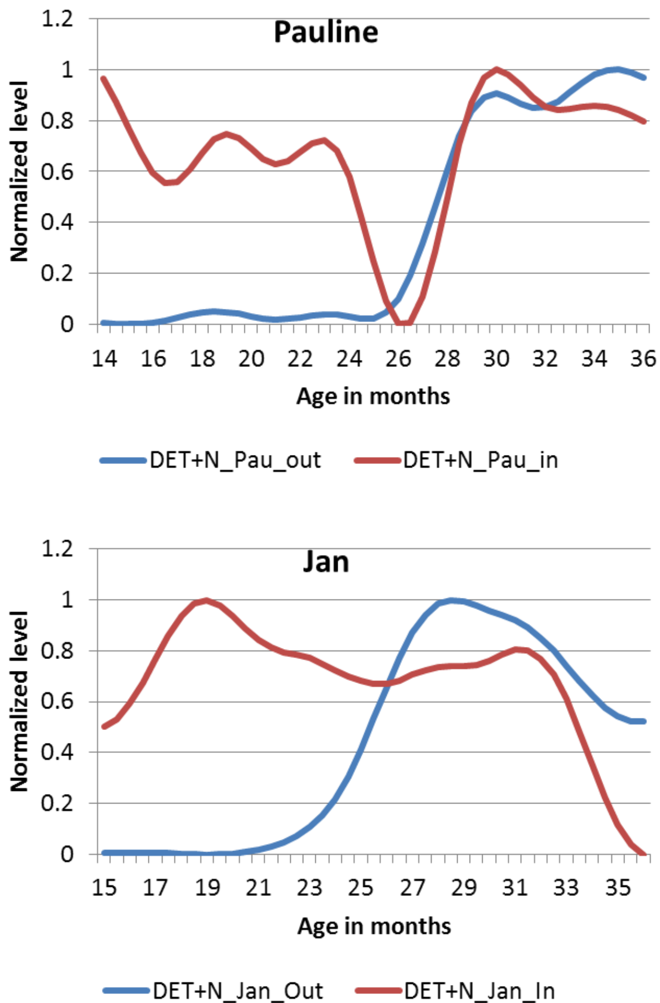

Figure 8 Smoothed and normalized curves of determiner use in two mothers and two children (Pauline and Jan). Patterns of co-adaptation are very different for these two children.

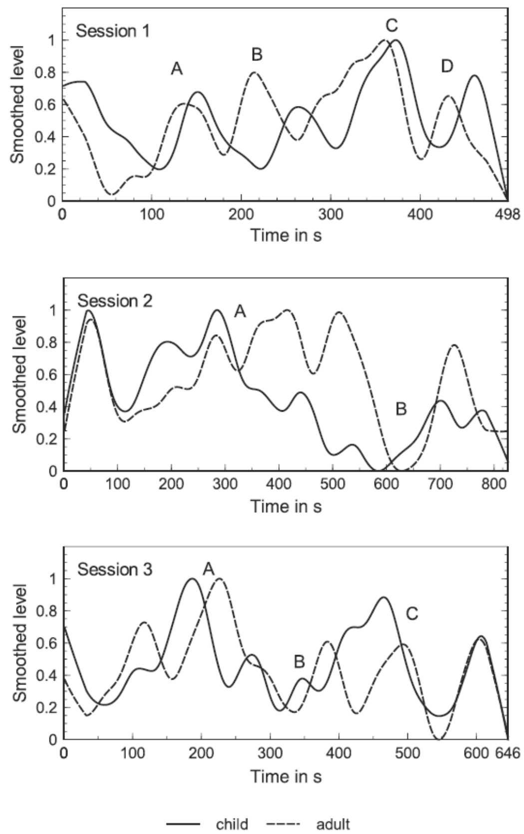

Figure 9. Normalized Loess curves of the complexity levels measured in the boy's answers (black line) and the researcher's questions (dashed line) of 3 sessions

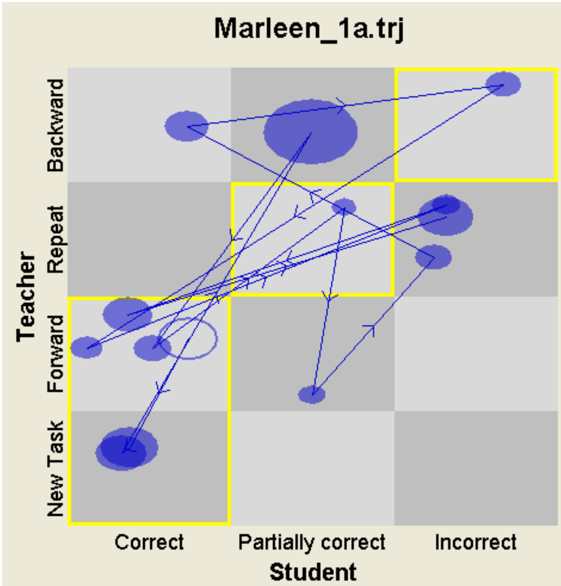

Figure 10. Example of a State Space Grid of one lesson. The states with a yellow line represent contingent scaffolding.

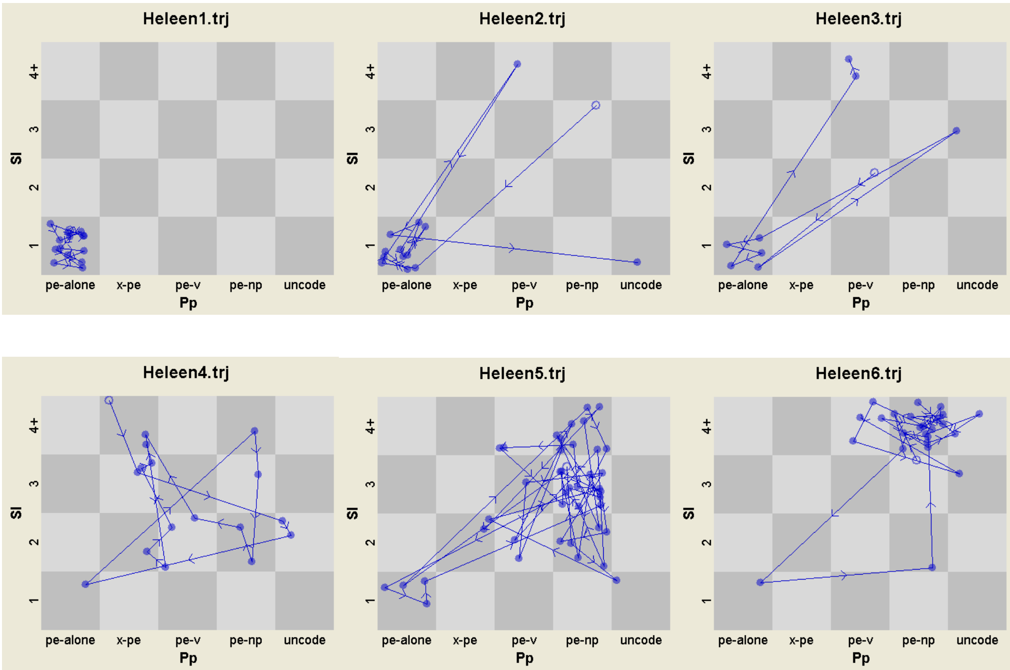

Figure 11. State Space Grids of the strategy of the prepositional phrase (horizontally) and sentence length (vertically) for the six sessions of Heleen.

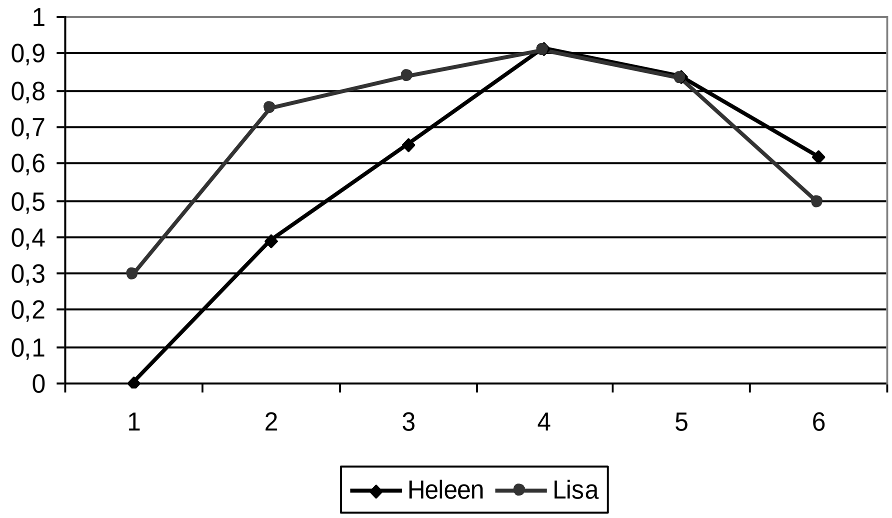

Figure 12. Dispersion across the grid of the states of MLU-preposition constructions in Heleen, compared with another child, Lisa.

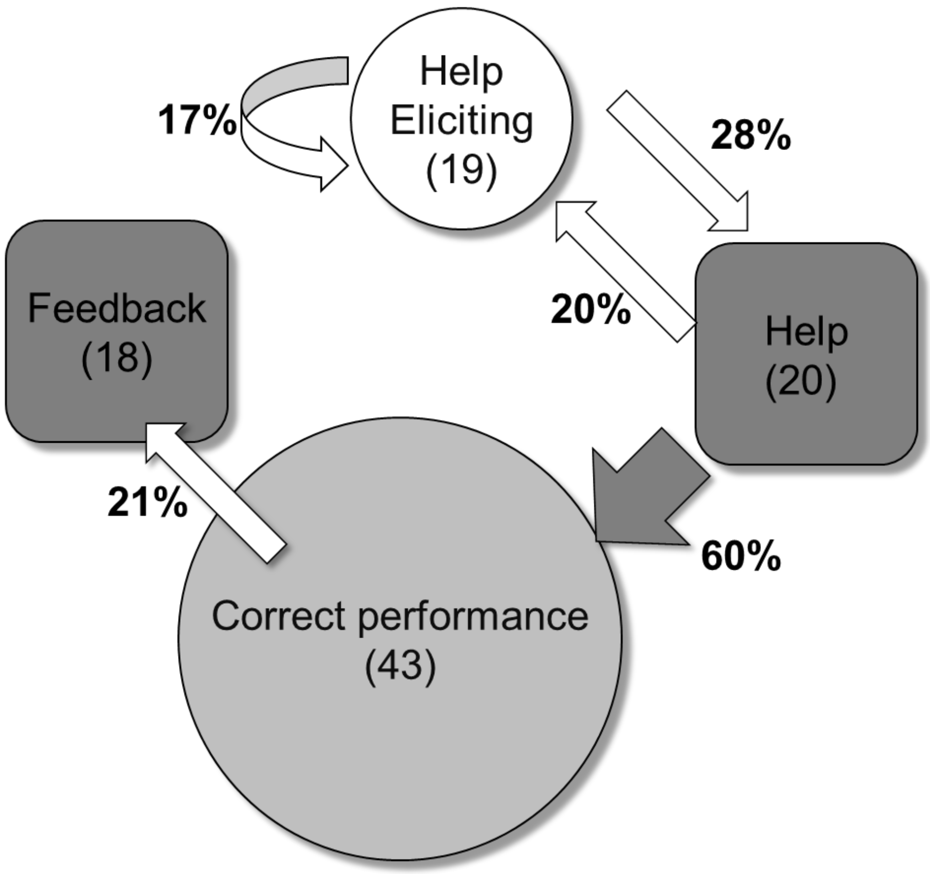

Figure 13: A state transition diagram of help eliciting behavior and correct performance of a student, and the associated help and feedback given by the teacher.