Bilingualism: Language and Cognition 10 (1), 2007, 7–21 © 2007 Cambridge University Press doi:10.1017/S1366728906002732

KEYNOTE ARTICLE

A Dynamic Systems Theory approach to second language acquisition*

KEES DE BOT
WANDER LOWIE
MARJOLIJN VERSPOOR
University of Groningen

In this article it is argued that language can be seen as a dynamic system, i.e. a set of variables that interact over time, and that language development can be seen as a dynamic process. Language development shows some of the core characteristics of dynamic systems: sensitive dependence on initial conditions, complete interconnectedness of subsystems, the emergence of attractor states in development over time and variation both in and among individuals. The application of tools and instruments developed for the study of dynamic systems in other disciplines calls for different approaches to research, which allow for the inclusion of both the social and the cognitive, and the interaction between systems. There is also a need for dense data bases on first and second language development to enhance our understanding of the fine-grained patterns of change over time. Dynamic Systems Theory is proposed as a candidate for an overall theory of language development.

Introduction

A major assumption underlying a great deal of L1 acquisition research has been that the acquisition of a language has a clear beginning and end state, and a somewhat linear path of development for each individual. Similarly, in much SLA research, an L2 learner, no matter what his/her L1, is predicted to go through highly similar stages in acquiring the L2. Such a view of language learning or processing is often associated with an INFORMATION PROCESSING (IP) model.

On the other hand, there have also been numerous linguistic and language acquisition studies that have not adhered to the linear view. They have shown that language, language acquisition, and language attrition are much more intricate, complex, and even unpredictable than a linear position would allow. Linguistic theories such as COGNITIVE LINGUISTICS and FUNCTIONAL LINGUISTICS, acquisition theories such as EMERGENTISM, and processing theories such as the COMPETITION MODEL recognize that there are many interdependent variables, not only within the language system, but also within the social environment and the psychological make-up of an individual. What these theories have in common is that they recognize the crucial role of interaction of a multitude of variables at different levels: in communication, in constructing meaning, in learning a language and among the languages in the multilingual mind. However, even though practitioners of these non-linear approaches seem to recognize some overlap and compatibility between the different theories, many of such theories still stand apart for lack of one overarching theory that allows to account for these ever interacting variables, non-linear behaviour, and sometimes unpredictable outcomes, a theory that does not regard real-life messy facts as "noise" but as part of the "sound" you get in real life.

We feel that DYNAMIC SYSTEMS THEORY (DST), even though admittedly with some unresolved issues, could be such a theory. The aim of this article is to explain how DST has developed, what some of its main characteristics are, how it has been applied to human and non-human communication, and how several common SLA features could be reinterpreted from a DST perspective. It is our claim that because DST takes into account both cognitive and social aspects of language development, it can provide a coherent approach to various issues in SLA.

Dynamic Systems Theory and its applications

The literature on the application of DST in SLA is still fairly limited. After the pioneering work by Larsen-Freeman in 1997, it remained silent for five years, until Herdina and Jessner published their book *A dynamic model of multilingualism* (2002) and Larsen-Freeman added to her earlier work in 2002. Inspired by this work and by Paul Van Geert's work on L1 acquisition, we have developed an interest in this topic, which has led to a number of publications (Verspoor, De Bot and Lowie,

* The authors are grateful to Paul Van Geert for his support, inspiration, and comments and for allowing us to quote excessively from his work, and to Diane Larsen-Freeman and Heike Behrens for comments on an earlier version. They are also grateful to three anonymous reviewers for their useful comments and suggestions and to Sybrine Bultena for her help with finalizing the manuscript.

Address for correspondence:
Prof. Dr. Kees de Bot, University of Groningen, Faculty of Arts, Department of Applied Linguistics, PO Box 716, 9700 AS Groningen, the Netherlands
E-mail: *c.l.j.de.bot@rug.nl*

https://doi.org/10.1017/S1366728906002732 Published online by Cambridge University Press

2004; De Bot, Lowie and Verspoor, 2005a, b; De Bot and Makoni, 2005). For those unfamiliar with the theory, we begin with a brief description of the theory and some examples of how it may apply to SLA.

Basic aspects of DST

DST, which developed as a branch of mathematics, is originally about very simple systems such as the two coupled variables in a double pendulum. Even though such a system has only two interacting variables or degrees of freedom, the trajectory of the system is complex.[1] When applied to a system that is by definition complex, such as a society or a human being, where innumerable variables may have degrees of freedom, DST becomes the science of complex systems. The major property of a DS is its change over time, which is expressed in the fundamental equation $x_{(t+1)}=f(x_{(t)})$, for any function describing how a state x at t is transformed into a new state x at time $t+1$.

While calculations may be the core of DST within the field of mathematics, they are not needed to grasp the general principles behind a dynamic system (Thelen and Smith, 1994; Van Gelder, 1998; Shanker and King, 2002). Complex systems such as a learning person are sets of interacting variables. Dynamic systems are characterized by what is called COMPLETE INTERCONNECTEDNESS: all variables are interrelated, and therefore changes in one variable will have an impact on all other variables that are part of the system. In many complex systems, the outcome of development over time can therefore not be calculated exactly; not because we lack the right tools to measure it, but because the variables that interact keep changing over time and the outcome of these interactions, unless they take place in a very simple system, cannot be solved analytically. To follow a dynamic trajectory, the system has to be simulated by doing the iterations, for there is no equation that will directly give a value of the system at some later time.

Dynamic systems are nested in the sense that every system is always part of another system, going from sub-molecular particles to the universe, with the same dynamic principles operating at all levels. As they develop over time, dynamic sub-systems appear to settle in specific states, so-called ATTRACTOR STATES, which are preferred but not necessarily predictable. Examples of attractor states are the two different ways horses may run: they either trot or gallop, but apparently there is no in-between way of running. States that are clearly not preferred are so-called REPELLER STATES. Attractors can be simple or complex and for some systems chaos can be the attractor state. Attractor states are by definition temporary and not fixed, but depending on the strength of the attraction, more or less energy is needed to make the system move on to another attractor state. The notions of development and attractor states are somewhat analogous to a ball rolling over a surface with holes and bumps, with the ball's trajectory as development, the holes as attractor states and the bumps as repeller states. The holes can be shallow or deep, and the deeper the hole is, the more energy is needed to get the ball out of the hole and make it move on to the next hole.[2]

Because the development of some dynamic systems appears to be highly dependent on their initial state, minor differences at the beginning may have dramatic consequences in the long run. This is called THE BUTTERFLY EFFECT, a term proposed by the meteorologist Lorenz, who wanted to account for the huge impact small local effects may have on global weather. Related to this is the notion of non-linearity, which means that there is a non-linear relation between the size of an initial perturbation of a system and the effects it may have in the long run. Some minor changes may lead to huge effects, while major perturbations may be absorbed by the system without much change. The sensitivity to initial conditions may depend on one or more critical parameters. Very similar systems may be variably sensitive to initial conditions, which tend to become especially relevant when the system is in a chaotic state.

Regardless of their initial states, systems are constantly changing. They develop through interaction with their environment and through internal self-reorganisation. Because systems are constantly in flow, they will show variation, which makes them sensitive to specific input at a given point in time and some other input at another point in time. In natural systems, development is dependent on resources: while the frictionless double swing presented earlier will make its tracks till eternity, all natural systems will tend to come to a still stand when no additional energy is added to the system.

DST, UG, and creativity in language use

Much of the debate on the application of DST in cognitive processing and language acquisition has taken place in the journal *Behavioral and Brain Sciences*. Two lead articles, one by van Gelder (1998) and one by Shanker and King (2002), argue for the application of DST in cognitive science and ape-language research, respectively. The peer comments on these two articles provide a rich picture of how the emergence of the DST paradigm has been received. In this section, we will discuss some of the issues raised that are particularly pertinent to the role of DST in SLA research, namely to what extent DST is different

[1] See <http://www.maths.tcd.ie/~plynch/SwingingSpring/double-pendulum.html> for an illuminating illustration.

[2] Even though this is only a metaphor in our case, in many cases it may be more than just a metaphor in that the proportions of that landscape, e.g. how deep the hole is, can be calculated.

https://doi.org/10.1017/S1366728906002732 Published online by Cambridge University Press

from an information processing view of language, how a DST approach can account for creativity in language use, and how language acquisition may be seen as a reflexive activity.

Fax machine or dance?

One of the main arguments made by Shanker and King is that a DST approach differs fundamentally from the information processing approach that has dominated the cognitive sciences and accordingly psycholinguistic aspects of bilingualism and SLA in the last decades. To contrast the two approaches, Shanker and King use two metaphors for the process of communication: fax machines and a dance.

The information processing approach is like two fax machines exchanging information: when one is sending, the other is receiving. The information to be communicated is coded in a message, which is then sent and decoded by the receiver. In verbal communication, speech is the code used to transfer information between a sender and a receiver. So both sender and receiver encode and decode information, using the same coding system to encode and decode the message. "The transmission metaphor treats communication as a *sequential* process in which partners take turns emitting and processing one another's messages" (Shanker and King, 2002, p. 605).

In contrast, the DST approach views communication as a dance. The metaphor originates from Sue Savage-Rumbaugh's work on the linguistic abilities of great apes, where she observes that language comprehension is based on inter-individual routines that are like "a delicate dance with many different scores, the selection of which is being constantly negotiated while the dance is in progress, rather than in advance" (Savage-Rumbaugh et al., 1993, p. 27). The dance metaphor is particularly well chosen to explain some of the basics of the DST approach. Based on very simple procedures (steps) carried out in coordinated fashion in dyads, complex patterns emerge from the interaction between the two dancers, and even increasingly more complex and unpredictable patterns will emerge over time when one pair of dancers interacts with other pairs on the dance floor.

The interaction in dyads is multimodal: voice, rhythm and facial expressions interact to create mutual understanding and agreement on steps to take. There is constant adaptation and change, but it is often unclear which partner is the initiator of change. Perfect dancers show what in developmental studies has become known as INTERACTIONAL SYNCHRONY, the seamless understanding between partners that are mutually attuned in the interactional process. Like in all other forms of communication, dancers go through waves of synchrony and asynchrony, and they are constantly adapting to repair asynchrony. But as Shanker and King point out, interactional synchrony is quite different from co-regulation in that in the latter the partners are not just in sync but also make meaning by acting together. "The dance metaphor focuses on *co-regulated* interactions and the emergence of creative communicative behaviors within that context" (2002, p. 605).

The comments in *BBS* on the Shanker and King 2002 article are reminiscent of a 1998 article by Van Gelder in the same journal, in which he argues strongly for the application of DST to describe and explain cognitive processing and in which he provides numerous examples of research that have applied principles of DST on a range of cognitive processes. In their reactions to the Shanker and King article and the one by Van Gelder, several commentators have argued against the idea that DST and an information processing view are incompatible, but the arguments are not convincing. Whereas Zentall (2002) argues that in an information processing approach there can be multiple channels to transfer information, which undermines the argument that not enough information or no information of different kinds can be transmitted, Waters and Wilcox (2002) point out that the main failure of the IP model is the assumption that words seen as the main carriers of meaning also contain that meaning. Conversely, in their view, meaning is constructed: "Signals, words, gestures and expressions do not *mean*: they are prompts for the construction of meaning" (Waters and Wilcox, 2002, p. 644). Giving meaning is not achieved through the transmission of bits of information, but through co-construction. Along similar lines, Westbury uses the term SILICONCENTRIC to typify the current computer-based metaphors of the IP paradigm. "The paradigm tends to judge perceptually mediated information about objects as more important or even more real than socially mediated information about conspecifics" (Westbury, 2002, p. 645), but he also argues that different types of information can be mediated in different ways, leaving room for both the fax machine and the dance metaphor.

Thompson and Valsiner (2002), who strongly support the use of DST and the dance metaphor, stress the role of agency, an issue that seems to be underdeveloped in discussions on language and cognition as dynamic systems. They argue as follows:

> To call a social interaction a dance is to stress the peraction of social agents. When agents peract, they act through or by means of one another. Each has a state of affairs towards which his or her behavior is directed, and that state of affairs requires certain actions on the part of the social partner. The behavior of each actor is therefore directed toward using the other as a tool to produce a particular desirable result. (Thompson and Valsiner, 2002, p. 641)

We may conclude that whereas the Information Processing model looks at communication as a linear, binary sequence of events, the dynamic systems model

https://doi.org/10.1017/S1366728906002732 Published online by Cambridge University Press

looks at the relation between behaviours and how the whole configuration changes over time. The information processing approach is linked to terms like signal and response, sending and receiving, encoding and decoding, and rule-governed behaviour, while the DST approach is based on terms like engagement and disengagement, synchrony and discord, breakdown and repair in interaction, and the properties that emerge from it.

Creativity in language

The IP model is often associated with a UG approach to language, which set itself off against behaviourism by assuming that creativity in language use cannot be accounted for without some innate mechanisms particular to language learning. Shanker and King formulate this contrast as follows: "Whereas the information-processing paradigm sees creativity as a property of the language system itself, . . . dynamic systems theory views creativity as a property of agents' behavior in co-regulated interactions" (2002, p. 608). However, several attempts have been made to explore the compatibility between DST and a UG approach to language. An early contribution to the application of DST in language is Mohanan (1992), who relates a dynamic perspective to UG by viewing universal principles as FIELDS OF ATTRACTION and argues that a DST perspective can explain the emergence of complexity in phonological development. Cooper's (1999) view is in line with Mohanan's work in that he sees universal aspects as attractors, resulting from random processes rather than constraining development or change. He introduces DST notions in his study of diachronic change in English, showing how changes at the individual level lead to changes in the system of the English language. His main conclusion is that it is possible to set up an "attractor grammar" with grammar rules seen as the BASINS OF ATTRACTION. A more traditional UG-based approach is used in a series of articles by Nowak and his colleagues (Nowak, Komarova and Niyogi, 2001, 2002), in which the necessity of UG is supported with evidence from formal language theory, learning theory, and evolution. Their computational treatment of the evolution of UG has strong dynamic components. Using their model, they account for the fact that individuals are able to select the right language by assuming a dynamic interaction between, among other factors, the communicative payoff of using a language structure, the fitness of a particular language and a learning algorithm. Referring to deterministic population dynamics, they argue that a succession of UG's has evolved from a system of early animal communication to the UG of human beings today.

Even though UG with universal principles seen as innate properties constraining the hypothesis space in acquisition is not by definition rejected in a DST approach, a DST approach does not require innate linguistic properties as a necessary condition for language acquisition because in DST complexity and therefore creativity emerges from the iterations. As Smith, Kirby and Brighton (2003) argue, language acquisition is probably not bias-free, but the bias is not necessarily specific for language.

Larsen-Freeman (2002), who initially suggests that UG may be seen as part of the initial condition of the developing dynamic language system, agrees that the two perspectives are complementary and could exist side by side with their own research traditions and communities. However, she argues for the application of DST to accommodate both social and cognitive approaches to SLA because in DST development is seen as a process that takes place through interaction between the individual and her environment, and she leaves little doubt about her own view on nativism when she refers extensively to Hopper's (1998) emergentist views:

> Grammar is regarded as epiphenomenal, a by-product of a communication process. It is not a collection of rules and target forms to be acquired by language learners. Language, or grammar, is not about having; it is about doing: participating in social experiences. (Larsen-Freeman, 2002, p. 42)

In a recent plenary address, Bybee sharply defined her position along similar lines through her title: "The impact of use on representation: Grammar is usage and usage is grammar".[3]

Also leading DST researchers, such as Thelen and Smith (1994), leave little room for nativist ideas on language acquisition. Shanker and King's view of language is even less traditional:

> We are not concerned here with what might or might not have gone on 'inside Kanzi's head' that enabled him to develop language skills, nor is language viewed as a combinatorial system whose 'structure' he had to 'grasp'. Rather language is viewed as a particular type of reflexive activity in which Kanzi was enculturated. (Shanker and King, 2002, p. 619)

"Reflexive" is used in the sense that language is not an abstract autonomous entity itself, but is used to communicate in a real world and addresses real wants and needs of participants.

Language as a reflexive activity

Empirical support for such a reflexive activity with a dynamic and interactionist point of view comes from Gogate, Walker-Andrews and Bahrick (2001), who present an overview of research on the interaction between information provided by adults in synchronous bimodal (visual–auditory) presentation of objects and their names, and the infant's reaction to that. There is clear evidence

[3] See < http://www.unm.edu/~jbybee/Bybee%20plenary.pdf>.

https://doi.org/10.1017/S1366728906002732 Published online by Cambridge University Press

of self-regulating processes of word acquisition in the child and dove-tailing in the reactions and anticipations of adults in that process. The child's perceptual skills and memory development appear to be dynamically related in the sense that actions by the child lead to affordances by the adult, for example, by naming objects the child is holding or pointing to, and that adult bimodal pairing of words and objects, such as moving an object and synchronously naming it, allow the child to associate the two and remember that link. The interaction is dynamic in the sense that the adult reacts to the child and the child reacts to the adult with the adult trying to calibrate her reactions to what the child can handle. There is also a dynamic development in the interactional patterns. In the early stages, the adult synchronous bimodal presentation is needed for such pairing to take place, while in later stages the child appears to "guess" the relation between words and objects even when not presented synchronously. Not only the interaction between caretaker and child but also the interaction between hearing a word and perceiving the properties of an object is dynamic. Empirical evidence for the simultaneous acquisition of words and perceptual properties is found in an acquisition study by Yoshida and Smith (2005). In teaching English perceptual cues to Japanese children, the authors show the following:

> By teaching associations between words and perceptual properties, one will change not only what is known about the words, but also what is known about the correlations among the perceptual properties. (Yoshida and Smith, 2005, p. 94)

A final point here is to what extent language is a special aspect of cognition, or just another type of behaviour. While in the mentalist tradition language is seen as a special and probably uniquely human facility, the different approach to language presented above leaves little room for such a special position for language and we tend to agree with Cowan when he claims that "[l]anguage comprehension and production function more like problems to be solved in other domains" (Cowan, 2003, p. 441). If indeed language comprehension and production require specific problem-solving skills, we may ask how young immature minds can ever solve language puzzles. For this Shanker and King refer to a hypothesis put forward by Newport (1991) and elaborated on by Deacon (1997), which can be summarized as "less is more". Language must have evolved in such a way that the immature brain can acquire it and Deacon provides the following evolutionary argument to support this:

> Language structures may have preferentially adapted to children's learning biases and limitations because languages that are more easily acquired at an early age will tend to replicate more rapidly and with greater fidelity from generation to generation than those that take more time or neurological maturity to be mastered. (Deacon, 1997, p. 137)

To summarize this section, a DST approach to communication is incompatible with an Information Processing model and even though it does not necessarily reject innate principles, it does not need specifically linguistic principles to account for the creativity in language use. From a DST perspective, language acquisition emerges through interaction with other human beings within a social context. For example, Kirby and his colleagues, who have developed the ITERATED LEARNING MODEL, have been able to model "the process by which the output of one individual's learning becomes the input of other individuals' learning" (Smith et al., 2003, p. 371). In other words, they see language as a culturally transmitted system, which means that learning as an iterative process works both within the individual and between individuals at the social level. In this view, language learning is both individual learning and learning through interaction. In the next section we will examine DST and the language learning process in more detail.

DST and language learning

In language evolution research, there is extensive literature on modelling language learning and language evolution as co-evolving processes, which are compatible with DST (cf. Nowak et al., 2001, 2002). However, for our discussion of SLA we will take a model as developed by Van Geert (1991) as our starting point to describe what constitutes language learning. In this model of L1 learning, growth is defined as follows:

> A process is called growth if it is concerned with the increase or decrease (i.e. negative increase) of one or more properties, and if that increase is the effect of a mechanism intrinsic to that process. (Van Geert, 1995, p. 314)

For growth to take place there are a number of requirements a system has to meet:

- There must be something that can grow. Van Geert calls this the MINIMAL STRUCTURAL GROWTH CONDITION.
- There must be resources to keep the process of growth going. A distinction is made between
  - internal resources, resources within the learning individual: the capacity to learn, time to learn, internal informational resources such as conceptual knowledge, and motivational resources; and
  - external resources, resources outside the learning individual: spatial environments to explore, time invested by the environment to support learning, external informational resources such as the language used by the environment, motivational resources such as reinforcement by the environment and material resources such as books, and TV's.

Resources in growth systems have two main characteristics: they are limited and they are interlinked

https://doi.org/10.1017/S1366728906002732 Published online by Cambridge University Press

in a dynamic system. The limitations hold for all internal resources. Memory capacity is limited, as is the time available to spend on learning, the available knowledge, and the amount of motivation to learn. The same goes for external resources. Both the number of different types of environments to which the child is exposed and the caretaker's willingness to invest time and energy in learning support are limited. The resources must have some minimum value for learning to take place. Without memory, input, internal informational resources, or motivation, there will be no learning. At the same time, there are compensatory relations between different types of resources. Effort can compensate for lack of time, or motivation can compensate for limited input from the environment. Because resources, both internal and external, are part of an interlinked dynamic structure, a growth in a child's informational resources will lead to a change in the interaction with the environment through a demand for more demanding tasks and environments.

Such a change in the interaction with the environment is evident from, for example, findings by Hirsh-Pasek, Golinkoff, and Hollich (1999). Using their 1996 coalition model as a base, they examine the major transitions in development. Assuming that a process of distributional learning guides language acquisition, they find that infants are differentially biased to attend to particular stimuli over others at different times in their development. Within the first nine months of life, babies learn to segment prosodic and phonological information, which allows them to recognize "words". However, these sound segments do not start out as symbols that stand for what they represent, but are mere sound–object associations. Also the cues on which they rely to associate these sounds and objects differ according to age. Whereas at about 12 months the baby is especially sensitive to perceptual salience, by 19–24 months of age the baby pays more attention to social cues such as eye gaze. In sum, the primes that feed into word learning are the immature word-learning principles associated with the phonological forms that have emerged from the prior phase of phonological and prosodic analysis. There is also evidence that children begin learning grammar when armed at least with the developmental primes of grammatical morphemes and sensitivity to order.

Even though resources are limited and have to be distributed over different subsystems that grow, not all subsystems require equal amounts of resources. Some CONNECTED GROWERS as Van Geert calls them, support each other's growth. An example could be the relation between the lexical development and the development of listening comprehension: with increased listening comprehension words are understood and interpreted more easily, stimulating development of lexical skills, while knowing more words makes the understanding of spoken language in turn easier. In this way the two connected growers need fewer resources than two growers that are unconnected. On the other hand, conditions also need to be right for development to take place:

> Some conditions of growth and development are simply unsuccessful, not because de developmental mechanisms are not operating, or because the growth rates are too low, but because the mechanisms themselves create conditions that lead to inadequate forms of interaction. (Van Geert, 1994, p. 358)

Finally, the concept of CARRYING CAPACITY is particularly relevant for SLA. Since growth is resource dependent and resources are limited, growth is by definition limited. The carrying capacity refers to the state of knowledge that can be attained in a given child's interlinked structure of resources, referred to as the COGNITIVE ECOSYSTEM (Van Geert, 1994, p. 314). For example, the emergence of the multi-word sentence coincides with a deceleration of lexical growth in the one-word phase. While in earlier phases all resources could be used to develop the lexicon through the linking of different types of sensory information, in the next phase more and maybe different resources are needed to develop the grammatical system that governs the functional distribution of information in multiple-word utterances.

One empirical study to support this view was done by Robinson and Mervis (1998), who used data from a case study of a child that was recorded daily over a 13-month period beginning at 10.5 months of age. On the basis of daily recordings, the total number of words acquired during the weeks was established. A logistic function appeared to describe the developmental curve best. After a slow start, there is a spurt between weeks 30 and 40, which then levels off, as Figure 1 illustrates.

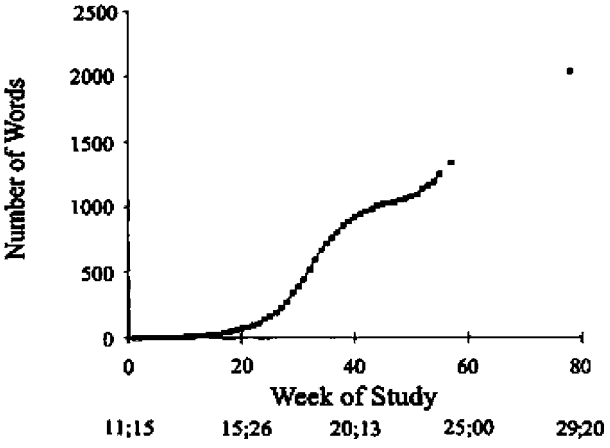


Figure 1. Ari's vocabulary growth based on the daily diary data (connected data points) and the final two 10-day vocabulary size estimates (from Robinson and Mervis, 1998, Figure 1).

https://doi.org/10.1017/S1366728906002732 Published online by Cambridge University Press

For the grammatical development, as measured by MLU data, the developmental curve is quite different: from 10 to 21 months there were basically only one-word sentences, and after this period there was a linear development of MLU. Another measure was the proportion of plurals used in obligatory contexts. Again, there was hardly any growth till week 30, and then a very rapid incline between weeks 30 and 40. Using a DST approach, Robinson and Mervis then tried to link the two variables. For the 10–21-month period, the correlation between MLU and vocabulary size was low ($r = .23$, $p = .517$), while for the 21–31-month period the correlation increased to.99 ($p < .001$). A similar pattern was found for the relation between the proportion of plurals in obligatory contexts and vocabulary size. A plot of the number of new words per week and proportion of plurals showed an interesting relation between the two developmental processes. Robinson and Mervis's Figure 6 (see our Figure 2) shows a nearly perfect negative relation between vocabulary growth and plural use. To combine these two developmental curves, Robinson and Mervis use a PRECURSOR MODEL as proposed by Van Geert (1995). In a precursor model, there are two variables, a predecessor and a successor. Growth in the successor is initially suppressed by growth in the predecessor, which is a form of competition, until a threshold level is reached in the predecessor. After this threshold is reached, growth in both variables shares any of the logically possible relations defined by competition and support. The change in relationships between the variables is captured in Robinson and Mervis's Figure 7 (see our Figure 3):

Van Geert has developed various other models to simulate different aspects of L1 development. As mentioned before, iterations play a crucial role in DST. This means that the same operation is carried out with the output of the previous operation as the input for the next. In other words, the present growth level depends on the previous growth level plus the interaction between that level and the resources available at that point. Here we present a simple model developed by Van Geert in his 1995 chapter. The basic equation is:

(1) $L_{n+1} = L_n(1 + R_n)$ for $L_0 > 0$

In this equation, L_n represents the growth level at point n after iteration, while R_n is the level of available resources at point n. As indicated above, the total amount of available resources will restrict the degree of growth and when resources are depleted, the growth will decrease. To model this, the variable R can be rewritten as:

(2) $R_n = r - a.L_n$

where r is a growth factor and a is a decline factor. This equation reflects that with more growth, resources will be depleted up to the point that r equals a and growth stops. Of course, as both R_n and L_n are in themselves complex

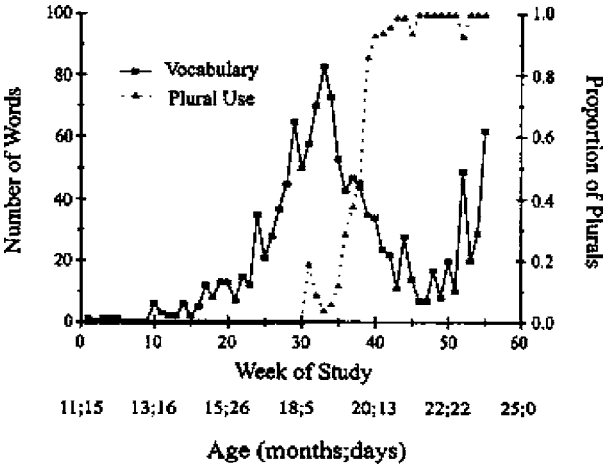


Figure 2. Number of new words added to vocabulary per week and proportion of plurals used in obligatory contexts (from Robinson and Mervis, 1998, Figure 6).

sets of variables, the equation is not as simple as its form might suggest. The equation mentioned in (1) and (2) is the growth equation that is used frequently in DST since it has shown to be able to model a wide range of developmental patterns in different areas. Depending on the initial state and the size of the growth rate, the equations may show various (unpredictable) growth patterns. Whereas convergence to a stable position in a state phase (a single attractor) is one possibility, convergence toward a cycle of 2, 4, 8 or more different attractor states or growth towards a chaotic pattern may occur just as easily.

The possibility of different growth patterns is crucial for our understanding of language development because it means that the same learning operation may lead to very different outcomes in the long run, depending on the starting point and the learning rate. So a similar

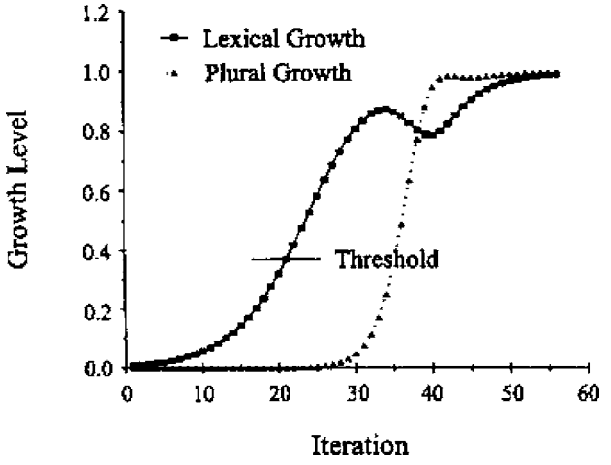


Figure 3. Precursor model of Ari's lexical growth and plural use in obligatory contexts (from Robinson and Mervis, 1998, Figure 7).

https://doi.org/10.1017/S1366728906002732 Published online by Cambridge University Press

"learning" procedure, rather than having a homogenizing effect, could actually lead to highly diverging patterns of development. Therefore, development may be gradual or it may show sudden changes, often reflecting a restructuring of the system (Van Dijk and Van Geert, 2005). Also some structures may appear and be used for a while and then disappear when they have lost their function. An example could be the fixed auxiliary + negation as typically found in the development of negation in L1 or L2.

Variability as part of development

The starting point of Dynamic Systems Theory is that a developing system is maintained by a flux of energy. Every developing cognitive system is constrained by limited resources, such as memory, attention, motivation, and so on. The system is in constant complex interaction with its environment and internal sources. Its multiple interacting components produce one or many self-organized equilibrium points, whose form and stability depend on the system's constraints. Growth is conceived of as an iterative process, which means that the present level of development depends critically on the previous level of development (Van Geert, 1994) and variation is not seen as NOISE but as an inherent property of a changing system:

> The theory radically rejects the automatic retreat to the error hypothesis and claims that variability bears important information about the nature of the developmental process. Dynamic systems theory stresses the importance of the context in which the behavior is displayed. Development takes place in real time and is considered highly context dependent. Therefore, it can be compared with an evolutionary process, which is also mindless and opportunistic. Thelen and Smith agree with the classical Darwinian emphasis on variability as the source of new forms. They state: "we believe that in development, as in evolution, change consists of successive make-do solutions that work, given abilities, goals and history of the organisms at the time" (1994, p. 144). Variability is considered to be the result of the systems' flexibility and adaptability to the environment. From a dynamic systems angle, variability has been viewed as both the source of development and the indicator of a specific moment in the developmental process, namely in the presence of a developmental transition. (Van Dijk, 2003, p. 129)

Intrinsic to this view is the idea that individual developmental paths, each with all its variation, may be quite different from one another, even though in a GRAND SWEEP view these developmental paths are quite similar. While the statistics of the TRUE SCORE approach are well developed and offer researchers the comfort of clear demarcations of what is "significant" and what is not (or so it seems), methods to look at variation as a source of information from a DST perspective are only beginning to be developed (see Van Dijk and Van Geert, 2005 for several interesting techniques that are also accessible for the less statistically sophisticated individual). The differences in techniques also reflect a different way of looking at developmental data. While the traditional approach is based on mathematical principles of chance and variation, the DST approach is much more geared towards visualization to see developments rather than to test them.

In our description so far we have stressed the unpredictability complex systems may show. However, we have to be careful not to overstress this point and confuse the terms COMPLEXITY and CHAOS with randomness. In development there is also a great deal of striking similarity, which may be explained by limitations in variability and change. This may reflect similarity in beginning conditions, but also similarity in the input and interaction with the environment. Much research on language development, while acknowledging the complexity of interacting variables, tries to disentangle the components of multifactor systems. A relevant example here is the work on the Competition Model for language development (see MacWhinney and Bates, 1989; and in particular MacWhinney, 1997 on SLA).

To conclude this section, from a DST perspective, a language learner is regarded as a dynamic subsystem within a social system with a great number of interacting internal dynamic sub-sub systems, which function within a multitude of other external dynamic systems. The learner has his/her own cognitive ecosystem consisting of intentionality, cognition, intelligence, motivation, aptitude, L1, L2 and so on. The cognitive ecosystem in turn is related to the degree of exposure to language, maturity, level of education, and so on, which in turn is related to the SOCIAL ECOSYSTEM, consisting of the environment with which the individual interacts. For any system to grow, a minimal amount of force or resources is needed. In addition, resources are compensatory. For instance, a low aptitude may be compensated by high motivation or vice versa. Each of these internal and external subsystems is similar in that they have the properties of a dynamic system. They will always be in flux and change, taking the current state of the system as input for the next one. A small force at a particular point in time may have huge effects (butterfly effect) and a much stronger force at another point in time may not have much effect in the long run. Each system has its own attractor and repeller states; however, variation is inherent to a dynamic system, and the degree of variation is greatest when a (sub) system moves from one attractor state to the other. Flux – growth or decline – is non-linear and cannot be predicted exactly.

DST and its applications to SLA theory

Most applied linguists will probably agree that SLA is an inherently complex process and recognize that many factors such as motivation, aptitude, degree of input, and L1 are all interrelated and have an effect on the L2 learning process. Nevertheless, many key issues in the

https://doi.org/10.1017/S1366728906002732 Published online by Cambridge University Press

SLA literature have been dealt with in clear cause-and-effect models and imply a linear point of view. In this section, we will explore how four key DST constructs– the role of initial states, attractor states, variation, and non-linearity – may apply to SLA.

Butterflies and attractor states in SLA

As indicated earlier, one of the most essential characteristics of dynamical systems is what has been called the SENSITIVE DEPENDENCE ON INITIAL CONDITIONS, or the butterfly effect. This effect refers to the unpredictability of the development of dynamical systems. Applied to meteorology, the question that was put forward by Lorenz in 1972 was: Does the flap of a butterfly's wings in Brazil set off a tornado in Texas? The comparison to SLA may not be obvious at first sight, but there is a growing body of evidence that suggests that initial conditions are precursors of the development of a second language.

As many SLA studies are cross-sectional and can therefore not give sufficient insight in the blur of interacting factors affecting the acquisition process, it is very difficult if not impossible to determine the effect of initial conditions on L2 development. Apart from indications of levels of proficiency in intervention experiments with a strongly limited scope, initial conditions are often ignored and assumed to be irrelevant. However, there is one line of research that points toward the occurrence of butterfly effects in SLA. It has been convincingly demonstrated that phonological awareness is one of the best predictors of reading acquisition in the native language (see, for instance, Stanovich, 1998; Sparks, Ganschow and Javorsky, 2000). Problems related to phonological awareness may be due to overt speech-related difficulties (like language delay, otitis media, etc.) in early childhood (Sparks and Ganschow, 1991). Furthermore, it has been argued repeatedly that a problem in one particular area of language learning affects other areas.

Consequently, the effect of phonemic coding difficulties may not be limited to reading and writing skills, but is likely to spread to the development of oral language, in both perception and production (see, for instance, Sparks, Ganschow and Patton, 1995). This spreading chain reaction will affect the overall ability to use the first language. It has also been shown that native language literacy is a crucial condition for the successful acquisition of a second language (see, for instance, Sparks and Ganschow, 1991; Dufva and Voeten, 1999) and that phonological awareness and word recognition skills in L1 affect word recognition in L2 (Durgunoglu, Nagy, and Hancin-Bhatt, 1993). From this evidence it can be tentatively inferred that very subtle and overt problems in early childhood, like a middle-ear infection, may have a long lasting effect at all levels of second language acquisition. This assumption is obviously rather speculative, but the growing body of evidence pointing at the causal relation between problems in L1 acquisition and the acquisition of a second language is a strong indication that difficulties in SLA are at least partly due to initial conditions butterflying their way throughout the process of second language acquisition.

Labov (1996) puts forward a similar suggestion about the causes of fossilization in L2 acquisition. He points out that one possible cause of L2 fossilization is an initial misperception of sounds. For example, Hispanic learners of English often simplify /rd/ clusters, perhaps because such clusters are difficult to perceive before an obstruent as in *card game* or *card table* and are therefore stored as *car game* or *car table*. If we assume some similarities between the L1 and L2 acquisition processes, it is very well possible that in the initial stages of L2 learning, the L2 learner is apt to store misperceived sounds, which quickly become entrenched in his or her speech and reach an attractor state that is difficult to get out of. So, for example, the L2 learner may not be able to perceive the past tense *-ed* marker as in *He talked to me* or the *to* in *I want to go* and simplify them, resulting in non-target forms not only in pronunciation but also grammar, commonly known as fossilization.

Larsen-Freeman (in press) also agrees that FOSSILIZATION reflects an attractor state. In her overview of some of the problems related to the concept of fossilisation, she shows there is basically no explanation (and even no adequate description) of fossilisation. By referring to work by MacWhinney and Tarone, she shows that fossilisation is both a cognitive phenomenon and a social phenomenon. Descriptions or explanations that focus on fossilisation as an end state in development miss the point of constant change that is typical of languages in use. She concludes: "If language is a dynamic system, then variability of performance and indeterminacy of speakers' intuitions would naturally follow" (2005, p. 10). Over time, the language systems settles in states that may reflect structures of the first language, overgeneralizations from the second language, but also in states that cannot be predicted nor explained by such influences. Such attractor states may be the unpredictable outcome of the constant and changing interaction between variables in the individual learner. But they are, as Larsen-Freeman points out cogently, not an end state in any sense and from a teaching perspective should not be seen as end states but rather as a reflection of the "boundlessness of potentiality".

Variation and morpheme order studies

L2 learner data usually display much variation. Researchers in second language acquisition have traditionally attempted to explain away seemingly FREE VARIATION by assuming full systematicity of variation in learner data. These explanations are often found in individual learner differences or the learner's linguistic

https://doi.org/10.1017/S1366728906002732 Published online by Cambridge University Press

or situational environment, but also in an assumed innate language specific endowment. The systematicity of this innate system used to be hidden in a black box that arouses language structures in a fixed linear order. Perhaps the best example of an assumed linear development in SLA is found in the morpheme order studies (Dulay and Burt, 1974). The publication of the article in which Dulay and Burt claim a fixed NATURAL SEQUENCE of the acquisition of grammatical morphemes regardless of the learner's L1 background initiated a lively debate on the issue. Apart from methodological points of criticism regarding the elicitation method used, the Bilingual Syntax Measure (BSM), the fact that the authors regarded accuracy to represent acquisition order met a lot of criticism.

Follow-up studies, with learners from different L1 backgrounds, using different elicitation methods measuring different levels of language and using adult participants, claimed to yield the same linear order of acquisition. The orders found, however, were not consistent in all studies. In a large-scale study conducted by Larsen-Freeman (1975), the orders found in the reading and the writing task did not comply with Dulay and Burt's order of acquisition. Hakuta (1974) found lower accuracy scores on articles for a Japanese learner of English, which might be due to the absence of articles in the learner's L1, but disturbed the neat fixed natural order of acquisition. A solution to the inconsistency in the orders found in different studies was found in creating clusters of morphemes. Referring to clusters, Krashen (1977) was able to account for the differences in the acquisition of morphemes.

Looking back at the morpheme order studies, we may now be able to account for orders found from a different perspective. The most important pitfall of the morpheme order studies was their predetermined desire to diminish the role of the learner's first language and to demonstrate linear development as a result of some innate mechanism. Grouping the data from many different learners is a perfect way to arrive at an overall picture of linear development. To illustrate this, consider the grouped versus individual data of 11 learners from a recent study into the development of L2 writing (Verspoor, Lowie and De Bot, in preparation).

In this study, a naturally occurring group of low intermediate high school students of English as a Foreign Language were asked to write a short journal entry at the beginning of each regular English lesson for about six weeks in a row. Depending on how many times students were present in class, there were 8 to 10 entries per student.

One of the variables under study was the development of the AVERAGE SENTENCE LENGTH of each story, represented for 11 students in Figure 4. As the solid gray line shows, the average sentence length fluctuates mildly around 12. However, when we study the developmental curves of the individual students a different picture emerges. In Figure 5a we have highlighted two individual students (TX and F) and in Figure 5b two other individual students (J and L). The difference is obvious. The two students in Figure 5a roughly mirror the pattern of the group as a whole: a mild oscillation around the initial value, in this case, of 10. Average sentence lengths in later observations are no better than those in the beginning of the curve. We can therefore describe this pattern as relatively "stable". However, subjects J and L in Figure 5b display a very different picture. First, while their "weakest" performance (around the value of 10) was roughly the same as the other students, they have much higher peak performances. Both students have one instance with an average sentence length of around 22, which is almost double the peak performance of many of the other students. This indicates that these two subjects are at least CAPABLE of producing much longer sentences, although they do not apply it all the time.

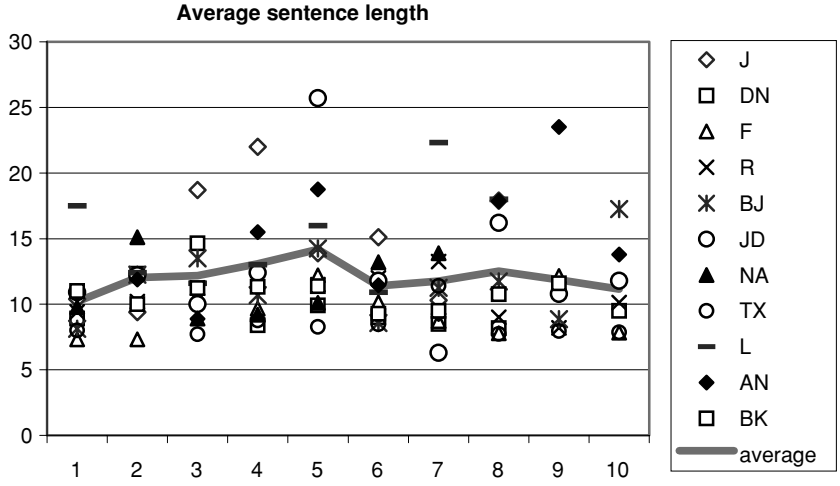


Figure 4. Average sentence length of 11 intermediate learners in 10 entries.

https://doi.org/10.1017/S1366728906002732 Published online by Cambridge University Press

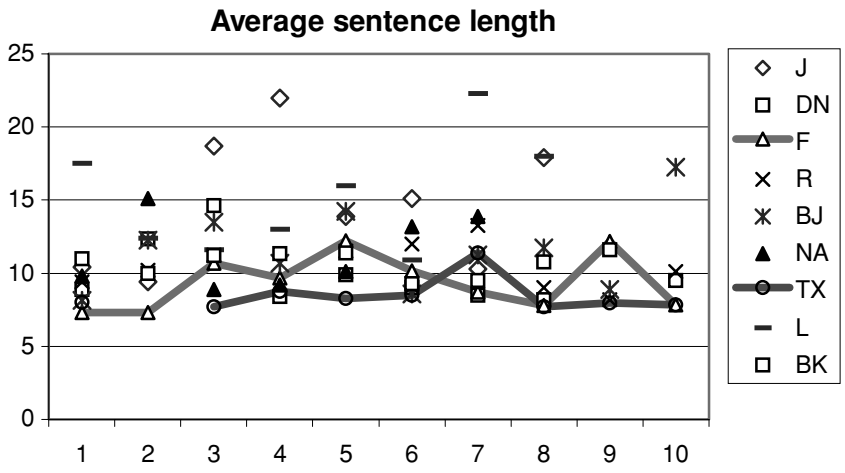


Figure 5a. Average sentence length of 11 intermediate learners in 10 entries with F and TX highlighted.

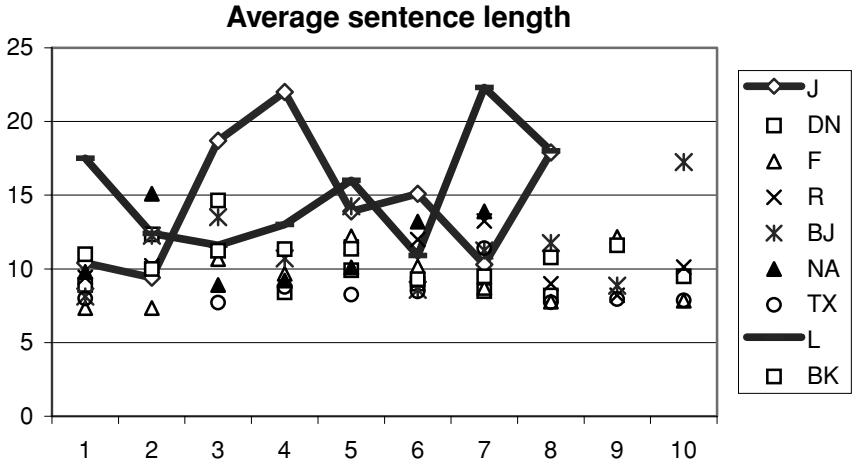


Figure 5b. Average sentence length of 11 intermediate learners in 10 entries with J and L highlighted.

This data illustrates the degree of both inter-individual and intra-individual differences in a supposedly rather homogenous group of learners during a short period of time, but these differences are concealed when averaged out as illustrated by the gray line in Figure 4. Moreover, if all the data points of all the individual learners had been clustered into one value, the average number of words per sentence is about 12. If we were interested in measuring the development of the group as a whole, it would not be unlikely that a few months later the same group would have an average sentence length of about 13, then 14 and 15. This is in fact the kind of a linear order of morpheme development Krashen (1977) achieved by not only clustering the data of different language users over time but also clustering different kinds of morphemes into one class.

Another problem of the early morpheme order studies was the lack of a theoretical explanation for the order, other than emerging from the black box. Researchers began to realize this and consequently the direction of the morpheme order studies shifted to the possible causes of the orders found. The causes mentioned included the syntactic complexity of the grammatical morphemes, earlier pointed to by Brown (1973) for first language acquisition; the use of similar ESL texts among English L2 learners (Bailey, Madden and Krashen, 1974); perceptual saliency; and phonological complexity. Based on the results of a detailed study, Larsen-Freeman (1976) argued that the main cause of the order was to be found in the frequency of occurrence in the input.

The search for an explanation of the morpheme orders usually starts from an all-or-nothing position. Some researchers stress the importance of one factor, while others emphasize another factor and a third investigation point to yet another one. From a DST perspective, however, causal factors need not be mutually exclusive. As Hirsh-Pasek et al. (1999) have shown for L1 acquisition, syntactic complexity, phonological complexity and frequency may

https://doi.org/10.1017/S1366728906002732 Published online by Cambridge University Press

be separate but dynamically interacting forces (attractors and repellors) shaping acquisition. This implies that any account that focuses on one aspect only cannot but provide a gross oversimplification of reality. Only an account that incorporates the dynamic interaction of all factors can form an appreciation of the actual complexity. At the same time, it is a matter of fact that it is very difficult to get a grip on complex interactions, so the optimal approach would be one in which the representation of the full complexity of the systems is linked to attempts to reduce that complexity separating highly relevant information from arguably less relevant information.

Non-linearity and attrition

As Herdina and Jessner (2002) point out, growth and decline are normal phenomena in developing systems, both are developmental, but the direction of change depends on the impact of internal and external resources. Research on both L1 and L2 attrition (Hansen, 2001; Schmid and Köpke, 2004) shows that language use and language input are vital for language maintenance and to account for the automatic decline with non-use, some DECLINE FACTOR has to be built in a language development model.

From a DST perspective one would not expect the decline to be linear nor to be similar across different individuals. However, this would be difficult to prove, because it is impossible to control for all variables. That is why computer simulations are useful. One of the few areas of SLA in which both network simulation data and empirical data are available is in language attrition research. Paul Meara, one of the pioneers of vocabulary acquisition research generally and the application of DST principles to lexical development, has recently (2004) presented work on attrition in lexical networks. He set up networks in which words are connected to two other words and all elements can be either on or off ("K = 2 random autonomous Boolean network"). Activation of a particular word depends on the level of activation of the words it is connected to. So changes in the activation status of one word will lead to a cascade of activation/deactivation until the network settles in a new attractor state. By varying the number of words that are deactivated (which are called ATTRITION EVENTS in his approach) the changes in the whole 2500 word network can be studied. The iterations with increasing numbers of switch-off words lead to remarkable differences in attrition patterns. While in some cases there was already a clear attrition effect after a fairly small number of attrition events, in other cases the network remained stable despite the large number of attrition events.

Of course, Meara's networks are much more rudimentary than a real lexicon, as there is considerable evidence that words are not simply switched on or off. For example, they show varying degrees of activation, which is reflected in latencies in reaction time experiments (Hakuta and d'Andrea 1992; Hulsen, De Bot and Weltens, 2002). Work using the savings paradigm, which tests residual knowledge through relearning, has shown that word knowledge can be retained at a very low level (De Bot and Stoessel, 2000). Therefore, a more realistic model would be one in which the lexicon consists of elements with varying degrees of activity. A general deactivation effect of network connections due to lack of use leads to a switching off of elements once they have reached a certain critical threshold.

Although Meara also emphasizes that the artificial network cannot be compared to what happens in a real lexicon, some of the findings are suggestive of how a dynamic systems approach may help us to interpret empirical data in language attrition. Firstly, the simulations show that attrition in such networks is clearly nonlinear: the effects of steadily increasing numbers of attrition events lead to very different patterns of attrition with some iterations leading to stable networks without much attrition and others showing heavy decline rapidly. Secondly, the data underline the earlier point of a need for a different approach to variation: while the average patterns show a gradual decline, none of the individual cases show this pattern.

Still, the findings in Meara's rudimentary attrition network mirror some of the crucial findings in attrition research. The first is considerable individual variation among individuals that are in more or less similar attrition settings. The work on Dutch and German in Australia (De Bot and Clyne 1994; Waas, 1996; Ammerlaan 1997) has shown that while some individuals already show considerable attrition after a few years in the L2 environment, others show a remarkable maintenance of language skills even after more than 25 years. The second is that there seems to be a part of the language system that, once acquired, is highly resistant to loss. This finding first emerged in work on the attrition of Spanish in the USA (Bahrick, 1984), which led Neisser (1984) to suggest a PERMASTORE of knowledge that is stable over time despite limited or no contact with the language. This finding has been supported by a host of research by various researchers (Hedgcock, 1991; Hansen, 2001).

In short, the combined work on simulation of dynamic lexical networks and empirical data on language attrition show the relevance of some DST ideas for SLA, in particular the non-linearity of development and the variation between individuals that follow from that.

Conclusion

The present article is an attempt to apply some DST concepts to SLA. We have argued that for some of the core issues in SLA a DST approach may help us develop a more realistic idea of what goes on in the learner's mind than other theories have done so far. In our view, the strongest

https://doi.org/10.1017/S1366728906002732 Published online by Cambridge University Press

point of a DST approach to SLA is that it provides us with a framework and the instrumentation that allows us to merge the social and the cognitive aspects of SLA and shows how their interaction can lead to development.

We have argued that languages and accordingly second languages behave like complex, dynamic systems. The major property of a Dynamic System is change over time. Through iterations of simple procedures that are applied over and over again with the output of the preceding iteration as the input of the next, complexity in language emerges. The dance metaphor is used to make clear that cognitive, social and environmental factors continuously interact, resulting in co-regulated interactions and the emergence of creative communicative behaviours. The developmental process of a complex system is determined by the initial state, the type of carrying capacity of the system, and the resources in both the cognitive system and the environment. In the developmental process certain sub-systems are precursors of other sub-systems. Not all sub-systems require an equal amount of energy because there are also connected growers, as may be shown in the dispersion of growth in the lexicon and grammar.

A DST view entails that an individual's language system with its numerous sub-systems is in constant flux, that the system as a whole and the sub-systems will show a great deal of variation, that small differences between individuals at a given point of time may have a great effect and that there is no such thing as an end state. Implicit in this view of language systems is that there is no need for a pre-existing Universal Grammar in the mind of any individual, but that a human disposition for language learning is required.

What DST provides is a set of ideas and a wide range of tools to study complex systems. We can no longer work with simple cause-and-effect models in which the outcome can be predicted, but we must use case studies to discover relevant sub-systems and simulate the processes. As the papers in volumes like Port and Van Gelder (1995) show, many aspects of human cognition can be modelled according to DST principles, provided we accept that sub-systems can be studied more or less in isolation and we are willing to accept some form of reductionism in our research. In a sense, DST can bridge the gap between holistic and reductionist views on SLA: it recognizes the fact that all aspects of human behaviour are connected and that the brain is not isolated and cognition is both embodied and situated as holisticists would argue, but at the same time it does aim at the full quantification that is the ultimate goal of the reductionists.

If this view of language is appropriate, it would entail that we should widen our scope in our quest to understand the SLA process. We should look to see whether individuals really have similar L1 systems. Considering the strong interaction with cognitive, social, and environmental factors, it is doubtful that their L1 systems are as similar as we may have assumed thus far. A DST approach would also predict that the cognitive and social skills apparent in the L1 affect the L2 learning process. By looking at dense corpora, we should also try to discover which sub-systems are precursors of other sub-systems and which sub-systems are connected growers. Such information could help us improve our teaching techniques and help avoid early entrenchment of non-target patterns. We should also look more at these factors in the attrition process.

Most importantly, though, we should look at our data with a more open mind. Traditional statistics is meant to reveal how a group performs as a whole and may be useful to see the grand sweep of things, but if we really want to know what happens in the actual process of language acquisition we should also look at the messy little details, the first attempts, the degree of variation at a developmental stage, and the possible attrition. It is very well possible that if we look closely enough, we find that the general developmental stages that individuals go through are much less similar than we have assumed thus far.

References

Ammerlaan, T. (1997). "Corrosion" or "loss" of immigrant Dutch in Australia: An experiment on first language attrition. In J. Klatter-Folmer & S. Kroon (eds.), *Dutch overseas: Studies in maintenance and loss of Dutch as an immigrant language*, pp. 69–97. Tilburg: Tilburg University Press.

Bahrick, H. (1984). Fifty years of second language attrition: Implications for programmatic research. *Modern Language Journal*, 68, 105–118.

Bailey, N., Madden, C. & Krashen, S. (1974). Is there a "natural sequence" in adult second language learning? *Language Learning*, 21, 235–243.

Brown, R. (1973). *A first language: The early stages*. Cambridge, MA: Harvard University Press.

Cooper, D. (1999). *Linguistic attractors: The cognitive dynamics of language acquisition and change*. Amsterdam & Philadelphia: John Benjamins.

Cowan, N. (2003). Comparisons of development modeling frameworks and levels of analysis in cognition: Connectionist and dynamic systems theories deserve attention, but don't yet explain attention. *Developmental Science*, 6 (4), 440–447.

De Bot, K. & Clyne, M. (1994). A 16-year longitudinal study of language attrition in Dutch immigrants in Australia. *Journal of Multilingual and Multicultural Development*, 15 (1), 17–28.

De Bot, K., Lowie, W. & Verspoor, M. (2005a). *Second language acquisition: An advanced resource book*. London: Routledge.

De Bot, K., Lowie, W. & Verspoor, M. (2005b). Dynamic systems theory and applied linguistics: The ultimate "so what"? *International Journal of Applied Linguistics*, 15 (1), 116–118.

https://doi.org/10.1017/S1366728906002732 Published online by Cambridge University Press

De Bot, K. & Makoni, S. (2005). *Language and aging in multilingual societies: Dynamic perspectives*. Clevedon: Multilingual Matters.

De Bot, K. & Stoessel, S. (2000). In search of yesterday's words: Reactivating a long forgotten language. *Applied Linguistics*, 21 (3), 364–388.

Deacon, T. (1997). *The symbolic species: The coevolution of language and human brain*. London: Penguin.

Dufva, M. & Voeten, M. (1999). Native language literacy and phonological memory as prerequisites for learning English as a foreign language. *Applied Psycholinguistics*, 20 (3), 329–348.

Dulay, H. & Burt, M. (1974). Natural sequences in child second language acquisition. *Language Learning*, 24, 37–53.

Durgunoglu, Y., Nagy, W. & Hancin-Bhatt, B. (1993). Cross-language transfer of phonological awareness. *Journal of Educational Psychology*, 85, 453–465.

Gogate, L., Walker-Andrews, A. & Bahrick, L. (2001). The intersensory origins of word comprehension: An ecological-dynamic systems view. *Developmental Science*, 4, 1–37.

Hakuta, K. (1974). A preliminary report on the development of grammatical morphemes in a Japanese girl learning English as a second language. *Working Papers in Bilingualism*, 3, 18–43.

Hakuta, K. & d'Andrea, D. (1992). Some properties of bilingual maintenance and loss in Mexican background high-school students. *Applied Linguistics*, 13 (1), 72–99.

Hansen, L. (2001). Language attrition: The fate of the start. *Annual Review of Applied Linguistics*, 21, 60–73.

Hedgcock, J. (1991). Foreign language retention and attrition: A study of regression models. *Foreign Language Annals*, 24, 43–55.

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

Hirsch-Pasek, K., Golinkoff, R. & Hollich, G. (1999). Trends and transitions in language development: Looking for the missing piece. *Developmental Psychology*, 16, 139–162.

Hopper, P. (1998). Emergent grammar. In M. Tomassello (ed.), *The new psychology of language*, pp. 155–175. Mahwah, NJ: Lawrence Erlbaum.

Hulsen, M., De Bot, K. & Weltens, B. (2002). Between two worlds. Social networks, language shift and language processing in three generations of Dutch migrants in New Zealand. *International Journal of the Sociology of Language*, 153, 27–52.

Krashen, S. (1977). Some issues relating to the monitor model. In H. Brown, C. Yorio & R. Crymes (eds.), *On TESOL '77*, pp. 144–158. Washington, DC: TESOL.

Labov, W. (1996). Some notes on the role of misperception in language learning. In R. Bayley & D. Preston (eds.), *Second language acquisition and linguistic variation*, pp. 245–252. Amsterdam & Philadelphia: John Benjamins.

Larsen-Freeman, D. (1975). The acquisition of grammatical morphemes by adult ESL learners. *TESOL Quarterly*, 9, 43–49.

Larsen-Freeman, D. (1976). An explanation for the morpheme acquisition order of second language learners. *Language Learning*, 26, 125–134.

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

Larsen-Freeman, D. (2002). Language acquisition and language use form a chaos/complexity theory perspective. In C. Kramsch (ed.), *Language acquisition and language socialization: Ecological perspectives*, pp. 33–46. London: Continuum.

Larsen-Freeman, D. (in press). Second language acquisition and the issue of fossilization: There is no end and there is no state. In Z. Han & T. Odlin (eds.), *Studies of fossilization in second language acquisition*. Clevedon: Multilingual Matters.

Lorenz, E. (1972). Predictability: Does the flap of a butterfly's wing in Brazil set off a tornado in Texas? Paper presented at the annual meeting of the American Association for the Advancement of Science, Washington, DC.

MacWhinney, B. (1997). Second language acquisition and the competition model. In A. de Groot & J. Kroll (eds.), *Tutorials in bilingualism: Psycholinguistic perspectives*, pp. 113–142. Mahwah, NJ: Lawrence Erlbaum.

MacWhinney, B. & Bates, E. (1989). *The crosslinguistic study of sentence processing*. New York: Cambridge University Press.

Meara, P. (2004). Modeling vocabulary loss. *Applied Linguistics*, 25 (2), 137–155.

Mohanan, K. P. (1992). Emergence of complexity in phonological development. In C. Ferguson, L. Menn & C. Stoel-Gammon (eds.), *Child phonological development*, pp. 635–662. Timonium, MD: York Press, Inc.

Neisser, U. (1984). Interpreting Harry Bahrick's discovery: What confers immunity against forgetting? *Journal of Experimental Psychology: General*, 113, 32–35.

Newport, E. (1991). Contrasting conceptions of the critical period for language. In S. Carey & R. Gelman (eds.), *The epigenesis of mind: Essays on biology and cognition*, 111–130. Mahwah, NJ: Lawrence Erlbaum.

Nowak, M., Komarova, N. & Niyogi, P. (2001). Evolution of universal grammar. *Science*, 291, 114–118.

Nowak, M., Komarova, N. & Niyogi, P. (2002). Computational and evolutionary aspects of language. *Nature*, 417, 611–617.

Port, R. & Van Gelder, T. (1995). *Mind as motion: Exploration in the dynamics of cognition*. Cambridge, MA: MIT Press.

Robinson, B. & Mervis, C. (1998). Disentangling early language development: Modeling lexical and grammatical acquisition using and extension of case-study methodology. *Developmental Psychology*, 34, 363–375.

Savage-Rumbaugh, S., Murphy, J., Sevcik, R., Brakke, K., Williams, S. & Rumbaugh, R. (1993). *Language comprehension in ape and child*. Chicago: University of Chicago Press.

Schmid, M. & Köpke, B. (2004). *First language attrition*. Amsterdam & Philadelphia: John Benjamins.

Shanker, S. & King, B. (2002). The emergence of a new paradigm in ape language research. *Behavioral and Brain Sciences*, 25 (5), 605–656.

Smith, K., Kirby, S. & Brighton, H. (2003). Iterated Learning: A framework for the emergence of language. *Artificial Life*, 9, 371–386.

https://doi.org/10.1017/S1366728906002732 Published online by Cambridge University Press

Sparks, R. & Ganschow, L. (1991). Foreign language learning differences: Affective or native language aptitude differences. *Modern Language Journal*, 75 (1), 3–16.

Sparks, R., Ganshow, L. & Javorsky, J. (2000). Déjà vu all over again: A response to Saitu, Horwitz and Garza. *Modern Language Journal*, 84, 251–255.

Sparks, R., Ganschow, L. & Pattton, J. (1995). Prediction of performance in first-year foreign language courses: Connections between native and foreign language learning. *Journal of Educational Psychology*, 87, 638–655.

Stanovich, K. (1998). Twenty-five years of research on the reading process: The grand synthesis and what it means for our field. In T. Shanahan & F. Rodriguez-Brown (eds.), *Forty-seventh yearbook of the National Reading Conference*, pp. 44–58. Chicago: National Reading Conference.

Thelen, E. & Smith, L. B. (1994). *A dynamic systems approach to the development of cognition and action*. Cambridge, MA: MIT Press.

Thompson, N. & Valsiner, J. (2002). Doesn't a dance require dancers? *Behavioral and Brain Sciences*, 25, 641–642.

Van Dijk, M. (2003). *Child language cuts capers: Variability and ambiguity in early child development*. Ph.D. dissertation, University of Groningen.

Van Dijk, M. & Van Geert, P. (2005). Disentangling behavior in early child development: Interpretability of early child language and the problem of filler syllables and growing utterance length. *Infant Behavior and Development*, 28, 99–117.

Van Geert, P. (1991). A dynamic systems model of cognitive and language growth. *Psychological Review*, 98, 3–53.

Van Geert, P. (1994). Vygotskian Dynamics of Development. *Human Development*, 37, 346–365.

Van Geert, P. (1995). Dimensions of change: A semantic and mathematical analysis of learning and development. *Human Development*, 38, 322–331.

Van Gelder, T. (1998). The dynamical hypothesis in cognitive science. *Behavioral and Brain Sciences*, 21, 615–656.

Verspoor, M., De Bot, K. & Lowie, W. (2004). Dynamic systems theory and variation: A case study in L2 writing. In H. Aertsen, M. Hannay & R. Lyall (eds.), *Words in their places: A Festschrift for J. Lachlan Mackenzie*, pp. 407–421. Amsterdam: Free University Press.

Verspoor, M., Lowie, W. & De Bot, K. (in preparation). Variation in L2 development: Seeing the noise from the sound.

Waas, M. (1996). *Language attrition downunder*. Frankfurt: Peter Lang.

Waters, G. & Wilcox, S. (2002). Making meaning. *Behavioral and Brain Sciences*, 25, 644–645.

Westbury, C. (2002). Blind men, elephants and dancing information processors. *Behavioral and Brain Sciences*, 25 (5), 645–646.

Yoshida, H. & Smith, L. (2005). Linguistic cues enhance the learning of perceptual cues. *Psychological Science*, 16, 90–95.

Zentall, T. (2002). A multichannel information-processing system is simpler and more easily tested. *Behavioral and Brain Sciences*, 25, 646.

Received July 20, 2005
Revisions received January 23, 2006; April 4, 2006
Accepted April 7 2006

https://doi.org/10.1017/S1366728906002732 Published online by Cambridge University Press