

University of Groningen

## Variability and Variation in Second Language Acquisition Orders

Lowie, Wander; Verspoor, Marjolijn

*Published in:*  
Language Learning

*DOI:*  
[10.1111/lang.12093](https://doi.org/10.1111/lang.12093)

**IMPORTANT NOTE:** You are advised to consult the publisher's version (publisher's PDF) if you wish to cite from it. Please check the document version below.

*Document Version*  
Publisher's PDF, also known as Version of record

*Publication date:*  
2015

[Link to publication in University of Groningen/UMCG research database](#)

*Citation for published version (APA):*

Lowie, W., & Verspoor, M. (2015). Variability and Variation in Second Language Acquisition Orders: A Dynamic Reevaluation. *Language Learning*, 65(1), 63-88. <https://doi.org/10.1111/lang.12093>

### Copyright

Other than for strictly personal use, it is not permitted to download or to forward/distribute the text or part of it without the consent of the author(s) and/or copyright holder(s), unless the work is under an open content license (like Creative Commons).

The publication may also be distributed here under the terms of Article 25fa of the Dutch Copyright Act, indicated by the "Taverne" license. More information can be found on the University of Groningen website: <https://www.rug.nl/library/open-access/self-archiving-pure/taverne-amendment>.

### Take-down policy

If you believe that this document breaches copyright please contact us providing details, and we will remove access to the work immediately and investigate your claim.

*Downloaded from the University of Groningen/UMCG research database (Pure): <http://www.rug.nl/research/portal>. For technical reasons the number of authors shown on this cover page is limited to 10 maximum.*

# Variability and Variation in Second Language Acquisition Orders: A Dynamic Reevaluation

Wander Lowie<sup>a,b</sup> and Marjolijn Verspoor<sup>a,b</sup>

<sup>a</sup>University of Groningen and <sup>b</sup>University of the Free State, South Africa

The traditional morpheme order studies in second language acquisition have tried to demonstrate the existence of a fixed order of acquisition of English morphemes, regardless of the second language learner's background. Such orders have been taken as evidence of the preprogrammed nature of language acquisition. This article argues for a process-based, dynamic explanation of development, in which each developmental step is based on the dynamic interaction of all processes involved. Due to the complexity of these interactions, the developmental process cannot be predetermined and fixed. Although stages of development like the acquisition order of morphemes are commonly observed as a grand sweep effect at the group level, these stages may be meaningless at the level of the individual language learner. This paradox shows we can only make the observations that our method allows us. If we are interested in grand sweep effects that may be generalizable to large populations of learners, we will have to carry out group studies with representative samples that can be analyzed using Gaussian statistics based on the normal distribution. But if we are interested in how an individual learner progresses over time as a result of changing variables in a changing context, we will have to conduct longitudinal studies and use nonlinear methods of analysis.

**Keywords** variability; morpheme orders; dynamic systems; second language acquisition

## Introduction

Second language acquisition (SLA) research in the late 1970s and early 1980s studied the sequence of acquisition of morphosyntactic characteristics of English by second language (L2) learners with different first language (L1)

---

Correspondence concerning this article should be addressed to: Wander Lowie, University of Groningen, Department of English, P.O. Box 716, Groningen, 9700 AS, Netherlands. E-mail: w.m.lowie@rug.nl

backgrounds. The purpose of these studies was to see if the L2 would develop in a fixed order, irrespective of the learner's L1. If this were the case, it could be argued that L2 learning is a universal and therefore predetermined process, which would point to both the innateness and the systematicity of syntax and grammar. The studies showed that approximately the same order of acquisition of English morphosyntax was found for learners with very different mother tongues. This caused a great deal of discussion about the interpretation of these findings, and even though our thinking about language acquisition has changed over the past 40 years, the morpheme order studies are still under debate. Given the symposia and colloquia at major recent conferences in the field, it could even be argued that there is an upsurge of interest in these studies. The discussions have centered around two questions. The first question is whether a fixed order of acquisition has convincingly and consistently been demonstrated. Not surprisingly, there is still disagreement about this point. Second, if we can speak of a fixed order, the question is what causes the order. Is a fixed order indeed evidence of the universal nature of acquisition that is the same for all language learners, all types of language learning, and irrespective of the learner's L1? Or are there other explanations for the observed similarities?

In this article we address both of these questions by discussing the possible existence of acquisition orders and their implications about the underlying learning process from a Dynamic Systems Theory (DST) point of view. We explain how DST can provide new insights into these questions by studying the type and amount of variability in language use. The objective of DST applications to SLA is to acknowledge the fact that language development should be seen as an individually owned process rather than a product and that this process is shaped by the nonlinear relationships of changing components over time. We argue that the process-based view and its consequences for the type of analyses we perform have important implications for the way we interpret language learning data. In this article we first explore the relevance of variability for the explanation of development, alongside a dynamic reinterpretation of the morpheme order studies. Subsequently, we discuss the theoretical and practical implications of these observations in view of both theory formation and empirical studies in the field of SLA. Finally, we try to point out the interconnections with the other contributions in this Special Issue.

### **Toward a Process-Based Explanation of Development**

The prevailing approach to SLA up to the beginning of this century has been to employ the product of language use as the measure of evaluation, usually

in combination with one or more independent variables, such as motivation, aptitude, anxiety, et cetera. The question then is in what way and to what extent the dependent measure (the product of language learning) can be accounted for by one or more of the independent variables. Using analyses of variance, regression analyses and, more recently, mixed-model analyses of means comparisons of group results, conclusions are drawn about the nature of the learning products and the factors that have played a role in shaping these products. When representative samples are assumed to have been drawn, the results are then generalized toward bigger populations of learners, which are in the end supposed to inform the process of language learning in general. This approach is well attested and generally applied in most fields of research and as long as basic assumptions regarding normality of the distribution, homogeneity of variance, representativeness of the sample, and careful operationalization are certified, the validity of the outcomes is assumed to be warranted (Lowie & Seton, 2013). This approach allows us to test explicit hypotheses about the factors influencing the language products in a very straightforward and objective way, similar to the way in which phenomena in nature can be tested against the laws of nature. Like the influence of gravity and resistance on dropping objects, the influence of individual factors on language use is supposedly tested accurately and objectively in this way. Due to the frequent application of quantitative statistical methods to language learning data and due to the assumed objectivity of hypothesis falsification, many researchers explicitly or implicitly assume that this method is the best way to come to an understanding of the process of language learning.

However, in the wake of Larsen-Freeman's (1997) argument in favor of considering language learning as a complex system and de Bot, Lowie, and Verspoor's (2007) argument for a dynamic systems approach to L2 development, an increasing number of studies have argued against the exclusive use of Gaussian statistics (based on the normal distribution) to evaluate the process of language learning. Three crucial points have been made by the complexity and dynamic systems approaches in order to emphasize the limitation of the interpretation of group means as a tool that tests hypotheses about factors affecting language learning.

The first point is that learning must be seen as a process rather than as a product. This implies that studying performance at one point in time may provide an inaccurate or at least an incomplete picture of language development. Although traditional intervention experiments with a pretest, a posttest, and often a delayed posttest can give us valuable information about the influence of individual factors or conditions on the state of the L2 system, they cannot

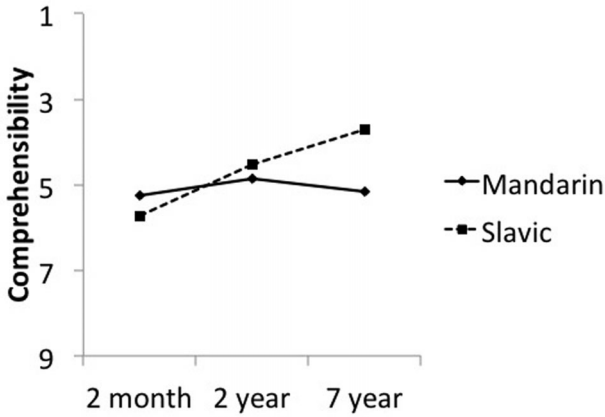
truly represent the individually owned process of development of that system. An illustration of this is provided by Lowie (forthcoming) on the acquisition of the phonology of Dutch as an L2. This case study followed the development of the Dutch phonological system by two American English children during their 1-year stay in the Netherlands. The data consisted of weekly recordings of several tasks that were analyzed in terms of a number of phonological correlates. When comparing the first and the last recordings of this longitudinal case study, the conclusion could be drawn that targetlike productions of Dutch vowels had been attained at the final data point. However, this conclusion would fail to give insight into the actual process, as the dense longitudinal data shows highly variable productions that indicate a very unstable vowel system. Moreover, any measurement before or after the focal point could have given a different picture. In other words, conclusions about the eventual attainment are strongly dependent on the coincidental time of the measurement. This study illustrates that the developmental process can only be approximated by extended time series and cannot be inferred by measurements at one or two points in time.

When we apply this observation to the order of acquisition of the morphosyntactic characteristics of English reported in the early morpheme order studies (see Ellis, this Special Issue for other views), we must conclude that the observed data are not in fact orders of acquisition, but rather accuracy scores at one moment in time. Whether accuracy scores can actually be associated with the sequence of acquisition is doubtful, as they do not provide information about the actual sequence in which the morphemes are acquired. In later studies in search of universal orders (e.g., Pica, 1983), the focus has been on acquisition sequences rather than accuracy orders, which were inferred from a limited number of cross-sectional groups. Although this approach may be a slightly closer approximation of the developmental trajectory, these few cross-sections cannot be regarded as truly representative of the process of acquisition, and they cannot answer the question of whether the order of acquisition is fixed regardless of the learner's native language. In fact, there are a number of longitudinal case studies that focus on sequences of acquisition (e.g., Cancino, Rosansky, Schumann, & Hatch, 1978). These studies typically show a high degree of variability within the same learner and large variation among learners;<sup>1</sup> therefore, they do not show a clear and fixed developmental sequence irrespective of the learner's age, learning context, and L1 background. For instance, when Rosansky (1976) discusses one of the six L1 Spanish speakers (Jorge, 13 years old, immersed in an English school context) in her study on the development of morphemes, she observes that the rank orders of the morphemes "fluctuate from month to

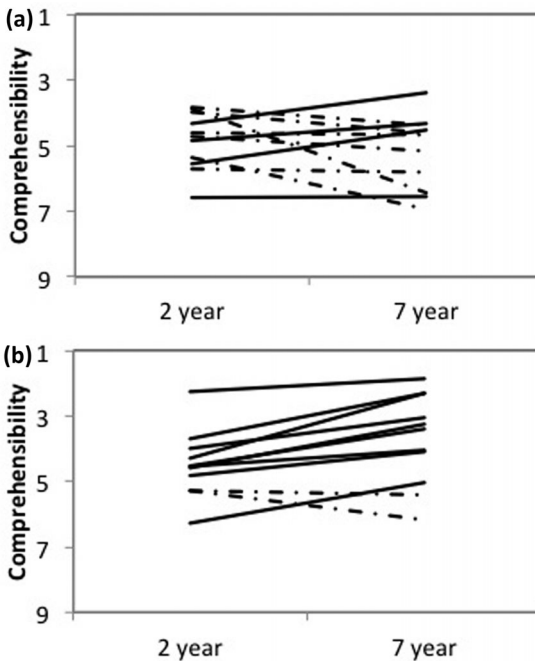
month” and do not resemble the overall longitudinal trend (Rosansky, 1976, p. 423). The same explicit observation is made by what could perhaps be seen as the original source of the morpheme order studies: Brown (1973) asserts that “there is by no means a perfectly regular progression in the data; there are uncountable regressions and unexplained abrupt advances” (p. 388). More recent case studies and corpus studies on the development of constructions in L2 (e.g., Eskildsen, 2012; Yuldashev, Thorne, & Fernandez, 2013) also report considerable variability and provide further evidence against a predetermined fixed order of acquisition. Summarizing, we would claim that measuring learners’ accuracy scores does not do justice to the process of language development. This is confirmed by longitudinal morpheme order studies (e.g., Hakuta, 1976) that refer to the actual sequence of acquisition, which apparently deviates from the order of acquisition inferred on the basis of accuracy counts.

A second point that has been emphasized by advocates of a dynamic systems or complexity approach to L2 development is that most studies making strong claims about fixed orders of acquisition have only looked at the overall developmental trend for groups of learners that may not be representative of any of the individual learners in the group. This problem is clearly illustrated by the results of a longitudinal study on L2 accent (Derwing & Munro, 2013; Derwing, Thomson, & Munro, 2006). These authors investigated the development of foreign accentedness and comprehensibility of groups of English L2 learners with a Slavic or Mandarin L1 background in a naturalistic environment. A comparison of the data at the onset of the study to the data after 2 years and 7 years of exposure showed that there was a significant improvement in accent for the Slavic participants, but not for the Mandarin participants (see Figure 1). However, as Derwing and Munro (2013) showed in a closer examination of individual development, some learners became more comprehensible while others had a stronger foreign accent after 12 years and became less comprehensible. None of the individual participants showed the mean pattern of the group (see Figure 2). In cases like this, we could and should doubt the relevance and meaningfulness of the mean trend, in spite of the fact that this pattern was found to be significant.

The problem for the generalization of longitudinal group data is explained by van Geert (2014). He argues that different dimensions can be distinguished in studying development of a L2. The first two dimensions comprise the interaction of several variables (dimension 1) as they shape the developmental trajectory of an individual learner (dimension 2). Taken together, these dimensions can be represented as a dynamic causal network of positive and negative couplings of the variables as they change over time. The



**Figure 1** Display of group means on the development of L2 English comprehensibility by Mandarin and Slavic learners (from Derwing & Munro, 2013, p. 174).



**Figure 2** Display of individual evidence on the development of comprehensibility (Derwing & Munro, 2013, p. 175). Picture 2a represents the development of the 11 Mandarin L2 speakers; Picture 2b represents the development of the 11 Slavic L2 speakers?

developmental patterns these networks produce are typically characterized by individual variability, from which a number of principal components can be derived. Yet, there is no evidence that the interaction of variables over time is the same for different individuals. Adding a third dimension of different individuals that are dealing with the same coupled variables will therefore lead to a different dynamic network. The assumption in Gaussian statistics is that the different individuals are of the same kind (e.g., intermediate learners of English with the same L1) and that therefore generalizations can be made for the entire population they represent. However, these generalizations are not warranted for development, as the dynamic network of coupled variables affecting the developmental trajectory of an individual cannot be equated with interindividual variation at one moment in time. Modern mixed-model designs can neutralize individual variation within the group by including it as a random effect in the analysis to allow for adequate generalizations in the frequency domain. But such generalizations are not warranted in the time domain. The observation that interindividual variation cannot be equated with intraindividual variability over time is further supported by Molenaar and colleagues (Molenaar, 2008; Molenaar, Huizenga, & Nesselroade, 2003), who convincingly showed it is mathematically wrong to equate individual behavior with group behavior.

The problem of the generalization of observations just discussed is clearly reflected in the early morpheme order studies as well. Based on the data in her longitudinal study, Rosansky (1976) doubts the generalizability of morpheme orders for groups of language learners: “With variance this large one must ask whether the sample means in these studies (my own included), are reasonable estimates of the population means, that is, are we accurately describing the language performance of the population?” (p. 418). From a dynamic systems point of view, the answer to this question will have to be negative. Because the focus of attention is the time dimension, the generalization of orders or even sequences of development relating to groups of individual learners is not warranted.

The third point raised by advocates of a dynamic systems or complexity approach is that the nature of the relationships among factors involved in L2 development may be nonlinear. Specifically, the assumption underlying the use of parametric Gaussian statistics— that the relationships under investigation are linear relationships— is fundamentally challenged by recent studies that have used a dynamic perspective and which show the relationships between subsystems of a developing language system are nonlinear (de Bot et al., 2007; Larsen-Freeman & Cameron, 2008). That is, due to the potentially continuous change of all interacting subsystems over time, and the constantly changing



interaction of all variables over time, the end result is not predetermined. The essence of nonlinearity of a relationship is that the outcome of the interaction of that relationship is not directly proportional to the input of the individual components. Consequently, some aspects of nonlinear systems appear to be unpredictable and chaotic. This does not imply that the behavior of a nonlinear relationship is random, but it does mean that the outcome is not predetermined. Nonlinearity of relationships can be found in all the three dimensions outlined by van Geert (2014) and discussed above, including the time dimension, so that we can indeed speak of nonlinear change or nonlinear development. Typical examples of nonlinear change are complex systems in the physical world, like the weather and our climate. Due to the nonlinear relationship between the subsystems that shape the weather conditions, and due to the iterative nature of the changes, small changes at one point in time may have unexpectedly large consequences on the long term, so that specific long-term predictions (as opposed to relatively accurate daily forecasts) are highly unreliable. This does not mean that the changes occur randomly. There may be global stages or repeating cycles at different time scales, like night and day or summer and winter. We know that within a 3-week period the cherry trees will blossom. But we cannot predict when an individual flower bud will open up or whether it will die.

Apart from some notable exceptions that will be discussed below, the large majority of classic morpheme order studies and other studies that focus on orders or sequences of L1 and L2 development have based their conclusions on group data at one moment in time, on measurements at two data points, or on cross-sections comparing groups of learners at different levels. If language development is a nonlinear, complex dynamic process affected by constantly changing interactions between subsystems at all levels and time scales, the conclusions from these studies may not be warranted. Two questions logically follow from this: Is language indeed a dynamic system? And if so, how does one investigate language development as a dynamic process?

### **Nonlinearity in Development, from Cognition to Language**

In the past decade or so, an increasing number of researchers have made the claim that human development is a nonlinear process. Thelen and Smith (1994; see also Smith & Thelen, 2003) have made a very strong claim that cognition must be regarded as a dynamic system and human learning as nonlinear development. They emphasize the point that ontogenetic processes emerge from the ongoing complex interaction of subsystems that are intrinsic to the system

rather than following an inevitable and predetermined route toward maturity. As an example they refer to the process of learning to walk. Early theories of the development of this ability regarded it as a “linear, stagelike progression through a sequence of increasingly more functional behaviors, driven toward adult forms by a grand plan” (Thelen & Smith, 1994, p. 6). For example, Konner (1991) maintains that motor development must be explained as a genetically programmed sequence of stages. His evidence for this claim originates in the assumed phylogenetic nature of these motor skills, based on the observed universality that all humans learn to walk in the same way. This, incidentally, is strongly reminiscent of a Universal Grammar (UG)-inspired view of language acquisition and the idea of a fixed order of acquisition of grammatical morphemes. Contrary to the idea of motor development as a predetermined process of stages that can be explained by a single cause—the maturation of the brain—Thelen and Smith demonstrate that learning to walk is a continuous process that is shaped by contextualized performance. They do so by demonstrating that very early limb movements, long before walking is observed, are essentially the same as the limb movements required for walking. On the treadmill, children show a remarkable ability to walk long before they reach the walking stage. Walking emerges from the continuous, coupled changes in a large number of embedded subsystems, including the brain, postural stability, increasingly strong muscles and bones, a motivation to move forward, and a walkable surface. Learning to walk is not a predetermined process of maturation, but a process that emerges from microscopic changes in all the subsystems involved in the holistic, self-organizing complex system.

What is true for locomotion, Thelen and Smith (1994) argue, is true for cognition. In spite of the bird’s-eye view that shows distinct developmental stages, as for instance worked out by Piagetian thinking, cognition does not develop in predetermined and possibly innate stages. The logical fallacy is in the assumption of a single cause. Cognition cannot only be accounted for by maturation of the brain but is formed by multiple causes. The Piagetian stages do not take into account the complex interactions that change over time in shaping the cognitive system. This is very clearly demonstrated in the A-not-B error, explained in detail by Smith and Thelen (2003). The A-not-B error is typically made by infants at a certain stage of development (substage 4 of the sensorimotor stage, around 10 months). It can be demonstrated by two types of trials. In trial 1 the experimenter hides a baby’s favorite toy under a box. The baby will search for the toy and will find it under the box (box A). Trial 1 is repeated a number of times. In trial 2, the experimenter hides the toy under a different box (box B). Children of 10 months will try to find the toy under box

A, while children of 12 months will ignore box A and find the toy under box B. This is a rather consistent finding for children at this stage of development and has been used as evidence for the existence of developmental stages that are universal, as all children at this stage make this error. However, in a series of experiments led by Thelen and Smith, it has been shown that the single causality of maturation is flawed and that a large number of factors can affect the error, regardless of the child's age; namely, the error is dependent on the timing between the trials, the position and the properties of the covers, and the number of repetitions of the first trial (Diedrich, Highlands, Spahr, Thelen, & Smith, 2001; Smith, Thelen, Titzer, & McLin, 1999). Changing the posture of the baby during the trials made the error disappear even for 8-month-old babies (Smith et al., 1999), as did attaching weights to the babies' wrists (Thelen, Schöner, Scheier, & Smith, 2001). These experiments clearly show the multicausality of the error and the wide range of cognitive, physical, and environmental factors that may affect it. While the grand sweep view of development may seem to show the existence of universally occurring stages, more detailed observations at shorter time scales and in different domains will show that new behavior emerges from the continuity of interacting variables over time. Minute changes in seemingly unrelated domains may lead to major changes in the self-organizing system and can lead to relevant individual differences. Summarizing their argument to see cognition as a dynamic system, Smith and Thelen (2003) conclude:

In human development, every neural event, every reach, every smile, and every social encounter sets the stage for the next and the real-time causal force behind change. If this is so, then we will gain a deeper understanding of development by studying multicausality, nested timescales and self-organization. (p. 347)

What is true for cognition has also been shown to be true for language development (de Bot et al., 2007; Larsen Freeman, 1997; van Geert, 1991). Maybe more than any other type of cognition, the language system is integrated in numerous subsystems in complete interconnectedness. While traditional analyses of variance and regression analyses have investigated the influence of motivation as a factor affecting language learning, a dynamic multicausality approach regards motivation as an integral part of a causal network, just like the environment in which language learning takes place. The context in which language learning takes place is coupled with self-perception, which is coupled with motivation, which is coupled with language learning (Dörnyei, 2009).

Language is also possibly the most obvious example of embodied cognition, which is both formed by and forms an integral part of social interaction. De Bot et al. (2007) emphasize this interconnectedness by referring to the dance metaphor applied to language interaction. Rather than a sequential interaction in which one isolated dyad reacts to another dyad, the dance partners are peracting agents. As Thompson and Valsiner (2002) put it:

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. (p. 641)

This type of interconnectedness accounts for perturbations of the language system in real time, but it is also the foundation of a usage-based account of ontogenetic language development at a larger time scale, and an explanation for phylogenetic language change at yet another time scale. Real-time processing, development, learning, and evolution have traditionally been considered as distinct processes. In a DST perspective they are essentially the same, though at different time scales. These dynamic time scales are fully interwoven. Real-time processing uses the dynamic history of components, while the real-time action itself becomes part of the dynamic history of the system (Smith & Thelen, 2003, p. 74).

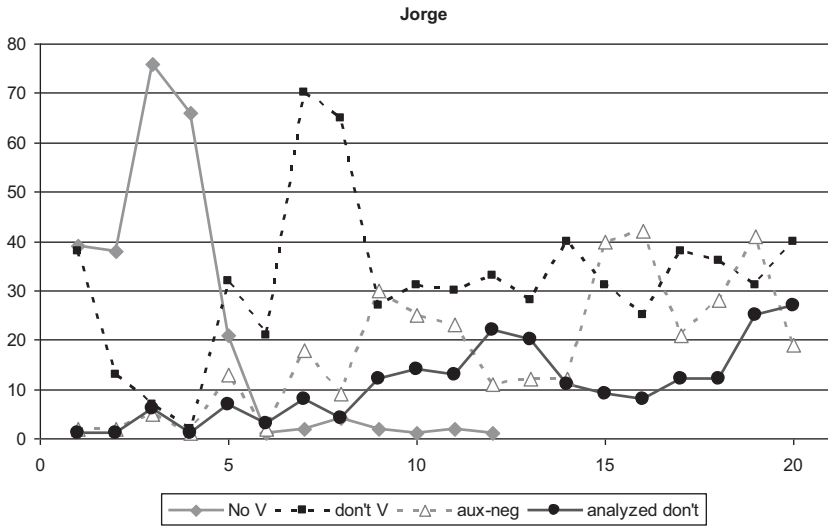
Similar to dynamic systems in the physical world, languages consist of many embedded subsystems. Linguistic subsystems are for instance the phonology, the morphosyntax, and the semantics of a language, which in turn consist of subsystems for different languages in a multilingual speaker (see Lowie & Verspoor, 2011). The language system itself is embedded as a subsystem in the larger cognitive system, which is embedded in a person's body and mind. An individual person is an embedded subsystem of a small circle of language users, embedded in a language community, et cetera. An important distinction between traditional models of language processing and the dynamic approach is that the subsystems are open modules. All changing subsystems can potentially and continuously interact with all other changing subsystems. Viewing language as an integral part of dynamically embedded subsystems has a strong explanatory power for language development as a semiotic system and embodied cognition. Like the smaller subsystems of language, language itself is embedded in the body. Embodiment in language production is evident from the fact that we need the muscles related to our speech organs to produce speech. Holistic task dynamic models of articulation (e.g., van Lieshout, 2004) assume that

speech production requires the coordinated action of all articulators involved in producing a speech gesture. But language and meaningful interaction are not limited to speech but are strongly integrated with facial expression and gestures of the hands and the rest of the body. Gesture research has demonstrated the strong interconnectedness of speech and body in both language use and language acquisition (de Bot & Gullberg, 2010). This observation is completely in line with a growing literature that suggests that “much human intelligence resides in the interface between the body and the world” (Smith, 2005, p. 286).

Considering the nature of language development, we can indeed conclude that this is a complex dynamic and nonlinear process. It is both complex and nonlinear, as language learning is not the product of just one process. Many social, psychological, physical, and environmental factors are involved, and all these factors are interconnected. Similar to what has been demonstrated by research in children’s cognitive development, there is no single cause of the ontogenetic process, and there is no sequence of influences at different stages of development. Instead, there are multiple and continuous interactions that form a multidimensional causal network. All dimensions of the network can change over time, but each dimension may have its own time scale. The emerging picture is a view of language as a dynamic and multidimensional state space that is essentially unique for each and every individual. Reducing research into this complex dynamic system to a search for a universal order of acquisition as the result of an innate grand scheme affected by multiple variables and at a single moment in time does not do justice to the complex interconnectedness of the process of language development. Human behavior, including language development, cannot be expected to obey the theorem of stationary data. To understand the process of language acquisition, the starting point should therefore be the individual process line as it unfolds in changes over time.

### **Alternatives Methods of Analysis: Variability as the Motor of Change**

Variability has a long history in SLA studies but, as Verspoor, Lowie, and van Dijk (2008) point out, many studies were concerned with explaining variability in interlanguage, which was assumed to be rather systematic. However, there were a few early studies that recognized that interlanguage is not necessarily systematic and that not all variability could or should be explained. After eliminating factors that would contribute to systematic variation, Ellis (1994)



**Figure 3** Cancino et al. (1978) data for Jorge’s (age 13) development of negative constructions. The figure shows the frequency of occurrence of the different negations (y axis) over time in weeks (x axis). (Reprinted with permission from van Dijk et al., 2011.)

found that there was still some degree of free variation, that is, variability that could not be attributed to any known linguistic, situational, or psychological factor. He refers to Cancino et al. (1978) and Gatbonton (1978) and concludes that a general finding of these studies is that “free variation occurs during an early stage of development and then disappears as learners develop better organized L2 systems” (p. 137). It is this type of variability, in conjunction with systematic variability, that researchers working from a DST perspective wish to examine in more depth. The reasoning is that the amount of variability will be relatively high when the system is reorganizing and low in a more stable system. Because variability is an inherent property of a self-organizing system, the degree and patterns of variability can inform us about the principal components in the developmental process.

For example, several studies have shown that increased variability coincides with a developmental jump (Spoelman & Verspoor, 2010; van Dijk, Verspoor, & Lowie, 2011). As Ellis (2015) discusses, van Dijk et al. (2011) reanalyzed the Cancino et al. (1978) data from a dynamic perspective and showed the emergence of the negative verb constructions produced by Jorge in detail. The result of their reanalysis is shown in Figure 3. As far as order of acquisition

is concerned, the expected sequence found in L1 acquisition was indeed recognizable in the L2 data as well. However, by only looking at the order of acquisition, other interesting developmental phenomena were overlooked in the original analysis. First of all, even at the very first data session, all four negative constructions occurred, and each construction seemed to develop in a wavelike form, with the nontargetlike construction first showing a high wave, the *don't* construction showing a second high wave, and the other two targetlike constructions exhibiting somewhat lighter waves. Second, the nontargetlike constructions would keep occurring even when targetlike constructions had increased. Finally, when van Dijk et al. tested the trajectory of the *don't* construction in a Monte Carlo simulation, the peak at around data point 7 was statistically significant, indicating that it was a developmental peak.

From a logical point of view, the relevance of variability makes good sense. If there is no variability, there can be no development. A large amount of variability signals that the learner is apparently trying things out and that the subsystem under consideration is unstable. The connection between learning and variability is also observed in the physical world. A child that is learning to ride a bike will initially sway heavily to keep its balance. With increasing experience, corrections will be timed better, resulting in less variability. If the riding conditions become more challenging, like riding on gravel instead of a smooth surface or riding with the hands off the handlebars, the child will initially sway more heavily again until this stage is mastered and variability decreases again. This example points to the meaningfulness of variability and its usefulness in understanding the dynamic learning process. Applied to grammatical development, the learner may initially use many different forms rather randomly but become increasingly sensitive to using the most conventional forms in a certain context. Initially this more balanced use may be disturbed by conditions like stress, but in the course of time the stability of the language system is likely to increase. In other words, variability is not a meaningless byproduct of development but is a driving force and a motor of change. Detailed studies of variability in L1 development (Bassano & van Geert, 2007) and in L2 development (Spoelman & Verspoor, 2010) have shown the relevance of variability in accounting for the process of development.

At a shorter time scale, analyses of variability have been used to interpret the coordination of subsystems during L1 and L2 use in real-time processing. Based on methods of analysis in adjacent fields of study, several researchers (see below) have investigated the pattern of variability and the long-range dynamics between data points in a time series to reveal information about the blend of mechanisms underlying language processing. The pattern of interest in this

body of work has been referred to as “pink noise,” which is a semiregulated variability pattern that signals a state of the dynamic system that is optimally operative between automatized behavior and adaptability to the changing environment. Pink noise has been shown to indicate coordinated behavior of dynamic subsystems and is seen as an important organizing principle of human cognition (Kello, Beltz, Holden, & van Orden, 2007) and brain activation (Kello, Anderson, Holden, & van Orden, 2008). The variability in these studies is analyzed by relating the magnitude of the changes in the system to the frequency with which these changes occur. This is done by running a spectral analysis of the data. When all sizes of change occur equally frequently, the variability pattern is completely random and associated with “white noise.” When the size of changes is fully (inversely) proportional to the frequency with which they occur, we are facing overregulated behavior, referred to as “brown noise.” The optimal type of organization is in between brown noise and white noise. In this situation, pink noise, the scaling relation of the size of the changes, approaches an inversely proportional relationship with the frequency in which they occur ( $1/F$ ). The intriguing observation about  $1/F$  scaling relations is that they are pervasive in all dynamic systems and are even seen as critical for dynamic systems (Kello et al., 2008). Examples are found in patterns of variability in different sizes of avalanches in a pile of rice (Bak, Tang, & Wiesenfeld, 1987), variability of heart rate in healthy individuals (Pagani et al., 1986), variability in repeated arm movements (Diniz, et al., 2011), and variability in word naming tasks (van Orden, Holden, & Turvey, 2003). According to this literature, the study of scaling relations can be used to interpret optimal coordination in dynamic systems. When the variability pattern of heart rate deviates from the optimal  $1/F$  scaling relation, being either too regular or too irregular, there is serious cause for concern, as these deviations are mostly found in patients at high risk of sudden death (Peng et al., 1995). Similarly, the optimal scaling relation of arm movements typically breaks down in people suffering from Parkinson’s disease (van Orden, Kloos, & Wallot, 2009).

An intriguing question is whether such differences in scaling relations can be found between L1 and L2 processing and, if so, whether this is dependent on proficiency levels, speaking conditions, and other factors that may influence the optimal dynamic coordination of the subsystems. These questions have recently been addressed in two studies that have applied spectral analyses to long time series (of at least 512 data points) of L1 and L2 use on short timescales (millisecond changes over a duration of maximally 20 minutes). Lowie, de Bot, and Plat (2014) report on a study in which a single participant repeatedly takes



part in a naming task in the L1 and the L2 over a period of 6 years. Their data show significantly more coordinated behavior in the L1 than in the L2 and show that L2 coordination increases with increased exposure to the L2. Surprisingly, relatively brief periods of temporary immersion in an L2 context negatively affected coordinated behavior in the L1. Similar results with regard to the difference between L1 and L2 processing were found in a study by Waegenmaekers, Lowie, Schoonen, Plat, and de Bot (2014).

Summarizing the evidence from variability studies at different time scales, we may conclude that intralearner variability can indeed be very meaningful. Because the analysis of variability values the dynamic nature of language development, it is an indispensable type of research within the dynamic approach to L2 development. The methods and techniques used to carry out variability analyses are rapidly developing and have now reached a stage of maturity, as the recent research we have reviewed here shows. Where initially the application of variability analyses was limited to the descriptive level, we are now able to evaluate patterns of variability in a meaningful way and use them to explain dynamic relationships between changing subsystems. The current techniques also enable us to test local hypotheses about change over time and changing relationships over time. Advanced techniques of variability analyses are therefore a crucial precursor to the gradual transition from the use of the dynamic systems paradigm as a powerful metaphor to its explanation of empirically observed dynamic relationships.

## **Theoretical and Empirical Implications**

All of the most relevant questions about SLA, including the age issue, L1 influence, individual differences, implicit versus explicit learning, the role of input, intentional versus incidental learning, and of course the order of acquisition of morphosyntax, are implicitly or explicitly about change over time. The reason is that learning is a process, not a product. In the history of SLA, numerous attempts have been made to infer the developmental process from measurements at one or two points in time. A group trend, however significant it may be, cannot accommodate the dynamic multicausality of the emerging language system of an individual, and interpreting a developmental phenomenon on mean trends and variance of group scores at one point in time underestimates the complexity of the developmental process.

Moreover, the conclusions about these group observations are misguided by our unsubstantiated and passionate desire to find universalities in human

cognition, which are supposedly manifested as predetermined stages in development. But similar to the universality fallacies in locomotion and cognition that were exposed by Thelen, Smith, and their colleagues, (e.g., Thelen & Smith, 1994) assumed universalities and predetermined stages in language development are an artifact of the method of investigation. In first language acquisition the idea of an innate UG as a grand scheme and single cause to account for development has had many proponents. However, more and more researchers in this field are starting to realize that an innate system is not only a redundant assumption to account for language development (Lewis & Elman, 2002; O'Grady, 2015), but also that language emerges from continuous interactions on all time scales. In second language acquisition the idea of fixed, predetermined stages as part of a grand scheme and single cause of development stands in sharp contrast with continuity approaches that emphasize the multicausality of language development by complex dynamic interactions. There is no denying that stages can be observed when we take a bird's-eye perspective of group behavior, especially if the groups consist of similar learners that acquire the language in a similar context. But is that relevant in accounting for the development of the individual learner? Another example is the critical period hypothesis for L2 learning. Looking at sufficiently large groups of learners from a sufficiently large distance, there may be a clear cut-off point, for example at age 12, after which it may be problematic to start learning a L2. But at the individual level that cut-off point like any other cut-off point is not predetermined or fixed. In other words, the two perspectives, the dynamic development of the individual and the grand sweep development of groups, can exist side by side. The perspective taken depends on the preferred methodology and the preferred theoretical framework.

The same is true for the morpheme order studies. The overall development, especially of clusters of grammatical morphemes, can clearly be observed by investigating large groups of learners. The observed general order or sequence of acquisition independent of L1 influence, although disputed by some (Luk & Shirai, 2009; Murakami, 2013), should not be confused with the implied cause of this order. The meta-analysis by Goldschneider and DeKeyser (2001) shows the pooled result of 924 subjects in total and reveals the multifactorial forces behind the grand sweep of development: "the combination of perceptual salience, semantic complexity, morphophonological regularity, and frequency does account for a very large portion of the total variance in the accuracy scores for grammatical functors" (pp. 34–35). This is a useful observation as it refines the possible causes of the grand sweep of accuracy scores. But it also clearly indirectly points to the real complexity of the individual's development over

time. Probably all of these factors affect each individual in a different way and to a different extent at different moments in time. Similar to the individuals in the Slavic group of Derwing et al.'s (2006) study, there may be no single individual that shows the exact order of the group. This impression is reinforced by the observations made by, for example, Rosansky (1976) that there was a large difference between the universally observed order and the sequence of acquisition of the individuals in her sample. Real multicausality can only be investigated by including the time dimension, which at the same time blurs the neat two-dimensional results of group studies.

This brings us to another fundamental issue. It has been asked (e.g., in the discussion with the audience at the 2013 *Language Learning* Roundtable where the arguments in this article were first presented) whether and to what extent DST can be considered a theory, because it cannot be falsified. Another concern that is regularly raised is that investigating individual case studies of development over time has limited value, because the findings from these studies cannot be generalized. These issues and concerns are not new and are not the only objections that are raised against the dynamic hypothesis (see van Gelder, 1998, for an overview and a systematic rebuttal of each of these objections). With regard to the unfalsifiability of the dynamic hypothesis as such, it can be said that DST is not very different from other approaches in this sense. UG and behaviorism cannot be falsified either. Yet there are definitely aspects within the dynamic approach that can be tested against empirical observations. One example is dynamic modeling as worked out in Caspi (2010), who demonstrated that dynamic hypotheses can be tested by creating computer simulations and that definitely not all possible scenarios can be accounted for in a dynamic model. If we create an unrealistic parameter setting, it cannot be dynamically modeled. Likewise, if pink noise is a manifestation of the dynamic coordination of subsystems, as is argued by many researchers (e.g., Bak et al., 1987; Kello et al., 2008), the occurrence of pink noise signals the existence of a dynamic system. Because pink noise scaling relations are found in acknowledged physical dynamic systems as well as in cognition, including language behavior, the conclusion can be drawn that language is a dynamic system. In other words, based on DST, specific hypotheses can certainly be drawn up and falsified. The difference with nondynamic approaches is that DST hypotheses will always concern change of systems over time.

Yet, because DST is a broad, all-encompassing, and multidimensional view on reality, it is not falsifiable as a whole. DST is not unique in this respect. Van Gelder (1998) points to the evolutionary hypothesis, which asserts that the development of species is the result of natural selection. Even though this

overarching idea is well accepted, it cannot be falsified in the short term. But if in the end it does not appear to account for the results of the existence of species, the theory will be falsified. The same applies to DST:

[T]he Dynamic Hypothesis will be known to be false if, after an extensive period of investigation, cognitive scientists have in practice rejected dynamical approaches in favor of some other modeling framework. (van Gelder, 1998, p. 659)

For the time being, the argument could be turned around: Aspects of theories like Processability Theory (see Pienemann, 2015; Lenzing, 2015) or UG can be falsified, but only within the two-dimensional scope of the method in which it is investigated. Such a scope is truly limited, as it fails to incorporate potentially nonlinear and dynamically changing relationships. If the strength of a theory is determined by its explanatory power, DST stands by far the best chance of being able to account for the development in the context of complex interactions over time. DST provides an explanation of human behavior that focuses on change over time and is embodied, fully situated, and ecologically valid.

The argument of generalization can also be refuted. As has been argued in the previous section, generalizations are only useful when we are interested in the grand sweep of development. As Thelen and Smith (1994) have demonstrated, generalizations can lead to a predisposed view of reality that claims the existence of universal stages of development that represent at the most a partial truth. At the individual level of development, there are no universal stages. Similar to the observation that results from case studies cannot be generalized, the results from group studies cannot be individualized. Assuming that language learning is fundamentally a highly individual process, the projection of clustered group results onto the individual level in research is flawed (also see Molenaar, 2008).

In addition to these theoretical implications of DST approaches to SLA, there are some methodological implications for the research we carry out. Because DST is inherently about change over time, DST-based research will have to include a time dimension. In agreement with earlier pleas for more longitudinal studies in SLA (Ortega & Byrnes, 2008; Ortega & Ibarra-Shea, 2005), longitudinal studies will have to be conducted to create meaningful time series. The number and density of the observations in the time series will have to be adjusted to the expected time scale of development. When the focus is on early learning, dense data over a relatively short period of time may be sufficient. When the focus is on L1 attrition by elderly participants, it may be

advisable to measure less frequently, but over a longer period. When the focus is on real-time processing, very dense data will have to be gathered (every second) in a relatively short period. For variability analyses, an extensive set of data points will have to be gathered, which can be problematic or unfeasible. While traditional statistics are not very suitable to analyze complex longitudinal data, nonlinear methods of analysis can be used such as dynamic variability analysis and dynamic modeling (see Verspoor, Lowie, & de Bot, 2011) or spectral analyses (Lowie et al., 2014). Although the methodologies for dynamic investigations are rapidly developing, more methods will have to be developed by which multiple data points can be gathered in an ecologically valid context and which include many different aspects of embedded and embodied language use. This may currently be the biggest challenge for the development of dynamic research.

However, the type of DST studies we have presented also has limitations. The methods and techniques require a dense corpus of quantifiable variables and not all SLA research lends itself to this. Moreover, it is impossible to capture all relevant variables in one longitudinal study. Similar to group studies, DST studies can only focus on a limited number of variables and additional research is needed to complement our findings or give us theoretical insights to investigate further. O'Grady (2015) zooms in on the role of processing and argues that the learner does not need an innate grammar but can detect regularities on his or her own. This is clearly a bottom-up process involving various interacting variables such as input and exposure playing a role in the emergence of a particular construction. It would be interesting to follow a few children over the course of a few years on the use of the pronouns O'Grady discusses and discover their actual process of acquisition. Eskildsen (2015) zooms in on a number of interacting variables such as the context with affordances, uses, and gestures needed to acquire various constructions and, in line with our findings, shows that there is not a neat sequence of acquisition, but that various targetlike and nontargetlike constructions cooccur from the very beginning and slowly give way to more targetlike constructions. The theories presented by Pienemann (2015) and Lenzing (2015), however, are not compatible with DST. Even though the presumed stages may be in line with the empirical data, both positions assume a genetically programmed sequence of predetermined stages, resulting in rule-like behavior with hypothesis testing, which is not needed in a dynamic, usage-based, emergentist account.

Ellis (2015) argues that, although the four different accounts of the sequence of acquisition of negative verb constructions start from different models, they do not actually contradict what he calls Long's (1990) law. We agree. But we

also think that it takes insufficient account of an important factor—variability. If we seek to present the findings of SLA in such a way that they are accessible to teachers and teacher educators, we must simplify. Therefore, we can tell them that, even though there are similar sequences for all learners, not a single learner will follow these exactly, and variability is functional, especially for the early stages. While learning, learners will have targetlike forms one day and not the next. They may even have peaks of overuse in nontargetlike forms after they have already shown some use of targetlike forms. This variability is part and parcel of the learning process and teachers should not try to eliminate it.

Zhang and Lantolf (2015) show that the teachability of language is not limited to fixed and predetermined sequences. Using meaning-focused instruction organized according to Vygotskian principles of developmental education, the participants in their study successfully learned grammatical forms that were well beyond their expected developmental stage. This observation is fully compatible with a DST approach to language development. Providing meaningful instruction that requires active engagement may lead to perturbations of the system. Following a perturbation, the system may reorganize to incorporate the new information. The effect of a perturbation depends on the state of the emerging dynamic network of coupled variables rather than on fixed cognitive processing constraints.

## Conclusion

L2 development is a process. The influential morpheme order studies that were first conducted in the 1970s and that were followed up over subsequent decades seem to acknowledge this observation by focusing on the order or the sequence of acquisition. The main conclusion from an extensive series of such studies is that, in spite of some local and individual irregularities, there is evidence for a universal and predetermined order of acquisition of English grammatical morphemes, regardless of the learner's L1. And although there has been some discussion about the methods used and about the possible causes of the observed order, the existence of the order is generally accepted as a robust outcome of L2 research. In this article we have argued that this claim may be true if we look at the grand sweep of phenomena but is seriously flawed in several ways if we seek insight into the process of development. An important flaw is that the orders are often inferred from accuracy scores instead of longitudinal observations. But more importantly, the orders are based on clustered observations of mean scores for groups of learners. A universal order would imply that the order is a valid representation of the developmental trajectory of individual learners. We have argued that generalized group means

can never make this claim, because the method does not acknowledge the dynamic interconnectedness and the embodied nature of language development. Language development is an inherently individual and dynamic process and there can be no logical expectation that the pattern found in generalizations at the group level is the same as the actual development of the individual learner. As we have pointed out by referring to data of other studies, there may not be a single individual whose development is identical to the presumed universal sequence. Our collective desire to organize cognitive development in neat and well-defined stages is found at many levels of research, from Piagetian stages of cognitive development to ideas about age effects in L2 learning, developmental stages in terms of processability, and morpheme orders. But given the complex dynamic nature of real development, the conclusions about the actual existence of these predetermined stages at the individual level are not warranted.

We can only make the observations our method allows us. This implies that we should adjust our method of investigation to the phenomena and questions we are interested in within the context and timescale of our focus. When we are interested in the global patterns of L2 learning and linearly related factors that affect it at one moment in time, group studies using Gaussian statistics are the most appropriate method. But if we want to understand the multicausality of the developmental process, we will need to use longitudinal case studies with frequent measurements over a long period of time and use variability as our measure of meaningful, context-dependent change and interaction among multiple subsystems and forces. Just as the seemingly robust stages of the A-not-B error turn out to be fuzzy and time and context dependent after scrutinizing contextualized development at the individual level, the robustness of globally observed orders of acquisition, critical periods, and stages of development may turn out to be indistinct and ill-defined at the individual level. The observed individual variation and the inconsistent outcomes of the morpheme order studies at least point in that direction.

Final revised version accepted 2 October 2014

## Note

- 1 In this article we have used the term “variability” for intra-learner variability over time and “variation” for inter-learner variation.

## References

- Bak, P., Tang, C., & Wiesenfeld, K. (1987). Self-organized criticality: An explanation of  $1/f$  noise. *Physical Review Letters*, *59*, 381–384.

- Bassano, D., & van Geert, P. (2007). Modeling continuity and discontinuity in utterance length: A quantitative approach to changes, transitions and intra-individual variability in early grammatical development. *Developmental Science*, *10*, 588–612.
- Brown, R. (1973). *A first language: The early stages*. Cambridge, MA: Harvard University Press.
- Cancino, H., Rosansky, E., Schumann, J., & Hatch, E. (1978). The acquisition of English negatives and interrogatives by native Spanish speakers. In E. Marcussen Hatch (Ed.), *Second language acquisition: A book of readings* (pp. 207–230). Rowley, MA: Newbury House.
- Caspi, T. (2010). *Dynamic aspects of lexical development*. Unpublished doctoral dissertation, University of Groningen, Netherlands.
- de Bot, K., & Gullberg, M. (Eds.). (2010). *Gestures in language development*. Amsterdam: John Benjamins.
- de Bot, K., Lowie, W., & Verspoor, M. (2007). A dynamic systems theory approach to second language acquisition. *Bilingualism*, *10*, 7–21.
- Derwing, T., & Munro, M. J. (2013). The development of L2 oral language skills in two L1 groups: A 7-year study. *Language Learning*, *63*, 163–185.
- Derwing, T., Thomson, R. I., & Munro, M. J. (2006). English pronunciation and fluency development in Mandarin and Slavic speakers. *System*, *34*, 183–193.
- Diedrich, F. J., Highlands, T. M., Spahr, K. A., Thelen, E., & Smith, L. B. (2001). The role of target distinctiveness in infant perseverative reaching. *Journal of Experimental Child Psychology*, *78*, 263–290.
- Diniz, A., Wijnants, M. L., Torre, K., Barreiros, J., Crato, N., Bosman, A. M. T., et al. (2011). Contemporary theories of 1/f noise in motor control. *Human Movement Science*, *30*, 889–905.
- Dörnyei, Z. (2009). Individual differences: Interplay of learner characteristics and learning environment. *Language Learning*, *59*(S1), 230–248.
- Ellis, R. (1994). *The study of second language acquisition*. Oxford, UK: Oxford University Press.
- Ellis, R. (2015). Researching acquisition sequences: Idealization and de-idealization in SLA. [Special Issue]. *Language Learning*, *65*, 181–209.
- Eskildsen, S. W. (2012). L2 negation constructions at work. *Language Learning*, *62*, 335–372.
- Eskildsen, S. W. (2015). What counts as a developmental sequence? Exemplar-based L2 learning of English questions. [Special Issue]. *Language Learning*, *65*, 33–62.
- Gatbonton, E. (1978). Patterned phonetic variability in second language speech: A gradual diffusion model. *Canadian Modern Language Review*, *34*(3), 335–347.
- Goldschneider, J., & DeKeyser, R. (2001). Explaining the “natural order of L2 morpheme acquisition” in English: A meta-analysis of multiple determinants. *Language Learning*, *51*, 1–50.
- Hakuta, K. (1976). A case study of a Japanese child learning English as a second language. *Language Learning*, *26*, 321–351.



- Kello, C. T., Anderson, G. G., Holden, J. G., & van Orden, G. C. (2008). The pervasiveness of 1/f scaling in speech reflects the metastable basis of cognition. *Cognitive Science*, 32, 1217–1231.
- Kello, C. T., Beltz, B. C., Holden, J. G., & van Orden, G. C. (2007). The emergent coordination of cognitive function. *Journal of Experimental Psychology: General*, 136, 551–568.
- Konner, M. (1991). Universals of behavioral development in relation to brain myelination. In K. R. Gibson & A. C. Petersen (Eds.), *Brain maturation and cognitive development: Comparative and cross-cultural perspectives* (pp. 181–223). Piscataway, NJ: Transaction Publishers.
- Larsen-Freeman, D. (1997). Chaos/complexity science and second language acquisition. *Applied Linguistics*, 18, 140–165.
- Larsen-Freeman, D., & Cameron, L. (2008). *Complex systems and applied linguistics*. Oxford, UK: Oxford University Press.
- Lenzing, A. (2015). Exploring regularities and dynamic systems in L2 development [Special Issue]. *Language Learning*, 65, 89–122.
- Lewis, J. D., & Elman, J. L. (2002). Learnability and the statistical structure of language: Poverty of stimulus arguments revisited. In B. Skarabela, S. Fish, & A. H.-J. Do (Eds.), *Proceedings from the 26th Annual Boston University Conference on Language Development* (Vols. 1–2, pp. 359–370). Somerville, MA: Cascadilla Press.
- Long, M. H. (1990). The least a second language acquisition theory needs to explain. *TESOL Quarterly*, 24, 649–666.
- Lowie, W. (forthcoming). L2 phonological development: A plea for a dynamic, process based methodology.
- Lowie, W., de Bot, K., & Plat, R. (2014). Pink noise in production: A nonlinear approach to the multilingual lexicon. *Ecological Psychology*, 26, 216–228.
- Lowie, W. M., & Seton, B. J. (2013). *Essential statistics for applied linguistics*. Basingstoke, UK: Palgrave Macmillan.
- Lowie, W., & Verspoor, M. (2011). The dynamics of multilingualism: Levelt's speaking model revisited. In M. S. Schmid & W. Lowie (Eds.), *Modeling bilingualism: From structure to chaos* (pp. 2067–2088). Amsterdam: John Benjamins.
- Luk, Z. P., & Shirai, Y. (2009). Is the acquisition order of grammatical morphemes impervious to L1 knowledge? Evidence from the acquisition of plural -s, articles, and possessive 's. *Language Learning*, 59, 721–754.
- Molenaar, P. C. M. (2008). On the implications of the classical ergodic theorems: Analysis of developmental processes has to focus on intra-individual variation. *Developmental Psychobiology*, 50, 60–69.
- Molenaar, P. C. M., Huizenga, H. M., & Nesselrode, J. R. (2003). The relationship between the structure of interindividual and intraindividual variability: A theoretical and empirical vindication of developmental systems theory. In U. M. Staudinger &

- U. Lindenberger (Eds.), *Understanding human development: Dialogues with lifespan psychology* (pp. 339–360). New York: Springer.
- Murakami, A. (2013). *Individual variation and the role of L1 in the L2 development of English grammatical morphemes: Insights from learner corpora*. Unpublished doctoral dissertation, University of Cambridge, UK.
- O'Grady, W. (2015). Processing determinism [Special Issue]. *Language Learning*, *65*, 6–32.
- Ortega, L., & Byrnes, H. (Eds.). (2008). *The longitudinal study of advanced L2 capacities*. New York: Routledge.
- Ortega, L., & Ibarra-Shea, G. (2005). Longitudinal research in SLA: Recent trends and future directions. *Annual Review of Applied Linguistics*, *25*, 26–45.
- Pagani, M., Lombardi, F., Guzzetti, S., Rimoldi, O., Furlan, R., Pizzinelli, P., et al. (1986). Power spectral analysis of heart rate and arterial pressure variabilities as a marker of sympatho-vagal interaction in man and conscious dog. *Circulation Research*, *59*, 178–193.
- Peng, C., Havlin, S., Hausdorff, J., Mietus, J., Stanley, H., & Goldberger, A. (1995). Fractal mechanisms and heart rate dynamics: Long-range correlations and their breakdown with disease. *Journal of Electrocardiology*, *28*, 59–65.
- Pica, T. (1983). Adult acquisition of English as a second language under different conditions of exposure. *Language Learning*, *33*, 465–497.
- Pienemann, M. (2015). An outline of processability theory and its relationship to other approaches to SLA. [Special Issue]. *Language Learning*, *65*, 123–151.
- Rosansky, E. J. (1976). Methods and morphemes in second language acquisition research. *Language Learning*, *26*, 409–425.
- Smith, L. B. (2005). Cognition as a dynamic system: Principles from embodiment. *Developmental Review*, *25*, 278–298.
- Smith, L. B., & Thelen, E. (2003). Development as a dynamic system. *Trends in Cognitive Sciences*, *7*, 343–348.
- Smith, L. B., Thelen, E., Titzer, R., & McLin, D. (1999). Knowing in the context of acting: The task dynamics of the A-not-B error. *Psychological Review*, *106*, 235–260.
- Spoelman, M., & Verspoor, M. (2010). Dynamic patterns in the development of accuracy and complexity: A longitudinal case study on the acquisition of Finnish. *Applied Linguistics*, *31*, 532–553.
- Thelen, E., Schöner, G., Scheier, C., & Smith, L. B. (2001). The dynamics of embodiment: A field theory of infant perseverative reaching. *Behavioral and Brain Sciences*, *24*, 1–86.
- 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., Verspoor, M. H., & Lowie, W. (2011). Variability and DST. In M. Verspoor, K. de Bot, & W. Lowie (Eds.), *A dynamic approach to second language development: Methods and techniques* (pp. 55–84). Amsterdam: John Benjamins.
- van Geert, P. (1991). A dynamic systems model of cognitive and language growth. *Psychological Review*, 98, 3–53.
- van Geert, P. (2014, March). *L2 acquisition and the study of change in complex systems: A perspective on universals and variability*. Paper presented at the Invited Colloquium, *From Universality to Variability in Second Language Development*, K. de Bot and D. Larsen-Freeman (Converners). American Association for Applied Linguistics Conference, Portland, OR.
- van Gelder, T. (1998). The dynamical hypothesis in cognitive science. *Behavioral and Brain Sciences*, 21, 615–656.
- van Lieshout, P. (2004). Dynamical systems theory and its application in speech. In B. Maassen, R. Kent, H. Peters, P. van Lieshout, & W. Hulstijn (Eds.), *Speech motor control in normal and disordered speech* (pp. 51–82). Oxford, UK: Oxford University Press.
- van Orden, G., Holden, J., & Turvey, M. (2003). Self-organization of cognitive performance. *Journal of Experimental Psychology: General*, 132, 331–350.
- van Orden, G. C., Kloos, H., & Wallot, S. (2009). Living in the pink: Intentionality, well-being, and complexity. In C. Hooker (Ed.), *Handbook of the philosophy of science, Vol. 10: Philosophy of complex systems* (pp. 639–683). Amsterdam: Elsevier.
- Verspoor, M., Lowie, W., & de Bot, K. (Eds.). (2011). *A dynamic approach to second language development: Methods and techniques*. Amsterdam: John Benjamins.
- Verspoor, M. H., Lowie, W., & van Dijk, M. (2008). Variability in second language development from a dynamic systems perspective. *Modern Language Journal*, 92, 214–231.
- Waegenmaekers, E., Lowie, W., Schoonen, R., Plat, H., & de Bot, K. (2014). The dynamics of bilingual lexical processing. Manuscript submitted for publication.
- Yuldashev, A., Thorne, S. L., & Fernandez, J. (2013). Second language learners' contiguous and discontinuous multi-word unit use over time. *Modern Language Journal*, 97(S1), 31–45.
- Zhang, X., & Lantolf, J. P. (2015). Natural or artificial: Is the route of L2 development teachable? [Special Issue]. *Language Learning*, 65, 152–180.