Chapter 3 Investigating language development from a dynamic perspective

3.1 Introduction
As the previous chapter has shown, the dynamic approach diverges from the prevailing methodology in applied linguistics by focusing on research questions that concern individual pathways of growth and componential interaction rather than linear cause-and-effect. Therefore, before describing the two empirical studies that are the core of this dissertation, this chapter explains some basic methods involved in DST-oriented research. It begins by addressing the affinity of the dynamic approach with case study methodology, and then focuses on variability analyses and mathematical modeling. The chapter is rather technical and detailed in nature, because at this stage the empirical dynamic approach is rather a “do it yourself” methodology – there is no conveniently assembled PC-package for dynamic analyses (although, rather exceptionally, the study described in Chapter 4 uses a model which has been preprogrammed by Paul van Geert). Moreover, there are as of yet no clear guidelines for applying DST in linguistic studies (but see van Dijk, 2003; van Geert & van Dijk, 2002; van Geert, 1994). Therefore many of the practices associated with the empirical dynamic approach are rather novel and daunting, not just in terms of exact procedures, but also conceptually. While this dissertation aims for transparency in explaining the procedures that it applies, their descriptions in this chapter may be too abstract out of the context of an empirical study. Thus as mentioned, this chapter may be cross-referenced from the following chapters, rather than read in entirety.

3.2 General study design
A longitudinal and detailed case study design is highly suitable for investigating SLA from a dynamic perspective, because it facilitates inspection of development as temporal change, and enables the relation and comparison of such change across various levels of the data. Case studies are used extensively in educational and developmental psychology; yet despite the proximity of these fields to applied linguistics, they are nowhere near as prominent in its literature (Meara, 1995). As van Lier points out,
Case study methodology has been extremely influential in shaping the way we talk about education, yet it has been traditionally regarded as somewhat of a soft and weak approach when compared to studies that have been deemed more rigorous, randomized, or experimental in nature (2005, p. 195).

Singer and Willet (2003) suggest a framework for case study methodology, which has been adopted by the studies in this thesis. It is comprised of six basic stages:

- Data collection
- Data description
- Data exploration
- Model specification
- Model fitting (assessment and possible optimization)
- Considering extensions

The following subsections describe each stage in the context of the dynamic approach. Sections 3.3 and 3.4 then lists the practical procedures that the stages entail.

### 3.2.1 Data collection

Case studies require longitudinal, systematic, and detailed data gathering. In choosing this data, “one can – indeed, one needs to – foreground a focal point, while allowing the background to continue on its dynamic trajectory” (Larsen-Freeman & Cameron, 2008, p. 234). In other words, suitable data need to be extracted from specific areas of language development, a process requiring the researcher to “determine the ecological circuit in which one is interested” (Larsen-Freeman & Cameron, 2008, p. 235). In the studies included in this dissertation, the focus is increasingly narrowed until it zooms in on specific parameters representing the phenomena of interest. For example, the writing study in Chapter 5 is concerned with text-level performance, within which it distinguishes the categories of complexity and accuracy, further specified in both the lexical and syntactic dimensions. In each of these categories, the study singles out specific representative measures.

### 3.2.2 Data description and exploration (variability analyses)

A dynamic standpoint “requires us to look for change and for processes that lead to change, rather than for static, unchanging entities. Furthermore, data are not cleaned up before analysis to get rid of inconvenient ‘noise’” (de Bot et al., 2007, p. 16). The second stage, data description, thus involves inspecting raw data as change trajectories over time. Plotting data values as a function of time visually displays the route of their development, including local dips and jumps. It also enables to compare
developmental trajectories across various components of language (or across participants, cf., Larsen-Freeman, 2006b). Often linear or other central trends are added to the data in order to visualize the overall direction of its growth. Variability around this trend can then be distinguished and inspected in further detail.

This stage, data exploration, involves investigating the variability patterns in the data in a similar manner, as trajectories of change. Like the raw data, variability can be inspected in a single component of language, or compared across several components. The two procedures correspond with two emphases on the value of variability, as discussed in section 2.1.1.6. First, variability in the development of one component can be analyzed from a general evolutionary or developmental perspective. In such cases, increased variability is considered a predecessor to qualitative shifts, which are frequently misidentified as discrete developmental jumps rather than the outcomes of continuous change (Ruhland, 1998; van Dijk, 2003). From this angle, variability has a facilitative role in development as enabling systemic adaptation. Second, variability patterns can also be compared across systemic components. In this context, shifts in variability are seen as indicators of change in underlying systemic interactions. The interpretation of findings from the analyses described so far may then be incorporated in a mathematical model, which is also based on the structural ordering of the data, and is therefore theoretically motivated.

3.2.3 Model specification
A dynamic model is comprised of two parameter types. The first is order parameters, which define the number of its components (growers) and their hierarchy, that is which growers act as precursors, and which growers are dependent on the former. The second type is control parameters, which correspond with factors that influence the course of development in each grower. In other words, order parameters define the structure of the model, while control parameters are “those aspects that cause the process to behave as it does” (van Geert, 2003, p. 663). Control parameters can in turn be divided into property and relational parameters. Property parameters specify the initial value, growth rate, carrying capacity (optimal attainable value), any delay in onset, and in some cases the amount of random variation in each grower. Relational parameters specify the interactions between a given grower and other growers in the model. They can be further distinguished as support, competition, and a conditional threshold that enables the onset of growth in a dependent grower and thereby also the
onset of competition and/or support towards it from other growers. This threshold value thus corresponds with the hierarchical order of the model, and relates to the role of the particular grower as a precursor. Any model of a given grower can incorporate these three parameter types as they refer to one or more other growers in the system, as well as interactions generated by these growers towards the specific grower. However, not all of these parameters need to be specified, and their inclusion in a model depends on the particular phenomena that it is intended to simulate. Both support and competition can be specified as either by level or by change interactions. By level means that their value changes in direct relation to the value of the grower that generates them; by change implies that their value is relative to the amount of change in the grower between the current and previous time points (see section 3.4 for the equations depicting these specifications).

No phenomena can be modeled without relying on theory, empirical findings, or both for defining the model parameters and setting their values. The model configuration is a means of hypothesis testing, since it evaluates both the interpretation of the data analyses and the structural theory that refers to it. In each of the two studies in this dissertation, which apply dynamic models to L2 vocabulary knowledge and writing performance, the model settings are informed by the relevant background literature and preceding findings. Section 3.4 introduces some dynamic models which have been previously applied to (predominantly) L1 data and serve as basis for those used in the current studies.

3.2.4 Model fitting
The fit of the model to the data can be evaluated visually, in which case it should be “convincingly similar to the outcomes of the real-world system” (Larsen-Freeman & Cameron, 2008, p. 41). Additionally, the model can be assessed by a parameter that denotes its goodness-of-fit. Some fit assessments are regression-based, comparing the linear regressions of the model to those of the data, correlating the model with the data values, or comparing the internal pattern of correlations in the model with that of data. Others are based on iterative weighted least squares. The present studies incorporate both methods, as suggested in previous applications of dynamic models to developmental data (van Geert & Steenbeek, 2005).

A model can also be fitted to data by optimization procedures, which arrive at the optimal configuration of its control parameter values while maintaining its original
structure as specified in the order parameters. The optimized parameter values can then be compared to the interpretation of the data analyses and the pertinent theory, while the model outcome is again compared with the data. Thus, the hypotheses derived from both theory and data interpretation, which have informed the model configuration, can be supported further, if the optimization shows that the best fit retains parameter values that are congruent with them. On the other hand, if the best fit is achieved by altering these parameters, these hypotheses will be refuted.

### 3.2.5 Considering extensions

In summary, an empirical dynamic study is comprised of three main stages. First, the data description determines a basic structural hierarchy (number and order) of components. Second, the data exploration may indicate the nature of the interactions (relational control) within this hierarchy. Third, the model configuration and optimization may test both the primary and secondary hypotheses (concerning the order and relational control parameters), which are based on the theory of the field and on the two preceding procedures. These steps constitute a kind of “three-in-one” methodological design. In such a study, the detailedness of the results and the need for its visual representation would require some form of accompanying, rather than retrospective, interpretation and discussion.

At any of the three stages, additional research questions can be posed with regard to the nature of development and its underlying mechanisms. Finally, depending on the outcomes of this methodology, extensions to the data collection, the analyses or the model should be considered, as in any other empirical study. In the dynamic approach this last stage is particularly important, since the focus on longitudinal and individual data cannot be generalized to the overall population.

### 3.3 Methods of growth and variability analysis

Several techniques have been used by dynamic-oriented studies to represent and inspect intra-individual growth and variability. Van Geert and van Dijk (2002) and van Dijk (2003) review a number of such measures, among them the moving range (min-max), which visualizes local variability peaks that precede developmental jumps. However, since the current studies investigate variability as indicative of systemic self-organization, the following sections describe only techniques that pertain to this role and are featured in these studies. These are trajectory plots, residual plots and moving correlations. Additionally, the studies use a technique of
local data smoothing called spline interpolation, which incorporates a certain amount of variability with locally regressed trend.

### 3.3.1 Growth trajectory plots
A trajectory plot displays change as a sequence of values plotted along the x-axis, which denotes time. Such a simple graphic presentation of data can convey meaningful information about the nature of its development. For example, a growth trajectory plot can illustrate the degree of consolidation of a newly-acquired linguistic feature, with the trajectory representing change in the amount of its usage. Adding a regression line to the plotted trajectory shows the overall direction (increase vs. decrease) of development, and visualizes the differences between the raw data values and their central trend.

Each segment in a growth trajectory is characterized by its degree of change in relation to adjacent segments. The plotted trajectories show the patterns of change between observations. Trajectories are more informative when presented in conjunction with those of other data variables: plots of two or more trajectories can expose their interactions over time. If shifts in one trajectory tend to coincide with parallel or inverse shifts in the other, this may be interpreted as indicative of a supportive or competitive interaction between them, respectively. However, such an interpretation would be speculative, and needs to be reaffirmed by subsequent procedures.

Figure 2 contains two growth trajectories, one of which shows a general linear increase, and the other a decrease. Both trajectories exhibit a high amount of variability around these trends. The patterns of this variability are mostly parallel, but peaks (outliers) in one index are usually accompanied by dips in the other (for example in weeks 6 and 10). The potential relatedness between the variability patterns of the two indexes can be inspected in further analyses, such as de-trending the data values and plotting their residuals.
3.3.2 De-trended data values (residuals)

De-trending data allows for inspecting variability independently of the linear trend. The residuals, or de-trended values, are obtained by subtracting the linear trend from the data series. The trend is calculated for each data value as the sum of the intercept and slope of the entire data series, multiplied by the number of measurements until that particular time point. The residuals can then be plotted and compared between two or more data variables. Residuals can also be correlated in static or moving correlations (see the following section).

The advantages of de-trending data are that local divergences from the trend can be revealed and displayed as a function of time. Thus, even when the overall trend is robust (i.e., strong increase or decrease, manifested in steep slope values), the residuals can reveal temporal patterns in the departures from it. Another advantage is that residuals may reveal the concurrence of such local fluctuations with similar or inverse patterns in another index. In this way, residual plots can indicate local interactions which may be obscured by the overall trend when it is incorporated in the data. For instance, while the central trends of two variables may be similar, their residuals may exhibit inverse patterns, implying a potentially competitive interaction. Figure 3 depicts the residuals of two indexes, showing their relative patterns.
3.3.3 Moving correlation
A common procedure in L2 research is correlating two developing indexes. Such a static correlation coefficient can be supplemented with a moving correlation, which shows temporal changes in the coefficient values in a moving window of several observations. Each window overlaps with the preceding window on all but the first measurement value. For example, in Figure 4, the first window features the correlation coefficient value in weeks 1-5, the second includes the coefficient in weeks 2-6, and so forth. Thus the correlation can be viewed as a function of time (van Geert & van Dijk, 2002). The moving correlation in Figure 4, changing over a 36-week period divided into overlapping windows of 5 measurements, shows repeated alternations between strong positive and weak negative coefficient values. If the interaction between the two variables was only summarized as a single coefficient value, these shifts would be obscured. Moreover, if the correlation is not statistically significant, it would likely not be noted, whereas the moving correlation shows that there might be a systematic pattern that underlies this correlation. While this pattern may prevent the correlation from becoming sufficiently high to reach significance, it may nonetheless be informative with regard to temporal changes in the interaction between the two variables.
Investigating language development from a dynamic perspective

Figure 4. A moving correlation between two indexes of L2 writing performance

Like the trajectory and residual plots, moving correlation plots can include more than one correlation, allowing for a visual comparison of the temporal patterns in one correlation with those in another. It should be noted that all of these techniques (and other forms of variability analyses which are not included in this thesis) are complementary. Using them conjointly with linear analyses can achieve a fuller representation of the developmental phenomena.

3.3.4 **Smoothing by local regression: spline interpolation**
Smoothing by local regression combines local trend with part of the data variability. It thus bridges central trend analyses like linear or polynomial regression with variability analyses. Local smoothing can be performed in several ways. The simplest and perhaps most common technique is a moving average, which is simply an average calculated in partially overlapping time (or measurement) windows of a fixed size, like the moving correlation technique. Although the moving average is very popular in longitudinal studies, another smoothing method, the spline function, is preferred in the current thesis. This is because in a moving average, the past (previous values) is more influential than the future (upcoming values). While this pitfall can be corrected by weighing, moving averages are also more susceptible to the influence of extreme values. Finally and perhaps obviously, averages do not always represent actual data values (there are no families with 2.5 children, for example); therefore they are a rather crude means of representing trend.\(^3\)

---

\(^3\) At the other end of the spectrum of smoothing methods is a B-spline interpolation, which is highly sensitive to local values but requires complex recalculations that are computationally costly, and is therefore not featured in this dissertation.
The term spline refers to a wide class of piecewise polynomial functions used to minimize data roughness by interpolation. Because spline functions are segmented, they avoid the high oscillations produced by polynomial functions. The natural cubic spline, which is the most popular spline function, is termed as such because it provides the closest approximation to the curve produced by the original spline device used in ship building. In this function, the data is divided into three segments, for each of which the equation \( \tilde{r}(t) = r_0 + at + bt^2 + ct^3 \) (the cubic polynomial) is calculated with different parameter values. The intersections between these segments are called knots. Continuity between the polynomial functions is maintained by matching their derivatives at each knot.

In the equation, \( t \) represents the time dimension, \( r \) represents the rate (configured as the onset data value at point 0, and the last value of the preceding section in the data for each knot), and the parameters \( a, b, \) and \( c \) comprise the cubic polynomial. After each knot, a new cubic polynomial is calculated, with the cumulative number of data points subtracted from the time term. For example, in the second segment of the spline in Figure 5 (below), \( t \) is the current week number minus the total number of weeks in the previous segment (12 and 24 for the first and second segments, respectively). In the plot generated by this equation (see Figure 5), the rate sub-value of the first segment is the y-axis intercept.

In this dissertation, spline functions are visually displayed in conjunction with optimized model outcomes (as well as with raw data values and their linear trends). This facilitates comparing the model plots to the data that they simulate, since optimized models cannot include random variation, and are therefore predictably smoother.
Data

Spline

Figure 5. Raw data and natural cubic spline for four vocabulary knowledge levels

3.3.5 Pitfalls of variability analyses
Investigating variation poses several problems. The first is that estimations of variability, even more than those of central tendency, require substantial sample sizes (Bates et al., 1995). For reasons of labor intensity, such samples, specifically in language development studies, can only be extracted from a small number of case studies. Thus, case studies in general, and those involving variability analyses in particular, are usually investigations of single learners (e.g., van Dijk, 2003; see also Meara, 1995).

A second problem concerns the choice of indexes for variability analyses. Averaging is inherent to many measures of language development, for instance to the widely-used index of mean utterance length (MLU), as its name attests. The representation of variability may be compromised if it is investigated in such averaged data values. A plausible solution for this problem is employing indexes that are based on ratios, rather than averages.
A more-general problem is a risk of “missing the forest for the trees”, while closely monitoring variability in a restricted number of variables. As Ellis warns, “it’s not enough to highlight individual variability (...). We still have to explain the regularities. And if we find it difficult to credit these as innately given, then we have to come up with some viable alternative, for we know that input will not suffice” (2007, p. 23). However, when variability analyses are used to supplement rather than replace central trend analyses, the problem of over-emphasizing irregularities is minimized. Moreover, both types of data analyses can be used in conjunction with models to explain general growth patterns of linguistic data, given unspecified input (cf., Bassano & van Geert, 2007). Therefore, at least in terms of intention (if not always execution), the dynamic perspective on language development can be aligned with more established linguistic approaches and their corresponding methodologies.

Variability analyses can provide valuable information about processes of language development, but usually require support by further procedures. Some studies use permutation techniques, such as a Monte Carlo simulation, to corroborate their interpretations of variability patterns as developmental indicators (cf., van Dijk, Verspoor, & Lowie, in press). As mentioned, such interpretations of variability analyses can also be supported (or contradicted) by mathematical models. The practical side of this approach, which is taken up by the current studies, is discussed in detail in the following sections.

### 3.4 Modeling a dynamic system: general considerations

Dynamic models are mathematical descriptions of the parameters that determine the development of a system. The values attached to these parameters are not meaningful in themselves. For example, an initial growth rate of 0.01 does not imply that the phenomenon at hand actually grows at this precise pace in real life. Rather, these values are relative to each other, so that if the value of a given grower increases at double the rate of another, this fact is meaningful. However, it should be kept in mind that only basic models or artificial phenomena rely on a fixed growth rate. In dynamic models, the growth rate is an onset value, which changes iteratively as a function of the interactions between the growers (relational control parameters). This interaction in turn is a function of both growth and carrying capacity (implied by resource limitations). In general, the configuration of order parameters relies more heavily on

---

4 as well as any other (property) control parameters, such as developmental delay
theory, whereas that of control parameters is also based on outcomes of data analyses. Thus constructing a mathematical model – both the combination of mathematical operators, which denote the order parameters, and the configuration of their values, which denote the control parameters – draws on both theory and findings. This does not mean that the theory need always be established and consensual. In fact, models can act as an exploratory tool that can substantiate less popular theories and show which axiomatic ones can benefit from revision. The following sections explicate several generic dynamic models, which have been so far applied to (mostly) L1 data. These models serve as a basis for the models used in the studies described in the subsequent chapters.

3.4.1 A basic growth model

Section 2.1.1.1 has shown how the logistic growth equation can replicate the typical sigmoid “learning curve”. This curve can be observed across numerous developmental phenomena, and is characterized by rapid initial growth, followed by a plateau. The logistic equation specifies growth as a joint product of the current value (growth level) of the grower and the available internal and external systemic resources. In other words, the equation combines exponential iterative growth with a delimiting factor, which renders it nonlinear. Equation 1 (below) is the difference form of the logistic equation as it describes the value $L_{t+\Delta t}$ of a grower $L_t$.

$$L_{t+\Delta t} = L_t + r \Delta t (1 - L_t / K)$$

Equation 2. A logistic growth function (after van Geert, 1994)

As previously mentioned, the original logistic equation, known as the Verhulst formula, was used to depict population dynamics on the basis of two central premises: that the rate of growth will always be proportional to the current value, and that it will also depend on the amount of available resources. Although these resources are not directly configured in the equation, they are implied through the $K$ parameter, which signifies the carrying capacity. The growth rate, typically marked as $r$, is the proportional increase in one time unit. $\Delta$ is the general difference symbol, which can also denote change, in this case change in time. Thus, the value $L_{t+\Delta t}$ is based on the preceding value $L_t$. 
Van Geert (1991) used the difference form of the logistic growth equation to model the growth of early L1 lexicon. Figure 6 shows the growth curve obtained from observations of a child’s lexical production together with a logistic growth function.

Figure 6. Lexical data (derived from Dromi, 1986, as cited in van Geert, 1994) vs. a logistic model. Based on van Geert (1994)

Depending on the value of the growth rate $r$, this function can yield three distinct growth patterns. The first is S-shaped increase towards a maximal value; the second is cyclic oscillations between various states (as shown in the bifurcation plot in Figure 1); and the third is irregular and seemingly-unpredictable oscillation. What determines the growth rate and its consequent effect on the shape of development are not only the current value of the grower (that is, its distance from the carrying capacity), but also various influences that relate to the process at hand, which can be expressed in different extended versions of the logistic equation. In construing these versions, the basic logistic function can be expanded by a range of terms that make the model more complex and specific, thus yielding a better fit. For example, van Geert (1993) points out that to successfully simulate real-life data, it is often necessary to add a feedback delay or resource oscillations as a further damping factor on the growth rate $r$, which determines the steepness of the S-shaped curve. Such delays or oscillations can be directly added to an equation of a single grower, or can be implied in coupling this equation with an equation that describes another grower.

With regard to the second option, all available resources are usually treated as a single and unchanging entity in dynamic models. However, as the previous chapter has discussed, the availability of resources is a function of interactions between co-
developing components of the system, and no grower is isolated (although, of course, it may be addressed as such from theoretical or methodological considerations). Changes in the amount of resources available to a specific grower therefore reflect systemic self-organization in terms of internal interaction and subsequent reallocation of resources. From this point of view, resource oscillations are inevitable, since the limited nature of the resources implies that co-developing growers compete for their share of these resources, even while supporting each other in terms of conjoined development. Thus, instead of configuring the effect of resource oscillations as a random and periodic damping factor on the growth rate, it can be expressed by incorporating additional systemic components in the model, thereby turning it into a model of connected growers.

3.4.2 Modeling connected growth
Dynamic modeling involves more than just fitting logistic growth equations to developmental trajectories of empirical data. By combining several developing components of a given phenomena in a hierarchy of interacting levels, it can simulate the effect of interactions between connected growers. In practice, this combination necessitates configuring coupled logistic equations. Equation 3 describes two co-developing growers, with no specific hierarchy between them.

\[
A_{n+1} = A_n \left( 1 + r_A \frac{A_n}{K_A} + s_A B_n - c_A B_n \right) \\
B_{n+1} = B_n \left( 1 + r_B \frac{B_n}{K_B} + s_B A_n - c_B A_n \right)
\]

Equation 3. Coupled growers with support and competition by level (based on van Geert, 1995)

The growers, A and B, support each other’s growth, but also compete for resources. Therefore their interactions are defined as bidirectional. The equation parameters are \(K\) = carrying capacity; \(n\) = number of observations, which is equivalent to time \(t\) in the previous equations, with 1 replacing \(\Delta\) as a fixed unit of change; \(r\) = growth rate; \(c\) = competition; and \(s\) = support. The first part of each equation is equivalent to the logistic growth equation, while the second part specifies the damping factors on this logistic growth, which are the competition and support from the other grower towards the one described by the equation. Support from the precursor A to the dependent B (\(s_B\)) or from B to A (\(s_A\)) is multiplied by the current value of the grower.
that generates it, meaning that it is by level. The contribution of support from another grower to the growth rate $r$ of its counterpart is curbed or dampened by a competition parameter $C$. The growth of $C$ is also by level, in direct relation to the value of the grower that generates it. Considering that the equations for both growers $A$ and $B$ are identical, and that there is no damping parameter on the onset of their growth onset or their interaction, iterating the equations would simply yield identical growth curves (given of course that the parameter values for each grower are equal). In other words, there is no hierarchy (order) between the two connected growers.

As discussed in the preceding chapter, a generic connected growth model which is particularly suitable for simulating cognitive and linguistic growth is the precursor model. In a precursor interaction, the development of one grower is a prerequisite to that of another. Various versions of the precursor model can be set up, with conditional unidirectional or bidirectional support and/or competition, either by level or by change. The following section contains the equations for a basic precursor interaction, followed by examples of versions of this model used to simulate linguistic development in several areas.

3.4.3 Precursor interactions: unidirectional support

A precursor interaction between growers can be specified as such by setting a threshold value that the precursor, i.e., the earlier-developing grower, needs to reach before it enables the development of its dependent. This development need not necessarily depend on explicit support from the precursor, but can simply ensue on its own as a result of resource reallocation (since the threshold value assumes that the precursor has ceased to tax these resources to the degree that they are unavailable for the dependent). However, usually some form of support from the precursor to the dependent is implied, since the notion of connected growth is related to the hierarchy of the system, in which lower “species” actively support the emergence of higher ones, as illustrated by the island metaphor (section 2.1.1.5.1).

Equation 4 configures two connected growers $A$ and $B$. The notion of a precursor interaction, in which the development of a given grower is a prerequisite for that of another, can be added to a connected growth model by an additional parameter $P$, which is a binary on/off variable. $P=0$ in if the current value of the first grower (the precursor) is lower than a specified threshold. Once that threshold is reached, then $P=1$. For example, if two-word utterances in early L1, denoted as $B$, can be generated
only once vocabulary, denoted as $A$, has reached a certain size, then $P=1$ when $A_n$ is equal or higher to this prerequisite threshold value (Bassano & van Geert, 2007). Thus $P_B$ thereby marks the onset of development in $B$ (since the model does not include any feedback delay). Support from the precursor $A$ to the dependent $B$ is therefore in effect conditional, as well as by level (directly related to the value of $A$).

\[
A_{n+1} = A_n \times \left[ 1 + r_A - \frac{r_A \times A_n}{K_A} \right]
\]

\[
B_{n+1} = B_n \times \left[ 1 + (r_B - \frac{r_B \times B_n}{K_B} + s_B \times A_n) \times P_B \right]
\]

Equation 4. Two growers in a precursor interaction with unidirectional support by level

Figure 7. Four growers in a precursor interaction with unidirectional support by level

Figure 7 shows the pattern that this type of precursor model would generate in a system of four connected growers, given equal growth rates and carrying capacities, with different conditional threshold values for the onset of support. The model is somewhat idealized, lacking the gradedness and variability that can often be seen in the emergence of developmental stages, which usually overlap. This is likely due to the fact that competition from the dependent is not included in the model. The following section demonstrates this extension, in a precursor model with unidirectional support (from the precursor to the dependent) and competition (from the dependent to the precursor).
3.4.4 Precursor interactions: unidirectional support and competition

\[
A_{n+1} = A_n \left[ 1 + \left( r_A - \frac{r_A * A_n}{K_A} - c_A * B_n \right) \right]
\]

\[
B_{n+1} = B_n \left[ 1 + \left( r_B - \frac{r_B * B_n}{K_B} + s_B * A_n \right) * p_B \right]
\]

Equation 5. Two growers in a precursor interaction with unidirectional (precursor to dependent) support and unidirectional (dependent to precursor) competition, both by level

Equation 5 incorporates competition as well as support: precursor A grows to a threshold value before enabling the emergence of B. A then supports B, while B competes with A by level. This equation is suitable for simulating interactions in which there is an eventual decline in the precursor due to increased competition from the dependent, as the value of the latter grows (in line with increased support from the precursor). In such a model, the dependent will ultimately overrun the precursor. The model can be used to capture the manner in which early strategies in child learning diminish over time, while more advanced strategies emerge, as suggested by Fischer (1980). It was also applied to the emergence of 1-, 2-, or 3-word utterances in early L1. In this data, the “holophrasic” one-word stage, the precursor to the other stages, is eventually replaced by the multi-word stage (Bassano & van Geert, 2007). This replacement would not normally take the shape of a smooth transition, in which dependents replace precursors sequentially. Rather, it is likely that the high amount of variation in developmental and linguistic data would yield a pattern of cyclic decline and recovery of earlier strategies, until their eventual disappearance.

Figure 8 depicts such a model as it is extended to four equations describing growers (levels) A, B, C and D. Beginning at A, each grower is precursor to a dependent, which in turn is a precursor to the following grower. The competition increases in accordance with the value attained by each grower. The model produces a typical pattern: within a thousand iterations, level A declines to a minimum, level B also declines, although not to that extent, and levels C and D increase towards their carrying capacity (configured as 1 in this version).
Investigating language development from a dynamic perspective

Figure 8. Four growers in a precursor interaction with unidirectional (precursor to dependent) support and unidirectional (dependent to precursor) competition, both by level

It is also possible to specify competition (or support) as increasing or decreasing in line with the local positive or negative growth in the grower that generates it (by change), rather than in accordance with the current value of this grower (by level). In this version of the precursor model, the precursor does not necessarily diminish over time. The next section addresses this possibility.

3.4.5 Precursor interactions: bidirectional support by level; unidirectional competition by change

In the preceding coupled grower equations, support and competition are by level, and thus simply linear functions of the previous value of the competitive or supportive grower. However, it is likely that there are developmental stages in which competition changes as a function of the growth process in the grower that generates it, reflecting the amount of resources invested in its local change between time units. This type of competition is referred to as competition by change. According to van Geert (2003), an example of such an interaction can be seen between early reading skills and vocabulary acquisition. While reading skills depend on prerequisite vocabulary knowledge, and while these skills in turn support the acquisition of new vocabulary, the process of reading skill acquisition may interfere with the process of acquiring new vocabulary items. Thus support between these two constructs is by level (and conditional on a threshold value when generated by the precursor, vocabulary knowledge, towards the dependent, reading skills). In contrast, competition from reading skills to vocabulary knowledge is by change, and depends on the amount of resources invested locally in reading skills acquisition. The following chapter
readdresses this example in the context of L2 vocabulary development. Equation 6 depicts this version of the precursor model.

\[
A_{n+1} = A_n + \left( r_A \frac{A_n}{K_A} + s_A B_n - c_A (B_n - B_{n-1}) \right)
\]

\[
B_{n+1} = B_n + \left( r_B \frac{B_n}{K_B} + s_B A_n \right) P_B
\]

Equation 6. Two growers in a precursor interaction with bidirectional support by level, and unidirectional (dependent to precursor) competition by change

Depending on the values of the support and competition parameters, such a model can increase to a maximal value on all levels, or only on some. Figure 9 depicts an extension of this model to a set of four connected growers, which all ultimately increase to their carrying capacities.

An alternative to specifying weak or strong support and/or competition values in the model is specifically configuring competition as diminishing over time, as a function of both the change process and the value of the grower that generates it (in other words, by change and by level, simultaneously). This is done by dividing the difference (change) between the present and the previous value of the grower by its current value (level), and multiplying the competition parameter value with the outcome. In this context, van Geert refers to two simultaneously competing growers \(M\) and \(L\), and the idea that the attention and effort spent to learn a new skill, principal, or strategy is considerably greater at the earlier states than it is in the later ones. The reasoning behind this assumption is that once the learning task becomes more familiar, less resources will have to be invested to achieve a similar amount of
progress. In this alternative version of the model, the competition factors \( C_L \) and \( C_M \) are multiplied by \((M_n - M_{n-1}) / M_n\) and \((L_n - L_{n-1}) / L_n\) respectively (1995, p. 322).

The coupled equations would thus take the following form:

\[
A_{n+1} = A_n \left(1 + \left[ r_A - \frac{r_A \cdot A_n}{K_A} - \frac{c_A \cdot (B_n - B_{n-1})}{B_n} + s_A \cdot B_n\right]\right)
\]

\[
B_{n+1} = B_n \left(1 + \left[ r_B - \frac{r_B \cdot B_n}{K_B} - \frac{c_B \cdot (A_n - A_{n-1})}{A_n} + s_B \cdot A_n\right]\right)
\]

**Equation 7. Connected growers in a precursor interaction with bidirectional support and bidirectional decreasing competition (based on van Geert, 1995, p. 323)**

The revised model ensures that ultimately, all growers will reach their carrying capacities, regardless of their parameter values.

### 3.4.6 Aggregated support and competition

Support and competition can be aggregated as a single influence on the growth rate of each grower. It is therefore possible to configure a general damping factor for each grower, which summaries a particular interaction between this grower and another. Such damping factors can in turn be summarized in a matrix that contains all the combinations of growers in the model. For a given system, the interactions by level and those by change need to be summarized in separate matrices.

<table>
<thead>
<tr>
<th></th>
<th>R</th>
<th>A</th>
<th>B</th>
<th>C</th>
<th>D</th>
</tr>
</thead>
<tbody>
<tr>
<td>A</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>B</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>C</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>D</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Table 1. Aggregated influences on the growth of four connected growers (extracted from van Geert, 1993, p. 232)

In Table 1, as in the equations in this chapter, \( r \) denotes the growth rate and \( K \) denotes the carrying capacity of each grower. \( d \) is the general damping factor which is the sum of influences (support and/or competition) from one grower to another in the specific version of the model. On the basis of such a matrix, van Geert (2003) has developed a VBA-based Excel spreadsheet program that iterates combinations of coupled grower equations like those featured in this chapter, and can thus accommodate various versions of the precursor model. This program was used as a basis for the simulation of L2 vocabulary development described in the next chapter.
3.5 Summary
Following the introduction to the theoretical dynamic perspective on L2 development, this chapter has focused on the practical steps involved in applying this approach in an empirical study. It first discussed the general compatibility of case study methodology with the dynamic approach, and then detailed the steps involved in using such a design in the context of a DST-oriented study. Next the chapter demonstrated some techniques of variability analyses, with a focus on revealing inter-componential interactions, which are applied in the current thesis. Finally, it showed how logistic growth equations can be used as a basis for modeling development, and how they can be adapted to simulate interactions between connected growers, particularly when such growers are defined as precursors and dependents. Models of growers in a precursor interaction depict a clear hierarchy, which should be observable in the data at least initially, when its components emerge in sequence. Such models should therefore be based on theory and recurrent empirical findings.

The following chapters apply the dynamic perspective to longitudinal data in two areas of L2 development: vocabulary knowledge in the fourth chapter, and writing performance, expressed in the accuracy and complexity of lexicon and syntax, in the fifth. These studies follow the procedures outlined so far: longitudinal and detailed data collection, growth and variability analyses, and data simulations on the basis of the outcomes of these analyses and the background literature. While each study has its own set of domain-specific research questions, they can be collapsed as two general questions. The first question is whether variability analyses can reveal complex dynamic interactions in the data. The second question is whether such interactions, when configured in a model depicting the data components as a hierarchy of connected growers, can adequately simulate the data and thereby reinforce the interpretations of the variability analyses. In other words, the two studies put the theoretical and empirical aspects of the dynamic approach to the test.