Caspi, Tal; Lowie, Wander

The Dynamics of L2 Vocabulary Development: A Case Study of Receptive and Productive Knowledge

A dinâmica do desenvolvimento do vocabulário L2: Um estudo de caso do conhecimento receptivo e produtivo

Tal Caspi*
University of Groningen - Holanda
Groningen - Holanda

Wander Lowie**
University of Groningen - Holanda
Groningen - Holanda
University of the Free State - África do Sul
Free State - África do Sul

RESUMO: Muitos estudos têm mostrado que a transferência do vocabulário L2 receptivo para a produção está longe de ser linear, se a produção é induzida por uma tarefa ou espontânea. A natureza e as causas dessa lacuna entre o conhecimento de vocabulário receptivo e produtivo têm sido amplamente debatidas. Este artigo tenta uma nova abordagem a este tema, investigando o desenvolvimento do conhecimento de vocabulário em um estudo de caso detalhado. Quatro níveis de conhecimento de vocabulário de ESL são identificados ao longo de um período de 36 semanas de intensa exposição à língua-alvo. Análises detalhadas revelam que esses níveis receptivo e produtivo de conhecimento interagem de forma complexa ao longo do tempo, simultaneamente competindo pelos recursos do aprendiz de língua enquanto também condicionalmente se apoiam mutuamente. Um modelo baseado em equações de crescimento dinâmico corrobora a hipótese de que as interações complexas entre os níveis de conhecimento de vocabulário contribuem para a lacuna entre vocabulário receptivo e produtivo.

PALAVRAS-CHAVE: Desenvolvimento do vocabulário.

* t.caspi@rug.nl

** w.m.lowie@rug.nl

ABSTRACT: Many studies have shown that the transfer of receptive L2 vocabulary into production is far from linear, whether production is elicited by a task or spontaneous. The nature and causes of this gap between receptive and productive vocabulary knowledge have been widely debated. This article attempts a novel approach to this topic by investigating vocabulary knowledge development in a detailed case study. Four knowledge levels of ESL vocabulary are traced across time during a 36-week period of intensive exposure to the target language. Detailed analyses reveal that these knowledge levels interact in a complex way over time, simultaneously competing for learner resources and conditionally supporting each other's growth. A model based on dynamic growth equations supports the hypothesis that complex interactions between vocabulary knowledge levels give rise to the receptive-productive gap.

KEYWORD: Vocabulary development.

Introduction

The distinction between receptive and productive vocabulary knowledge is widely accepted (MELKA, 1997), with numerous studies noting that in second language (L2) vocabulary the transition from receptive knowledge to production is not immediate, linear, or predictable (SCHMITT; MEARA, 1997; LAUFER, 1998; LAUFER; PARIBAKHT, 1998). This finding, commonly referred to as the receptive-productive gap, is subject to debate. The size and stability of the gap vary across studies, as do the definitions and operationalizations of vocabulary reception and production (FAN, 2000; WEBB, 2008; LAUFER, 2003, 2009). Studies concerned with the receptive-productive gap have analyzed differences across learner populations, or between vocabulary items of different frequencies, taking measurements either once or pre- and post-treatment. Consequently, not much is known about the process of change in the gap over time. A detailed case study, frequently used in L1 research, can shed light on the gap as a developmental phenomenon. The current study combines this approach with the dynamic/complexity perspective, which considers language to be an open and resource-limited system, in which interaction between components is inherent to their development.

The study encompasses two stages: (1) Data analyses focusing on both general trend and variability and (2) Mathematical simulation testing the interpretation of these analyses. These procedures correspond to two research questions: (1) How stable is the gap over time? and (2) How can the complexity/dynamic perspective contribute to its understanding? More specifically, can the precursor model, a basic model based on generic principles of development in a dynamic system, account for growth patterns in empirical data.

Rather than positing a dichotomy of vocabulary reception vs. production, we specify a four-level continuum which merges two validated paradigms (LAUFER; NATION, 1999; LAUFER *et al.*, 2004). Ranking from least to most productive, the continuum encompasses word recognition, word recall, controlled production, and free production. The following Background section reviews some of the literature on the receptive-productive gap, before describing the dynamic/complexity perspective and its relevance to L2 vocabulary development.

Background

The receptive-productive gap

Numerous paradigms have attempted to capture the complex and multidimensional nature of vocabulary knowledge (MELKA, 1997; SCHMITT *et al.*, 2001). Several of these distinguish breadth from depth, with breadth referring to the number of known words and depth to the degree of their knowledge (QIAN, 1999; READ, 2000; LAUFER; GOLDSTEIN, 2004). However, apart from a general consensus that production is a more sophisticated skill than reception, the operational definitions of reception and production vary greatly between studies (MELKA, 1997). For instance, some studies define production as the ability to recall a word, as measured by its elicitation through L1 translation (FITZPATRICK *et al.*, 2008; WEBB, 2008). Others consider production as the ability to write a word in context, which is either elicited (for instance, by a gap-fill task, cf. LAUFER; NATION, 1995, 1999), or spontaneous (LAUFER; PARIBAKHT, 1998; LAUFER, 1998). This distinction between controlled and spontaneous/free production is essential, as "not all learners who use infrequent vocabulary when forced to do so will also use it when left to their own selection of words" (LAUFER, 1998, p. 257).

Their differences notwithstanding, researchers in applied linguistics generally agree that receptive knowledge does not transfer directly or systematically to production. For example, Fan (2000) reported that only about 75% of receptively-known words are known productively. She defined receptive knowledge as the ability to relate word form to meaning, in accordance with Nation's Levels Test (1990), and production as controlled elicitation, assessed by the gap-fill Productive Levels Test (LAUFER; NATION, 1995, 1999). After testing participants in nine proficiency levels, Fan reported "no consistent relationship between language proficiency and the two types of vocabulary

knowledge" (*Ibid.*, p. 117). Moreover, she stipulated that "with such big variations among the 9 groups of students, it can be predicted that this gap will be more varied with regard to individual students" (*Ibid.,* p. 117). Due to this salient variability, Fan suggested that case studies would contribute to understanding the gap, in congruence with other recommendations (cf., SCHMITT; MEARA, 1997; HORST; MEARA, 1999).

An earlier study (WEBB, 2008) found the gap to be much smaller, with less than 10% of receptive words not known at a productive level (versus Fan's 25%). However, the two studies differed in their operationalization: Webb defined reception as L2 (English) to L1 (Japanese) translation ability and production as L1 to L2 translation. He used two alternative scoring methods, sensitive and strict, including and excluding spelling errors, respectively. Although the sensitive scoring revealed a smaller gap than the strict one, as well as large individual differences, Webb concluded that receptive vocabulary is larger than productive vocabulary and that the gap is larger in lower frequency words.

The incongruence between the two studies may stem from differences in their operationalization and assessment, tested word frequency, or context and duration of L2 learning. However, even single studies frequently note a high degree of inter-learner variability in gap size (MELKA, 1997). For example, prolonged L2 immersion has a positive influence on the transfer of receptive word knowledge into production; however, this effect is far from uniform in different learners (SCHMITT; MEARA, 1997; LAUFER; PARIBAKHT, 1998). In some cases, production in certain word frequencies even decreased after L2 immersion (SCHMITT; MEARA, 1997). Despite robust variability in findings and repeated recommendations for longitudinal methodology, which is highly effective and popular in L1 development studies, most studies of L2 vocabulary focus on cross-sectional data, with only single or pre- and post-treatment measurements. The few exceptions did not explicitly focus on the gap, referring to it only in passing (HORST; MEARA, 1999; FITZPATRICK *et al.*, 2008). So far, no study has focused on the gap as an inherent aspect of L2 vocabulary development or on changing interactions between knowledge levels.

A dynamic approach to vocabulary knowledge

The present study combines longitudinal methodology with the dynamic/complexity approach to language development (VAN GEERT, 1991, 2008; LARSEN-FREEMAN, 2006; DE BOT, 2008; DE BOT; LOWIE;

VERSPOOR, 2007; VAN GEERT, 2009). Dynamic systems theory (DST) is concerned with describing and explaining change throughout numerous complex systems, such as climate change, mineral formation, and animal and human behavior. Since the 1990s, DST has been extended to language development studies, in which it emphasizes process, whether individual or social-historical, rather that product (VAN GEERT, 1991; LARSEN-FREEMAN, 1997; ROBINSON; MERVIS, 1998; VAN DIJK; VAN GEERT, 2007; DE BOT, 2008; DE BOT; LOWIE; VERSPOOR, 2007; VERSPOOR *et al.*, 2008; LARSEN-FREEMAN; CAMERON, 2008; VAN GEERT, 2009). Development relies not just on input, but on principles that underlie natural complex systems: constant inter-componential interactions, shaped by finite resources (VAN GEERT, 1994, 2003). By considering changes in interactions between components of language, the dynamic approach to language development considers variability as indicative of change processes, not merely measurement error (VAN GEERT; VAN DIJK, 2002). This approach diverges from prevailing methodology in applied linguistics, which focuses on group effects and strives to determine statistically significant and static relations between various factors and language proficiency or between language modalities (DE BOT, 2008; VAN GEERT, 2008). Likewise, most vocabulary studies address interactions between various vocabulary knowledge levels as linear and stable (e.g., WEBB, 2008).[1] Implicit in this view is that a score on one knowledge level can predict the score on another. However, while such analyses can effectively show the general growth trend and correlations between knowledge levels, they cannot reveal temporal patterns of change and interaction.

Dynamic studies often complement such central trend analyses with variability analyses that reveal micro-level interactions between data variables. Several such studies showed that variability patterns change meaningfully and systematically over time, possibly indicating underlying processes, such as inherent competition between various dimensions of language (LARSEN-FREEMAN, 2006;VERSPOOR *et al.*, 2008; SPOELMAN; VERSPOOR, 2010; VAN DIJK *et al*, 2011). The present study expands this strategy. It first inspects variability as a potential indicative of changing interactions between co-developing

---

[1] A notable exception are Paul Meara's models (2001, 2006), which address vocabulary knowledge as a network of activation in which vocabulary items are connected to each other and development is iterative. These models operate at the word level, and are therefore not directly related to the current study.

vocabulary knowledge levels, then configures these interactions in a mathematical model aimed at testing the interpretation of these variability patterns.

Dynamic models simulate patterns of self-organization in complex systems. In such models, development is iterative (each systemic state is based on a previous state), curbed by a maximal value (carrying capacity) and by resource limitations, and emerges from interactions between interconnected systemic components (referred to as *connected growers* due to their continuous development). Any dynamic model assumes that development is enabled by systemic resources and curbed by their limited nature, which shapes the interactions between growers.

These interactions encompass competition, support, or simultaneous combinations of both. For instance, various species compete for territory and water in a given ecosystem, but may also support each other in predator-prey or other interactions. Therefore, growers develop nonlinearly, with the term *growth* encompassing both increase and decline (VAN GEERT, 1993, 1994, 2003, 2011).

In the context of language development, dynamic models are based on the premise that learning new skills or aspects of knowledge is "proportional to the skills or knowledge already attained, proportional to the difference between what is already learned and what has yet to be learned, and proportional to the available resources (e.g., the quality of the teaching, the student's motivation, etc.)" (VAN GEERT, 2008, p. 190). Whether these resources are internal, such as attention or memory capacity, or external, such as input, their limited nature determines the interactions between the growers in the system. These interactions, in turn, affect the growth rates of language growers (VAN GEERT; STEENBEEK, 2005; LOWIE *et al.*, 2011).

In language development, mathematical models can illuminate underlying processes and test hypotheses about the causes and course of change, thereby supplementing cross-sectional studies. For example, Van Geert *et al.* (2011) explored teacher-learner scaffolding by using coupled growth equations, showing that development is shaped by dynamic and changing interactions. The following section elaborates on the basic equations used to model language development.

A basic model of dynamic growth in a complex system

In a complex system, certain growers are *precursors*, which must attain a threshold value to enable the growth of their *dependents*, first by freeing

some resources and second by possibly acting as building blocks to these dependents (VAN GEERT, 1991). Early L1 lexicon has been identified as a precursor to syntax, because the latter's development is enabled only once a minimal vocabulary size has been reached. Once this prerequisite threshold enables syntactic development, a dynamic interaction between lexicon and syntax ensues. Initially, early syntax may compete for resources with the ongoing process of vocabulary acquisition, thereby curbing lexical growth. Over time, the interaction between these growers shifts from competition to support as both stabilize (VAN GEERT, 1991). In other words, emerging skills can support each other by the value that they attain even while competing locally through their growth processes for limited resources. This interpretation was further corroborated by a mathematical model of early L1 lexicon and syntax in a precursor interaction, which has successfully replicated empirical data (ROBINSON; MERVIS, 1998). Verspoor *et al.* (2008) have suggested that written L2 lexicon and syntax also have a similar precursor interaction, identifying complex and shifting interactions between lexical and syntactic indexes in the variability patterns of longitudinal data. Thus, the precursor model may be relevant to various aspects of language.

In the model, support between two growers changes as a function of the value of the grower that generates it, affecting the growth rate of the other grower. In turn, this rate is relative to the ratio between the grower's current and optimal value. The optimal value is known as the grower's *carrying capacity*, its maximal value given optimal resource allocation. Another impact on growth rate is competition between the growers, which can change as a function of either the current value of the grower generating it (*by level*) or as a function of the growth process itself (*by change*). The interconnectedness of growers, the precursor hierarchy between them, and the effects of support and competition on their respective growth rates can be described by coupled equations (with more equations added as per the number of described growers).

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

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

EQUATION 1. Coupled growth equations for connected growers in a precursor interaction, with bidirectional support by level and competition by change, based on Van Geert (1991)

Equation 1 describes a simple system consisting of two connected growers, *A* and *B*. For each grower, the parameters in the equation are K = carrying capacity; r = growth rate; c = (level of) competition from the other grower towards it; s = (level of) support from the other grower; n = time, with the model describing each grower's value at time point *n+1* as iteratively based on its previous value at time point *n*. The first half of each equation consists of the logistic growth equation, which depicts nonlinear growth as curbing exponential growth in accordance with the distance between the current value of the grower and its carrying capacity. The second half of each equation specifies the interaction: support and competition received from the other grower. Growth in the support parameter *s* is *by level*, whereas growth in the competition parameter *c* is *by change*. This means that support changes in line with the value of the grower that generates it. For example, s_A (support from *B* towards *A*) is linearly related to the current value of the grower (B_n). Competition, on the other hand, changes in line with the change that occurs in the value of the grower that generates it between time points *n-1* and *n*, the two occasions preceding the current time point *n+1*. For example, C_A changes in accordance with the difference between the values B_n and B_{n-1}. Additionally, grower *A* is defined as a precursor to grower *B*, which means that grower *A*'s development is a prerequisite to the development of another grower, given that *A* reaches a threshold value. This is reflected in *P*, which is a binary on/off variable: *P*=1 when *A* =threshold value, whereas *P*=0 up to that point.

Equation 1 depicts one version of the precursor model. Alternative versions can be configured, in line with the described phenomena and relevant theory. In this way, dynamic models can be used to assess the interpretation of empirical results and test hypotheses about interactions between components and their effect on development.

Research questions and predictions

This study formulates two key questions: 1) How stable is the receptive-productive gap? 2) Can change in the gap over time (or lack thereof) be accounted for by dynamic precursor interactions between vocabulary knowledge levels?

Based on previous studies, we expected that over a period of L2 immersion all vocabulary knowledge levels should increase but that their growth would be disparate and, therefore, the gap would continue to be manifested. Furthermore, in line with the complexity/dynamic <u>approach</u>, we

stipulate that L2 vocabulary knowledge is a dynamic system of precursor and dependent levels. Within this interconnected hierarchy, development is shaped by constant and shifting interactions determined by resource limitations. More specifically, we expected to find competitive interactions between the lower (more-receptive) levels and the higher (more-productive) levels within the vocabulary knowledge continuum. We assumed that inherently limited cognitive resources were necessary to acquire new receptive vocabulary and maintain existing receptive vocabulary knowledge, as well as to shift receptive knowledge into production. This competition offsets the support from receptive vocabulary knowledge towards its production, which is derived from the essential role of receptive knowledge as a precursor to production. A certain amount of receptively-known vocabulary should be a prerequisite for production, and interactions between lower, more-established receptive knowledge levels and higher, less-established productive levels would be simultaneously supportive and competitive.

Methodology: participants, materials and procedures

The data described here was derived from a 24-year-old native speaker of Brazilian Portuguese, a female university student enrolled in an intensive English-speaking Master's degree program. Data collection began at the onset of the immersion period and lasted 36 weeks. We assumed that the student's intensive exposure to spoken and predominantly written academic vocabulary would lead to vocabulary gain, as reported by studies of vocabulary acquisition through reading (HORST, 2005). Accordingly, we focused on academic English as the target vocabulary, as defined by the University Word List (UWL) (XUE; NATION, 1984) and the Academic Word List (AWL) (COXHEAD, 2000). These lists, which overlap to a degree, were compiled by extensive analyses of academic corpora from various disciplines, and exclude both specialized and general use vocabulary. The words in the lists are distributed across different frequencies, thereby avoiding potential frequency effects on receptive-productive transfer, as suggested by previous studies (e.g., LAUFER; PARIBAKHT, 1998; WEBB, 2008).

Recognition, recall, and controlled production of academic vocabulary were measured by the Longitudinal Academic Vocabulary Tests (LAVT) (CASPI; LOWIE, 2010), which merges adaptations of two vocabulary tests. Word recognition and word recall are based on the Computer Adaptive Test of Size and Strength (CATSS) (LAUFER *et al.*, 2004; LAUFER; GOLDSTEIN,

2004), while controlled production is based on the productive version of the Levels Test (LAUFER; NATION, 1995).

Word recognition, the lowest knowledge level, is operationalized as the ability to identify a target word when supplied with its meaning, given a choice of four alternatives. For example:

> (target word: **preliminary**):
>
> "coming before a more important action or event, especially introducing or preparing for it"
>
> a) applicable   b) preliminary   c) immense   d) extended

Word recall is measured as the ability to recall the word when presented with its meaning and initial letter. For example:

> (target word: **cooperate**):
>
> "to act or work together for a particular purpose, or to help someone willingly when help is requested"   c_____

Controlled production is elicited by the ability to use a word in a gap fill exercise, given the word's initial letter. Two sentences are used per word, in order to capture its use in varied contexts. The sentences were derived from news articles rather than academic contexts, because in the latter the target word would likely to be surrounded by lower-frequency vocabulary that might affect its elicitation. For example:

> (target word: **norm**):
>
> "One child per family is fast becoming the n_____ in some countries; Many immigrants find it hard to adjust to European cultural n_____s".

The LAVT is scored by calculating the ratio of correct to total items. It encompasses a database of 1,000 test items per each vocabulary knowledge level, each testing a different word. From this database, we randomly drew weekly test versions containing 30 items per knowledge level. This procedure was shown to have significant within-versions reliability (CASPI; LOWIE, 2010; see also Table 1).

TABLE 1
Correlations between two LAVT versions

Test part	Pearson's r coefficient	Number of participants
Controlled production	0.775	27
Recall	0.844	32
Recognition	0.733	31

The highest level in the knowledge continuum is free production, spontaneous written word use. Free production was measured by freely written essays on equivalent topics. Accordingly, the choice of essay topics included varied non-academic subjects ranging from the environment, economy, and education to personal experience and opinion, taken from a list of potential TOEFL essays. To encourage the participant's motivation to write and promote her creativity, a choice between three topics was given each week.

Previous studies calculated free production as the proportion of target vocabulary to another type of vocabulary (LAUFER, 1998; LAUFER; PARIBAHKT, 1998). This calculation omits the dimensions of complexity and accuracy that are inherent in testing the lower vocabulary knowledge levels. In effect, it measures free production on a different scale. We corrected for this disparity by calculating free production as the ratio of correct academic (UWL and AWL) word families to the total number of academic word tokens, divided by the total number of correct content words and multiplied by the general family/token ratio. In this way, both complexity (family/token ratio) and accuracy (ratio of correct items) were incorporated in the free production index. When free production is calculated as the ratio of the target vocabulary to other words it cannot possibly reach an optimal 100% value attainable by lower vocabulary knowledge levels. Our calculation also minimizes this disparity.

Results

We first plotted the ratios of known vocabulary per level and their linear trends in order to observe the course of development and compare its trajectory between levels.
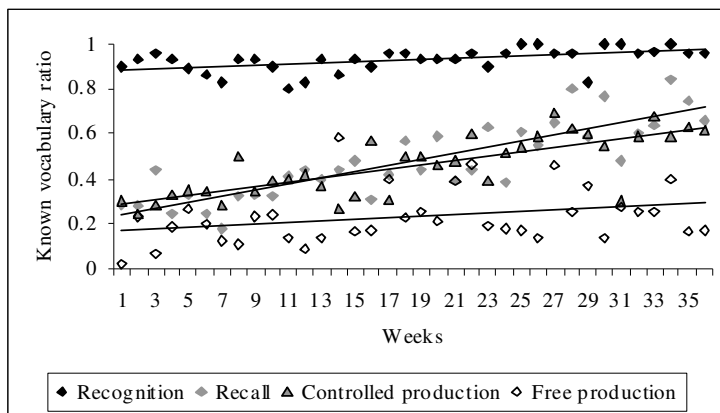
FIGURE 1. Raw data values and linear trends

Although all knowledge levels show a general linear increase, it is more prominent in recall and controlled production than in free production and recognition. Predictably, the onset value of recognition is much higher than that of the other knowledge levels; therefore, its growth is inherently restrained. Free production, which as the least-established knowledge level, has the lowest onset value and exhibits the highest amount of variability, as seen in the distribution of its values around the trend. At some points, such as week 14 and 23, the value of free production even exceeds that of the lower knowledge levels, while at other points it drops to much lower values.

To inspect both growth trends and surface interactions in the data, we next correlated each knowledge level with both time and the other three levels. Because individual growth data are unlikely to be normally distributed, we calculated Spearman's *rho* coefficient, a nonparametric measure of correlation.

TABLE 2

Correlations between the knowledge levels and time and
within the continuum

Correlation	Coefficient value
Recognition-weeks	0.628 (p<0.01)
Recall-weeks	0.849 (p<0.01)
Controlled production-weeks	0.769 (p<0.01)
Free production-weeks	0.347 (p<0.05)
Recognition-recall	0.599 (p<0.01)
Recall-controlled production	0.658 (p<0.01)
Controlled-free production	Not significant
Recognition-controlled production	0.439 (p<0.05)
Recognition-free production	Not significant
Recall-free production	Not significant

Table 2 shows strong positive correlations between the three lower knowledge levels and between those levels and time. This finding corresponds to the visual impression from the data and linear trends in Figure 1. The linear increase in free production (its correlation with the number of weeks) was lower than that of the other knowledge levels. This is not surprising considering both its high degree of variability and its different operationalization and measurement. However, while all knowledge levels, including free production, showed a statistically significant increase during the study period, none of the correlations between free production and the other levels were significant. This finding encapsulates the receptive-productive gap.

To further inspect the data variability and how interactions between the knowledge levels changed over time, we subtracted the linear trends from the data to calculate its residuals, the "de-trended" values, which we plotted as a time series for each pair of consecutive knowledge levels.
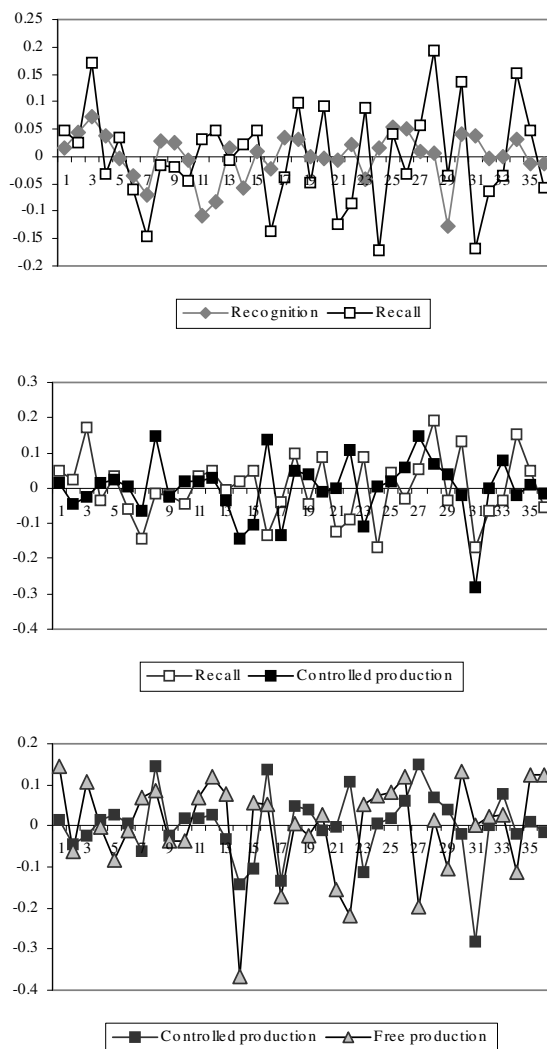
FIGURES 2. Residuals of pairs of adjacent levels on the vocabulary knowledge continuum. The x-axes denote the number of weeks; the y-axes denote the value of the residuals

In the top plot in Figure 2, the residuals of recognition and recall exhibit mainly parallel variability patterns, and few inverse ones, in which an increase above the linear trend in one measure is accompanied by a decrease in the other (e.g., in weeks 10–14). Recall and controlled production, in the middle plot, show both parallel variability patterns, as in weeks 1–13 and 30–33, and inverse patterns in which an increase in one level is accompanied by a decrease

in the other. In the bottom plot, controlled and free production show mostly inverse patterns, particularly in the second half of the period. Taken together, these residual patterns can be interpreted as indicative of a predominantly supportive interaction between recognition and recall, moderate competition between recall and controlled production, and strong competition between controlled and free production. However, at this point, these interpretations are based solely on the visual impression of variability patterns and are therefore a speculation that should be corroborated by further analysis and modeling.

We next plotted the correlations between the knowledge levels as functions of time (moving correlations), to further inspect variability patterns not only in the trajectories of each level, but also in their interactions, and to check our interpretation of the residual patterns in Figure 2. We correlated the residuals of the levels in a moving window of 5 observations, between observations 1–5, 2–6, and so forth. This technique visualizes gradual change in the interaction between the levels, which was expressed only as single coefficient values in Table 2.
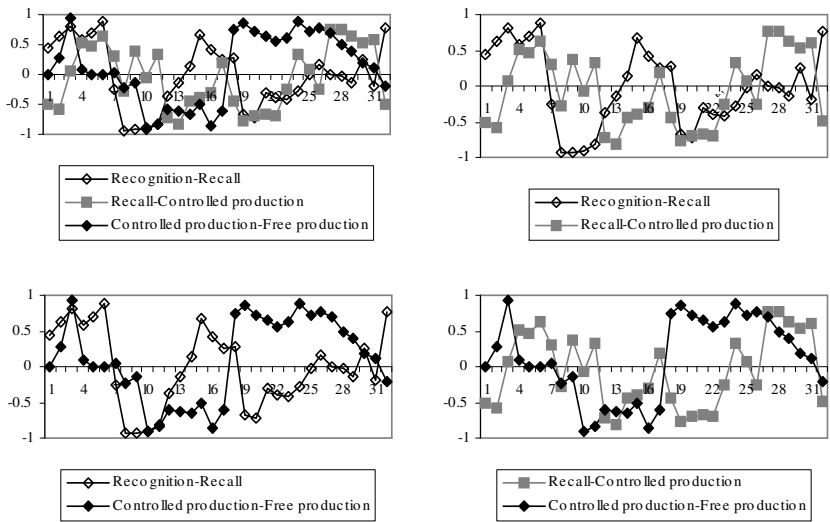

FIGURE 3: Moving correlations between consecutive knowledge levels, plotted all together (top left), and in pairs. The y-axes denote Spearman's *rho* values; the x-axes denote time in weeks

As Figure 3 shows, these correlations mostly fluctuate between highly positive and highly negative values. For instance, in the first half of the immersion period, the recall-controlled production correlation shifts between near-maximum negative and positive coefficient values. This finding is remarkable because, at the macro-level, the interaction between recall and

controlled production appears to be purely supportive, as expressed both in their nearly identical linear trends and in their strong and positive correlation (see Figure 1 and Table 1), while their residual plot (in Figure 2) has suggested some form of competition. The shifts from positive to negative coefficient values in the moving correlation can be interpreted as expressing shifts between support and competition typical of the precursor model. In this interpretation, vocabulary recall is a precursor to controlled production, preceding and initially inhibiting the latter's growth. This competition not stable, but localized in given moments (by change), in conjunction with the resources required for the developmental process; however, as the amount of vocabulary known at the recall level increases in size, it generally supports the controlled production of the same vocabulary (by level).

Furthermore, juxtaposing pairs of moving correlations shows inverse patterns: when one interaction is positive, the other is often negative. Plotting the recall-controlled production correlation alongside the recall-recognition correlation (top right in Figure 3) shows that almost every alteration in one of these correlations is accompanied by an inverse shift in the other. Similarly, comparing the recognition-recall and controlled-free production correlations (bottom left in Figure 3), as well as the recall-controlled production and controlled-free production correlations (bottom right in Figure 3), demonstrates that fluctuations from negative to positive values in one interaction are often accompanied by an inverse shift in the other.

The residual and moving correlation plots thus strengthen the impression of complex interactions between the knowledge levels, encompassing cycles of competition and support. In line with the analyses so far, as well as with the notion of the vocabulary knowledge continuum as a hierarchy, we specified these interactions as support between recognition and recall, weak-moderate competition between recall and controlled production, and strong competition between controlled and free production. We next configured these interactions in a mathematical model of the participant's vocabulary knowledge.

The precursor model simulation

To simulate the data, we used a model programmed in Excel VBA-code by Paul van Geert. The model is based on paired equations, like those in Equation 1, in which, for any iteration, the value of each grower is the outcome of a logistic function based on its previous value and its interactions with the other growers in the system. The parameter configuration in the model relied on the findings

of the variability analyses, in combination with our interpretation of the gap as an expression of precursor interactions within the receptive-productive continuum. The order parameters of the model specify four growers in a hierarchy of precursors and dependents that corresponds with the ascending order of the continuum: recognition is a precursor to recall, which acts as a precursor to controlled production, which in turn acts as a precursor to free production.

In addition to this hierarchical order, we also defined the property control parameters of initial growth rate and initial value. For each precursor, we specified a threshold value, referred to as the *precursor level*, which it needs to reach before enabling the development of its dependent. Another type of control parameter specified the three key interactions between the three combinations of precursor-dependent dyads in the hierarchy. For each grower, these interactions, namely the support and competition parameters, were aggregated as two neutral factors impacting its growth. One parameter tallied the interactions by change; the other of those by level. The revised equation refers to this aggregated parameter as *s* for support, simply because it is added to the growth rate. However, since this parameter can also take on a negative value, it can also express competition, or a negative "support". Apart from the initial values, which were configured as the onset values in the data, all control parameters were assigned with uniform default values, which were then converged via the optimization procedure.

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

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

EQUATION 2. Connected growers in a precursor interaction with unidirectional "support" by change and "support" by level

Model optimization

The next step was to assign default values to all remaining control parameters, namely the precursor level (the conditional threshold determining the value of P as 0 or 1), the growth rates, and the relational control parameters specifying the between-level interactions. The values of the parameters describing interactions between the levels in the model were converged via the simplex algorithm (NELDER; MEAD, 1965), available through the Excel Solver AddIn feature. This optimization is based on minimizing the outcome of a function, in

this case, the sum of squared differences between the model and the data. The method generates positions of trail matrixes derived from configuring different values in the parameters of the function and extrapolates each outcome of the re-configured function at points arranged as a simplex (a function whose derivative is a shape of N+1 vertices in N dimensions, such as a triangle, which has 3 vertices and 2 dimensions; the number of dimensions is determined by the number of parameters in the optimized function). The matrix of squared differences is recalculated up to 10,000 times until an optimal solution is reached, with a tolerance level of 0.00001 (meaning that minimizing the function outcome by this or any higher number is criteria for improvement).

By optimizing the parameter values, namely those of growth rate and the interactions between the knowledge levels, it was possible to assess the overall fit of the precursor model to the data. First, this was done by checking the sum of squared differences between the model and the data (rendered sum of least squares via the optimization procedure). This fit had a value of 0.546191, suggesting a strong similarity. We then plotted and visually compared the generated values from the optimized model with the data, its linear trends, and a smoothed function of the data, which incorporates both trend and variability (spline interpolation).
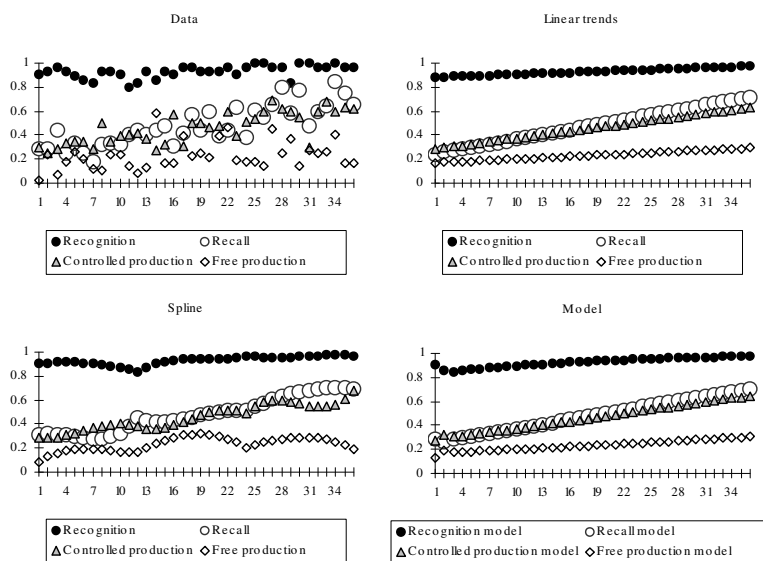

FIGURE 4: (clockwise from top left) data, linear trends, spline and optimized model (fit: 0.546191); the y-axes denote the ratio of known vocabulary; the x-axes denote weeks or, in the model, the number of iterations

Figure 4 shows a great likeness between the optimized model and the data trends, as expressed by the relatively low sum of squared differences between the model and the data. A linear regression also shows that the free production values generated by the model are significant predictors of the data at $F(1,34)=4.197$, $p<0.05$, confirming that the sum-of-squares fit value is good.

The use of the model demonstrates that the data, and thereby the receptive-productive gap that it reflects, can be simulated quite accurately on the basis of precursor interactions iterated over time, without necessarily assuming the influence of external factors on the internal dynamics of the vocabulary knowledge continuum.

We then looked at the optimized configuration of parameter values, which can reveal if the good fit of the model was achieved by maintaining our original configuration of precursor interactions, or whether these parameters needed to be changed in order to achieve an optimal model fit.

TABLE 3
Optimized model parameters

Growth rate recognition	0.057057
Growth rate recall	0.051112
Growth rate controlled production	0.04063
Growth rate free production	0.021853
Recognition to recall, *by change*	0.000484
Recall to controlled production, *by change*	**-0.00156**
Controlled to free production, *by change*	**-0.00279**
Recognition to recall, *by level*	0.152386
Recall to controlled production, *by level*	0.166983
Controlled to free production, *by level*	0.137332
Precursor value, recognition to recall	0.125696
Precursor value, recall to controlled production	0.055095
Precursor value, controlled to free production	0.141572

The values in Table 4 verify our assumption of a hierarchy of precursors and dependents. Within this continuum, support from precursors to dependents is instigated when precursors reach a threshold value. The optimized values show that the configuration of the model which obtained an optimal fit consists of a negative value of the *by change* interactions and a positive value of the *by level* interactions from recall to controlled production

and from controlled production to free production. This indicates complex interactions, comprised of simultaneous competition by change and support by level, from these precursors to their dependents, i.e., from recall and controlled production, the more-receptive levels, to controlled production and free production, the higher and more-productive levels, respectively. In contrast, the interactions between recognition and recall have positive values, implying that recognition does not compete with recall.

Discussion

The objective of this study was to investigate how interactions between levels of vocabulary knowledge affect the transfer of receptive vocabulary to free production, by applying a detailed longitudinal analysis and a process-oriented, dynamic approach. The study focused on the unequal development of receptive and productive L2 vocabulary knowledge known as the receptive-productive gap (LAUFER, 1998). The literature on vocabulary knowledge differs in its definitions and operationalizations of reception vs. production (SCHMITT, 2010) and suggests several explanations for this gap. Among these are effects of word frequency, learner proficiency or learning context. Regardless of their suggested cause, most studies have implicitly treated the gap as a static product that arises from linear interactions between vocabulary knowledge levels and other factors (SCHMITT; MEARA, 1997; LAUFER, 1998; LAUFER; PARIBAKHT, 1998; WEBB, 2008). By contrast, we focused on the development of the continuum of four knowledge levels of L2 vocabulary: recognition, recall, controlled production, and free production. The target vocabulary was academic English, in which an advanced learner was immersed during a 36-week period. Free production was recorded weekly as spontaneous, accurate, and complex usage of the target vocabulary in written essays; recognition, recall and controlled production were measured in a test with established equivalent forms reliability.

We hypothesized that L2 vocabulary knowledge may develop in accordance with the general principles that govern dynamic systems: internal interactions between co-developing components are shaped by resource limitations. Overall, the data showed an increase on all knowledge levels, with a high degree of variability, particularly in free production. However, despite this increase, the gap between the levels and particularly between free production and the other levels remained robust while showing a high degree of variability. This finding matches and corroborates prior studies (SCHMITT; MEARA, 1997;

LAUFER, 1998; LAUFER; PARIBAKHT, 1998; FAN, 2000). Correlation analyses showed that all levels increased significantly with time, but that the growth trend of free production was the weakest. Following this step, variability analyses – residual plots and moving correlations between these residuals – showed that pairs of consecutive knowledge levels exhibit patterns of competition and support associated with the precursor model (VERSPOOR *et al.*, 2008). We thus speculated that lower, more-receptive levels within the knowledge continuum act as conditional precursors for higher, more-productive levels. Once crossing a threshold value, the precursors support the development of their dependents; conversely, the dependents compete with their precursors locally, through the process of their development.

To test whether this version of the precursor model can indeed replicate and thus account for the data growth patterns, we simulated the system on the basis of coupled logistic growth equations (VAN GEERT, 1991, 1994). This model was optimized and its outcome visually compared to the data trends. These procedures showed that precursor interactions across the continuum, in combination with simultaneous competition *by change* and support *by level* between recall and controlled production and between controlled and free production, can generate patterns that are very similar to the data. This implies that, although in general each lower level in the continuum supports its consecutive, more-productive dependent, it also competes with it for resources through the process of change (growth) itself. A comparison of the model with the data and its growth trends and the model fit value was accordingly good.

The dynamic perspective has previously been applied to L2 development by focusing on variability patterns (cf. VERSPOOR *et al.* 2008; SPOELMAN; VERSPOOR, 2010). We have attempted to take this analysis further by not only describing data variability, but also by attempting to simulate development in a dynamic model. The added value of using simulations is that the hypothesized complex dynamic interactions of the variables can be tested by configuring them as numerical values in equations that describe iterated and connected growth. In this case, this meant defining receptive knowledge levels as precursors to productive levels. The simulation indicated that the precursor model, a generic paradigm of development in complex systems, has strong potential for explaining the receptive-productive gap by shedding light on the interconnected and dynamic interactions that underlie vocabulary knowledge and, more specifically, free vocabulary production.

Although DST is not a linguistic theory, it does offer a framework through which such theories can be examined when accounting for development. We suggest a new direction for exploring L2 vocabulary development, but would like to emphasize that the study cannot be considered fully conclusive; applying the dynamic perspective to L2 development is still in its infancy, and our study has several drawbacks.

First, any model is by definition an abstraction. Dynamic modeling enables us to describe, demonstrate, and explain effects that may otherwise remain obscure. Despite these advantages, it has inevitable limitations. Language development is shaped by the dynamic interactions of an infinite number of variables, which can never be simultaneously entered in a single model. However, this model could be extended into a more intricate hierarchy of vocabulary knowledge levels and their interactions.

Second, modeling data on the basis of single cases cannot be generalized to larger populations. However, the inverse limitation holds for cross-sectional studies, which are often not generalizable to individuals (cf. SCHMITT; MEARA, 1997).

Finally, the operationalization of free production, while in line with previous studies and further corrected for accuracy and complexity, is still not optimal. According to dynamic/complexity theory principles, the higher variability in free production in relation to the other knowledge levels is inherent and expected, as it is a lesser-established knowledge mode. However, measuring free production might be inherently less reliable than measuring more-receptive vocabulary knowledge levels, as by definition, free production cannot be elicited. Yet, the fact that the participant *has* shown free production of academic vocabulary in her (non-academic) writing shows that receptive knowledge of vocabulary items does transfer into production even when not specifically required. It is predominantly the nonlinear and variable nature of this use that has puzzled researchers (MELKA, 1997; LAUFER, 1998; WEBB, 2008; SCHMITT, 2010).

Our study shows the benefit of applying the dynamic/complexity perspective to empirical data to better understand essentially developmental and natural phenomena. Although additional data from various stages of L2 vocabulary development are necessary to establish the explanatory power of the dynamic approach more firmly in this area, it is clear that the merit of the approach is found in its link to existing theory and empirical data. This is a collaborative, iterative process to which we hope our study contributes.

Acknowledgements

References

CASPI, T.; LOWIE, W. A dynamic perspective on academic English L2 lexical development. In: CHÁCON-BELTRÁN, R.; ABELLO-CONTESSE, C.; TORREBLANCA-LÓPEZ, M. d. M.; LÓPEZ-JIMÉMENEZ, D. (Ed.). *Further insights into non-native vocabulary teaching and learning*. Bristol: Multilingual matters, 2010. p. 41-60.

COXHEAD, A. A new academic word list. *TESOL Quarterly*, TESOL, v. 34, n. 2, p. 213-238, 2000.

DE BOT, K. Introduction: Second language development as a dynamic process. *The Modern Language Journal*, Blackwell Publishers, v. 92, n. 2, p. 166-178, 2008.

DE BOT, K.; LOWIE, W. M.; VERSPOOR, M. H. A Dynamic Systems Theory approach to second language acquisition. *Bilingualism: Language and Cognition*. Cambridge University Press, v. 10, n. 1, p. 7-21, 2007.

FAN, M. How big is the gap and how to narrow it? An investigation into the active and passive knowledge of L2 learners. *RELC Journal*, SAGE Publications, v. 31, n. 2, p. 105-119, 2000.

FITZPATRICK, T.; AL-QARNI, I.; MEARA, P. Intensive vocabulary learning: a case study. *Language Learning Journal*, Routledge, v. 36, n. 2, p. 239-248, 2008.

HORST, M. Learning L2 vocabulary through extensive reading: a measurement study. *The Canadian Modern Language Review*, University of Toronto Press, v. 61, n. 3, p. 355-382, 2005.

HORST, M.; MEARA, P. Test of a model for predicting second language lexical growth through reading. *The Canadian Modern Language Review*, University of Toronto Press, v. 56, n. 3, p. 308-328, 1999.

LARSEN-FREEMAN, D. The emergence of complexity, fluency and accuracy in the oral and written production of five Chinese learners of English. *Applied Linguistics*, Oxford University Press, v. 27, n. 4, p. 590-619, 2006.

LARSEN-FREEMAN, D.; CAMERON, L. *Complex systems and applied linguistics*. Oxford: Oxford University Press, 2008. 287p.

LAUFER, B. The development of passive and active vocabulary in a second language: Same or different? *Applied Linguistics*, Oxford University Press, v. 19, n. 2, p. 255-271, 1998.

LAUFER, B. Vocabulary acquisition in a second language: Do learners really acquire most vocabulary by reading? Some empirical evidence. *Canadian Modern Language Review*, University of Toronto Press, v. *59*, n. 4, p. 567-587, 2003.

LAUFER, B.; ELDER, C.; HILL, K.; CONGDON, P. Size and strength: Do we need both to measure vocabulary knowledge? *Language Testing*, SAGE Publications, v. 21, n. 2, p. 202-226, 2004.

LAUFER, B.; GOLDSTEIN, Z. Testing vocabulary knowledge: Size, strength, and computer adaptiveness. *Language Learning*, Blackwell Publishing, v. 54, n. 3, p. 399-436, 2004.

LAUFER, B.; NATION, I. S. P. Vocabulary size and use: Lexical richness in L2 written production. *Applied Linguistics*, Oxford University Press, v. 16, n. 3, p. 307-322, 1995.

LAUFER, B.; NATION, I. S. P. A vocabulary-size test of controlled productive ability. *Language Testing*, SAGE Publications, v. 16, n. 1, p. 33-51, 1999.

LAUFER, B.; PARIBAKHT, T. S. The relationship between passive and active vocabularies: Effects of language learning context. *Language Learning*, Blackwell Publishing, v. 48, n. 3, p. 365-391, 1998.

LOWIE, W.; CASPI, T.; VAN GEERT, P.; STEENBEEK, H. Modeling development and change. In: VERSPOOR, M.H.; DE BOT, K.; LOWIE, W. (Ed.). *A dynamic approach to second language development*: Methods and techniques. Amsterdam, Philadelphia: Benjamins, 2011. p. 22-122.

MEARA, P. The mathematics of vocabularies. In: GILL, M.; JOHNSON, A.; KOSKI, L.; SELL, R.; WÅRVIK, B. (Ed.). *Language, learning, literature*: Studies presented to Håkan Ringbom. Åbo: Åbo Academy, 2001. p. 151-167.

MEARA, P. Emergent properties of multilingual lexicons. *Applied Linguistics*, Oxford University Press, v. 27, n. 4, p. 620-644, 2006.

MELKA, F. Receptive versus productive aspects of vocabulary. In: SCMITT, N.; MCCARTHY, N. (Ed.). *Vocabulary description, acquisition, and pedagogy*. New York: Cambridge University Press, 1997. p. 84-102.

NATION, I. S. P. *Learning and teaching vocabulary*. New York: Newbury House, 1990. 275p.

NELDER, J. A.; MEAD, R. A simplex method for function minimization. *Computer Journal*, Oxford University Press, v. 7, n. 4, p. 308-313, 1965.

QIAN, D. D. Assessing the roles of depth and breadth of vocabulary knowledge in reading comprehension. *The Canadian Modern Language Review*, University of Toronto Press, v. 56, n. 2, p. 283-307, 1999.

READ, J. *Assessing vocabulary*. Cambridge: Cambridge University Press, 2000. 293p.

ROBINSON, B. F.; MERVIS, C. B. Disentangling early language development: Modeling lexical and grammatical acquisition using an extension of case-study methodology. *Developmental Psychology*, American Psychological Association, v. 34, n. 2, p. 363-375, 1998.

SCHMITT, N. *Researching vocabulary*: A vocabulary research manual. Basingstoke: Palgrave Macmillan, 2010. 408p.

SCHMITT, N.; MEARA, P. Researching vocabulary through a word knowledge framework. *Studies in Second Language Acquisition*, Cambridge University Press, v. 19, n. 1, p. 17-36, 1997.

SCHMITT, N.; SCHMITT, D.; CLAPHAM, C. Developing and exploring behaviour of two new versions of the Vocabulary Levels Test. *Language Testing*, SAGE Publications, v. 18, n. 1, p. 55-88, 2001.

SPOELMAN, M.; VERSPOOR, M. Dynamic patterns in the development of accuracy and complexity: A longitudinal case study on the acquisition of Finnish. *Applied Linguistics*, Oxford University Press, v. 31, n. 4, p. 532-553, 2010.

VAN DIJK, M.; VAN GEERT, P. Wobbles, humps and sudden jumps: a case study of continuity, discontinuity and variability is early language development. *Infant and Child Development,* Blackwell Publishing, v. 16, n. 1, p. 7-33, 2007.

VAN DIJK, M.; VERSPOOR, M. H.; LOWIE, W. Variability and DST. In: VERSPOOR, M.; DE BOT, K.; LOWIE, W. (Ed.). *A dynamic approach to second language development*: Methods and techniques. Amsterdam, Phildelphia: Benjamins, 2011. p. 55-84.

VAN GEERT, P. A dynamic systems theory model of cognitive and language growth. *Psychological Review*, American Psychological Association, v. 98, n.1, p. 3-53, 1991.

VAN GEERT, P. *Dynamic systems of development*: Change between complexity and chaos. London: Harvester Wheatsheaf, 1994. 300p.

VAN GEERT, P. Dynamic systems approaches and modeling of developmental processes. In: VALSINER, J.; CONOLLY, K. J. (Ed.). *Handbook of developmental psychology*. London: SAGE Publications, 2003, p. 640-672.

VAN GEERT, P. The dynamic systems approach to the study of L1 and L2 acquisition: an introduction. *The Modern Language Journal*, Blackwell Publishers, v. 92, n. 2, p. 179-199, 2008.

VAN GEERT, P. A comprehensive dynamic systems theory of language development. In: DE BOT, K.; SCHRAUF, R. W. (Ed.). *Language development over the life span*. New York: Routledge, 2009. p. 60-104.

VAN GEERT, P.; STEENBEEK, H. The dynamics of scaffolding. *New Ideas in Psychology*, Elsevier BV, v. 23, n. 3, p. 115-128, 2005.

VAN GEERT, P.; STEENBEEK, H.; VAN DIJK, M. A dynamic model of expert-novice co-adaptation during language learning and acquisition. In: SCHMID, M. S.; LOWIE, W. (Ed.). *Modeling bilingualism*: From structure to chaos. Amsterdam, Philadelphia: Benjamins, 2011. p. 235-266.

VAN GEERT, P.; VAN DIJK, M. Focus on variability: New tools to study intra-individual variability in developmental data. *Infant Behavior and Development*, Elsevier BV, v. 25, n. 4, p. 340-374, 2002.

VERSPOOR, M. H.; LOWIE, W. M.; VAN DIJK, M. Variability in L2 development from a dynamic systems perspective. *The Modern Language Journal*, Blackwell Publishers, v. 92, n. 2, p. 214-231, 2008.

WEBB, S. Receptive and productive vocabulary sizes of L2 learners. *Studies in Second Language Acquisition*, Cambridge University Press, v. 30, n. 1, p. 79-95, 2008.

XUE, G.; NATION, I. S. P. A university word list. *Language Learning and Communication*, John Wiley & Sons, v. 3, n. 2, p. 215-229, 1984.