



**HAL**  
open science

## ‘Speaking volumes’: A longitudinal study of lexical and grammatical growth between 17 and 42 months

Florence Labrell, Paul van Geert, Christelle Declercq, Véronique Baltazart, Stéphanie Caillies, Marie Olivier, Sandrine Le Sourn-Bissaoui

### ► To cite this version:

Florence Labrell, Paul van Geert, Christelle Declercq, Véronique Baltazart, Stéphanie Caillies, et al.. ‘Speaking volumes’: A longitudinal study of lexical and grammatical growth between 17 and 42 months. *First Language*, 2014, 34, pp.97 - 124. 10.1177/0142723714526573 . hal-03201806

**HAL Id: hal-03201806**

**<https://hal.science/hal-03201806>**

Submitted on 19 Apr 2021

**HAL** is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L’archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d’enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

# 'Speaking volumes': A longitudinal study of lexical and grammatical growth between 17 and 42 months

First Language

1–28

© The Author(s) 2014

Reprints and permissions:

sagepub.co.uk/journalsPermissions.nav

DOI: 10.1177/0142723714526573

fla.sagepub.com



**Florence Labrell**

Université de Reims, France; U669 INSERM, France

**Paul van Geert**

Heijmans Institute, Netherlands



**Christelle Declercq**

Université de Reims, France



**Véronique Baltazart**

Université de Reims, France



**Stéphanie Caillies**

Université de Reims, France



**Marie Olivier**

Université de Reims, France



**Sandrine Le Sourn-Bissaoui**

Université Rennes 2, France



## Abstract

Dynamic analyses of language growth tell us how vocabulary and grammar develop and how the two might be intertwined. Analyses of growth curves between 17 and

---

### Corresponding author:

Florence Labrell, Université de Reims – Champagne Ardenne, UFR Lettres et Sciences Humaines, Département de Psychologie, 57, rue Pierre Taittinger, F-51096 Reims cedex, France.

Email: [florence.labrell@univ-reims.fr](mailto:florence.labrell@univ-reims.fr)

42 months, based on longitudinal data for 34 children, revealed interesting patterns of vocabulary and grammatical developments. They showed that these patterns were nonlinear, but with coinciding peaks of growth, suggesting a bilateral relationship between acquisition of vocabulary and grammar. A more detailed analysis of specific components of vocabulary (nouns, verbs, grammatical words) and grammar showed that each followed its own developmental course, but that its growth rates were likely to be negatively or positively correlated with those of other components. For example, a faster rate for acquiring nouns coincided with a slower rate for verbs. Last, an assessment of intra-individual variability in three children showed that mean scores obscure individual profiles.

### **Keywords**

Dynamic systems, grammatical development, lexical development

### **Introduction**

Over the past few decades, an impressive body of research has been devoted to explaining the early trajectory of language development, which is characterized by rapid initial learning of the lexicon (first words typically occurring between 8 and 14 months of age), followed by multiword utterances and grammatical development around the child's second birthday. The question of the interdependence of lexical and grammatical growth was first raised more than a decade ago (Bates & Goodman, 1999), but the hypothesis that grammatical acquisition is underpinned by lexical acquisition, while attractive, is still under debate (LeNormand, Moreno-Torres, Parisse, & Dellatolas, 2013).

A widely held view in the literature is that vocabulary initially develops slowly, but once the first 50 words have been acquired, there is an explosive increase – a phenomenon known as *the vocabulary spurt* (Bloom, 1973; Dromi, 1987; Nelson, 1973; for a review, see Dapretto & Bjork, 2000). Thus, there appears to be a switch from a period of continuous progression to a discrete episode of faster growth, characterized by the acquisition of as many as 4–10 new words in a single day. This switch has been variously attributed to cognitive, conceptual and linguistic changes (Li, Zhao, & MacWhinney, 2007; Nazzi & Bertoni, 2003). Several authors, however, have questioned the scope and universality of this spurt (Ganger & Brent, 2004; Reznick & Goldfield, 1992). For example, when Ganger and Brent (2004) followed 20 children in a longitudinal study of early lexical development, using the parental report method, they only found a vocabulary spurt for five of these children. For the other children, the rate of new word production was more gradual. Thus, early lexical development appears to be characterized by inter-individual variability, and the rate at which young children acquire new words may not be as uniform as suggested in the vocabulary spurt literature. There are several alternatives to the theory of a vocabulary spurt. For instance, lexical growth may be just a gradual process, or else there may be not one, but several spurts, with the rate of acquisition varying across the period of interest.

For their part, Boelens and Mollers (2009) noted that the detection of lexical growth may require frequent sampling and that variations may be overlooked if measures are only made, say, once a week. When they studied the development of both the lexicon and grammar in a Dutch boy between the ages of 1;0 and 2;6, they failed to find any evidence for competition between lexical and grammatical growth, on the basis of either monthly or weekly measures. They concluded that the lexicon and grammar may only compete when the lexicon is in a phase of rapid initial growth, and that their measures were insufficiently frequent to highlight such transitions in language development.

The types of early words produced by the children are also a relevant issue in lexical development. Four kinds of early words can be isolated in most languages, including French: paralexical words (pragmatic words used for early communication, such as 'Hello!'), nouns, predicates (verbs and adjectives) and grammatical words (Bassano, Mailllochon, & Eme, 1998). Their respective proportions have been found to change during the early stages of production. More specifically, Bates et al. (1994) identified three waves of reorganization in the composition of the lexicon in English. Nouns predominate in the first wave, which these authors interpreted as an emphasis on reference. Research has shown that nouns are acquired earlier and more rapidly than verbs (Bates, Bretherton, & Snyder, 1988; Bates et al., 1994; Benedict, 1979; Gentner, 1982, 2006; Nelson, 1973). This noun predominance is not universal. For instance, it has not been observed in children acquiring Korean (Choi & Gopnik, 1995; Gopnik & Choi, 1995) or Mandarin (Tardif, 1996; Tardif, Shatz, & Naigles, 1997). It is, however, present across several different languages, including French, Dutch and Spanish (Bassano, 2000; Bornstein et al., 2004; Jackson-Maldonado, Thal, Marchman, Bates, & Gutierrez-Clellen, 1993). Bates et al. (1994) noted that once the English lexicon has grown to about 100 words, the proportion of nouns decreases and the proportion of predicates increases. This is the second wave, corresponding to an emphasis on predication. The third and final wave, indicating a shift toward grammar, begins after the lexicon has grown to around 400 words, and is characterized by a radical increase in the proportion of grammatical words. Studies concerning other languages, such as French or Italian, have corroborated these observations (Bassano, Eme, & Champaud, 2005; Bassano et al., 1998; Bassano, Labrell, Champaud, Lemétayer, & Bonnet, 2005; Caselli et al., 1995).

In summary, the studies discussed above suggest that grammatical development is contingent upon the development of vocabulary. A considerable body of empirical evidence has been amassed to support this hypothesis (Bates & Goodman, 1999; Dromi, 1987). This developmental relationship is explained by theories of language acquisition and demonstrated by computational models that provide evidence for the psychological reality of the mechanisms they embody (Cohen & Chaput, 2002). It has been suggested that the dependency of grammar on the lexicon is subtended by a strong causal relationship (Bates, Dale, & Thal, 1995; Bates et al., 1988; Bates & Goodman, 1997, 1999), and that a modular distinction between grammar and the lexicon has been overstated (Bates & Goodman, 1997). Rather, the various bursts characterizing vocabulary and grammatical growth might be viewed as different phases in a nonlinear wave starting with the child's first word and extending to the appearance of a grammar. Whatever the link between the two acquisition processes, the notion of a critical mass of vocabulary being reached before the grammar's growth is supported by a number of arguments. For a start,

whatever their native language, children can only begin combining words once they have compiled a sufficient lexical stock, following the vocabulary spurt (Bates et al., 1995). In addition, studies using parental reports have found that the size of the lexicon is related to grammar production (Fenson et al., 1994; Marchman & Bates, 1994). Similar findings have been reported for late-talking children and children with atypical development resulting from Williams syndrome or a specific language impairment (SLI) (Bates et al., 1997; Harris, Bellugi, Bates, Jones, & Rossen, 1997; Moyle, Weismer, Evans, & Lindstrom, 2007), as well as for children speaking a language other than English, such as French, Italian, or Spanish (Bassano & van Geert, 2007; Caselli, Casadio, & Bates, 1999; Caselli et al., 1995; Devescovi et al., 2005; Jackson-Maldonado et al., 1993), and bilingual children (Simon-Cerejido & Gutiérrez-Clellen, 2009). Thus, there is a substantial body of literature to support the hypothesis that the production of grammatical structures depends on the extent of expressive vocabulary.

Nonetheless, numerous experiments have indicated that early receptive grammar knowledge is present before grammatical structures are used in production, and even before children produce their first words. These experiments have shown that children aged 16–18 months are sensitive to syntactic elements such as word order, anaphora and the inflectional properties of familiar words (Golinkoff, Hirsh-Pasek, Cauley, & Gordon, 1987; Lidz, Waxman, & Freedman, 2003; Soderstrom, White, Conwell, & Morgan, 2007). Moreover, research conducted within the theoretical framework of bootstrapping has shown that infants exploit this grammatical knowledge to acquire word meaning (Christophe, Millotte, Bernal, & Lidz, 2008; Naigles, 1990). Empirical data demonstrate that infants and toddlers recruit their knowledge of the syntactic frame in which a novel word appears to deduce its meaning (Gillette, Gleitman, Gleitman, & Lederer, 1999; Landau & Gleitman, 1985; Naigles, 1990). In addition to this generativist view of the early use of grammatical categories (Conwell & Demuth, 2007), there is also the idea put forward by Ninio (2006, 2011), whereby children initially possess only a surface knowledge of grammatical relationships – a rudimentary use of grammatical relationships that will later be enriched by cognition and social experiences. A recent study using a part-of-speech tagging analysis coupled with multiple regression analyses showed that between 2 and 4 years of age, grammatical categories are the best predictors of mean length of utterance (MLU) (LeNormand et al., 2013). In other words, grammatical diversity is the best predictor of general language complexity (number of word types) between 2 and 4 years of age, rather than lexical diversity.

Finally, studies of early production have introduced the idea of a bilateral relationship between vocabulary and grammar (Dromi, 1987). In a cross-sectional study investigating grammar and vocabulary with adapted versions of the MacArthur Communicative Development Inventories (CDIs, see Fenson et al., 1994), Dixon and Marchman (2007) showed that the lexicon and grammar develop synchronously in children between 16 and 30 months. Using the same tool in a cross-lagged genetic design, Dionne, Dale, Boivin, and Plomin (2003) showed that lexical growth and grammar correlate strongly in same-sex twins between 2 and 3 years of age. However, as Dixon and Marchman (2007) used means to measure lexical and grammatical growth, they obscured individual patterns and the existence of any temporal dependencies in real-time series. The study by Dionne et al. (2003) similarly failed to demonstrate the existence of

time-series dependencies (cross-sectional and group-based longitudinal studies are only relevant if the timing of growth patterns is extremely similar among individuals, but this condition is not met in real-life data). In contrast, Robinson and Mervis (1998) chose to study the lexical and grammatical development of a single male child (Ari), English speaking, between the ages of 10 and 30 months, focusing specifically on the development of the use of plurals. Using dynamic-systems modelling procedures based on a hypothetical precursor model (van Geert, 1991), these authors showed that the acquisition of plural forms only started once a threshold had been reached in the size of the lexicon. Furthermore, lexical growth slowed as the use of plurals increased, probably because of the additional attention needed to produce words in the plural form. Therefore, the precursor model developed for Ari's lexical and grammatical development showed strong competitive relations between lexical growth and early use of plurals, possibly based on shared learning mechanisms.

To sum up, there have been numerous studies of productive language in children during their first few years of life, yielding different results regarding the links between lexical and grammatical development. Some of the longitudinal studies explored the development of both the lexicon and grammar via direct observations (Bassano et al., 1998; Boelens & Mollers, 2009; Dromi, 1987; Robinson & Mervis, 1998) or parental reports (Dionne et al., 2003; Dixon & Marchman, 2007). Other longitudinal studies considered the lexicon (Ganger & Brent, 2004; Reznick & Goldfield, 1992) but not the onset of grammar, mainly because the duration of the study did not allow later grammatical forms to be measured. The overall aim of these studies was to see whether there is some kind of conditional or bidirectional relationship between lexical and grammatical development. In order to highlight the existence of such a relationship, a longitudinal design is needed, featuring measurements of varying frequency that cover both the lexicon and grammar.

In addition to the existence of general developmental trends, intra-individual variability in lexical development is an important issue, particularly when adopting a dynamic-systems approach (Bassano & van Geert, 2007; van Geert & van Dijk, 2002). Instead of being dismissed as a measurement error, variability should be treated as a potential driving force for development, a source of linguistic experimentation and possible feedback. Variability can also be regarded as an indicator of development, with temporal increases in variability possibly reflecting an underlying developmental transition. It is important to take individual trajectories into account, as classic analyses of group trajectories cannot, in theory, capture individual curves. Intra-individual variations cannot be inferred from statistics pooling inter-individual data without violating ergodicity, namely population homogeneity (i.e. the same statistical model can be applied to each individual), and stationarity (i.e. variance and average are stable across time points) (Molenaar & Campbell, 2009).

Accordingly, the present study explored lexical growth with a focus on individual trajectories within the framework of dynamic-systems theory. This theoretical approach is particularly appropriate for measuring individual differences in language development, especially lexical and grammatical growth (Bassano & van Geert, 2007; Robinson & Mervis, 1998; Ruhland & van Geert, 1998; van Dijk & van Geert, 2005; van Geert, 1991; van Geert & van Dijk, 2002).

When it is applied to aspects of cognition (e.g. language), dynamic-systems theory is not a unitary concept. Rather, it consists of a set of approaches that have a number of basic principles in common relating to cognition and its development, some of which are shared by other approaches, including connectionism (Spencer, Thomas & McClelland, 2009; van Geert & Fischer, 2009). To explain change, dynamic-systems theory focuses on the continuities and discontinuities that occur in systems throughout the course of development. Van Geert (2008) defined a system as ‘any collection of identifiable elements – abstract or concrete – that are somehow related to one another in a way that is relevant to the dynamics we wish to describe’ (p. 180). There have been interesting interpretations of cognitive and language development in terms of dynamic systems (van Geert, 1991, 2008; van Geert & Steenbeek, 2005). Under this theoretical approach, development is described in a state space (i.e. a space of descriptive dimensions) to which a set of dynamic rules is assigned (e.g. the strongly competitive relations between lexical growth and early use of plurals demonstrated in Ari’s case; Robinson & Mervis, 1998). For example, van Geert (1991) described the expansion of vocabulary during the one-word stage in a state space with four dimensions, namely number of acquired words, growth rate (e.g. 0.5–2 words per day), carrying capacity (1–500 words) and feedback delay. Note that *carrying capacity* is the sum of the resources (ranging from linguistic input to memory, motivation, etc.) that support the growth of language in a particular child and, in principle, lead to a stable state of acquisition.

The dynamic approach to development therefore views the next step in a system of variables as a function of the preceding step. A dynamic system changes because the components of that system, and also those of other systems to which it is connected, interact with one another. Specifying the way in which these components interact and hence change, is one of the purposes of dynamic models. A major property of change relates to the iterative nature of development; that is, the principle whereby each successive state is a consequence of the preceding one (van Geert & Steenbeek, 2005). One of the consequences of this principle is that dynamic systems are characterized by self-organization and nonlinearity. Language acquisition is one such self-organizing process (Bassano & van Geert, 2007; van Geert, 2008, 2009).

A child’s linguistic system can be broken down into several components, including phonology, the lexicon and syntax. Each of these components can be described as having an initial state, and its development can be defined as a function of its dependencies on other components. For the purposes of the present study, the components of interest were lexical knowledge and syntactic knowledge. Drawing on longitudinal data for French children between their second and fourth years, we sought to capture the dynamics of lexical growth and its relationship with grammatical knowledge. In line with the principles of dynamic-systems theory, we not only looked at general longitudinal trends in the data, but also focused on individual trajectories, in order to account for a picture of the interplay between our two variables (lexicon and grammar). This longitudinal study spanning 25 months was designed to answer the following questions:

1. *What is the pattern of lexical development between the second and fourth years of life? Does it exhibit a spurt, and if so, is there one spurt or several spurts?*

Except for the question of the vocabulary spurt, to our knowledge no study has ever been conducted to study variability in the rate of new word production in the course of early lexical development. The primary purpose of our research was to investigate the dynamics of vocabulary growth in French children between 17 and 42 months, paying particular attention to growth in different word categories. Until now, research concerning the vocabulary spurt has been predominantly based on nouns or verbs. We therefore sought to include other categories, such as predicates and grammatical words, in order to investigate whether they display a gradual increase, a single spurt or several spurts, as identified by Bates et al. (1994).

2. *What is the pattern of grammatical growth across this period? Is there one spurt, several spurts or no spurt at all?*
3. *Are there any links between lexical and grammatical growth patterns observed over this period?*

The increase in grammatical knowledge may well depend on lexical knowledge. For example, a child needs to attain a threshold number of words in order to be able to combine words. However, lexical growth may momentarily slow down when grammatical knowledge increases, as shown by Robinson and Mervis (1998). Lexical growth may then pick up again, as grammatical knowledge makes it easier to express more complex meanings. In order to study the possible interweaving of lexical and grammatical development, we analysed children's grammatical productions to investigate whether potential increases in the lexicon are related to increases in grammar production. For each child with complete data, we investigated whether increases or decreases in lexical growth acquisition coincided with, preceded, or followed increases in grammatical knowledge.

4. *What is the nature of inter-individual variability in lexical and grammatical growth development? Does every child show the same pattern of lags or interdependencies between the lexicon and grammar, or are these patterns mainly idiosyncratic?*

Since vocabulary growth is a process that, by definition, occurs in individual children, we investigated the similarities and differences between individual curves and the group-based curves. It is likely that individual development curves show highly idiosyncratic patterns that are not reducible to a common underlying trend shown in group data (Molenaar, 2004; Molenaar & Campbell, 2009). In order to check this possibility, we compared the curves of seven individual children with the group curves based on all the observations (34 observations in all). The seven children for whom the individual curves are described are those with no missing data, which enabled the most reliable estimation of their individual growth curves.

## Method

### *Participants*

We studied 34 first-born children (16 girls and 18 boys) between the ages of 17 and 42 months in this longitudinal study. The first evaluation always took place when the child



was about 17 months old (to the nearest 2 weeks). Children were recruited via the list of births issued by the Reims city registry office. The inclusion criteria were being French monolingual with no reported speech, hearing or serious health problem. The majority of families were middle to upper middle class, but the sample included families from nearly the full socioeconomic range.

## Material

We used the DLPF measure (*Développement du Langage de Production en Français*, DLPF; Bassano, Labrell, et al., 2005; Labrell, Bassano, Champaud, Lemétayer, & Bonnet, 2005), a reliable and standardized parental report.<sup>1</sup> The general principle behind this report is the same as in the MacArthur–Bates CDIs, but there are several specific differences. First, the DLPF concerns a different age range, as it assesses the lexical development of children aged 18–42 months. Because of this age range, the DLPF is an appropriate instrument for studying lexical development between the second and fourth years of life. Second, in addition to lexical production, it assesses grammatical and pragmatic abilities, making it possible to investigate the relationship between lexical and grammatical development. Third, unlike the French-language adaptations of the CDIs, the DLPF was constructed on the basis of longitudinal corpora yielded by earlier observational studies of French-speaking children (Bassano et al., 1998). The DLPF was validated with 468 children aged 18–42 months, according to the criteria used by Dale, Fenson, and Thal (1993). Standardization measures were also computed with 468 children aged 18–42 months (Bassano, Labrell, et al., 2005). Four versions were constructed according to the children's age. Given our objectives, we only used the vocabulary section and the sentences and grammar section.

The vocabulary section consists of a checklist of words divided into four categories: socio-pragmatic items (*sound effects and animal sounds, games and routines*), nouns, organized into 13 semantic subcategories (*people, animals, vehicles, toys, food and drink, clothing, body parts, furniture, rooms, natural elements, locations, feelings and emotions, other abstract nouns*), predicates (verbs and adjectives) and grammatical words, organized into five subcategories (*adverbs, question words, articles, pronouns, conjunctions*). The number of words in this section depends on the version: Version 1 comprises 581 words, Version 2 comprises 934 words, Version 3 1232 words and Version 4, 1477 words.

The grammar section, based on the MacArthur–Bates CDIs, comes in three parts: grammatical forms, sentences and complexity. The grammatical forms part checks whether children use grammatical morphology for nouns (determinants) and verbs (inflectional markers). The sentences section simply checks whether children combine words. The complexity section comprises two subsections. In the first one, parents are presented with 32 pairs of utterances, the second utterance in each pair being slightly more complex than the first. Within each pair, parents are asked to choose the utterance that most closely resembles the way their child is talking at that time. The second subsection contains 13 items in Versions 1–3 and 17 items in Version 4. Their purpose is to check whether children produce negation markers (e.g. 'Maman (ne) vient **pas**' [Mummy isn't coming], 'Je n'aime **ni** la salade **ni** les épinards' [I like **neither** lettuce **nor**

spinach]) and grammatical words connecting sentences (e.g. 'Bébé dort **parce qu'**il est fatigué' [Baby is sleeping **because** he's tired]).

## Procedure

The parents completed the questionnaire once a month when the children were aged between 17 and 35 months, and once every two months when the children were aged between 36 and 42 months. To familiarize the parents with completing the DLFP forms, an experimenter was always present the first time, to answer any parental question. Parents were informed that we aimed to collect vocabulary and grammar from children at several ages. That is why their children might not know all the items. From the second session, questionnaires were brought to the parents' homes. For the vocabulary section, parents were asked to tick the boxes corresponding to the words that their child had started to say. For the sentence and grammar sections, they completed the questionnaire in two different ways. Either they indicated if their child had produced a given utterance, or they indicated whether their child had not yet (or never), sometimes, or often, produced a given utterance.

For each child, we had 23 measures. The average number of missing values was six (26%). Five children had no missing values at all, two had only one. The one missing value was replaced by the preceding measurement. By doing so, we obtained a set of seven that were treated as (near) complete cases or 'no missing value cases' (the number of missing values for this group was 1%).

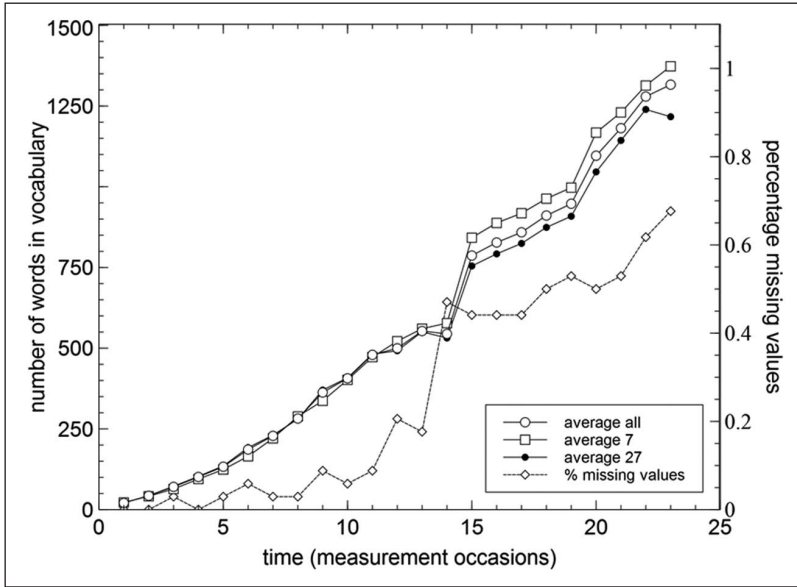
There could be an effect of repeated testing on the parents' evaluations. However, in line with Bates and Goodman (1997) on this potential effect, it can be thought that none of the parents was aware of the hypothesis and they could not be suspected of giving feedback or of encouraging their child.

## Results

### *Analysis of vocabulary data*

Our main question concerned the shape of the vocabulary growth curve based on the means of the 23 measures, and any differences or similarities between this group-based curve and the curves of the individual children. The variables were as follows: nouns, verbs, adjectives, grammatical forms and total vocabulary (i.e. the sum of all the preceding variables). We wanted to find out whether vocabulary growth follows a linear or a nonlinear pattern, and if the pattern is nonlinear, whether it is simple or composite. By *simple pattern* we mean that it follows a continuous, nonlinear pattern, such as an S-shaped curve. A *composite pattern* contains discontinuities, such as spurts and temporary plateaus.

Before presenting an analysis of the individual growth curves, which are based on the data from seven children for whom there are (virtually) no missing values, we need to check whether this group is sufficiently representative of the whole group of children. In order to do so, we present the cumulative average growth curves for this group of seven (1% missing values), for the group of 27 children for whom on average 33% of



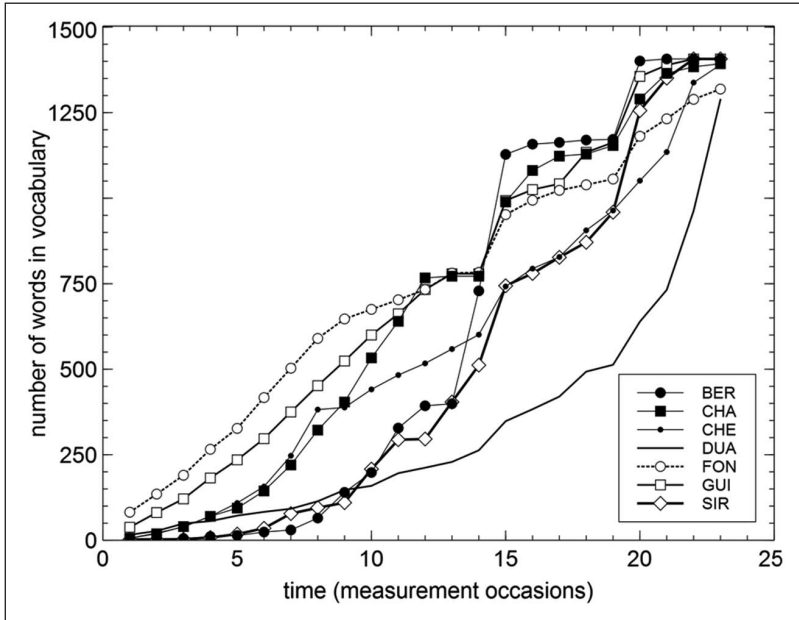
**Figure 1.** Raw data of average vocabulary growth for all children ( $N = 34$ ), compared with average vocabulary growth for the groups of children with and without missing data ( $n = 27$  and  $n = 7$  respectively). The secondary axis (left side) shows the percentage missing values over the course of the 24 measurement occasions.

the measures are missing, and for the whole group of 34 children (26% missing values; see Figure 1).

The graph shows a rapid increase in missing values at measurement 14. This increase in missing values precedes a rapid increase in vocabulary. However, this rapid vocabulary increase cannot be an artefact of the rapid increase in missing values, since it also occurs in the seven children for whom there are no missing values at all. In fact, the similarity between the vocabulary curve based on the children without and those with missing values is striking, especially if one focuses on the qualitative patterns of stepwise changes in the curves. It should be noted however that from measurement 15 onwards, the average curve for children with missing values lags somewhat behind the curve based on complete observation sets, which is only due to the fact that the jump from measurement 14 to 15 is greater in the complete dataset than in the incomplete dataset. In short, from the viewpoint of the underlying growth pattern, there is no reason to believe that the group of children with a complete dataset are clearly different from the group of children with missing data.

Overall, the first growth spurt peaked at around measure 8 and levelled off at around measure 12, while the second peaked at around measure 15 and levelled off at around measure 17, and the third peaked at around measure 20 and levelled off at around measure 23.

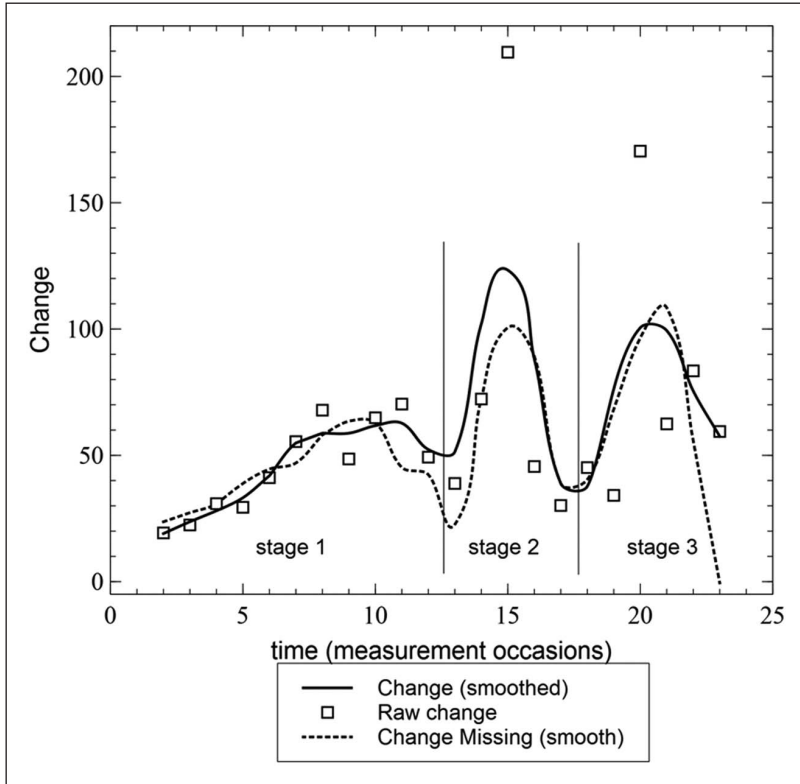
We now present the raw data (cumulative vocabulary) for the children for whom we have a complete dataset, i.e. 23 consecutive measures: BER, CHA, CHE, DUA, FON, GUI and SIR. Visual inspection of the data (see Figure 2) shows that most of the curves



**Figure 2.** Individual growth curves of vocabulary of the seven children who completed the 23 measurements.

follow a pattern of successive spurts and plateaus, although the plateaus were short, and therefore suggesting temporal increases in the growth rates. The curve based on the means of the seven values for each measurement point had three S-shaped components, in the form of a classic, stepwise curve (see Figure 1). A similar pattern occurs in the curve based on the means of the 27 children with missing values. However, a comparison of the group-based curves with the seven individual curves shows that the latter are characterized by idiosyncratic patterns that, at face value, are not all easily reconciled with the group-based curve.

Before discussing the statistical test used to check whether the shape of the curve did indeed correspond to a sequence of three growth spurts, rather than being a simple linear curve with random variations occurring within statistical limits, we need to describe a method for specifying the structure of the curve based on an analysis of the rates of vocabulary change. Since the vocabulary curves were cumulative, the raw rates of change were equal to the differences between successive vocabulary levels, with a minimum of zero. Raw rates of change were calculated for the seven children with all 23 measures. The data were then smoothed by means of the LOESS nonlinear smoother, with a smoothing window of five measures. The smoothing function is designed to reduce small local fluctuations. Figure 3 shows the raw change rates, as well as the smoothed change rate curve, for the seven pooled children. The peaks in the curve correspond to periods of rapid change, whereas the valleys correspond to slow change. The combination of a peak and a valley corresponds to a ‘stage’ in the growth pattern



**Figure 3.** Raw and smoothed change scores of vocabulary growth in seven children without missing values, compared with the smoothed change scores of vocabulary growth in 27 children with missing values.

reflecting a sequence of slow–rapid–slow growth rates. The figure shows that there are three such sequences, corresponding to a three-stage model of growth, with a longer initial stage characterized by less rapid growth, followed by two shorter stages marked by more rapid growth. Figure 3 also shows that the smoothed curve of the change rates based on the group of children with missing values corresponds to the same pattern of three consecutive stages in the growth of vocabulary.

The statistical question is whether the three-stage pattern is real or whether the observed maxima and minima in the fitted curve actually stem from accidental variations within the limits of statistical error. In short, our hypothesis is that the growth curve is a composite curve, consisting of three clearly distinguishable sub-curves each resembling an S-shaped form. In contrast, according to the null hypothesis, the growth curve is a simple, non-composite pattern of vocabulary growth, with variations within a bandwidth of error. According to this null hypothesis, the stages we saw in the data actually arose from random statistical variation. In order to statistically test this null hypothesis, a statistical simulation procedure was required, i.e. a Monte Carlo test which is based on the

individual growth curves. The statistical simulation procedure is explained in the Appendix. On the basis of this statistical simulation, we can conclude that the indicators suggesting three consecutive stages of local growth spurts were unlikely to result from normal statistical variation around a non-composite growth pattern ( $p < .002$ ).

*Summary.* We can conclude that vocabulary growth across the 23 measures followed a nonlinear pattern featuring three local growth spurts, each characterized by a sigmoid increase and a temporary levelling off. This conclusion is based on the mean vocabulary scores of the whole sample, as well as the scores of the set of seven children who had no missing values. Thus, vocabulary growth occurs in several S-shaped patterns, rather than one single one, forming a series of superimposed S-shapes. Dynamic growth theory (van Geert, 1991, 1994, 1998, 2003) explains these patterns on the basis of discontinuities in the underlying carrying capacity of the system (by *system* we mean the language learner and her or his environment). In normally developing systems, the carrying capacity changes as new skills, abilities or opportunities emerge. The emergence of these new supporting resources is reflected in discontinuous changes in carrying capacity, which can result in local growth spurts. The current data suggest that over the course of the 23 measures, two major changes in carrying capacity occurred, one after measurement 12 and the other after measurement 17.

### *Analysis of the grammatical data*

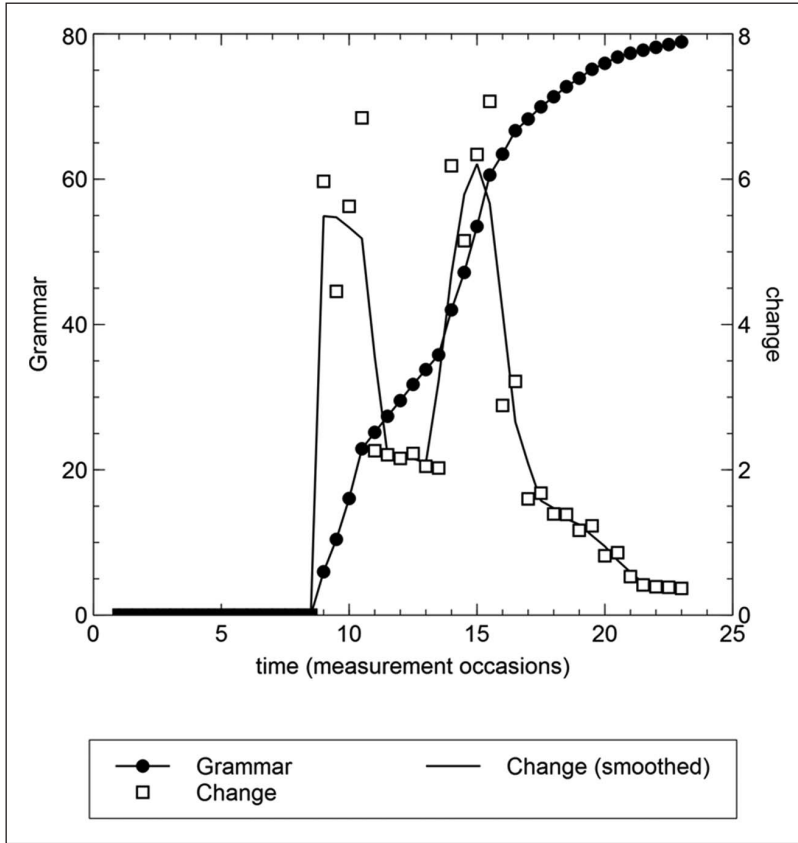
*Based on the seven children.* A comparable analysis of the grammar data for the seven children with full data provided a very similar picture to that yielded by the vocabulary data, in that the first growth peak was smaller than the second. However, the third growth peak observed for vocabulary was absent in the grammar data for the seven children.

*Based on all the children.* The grammatical growth curve based on data for the whole sample was very similar to the curve established for the seven children. Figure 4 shows the grammar growth curve for all the children, together with the curve of the change scores, that is, of the changing growth rates. The peaks shown in this figure correspond to the first and second growth peaks observed for vocabulary. There was no peak corresponding to the third peak in vocabulary.

The growth peaks were again established by taking the difference score of the means (for each measurement point) of the curves of all the children. This difference score was smoothed in order to reduce small local fluctuations and to produce a clearer visual pattern.

### *Links between lexical and grammatical development*

As we were focusing on quantitative time-series data, our analysis of the links between lexical and grammatical development had to be confined to searching for relationships among the quantitative, nonlinear properties of the growth curves. Dynamic growth models (van Geert, 1991, 2008; van Geert & Steenbeek, 2005) can only bring to light a limited number of possible dynamic relationships between any two variables in a network of interacting variables (in the present study, this network consisted of four nodes,

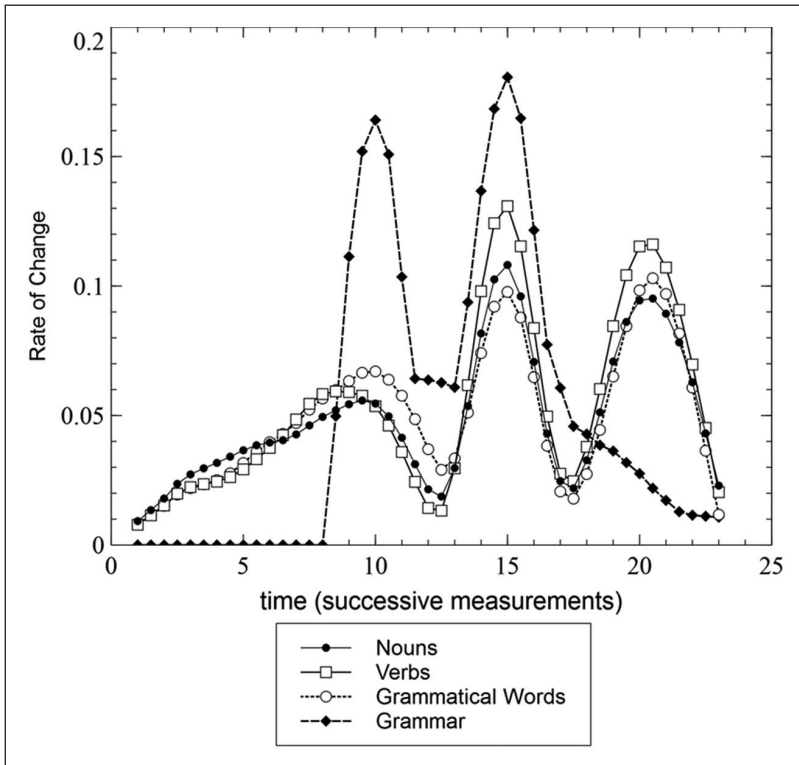


**Figure 4.** Average smoothed growth curves of grammatical data for all children, with smoothed change scores (rate of change) superposed.

The straight line up to about measure 7 reflects the fact that all the measurements before measurement 10 were non-existent (i.e. equal to zero). Smoothing had the effect of extending the early phase of the acquisition slightly towards the measurement points where no measures were actually made.

namely nouns, verbs, grammatical words and grammar). A variable can have a positive, neutral or negative effect on another variable. If this effect is asymmetrical, an increase in the first variable usually precedes an increase in the second if the relationship is positive, and a decrease in the second if the relationship is negative. If there is no relationship between two variables, increases or decreases in the two variables are not coordinated. If the relationship is symmetrical, for instance if one variable has a positive effect on another and vice versa, increases and decreases in the variables occur more or less simultaneously. Hence, the co-occurrence of particular nonlinear properties, such as the moment of maximum change in a variable, can be used as an indicator of an underlying relationship.

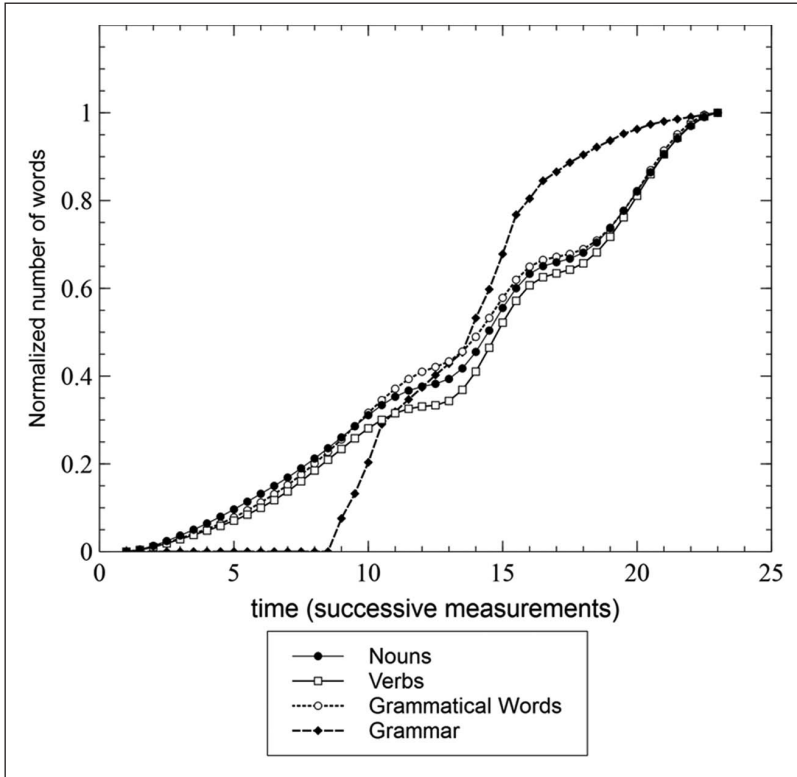
Co-occurrence can be visualized by plotting the growth rate curves of the four variables of interest. Since the magnitude of these four variables differs considerably, we need



**Figure 5.** A comparison of the timing of the growth peaks in nouns, verbs, grammatical words and grammar. The rate of change is based on normalized values of the growth curves, ranging from 0 to 1. In order to determine the absolute value of the rate of change for each variable, the rate must be multiplied by the absolute amount of change, which is 74 for nouns, 269 for verbs, 184 for grammatical words and 79 for grammar.

to normalize the growth curves in order to make a qualitative comparison possible (the total increase in the number of nouns is 740, 269 for verbs, 118 for grammatical words and 79 for grammar). Normalization is performed by converting all growth curves into values between 0 and 1. If we inspect the curves for the rate of change (see Figure 5), we can use the local maximum values as an indicator of the maximum rate of change, and compare the times at which these values occur in order to find potential co-occurrence, sequentiality or delay between the variables. Inspection of the group curves shows that the first acceleration occurs between measurement points 8 and 9 for verbs, between 9 and 10 for nouns, and at measurement point 10 for both grammar and grammatical words. The second acceleration takes place at measurement point 15 for all four variables, and the last between measurement points 20 and 21 for nouns, verbs and grammatical words. In short, the mean curves exhibit a high level of coherence in an important nonlinear indicator, namely the point of maximum change, suggesting a relationship of strong mutual support.

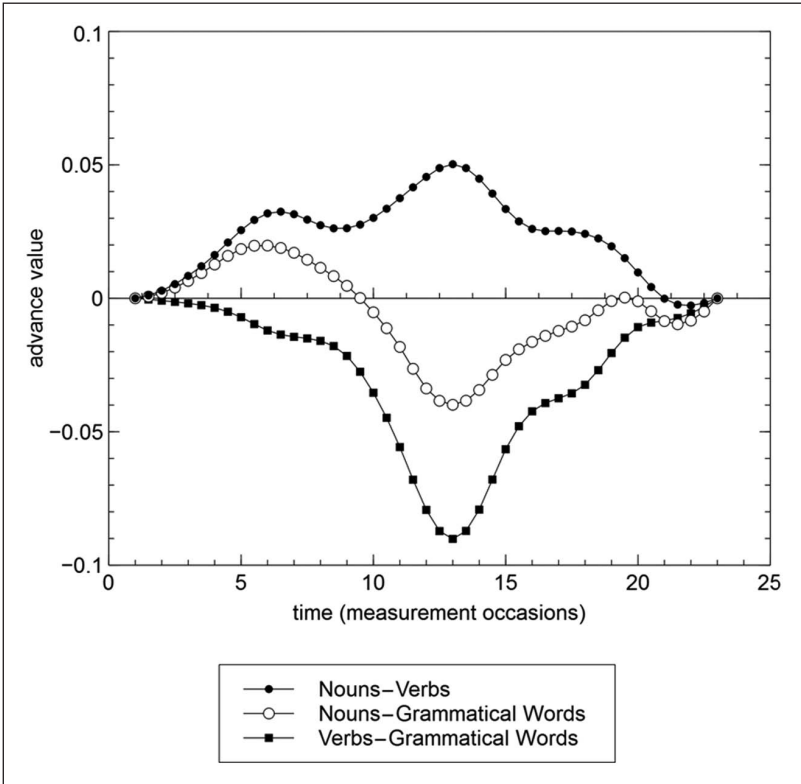




**Figure 6.** Normalized curves of the growth of nouns, verbs, grammatical words and grammar, used for the calculation of the advance value.

However, co-occurrences of nonlinear indicators, such as points of maximum growth, are not the only way in which relationships between variables can be described. Although the point of maximum growth may be the same for two variables (e.g. nouns and verbs), one variable may progress considerably faster than the other, thus in a sense taking the lead over the other. We can define a simple metric expressing the advance of one variable over another, namely the *advance value*, calculated for each pair of variables. The advance value of variable A over B is the difference between the normalized values of A and B for each measurement point. Since the magnitude of the advance values is necessarily based on comparisons between dimensionless variables (variables reduced to values between 0 and 1), the meaning of these values is inferred from the graph in Figure 6, which compares the growth patterns for the four variables of interest by means of normalized growth curves.

Figure 7 shows the advance values for the variables nouns over verbs, nouns over grammatical words and verbs over grammatical words respectively. Nouns showed a faster rate of development than verbs across all 23 measures. Nouns also developed faster than grammatical words up to measurement point 10, after which grammatical



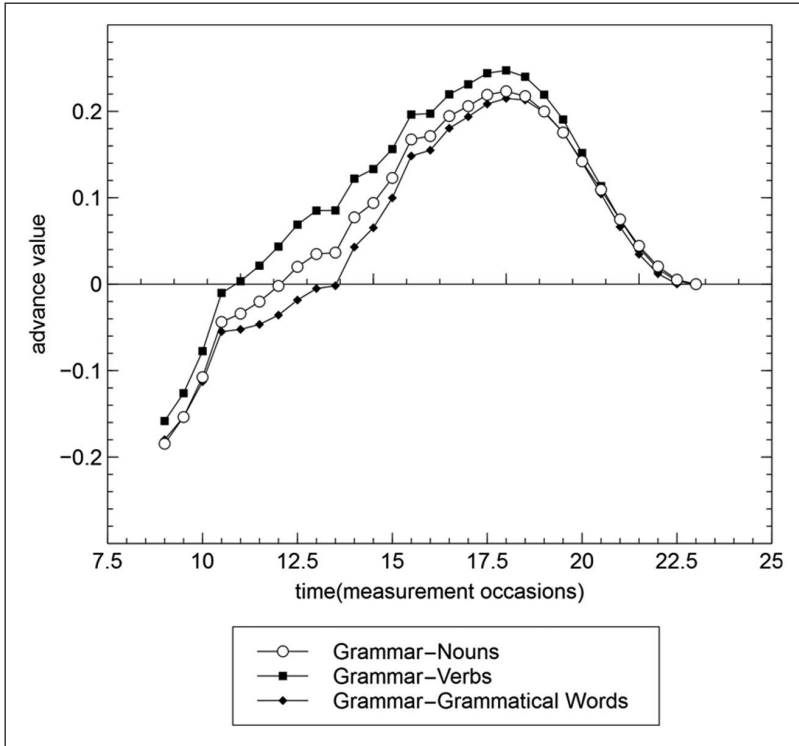
**Figure 7.** Advance values for the comparison nouns–verbs, nouns–grammatical words and verbs–grammatical words.

words tend to develop faster than nouns for the rest of the measurement period. Finally, verbs developed at a slower pace than grammatical words across all 23 measures.

Figure 8 presents a similar comparison to identify the relationship between grammar and the three vocabulary variables.

However, as the measurement of grammatical development only starts at measurement point 9, we can only compare the variables for the period between measures 9 and 23. Although grammar lags behind nouns, verbs and grammatical words for the first four measures (after point 9), grammar then takes the lead. This initial lag is, in a sense, an artefact of the late start of the grammar measures, compared with the vocabulary variables, but the rapid growth of grammar compared to that of nouns, verbs and grammatical words is entirely genuine, showing that, relative to the progressing of the vocabulary variables, grammar undergoes a significant growth spurt beginning at around measure 13.

The method of advance values – indicating which variable is running ahead of which other one – can also be applied to the normalized change scores, to determine which variable was developing faster than another at a particular moment in time. With the four

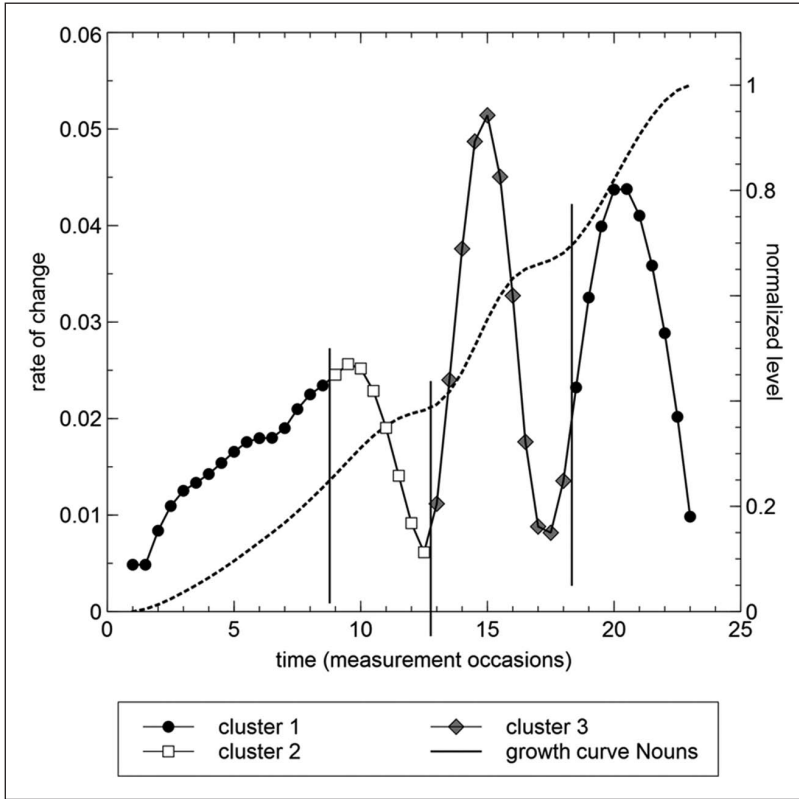


**Figure 8.** Advance values of grammar in comparison with three vocabulary variables: nouns, verbs and grammatical words.

variables (nouns, verbs, grammatical words and grammar), we obtained six time series of advance value variables (e.g. N-V). We can wonder whether these time series show some sort of underlying structure, in the sense of patterns of relationships between these time series that change over the course of the 23 measures. In order to find out whether such patterning occurs, we included these variables in a hierarchical cluster analysis. The analysis yielded three clusters as the best possible solution, corresponding to qualitatively different phases in the growth of nouns, verbs, grammatical words and grammar (see Figure 9). Clustering was carried out with the TANAGRA software package (Rakotomalala, 2012).

The first phase begins at the first measurement point and ends before measure 9, the second runs from measure 9 to 13, the third from 14 to 18, and the fourth from 19 to 23. The first and fourth phases are qualitatively similar, in that they belong to the same cluster (we discuss why this is so in our description of the cluster properties).

If we project the clusters onto the graph of the growth of a particular variable (e.g. nouns), we can see that the clusters correspond quite nicely to the flex points in the graph. The one exception concerns the period between measures 1–13, where there is a succession of two clusters, the first of which probably stems from the fact that



**Figure 9.** Clustering of advance value time series for nouns, verbs, grammatical words and grammar. Hierarchical cluster analysis yields three clusters as the best possible solution. The first and last clusters are qualitatively similar. The first and second cluster correspond to the first growth spurt, the third cluster corresponds to the second growth spurt, and the last cluster corresponds to the third growth spurt.

grammar was not measured before week 9. It is the combination of clusters 1 and 2 that corresponds to the first growth spurt period in the growth of the four vocabulary variables.

The clusters are characterized by the following properties. The first cluster, which appears during the first and last stages, is characterized by a low rate of change in grammar, compared to the vocabulary variables. This is why the first cluster applies to the first as well as to the last stage, as grammatical development was not measured before measure 9 and was thus represented by zero change. It also shows a characteristic decline in the growth rate toward the end of the measurement period. The second phase, which begins with the measurement of the growth of grammar (measure 9), is characterized by a rapid growth of grammar relative to the vocabulary variables, and by rapid growth in grammatical words relative to nouns and verbs. The third phase corresponds to the third cluster and is characterized by rapid growth in grammar relative to the

vocabulary variables, relatively slow growth in grammatical words relative to the growth in verbs, nouns and grammar, and faster growth in verbs relative to the growth in nouns. The appearance of these clusters, corresponding to change points (or flex points) in the quantitative data, suggests that dynamic connections between variables – for instance, the amount of support afforded by one variable to another – may change in the course of development. Exactly how and why these relationships change will be a matter for further investigation.

### *Individual patterns of relationships*

As explained earlier for ergodicity, we know that statistical patterns derived from group data are not necessarily applicable to the individuals who make up the sample (Molenaar & Campbell, 2009). In order to compare the group data with those of individual children, we analysed two randomly selected cases, based on the alphabetical order of the participants' names, and performed the same analyses as with the group data (see Figure 10).

Our comparison of the growth acquisition rates of the two children with those of the group as a whole revealed a number of similarities and differences. For instance, the first child (C) displays a relatively chaotic pattern, with relatively little coherence in the growth peaks for the variables. The second child (child B in the figure), on the other hand, exhibits a highly coherent pattern of peaks. However, if we compare this child with the group-based curves of the growth peaks, we observe that the first peak that characterizes the group data is virtually absent in the second child. This second child thus provides a good illustration of the idiosyncratic nature of the growth curves. Finally, a cluster analysis based on the advance values of the growth rates yielded clusters for each of the two children that were comparable to those of the group, but there were characteristic differences in the individuals as to the exact positioning of the clusters along the time axis.

This temporal shift was particularly noticeable for the second and third clusters. A full exploration of early language development, in particular with regard to the coherence of key development phenomena such as peaks, requires an analysis of individual cases, of which we have presented two examples. However, a further analysis of the range of meaningful individual differences exceeds the scope of the present article.

## **Discussion**

The aim of the study was to capture the dynamics of lexical growth and its relationship with grammatical knowledge. Analyses conducted within the framework of dynamic-systems theory provided the following answers to our questions concerning vocabulary and grammar development, and the relationship between the two.

First, the results provided interesting outcomes about lexical and grammatical growth patterns, in terms of various spurts. In this study, by envisaging these developments simultaneously, we have managed to focus on the connections between lexical and grammatical growths. By doing so, we were able to combine the first and second aims of the present study. Vocabulary growth across the 23 measures followed a non-linear pattern, comprising three stages in the form of local growth spurts,

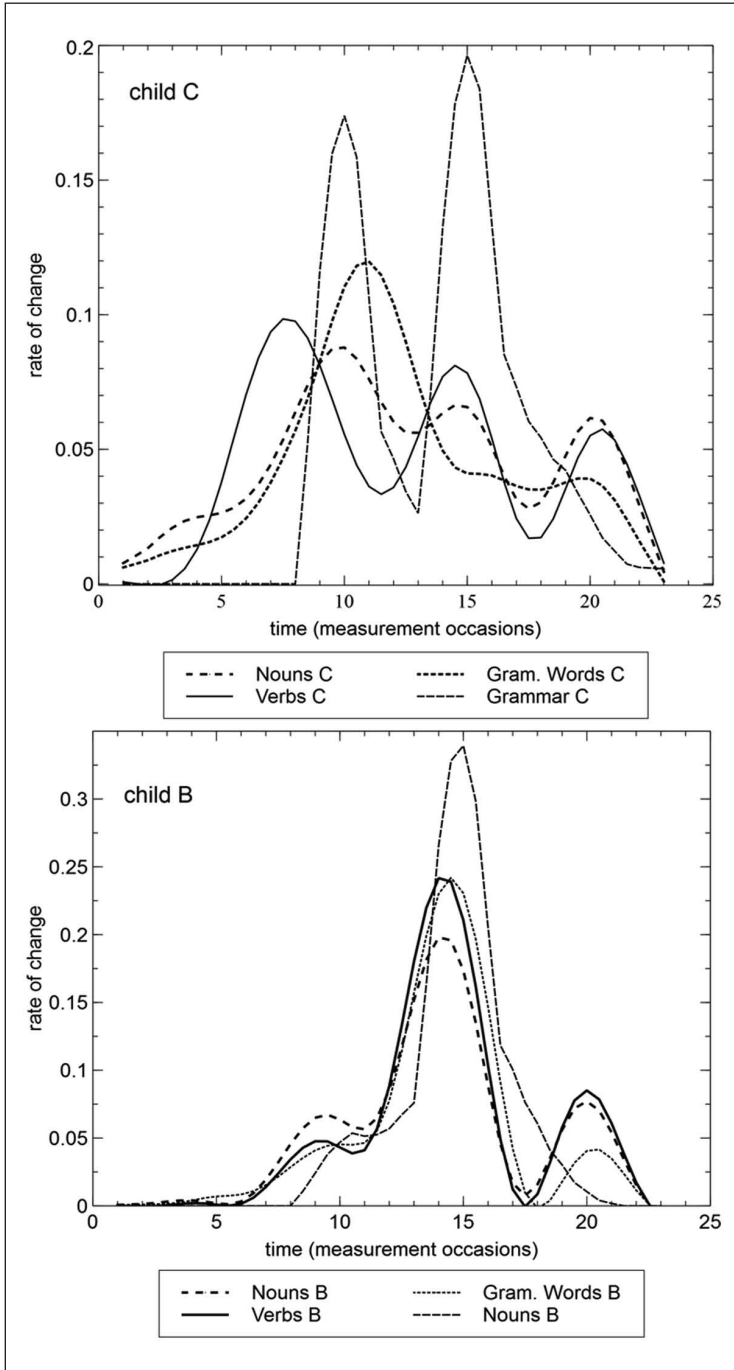


Figure 10. Growth rate patterns in two individual children.

each characterized by a sigmoid increase and temporary levelling off. Grammatical development showed two growth peaks, comparable to those for vocabulary, the first being lower than the second. Results suggested that over the course of the 23 measures, two major changes in carrying capacity occurred, one after measure 12, the other after measure 17, that is, around 28 and 33 months, respectively. We suggest that a vocabulary spurt occurs around 28 months, followed by a grammar spurt around 33 months, combined with a slowing in vocabulary acquisition, which is consistent with previous results based on direct observations in both French (Bassano, Eme, & Champaud, 2005) and American studies (Bates & Goodman, 1999; Dromi, 1987). However, the coincidence of the lexical and the grammatical spurts points to strong mutual links between both growth trajectories, which are consistent with the hypothesis of a bilateral relationship between vocabulary and grammar (Dionne et al., 2003; Dixon & Marchman, 2007; Robinson & Mervis, 1998).

Unlike the majority of longitudinal studies on lexical and grammatical growths, our study focused on the analysis of various vocabulary components separately, but also in connection one with another. It revealed a faster rate of development for nouns, which may however be overestimated in parental reports (Bassano, Labrell, et al., 2005). This was accompanied by a slower growth rate for verbs, which is coherent with the notion of noun–verb asynchrony (Bassano, Eme, & Champaud, 2005; Bates et al., 1988, 1994; Benedict, 1979; Gentner, 1982, 2006; Nelson, 1973). When the growth acquisition rate for nouns started to decrease, grammatical words, required for sentence organization, started to increase faster.

Our third aim concerned the possible links between lexical and grammatical development between 17 and 42 months. Although our indirect data did not enable us to measure the potential links statistically, we were nevertheless able to analyse similarities in the growth curves and the growth rate curves for the four variables. These analyses showed that verbs, nouns, grammatical words and grammar are fairly unlikely to be independent variables. If they were independent, the patterns of change in growth rates would also be independent from one another. And if this were so, the observed coherence between the growth patterns would be hard to explain. This coherence could possibly be ascribed to some underlying, and as yet unknown, latent variable driving the development of each of these variables, and also responsible for the small differences in timing, unless these differences are deemed to be entirely accidental. Although this alternative explanation cannot be ruled out on the basis of the data, an explanation based on a pattern of dynamic interactions between the variables is more in line with what we know about the relationships between lexical and grammatical growth. In other words, our results are consistent with the hypothesis of strong competitive relationships between lexical and grammatical acquisitions, as shown previously in the longitudinal study on children aged 10–30 months by Robinson and Mervis (1998).

The fourth aim of the study was to look for inter-individual variability in lexical and grammatical growth based on dynamic analysis. Results showed that the cluster analysis based on the advance values of the growth rates of two children randomly chosen in the database was sufficiently similar to the pattern observed in the group data to warrant the use of group-based temporal patterns as approximate models of individual trajectories. However, the dynamic analysis of the individual patterns also revealed characteristic

idiosyncrasies in addition to the general trends of lexical and grammatical development in the current sample.

Nevertheless, the possible limitations of the study need also to be envisaged, as well as further investigations to overcome them. First, the indirect data collection via parental reports may have given a distorted picture of actual vocabulary growth. However, the fact that the current findings are coherent with direct observations collected previously and across shorter time spans provides support for the reliability of the parental report data. Nonetheless, to analyse more precisely how vocabulary and grammar are linked, our quantitative approach needs to be supplemented with qualitative and content-based measures of lexical and grammatical production, and indirect assessments should contain more details about the first grammatical words used by children. For instance, a recent study highlighted the role of specific grammatical categories, such as personal pronouns (accounting for 66% of the variance in MLU), prepositions (accounting for 62%) and determinants (accounting for 56%) (LeNormand et al., 2013). Even though these word categories are more difficult to identify, they would be useful indicators.

Another limitation concerns the methodology directly inspired from dynamic systems, which are defined as systems of variables whose interactions determine their respective growth patterns. In order to actually test whether the patterns we observed were the result of specific dynamic relationships between the variables, we would have to extend the current approach, first with a far more refined analysis of the children's actual linguistic productions, and second with the construction of a dynamic model in which the hypothesized relationships could be implemented. Finally, the study of curves based on data aggregated across many individuals is only the first step towards understanding the nature of the dynamic relationships between different variables. The next step is to analyse and describe the growth pattern of each individual and to compare them one with another, and with the aggregated growth curve, in order to check the universality of the underlying dynamic model. In the present article, we provide an example of two individual growth patterns to illustrate how this analysis might be done. One of the advantages of dynamic models is that they can explain a variety of individual trajectories based on one underlying model of dynamic interactions. In future analyses of our data, we hope to address these important issues.

## Funding

This work was supported by funding from the University of Reims (2006), and funding from the Champagne Ardenne regional council (2008), obtained by the first author.

## Note

1. The development of the DLPF questionnaire was funded by the French Research Ministry (Cognitique Programme).

## References

- Bassano, D. (2000). Early development of nouns and verbs in French: Exploring the interface between lexicon and grammar. *Journal of Child Language*, 27, 521–559.



- Bassano, D., Eme, E., & Champaud, C. (2005). A naturalistic study of early lexical development: General processes and inter-individual variations in French children. *First Language*, 25, 67–101.
- Bassano, D., Labrell, F., Champaud, C., Lemétayer, F., & Bonnet, P. (2005). Le DLPF: Un nouvel outil pour l'évaluation du langage de production en français. *Enfance*, 57, 171–208.
- Bassano, D., Maillachon, I., & Eme, E. (1998). Developmental changes and variability in early lexicon: A study of French children's naturalistic productions. *Journal of Child Language*, 25, 493–531.
- Bassano, D., & van Geert, P. (2007). Modeling continuity and discontinuity in utterance length: A quantitative approach to changes, transitions and intraindividual variability in early grammatical development. *Developmental Science*, 10, 588–612.
- Bates, E., Bretherton, I., & Snyder, L. (1988). *From first words to grammar*. New York, NY: Cambridge University Press.
- Bates, E., Dale, P. S., & Thal, D. (1995). Individual differences and their implications for theories of language development. In P. Fletcher, & B. MacWhinney (Eds.), *Handbook of child language* (pp. 95–151). Oxford, UK: Basil Blackwell.
- Bates, E., & Goodman, J. C. (1997). On the inseparability of grammar and the lexicon: Evidence from acquisition, aphasia, and real-time processing. *Language and Cognitive Processes*, 12, 507–584.
- Bates, E., & Goodman, J. C. (1999). On the emergence of grammar from the lexicon. In B. MacWhinney (Ed.), *The emergence of language* (pp. 29–79). Mahwah, NJ: Lawrence Erlbaum.
- Bates, E., Marchman, V., Thal, D., Fenson, L., Dale, P. S., Reznick, J. S., . . . & Hartung, J. (1994). Developmental and stylistic variation in the composition of early vocabulary. *Journal of Child Language*, 21, 85–123.
- Bates, E., Thal, D., Trauner, D., Fenson, J., Aram, D., Eisele, J., & Nass, R. (1997). From first words to grammar in children with focal brain injury. *Developmental Neuropsychology*, 13, 275–343.
- Benedict, H. (1979). Early lexical development: Comprehension and production. *Journal of Child Language*, 6, 183–200.
- Bloom, L. (1973). *One word at a time: The use of single word utterances before syntax*. The Hague, The Netherlands: Mouton.
- Boelens, H., & Mollers, J. (2009). In search for competition between lexical and grammatical growth. *Psychological Reports*, 104, 407–417.
- Bornstein, M. H., Cote, L. R., Maital, S. P. K., Park, S. Y., Pascual, L., Pêcheux, M. G., . . . Vyt, A. (2004). Cross-linguistic analysis of vocabulary in young children: Spanish, Dutch, French, Hebrew, Italian, Korean, and American English. *Child Development*, 75, 1115–1139.
- Caselli, M., Bates, E., Casadio, P., Fenson, J., Fenson, L., Sanderl, J., & Weir, J. (1995). A cross-linguistic study of early lexical development. *Cognitive Development*, 10, 159–199.
- Caselli, M., Casadio, P., & Bates, E. (1999). A comparison of the transition from first words to grammar in English and Italian. *Journal of Child Language*, 26, 69–111.
- Choi, S., & Gopnik, A. (1995). Early acquisition of verbs in Korean: A cross-linguistic study. *Journal of Child Language*, 22, 497–529.
- Christophe, A., Millotte, S., Bernal, S., & Lidz, J. (2008). Bootstrapping lexical and syntactic acquisition. *Language and Speech*, 51, 61–75.
- Cohen, L. B., & Chaput, H. H. (2002). Connectionist models of infant perceptual and cognitive development. *Developmental Science*, 5, 173–175.
- Conwell, E., & Demuth, K. (2007). Early syntactic productivity: Evidence from dative shift. *Cognition*, 103, 163–179. doi:10.1016/j.cognition.2006.03.003

- Dale, P. S., Fenson, L., & Thal, D. (1993). *Some suggestions for the adaptation of the MacArthur Communication Development Inventories to additional languages*. Unpublished manuscript.
- Dapretto, M., & Bjork, E. L. (2000). The development of word retrieval abilities in the second year and its relation to early vocabulary growth. *Child Development, 71*, 635–648.
- Devescovi, A., Caselli, M. C., Marchione, D., Pasqualetti, P., Reilly, J., & Bates, E. (2005). A crosslinguistic study of the relationship between grammar and lexical development. *Journal of Child Language, 32*, 759–786.
- Dionne, G., Dale, P. S., Boivin, M., & Plomin, R. (2003). Genetic evidence for bidirectional effects of early lexical and grammatical development. *Child Development, 74*, 394–412.
- Dixon, J. A., & Marchmann, V. A. (2007). Grammar and the lexicon: Developmental ordering in language acquisition. *Child Development, 78*, 190–212.
- Dromi, E. (1987). *Early lexical development*. Cambridge, UK: Cambridge University Press.
- Fenson, L., Dale, P. S., Reznick, J. S., Bates, E., Thal, D. J., & Pethick, S. J. (1994). Variability in early communicative development. *Monographs of the Society for Research in Child Development, 59*(5), 1–173; discussion: 174–185.
- Ganger, J., & Brent, M. R. (2004). Reexamining the vocabulary spurt. *Developmental Psychology, 40*, 621–632.
- Gentner, D. (1982). Why nouns are learned before verbs: Linguistic relativity versus natural partitioning. In S. A. Kuczaj (Ed.), *Language development* (Vol. 2, Language, thought and culture, pp. 301–334). Hillsdale, NJ: Lawrence Erlbaum.
- Gentner, D. (2006). Why verbs are hard to learn. In K. Hirsh-Pasek, & R. M. Golinkoff (Eds.), *Action meets word: How children learn verbs* (pp. 544–564). New York, NY: Oxford University Press.
- Gillette, J., Gleitman, H., Gleitman, L., & Lederer, A. (1999). Human simulations of vocabulary learning. *Cognition, 73*, 135–176.
- Golinkoff, R. M., Hirsh-Pasek, K., Cauley, K. M., & Gordon, L. (1987). The eyes have it: Lexical and syntactic comprehension in a new paradigm. *Journal of Child Language, 14*, 23–45.
- Gopnik, A., & Choi, S. (1995). Names, relational words, and cognitive development in English and Korean speakers: Nouns are not always learned before verbs. In M. Tomasello, & W. E. Merriman (Eds.), *Beyond names for things: Young children's acquisition of verbs* (pp. 63–80). Hillsdale, NJ: Lawrence Erlbaum.
- Harris, N. G. S., Bellugi, U., Bates, E., Jones, W., & Rossen, M. (1997). Contrasting profiles of language development in children with Williams and Down syndromes. *Developmental Neuropsychology, 13*, 345–370.
- Jackson-Maldonado, D., Thal, D., Marchman, V., Bates, E., & Gutierrez-Clellen, V. (1993). Early lexical development in Spanish-speaking infants and toddlers. *Journal of Child Language, 20*, 523–549.
- Labrell, F., Bassano, D., Champaud, C., Bonnet, P., & Lemétayer, F. (2005). L'évaluation du développement lexical entre 1 et 4 ans: Présentation du DLFP. In F. Grossmann, M. A. Paveau, & G. Petit (Eds.), *Didactique du lexique: Langue, cognition, discours* (pp. 51–62). Grenoble, France: ELLUG.
- Landau, B., & Gleitman, L. (1985). *Language and experience: Evidence from the blind child*. Cambridge, MA: Harvard University Press.
- LeNormand, M. T., Moreno-Torres, I., Patisse, C., & Dellatolas, G. (2013). How do children acquire early grammar and build multiword utterances? A corpus study of French children aged 2 to 4. *Child Development, 84*, 647–661.
- Li, P., Zhao, X., & MacWhinney, B. (2007). Dynamic self-organization and early lexical development in children. *Cognitive Science, 31*, 581–612.

- Lidz, J., Waxman, S., & Freedman, J. (2003). What infants know about syntax but couldn't have learned: Experimental evidence for syntactic structure at 18 months. *Cognition*, *89*, B65–B73.
- Marchman, V. A., & Bates, E. (1994). Continuity in lexical and morphological development: A test of the critical mass hypothesis. *Journal of Child Language*, *12*, 339–366.
- Molenaar, P. C. (2004). A manifesto on psychology as idiographic science: Bringing the person back into scientific psychology, this time forever. *Measurement*, *2*, 201–218.
- Molenaar, P. C., & Campbell, C. G. (2009). The new person-specific paradigm in psychology. *Current Directions in Psychological Science*, *18*, 112–117.
- Moyle, M. J., Weismer, S. E., Evans, J. L., & Lindstrom, M. J. (2007). Longitudinal relationships between lexical and grammatical development in typical and late-talking children. *Journal of Speech, Language, and Hearing Research*, *50*, 508–528.
- Naigles, L. G. (1990). Children use syntax to learn verb meanings. *Journal of Child Language*, *17*, 357–374.
- Nazzi, T., & Bertoncini, J. (2003). Before and after the vocabulary spurt: Two modes of word acquisition? *Developmental Science*, *6*, 136–142.
- Nelson, K. (1973). Structure and strategy in learning to talk. *Monographs of the Society for Research in Child Development*, *38*(1–2), 1–135.
- Ninio, A. (2006). *Language and the learning curve: A new theory of syntactic development*. Oxford, UK: Oxford University Press.
- Ninio, A. (2011). *Syntactic development, its input and output*. Oxford, UK: Oxford University Press.
- Rakotomalala, R. (2012). *TANAGRA: Un logiciel gratuit pour l'enseignement et la recherche*. In Actes de EGC'2005, RNTI-E-3, Vol. 2, pp. 697–702, 2005. Retrieved from <http://eric.univ-lyon2.fr/~ricco/tanagra/en/tanagra.html>
- Reznick, J. S., & Goldfield, B. A. (1992). Rapid change in lexical development in comprehension and production. *Developmental Psychology*, *28*, 406–413.
- Robinson, B. F., & Mervis, C. B. (1998). Disentangling early language development: Modeling lexical and grammatical acquisition using an extension of case-study methodology. *Developmental Psychology*, *34*, 363–375.
- Ruhland, R., & van Geert, P. (1998). Jumping into syntax: Transitions in the development of closed class words. *British Journal of Developmental Psychology*, *16*, 65–95.
- Simon-Cerejido, G., & Gutiérrez-Clellen, V. F. (2009). A cross-linguistic and bilingual evaluation of the interdependence between lexical and grammatical domains. *Applied Psycholinguistics*, *30*, 315–337.
- Soderstrom, M., White, K. S., Conwell, E., & Morgan, J. L. (2007). Receptive grammatical knowledge of familiar content words and inflection in 16-month-olds. *Infancy*, *12*(1), 1–29.
- Spencer, J., Thomas, M. S. C., & McClelland, J. L. (2009). *Toward a new unified theory of development: Connectionism and dynamical systems theory re-considered*. Oxford, UK: Oxford University Press.
- Tardif, T. (1996). Nouns are not always learned before verbs: Evidence from Mandarin speakers' early vocabularies. *Developmental Psychology*, *32*, 492–504.
- Tardif, T., Shatz, M., & Naigles, L. (1997). Caregiver speech and children's use of nouns versus verbs: A comparison of English, Italian, and Mandarin. *Journal of Child Language*, *24*, 535–565.
- van Dijk, M., & van Geert, P. (2005). Disentangling behavior in early child development: Interpretability of early child language and its effect on utterance length measures. *Infant Behavior & Development*, *28*, 99–117.
- van Geert, P. L. C. (1991). A dynamic systems model of cognitive and language growth. *Psychological Review*, *98*, 3–53.

- van Geert, P. L. C. (1994). *Dynamic systems of development: Change between complexity and chaos*. New York, NY: Harvester.
- van Geert, P. L. C. (1998). A dynamic systems model of basic developmental mechanisms: Piaget, Vygotsky and beyond. *Psychological Review*, *105*, 634–677.
- van Geert, P. L. C. (2003). Dynamic systems approaches and modeling of developmental processes. In J. Valsiner, & K. J. Conolly (Eds.), *Handbook of developmental psychology* (pp. 640–672). London, UK: Sage.
- van Geert, P. L. C. (2008). The dynamic systems approach in the study of L1 and L2 acquisition: An introduction. *The Modern Language Journal*, *92*, 179–199.
- van Geert, P. L. C. (2009). Transmission, self-organization, and the emergence of language: A dynamic systems point of view. In U. Schonpflug (Ed.), *Cultural transmission: Psychological, developmental, social, and methodological aspects* (pp. 48–69). New York, NY: Cambridge University Press.
- van Geert, P. L. C., & Fischer, K. W. (2009). Dynamic systems and the quest for individual-based models of change and development. In J. P. Spencer, M. S. C. Thomas, & J. McClelland (Eds.), *Toward a new grand theory of development? Connectionism and dynamic systems theory reconsidered* (pp. 313–336). Oxford, UK: Oxford University Press.
- van Geert, P. L. C., & Steenbeek, H. (2005). Explaining after by before. Basic aspects of a dynamic systems approach to the study of development. *Developmental Review*, *25*, 408–442.
- van Geert, P. L. C., & van Dijk, M. (2002). Focus on variability: New tools to study intra-individual variability in developmental data. *Journal of Infant Behavior and Development*, *24*, 1–35.

## Appendix

According to the null hypothesis, apart from being non-composite, the growth curve for vocabulary can take on a variety of forms. It might be linear, exponentially increasing, exponentially decreasing, or S-shaped. This family of possible non-composite growth curves might be fitted by a single mathematical format, namely that of a sigmoid curve. Thus, in order to test whether the observed growth curve is statistically indistinguishable from a non-composite growth curve, we first fitted a sigmoid model to each of the seven children for whom there are no missing data.

Sigmoid curves can capture a variety of curves, ranging from S-shaped curves, to exponentially increasing or decreasing curves, and linear patterns, which are the patterns that are likely to occur if the growth trajectory is non-composite. For each of the fitted curves, we calculated the residuals (difference between the observed and predicted values), which were symmetrically distributed along the predicted values. These residuals provide an estimation of the measurement error for each of the seven children separately, under the null-assumption that the real growth curve is non-composite. Next, we fitted a model of the residuals, with vocabulary level as the predictor. This model yielded the expected values for the residuals and for the standard deviations of the residuals. Based on the estimated non-composite growth model for each child and the corresponding residuals model, we were able to simulate 200 stochastic growth curves for each child, by adding (or subtracting) a random residual drawn from the residuals model to each of the expected values for vocabulary (each of the simulated growth curves represents the child's 'real' growth curve under the null hypothesis plus or minus measurement error).

The statistical test for the occurrence of sub-stages (characterized by significant reductions in the growth rate) proceeded as follows. For each of the seven children, we randomly selected one out of the 200 simulated datasets (it can be recalled that these datasets represent measurement error variations around a child's presumed real, non-composite growth curve). For each of these seven simulated datasets, we then calculated the average change score curve, just as we had done for the real, observed scores. This process of selection of seven datasets was repeated 500 times, thus yielding 500 change curves based on the null hypothesis model. The 500 curves were then smoothed in exactly the same way as the observed change curve, thus statistically reducing local fluctuations in the same way as we did for the observed data. We then checked whether the simulated curves, based on the null-hypothesis model of non-composite growth, showed statistical variations in the growth rate comparable to those that we took as evidence for the existence of sub-stages. The occurrence of significant and systematic variations in the change rate, indicative of the occurrence of distinguishable growth stages, was signalled by the magnitude of the local slopes of the change curve (*local* was defined here as covering three or four measurement points). If the local slopes in the change curves based on the null hypothesis were of a similar magnitude to those in the observed change curve, we could conclude that our indicators for growth division resulted from normal statistical variation around non-composite curves (i.e. curves not exhibiting these sub-stages). The number of occurrences needed to conclude that the sub-stage indicators were the result of just such random variation was based on normal significance testing, and was defined as  $p \geq .05$ .

Calculation of local slopes in the change curves based on the null-hypothesis model revealed that none of the local slopes that occurred in the 500 curves was of the same magnitude as those observed in the real change curves. We could thus conclude that the growth stage indicators were unlikely to result from normal statistical variation around non-composite growth curves ( $p < .002$ ).