Grammar and the Lexicon: Developmental Ordering in Language Acquisition

James A. Dixon
University of Connecticut

Virginia A. Marchman
Stanford University

Recent accounts of language acquisition propose that the knowledge structures that comprise language develop within a single, unified system that shares computational resources and representations. One implication of this approach is that developmental relations within the system become central to theorizing about language acquisition. Previous work suggested that lexical development preceded grammatical development, a developmental ordering with strong theoretical implications. One purpose of the current article is to test this developmental ordering hypothesis. Results showed that children (aged 16–30 months) developed lexicon and grammar synchronously. The second purpose is to demonstrate a recently developed method for testing developmental ordering, the nonlinear-mapping approach, and show how the method can be extended to capitalize on multiply determined developmental systems, such as language.

Developmental ordering is a crucial link between developmental theory and data. The order in which skills, abilities, structures, and so forth develop is a fundamental source of evidence in developmental science. Developmental theories are constrained by the ontogenetic orderings that are empirically observed. Likewise, our theories and hypotheses are also often motivated by the order in which items arise. For example, children acquire basic-level concepts before either superordinate or subordinate ones, an ordering that constrains theories of semantic cognition (McClelland & Rogers, 2003). In memory development, Courage and Howe (2004) proposed that autobiographical memory and the concept of “self” develop synchronously, thus placing a lower limit on early autobiographical recollections. In the development of early verbal behavior, Ejiri and Masataka (2001) proposed that rhythmic motor action and nonverbal vocalizations are important precursors to canonical babbling. In general, developmental ordering allows researchers to address fundamental questions about relations within the developing system, for instance, whether development of different aspects of the system stems from a common source or whether the development of one aspect of the system is contingent upon the development of another.

In the domain of language, developmental ordering relations have played a particularly important role in recent theorizing. Across the first years of life, children’s language abilities typically progress through several phases that appear to reflect an asynchronous sequence of events. Babbling and sound play characterize much of the progress within the first year, followed later by the onset of recognizable words around a child’s first birthday. It is typically only after a relatively extended period of single words or short simple phrases (e.g., “mommy sock”) that children’s productions reflect the grammar of their native language, increasing in overall length as well as the inclusion of closed-class forms (e.g., inflections like plural—s or past tense—ed). While there are well-documented individual differences in the timing of this sequence (e.g., Bates, Dale, & Thal, 1995), the transitions from sounds to words to grammar have traditionally been viewed as demarcating three distinct and developmentally autonomous phases (e.g., Pinker, 1999).

Others have instead proposed that the knowledge structures that comprise these fundamental language domains are constructed by the child within a unified developing system using a common set of domain-general learning mechanisms and computational or representational resources (Bates & Goodman, 1999; Elman, 2004; MacWhinney, 2001, 2004; Marchman, 1997; Zevin & Seidenberg, 2004). According to this view, the wide array of knowledge structures that comprise language are emergent phenomena; they do not reside in a preformed state

Portions of this work were supported by a Grant from the National Institutes of Health (HD 42235). The authors would like to thank the members of the MacArthur–Bates CDI Advisory Board for permission to use the norming data for the current project. We also thank Donna Thal, Jeff Elman, and two anonymous reviewers for helpful comments on a previous version of the manuscript.

Correspondence concerning this article should be addressed to James A. Dixon, Department of Psychology, 406 Babbridge Road, University of Connecticut, U-1020, Storrs, CT 06269-1020. Electronic mail may be sent to james.dixon@uconn.edu.

© 2007 by the Society for Research in Child Development, Inc. All rights reserved. 0009-3920/2007/7801-0011
within the organism, nor are they simply learned from the environment. Rather, they emerge through complex interactions between domain-general learning mechanisms and an intricately structured, multiply faceted world. As the system organizes in response to regularities at the different levels of language, various structures emerge at higher levels of organization. For example, Plaut and Kello (1999) showed that phonological representations may emerge from the confluence of semantic, acoustic, and articulatory factors. According to this view, phonological representations form through the repeated interactions with these other levels; such interactions can only occur if all the levels are cast in a single system. Similarly, it has been proposed that abstract grammatical abilities (e.g., productive rule-like use of inflectional morphemes, like “daddy goed”) emerge over the course of building a lexical system (e.g., Bates & Goodman, 1999). Here, grammatical rules and principles are derived via the tracking of frequency-weighted regularities that occur within and across lexical forms.

For these theories of acquisition, then, it becomes paramount to understand the developmental ordering relations within that system. These orderings can reveal contingencies in how development takes place. For example, developmental change within one level of language (e.g., the lexicon) may be contingent upon prior developmental change within another level (e.g., phonology). Thus, developmental orderings place important constraints on theory and provide the central phenomena that are addressed by computational models. That is, computational models of language acquisition predict the order in which language structures emerge based on the principles of acquisition that are embedded in the models. The order of the emergence of language structures is one of the model’s key benchmarks; the performance of the model is compared with what is known about the developmental sequences observed in children (Elman, 2001; Marchman, Plunkett, & Goodman, 1997; Munakata & McClelland, 2003; Plunkett & Marchman, 1993; Thomas & Karmiloff-Smith, 2003). Models are supported to the degree that they show the same developmental orderings that have been observed empirically.

A major goal of the current article is to test one such ordering hypothesis, specifically, that lexical development precedes the development of grammar. A considerable body of empirical evidence has been presented that supports this hypothesis (Bates & Goodman, 1999; Dromi, 1987). This developmental relation must be explained by theories of language acquisition and demonstrated by computational models of the process claiming to provide evidence for the psychological reality of the mechanisms they embody (Cohen & Chaput, 2002). For example, in the acquisition of the English past tense, children typically produce overregularization errors after the irregular verbs had been produced correctly, that is U-shaped development (e.g., Cazden, 1968, but see Marcus et al., 1992, for an alternative view). Plunkett and Marchman (1993) presented a computational model of the development of the English past tense in which the model’s behavior also followed a U-shaped course. While the model’s behavior did not mimic children in all regards, the claims made by Plunkett and Marchman (1993) regarding the computational mechanisms underlying the transition from rote production of lexical items to productive language use would have been greatly undermined if the model did not follow the same developmental ordering.

A second major goal of the article is to demonstrate and extend previous work on a general method for testing developmental ordering hypotheses. The developmental ordering of items (e.g., structures, abilities, skills, behaviors) is a key theoretical prediction about any developing system. However, given the types of measures researchers typically have available, testing these hypotheses has been notoriously difficult. The current article provides a detailed example of how a developmental ordering hypothesis can be tested using an approach presented by Dixon (2005), which we call the non-linear-mapping approach. It also demonstrates how this approach can be extended for developmental systems that contain multiple influences, such that converging evidence is brought to bear on the issue.

Developmental Ordering of the Lexicon and Grammar

A large body of work shows that development of the lexicon and grammar are strongly correlated, a cross-domain association that has been taken as evidence that learning in these domains is paced by the same computational or representational mechanisms. Nevertheless, the developmental asynchrony between lexical learning and grammar learning suggests that lexical development may initially outpace the development of grammar (see Bates & Goodman, 1999 for a review). One central piece of evidence for this latter hypothesis concerns the shape of the relation between measures of grammar and vocabulary size. When children’s scores on a measure of grammar are plotted as a function of their scores on a measure of vocabulary size, the relation appears strikingly curvilinear. Figure 1 shows an
example of this relation using normative data from the vocabulary checklist and grammar complexity sections of the MacArthur-Bates Communicative Development Inventories (CDI; Fenson et al., 1993, in press). As can be seen in the figure, lexical development appears to initially occur more rapidly than grammatical development (i.e., there is more change in the size of the lexicon early on than in grammar). Note that this pattern of data cannot be explained by the logically required fact that grammar must have some minimum set of words on which to be used. Taken alone, this fact only specifies the location of the intercept; some minimum number of words is necessary for grammatical development to begin. Beyond that minimum number of words, the relation between lexicon and grammar could be negligible. Thus, this fact does not imply that grammar and the lexicon will be strongly related over their respective developmental courses (and is completely mute on what shape that relationship might take).

Bates and Goodman (1999) noted that this nonlinear relationship between the lexicon and grammar has been demonstrated both cross-sectionally and longitudinally for English-speaking children (see also Bates et al., 1994, Dale, Dionne, Eley, & Plomin, 2000; Fenson et al., 1994; McGregor, Sheng, & Smith, 2005). Studies have also found similar relationships in Italian (Caselli, Casadio, & Bates, 1999), Hebrew (Maital, Dromi, Sagi, & Bornstein, 2000), Icelandic (Thordardottir, Weismer, & Evans, 2002), and Spanish (Jackson-Maldonado et al. 2003). Bates and Goodman (1999) examined several possible methodological artifacts that could account for these effects, and demonstrated that the relationship held even if words that are related to grammatical complexity (i.e., grammatical function words such as prepositions and conjunctions) were omitted from the vocabulary count. When the growth curves of individuals were examined, their patterns were similar to the average patterns. Furthermore, a recent study with children learning both English and Spanish at the same time (Marchman, Martinez-Sussmann, & Dale, 2004) indicated that similarly strong relations are seen within each language (i.e., English lexicon to grammar; Spanish lexicon to grammar), even though across language relations (i.e., English lexicon to Spanish lexicon; English grammar to Spanish grammar) were weak. These same relations were found using measures of vocabulary and grammar based on naturalistic language performance, not just those relying on checklist reports from parents and/or other caregivers.

A considerable amount of theorizing has been carried out to explain this developmental ordering relationship. Perhaps the most prominent idea is that grammar emerges from the mechanisms involved in acquiring the lexicon itself. Under this hypothesis, as the lexicon increases in size, grammar becomes organized into increasingly complex forms. Put another way, grammatical forms are emergent properties of the unified language system and are most strongly tied to changes in the lexicon within that system. As the size of the lexicon grows, the quality of the emergent grammar changes. In this way, a “critical mass” of vocabulary is required for the development of a particular level of grammatical complexity (Marchman & Bates, 1994; Plunkett & Marchman, 1993). In another version of this view, learning words involves learning their grammatical as well as lexical–semantic properties, including requirements regarding the construction in which those words can legally appear and which inflectional morphemes are required. Others have pointed out that early word combinations tend to be highly routinized and situation-specific, suggesting that learning grammar, like learning words, may be driven by processes that are item-specific and frequency dependent. It is only over the course of learning an increasingly sophisticated and interrelated lexicon that grammatical structures become encoded in terms of their abstract syntactic form (e.g., Akhtar, 1999; Braine, 1976; Lieven, Pine, & Baldwin, 1997; Tomasello, 2003). All of these proposals presume strong associations between the learning principles involved in lexical and grammatical acquisition, and are compatible with an increased integration of lexical–semantics and grammar in theories of linguistics (e.g., Bresnan, 1982; Goldberg, 1999). Should they bear out empir-
ically, each places important constraints on the architecture of the language system and causal relations within that system.

Reconsidering Developmental Ordering and Functional Form

The shape of the relationship between two measures provides a potentially powerful indicator of the type of developmental relation between the underlying variables. However, the interpretation of that shape or functional form depends crucially on the mapping between the underlying variables and their respective measures. Consider, for example, the developmental relationship between the two items, A and B, shown in the upper panel of Figure 2. The line in the center of the figure, labeled “Development,” shows that development proceeds from the bottom of the panel to the top. Each item is represented by a bar that increases in saturation as the item develops. The period during which the developmental changes of interest occur for each item is marked by a rectangle enclosing the bar. In the current example, items A and B develop synchronously; they start and complete development together and develop at the same rate.

Each underlying item is assessed by a measure, shown next to it (and labeled arbitrarily with numbers from 0 to 100). The lines connecting the underlying variables to the measures illustrate the mapping between them. For both items, the measured values increase as the underlying level of the variable increases. Furthermore, the measure of item A captures developmental change in A equally well across its developmental span. That is, the measure is equally responsive to early and later changes in A (i.e., forms an interval-level scale; Stevens, 1951). However, the situation for item B is different. The measure of item B is more responsive to later changes than it is to earlier ones. The mapping between the underlying variable B and the measure of B is ordinal, but nonlinear.

The dashed lines show hypothetical individuals at different points in the developmental period. For each measure, an individual’s score can be found by following the dashed line to the underlying variable and then the solid line to the measure. The lower panel of Figure 2 shows the observed relationship between the measures of A and B. Note that, despite the fact that the underlying relationship is synchrony, the observed relationship shows a strong A before B, curvilinear pattern. The nonlinear mapping between the underlying variable B and its measure creates this illusory pattern. Although a nonlinear mapping, such as that exhibited by the measure of item B, may initially seem unlikely or even contrived, note that this situation only requires that the measure be more responsive at some levels of development than others. It seems likely that this is a very common, perhaps even modal, situation in developmental research. Certainly, all researchers hope that their measures capture developmental change across the span, but it is probably overly optimistic to expect that a measure captures the extent of developmental change equally well across the developmental range.

This analysis leads to the unsettling, general conclusion that the observed functional form seen in the data may not accurately reflect the actual form of the underlying relationship between the variables of interest. But, is it possible to tell the difference between relations that reflect a true developmental ordering from those that derive from measurement nonlinearities? Some recent work on this issue has indicated that it is. For example, Dixon (1998) showed how some developmental ordering hypotheses were eliminated by particular data patterns. In more recent work, Dixon (2005) showed that nonlinear mappings, such as the one discussed above, must create specific relationships between error and the predicted value of the measure, given standard ordinary least-squares (OLS) regression assumptions (Cohen, Cohen, West, & Aiken, 2003). Furthermore, the evaluation of these error patterns provides the tools to test directly whether a nonlinear mapping between the measure and the underlying variable is masking the true underlying relation among the constructs. That is, by examining the pattern of residuals (i.e., errors) across the predicted values of the measure used as the dependent variable, it is possible to determine whether an observed ordering relation between two variables (i.e., A before B) is actually an artifact of a nonlinear mapping between measure and construct. Because this approach features heavily in the analyses that follow, we first provide an overview of the logic and then present the approach in some detail using the synchrony relationship as an example.

An overview of the nonlinear mapping approach. The logic of the nonlinear mapping approach flows from a fundamental and widespread assumption about the nature of developmental systems in combination with some easily demonstrable consequences of nonlinear functions. We assume that development results from the interactions of a complex system and that, therefore, even the strongest relationships are not deterministic. One seemingly unremarkable implication of this assumption is that there is always
Figure 2. The synchrony relationship between items $A$ and $B$ is shown in the top panel. The measure of each item, with values that range from 0 to 100 (an arbitrarily chosen scale), is shown beside it. The lines connecting each underlying variable to its respective measure give a sense of the mapping between them. For example, the measure of item $A$ is linearly related to the underlying level of $A$. The measure of item $B$ is nonlinearly related to the underlying variable $B$; earlier developmental changes in underlying item $B$ show smaller increases in the measure relative to later ones. The lower panel shows the idealized data pattern that results from the situation in the top panel.
some unexplained variance in the relation between any two (or more) items. Note that this is an assumption about the underlying reality of developmental systems—it is presumed to be true of developmental relations regardless of whether anyone is measuring them. The nonlinear mapping approach capitalizes on this bit of unexplained variance to help diagnose the relation or mapping between underlying variables and our measures of them. If the mapping is nonlinear, then the unexplained variance must become systematically related to the measured variable. This follows from the fact that nonlinear functions create systematic relations among previously unrelated terms (the mapping from the underlying variable to its measure acts like a function). Therefore, by examining the relation between error and the measured variable, one can test whether an observed developmental relation was created through a nonlinear mapping. In the next section, we present the approach in more detail using the synchrony relation as an example. We use the familiar general linear model to represent developmental relations, thus facilitating the translation from hypothesized relations to well-known analytic methods, such as OLS regression.

Representing synchrony as a simple equation. The synchrony relationship can be represented as a simple equation: \( B_u = \alpha + \beta A_u \), where the intercept is denoted as \( \alpha \) and the slope as \( \beta \). The equation specifies that changes in \( B_u \) are linearly related to changes in \( A_u \) and, therefore, is another way of stating that the underlying variables \( A \) and \( B \) develop synchronously. Unexplained variance or error, in the form of individual-level differences and unmeasured influences, enters the model at this level: \( B_u = \alpha + \beta A_u + e_u \). In standard fixed-effect approaches, such as OLS regression, we assume that error, \( e_u \), is unrelated to the levels of the predictors. We also assume that \( e_u \) is normally distributed, has a mean of 0, and an unknown variance.

This simple model describes the relationship between \( A \) and \( B \) at the underlying level, a level we cannot observe directly—measures are all we have in hand. Assume that, consistent with Figure 2, our measure of \( A \), \( A_m \), is a linear function of \( A_u \). Similarly, assume that our measure of \( B \), \( B_m \), is a nonlinear function of \( B_u \) in the manner depicted in Figure 2. The relationship between the underlying variable \( B \) and its measure can be represented as a nonlinear function such as being raised to a positive power greater than 1: \( B_m = (B_u)^{\beta} \). Substituting the equation presented above for \( B_u \), we find that \( B_m = (\alpha + \beta A_u + e_u)^{\beta} + e_m \). The additional error term, \( e_m \), represents error that occurs at the measurement level (e.g., lapses in the participant’s attention, misunderstanding of a question, etc.).

The nonlinear mapping function above creates the curvilinear relationship between \( A_u \) and \( B_u \) from the linear relationship between \( A \) and \( B \). However, it also leaves a signature pattern in the data. Because the sum of \( A_u \) and underlying error, \( e_u \), is nonlinearly transformed, these terms that were previously independent now become related. This mathematical fact is easily demonstrated in the current context by expanding the equation: \( (\alpha + \beta A_u + e_u)^{\beta} \), which yields: \( \alpha^\beta + \beta \alpha^{\beta-1} \beta A_u + \beta \alpha^{\beta-2} \beta^2 A_u^2 + \beta \alpha^{\beta-3} \beta^3 A_u^3 + \cdots + \beta \alpha e_u + \beta e_u A_u + \beta e_u^2 \). For our current purposes, the interesting element is the last, multiplicative term, \( 2\beta A_u e_u \), because it shows that the nonlinear mapping function creates a relation between \( A_u \) and \( e_u \). (The remaining terms can be ignored for the moment.) Multiplicative terms, of course, imply that the effect of one variable will depend on the value of the other. All nonlinear mappings create relationships among the underlying terms; the powers are a convenient way to represent nonlinear mappings, but not essential to the approach. The major point here is that if there is a nonlinear mapping from the underlying variable to the observed measure, systematic relationships between the underlying variable and underlying error will be created.

Of course, the precise form of this relationship will depend on the particular type of nonlinear mapping. For example, variable–error relations will look different if there is a nonlinear mapping from variable to measure in only one of the variables versus if variable–measure nonlinearities exist in both variables. Therefore, given an observed developmental ordering (e.g., \( A \) before \( B \)), it is possible to test whether this ordering is an artifact of a nonlinear mapping between the measure and the underlying construct, and hence, does not actually reflect the true underlying developmental relationship.

An Alternative Explanation for the Lexicon–Grammar Relationship

In the current context, we consider two possible situations in which we would see a curvilinear relationship between the lexicon and grammar. First, it is possible that the underlying developmental relationship is indeed as it appears at the level of the measures; lexical development precedes grammatical development. In this case, both measures of lexicon and grammar would capture developmental change equally well across the period, that is, there would be no nonlinearities in measurement that would change the way the underlying
developmental ordering appears in the data. A second possibility is that the underlying relationship between lexicon and grammar is actually synchrony, but a nonlinear mapping between one or more of the underlying variables to its respective measure is driving the curvilinear relationship, masking the true situation that lexicon and grammar are developing synchronously. (For purposes of exposition, we postpone discussion of a third logical possibility, that grammatical development precedes lexical development, until later in the article.)

In panel (i) at the top of Figure 3, we plot the observed priority relation between lexicon and grammar. In the bottom panel (u), we plot the underlying relation between lexicon and grammar as synchrony. Now, consider how this synchrony relationship might give rise to the observed curvilinear relationship in more detail. There are only three types of nonlinear mapping situations that can create
the curvilinear pattern in panel (i) from an underlying synchrony relationship (panel (u)). The second row of panels in Figure 3 illustrates these three situations graphically. In the left-most panel (ii), the mapping between the lexicon and its measure is nonlinear such that early development of the lexicon gets more change in the measure relative to later development. The mapping between grammar and its measure, however, is linear. In the center panel (iii), the mapping between the lexicon and its measure is linear, but the mapping between grammar and its measure is nonlinear. The grammar measure is more responsive to later developmental changes than it is to earlier ones. The right-most panel (iv) shows both variables with measures that are nonlinear in the ways just described. Thus, the only ways to obtain the observed nonlinear relationship when the lexicon and grammar are actually developing in synchrony are (1) for the measure of the lexicon to be a nonlinear, decelerating function of the underlying lexicon, or (2) for the measure of grammar to be a nonlinear, accelerating function of underlying grammar, or (3) for both those nonlinear mappings to occur. Dixon (2005) demonstrated that these types of nonlinear mappings between measures and their underlying constructs make predictions about the patterns that will be observed at the level of the residuals.

Nonlinear mappings predict specific residual patterns. We illustrate the predictions for each type of nonlinear mapping using a simple simulation, adapted from Dixon (2005). The simulation allows us to show how error that is unrelated to the underlying variables (i.e., homoscedastic) becomes systematically related to the measured variables (i.e., heteroscedastic). Furthermore, the simulation concretely demonstrates that each type of nonlinear mapping shown in Figure 3 creates a specific relationship between error and predicted values that can be both visually inspected and tested statistically. We present the general forms of these predicted residual patterns graphically.

First, we show the hypothesized underlying synchrony relationship and residuals. This under-the-hood view of the developmental relations is, of course, never available to researchers. We present it here to show the relation graphically and to provide a point of comparison for examining the residual plots below. The simulated data consist of two underlying variables, which we label here \( L \) (lexicon) and \( G \) (grammar), related via the following equation:

\[
G_u = \alpha + \beta L_u + \epsilon_u.
\]

The values of \( \alpha \) and \( \beta \) were set at 5 and 1.5, respectively, and \( \epsilon_u \) had a mean of 0 and a variance of 20. The upper panel of Figure 4 plots the simulated relationship between lexicon and grammar when they are developing in synchrony. The middle and lower panels of Figure 4 are standard ways of presenting residuals. Deviations from the

Figure 4. The top panel shows the underlying relationship between lexicon and grammar specified in the simulation. The middle panel shows the residuals when grammar is predicted as a function of lexicon. The lower panel shows the residuals when lexicon is predicted as a function of grammar.
model predictions (the zero point on the y axis) are shown as a function of the predicted values. Note that the extent of scatter around the best-fitting line is constant in both plots, indicating that at the underlying level error is unrelated to the predicted values of grammar and lexicon, respectively (i.e., homoscedastic). However, if a nonlinear mapping obtains between the measure and the underlying construct, the appearance of these residual patterns will change dramatically in ways that reflect the particular type of nonlinearity imposed by the measure.

To illustrate, Figure 5 presents the patterns of residuals that are obtained when one or more of the constructs is nonlinearly related to its measure. In all cases, the observed relation between lexicon and grammar is priority, as plotted in the top-most panel (panel (i)). The left-most portion of the figure presents the situation in which the measure of the lexicon, \( L_m \), was a decelerating function of \( L_u \), illustrated graphically in panel (ii). To generate the residual plots, we created \( L_m \) as: \( L_m = (L_u)^2 + \epsilon_{m} \cdot G_m \) was a linear function of \( G_u \); \( G_m = G_u + \epsilon_{mg} \). The error terms were drawn from a normal distribution with a mean of 0 and a variance of 20. Panel (viii) presents the pattern of residuals when \( G_m \) and \( G_m^2 \) were used to predict \( L_m \). In this case, the model fit was very good, but the pattern of residuals was strongly heteroscedastic (i.e., the distribution is not uniform across the values of the dependent measure). In this case, the absolute values of the residuals were negatively related to the magnitude of the predicted values of \( L_m \) and the degree of scatter changes across the predicted values as well. Panel (v) plots residuals when \( L_m \) and \( L_m^2 \) were used to predict \( G_m \). Again, the pattern of residuals was heteroscedastic, although the relationship between the absolute value of the residuals and the predicted values was positive and somewhat weaker. These two plots (panels (viii) and (v)) illustrate that if the underlying relationship between the lexicon and grammar were synchronous but a nonlinear mapping between the underlying lexicon, \( L_u \), and its measure, \( L_m \), were driving the observed data pattern, the residuals would be related to the predicted values. This relationship is the strongest for predicted values of lexicon, because the measure of lexicon here is nonlinear. The nonlinear mapping has left a signature pattern in the residuals; a specific set of relationships between underlying error and the predictor variables has been induced by the nonlinearity of the measure.

Next consider what would happen if our measure of grammar, \( G_m \), was an accelerating function of underlying grammar, \( G_u \), illustrated graphically in panel (iii) of Figure 5. We created \( G_m \) such that \( G_m = (G_u)^2 + \epsilon_{mg} \cdot L_m \) was created as a linear function of \( L_u \), and the error terms were created as described above. When \( G_m \) and \( G_m^2 \) were used to predict \( L_m \), the fit was again very good, but the pattern of residuals was quite different from that in the previous case. As can be seen in panel (ix) of Figure 5, the pattern of residuals was curvilinear with two bends (i.e., cubic), but does not show the negative relationship presented in panel (viii). When \( L_m \) and \( L_m^2 \) were used to predict \( G_m \), the pattern of residuals was strongly heteroscedastic, as can be seen in panel (vi). Here, the residuals were positively related to the predicted value of \( G_m \).

Finally, consider what would occur if both measures, \( L_m \) and \( G_m \), were nonlinear functions as just described (i.e., \( L_m \) was an accelerating function of \( L_u \), \( G_m \) was a decelerating function of \( G_u \)), depicted graphically in panel (iv) of Figure 5. When \( G_m \) and \( G_m^2 \) were used to predict \( L_m \), panel (x) shows that the pattern of residuals takes on a curvilinear shape and that the residuals were negatively related to the predicted values of \( L_m \). Finally, when \( L_m \) and \( L_m^2 \) were used to predict \( G_m \), the residuals were again positively related to the predicted values of \( G_m \) (panel vii). Thus, each of the three nonlinear mapping situations that are capable of creating the observed priority relationship from underlying synchrony can be identified by a specific signature pattern in the plots of the residuals.

Although this might initially appear complex, the implications can be summarized as follows: first, if a curvilinear relationship is observed at the level of the measures (e.g., as has been shown for lexicon and grammar) but the true relation between lexicon and grammar is synchrony, it must be the case that there is a nonlinear mapping between one or more of the measures and their underlying constructs. Specifically, (a) the measure of the item that appears to be developing more rapidly, the lexicon in this case, may be a decelerating function of the underlying variable, and/or (b) the measure of the item that appears to be developing more slowly, grammar in this case, may be an accelerating function of the underlying variable.

Second, each type of nonlinear mapping will necessarily induce a systematic relation between underlying error and the predictor(s), a relation that is observable in the residual patterns. Measures that are decelerating functions of the underlying variable create negative relationships between error and the predicted value of the measure. Measures that are accelerating functions create positive relationships between error and the predicted values.
Figure 5. The top panel (i) shows an idealized version of the observed relationship between lexicon and grammar. The three different types of mapping situations that can produce this pattern are shown using the synchrony relationship as an example (panels (ii) – (iv)). The bottom two rows show residual patterns based on modeling the underlying relationship as synchrony and the nonlinear mappings presented in each column. To emphasize that these are predicted patterns from the simulation, rather than results from the actual measures, we have retained the arbitrary scale values (0 – 100). The patterns of residuals when lexicon and lexicon$^2$ are used to predict grammar are shown in the third row for each of the three nonlinear mapping situations (represented in the columns). The patterns of residuals when grammar and grammar$^2$ are used to predict lexicon are shown in the bottom row for each nonlinear mapping situation.
Third, as we show below, one can use available statistical methods to assess whether these predicted relationships (between error and predicted values) are obtained in the data, providing useful information regarding the validity of claims on developmental ordering. In this case, we illustrate this technique by examining whether the curvilinear relation between lexicon and grammar reflects an underlying situation of priority or is actually synchrony. We used this analytic approach using recent data from the MacArthur–Bates CDI. We present the results of this analysis, after reviewing the CDI measures and sample. Finally, we show how the approach can be extended to capitalize on the multiply determined nature of most developmental systems.

Method

Participants

Parental report data were compiled from typically developing children sampled cross-sectionally from 16 to 30 months (N = 1,461; 728 females) in the updated norming sample of the MacArthur–Bates CDI: Words & Sentences (CDI: W&S; Fenson et al., 2007). A minimum of 37 girls and 37 boys were represented at each age (range: \( n = 37 - 60 \)). All children were learning English as a first language, and none of the children were reported to have experienced serious birth complications, diagnosed developmental disabilities, hearing loss, or other medical problems. The sample consists of participants from Seattle, WA, San Diego, CA, Madison, WI, Dallas, TX, New Haven, CT, New Orleans, LA, and Providence, RI. The ethnicity breakdown of the sample reflects a dominance of Caucasian (White) participants, with African American, Asians, and Hispanic participants each representing somewhat smaller proportions of the sample. The distribution is similar to that in the United States at large, except for a lower percentage of Hispanics due to the requirement that children are learning English as a native language.

Measures and Procedure

As described in Fenson et al. (in press), parents completed the CDI: W&S, as well as a Basic Information Form, which provided basic demographic and health information.

Size of vocabulary is assessed using the vocabulary checklist from the CDI: W&S. This section asks reporters (typically mothers) to indicate which words (out of 680 possible) their child “understands and says.” These checklists capture the reporters’ general knowledge about whether or not the child “knows” a particular word, rather than specific information about the contexts in which a word is used or the accuracy of its pronunciation. The items are listed alphabetically in categories (e.g., animal sounds, vehicles). Vocabulary production scores are the total number of words that are reported.

Progress in grammar is assessed using the Grammatical Complexity section of the CDI: W&S. In this section, reporters read a series of 37 pairs of phrases, and are asked to indicate which one of a pair of phrases “sounds most like the way your child is talking right now.” The first is an example that lacks grammatical markers or is syntactically simple, whereas the second provides a more complex alternative (e.g., “Kitty sleep” vs. “Kitty sleeping”). The grammatical complexity score is the number of times the second sentence was selected (37 maximum).

Results

Table 1 shows the means and standard deviations for the reported vocabulary and grammar measure grouped by age in months. Both vocabulary production and grammatical complexity increased with age, \( r_s(1460) = .64, .60 \), respectively. An \( \alpha \) level of 0.05

<table>
<thead>
<tr>
<th>Age (months)</th>
<th>Number of words</th>
<th>Grammatical complexity</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>( M )</td>
<td>( SD )</td>
</tr>
<tr>
<td>16</td>
<td>58.86</td>
<td>59.08</td>
</tr>
<tr>
<td>17</td>
<td>68.24</td>
<td>83.95</td>
</tr>
<tr>
<td>18</td>
<td>107.07</td>
<td>110.54</td>
</tr>
<tr>
<td>19</td>
<td>167.04</td>
<td>138.94</td>
</tr>
<tr>
<td>20</td>
<td>181.27</td>
<td>134.47</td>
</tr>
<tr>
<td>21</td>
<td>208.15</td>
<td>157.12</td>
</tr>
<tr>
<td>22</td>
<td>256.61</td>
<td>166.92</td>
</tr>
<tr>
<td>23</td>
<td>313.44</td>
<td>162.37</td>
</tr>
<tr>
<td>24</td>
<td>307.29</td>
<td>171.03</td>
</tr>
<tr>
<td>25</td>
<td>338.18</td>
<td>178.11</td>
</tr>
<tr>
<td>26</td>
<td>382.52</td>
<td>180.70</td>
</tr>
<tr>
<td>27</td>
<td>408.20</td>
<td>176.24</td>
</tr>
<tr>
<td>28</td>
<td>414.14</td>
<td>158.40</td>
</tr>
<tr>
<td>29</td>
<td>433.06</td>
<td>174.74</td>
</tr>
<tr>
<td>30</td>
<td>518.64</td>
<td>125.23</td>
</tr>
</tbody>
</table>
was used for all significance tests, unless otherwise indicated.

**Predicting Grammar From the Lexicon**

Consistent with the literature reviewed above, the relationship between vocabulary production and grammatical complexity was strong and curvilinear. When vocabulary production was used to predict grammar, the model explained 68% of the variance, $F(1, 1459) = 3,080.43, B = 0.05, t(1459) = 55.50$. (To remove nonessential multicollinearity, all predictors were grand-mean centered before computing the interaction and quadratic terms; Cohen et al., 2003.) Adding the quadratic term to the model (i.e., vocabulary squared) significantly increased the proportion of explained variance to 74%. The parameter estimates for the linear term, $B = 0.04$, and the quadratic term, $B = 0.00008$, were both significantly different from 0, $ts(1458) = 51.78$, 18.07, respectively. The model predictions are shown in the upper panel of Figure 1.

Following the logic outlined above, we examined the pattern of residuals across the predicted values of grammar. As can be seen in the middle panel of Figure 6, the pattern of residuals was systematically heteroscedastic. A simple initial approach to evaluating this type of pattern is to examine the absolute values of the residuals as a function of the predicted values of grammar, $r(1460) = .53$. This relationship is probably underestimated, because the scale is clearly bounded in the upper and lower regions. However, the pattern is very similar to that in panel (vi) of Figure 5.

A more sophisticated approach allows us to capture heteroscedastic patterns, such as this one, more completely. Because residual patterns can reflect both the variance of the errors and their mean values changing as a function of the predicted values, alternative methods are necessary to test these relationships appropriately. Therefore, we adopted a very flexible approach developed by Breusch and Pagan (1979) and Cook and Weisberg (1983) (see also Fox, 1991). This approach allows one to model residual patterns with standard regression techniques using a simple transformation of the residuals: The residual score for each individual is squared and standardized (i.e., divided by a term that closely approximates the sample variance). The regression procedure yields a test statistic that is asymptotically distributed as chi-square. (We present the details of computing these standardized scores, along with more information about obtaining the test statistic, in Appendix A.) When the predicted values of grammar were used to predict the standardized residuals, the positive relationship seen in the raw residuals was confirmed, $\chi^2(1) = 568.69$.

**Figure 6.** The top panel shows the relationship between the number of words produced and grammatical complexity. The middle panel shows the residuals as a function of the predicted values of grammatical complexity when both number of words and its square are used as predictors. The lower panels show the residuals as a function of predicted values of lexicon when both grammatical complexity and its square are used as predictors.

**Predicting Lexicon From Grammar**

When grammar and the quadratic term (i.e., grammar squared) were used to predict vocabulary, the model results were similar to those described above; 73% of the variance was accounted for, $F(1, 1458) = 1,927.97$. Both parameter estimates were significantly different from 0, $Bs = 19.72, -.40$, for the linear and quadratic terms, respectively,
ts(1458) = 48.17, −15.81. The model predictions were nearly identical to those in the upper panel of Figure 1, so we do not plot them again.

The pattern of residuals is shown in the lower panel of Figure 6. As can be seen in the figure, the linear relationship between the absolute values of the residuals and the predicted values was very weak, r(1460) = −.073. Rather, the relationship appears to be strongly curvilinear. To describe the curvilinear nature of the relationship, we again used the familiar power polynomials. We used the predicted vocabulary values, and the square and cube of the predicted values, as predictors of the (signed) residual values. The parameter estimates for the linear, quadratic, and cubic terms were all significant, Bs = −0.287, −0.002, 0.00001, ts(1124) = −5.69, −9.97, 8.16. The pattern is curvilinear in much the same way as that in panel (ix) of Figure 5; it is cubic and there is very little linear relationship between the absolute values of the residuals and the predicted values of vocabulary.

This relationship can be tested more formally using the standardized squared residual scores described above. Squaring the raw residuals makes this cubic relationship become quadratic; the negative deviations are now positive. To test for this quadratic trend, we first entered the predicted values of vocabulary as a predictor of the squared residuals. Consistent with the descriptive analysis presented above, there was a weak, negative relationship between predicted vocabulary and the squared residuals, $\chi^2(1) = 4.65$. Adding the square of the predicted vocabulary scores dramatically increased the fit of the model, change in $\chi^2(1) = 34.08$. Also consistent with the descriptive analysis, there was a strong curvilinear relationship between the residuals and the predicted values.

In summary, these analyses of relations between lexicon and grammar in the CDI normative data reveal that the residuals are systematically related to the predicted values. The observed residual patterns are consistent with the possibility that lexicon and grammar develop synchronously, but that the measure of grammar is nonlinear such that early developmental changes are underrepresented relative to later ones. Put another way, the grammatical complexity measure appears to be systematically more responsive to the later developments in grammar. Therefore, the evidence considered thus far suggests that the lexicon and grammar may develop synchronously. These analyses indicate that the oft-cited curvilinear relationship between vocabulary size and grammar is actually a function of a nonlinear mapping between underlying grammar and its measure, and not a reflection of the actual underlying relationship at the level of the constructs. Next, we show how the nonlinear mapping approach can be extended to provide converging evidence on developmental ordering hypotheses.

### Extending the Approach to Multiply Determined Systems

Most developmental theories propose that many variables impact the developing system and, therefore, contribute to the development of the item under investigation. For example, the development of children’s theory of mind may depend on executive function, sibling relations, and the ability to relate multiple dimensions, as well as a number of other factors (e.g., Astington & Jenkins, 1999; Wellman, Cross, & Watson, 2001). In the current context, grammar has been hypothesized to emerge from the unified, developing language system. Lexical development may be playing a central role, but other codeveloping factors, such as working memory capacity, pragmatics, quality of social interactions, and so forth, would also be important. A common strategy in developmental work is to assume that these factors are largely, if imperfectly, age-related, and thus age can be used as a proxy for these additional codeveloping influences. For example, showing that one’s theoretically important predictors explain variance in the dependent measure, above and beyond any variance explained by age, is standard (and reasonable) practice (see e.g., Kail, 2000; Marchman et al., 2004).

Here, we capitalize on the ability of age to stand in for other developmental influences, but within a somewhat different strategy. Recall that we showed above that error becomes associated with the variables at the underlying level as a result of the nonlinear mapping. For example, if $G_u = \alpha + \beta L_u + \varepsilon_{wr}$ and $G_m$ are related to $G_u$ through a nonlinear mapping: $G_m = (G_u)^2 + \varepsilon_{may}$, then the error term, $\varepsilon_{wr}$, will become related to the $L_u$ and ultimately to the predicted values of $G_m$. The underlying error term captures unexplained differences in grammar, including influences from factors not currently in the equation. If we have measures of these other factors, we can, of course, remove them from the error term by adding them to the equation. In the current context, we take age, $A_u$, as a proxy for unmeasured developmental influences that affect the child’s underlying level of grammar; therefore, $\varepsilon_u = (A_u + \varepsilon_{wr})$, and $G_u = \alpha + \beta L_u + \beta_a A_u + \varepsilon_{wr}$. All that has changed here, of course, is that we have unpacked the error term by including age effects in the model. This small change allows us to extend the approach
in a significant way. Just as error became associated with \( L_u \) as a result of the nonlinear mapping, \( A_u \) will become related to \( L_u \): the nonlinear mapping creates the equivalent of a simple linear interaction term, \( A_u \times L_u \) (This can be confirmed easily by expanding the squared equation.) A nonlinear mapping between the underlying level of grammar and its measure, therefore, predicts that age and vocabulary will interact in predicting grammar. When the nonlinear mapping is an accelerating function, as in the example above, the interaction term will be positive. When the nonlinear mapping is a decelerating function, the interaction term will be negative.

To test this prediction in the current dataset, we added age and its linear interaction with vocabulary (i.e., Age \( \times \) Vocabulary) as predictors to the model for grammar (which already included both the linear and quadratic vocabulary terms). The results indicated that the interaction term was positive and contributed significantly to the model, \( B = 0.0032, t(1456) = 5.28 \); i.e., the effect of vocabulary increased with age. This effect is a straightforward consequence of a nonlinear construct-to-measure mapping. Terms that combine before the nonlinear mapping become positively related if the nonlinear mapping is an accelerating function.

It is also worth noting that the pattern of residuals remained systematically heteroscedastic across the predicted values of grammar when age and its interaction with vocabulary were included in the model. The correlation between the absolute value of the residuals and the predicted values was .40; the relationship between standardized residuals and the predicted values of grammar remained very strong, \( \chi^2(1) = 222.81 \). If a nonlinear mapping has occurred, error should remain associated with the predicted values, even when the set of true predictors are used in the model. While age is only a rough proxy for the set of true codeveloping variables, the fact that the heteroscedasticity is not ameliorated by its inclusion in the model further supports the interpretation that there is a nonlinear mapping between measure and underlying construct. Both the interaction between age and vocabulary, and the fact that including this term in the model did not eliminate the heteroscedasticity in the residuals are consistent with a situation in which there is a nonlinear mapping between the measured and actual levels of grammar.

_synchrony versus opposite priority_

Thus far, we have only considered synchrony as a potential alternative hypothesis to priority, but another form of developmental ordering might also give rise to the observed curvilinear relationship between lexicon and grammar. Specifically, grammar may initially develop more rapidly than the lexicon. This alternative, which Dixon (2005) referred to as “opposite partial priority,” posits that at the underlying level, the relationship between grammar and the lexicon is still curvilinear but opening downward, rather than upward, as seen at the observed level (i.e., concave vs. convex). In order for this opposite underlying relationship to give rise to the observed convex pattern, one or both of the measures must be dramatically nonlinear. Dixon (2005) showed that the same types of nonlinear mappings that produce the curvilinear relationship in a situation where the two underlying constructs were developing in synchrony would be necessary to produce a curvilinear relationship in the case of opposite priority. However, the nonlinearity would have to be even more extreme. For example, to go from even a fairly weak concave underlying relationship to a convex relationship, such as that observed between lexicon and grammar, would require a nonlinear mapping more extreme than the square.

As the nonlinearity between the underlying and measured levels increases (i.e., goes up the ladder of powers; Mosteller & Tukey, 1977), the relationships among the underlying terms change. Recall that when the nonlinear mapping approximates the square, it creates a product term involving lexicon and age, a simple linear interaction or moderator effect. However, if the nonlinear mapping approximates the cube, for example, then the interaction becomes quadratic. Moderators like other predictors may have simple linear effects or curvilinear effects (Baron & Kenny, 1986). The quadratic interaction prediction follows directly from the expansion of \((\alpha + \theta_1 \times L_u + \theta_2 \times A_u + e_u)^2 + e_{m-g} = G_{m-u} \). These equalities simply state that the measured values of grammar, \( G_{m-u} \), are related to the underlying values of grammar, \( G_u \), through a nonlinear mapping (i.e., the cube), and that lexicon, \( L_u \), and age, \( A_u \), are related to underlying grammar. (Because the expansion of the equation above is cumbersome, we present it in Appendix A, along with a description of a simple simulation of these relationships.) Testing for the presence of a quadratic interaction is straightforward; one creates a quadratic interaction term, which is then added to the model. In this case, the term is the product of lexicon and age squared: \( L_u A^2 \).

In summary, an opposite priority relationship at the underlying level could create the observed pattern, but only if a fairly extreme nonlinear mapping was at work. Such nonlinear mappings must create a
set of specific relationships among the underlying variables; a nonlinear mapping more extreme than the square will create a curvilinear, accelerating interaction. Therefore, we can test whether this type of nonlinear mapping is driving the observed relationship by testing for a positive, quadratic interaction.

To address this hypothesis we added $A_2^2$ and the quadratic interaction term to the model. ($A_2^2$ must be added so that the interaction term does not pick up on the quadratic effect of $A_m$ itself.) As noted above, if the observed relationship between lexicon and grammar was driven by a nonlinear mapping that is more extreme than the square, a significant, positive quadratic interaction is predicted. However, the parameter estimate for the quadratic interaction term was nonsignificant, $B = 0.00008$, $t(1455) = 0.16$. The power of this test is estimated at .97. Therefore, the hypothesis that grammatical development actually precedes lexical development is not supported.

Power and Sample Size Issues

The current data set is very large by most standards, which might raise concerns about the feasibility of this approach with the smaller data sets that researchers typically possess. Therefore, we conducted a power analysis (Cohen, 1988) to estimate the sample sizes necessary to obtain reasonable power (.80) for the more central tests of the nonlinear mapping explanation within the current data set.

First, consider the relationship between the standardized residuals and the predicted values of grammar. Given the magnitude of the observed effect, an $N$ of 50 would be necessary to achieve a power of .80, given an $\alpha$ of 0.05. This is a modest sample size by most standards. Next, consider the curvilinear relationship between the standardized residuals and the predicted values of vocabulary. An $N$ of 270 would be needed to obtain a power of .80, with $\alpha$ set at 0.05. This is a fairly large sample size, but then detecting relationships represented by higher order polynomials often requires larger samples. Finally, consider the effect of adding age and its interaction with vocabulary to the model for grammar, $f^2 = .087$, between a small and medium effect, according to Cohen. With $\alpha$ set at 0.05, an $N$ of 115 is needed to obtain a power of .80, a sample size that is within reach for many types of investigations.

Of course, the effect sizes observed in the current study may or may not be similar to those in other research domains. The main point here is that had this study been conducted with far fewer participants (i.e., <20% of the current sample size), it would still have had an excellent chance of reaching the same conclusions. The approach does not require the very large numbers necessary for norming an assessment instrument; rather it can be used with sample sizes more usually obtained in developmental studies.

Discussion

Previous work suggested that, across the first few years of life, the strong developmental relationship between the lexicon and grammar was curvilinear. This relationship was, at the very least, a phenomenon that required explanation. Some of these explanations have significant theoretical implications. For example, the curvilinear relationship led researchers to hypothesize that lexical development might drive grammatical development. According to one version of this hypothesis, the growing lexicon provides the foundation for grammar learning, and contributes fundamentally to the organization of increasingly complex grammatical forms. Specifically, the developmental precedence of lexicon over grammar was consistent with proposals in which grammatical principles emerge in a system that has built up a sufficient lexical base to support the further abstraction of grammatical regularities, that is, a critical mass (Marchman & Bates, 1994). This proposal derived from the converging evidence of the observed developmental asynchrony between lexical and grammatical development in naturalistic and parent report studies of children, as well as analyses of the computational principles captured in models of language learning (Bates & Goodman, 1997; Plunkett & Marchman, 1993).

However, the analyses presented here strongly suggest that the actual developmental relationship between reported vocabulary and grammatical complexity may be qualitatively different from that seen at that level of the measures. We found patterns of residuals consistent with the proposal that developments in lexicon and grammar actually occur synchronously. The observed relation between lexicon and grammar appears to be curvilinear because the mapping from the underlying grammar to its measure is nonlinear. Thus, while the original proposals required an explanation of the developmental priority of the lexicon over grammar, it appears that the curvilinear relationship is actually an artifact of a nonlinear mapping between underlying grammar and its measure.

Furthermore evidence for the curvilinear lexicon–grammar relation deriving from a nonlinear mapping at the level of the measures was obtained in
subsequent analyses that focused on the consequences of adding age to the model as a proxy for global developmental changes. According to the nonlinear mapping explanation, age and vocabulary should interact (positively and linearly) in predicting grammar. As predicted, the interaction between age and vocabulary was strong and accounted for significant variance. This interaction is a direct prediction from the mathematics of nonlinear mapping and, therefore, provides converging evidence for the proposal that the measure of grammar is an accelerating function of the underlying construct. The hypothesis that lexicon leads grammar cannot easily accommodate the interaction; there is nothing in this proposal that predicts that the effect of lexicon should increase with age.

What are the consequences of these findings for theories of lexical–grammar relations? We must first note that several previous studies have proposed that lexical and grammatical growth do indeed proceed synchronously. For example, Anisfeld, Rosenberg, Hoberman, and Gasparini (1998) reported on timing relations between changes in the growth of vocabulary size and the onset of word combinations examined at weekly intervals over the course of 8–10 months in five children. They found that accelerations in lexical growth (i.e., the vocabulary spurt) were coincident with the onset of word combinations and that lexical growth tended to decelerate with continued growth in word combinations (see also Dromi, 1987; van Geert, 1991). These close-timing synchronies were interpreted as evidence that grammar learning is well underway before the vocabulary spurt and that grammatical analyses facilitate further lexical learning (i.e., a type of syntactic bootstrapping, Gleitman & Gleitman, 1992; Naigles, 1990). Thus, lexical growth does not occur presyntactically, but rather is “part and parcel of the child’s transition to grammatical language” (Anisfeld et al., 1998, p. 166).

But, does evidence for developmental synchrony of lexicon and grammar undermine the notion that the developing lexicon provides a critical foundation for grammar learning? While it may seem at first blush to be the case, the notion of a critical mass is not in and of itself incompatible with the idea that systemic change is ongoing simultaneously in both lexical and grammatical domains. Even if a sufficient number of lexical forms may be required for the system to abstract and apply certain grammatical regularities, the computational principles guiding these abstractions are operating in the context of a highly interactive and multiply determined system. In such systems, changes are simultaneously occurring at multiple levels across the system, to a more or less degree, and are not limited to particular time windows. It would be unwarranted to assume that the child is not learning anything about grammar until grammatical principles are able to manifest themselves in observable ways (e.g., the production of overregularized forms). It is well known that children demonstrate sensitivity to grammatical regularities in comprehension before the point when they can reliably use those grammatical forms in production (e.g., Shipley, Smith, & Gleitman, 1969). Thus, the fact that parents report that children show little grammatical sophistication during periods of early lexical growth is likely to be more a function of the form of the knowledge that is being reported (i.e., production of closed-class inflections or multiword phrases), rather than an index of a lack of grammatical knowledge per se.

In addition, the mechanisms underlying critical mass effects are most reasonably explained in terms of links between precise lexical accomplishments and the emergence of very particular abstract patterns that form the basis for grammatical regularities. For example, it seems reasonable to propose that children’s productive use of past tense forms (i.e., overregularization errors) should be most dependent on learning a particular set of lexical items that provide the relevant cues from which to abstract the regular past tense pattern, rather than total vocabulary size. Indeed, rather than a general measure like total vocabulary size and grammatical complexity, the analyses presented in Marchman and Bates (1994) and the computational model of Plunkett and Marchman (1993) focused on relations between the number of regular verb forms and the productive use of past-tense inflections. In children, these more focused measures of lexical accomplishments (i.e., size of regular verb lexicon) are highly intercorrelated with “omnibus” measures of reported vocabulary. Yet, it would be naïve to assume that these strong intercorrelations indicate that critical mass effects are actually operating in a monolithic fashion across all of lexical development and all of grammar. Thus, it would be productive for future work to go beyond global measures of lexical and grammatical progress and to begin to map in more precise ways which particular features of children’s lexical knowledge do (and do not) serve as the foundational precursors for the child’s abstraction of specific grammatical regularities.

Finally, recent work has identified crosslinguistic differences in the shape of lexical–grammar relations that depend on the particular measures used. In Caselli et al. (1999), lexical–grammar relations in both English and Italian exhibited the characteristic nonlinear
function when grammar was measured using the complexity scale from the CDI, as analyzed here. However, linear relations were seen when grammar was indexed by the child’s reported use of function words, a finding that was interpreted to suggest that Italian-learning children might require a smaller lexical base of open class content words before they can begin to produce closed-class function words. Furthermore, in a recent follow-up study, Devescovi et al. (2005) reported that linear relations were again observed in Italian (but not English) when grammar was indexed by yet another measure, “mean of the three longest utterances” (M3L). Such crosslinguistic differences in the shape of lexical–grammatical relations remain compatible with a view in which lexical development drives the abstraction of grammatical regularities to the extent that these differences can be accounted for under the assumptions of a multiply determined and highly interactive system. Thus, linear lexical–grammatical relations would be more likely to occur in Italian than English, given that the “relatively rich, regular and consistently marked grammatical system in Italian may provide an easier target, requiring few exemplars (and smaller vocabularies) to support extraction of strong generalizations” (Devescovi et al., 2005, p. 783). These sorts of predictions fall naturally out of a view of language in which lexical and grammatical development are tightly yoked within the context of a unified learning system. Future studies should continue to examine the degree to which these developmental relations are obtained empirically in different languages and across a wider range of measures.

Of course, the conclusion that lexicon and grammar develop synchronously across the developmental period cannot escape another set of theoretical implications. Like the proposal that lexical development drives grammatical development, the developmental synchrony of lexicon and grammar could be a consequence of them both being driven by a single underlying factor. For example, if lexical and grammar learning occurs within a unified system, then particular features of the input to the system (i.e., a common set of environmental influences) may control both the acquisition of new words and new grammatical forms. It is well known that the features of the talk that children hear (i.e., quality and quantity of speech to the child) can substantially impact vocabulary learning (e.g., Hart & Risley, 1995; Huttenlocher, Haight, Bryk, & Seltzer, 1991), and have long-term consequences for performance on assessments of language and cognitive skill more generally. Thus, it is feasible to propose that the specific features of the language-learning environment (e.g., amount or quality of talk) that enhance children’s learning of words might also be those that enhance the mechanisms guiding the acquisition of grammar. Importantly, such an interpretation would not require that lexicon and grammar share key mechanisms or representations in any theoretically important way.

Alternatively, developmental change in the system may be governed by a single system–internal control parameter. For example, Thelen, Schöner, Scheier, and Smith (2001) showed that developmental change in the A-not-B error could be explained by changes in the “cooperativity” of the dynamic field, a single parameter that controls the field’s resting level. Thelen et al. described differences in this parameter as analogous to changing the weight of internal processes relative to the weight of external input. The synchrony relationship is consistent with the idea that these two aspects of language are controlled by a single, system-wide parameter - “cooperativity” might be one candidate, but clearly others would also be worth investigating.

One such candidate could be general cognitive or information processing efficiency (Kail, 2000). Several recent studies have suggested that improvements in the efficiency of spoken language understanding are associated with growth in expressive vocabulary size (Fernald, Swingley, & Pinto, 2001; Fernald, Perfors, & Marchman, 2006; Zangl, Klarmann, Thal, Fernald, & Bates, 2005). Indeed, it may be the case that children’s early success in language learning is in fact facilitated initially by their limited processing capacity (e.g., Elman, 1993, Newport, 1990). If the subsequent development of language is linked to global characteristics of a child’s developing information processing system, it is certainly reasonable to propose that processing factors could impact a child’s progress in both learning words and learning grammar, in the absence of any more specific links between the two.

The possibility that lexical–grammatical links, whether they be synchronous or otherwise, are entirely indirect, driven by the mutual impact of mechanisms, or representational requirements that operate outside the lexical and grammatical systems per se, has been addressed in several recent studies. For example, using multivariate behavioral genetic techniques, Dale et al. (2000) argued that there is a substantial genetic influence on the relationship between vocabulary and grammar and that general abilities lacking a strong verbal component are not likely to be responsible for pacing the developments in both domains. Furthermore, in a recent follow-up study, Dionne, Dale, Boivin, and Plomin (2003) replicated the heritability findings of the 2 years sample reported in Dale et al. (2000), but also attempted to
address the directionality of the effects using a cross-
lagged technique. They concluded that lexical
knowledge was related to grammatical level, as well
as grammatical level facilitating lexical learning (i.e.,
syntactic bootstrapping).

Finally, recent studies with children learning two
languages have demonstrated that lexical–grammar
links do not “spill over” from one language to the
other, but rather are closely yoked to the accomplish-
ments within each of the languages being learned
(Conboy & Thal, 2006; Marchman et al., 2004). In
Marchman et al.’s study of more than 110 children
learning both English and Spanish, the results indi-
cated that grammatical abilities were strongly tied to
lexical level in each language (i.e., Spanish grammar
scores were predicted by Spanish vocabulary; English
grammar scores were predicted by English vocabu-
lary). Furthermore, grammatical accomplishments in
each language were uniquely related to vocabulary
scores in each language, and hence, were not attrib-
utable to other factors that may have been guiding the
child’s general progress in language learning (e.g., age,
mother’s years of education, relative English-to-Span-
ish exposure, lexicon, and grammar in the other lan-
guage). Similar patterns were seen using naturalistic
language samples. These results suggest that lexical–
grammatical links are not an artifact of a global language-
learning ability, but rather that these accomplishments
are tied together in a very precise way in the context of
solving a very precise problem (i.e., becoming a pro-
ficient language user of a particular language).

One other type of explanation may be advanced to
explain the synchrony relationship—lexical and
grammatical development may be mutually and re-
ciprocally influential at a relatively fine time-scale.
That is, small gains in one result in small gains in
the other. According to this hypothesis, separate
systems of lexicon and grammar are involved in a
constant, cyclical interchange. These sort of bidirec-
tional and continuous effects might appear as syn-
chrony relations (although they could also create
other functional forms as well), but are actually the
consequence of temporally fine-grained interactions
across, rather than within, systems. To our knowl-
edge, studies designed to test this kind of continuous
relationship have yet to be undertaken.

While this is no exhaustive list of the hypotheses
capable of explaining the strong lexicon and gram-
mar associations without resorting to an emergentist
view, such hypotheses will generally either contain a
third factor that drives the relationship or some form
of reciprocal causation. Clearly, ruling out all possi-
bile alternative explanations is a daunting task.
However, the evidence for strong relations between
lexicon and grammar is strikingly robust across
studies that examine a variety of populations and
adopt several different methodologies. Furthermore,
recent studies have taken important steps toward
ruling out several of the leading alternative hypo-
theses—that is, artifacts of common environmental
influences or a general cognitive or language-learn-
ing skill. Future work should continue to explore a
broad range of alternative explanations, and their
implications for interpreting developmental ordering
as evidence of shared system-internal processes.

An Example of Testing Developmental Ordering

A second major purpose of the current article was
to provide a complete example of how the nonlinear-
mapping approach (Dixon, 2005) could be used to
test developmental ordering hypotheses. We showed
that the curvilinear relationship between lexicon and
grammar could result from synchrony at the
underlying level, but only under a very limited set of
nonlinear mapping conditions. Specifically, the
measure of grammar must be an accelerating func-
tion of underlying grammar and/or the measure of
lexicon must be a decelerating function of under-
lying lexicon. Because these nonlinear mappings
between the measure and underlying construct will
create specific relationships between error and the
underlying variables, they make predictions about
the patterns that will be observed in the residuals.

Although the model fits were quite good by most
standards (e.g., 73–74% of the variance explained),
the residuals were systematically heteroscedastic.
When vocabulary was used to predict grammar, the
pattern showed a pronounced wedge shape—the
variance increased as the predicted values of gram-
mar increased. To quantify this relationship de-
scriptively, we correlated the absolute value of the
residuals and the predicted values. This very familiar
procedure captures the pattern to some extent,
but a slightly more complex strategy is required to
model formally the relationship between residuals
and predicted values. Using a type of standardized
residual and appropriate test statistic (which col-
lectively handles the violations of distributional as-
sumptions that are problematic for testing residual
patterns), we showed a strong, positive relationship
between the predicted values and residuals.

A very different pattern of residuals emerged
when grammar was used to predict vocabulary.
The pattern appears curvilinear with two bends
(i.e., cubic). Both the familiar descriptive methods
(i.e., OLS regression) and the more formal analysis
of standardized residuals confirmed the visual
impression in the plot. The pattern of residuals was cubic and there was only a very weak negative relationship with the predicted values.

Both these residual patterns were consistent with those predicted by the hypothesis that the lexicon and grammar develop synchronously but that the measure of grammar was an accelerating function of the underlying level. The only other logical possibility (i.e., the measure of lexicon is a decelerating function of underlying lexicon) capable of creating the observed curvilinear data pattern predicted quite different residual patterns. Note that had we observed homoscedastic patterns or heteroscedastic patterns that conflicted with the predicted patterns (e.g., a negative relationship between the absolute values of the residuals and predicted values of grammar), the proposal that lexicon and grammar develop in synchrony would have been placed under considerable pressure. Because the nonlinear mapping must create the predicted residual patterns, a failure to observe them is good evidence against nonlinear mapping. Indeed, the only way to arrive at a homoscedastic pattern given a nonlinear mapping (and nontrivial error at the underlying level) is to have a true association between error and the variable of interest that runs in the opposite direction. Making inferences from correlational data requires the assumption that such compensatory relationships are not driving the results. In this way, failure to observe the predicted residual patterns would constitute correlational evidence against a nonlinear mapping between the measure and underlying construct.

**Extending the nonlinear mapping approach to multiply determined systems.** Most current approaches to developmental science assume that development is a complex, multiply determined (or multiply probabilistic) phenomenon. Researchers hope to identify many, perhaps even most, of the central factors that contribute to the development of a particular item (e.g., a structure, skill, ability), but rarely would one assert that they knew the full set of developmental influences. We often handle this multiplicity of developmental influences by using proxy variables, such as age, socioeconomic status, birth weight, and so forth; we assume that these global variables capture some of the variance attributable to unmeasured developmental dimensions. For example, age is often taken as a proxy for the host of dimensions that are believed to develop across time.

In the current context, we showed how the nonlinear mapping approach can be extended to capitalize on the multiply determined nature of the developmental system. We noted that the error term at the underlying level carries effects that are not currently in the model. Adding a proxy variable to the model essentially splits its effect from the underlying error term. Because the proxy variable will have its effects on the measure through the nonlinear mapping, it must become related to the other predictor (or predictors) in the model, just as the error term did. That is, the nonlinear mapping must create an interaction among the predictors. Therefore, a second level of evidence can be brought to bear on the proposal of a nonlinear mapping between the measure and construct—if a nonlinear mapping exists, variables that capture general or global effects (and are often of little theoretical interest) are predicted to interact with the other predictors in the model.

A corollary of this prediction is that the heteroscedastic nature of the residual patterns will remain after the proxy variable and the interaction term are added to the model. Because the heteroscedastic patterns are caused by the nonlinear mapping, adding additional terms to the model cannot remove the relationship between error and the predictors. Whatever error is left at the underlying level must still undergo the nonlinear mapping and, therefore, become related to the other terms in the model. This point is important to note, because an alternative explanation for heteroscedasticity is that a variable, which is correlated with the current predictors, is missing from the model. According to this alternative explanation, it should be possible to nearly eliminate the heteroscedasticity of the residuals. The nonlinear mapping account predicts that heteroscedasticity cannot be eliminated in this way.

**Sources of Nonlinear Mapping**

It is clearly disheartening to find out that a measure is not doing a consistent job of capturing change across the developmental period of interest (i.e., there is a nonlinear mapping between one’s measure and the construct of interest), as we have demonstrated for grammatical complexity, in the current case. How and why does this situation arise? It is likely that the reasons for a nonlinear mapping will vary depending on the features of the particular measures and the characteristics of the particular constructs under consideration. For parent report, it is not surprising that the kinds of grammatical changes observed early in development are less likely to be those that parents can easily report on. For other measures, the situation might arise due to different developmental requirements of the responses that are required by the child (e.g., drawing,
pointing at pictures). As mentioned above, studies using methodologies with different task demands (e.g., looking preference) have demonstrated that children do know something about grammar earlier than they can begin to produce early word combinations similar to those listed on the CDI (e.g., Naigles, 2002). Thus, future studies might find it advantageous to use a combination of techniques to assess changes that occur both early and late in the period, and how those changes relate to lexical growth.

The current findings do not discredit the CDI as a useful instrument to measure language progress, nor do they bear exclusively or particularly to measures that are derived from parent report (see Fenson et al., in press for a discussion of the strengths and limitations of parent report). This particular instrument, and the technique of parent report more generally, remain worthwhile methods for developmentalists to use. As with any measure, the data derived from the CDI must be interpreted in light of potential nonlinearities in measurement. We have demonstrated here that nonlinear mapping in measurement can have significant theoretical consequences. Nonlinearities of the sort uncovered for the grammatical complexity measure from the CDI are almost certainly not unique to this instrument; indeed, they are likely to be more common than not in our field. The techniques described here illustrate that knowing more, rather than less, regarding the relations between the measures and underlying variables can have important consequences for hypotheses about the developmental ordering of core theoretical constructs.

Summary and Conclusion

We examined the developmental relationship between lexicon and grammatical complexity using data from the norming sample of the MacArthur–Bates CDI: W&S. Results suggested a significant re-interpretation of the nonlinear relation between lexicon and grammar, a data pattern that was previously thought to imply that lexical development precedes grammar. Furthermore, we showed how to extend the nonlinear mapping approach (Dixon, 2005) to multiply determined systems, such that a second level of evidence was brought to bear on the question of developmental ordering. The analyses converge on the conclusion that the lexicon and grammar are actually developing in synchrony across the first few years of life. This finding has important implications for theories of language acquisition that view lexical and grammatical development within a unified system that shares important representational and computational resources.

References


McGregor, K. K., Sheng, L., & Smith, B. (2005). The precocious two-year-old: Status of the lexicon and

Appendix A

Modeling heteroscedasticity with standardized residuals. To compute the standardized squared residuals, one first creates a term closely related to the mean squared error from the current regression model. However, the sum of squared (SS) error is divided by $N$, rather than $N - k - 1$. This modified variance term is therefore: $\sigma^2 = \Sigma(e_i^2)/N$, where $e_i$ is the residual for individual $i$. This value is easily obtained because $\Sigma(e_i^2)$ is the SS error term from the regression that generated the residuals. Next, each residual term is squared and divided by $\sigma^2$. These standardized squared residuals are then subjected to a regular OLS regression analysis. For example, to test whether the predicted values of grammar were linearly related to the residual values, we used the predicted values of grammar as the predictor variable and the standardized squared residuals as the dependent variable. The final step is to obtain the sums of squares for the model from the regression on the standardized squared residuals. Dividing this value by $2$, $SS_{\text{model}}/2$, yields a test statistic that is asymptotically distributed as chi-square, on degrees of freedom equal to the number of predictors in the model. See Fox (1991) for a very accessible review of this approach and a discussion of alternative methods.

An example of a more extreme nonlinear mapping: The cube. First, we consider the expansion of $(x + \beta_1 * L_{w} + \beta_3 * A_{w} + e_{w})^3$. To increase the clarity of the presentation, we use the following notation: $L = \beta_1 * L_{w}$, $A = \beta_3 * A_{w}$, $E = e_{w}$. Assuming that $\alpha = 0$, the expansion of the cubed equation yields the following terms: $L^3 + 3A^2L + 3L^2A + 3L^3 + 3A^2E + 3AE^2 + 6LAE$. The expansion shows the quadratic interaction term, $3L^2$.

To show the quadratic prediction more completely, we created a simple model of this nonlinear mapping situation, given an underlying relationship in which the lexicon developed more slowly than grammar (i.e., the opposite partial priority hypothesis). We created a data set in which the underlying relationship between grammar, lexicon, and age was: $G_w = \alpha + \beta_1 * L_w^2 + \beta_3 * A_w + e_{w}$, where $\beta_1 = 20$, $\beta_3 = 1$, and $e_{w}$ was drawn from a normal distribution with a mean of 0 and variance of 50.

This model creates a curvilinear relationship such that grammar develops slightly more quickly than the lexicon; when the underlying relationship is plotted as in Figure 1, the curve is concave (i.e., opening downward) rather convex (i.e., opening upward). We explicitly choose to model a mild version of the opposite priority hypothesis, because it
requires a less extreme nonlinear mapping to create the observed, convex pattern. Less extreme nonlinear mappings will have smaller effects.

The nonlinear mapping between underlying grammar and the measure of grammar was represented as the cube: \( G_m = (G_u)^3 + \epsilon_m \). The measure of lexicon, \( L_m \), was a linear function of underlying lexicon, \( L_u \). The error term, \( \epsilon_m \), was drawn from a normal distribution with a mean of 0 and variance of 50. Finally, to simulate the developmental relationships between lexicon, grammar, and age more closely, the measure of lexicon was created such that it was correlated with grade (as well as with grammar). The correlations among the three variables: grammar, lexicon, and age, closely approximated the zero-order correlations in the CDI data set (.85, .67, and .64, for grammar–lexicon, lexicon–age, and grammar–age, respectively).

When \( L_m \) and \( L_m^2 \) were used to predict \( G_m \), the model fit was, of course, very good, \( R^2 = .88 \). Adding Age and the linear interaction term, lexicon \( \times \) age, contributed significantly to the model, increasing the fit by 7% and 4%, respectively. The crucial question for our current purposes, however, was whether adding the quadratic interaction term, lexicon \( \times \) age\(^2\), to the model would also improve the fit, as suggested by the expanded equation presented above. The quadratic interaction was significant, \( B = 11.06, t(993) = 12.88 \), when added to a model that included all the terms above and Age\(^2\).

This simple model demonstrates that, given even a weak underlying opposite-priority relationship (i.e., a concave curve), the nonlinear mapping necessary to create the observed relationship (i.e., a convex curve) also creates a quadratic interaction between the major predictor (i.e., lexicon) and the secondary variable (i.e., age). Therefore, testing for this quadratic interaction provides one way to distinguish between an underlying opposite-priority relationship and an underlying synchrony relationship.