ARTICLE

FIRST LANGUAGE

http://fla.sagepub.com Vol 29(3): 266–288 (200908)
DOI: 10.1177/0142723708101679

Parental input and connective acquisition: A growth curve analysis

Rosie van Veen, *Utrecht University*
Jacqueline Evers-Vermeul, *Utrecht University*
Ted Sanders, *Utrecht University*
Huub van den Bergh, *Utrecht University and Amsterdam University*

ABSTRACT

In this study, the influence of parental input on the acquisition of discourse connectives was investigated. Three factors were hypothesized to play a role in contributing to the course of language acquisition: first, an increase in age, and hence, an increase in conceptual abilities; second, short-term frequency effects (effects of parental input in the space of one recording); and third, long-term frequency effects (effects of the cumulative parental input over a longer period of time). The authors developed a growth curve analysis and used this to analyze data from a dense longitudinal corpus of a German boy aged 1;11.12–2;11.27. Results show that each factor has a significant effect on the acquisition of the German connectives *aber* 'but,' *damit* 'so that,' *und* 'and,' *weil* 'because,' and *wenn* 'when' and should always be taken into account when studying connective acquisition. Furthermore, growth curve analysis promises to be an innovative tool to study factors influencing the course of language acquisition.

KEYWORDS

Connectives; first language acquisition; frequency effects; growth curve analysis; parental input

A more or less classical approach to issues of first language acquisition is one in which linguists and psycholinguists aim at pinpointing the processes that influence the exact course of acquisition. Time and again, such studies have found support for a cumulative complexity theory of some kind, which proposes that this course of acquisition is influenced by factors that are related to the child's increase in age, such as an increasing ability to deal with words and structures that are cognitively more complex (see Brown, 1973; Clark, 2002; Clark & Clark, 1977; Piaget, 1969). In recent years, a growing number of studies – especially those taking a usage-based approach – have stressed the importance of the ambient language for the linguistic development of a child (see Diessel, 2004; Tomasello, 2006). The general claim is that the amount and type of input children receive influences the development of their linguistic abilities (see Borovsky & Elman, 2006; Kidd, Lieven, & Tomasello, 2006; Matthews, Lieven, Theakston, & Tomasello, 2005; Valian & Casey, 2003). Traditionally, cumulative complexity theory and the parental input theory have been studied by investigating two main factors, namely, increase in ability and effect of parental input respectively.

In first language acquisition studies of typically developing children, an increase in cognitive abilities is usually associated with an increase in age – though not necessarily in a linear relationship. In fact, researchers who are interested in the role of increasing cognitive abilities during the course of acquisition do not tend to take cognitive ability itself as their independent variable, but age (e.g., by comparing the linguistic capacities of children in different age groups). This need not be problematic: if increase in age indeed proves to be a prominent factor, it is more likely that age differences have to be associated with differences in cognitive abilities rather than with other developments related to age, such as neurological developments and physical maturation. In the current study, we also take age as an independent variable, claiming that this is a good indicator of the cognitive abilities of typically developing children.

Regardless of its importance, the age factor is usually not included in studies on the parental input theory; in fact, it is largely ignored. Still, even if one were interested in parental input as an explanation for language development, it is not entirely clear how this input theory should be operationalized. Studies investigating the relation between input and output usually suffice by studying overall parental frequencies or frequencies of use per developmental 'stage' (e.g., Aksu-Koç, 1998; Krasinski, 1995; Naigles & Hoff-Ginsberg, 1998; Sandhofer, Smith, & Luo, 2000; Shirai, 1994; Wijnen, Kempen, & Gillis, 2001). For example, Huttenlocher, Haight, Bryk, Seltzer, and Lyons (1991) found a correlation between the average frequency of content words in a sample of 11 mothers' child-directed speech and the mean age of first appearance in the children's output. In another study, aiming at investigating the effect of language input on the acquisition of root infinitives, Wijnen et al. (2001) did not group the data, but divided the corpora into four sections representing developmental phases, which allowed for a general overview of language development. This method is restricted, because dividing up data according to developmental stage does not allow for a linear overview of the data. Moreover, dividing up data in a different way (e.g., according to mean length of utterance [MLU] or age in months) will almost certainly lead to different results.

Thus, if we want to investigate the influence of parental input throughout the course of acquisition, we need to take a longitudinal perspective and present input and output data on the basis of a timeline. Furthermore, since typically developing children will also experience an increase in their cognitive abilities, it is important to link parental frequencies of use to specific ages on a longitudinal scale. The first goal of this article is to develop a method that allows us to investigate the contribution of the factors age and input frequency to the course of language acquisition.

In order to establish frequency effects of parental input, longitudinal studies have often operationalized input frequency by calculating the overall frequency of use in the adult language. For example, in a study focusing on verb acquisition, Naigles and Hoff-Ginsberg (1998) reported that total verb frequency and final position frequency were among the most important factors in the input determining the ease of acquisition. Similarly, where the development of conjoined clauses is concerned, Diessel (2004) related the age of emergence of the various types of conjoined clauses to their overall frequency in the input. In other words, these studies focused on long-term effects of input frequency, or the cumulative effect of parental input the child has been exposed to over a longer period of time. However, another way of operationalizing input frequency is by looking at the effect of parental input on the child's output in the space of one recording. This allows for a focus on short-term effects, such as imitation.

However, where input frequency is concerned, there are several ways to operationalize this factor. This gives rise to the second goal of this article: disentangling short-term from long-term effects. This can be explained by considering the single passage in (1). This passage could lead one to conclude that Leo's production of *aber* ('but') is related to, or even an imitation of, his father's previous use of the connective. This view is supported by certain accounts of the usage-based approach in which it is expected that children will, to a certain extent, imitate what their parents do (cf. Tomasello, 2003). Hence, this passage argues for a short-term effect of language input:

(1) Father: Ja okay, dann helf ich mit.
Father: **Aber** was machen wir denn mit Eule und Osterhase?
Leo: Eule kriegt dies
Leo: **Aber** Osterhase wollte keines. [2;09.00]
Father: Was wollte Osterhase nicht haben?
Leo: Kein Spielzeug

Father: Yes okay, I will help.
Father: **But** what shall we do with owl and Easter bunny?
Leo: Owl gets this.
Leo **But** Easter bunny wanted none [2;09.00]
Father: What didn't Easter bunny want to have?
Leo: No toys.

However, when we consider the passage in (2), which was taken from the same recording, we find that Leo is perfectly capable of producing *aber* of his own accord. This implies that Leo's use of *aber* at age 2;09.00 cannot be the mere result of short-term effects of the adult input, but rather a result of all the language Leo has ever been immersed in. In other words, Leo's production of *aber* in (2) could not only be influenced by the instances of *aber* he hears from his parents in this particular file, but also by all the other instances of *aber* he has heard before the age of 2;9.00. This could indicate a long-term effect of language input on Leo's use of *aber*. A similar idea was also put forward by Abbot-Smith and Tomasello (2006), who claim that 'every utterance a child hears and processes has a lasting effect on linguistic representations' (p. 283). A longitudinal analysis in which both short-term and long-term effects are taken into account should reveal whether this line of reasoning holds true.

(2) (*Context:* talking about a race between two model trains)
Father: Ja, wenn's über die Brücke, dann verlieren die sich gegenseitig.
Father: So gut ist die Verbindung dann doch nicht zwischen denen.
Leo: **Aber** auf e Abstellgleis gefahren. [2;09.00]

Father: Yes, when it crosses the bridge, it will surely lose.
Father: So, the connection between them is not good. [2;09.00]
Leo: **But** it drove over the sidetrack.

So far, we have examined three factors that could influence the course of language acquisition – cognitive growth effects related to age, short-term effects of language input, and long-term effects of language input – as if they operate independently of one another. However, it is also possible that parental input and the cognitive capacities of the child interact. This would be in line with ideas about 'audience design' (Clark & Murphy, 1983), which is a term similar to what is known as parental scaffolding (McCabe & Peterson, 1997; Rome-Flanders, Cronk, & Gourde, 1995; Stevens, Blake, Vitale, & Macdonald, 1998). As the term itself suggests, audience design involves the speakers' tailoring of the utterances to suit their audience. In child-directed speech this means that, as children develop, parents adapt their speech to the child's increasing language abilities. In so doing, they progressively increase the complexity of their language to encourage the child to produce slightly more complex structures. In their study on causality, McCabe and Peterson (1997) investigated parent's child-directed speech as a source of language development. They found that audience design played a prominent role in children's acquisition of causal links: 'parents scaffold their children's emergent causal language by asking questions, repeating or revising such questions, occasionally answering those questions themselves, or abandoning them' (p. 151). This scaffolding behavior preceded children's productions by five or six months. A positive relation is also reported by Stevens et al. (1998), who found a relation between maternal labeling and the vocabulary size of children aged 9 and 15 months. On the other hand, Rome-Flanders et al. (1995) did not find a link between infant language level and maternal scaffolding behaviors. It seems interesting, then, to explore the possible occurrence of audience design in the area of connective acquisition.

The current study focuses on separating age effects from two types of input effects: short-term effects of language input, i.e., an effect of parental input on the child's output in the space of one recording; and long-term effects of language input, i.e., a cumulative effect of the parental input the child has been exposed to over a longer period of time. We have just seen that the field seems to be in need of a method to study the role of both these factors influencing the language development of a child. For this reason, we set out to develop an innovative method that would enable us to do this and simultaneously shed new light on parental input theory. This is done by investigating two main research questions: (1) Does parental input affect the language development throughout the course of acquisition? and (2) Can we distinguish different types of effects of input frequency from effects of age during the course of acquisition? In addition, we perform a preliminary investigation into audience design as an explanatory factor. Our results are demonstrated by use of a case study on German connective acquisition consisting of a dense longitudinal corpus of spontaneous speech data.

Where the acquisition of connectives is concerned, both parental input and age have been stressed as factors influencing the course of development. Diessel (2004) found that parental input seems to play a role: 'the more frequent a specific conjoined clause appears in the mothers' data, the earlier it emerges in the children's speech' (p. 172). Note that this study did use longitudinal data, but based its analysis on overall frequencies. On the other hand, Evers-Vermeul and Sanders (2008) claim that the cognitive complexity of connectives and coherence relations sets the pace for acquisition, at least in part (see also Bloom, Lahey, Hood, Lifter, & Fiess, 1980; Evers-Vermeul, 2005; Spooren & Sanders, 2008). This complexity factor is closely related to increase in age: at first children will only be able to deal with relatively easy items, and as they grow older – and their cognitive abilities increase – they will be able to deal with increasingly complex items. For instance, Evers-Vermeul and Sanders (2008) found that children are able to express positive relations (*and*, *and then*) before negative ones (*but*, *however*) and that they are able to express additive relations (*and*) before they master causals (*because*). However, this study only took the first emergence of connectives into account and did not investigate the role of parental input.

For the current study, the connectives *aber* ('but'), *damit* ('so that'), *und* ('and '), *weil* ('because'), and *wenn* ('when') were investigated because they were expected to occur early on in the acquisition process. These connectives occur frequently in the adult language and they express relatively easy types of coherence relations (for a discussion of the relative complexity of coherence relations and connectives, see Evers-Vermeul, 2005; Evers-Vermeul & Sanders, 2008; Sanders, Spooren, & Noordman, 1992; Spooren & Sanders, 2008).

METHOD

Data

We analyzed the data from one parent–child pair in a dense longitudinal corpus. The corpus is that of a monolingual German boy, Leo, and was made available by the Max Planck Institute, Leipzig; the relevant background information was provided by Behrens (2006). Both parents have completed a higher education and speak dialect-free, clearly articulated standard High German. The current study only analyzed the files that resulted from the first year of recording, during which Leo was aged 1;11.12–2;11.27 (714–1094 days). In this period, five one-hour samples per week were made, which means that we are assured of a dense data set. In addition, his parents kept a daily diary in which Leo's newest and most complex utterances were noted. Files that did not contain any adult speech, because they were mostly based on diary recordings, were excluded from the analysis. In total, 257 files were found suitable for analysis.

The input during the first year consisted mainly of Leo's parents and a research assistant, though his mother was the primary caregiver of the child. Once a week the mother worked in another job, so the father took over childcare and the recordings. Also, once a week, or more often if requested, the research assistant took care of the recordings. Apart from the parents, the research assistant was also included in the analysis. Still, we refer to any results regarding the input as 'parental input'; we are aware of the possibility that the input produced by the research assistant may differ from that of the parents, just as much as the type and amount of input may vary per parent. As a rule, no other adults or children were present during recordings. In this year, Leo was still an only child and did not yet attend kindergarten. Moreover, he was not exposed to indirect input through TV in his home environment, though his parents did read to him quite often. As a result, the recordings have captured a representative part of the input available to Leo.

Procedure

For Leo and his three caregivers, the frequencies of *aber*, *damit*, *und*, *wenn*, and *weil* per file were computed by using Computerized Language Analysis (CLAN) (MacWhinney, 2000). In order to perform a time-efficient preliminary analysis, these connective frequencies were calculated on the basis of blind counts, which means that non-connective uses were also included. In addition, the word count per file was computed so that the length of the recording could be taken into account. This is important since this length influences the frequency with which connectives appear in the transcript. The analysis of the parental frequencies was carried out for all recorded utterances, including those that were not child-addressed.

For each connective, we performed a logistic regression analysis (Fienberg, 1977). A logistic regression enables us to investigate the relation between a linear explanatory variable (age) and a dichotomous explanatory variable (absence vs. presence of connective) (see Menars, 2002). We modeled the probability of occurrence of the particular connective in Leo's speech as a function of three predictors: Age, Short-term input, and Long-term input. In order to establish any short-term effects, the relative input frequency per file was related to the absence or presence of that connective in the child's language in that same file. In order to establish possible long-term effects, we related the cumulative input frequency up to each file to the presence or absence of the connective in question in the child's language. For example, when measuring the cumulative effect of *aber* at day 824, the input of *aber* in file 824 and in the 110 previous files was considered.

We have chosen a polynomial model in which Age^0, Age^1, Age^2, and so on are used as predictors.[1] An advantage of such polynomials is that they are extremely flexible; they can take almost any shape, depending on the number of powers and the value of the regression coefficients (see Goldstein, 1995; van den Bergh & Rijlaarsdam, 1996). Please note that the number of explanatory variables (powers of age) is considered an empirical matter. Extra predictors (powers of age) were only added when the previously added variables contributed significantly to the explanation of the occurrence of a connective. Results are significant if the ratio of the regression weight (β) and the standard error exceed 1.965 (and hence $p < 0.05$). For ease of interpretation, not the coefficients but the net results are presented in the growth curves; a complete statistical overview can be found in the Appendix.

We refer to our method as a growth curve analysis. This type of analysis has shown its usefulness in a number of developmental areas, such as onset of tense marking (Hadley & Holt, 2006), development of reading abilities (Francis, Shaywitz, Stuebing, Fletcher, & Shaywitz, 1996), and vocabulary growth (Huttenlocher et al., 1991; Rescorla, Misrak, & Singh, 2000). Growth curve analysis has also been used to evaluate different types of metrics for assessing the complexity of language samples from school-aged children (Kemper, Rice, & Chen, 1995).

ANALYSIS AND RESULTS

First, we show how we arrived at the best fitting models with the smallest number of parameters and the best fit to the data. For each connective, these optimal models reveal that all three predictors – Age, Short-term input, and Long-term input – make a significant contribution. Second, we discuss the short-term effects in greater detail. Finally, we report on a first analysis that tests the theory of audience design.

For each connective, three models were tested in a logistic regression analysis. In the first model, the occurrence of each connective is modeled as a function of the predictor Age only; the predictors Short-term input and Long-term input are taken into account in the second and third model respectively.

Figure 1 presents the probability of the selected connectives occurring in Leo's speech. The results of our first model show that age significantly affects the probability of a connective occurring in Leo's speech: the use of connectives changes with age. This is the case for every connective, though the slope of the growth curve can vary per connective. For instance, Fig. 1 shows a period of rapid increase in the probability of *und* ('and') occurring in Leo's speech. For *und* the probability of occurrence is 0.15 around 714 days, which means there is a 15% chance that *und* will appear in the recording. However, after around 840 days, the probability of Leo producing *und* remains 1, which means that the child uses *und* in every taped conversation. Also, we see that the probability of *wenn* ('when')

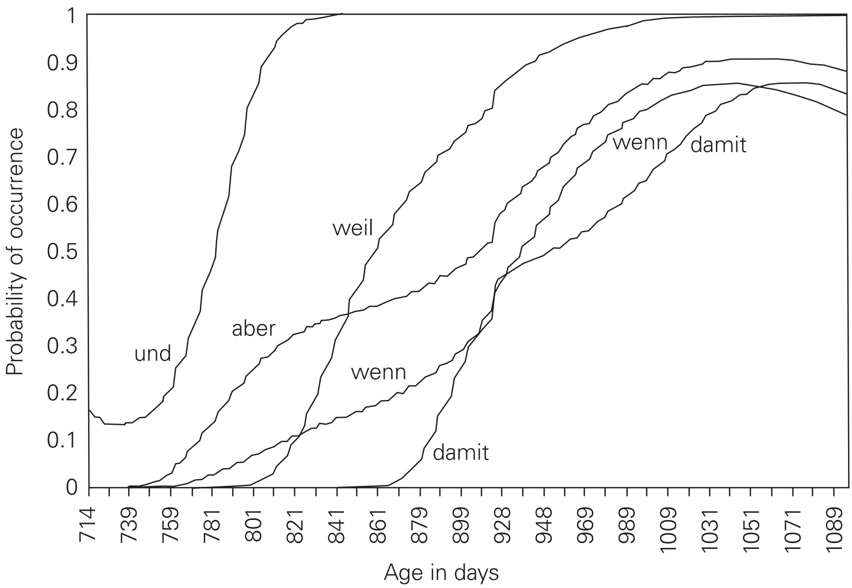

Figure 1 Probability of five connectives occurring (model with only Age as predictor)

occurring increases steadily from the first recording, whereas initially the probability of *damit* 'so that' occurring is very small, but rapidly increases after 870 days.

In the second model, the predictor Short-term input, the effect of parental use of the same connective on the same occasion, was added to the first model. The results show that the input of each connective has a significant effect on the probability of the connective occurring on top of the effect age has been shown to have (see the Appendix for relevant statistical details). This means that the amount of connectives in the language of Leo's parents has a short-term effect on Leo's connective use. For example, the use of *aber* in the input of a particular file influences Leo's production of *aber* in that same file (χ^2 (3) = 3573.6; $p < 0.001$).

In the third and final model, the predictor Long-term input was added to the second model. This predictor measures the cumulative effect of the input for each connective. The predictor Long-term input significantly influences the probability of a particular connective occurring in Leo's speech. For example, the cumulative frequency of *aber* in the input affects Leo's use of that connective (χ^2 (3) = 761.0; $p < 0.001$). Hence, this model was the best fitting model with the smallest number of parameters. Table 1 shows that the percentage of connective occurrences that are correctly predicted by the third model is relatively high for each of the connectives under investigation.

Effects of parental input

In the previous subsection, we found that the predictors Short-term input and Long-term input contribute significantly to each model. In the current subsection,

Table 1 Percentage of recordings that are correctly predicted to contain a connective

Connective	*Correctly predicted (in %)*
aber	77.2
damit	82.7
und	97.4
weil	91.6
wenn	80.7

we look at the nature of these effects. In order to do so, the growth curves resulting from the optimal models of *aber* ('but') and *weil* ('because') are discussed in greater detail (see Figs 2 and 3).

As Figs 2 and 3 illustrate, the parental input effects are characterized by steep peaks and dips in the growth curves and show a different pattern for each connective. The growth curves suggest that the parental influence is not continuous, but undergoes periods of little to no influence and periods of substantial influence. This pattern is roughly the same for each connective and can be divided into three stages: first, a period in which there is little influence; second, a period in which there is lots of influence; third, a final period in which there is little or no influence.

During the first period, the growth curves do not show any visible input effects; given the absence of peaks and dips, the input does not seem to directly influence the probability of a connective occurring.[2] This period of little influence is followed

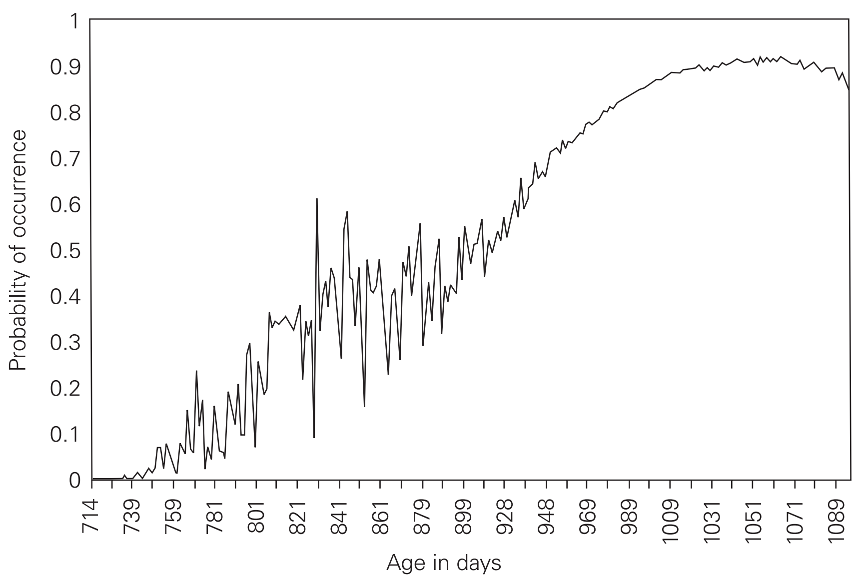

Figure 2 Probability of *aber* occurring

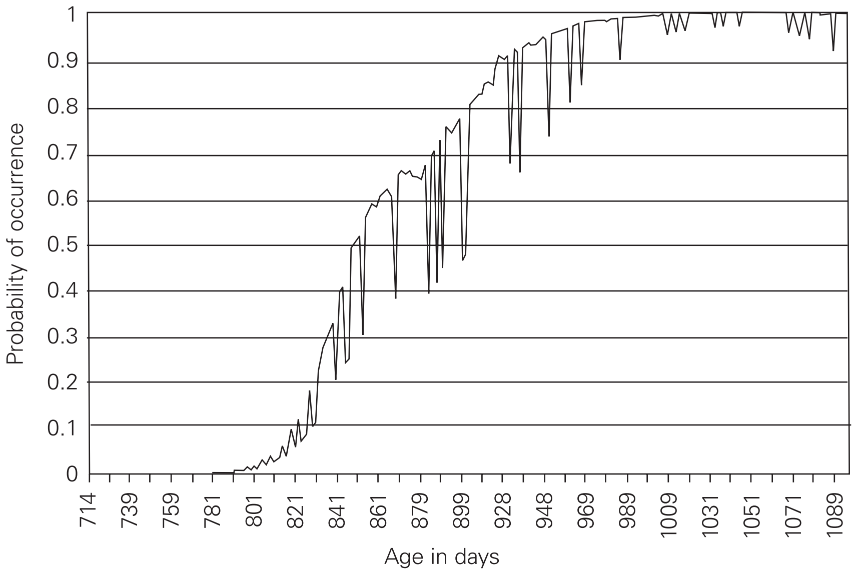

Figure 3 Probability of *weil* occurring

by a period with a large amount of influence. When we look at the growth curve for *aber* ('but') in Fig. 2, the raggedness of the curve shows that the direction of the influence differs per file. Thus, if the input has a large positive effect on the expected probability (i.e., the chance of Leo producing *aber*), this is indicated by a peak in the graph. Likewise, a negative effect on the expected probability is indicated by a dip in the graph. A positive effect is due to a high frequency of *aber* in the input; a negative effect is due to the lack of or low frequency of *aber* in the input. For example, when Leo is 828 days old we find that the probability of *aber* occurring in his speech suddenly drops to 0.09. When we look at the specific recording, we find that the parents produced four instances of *aber*, whereas in the previous 10 files they produced an average of 18 instances of *aber* per recording. Hence, according to the model, Leo's production of *aber* will be low too. Similarly, when Leo is 829 days, we find that the probability suddenly increases to 0.61. In this specific recording, the adults produce 40 instances of *aber*, which is the highest adult count of *aber* for all recordings analyzed. In this case, the model predicts Leo to show a relatively high production of *aber*.

Like the growth curve for *aber*, the growth curve for *weil* ('because') in Fig. 3 first shows a period of no parental influence, which is followed by a period with greater influence. After 821 days, the growth curve shows a rapid increase in the probability of *weil* occurring in Leo's speech. The effect of parental input is characterized by a number of large negative deviations. These sudden falls in the probability of *weil* occurring in Leo's speech are caused by a sudden decrease of *weil* in the input. This effect can be traced back to the actual recordings. For example, an inspection of the recording made at 890 days shows that it does not contain any instances of *weil* in the input, whereas an inspection of its surrounding files shows that they contain around five instances of *weil*. This finding is similar for every deviation in Fig. 3, but also for the sudden deviations in the growth curves of the other connectives. Hence, these large, sudden deviations are likely to be due to short-term influences of the input.

Apart from the first stage, the final stage in the growth curves also shows a lack of parental influence. For example, the curve for *aber* in Fig. 2 evens out from around 980 days. This smoothness shows that the effect of parental input becomes much smaller. Thus, the effect of parental input on Leo's production decreases and could be a sign of Leo's independent use of *aber*, indicating mastery of the connective. Similarly, the relatively smooth curve toward the end of the growth curve of *weil* in Fig. 3 indicates that, even though the probability of the connective occurring is high (around 100%), the influence of the input is small, suggesting mastery of the connective as well.

Note that Figs 2 and 3 do not provide visible cues that allow us to distinguish short-term effects from long-term effects. These can only be traced by comparing Figs 2 and 3 with the growth curves resulting from the earlier models, which include the predictor Short-term input, but not the predictor Long-term input. For reasons of exposition, we have chosen to only include the graphs from the final model. However, a comparison of these growth curves revealed that the short-term input was responsible for large peaks and dips, whereas the long-term input was responsible for flattening these peaks and dips. Apart from the visible output, the statistical analyses show that both factors contribute significantly to the model and that many of the peaks and dips in the graphs can be traced back to relatively high or relatively low frequencies in specific files (indeed, indicating short-term effects).

Audience design

The previous subsection revealed that the growth curves (except the growth curve of *und*) show an early period of little parental influence. For ease of discussion, such a period is illustrated again by the growth curve of *damit* ('so that') in Fig. 4. This growth curve shows a large period in which the child does not use *damit*. Only after around 850 days do we find that the parental input seems to affect Leo's use of *damit*.

One explanation for the occurrence of an early period of little parental influence could be that the connective simply did not occur in the adult input and could consequently not influence Leo's connective use. This would mean that the parents are somehow modifying their input to Leo's cognitive and linguistic capacities. This idea fits the theory of audience design (see Clark & Murphy, 1983). To test whether this is indeed the case, the relationship between the proportion of the connective use in the input and Leo's age was examined. The results of the connective *damit*, which experienced no effects from adult input up to 850 days, are illustrated in Fig. 5. This figure sets out age in days against the proportion of the connective in the input.

In the scatterplot – which is typical of the scatterplots of the other connectives – we find that the proportion of *damit* in the input remains stable over time. The fact

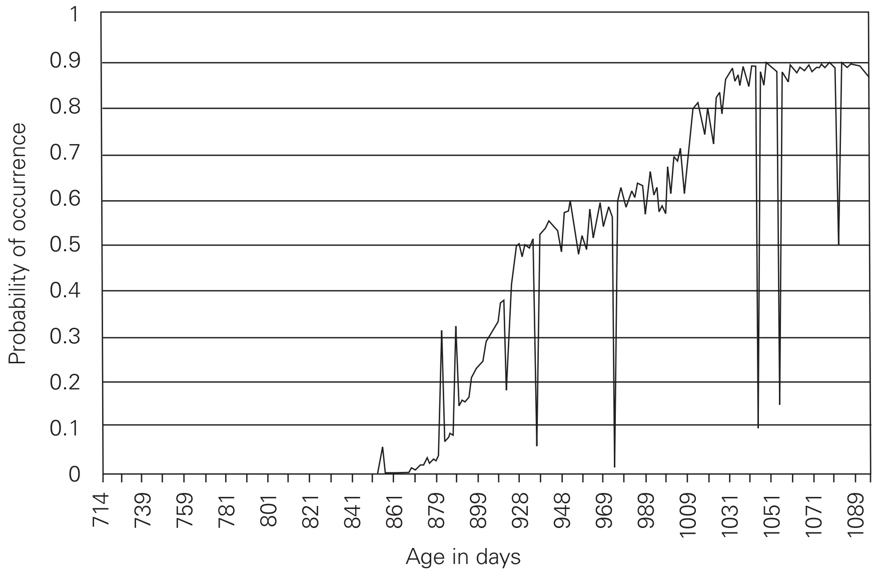

Figure 4 Probability of *damit* occurring

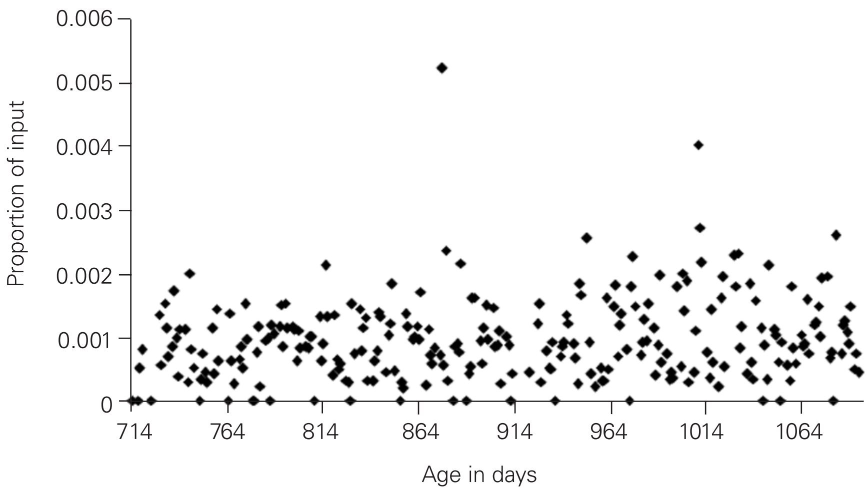

Figure 5 Proportion of *damit* in adult input

that input does not change over time indicates that Leo's parents did not adapt their use of *damit* to Leo's capacities. There do seem to be some outliers that stand apart from the clustered band around the proportion of 0.001. These outliers do not affect the absence of any trend. In fact, we find the same pattern for each connective in this case study (although these scatterplots are not reported here for reasons of space); the relative input frequency of each connective is stable across the recordings. Hence, it is unlikely that audience design is the reason for the lack of parental influence in the beginning of the acquisition process.

DISCUSSION

Overall, we find that – in addition to the factor Age – both the predictor Short-term input and the predictor Long-term input influence the acquisition of connectives. In the following subsections, we discuss how to interpret the significant contributions of each of these three predictors and return to the issue of audience design in the light of our correlational findings.

Increasing age and cognitive abilities

Each growth curve – with exception of the growth curve for *und* – starts with a period in which Leo does not produce the connective. This can be attributed to the predictor Age: a 3-year-old is expected to be able to use *damit* ('so that') or *wenn* ('when') to a certain extent, whereas a 2-year-old clearly is not. However, apart from the fact that the predictor Age contributes to the overall statistical model in a significant way, it is as yet unclear what this entails. We have suggested it involves an increase in the child's increasing ability to deal with words and structures that are cognitively more complex. This means that the initial period during which Leo does not produce a connective is explained by his not being able to deal with its semantic and/or syntactic complexity (see Bloom et al., 1980; Brown, 1973; Evers-Vermeul, 2005; Evers-Vermeul & Sanders, 2008). The idea that Leo's use of connectives is influenced by the relative complexity of the connectives would explain Leo's early use of *und* (see Fig. 1), which is assumed to be the least complex connective, and his late use of *damit*, which is considered to be the most complex connective we tested in this study.

An alternative explanation for an early period of little parental influence in terms of audience design – the connective simply did not occur in the parental input and could consequently not influence Leo's connective use – was not proven to be correct. In this study we could not show any instances of audience design. If that had been the case, we would have expected a gradual increase in the number of connectives in the parental input – in line with Leo's increasing age and cognitive abilities. Furthermore, we would have expected the input patterns to differ per connective, since connectives are known to vary in their syntactic and conceptual complexity (Evers-Vermeul, 2005). Instead, we found stable frequencies of parental connective use as well as similar input patterns for the connectives under investigation.

Three types of short-term effects

For each connective, we found a period in which there are a large number of sudden deviations from the growth curve. Consequently, it seems as though the short-term effects of parental input are of particular importance during this period. However, the growth curves do not indicate what causes the differences in the intensity of the peaks and the dips in the developmental curves for the connectives under investigation. One explanation for the prominent short-term effects in these periods could be that Leo in a way imitates what he hears in the input. This would be in line with assumptions held by proponents of the usage-based approach, who stress the importance of imitation in the process of language acquisition (see Diessel, 2004; Tomasello, 2003). An example of this direct type of parental influence can be found in the example (1), presented earlier. However, as the example in (2) shows, not all instances of Leo's connective use can be interpreted in terms of direct imitation.

The steepness of the peaks and dips in the growth curves of the connectives could also be accounted for by the possible existence of context effects; some connectives occur more often in certain contexts than others. For example, the growth curve for *und* ('and') shows neither steep peaks nor deep dips. This can be explained by the fact that the use of this connective is not restricted by the conversational context. The growth curve for the connective *weil* ('because'), on the other hand, is characterized by a large number of negative deviations. These even occur after the age at which the probability of Leo using *weil* has reached the 100% mark and we can expect that Leo is capable of producing *weil* independently. This finding can be accounted for by the fact that *weil* is a context-dependent connective; just like the English 'because,' it is generally used to argue, to give an account for performing a certain speech act, or to describe causal relations in the real world (for a discussion of different types of causal relations in child language, see Spooren & Sanders, 2008). Situations in which *weil* often occurs are, for instance, in explanatory settings and in *warum–weil* ('why–because' sequences) sequences during play. Thus, certain contexts require the use of *weil* leading to a high frequency of *weil*, though it is questionable whether there are any contexts in which *weil* cannot occur. In-depth qualitative analysis will have to reveal whether we can identify context-effects as such.

In addition to imitation and context effects, a third type of short-term effects can be distinguished: parental echoes. As is shown in example (3), the parental use of a connective does not always precede the child's connective use, but sometimes follows it:

(3)	Leo:	Da . . . der Zug hält.	
	Leo:	Der Zug hält.	
	Mother:	Der Zug hält?	
	Leo:	**Weil** das Signal nach oben ist.	[2;06.12]
	Mother:	**Weil** das Signal nach oben ist.	
	Mother:	Aha, gut.	
	Leo:	There . . . the train stops.	
	Leo:	The train stops.	
	Mother:	The train stops?	
	Leo:	**Because** the signal is up (is red).	[2;06.12]
	Mother:	**Because** the signal is up.	
	Mother:	Ah, good.	

By repeating Leo's answer to the *warum*-question, his mother emphasizes her understanding of his answer, as well as the correctness of the use of *weil*. In other words, she reinforces Leo's use of *weil*. Apart from reinforcements, we also find that recasts follow Leo's connective use, though they do not seem to occur as often as reinforcements in our corpus. Recasts differ from reinforcements in that they repeat the utterance with corrections (see Chouinard & Clark, 2003). In a way, however, these recasts reinforce the production of connectives as well, since the correction is typically aimed at a different part of the connective clause. Both types of parental echo could lead to an increased frequency of a certain connective. However, they cannot explain the sudden deviations in the growth curves, because we do not expect reinforcement and recasts to occur often in one file, without occurring in the next. Qualitative research is needed to find further support for these ideas and to find to what extent these three types of short-term effects – direct imitations, context effects, and parental echoes – contribute to connective acquisition.

Long-term effects and mastery

The logistic regression shows that adding Long-term input as a predictor of Leo's connective use leads to a significantly better fit. In addition, we found that the long-term input diminishes the short-term effects in specific files: it flattens the peaks and dips in the growth curves. Hence, it seems likely that utterances produced by parents indeed have a lasting effect on linguistic representations (Abbot-Smith & Tomasello, 2006). For connective use this means that, even though the parental influence is small in the first period, the child is in fact taking in the input he or she receives during this period. Thus, even when there is no short-term effect in that the child immediately puts the input to use, we can claim that Leo is indeed sensitive to his parents' connective use from the beginning of his linguistic development.

The presence of long-term effects sheds new light on the operationalization of the notion 'acquisition.' Researchers differ greatly in the way they operationalize this notion. Some researchers of connective acquisition have tried to determine the emergence of connectives (see Bloom et al., 1980; Diessel, 2004; Evers-Vermeul, 2005), while others have focused on full mastery or on the developmental stages between emergence and full mastery (see Braunwald, 1997; Piaget, 1969). The presence of fluctuations in the probability of occurrence of connectives in the child output indicates that acquisition should be regarded as a long-term process and not in terms of 'acquired' or 'not acquired.' Characterizing children's use of a connective in terms of either 'acquired' or 'not acquired' would not account for the 'more or less' considerations that we have found in our growth curves. In line with van den Bergh, Herrlitz and Klein Gunnewiek (1999), we have shown that 'language development can be described more adequately by a continuous scale on which the mastery is allowed to vary from "no mastery at all" to "complete mastery"' (p. 25).

In addition, the methodology used in this study gives rise to a new operationalization of the notion 'mastery.' A child can be said to have 'mastered' a linguistic item when he or she is able to use it autonomously – both in production and comprehension – and use it in a variety of functions, contexts, and registers (see Berman, 1996; Braunwald, 1997). The data in this study allow us to talk about mastery in terms of independent productive use, but do not tell us anything about mastery in terms of comprehension, various functions, contexts or register. In the following, we restrict our discussion of mastery to include independent productive use only.

Previous studies have established full mastery on the basis of an arbitrary proportion of the parental target input (see Brown, Cazden, & Bellugi-Klima, 1969; de Villiers & de Villiers, 1973). However, this method can result in an incorrect 'yes/no'-idealization of the developmental data. Moreover, it does not take into account that the probability of occurrence of connectives can vary per recording. The growth curve analysis allows us to establish mastery on the basis of the smoothness of the growth curves toward the end of the curves. This smoothness implies that, even though the probability of the connective occurring is high, the influence of adult input is small. This indicates that Leo is capable of using these connectives on his own accord, without taking notice of his parents' use of the connective. Hence, mastery can be established in the final, relatively smooth part of the growth curve, where no effects of parental input can be found.[3]

This view sheds new light on the examples in (1) and (2), which were recorded during the final period, when Leo was aged 2;09.00 (1005 days old). In this stage of Leo's acquisition of *aber*, his parents do not show a great influence. Thus, we can conclude that his use of *aber* in (1) and (2) was not a form of direct imitation (i.e., a short-term effect), but most likely an example of mastery, a product of the long-term effect of parental input. Of course, Leo's independent use of *aber* does not mean that he has mastered all its possible uses: his production and comprehension of *aber* will still need to mature for it to encompass all varieties of the word.

Correlational evidence and audience design

In this study, we have been able to establish a correlation between the parental connective input and Leo's connective use. So far, we have given a one-way interpretation to these results: the parents influence Leo (A > B). However, given the correlational nature of our results, it is also possible to argue that Leo's output is in fact influencing the parental input (B > A) or that we are dealing with a tight interaction between parental input and child output (A <> B). The idea that the input provided by Leo's parents is guided by Leo's cognitive and linguistic capacities is a type of audience design. Our results provide two arguments against this idea. First, we have shown that the proportion of connectives in the parental input remains stable over time (cf. the scatterplot in Fig. 5). This implies that parents do not adapt their connective frequency to the increasing abilities of the child. Second, we should note that the parents were using the connectives before the child first started to use them. This implies that the parent is teaching the child and not vice versa. In any case, the results from the current study could not show any instances of audience design at the quantitative level. These findings differ from those reported by McCabe and Peterson (1997) and Stevens et al. (1998), but are supported by the study by Rome-Flanders et al. (1995).

Proponents of audience design could claim that it is not the quantity of the input that is changing, but rather the quality of the connective input. The amount of parental connectives may remain stable, but parents could alter the type of connective utterances they use. For example, they could gradually increase the number of cause–consequence utterances while decreasing the number of consequence–cause clauses. Or they could – seeing that their child starts to reason more often – increase the number of claim–argument relations with respect to consequence–cause relations. Similarly, the type of interaction between parent and child could change. Instead of producing connective utterances that the child imitates, the parent could start producing reinforcements, imitating what the child says. In order to see whether this is an accurate explanation, one would have to perform a qualitative analysis of the data. Such a qualitative analysis might also provide cues for resolving the conflicting evidence in this area.

CONCLUSION

In this study, we have investigated the acquisition of discourse connectives. An ongoing debate in current studies of language acquisition centers around the question whether the acquisition of linguistic elements is determined by inherent complexity of linguistic and conceptual structures, or also by the language input children receive – as increasingly influential usage-based accounts propose. We have seen how both input-based and cognitive complexity explanations might hold for the acquisition of connectives, too. Several researchers have shown how English and Dutch children acquire connectives expressing more complex relations like negative (*but*) or causal (*because*) only after they have mastered less complex ones like additives (*and*) and temporals (*and then*) (Bloom et al., 1980; Evers-Vermeul & Sanders, 2008; Spooren & Sanders, 2008).

Although parental input is thought to be a serious alternative to cognitive complexity, current methods have proved inadequate to investigate its contribution to the course of acquisition. The field seemed to be in need of a method to study the role of both complexity and parental input as factors influencing the language development of a child. For this reason, we have developed an innovative method that would enable us to do this and simultaneously shed new light on parental input theory. We can now conclude that the application of growth curve analysis to child language data, as proposed in this article, is a successful method. It enables us to take into account both age – which is closely related to increasing complexity – and long- and short-term effects of parental input. In fact, the method produces a relation between input and output that is interpretable. Moreover, the percentage of recordings that are correctly predicted to contain a connective is relatively high for each of the connectives under investigation. This can lead us to conclude that ratio and interval (continuous) data might better allow researchers to disentangle and model the short- and long-term effects of parental input than categorical data would.

We have demonstrated that – in addition to the factor Age – frequency of use in the adult input significantly influences the acquisition of connectives. This influence is not continuous, but undergoes periods of little to no influence and periods of substantial influence. Future research should increase the generalizability of this study by including the analysis of more child–parent pairs. An additional advantage of the simultaneous analysis of data of several child–parent pairs is that it would enable us to determine the relative influence of the three predictors: Age, Short-term input, and Long-term input. This would refine the picture obtained in this study, in which we could establish the significance of the contribution of the three factors, but not the weight of their contribution. Also, it would be interesting to see whether parents differ from one another with regards to adapting their connective use to their child's needs. Evers-Vermeul (2005) has shown that this type of research is promising; she found that one and the same mother differentiated her frequency of use of the Dutch connective *toen* 'then/when', depending on which daughter she spoke to. With regards to the current study, this type of research would also reveal whether the input produced by the research assistant differs from that of the parents or whether children with and without siblings show different acquisition patterns.

A second way to improve the generalizability of our study is to take polysemy into account. The results presented in this study were based on counts of lexical forms only. Hence, part of the words that were labeled as connectives could have been discourse markers or have other non-connective uses. We expect that the polysemic uses of *und* and *damit* will have led to larger usage frequencies. In future research, this could be corrected for by differentiating between the possible uses of a connective form.

The current study focused on the frequency effects of the adult input. However, to get a more fine-grained picture of the influence of adult input, additional qualitative research should be carried out on the data (see Tomasello, 2000, p. 211). This might help us in distinguishing between three possible variants of short-term effects: direct imitation, reinforcement, and influence of the conversational context. Furthermore, only qualitative analyses can show whether children make mistakes in the use of certain connectives, or whether they use these connectives in a restricted range of all the meanings they can express. Finally, a qualitative analysis may refine the picture we obtained for audience design; it allows us to find out whether parents adapt their speech to the cognitive abilities of their children in a qualitative way.

To conclude, we have shown that a proper explanation for connective development in child language needs to take into account: (1) effects of increasing age and, hence, the child's increasing cognitive abilities; (2) short-term frequency effects of the adult input; as well as (3) long-term frequency effects of the adult input. Future research will have to reveal the relative contribution of each of these factors.

We trust that the growth curve analysis developed in this study will have a role to play in this kind of follow-up research, because it enables researchers to operationalize parental input hypotheses of language acquisition. The use of this method will be a helpful tool in answering crucial questions in the fascinating field of language acquisition, such as 'what is the role of complexity versus input?' We will continue this type of research in the field of discourse connectives, but we expect that growth curve analysis will also be a successful method to study the development of many other linguistic elements.

ACKNOWLEDGEMENTS

The first and the third author's research was enabled by the Netherlands Organization for Scientific research, through NWO-Vici-grant 277-70-003, awarded to Ted Sanders. Furthermore, we would like to thank the Max Planck Institute in Leipzig, Germany, and Heike Behrens for sharing their data with us. Finally, we would like to thank Kevin Durkin and the two anonymous reviewers for their valuable comments and suggestions.

NOTES

1. Age to the power zero is 1. This generates a vector of ones that is traditionally known as the constant. Age^1 introduces a linear relation between age and connective use into the model; Age^2 introduces an acceleration or deceleration in the growth, and Age^3 and so on modify the slope of the growth curve even further.
2. An exception can be found in the connective *und*, which is the only connective that lacks an early period of little parental influence. It is likely that such a period took place before Leo was 714 days old and was consequently not captured on record.
3. Growth curve analysis is not suitable for infrequent linguistic items for which 'growth' is difficult to establish, such as the word 'thunder' or 'pretty,' but is suitable for grammatical phenomena (e.g., plural –s), function words (e.g., connectives), and word groups (e.g., adjectives). Consequently, this method for establishing mastery does not apply to infrequent items.

REFERENCES

Abbot-Smith, K., & Tomasello, M. (2006). Exemplar-learning and schematization in a usage-based account of syntactic acquisition. *The Linguistic Review, 23*, 275–290.

Aksu-Koç, A. (1998). The role of input vs. universal predispositions in the emergence of tense-aspect morphology: Evidence from Turkish. *First Language, 18*, 255–280.

Behrens, H. (2006). The input–output relationship in first language acquisition. *Language and Cognitive Processes, 21*, 2–24.

Berman, R. A. (1996). Form and function in developing narrative abilities. In D. I. Slobin, J. Gerhardt, A. Kyratzis, & J. Guo (Eds.), *Social interaction, social context, and language* (pp. 343–367). Mahwah, NJ: Lawrence Erlbaum.

Bloom, L., Lahey, M., Hood, L., Lifter, K., & Fiess, K. (1980). Complex sentences: Acquisition of syntactic connectives and the semantic relations they encode. *Journal of Child Language*, *7*, 235–261. [Reprinted in L. Bloom (1991). *Language development from two to three* (pp. 261–289). Cambridge, UK: Cambridge University Press.]

Borovsky, A., & Elman, J. (2006). Language input and semantic categories: A relation between cognition and early word learning. *Journal of Child Language, 33*, 759–790.

Braunwald, S. R. (1997). The development of *because* and *so*: Connecting language, thought, and social understanding. In J. Costermans & M. Fayol (Eds.), *Processing interclausal relationships: Studies in the production and comprehension of text* (pp. 121–137). Mahwah, NJ: Lawrence Erlbaum.

Brown, R. (1973). *A first language: The early stages*. Cambridge, MA: Harvard University Press.

Brown, R., Cazden, C., & Bellugi-Klima, U. (1969). The child's grammar from I to III. In J. P. Hill (Ed.), *Minnesota symposia on child psychology* (Vol. 2, pp. 28–73). Minneapolis: University of Minnesota Press.

Chouinard, M. M., & Clark, E. V. (2003). Adult reformulations of child errors as negative evidence. *Journal of Child Language*, *30*, 637–669.

Clark, E. V. (2002). *First language acquisition*. Cambridge, UK: Cambridge University Press.

Clark, H. H., & Clark, E. V. (1977). *Psychology and language: An introduction to psycholinguistics*. New York: Harcourt Brace Jovanovich.

Clark, H. H., & Murphy, G. L. (1983). Audience design in meaning and reference. In J. F. LeNy & W. Kintsch (Eds.), *Language and comprehension* (pp. 287–299). Amsterdam: North-Holland.

De Villiers, J. G., & de Villiers, P. A. (1973). A cross-sectional study of the acquisition of grammatical morphemes in child speech. *Journal of Psycholinguistic Research, 2*, 267–278.

Diessel, H. (2004). *The acquisition of complex sentences*. Cambridge, UK: Cambridge University Press.

Evers-Vermeul, J. (2005). The development of Dutch connectives: Change and acquisition as windows on form–function relations. Unpublished Ph.D. Dissertation UIL-OTS, Utrecht University, Utrecht, LOT.

Evers-Vermeul, J., & Sanders, T. (2008). The emergence of Dutch and English connectives: How cumulative complexity explains the order of acquisition. *Journal of Child Language*.

Fienberg, S. E. (1977). *The analysis of cross-classified categorical data.* Cambridge, MA: MIT Press.

Francis, D., Shaywitz, S., Stuebing, K., Fletcher, J., & Shaywitz, B. (1996). Developmental lag vs. deficit models of reading disability: A longitudinal, individual growth curve analysis. *Journal of Educational Psychology*, *1*, 3–17.

Goldstein, H. (1995). *Multilevel statistical models* (2nd ed.). London: Edward Arnold.

Hadley, P. A., & Holt, J. K. (2006). Individual differences in the onset of tense marking: A growth curve analysis. *Journal of Speech, Language, and Hearing Research*, *49*, 984–1000.

Huttenlocher, J., Haight, W., Bryk, A., Seltzer, M., & Lyons, T. (1991). Early vocabulary growth: Relation to language input and gender. *Developmental Psychology*, *27*, 236–248.

Kemper, S., Rice, K., & Chen, Y.-J. (1995). Complexity metrics and growth curves for measuring grammatical development from five to ten. *First Language*, *15*, 151–166.

Kidd, E., Lieven, E. V. M., & Tomasello, M. (2006). Examining the role of lexical frequency in the acquisition and processing of sentential complements. *Cognitive Development*, *21*, 93–107.

Krasinski, E. (1995). The development of past marking in a bilingual child and the punctual–nonpunctual distinction. *First Language*, *15*, 239–276.

MacWhinney, B. (2000). *The CHILDES project: Tools for analyzing talk* (3rd ed.) Mahwah, NJ: Lawrence Erlbaum.

McCabe, A., & Peterson, C. (1997). Meaningful 'mistakes': The systematicity of children's connectives in narrative discourse and the social origins of this usage about the past. In J. Costermans & M. Fayol (Eds.), *Processing interclausal relationships: Studies in the production and comprehension of text* (pp. 139–154). Hillsdale, NJ: Lawrence Erlbaum.

Matthews, D., Lieven, E. V. M., Theakston, A. L., & Tomasello, M. (2005). The role of frequency in the acquisition of English word order. *Cognitive Development, 20*, 121–136.
Menars, S. (2002). *Applied logistic regression analysis* (2nd ed.). Series: Quantative Applications in the Social Sciences, 106. Thousand Oaks, CA: Sage.
Naigles, L. R., & Hoff-Ginsberg, E. (1998). Why are some verbs learned before other verbs? Effects of input frequency and structure on children's early verb use. *Journal of Child Language*, *25*, 95–120.
Piaget, J. (1969). *Judgement and reasoning in the child*. London: Routledge & Kegan Paul. [Translation of J. Piaget (1924). *Le jugement et le raisonnement chez l'enfant*. Neuchatel: Delachaux et Niestlé.]
Rescorla, L., Mirak, J., & Singh, L. (2000). Vocabulary growth in late talkers: Lexical development from 2;0 to 3;0*. Journal of Child Language, 27*, 293–311.
Rome-Flanders, T., Cronk, C., & Gourde, C. (1995). Maternal scaffolding in mother–infant games and its relationship to language development: A longitudinal study. *First Language, 15*, 339–355.
Sanders, T. J. M., Spooren, W. P. M., & Noordman, L. G. M. (1992). Toward a taxonomy of coherence relations. *Discourse Processes*, *15*, 1–35.
Sandhofer, C. M., Smith, L. B., & Luo, J. (2000). Counting nouns and verbs in the input: Differential frequencies, different kinds of learning? *Journal of Child Language*, *27*, 561–585.
Shirai, Y. (1994). On the overgeneralization of progressive marking on stative verbs: Bioprogram or input? *First Language*, *14*, 67–82.
Spooren, W., & Sanders, T. (2008). The acquisition of coherence relations: On cognitive complexity in discourse. *Journal of Pragmatics*, *40*, 2003–2036.
Stevens, E., Blake, J., Vitale, G., & MacDonald, S. (1998). Mother–infant object involvement at 9 to 15 months: Relation to infant cognition and early vocabulary. *First Language*, *18*, 203–222.
Tomasello, M. (2000). Do young children have adult syntactic competence? *Cognition, 74,* 209–253.
Tomasello, M. (2003). *Constructing a language: A usage-based theory of language acquisition*. Cambridge, MA: Harvard University Press.
Tomasello, M. (2006). Acquiring linguistic constructions. In D. Kuhn & R. Siegler (Eds.), *Handbook of child psychology* (pp. 255–298). New York: Wiley.
Valian, V., & Casey, L. (2003). Young children's acquisition of wh-questions: The role of structured input. *Journal of Child Language*, *30*, 117–143.
van den Bergh, H., Herrlitz, W., & Klein Gunnewiek, L. (1999). Cognitive co-ordination in foreign language learning: A sketch of a research paradigm. In J. Don & T. Sanders (Eds.), *Utrecht Institute of Linguistics OTS yearbook 1998–1999* (pp. 20–33). Utrecht: Utrecht Institute of Linguistics OTS.
van den Bergh, H., & Rijlaarsdam, G. (1996). The dynamics of composing: Modeling writing process data. In C. M. Levy & S. Randall (Eds.), *The science of writing: Theories, methods and individual differences and applications* (pp. 207–232). Mahwah, NJ: Lawrence Erlbaum.
Wijnen, F., Kempen, M., & Gillis, S. (2001). Root infinitives in Dutch early child language: An effect of input? *Journal of Child Language*, *28*, 629–660.

APPENDIX

Table A1 *Aber:* measure of error per model

Predictors in model	*–2 Log-likelihood*	χ^2	*d.f.*	*p*
1. Age	311302.1			
2. Age, Short-term input	307727.5	1–2 = 3573.6	3	< 0.001
3. Age, Short-term input, Long-term input	306966.6	2–3 = 761.0	3	< 0.001

Table A2 *Damit:* measure of error per model

Predictors in model	*–2 Log-likelihood*	χ^2	*d.f.*	*p*
1. Age	263848.7			
2. Age, Short-term input	248578.4	1–2 = 15270.3	3	< 0.001
3. Age, Short-term input, Long-term input	247917.3	2–3 = 661.1	3	< 0.001

Table A3 *Und:* measure of error per model

Predictors in model	*–2 Log-likelihood*	χ^2	*d.f.*	*p*
1. Age	46890.9			
2. Age, Short-term input	46804.8	1–2 = 86.2	3	< 0.001
3. Age, Short-term input, Long-term input	46455.1	2–3 = 349.6	3	< 0.001

Table A4 *Weil:* measure of error per model

Predictors in model	*–2 Log-likelihood*	χ^2	*d.f.*	*p*
1. Age	153762.2			
2. Age, Short-term input	145146.1	1–2 = 8616.1	3	< 0.001
3. Age, Short-term input, Long-term input	144901.2	2–3 = 244.9	3	< 0.001

Table A5 *Wenn:* measure of error per model

Predictors in model	*–2 Log-likelihood*	χ^2	*d.f.*	*p*
1. Age	298752.8			
2. Age, Short-term input	295834.3	1–2 = 2918.4	3	< 0.001
3. Age, Short-term input, Long-term input	293622.0	2–3 = 2212.3	3	< 0.001

ADDRESS FOR CORRESPONDENCE

Rosie van Veen
Utrecht Institute of Linguistics, Trans 10, NL – 3512 JK Utrecht, The Netherlands
E: r.vanveen@uu.nl