

The influence of input on connective acquisition: a growth curve analysis of English *because* and German *weil**

ROSIE VAN VEEN, JACQUELINE EVERS-VERMEUL,
TED SANDERS

Utrecht Institute of Linguistics OTS – Utrecht University

AND

HUUB VAN DEN BERGH

*Utrecht Institute of Linguistics OTS – Utrecht University; Graduate School
of Teaching and Learning – Amsterdam University*

(Received 7 September 2010 – Revised 20 March 2011 – Accepted 2 August 2012 –
First published online 30 November 2012)

ABSTRACT

The current study used growth curve analysis to study the role of input during the acquisition of the English causal connective *because* and its German counterpart *weil*. The corpora of five German and five English children and their adult caretakers (age range 0;10–4;3) were analyzed for the amount as well as for the type of connective use – imitated, elicited, and independent. The growth curves showed that children's elicited use developed faster than their independent use; imitations were rare. Adult connective input was not found to function as a scaffold of children's connective use. Rather, the adult *why/warum*-questions played an important role in the acquisition of *because* and *weil*. In turn, children also used *why/warum*-questions to elicit causal responses from their caretakers, which shows that children were responsible for a great part of their own input.

INTRODUCTION

In recent years, the study of first language development has benefited from the increasing availability of longitudinal data. As a result, a growing

[*] The first, second, and 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 Rosemarie Rigol, Heike Behrens, and the Max-Planck-Institute for Evolutionary Anthropology, for sharing their data with us. Address for correspondence: Rosie van Veen, Utrecht Institute of Linguistics, Trans 10, NL-3512 JK Utrecht, The Netherlands. tel. +31 30 253 8692; fax +31 30 253 6000; e-mail: rosievanveen@hotmail.com

number of studies use spontaneous language corpora to assess children's language development. With the growing interest in dense corpora comes the need for sophisticated methodologies to analyze large amounts of data, such as growth curve analysis (Goldstein, 1999). By applying growth curve analysis to longitudinal child language data, it is possible to statistically model the development of a certain linguistic item (e.g., particular grammatical phenomena, function words, or word classes) over time. This statistical modelling produces parameters (slope, acceleration, etc.) that describe the growth trajectory of the child's use of a certain linguistic item. In modelling this growth trajectory, growth curve analysis can take into account the effects of multiple (hypothesized) predictors of language development, such as parental input, the child's mean length of utterance, or vocabulary size (Menard, 2002).

An important advantage of growth curve analysis is that it takes into account the full effect of age as an independent continuous variable. By using age as a continuous variable, it is possible to consider each data point (i.e., each recording), which leads to a detailed account of children's linguistic development and/or the input they receive over time. Many studies, however, do not take such a detailed view of language development. For example, several studies focusing on the role of input on language acquisition – the interest of the current study – do not employ age as a continuous variable, but relate the total input frequency to language proficiency at a specific age (e.g., Kirjavainen, Theakston & Lieven, 2009; Naigles & Hoff-Ginsberg, 1998). Although these results may be very informative, the method has a drawback: it is impossible to track any changes in input over time. Other studies investigate frequencies of use per developmental 'stage' (e.g., Sandhofer, Smith & Luo, 2000; Theakston & Lieven, 2008; Wijnen, Kempen & Gillis, 2001). This method allows for a general overview of language development, but is restricted, because dividing up data according to developmental stage (e.g., mean length of utterance or age in months) does not allow for a linear overview of the data (cf. concerns raised by Stoll & Gries, 2009).

In this paper, we aim to use growth curve analysis to model children's language development over time. Indeed, as children grow older we expect an increase in their use of certain linguistic phenomena. Therefore, it makes sense to use time – measured as the increase in the child's age – as a predictor in growth curve analysis (for the use of time as a predictor, see also Goldstein, 1979; Van den Bergh & Rijlaarsdam, 1996).¹ Since children's

[1] Many studies use mean length of utterance (MLU) instead of age to indicate an increase in language development (Brown, 1973; de Villiers & de Villiers, 1973). Since many studies have found a significant linear relationship between age and MLU in typically developing children up to age five (Klee, Schaffer, May, Membrino & Mougey, 1989;

ability to deal with cognitively more complex words and structures is associated with an increase in children's age, we use the predictor age as an indicator of children's cognitive abilities.²

In the current study, we will apply growth curve analysis to the development of causal connectives, such as *because*. Connectives are interesting candidates for growth curve analysis, because input has been found to influence their course of development (Diessel, 2004). By relating the age of emergence of the various types of conjoined clauses to their overall frequency in the input, Diessel (2004) found that parental input seems to play a role in connective acquisition: 'the more frequent a specific conjoined clause appears in the mothers' data, the earlier it emerges in the children's speech' (p. 172). His study, however, does not track any changes in input over time, because language proficiency at a specific age is related to the total input frequency.

In order to track these changes over time, Van Veen, Evers-Vermeul, Sanders, and Van den Bergh (2009) used growth curve analysis to investigate the effects of parental input on the acquisition of the German discourse connectives *aber* 'but', *damit* 'so that', *und* 'and', *weil* 'because', and *wenn* 'when'. The predictor parental input was operationalized by distinguishing between short-term input and long-term input. Short-term input was defined as the parental input in the space of one recording and long-term input as the cumulative parental input the child has been exposed to over a longer period of time. Results showed that both factors have a significant effect on the acquisition of German connectives. In addition, they found that the influence of parental input is not continuous, but undergoes periods of little to no influence and periods of substantial influence. These findings reinforce the idea that a longitudinal perspective is necessary for a proper explanation of language acquisition.

However, the finding that the influence of input is not continuous raises the question of how the connective input children receive is actually distributed over time. We see three possible distribution patterns, which are presented graphically in the fictional examples in Figures 1a–c. A first possibility is that the input remains constant over time (see the horizontal line in Figure 1a), indicating that adults do not adapt their input to children's increasing ability to produce connectives (which is represented as

Miller & Chapman, 1981), we felt it was unnecessary to include both MLU and age as predictors.

[2] Of course, the relationship between age and cognitive abilities is not truly linear. We merely intend to use age as an indicator – not as a strict measure – of growing cognitive abilities. Ideally, children's cognitive development should be established on the basis of independent tests at different ages, but unfortunately such data were not collected for the longitudinal corpora under investigation.

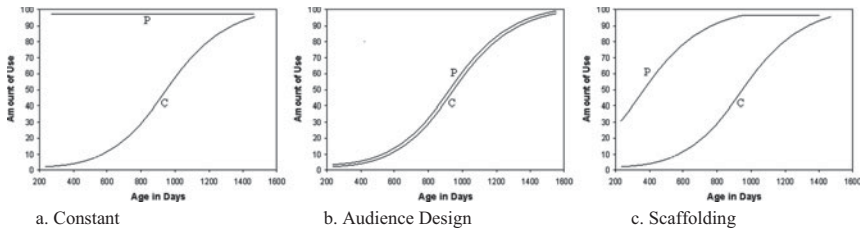


Fig. 1. Fictional growth curves of input (P) in relation to children's (C) connective development.

an s-shaped curve and is the same in all three figures). A second possibility is that adults adjust the quantity and quality of their connective use to the cognitive and linguistic capacities of the child. This would be a type of audience design (Clark & Murphy, 1982), which entails that adults will not use a connective if their child is unable to comprehend what it means. Such a tight relationship between adults and children leads to the audience design represented in Figure 1b: the adult connective use increases at the same rate as the children's development. A third and related possibility is that of scaffolding (Wood, Bruner & Ross, 1976), which refers to the way adults gradually increase the complexity of their language to encourage the child to produce slightly more complex words and constructions. Adults show this increase in the period PRECEDING the child's ability to produce these complex words and constructions. Scaffolding is illustrated in Figure 1c: the input increases ahead of the children's development. Like audience design, scaffolding emphasizes the adjusting of the type and amount of input to the child's ability. However, in terms of development, scaffolding can make the difference between the level a child is capable of reaching independently and the level a child can achieve under guidance of an adult (compare the similar idea behind Vygotsky's (1978) 'zone of proximal development').

In the current study, we will investigate two ways in which input can play a role in the acquisition of English *because* and its German counterpart *weil*: the AMOUNT of connective input, and the TYPE of connective input. By investigating the amount of the input, we aim to determine whether the connective input is constant over time, or whether it increases in parallel to the child's connective use (audience design) or ahead of the child's connective use (scaffolding). In their case study, Van Veen *et al.* (2009) plotted the relative input frequency of the connectives per recording and, on the basis of these plots, conclude that the connective input remains relatively stable over time, which indicates that parents use connectives, even when their child does not. However, they did not test the significance

of increases or fluctuations in the input frequencies. To solve this issue, the current study will use growth curve analysis to model the connective use of the children over time, as well as that of their ambient language.

Apart from adjusting their amount of input, adults may also scaffold a child's connective use by adjusting their type of connective use to the child's increasing ability. For the current study, we distinguish between three different types of connective use: imitations, elicited use, and independent use. First, adults may imitate their children by way of reinforcing the child's connective usage. This is a type of indirect feedback that often takes the shape of reformulations and that reinforces the child's connective use. For example, the father's use of *weil* in (1) reformulates Simone's previous utterance by correcting Simone's case marking of *keine Zähne*. Research by Chouinard and Clark (2003) shows that this type of feedback is successful: children pay attention to the information in adult reformulations in 10% to 50% of the cases (see also Farrar, 1992; Saxton, 2000). Note that a high number of adult reformulations will result in a strong correlation between adult input and child output (compare the findings by Van Veen *et al.*, 2009).

(1) (Simone 2;06.23)

- FATHER: Kann denn der Tobias schon Fisch essen?
'Can Tobias eat fish yet?'
- SIMONE: Nein.
'No.'
- FATHER: Warum kann der noch nicht Fisch essen?
'Why can't he eat fish yet?'
- SIMONE: **Weil** der noch kein Zähne hat.
'Because it does not have any teeth yet.'
- FATHER: **Weil** er noch keine Zähne hat.
'Because he does not have any teeth yet.'

Second, children may elicit adult connective use. This type of use generally occurs in response to a *how come* or a *why*-question (in German a *wieso* or *warum*-question). Adults are known to scaffold children's causal language by asking *why*-questions, and from age 2;2 children start to ask *why*-questions increasingly often (Hood, Bloom & Brainerd, 1979; McCabe & Peterson, 1997). For example, Sarah's mother's use of *weil* in (2) was elicited by Sarah's *why*-question.

(2) (Sarah 2;07.05/Situation: looking at Tony the pony)

- SARAH: Where fingers?
- MOTHER: Your fingers are right here.
- SARAH: Tony Pony fingers.
- MOTHER: Tony, Tony don't have any fingers.

- SARAH: Why?
 MOTHER: **Because** he's only a horse.
 SARAH: Horse?
 MOTHER: Yeah, horses don't have fingers.

Third, adults may provide positive evidence by using connectives in child-directed utterances that are not imitations or responses to *why*-questions, but that are produced spontaneously. We label this 'independent' use. By combining the qualitative analysis of the input with growth curve analysis, we can assess the timing, and thus the extent, of scaffolding, for each type of connective input. In addition, we will assess the children's usage of the three types of *weil* and *because*. This way, we can show how the different types of connective input are linked to the child's increase of independent and elicited connective use, presenting a clearer picture of the relation between adult and child connective use.

To summarize, the current study uses growth curve analysis to study the role of adult input in connective acquisition. In particular, we will investigate the development of English *because* and German *weil* in the longitudinal corpora of five German and five English children and their caretakers. By distinguishing between various types of uses of *because* and *weil* – imitated, elicited, and independent – we present a detailed view of the nature of the connectives adults and children produce.

By adding a cross-linguistic perspective, we can investigate whether the influence of input on connective acquisition extends to languages other than German (the language studied in Van Veen *et al.*, 2009), such as English. Cross-linguistically, *because* and *weil* are conceptually compatible: in both languages they are the most frequent connective to denote backward causal relations (consequence–cause) (Pit, 2003; Sanders, 2005).³ Thus, in the current study, cross-linguistic differences between the results for *because* and *weil* cannot be attributed to conceptual differences. However, developmental differences may be related to differences at the level of lexical competition: *because* occurs in related causal uses such as the non-clausal *because of* and the bromide *just because*, whereas German speakers may use various lexical items for these expressions: *wegen* and *darum*. Thus, the lexical uniformity found in English may lead to *because* developing earlier than German *weil*.

Because and *weil* also differ in terms of their accompanying word order. Syntactically, English typically requires an SVO word order (subject–verb–object) for its main clauses as well as for its subordinate

[3] We should point out that the German language distinguishes between *weil* and *denn* 'because'. *Denn* is also used as a backward causal, but since it is rarely found to be used in contemporary speech (Günthner, 1993) this is not an issue for our current interests.

clauses introduced with *because*, as in (3). German requires a verb-second word order for its main clauses, but its subordinator *weil* introduces a proposition with the finite verb at the end of the clause, the word order typical of German subordinating clauses, see (4).⁴ The cumulative complexity approach put forward by Evers-Vermeul and Sanders (2009) argues that clauses with a verb-final word order are cognitively more complex than clauses with the verb-second word order, because it forces the reader to impose a hierarchical ordering upon the clauses. This difference in complexity means that we may find that *because* appears at an earlier age than *weil*.

(3) (Nina 3;01.06)

NONNA: Nina, what did you have for supper last night?

NINA: I can't talk to you, **because** I have this in my mouth.

(4) (Simone 3;03.25/Situation: talking with her father Max)

FATHER: Maxe raucht mal noch mal ne Zigarette. Ne?

'Max smokes another cigarette. Okay?'

SIMONE: Ich will meine Schnullers, **weil** du ne Pfeife hast.

I want my dummies, because you a pipe have.

'I want my dummies, because you have a pipe.'

THE AMOUNT OF CONNECTIVE USE

METHOD

Participants and corpora

We analyzed the corpora of the spontaneous speech data of five German and five English children and caretakers (age range 0;10 to 4;3). Table 1 introduces the respective age range and number of files (with ages in years;months.days) for each child. All data were downloaded from the CHILDES database (MacWhinney, 2000).

The children were selected on the basis of their age range and the density of their data. In order to maintain a heterogeneous corpus in terms of sampling density, as well as allowing for a detailed view of development, the recordings in the corpora were no more than three weeks apart. Furthermore, the corpus ideally included recordings between age 2;2 and 3;0, the period during which we expect *because* and *weil* to be acquired (based on English data from Bloom, Lahey, Hood, Lifter & Fiess, 1980). In order to capture development after first use and to include the development

[4] German also allows for a verb second word order for clauses introduced with *weil* (Keller, 1995). However, this usage was very infrequent in our corpora. Of the total number of *weil* included for analysis a verb second word order was observed in eight per cent of the cases (4% of the children's cases and 10% of the adults' cases).

TABLE 1. *Overview of the corpora with age range and number of files per child*

Name	Corpus	Age range*	# Files
English children			
Abe	Kuczaj (1976)	2;04·24–4;02·19	167
Naomi	Sachs (1983)	1;02·29–3;08·19	90
Nina	Suppes (1974)	1;11·16–3;03·21	52
Sarah	Brown (1973)	2;03·05–4;02·16	96
Shem	Clark (1978)	2;02·16–3;02·02	40
German children			
Caroline	Von Stutterheim	0;10·01–2;10·26	232
Leo	Behrens (2006)	1;11·12–4;02·16	462
Pauline	Rigol	1;10·09–4;01·28	46
Sebastian	Rigol	2;01·12–4;01·27	43
Simone	Miller (1979)	1;09·11–4;00·06	74

of possible late-bloomers, we decided to focus our analysis on the period up to age 4;3. For this reason, any available recordings after age 4;3 were not included in the study. An exception was made for Caroline, for whom we excluded the final three transcripts because of the large interval between these recordings.

Measures

Once the corpus-data were selected, the frequencies of *because* or *weil* per file were computed by using Computerized Language Analysis (CLAN) (MacWhinney, 2000). In order to ensure that the dataset consisted of connective uses, only tokens where *because* and *weil* function as connectives, that is when they were part of a coherence relation (Sanders, Spooren & Noordman, 1992), were included. This meant that the use of bare *because* (without propositional content), or *because* occurring in an unintelligible utterance (often occurring with *xxx*) were excluded from analysis. Self-repetitions and corrections as well as false starts were also excluded. This resulted in the exclusion of 20 to 50% of the children's fragments, and of 3 to 18% of the parental data. Table 2 provides an overview of the total number of *because* or *weil* per corpus as well as the number of fragments included for analysis after filtering for non-connective uses.

Lastly, we used CLAN to compute the word count per file – with a separate count for adults and children – so that we could control for the length of the recording. This is important because the number of words spoken in a recording can influence the frequency with which the connectives appear in the transcript. The analysis of the adult frequencies was carried out for all recorded utterances, including those that were not child-addressed (as they are still part of the child's input). In total, only

TABLE 2. *Overview of the corpora with total number of because/weil and number of because/weil included for analysis (with % of total)*

Name	Age range	# Files	Child <i>because/weil</i>		Adult <i>because/weil</i>	
			Total	Included	Total	Included
English children						
Abe	2;04.24-4;02.19	167	584	469 (80%)	291	266 (91%)
Naomi	1;02.29-3;08.19	90	114	80 (70%)	61	50 (82%)
Nina	1;11.16-3;03.21	52	421	303 (72%)	249	242 (97%)
Sarah	2;03.05-4;02.16	96	82	45 (55%)	185	170 (92%)
Shem	2;02.16-3;02.02	40	439	319 (73%)	335	315 (94%)
German children						
Caroline	0;10.01-2;10.26	232	170	83 (49%)	331	309 (93%)
Leo	1;11.12-4;02.16	462	1224	852 (70%)	1301	1218 (94%)
Pauline	1;10.09-4;01.28	46	106	63 (59%)	146	124 (85%)
Sebastian	2;01.12-4;01.27	43	70	40 (57%)	212	200 (94%)
Simone	1;09.11-4;00.06	74	580	305 (53%)	348	331 (95%)

66 instances (2%) of *because* and *weil* were found not to be addressed to the child, so the effects of including (or excluding) them from analyses are negligible.

Data analysis

We performed growth curve analyses on the final dataset by means of a multilevel logistic regression (Goldstein, 1999), using the statistical software MLN. When performing growth curve analysis, it is important to realize that developmental data are inherently a form of nested data. This means that the data are nested within children – who vary in their age – and between children – who differ in their linguistic skills or socioeconomic status. As a result, there may be correlations between numerous features of the data, all of which can be taken into account by a multilevel approach to the data (Quené & Van den Bergh, 2004, 2008). In addition, a multilevel approach takes into account missing data and/or lack of fixed occasions. Thus, it allows individuals to vary in the number of occasions as well as in the timing of occasions (e.g., spontaneous language recordings taken at different ages and intervals). So, by taking a multilevel approach to growth curve analysis, we ensure a good assessment of children's individual development over time.

The fact that we used a logistic regression means that we characterized our dependent variable (the use of *because* and *weil*) as dichotomous: absence versus presence of the connective (see Menard, 2002). As a result, we used growth curve analysis to model the probability that a connective

occurred in a recording (yes/no). Hence, the analysis does not provide any details about the number of connectives that occur in a recording. By characterizing our data as dichotomous, we circumvent the question as to how often a connective should ‘normally’ be used. Since discourse connectives are optional – it is possible to express causal relations without marking them with *because* or *weil* – we cannot identify any situations in which a connective is compulsory (as opposed to many obligatory items, such as a plural marker). As a result, the total connective frequency will vary due to context differences or individual preferences and not necessarily due to linguistic competence. A dichotomous approach, merely establishing whether a connective was used, limits these variation effects. By reducing the total connective frequency to absence or presence of the connective in each recording, we also avoid the related issue as to how often children should use *because* and *weil* in order to establish mastery of the connective.

We used growth curve analysis to model the probability of occurrence of *because* or *weil* for English children, English adults, German children, and German adults in a single model. This way, we can make a direct comparison between the connective development for English and German. We did this as a function of the predictor Age (expressed in days), which was used to establish growth over time, and the covariates Child Word Count (Child WC) and Adult Word Count (Adult WC), which control for effects due to the length of the recording; we also took into account any interactions and random effects (for children and adults).⁵ In order to end up with a model that best fits (represents) the observed connective use, extra variables were only included when the previously added variables made a significant contribution to the model. Note that the intercept was not set at age zero. Rather, as a result of the centring of age around the grand mean, the intercept was set at 1,077 days (2;11.19). For ease of interpretation, age is transposed back to its original values in the figures.

We used a polynomial model with Age¹, Age², Age³ as predictors. Significant effects of Age¹ (with a positive parameter estimate) indicate a linear increase and are thus a measure of the overall slope of the growth curve. Significant effects of Age² denote a single change in the direction of the curve (a parabola; from a rise to a fall, or from a fall to a rise). Significant effects of Age³ indicate a cubic increase, and thus denote two changes in the direction of the curve (an s-shape; e.g., from rise to fall to rise). Note that extra powers of Age were only added when the previously

[5] We also explored the effects of MLU as a predictor of connective development. However, MLU did not make a significant contribution to the model in Table 3 (parameter estimate = 0.00018; se = 0.00011). So, with respect to age and word count, MLU does not provide an extra explanation for connective acquisition. We can also rule out an effect of collinearity, because MLU was not highly correlated with age ($r = .05$; $p < .001$) nor with word count ($r = .47$; $p < .001$).

TABLE 3. *Parameter estimates for the growth curve analysis of the timing of because and weil*

ID		Parameter estimates (se)
Fixed parameters		
1	German Child (intercept)	0.190 (0.210)
2	German Child*Age **	0.339 (0.114)
3	German Child*Age ²	-0.218 (0.032)
4	German Child*Age ³	0.035 (0.010)
5	German Adult (intercept)	1.665 (0.191)
6	German Adult*Age	0.239 (0.040)
7	German Adult*Age ²	-0.060 (0.13)
8	German Child WC	0.063 (0.016)
9	German Adult WC	0.057 (0.100)
10	English Adult WC vs. German Adult WC	0.053 (0.019)
Random parameters		
11	German Child (intercept)	0.289 (0.085)*
12	German Child*Age	0.053 (0.019)*
13	German Adult (intercept)	0.169 (0.057)*

NOTE: ID = identification number of the estimate.

* One-sided as variance cannot take negative values.

** Age in days is rescaled around the mean age (divided by 100). Due to this centring, the variable age ranges from -7.716 to 4.622, which corresponds to 305 to 1,539 age in days. At 305 days (the youngest age) the natural logarithm of the mean expected probability of a German child using *weil* equals $0.190 + 0.339*(-7.716) - 0.218*(-7.716^2) + 0.035*(-7.716^3) = -31.475$. This means that at the age of 305 days the probability of *weil* occurring is $< .001$ (the reverse logarithmic transformation of -31.475). At the mean age (i.e., 1,077 days) the variable age in the equation equals 0. Therefore, the natural logarithm of the expected mean probability of occurrence equals $0.190 + 0.339*(0) - 0.218*(0^2) + 0.035*(0^3) = 0.190$. This equals a .547 probability of *weil* occurring at 1,077 days in age. At the highest age of 1,539 days, the natural logarithm of the mean expected probability of *weil* equals $0.190 + 0.339*(4.622) - 0.218*(4.622^2) + 0.035*(4.622^3) = .540$. This equals a probability of .632 that a German child will use *weil*. So, for every day the expected mean probability of occurrence can be estimated based on the model parameters. Please note that there are individual parameters as well. Hence, for each individual, for each day, the use of *weil* can be approximated.

added power of Age contributed significantly to the explanation of the occurrence of the connective, and thus led to a better fitting model.

RESULTS

Amount of connective use

The results of the growth curve analysis are presented in Table 3, which shows the best fitting model with the smallest number of parameters. Table 3 only includes significant effects (results are significant if the ratio of the parameter estimate and the standard error exceed 1.965, and hence

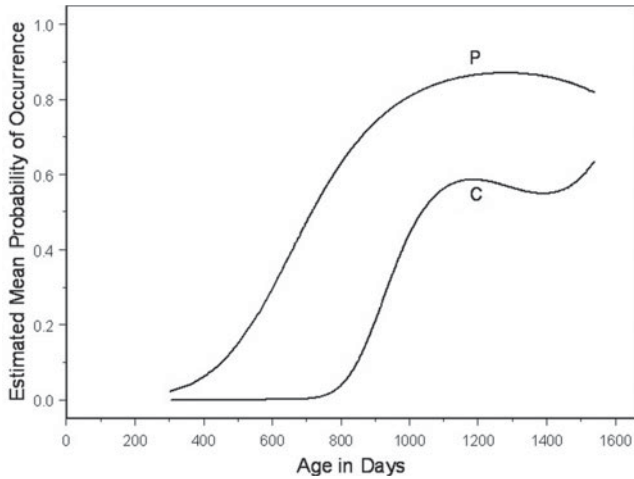


Fig. 2. Estimated mean occurrence of the use of *because* and *weil* by children (C) and parents (P).

$p < .05$). For clarity of reference, we will refer to the elements in the model by using the identification number associated with the relevant estimate.

As a result of parsimony restrictions, Table 3 only shows the fixed effects for German children (1–4). The parameter estimates for the corresponding effects for English children did not differ significantly from those for the German children. Thus, we do not find any differences between children's development of German *weil* and English *because*, and have to conclude that the average acquisition of *weil* is the same as the average acquisition of *because*.

For both languages, we find effects of Age (2–4), which show that the probability of *weil* and *because* occurring increases over time. Moreover, the effect of Age³ (4) indicates that the curve will have an s-shape. The resulting growth curve in Figure 2 shows that this is indeed the case: the overall development for children (C) increases from a probability of .00 at 300 days (0;10) to .60 at 1,500 days (4;02). Children start to use connectives at around 800 days (2;02), the point where the growth curve first starts to increase. The s-shape shows that growth slows down a little between 1,100 and 1,400 days (i.e., the probability of occurrence remains stable) before increasing again. Note that caution is required when interpreting the ends of growth curves. The model is sensitive to extremes at the beginning and end of the curves – the estimates are based on fewer data points – which means that a large peak or dip at the end of a curve is likely to be an artefact of the model.

As for the children, the parameter effects for the English adults did not differ from those of the German adults (5–7). So, with our current data, we cannot show that there are cross-linguistic differences in their use of *because* and *weil*. Both languages show an increase in use over time (6–7). The adults' overall growth curve in Figure 2 shows a steep linear increase in probability (6), slowing down at around 1,100 days, following a quadratic effect of Age (7). The growth curves show that adults increase their usage of the connective ahead of the children, which could be interpreted as a sign that adults are adapting their usage frequency to the children's increasing ability to produce connectives (i.e., a sign of scaffolding).⁶

Table 3 also shows significant effects of word count (8–10). The growth curves presented here control for file length by modelling the curves as if the word count is exactly the same in every recording for every child (set at the average Child Word Count of 1,570), and for every adult (set at the average Adult Word Count of 3,290). Nevertheless, it is possible to interpret the reported effects of word count: the positive estimates show that a larger word count has a positive effect on the probability of a connective occurring. The effect for German and English children is the same (8), but the effect is 0.053 larger for English adults (10) than for German adults (9). Thus, English adults with a high word count are more likely to use a connective than German adults with a high word count. A growth curve that allows the word count to vary per recording – taking into account the actual word count per recording – would look jagged with many peaks (caused by a large word count) and dips (caused by a small word count) in the connective's probability of occurrence, showing that word count plays an important role in the probability that a connective will be used. Note that word count can vary greatly for children and adults, which is why the extent of the peaks and dips will also vary.

The random parameters in the model show that there is variance between children (11) and adults (13) at the intercept (set at the average age in days of 1,077). The variance between children was found to increase as children grow older (12). This can be expected as young children do not yet use connectives, so for very young children the differences will be negligible. Once children grow older, and start to use connectives, the differences between children will be more marked. This is reflected in the individual growth curves, which we plotted in Figure 3 for German and in Figure 4 for English. The variation between children becomes clear when we look at the wide range of probability levels children reach: in Figures 3 and 4 we see

[6] A comparison of the parameter estimates for the children and the adults by means of a contrast test of the fixed parameters confirms that they differ from each other ($\chi^2(1) = 29.71$; $p < .001$). Thus, the adult growth curve cannot be described with the children's growth curve and vice versa.

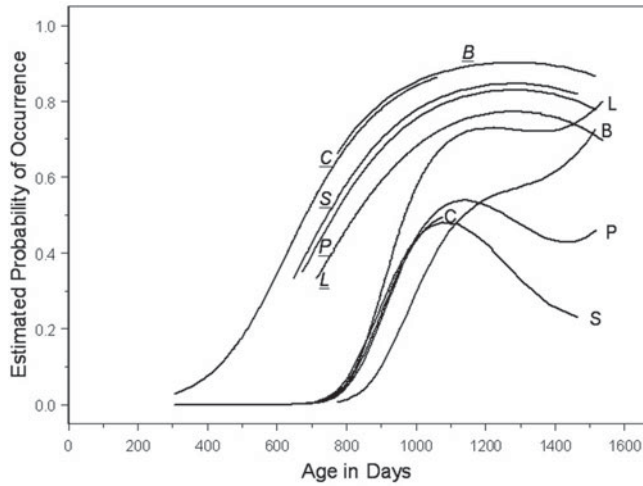


Fig. 3. Estimated mean occurrence of *weil* for individual German children and adults: Caroline (C); Caroline's parents (C); Leo (L); Leo's parents (L); Pauline (P); Pauline's parents (P); Simone (S); Simone's parents (S); Sebastian (B); Sebastian's parents (B).

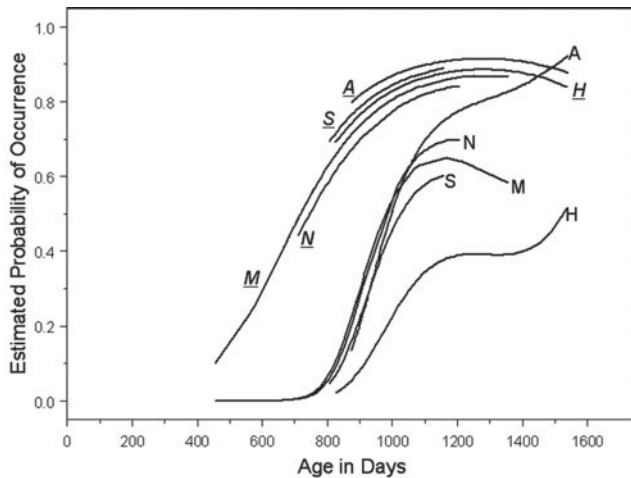


Fig. 4. Estimated mean occurrence of *because* for individual English children and adults: Abe (A); Abe's parents (A); Naomi (M); Naomi's parents (M); Nina (N); Nina's parents (N); Sarah (H); Sarah's parents (H); Shem (S); Shem's parents (S).

that Simone (S), Pauline (P) and Sarah (H) lag behind, while Abe (A) and Leo (L) are right ahead at the same level of probability as their caretakers are. Achieving the same chance of producing a connective as their caretakers

may be interpreted as a sign of complete mastery. The individual growth curves for the adults all follow a similar trend, but differ with respect to the level of probability they reach. For example, Abe's caretakers (*A*) show a .90 probability of using *because*, but Leo's caretakers (*L*) only reach a probability of .70.

THE TYPE OF CONNECTIVE USE

Apart from changing their amount of input, adults may also scaffold the child's connective use by adjusting the type of connective input to the child's increasing ability. In this section, we investigate the development of three different types of connective use: imitations, elicited use, and independent use.

METHOD

Measures

In order to ensure maximum comparability, we used the same data selection as for the amount of connective use. The selected utterances were annotated as one of three timing types: imitation, elicited use, and independent use. We did this by assessing the five preceding and five following utterances and, if necessary, the entire preceding discourse.

First, imitations encompass literal imitations or reformulations. As we saw in (1), adult imitations are often a type of reformulation. However, literal imitations may also occur. For example, in (5), we see that Leo's father uses an exact imitation of his son's utterance, although this time the purpose seems to be a matter of clarification rather than of reformulation.

(5) (Leo, 2;07-26)

FATHER: Wieso zittern die denn, die Fische?
'Why are they shivering, the fish?'

LEO: Ja. **Weil** die, **weil** die, **weil** das Wasser tief ist.
'Yes. Because the, because the, because the water is deep.'

FATHER: **Weil** das Wasser tief ist?
'Because the water is deep?'

Children may also imitate their caretakers. For example, Nina's utterance 'Because he loves Nina' in (6) is a literal imitation of her mother's previous utterance. In their study, Van Veen *et al.* (2009) suggest that short-term effects of input (i.e., effects within a single recording) may be attributed to children imitating their parents. The number of imitations will reveal whether this suggestion is plausible, or whether other explanations for short-term effects of input need to be sought.

(6) (Nina 2;02.28)

MOTHER: Is Heidi a big dog or a little dog?

NINA: Little dog.

MOTHER: He's a little tiny dog # isn't he? And when he sees Nina # he goes running up to Nina # doesn't he?
Because he loves Nina.

NINA: **Because** he loves Nina.

Second, children and adults may show elicited usage. This type of use generally occurs in response to a *how come* or a *why*-question (in German a *wieso* or *warum*-question). For example, Caroline's use of *weil* in (7) was elicited by her mothers *warum*-question (see also examples (2) and (5)).

(7) (Caroline 2;08.16/Situation: the butter is soft)

MOTHER: Gestern war sie glaube ich hart. Und warum war sie gestern hart?
 'I think it was hard yesterday. And why was it hard yesterday?'

CAROLINE: **Weil** die im Eisschrank war.
 'Because it was in the fridge.'

Third, the use of *because/weil* may be independent. In order for an utterance to be independent it must be unprompted and creative. That is to say, it may not be elicited nor may it imitate (part of) the previous utterance. For example, Shem's use of *because* in (8) is independent because he makes the causal relation on his own accord. Similarly, the investigator's *weil* in (9) is independent in that the causal relation was not (verbally) elicited by the child. By using *because* or *weil* in child-directed utterances that are not imitations or responses to *why*-questions, adult independent usage provides positive evidence. Compare also (3) and (4).

(8) (Shem 2;07.18)

SHEM: I want, I was got uh pee stomachache.

INVESTIGATOR: You have to pee?

SHEM: I'm going like this **because** I peeing in my pants.

INVESTIGATOR: You are?

SHEM: Yeah.

INVESTIGATOR: Oh, that's not a very good idea. Why don't we just go to the bathroom?

(9) (Pauline 2;09.18/Situation: playing shops)

PAULINE: Ein Zahnpasta kaufen.
 'Buy toothpaste.'

INVESTIGATOR: Ja, die muss man haben, das ist richtig.
 'Yes, one should have that, that's right.'

INVESTIGATOR: Und vielleicht etwas Spinat? **Weil** es jetzt draussen keinen mehr gibt.
 ‘And maybe some spinach? Because there isn’t any more outside.’

Data analysis

We performed the same type of growth curve analysis for the different types of connectives as for the timing of the connective: a multilevel logistic regression (Goldstein, 1999). In short, we modelled the probability of occurrence of independent and elicited use of *because* or *weil* for English children, English adults, German children, and German adults in a single model. We performed a growth curve analysis as a function of the predictor Age and the covariates Child Word Count (Child WC) and Adult Word Count (Adult WC). As we shall see in the next section, the use of imitations was rare. This meant that we did not have enough data to include imitations in the growth curve analysis.

RESULTS

Type of connective use

Table 4 presents an overview of the children’s uses of *because* and *weil*. The most noticeable result is the children’s low number of imitations: sharing a mere 10 instances between them, only six children were found to imitate their caretakers’ previous utterance. Thus, interpreting effects of input in the parroting sense of imitation is incorrect.

The overview also shows that adults often elicit causal responses through *why/warum*-questions: between 25 and 81 % of the children’s connective use is elicited. For four children elicitation is responsible for the largest part of their *because/weil* usage (Abe, Naomi, Sarah, and Simone). The remaining six children – four of whom are German – show a predominantly independent usage of *because/weil*.

The overview of the adult uses of *because* and *weil* in Table 5 shows that adults produce a larger percentage of independent connective use than their children. They also show a slightly larger number of imitations: between 1 and 24 % of their usage consists of reformulating or imitating their child’s previous utterance. The caretakers of Simone, Naomi, Shem, Caroline, and Leo show a particular preference for this usage, indicating distinct individual styles. Yet, as imitations/reformulations only make up a small portion of the total adult usage, their role is likely to be limited.

As they do themselves, adults are often asked *why/warum*-questions by their children. This shows that their children are interested in causality.

TABLE 4. *Child usage of because/weil specified for type of connective (with % of total)*

Name	Age range	# Files	Number of child <i>because/weil</i>			
			Total	Elicited by adults	Imitation	Independent
English children						
Abe	2;04·24-4;02·19	167	469	273 (58%)	1 (0%)	195 (42%)
Naomi	1;02·29-3;08·19	90	80	20 (25%)	0	60 (75%)
Nina	1;11·16-3;03·21	52	303	103 (34%)	1 (0%)	199 (66%)
Sarah	2;03·05-4;02·16	96	45	27 (60%)	1 (2%)	17 (38%)
Shem	2;02·16-3;02·02	40	319	162 (51%)	1 (0%)	156 (49%)
German children						
Caroline	0;10·01-2;10·26	232	83	28 (34%)	3 (4%)	52 (63%)
Leo	1;11·12-4;02·16	462	852	376 (44%)	3 (0%)	473 (56%)
Pauline	1;10·09-4;01·28	46	63	21 (33%)	0	42 (67%)
Sebastian	2;01·12-4;01·27	43	40	16 (40%)	0	24 (60%)
Simone	1;09·11-4;00·06	74	305	246 (81%)	0	59 (19%)

TABLE 5. *Adult use of because/weil specified by timing of the connective (with % of total)*

			Number of parental <i>because/weil</i>			
Name	Age range	# Files	Total	Elicited by child	Imitation	Independent
English children						
Abe	2;04·24-4;02·19	167	266	172 (65%)	9 (3%)	85 (32%)
Naomi	1;02·29-3;08·19	90	50	10 (20%)	5 (10%)	35 (70%)
Nina	1;11·16-3;03·21	52	242	23 (10%)	8 (3%)	211 (87%)
Sarah	2;03·05-4;02·16	96	170	56 (33%)	1 (1%)	113 (66%)
Shem	2;02·16-3;02·02	40	315	57 (18%)	24 (8%)	234 (74%)
German children						
Caroline	0;10·01-2;10·26	232	309	168 (54%)	24 (8%)	117 (38%)
Leo	1;11·12-4;02·16	462	1218	379 (31%)	141 (12%)	698 (57%)
Pauline	1;10·09-4;01·28	46	124	12 (10%)	5 (4%)	107 (86%)
Sebastian	2;01·12-4;01·27	43	200	56 (28%)	6 (3%)	138 (69%)
Simone	1;09·11-4;00·06	74	331	7 (2%)	80 (24%)	244 (68%)

Here too, we find many individual differences. On the one hand, Abe's and Caroline's caretakers often use *because/weil* in response to the children's questions (65% and 54%, respectively) and consequently have a small percentage of independent use (32% and 38%, respectively). As a result, Abe and Caroline are responsible for a large part of their own connective input. On the other hand, Nina's and Pauline's caretakers' use of

TABLE 6. *Parameter estimates for the growth curve analysis of different types of because and weil*

ID		Parameter estimates (se)
Fixed parameters		
1	G Child Independent (intercept)	-1.135 (0.162)
2	G Child Independent*Age	0.519 (0.058)
3	G Child Independent*Age ²	-0.089 (0.021)
4	G Child Elicited (intercept)	-0.489 (0.157)
5	G Child Elicited*Age	0.196 (0.083)
6	G Child Elicited*Age ²	-0.157 (0.027)
7	G Child Elicited*Age ³	0.027 (0.009)
8	G Adult Independent (intercept)	0.274 (0.157)
9	E Adult Independent (intercept) vs. G Adult Independent (intercept)	0.441 (0.225)
10	G Adult Elicited (intercept)	-1.670 (0.179)
11	G Adult Elicited*Age	0.475 (0.054)
12	E Adult Elicited (intercept) vs. G Adult Elicited (intercept)	-0.217 (0.254)
13	E Adult Elicited*Age vs. G Adult Elicited*Age	-0.221 (0.078)
14	G Child WC Independent	0.031 (0.014)
15	G Child WC Elicited	0.040 (0.013)
16	G Adult WC Independent	0.061 (0.006)
17	G Adult WC Elicited	-0.034 (0.008)
Random parameters		
18	G Child (intercept)	0.110 (0.035)*
19	G Adult (intercept)	0.070 (0.023)*

NOTE: ID=identification number of the estimate. German (G); English (E).

* One-sided as variance cannot take negative values.

because/weil is rarely elicited (10% each), but does show a larger percentage of independent use (87% and 86%, respectively).

The significant results of the growth curve analysis are presented in Table 6, which shows the best fitting model with the smallest number of parameters.

The model for the type of connective use presented in Table 6 shows similarities to the overall model in Table 3. As for the overall model, the random parameters show that there is variance between children (18) and adults (19) at the intercept, which is larger for the children than for the adults. In addition, we find effects of word count (14–17). The growth curves in Figures 5 and 6 control for file length by modelling the curves for the average word count. However, note that, in contrast to the overall model, the effects of word count are the same for English and German adults (which is why only the data of the German adults are included in the model).

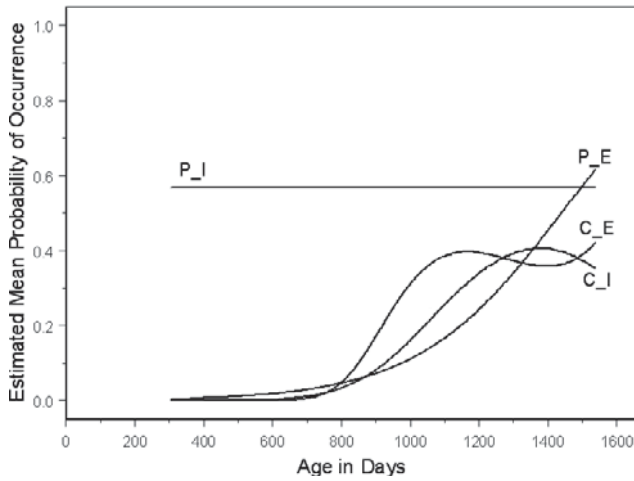


Fig. 5. Estimated mean occurrence of elicited (E) and independent (I) use of *weil* by German parents (P) and children (C).

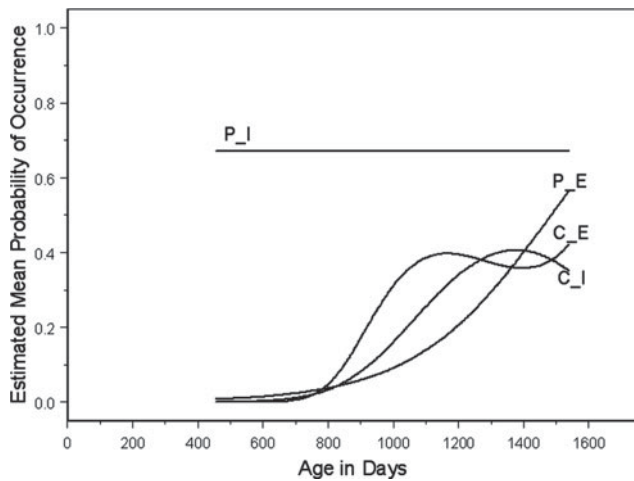


Fig. 6. Estimated mean occurrence of elicited (E) and independent (I) use of *because* by English parents (P) and children (C).

Turning to the independent (1–3) and elicited uses (4–7) in Table 6, we only see the German data in the model. This means that German and English children show the same developmental pattern, a finding that can also be derived from the children's growth curves in Figures 5 and 6, which

are identical. These curves indicate that the children's elicited use (C_E) develops faster than their independent use (C_I), but that children start to produce both elicited and independent use from 800 days.

The children's elicited use stabilizes at .40 at around 1,000 days (2;09), due to a quadratic effect of age (6). Then, a cubic effect of age (7) ensures that their development picks up again towards the end of curve, after 1,400 days. In comparison, the children's independent use increases at a slower rate (2), but eventually reaches a higher probability of occurrence of .45 at 1,300 days (3;07). It is likely that their independent use will increase again later, due to a cubic effect (creating an s-shape), but unfortunately our data do not capture that phase of their development.

Since the model in Table 6 does not directly compare parameter estimates for independent and elicited use, we need a contrast test of the fixed parameters to confirm that the resulting growth curves differ significantly from each other. This test confirms that they are indeed different: the growth curve for independent use cannot be described with the growth curve for elicited use, and vice versa ($\chi^2(1) = 6.23$, $p = .013$).

Moving on to adult independent use, we find that the results in Table 6 do not show any effects of age, but only a constant: the intercept (8–9). This means that their independent use does not develop over time. Therefore, their growth curves in Figures 5 and 6 consist of a horizontal line (P_I). The intercept for the German adults (8) is lower than for the English adults (9), a finding that is reflected in the growth curve: the English line lies at a higher probability than the German line.

For the adult elicited use (10–13), we only find linear effects of age. This effect is larger for German adults (11) than for English adults (13) and can be traced to their growth curves (P_E): the German adults have a steeper slope and reach a probability of .65, whereas the English adults reach a slightly lower probability of .60.

Importantly, a comparison of Figures 5 and 6 to the model for the amount of connective input (Table 3, Figure 2) reveals that the effects of age found in the overall model are not caused by the adults' independent use, but by their elicited use. Thus, the signs of scaffolding in Figure 2 (adults increase their connective use ahead of their children) were in fact triggered by the increasing amount of *why*- and *warum*-questions posed by the children.

DISCUSSION

The current study used growth curve analysis to study the amount and type of input during the acquisition of the causal connectives *because* and *weil* in German and English longitudinal corpora. In this section, we will (a) discuss the most striking findings for the children's connective

development, (b) interpret the findings on parental input in terms of scaffolding, and (c) discuss the merits and limits of using growth curve analysis.

The children's connective development

Children's connective acquisition is a long-term process: their production of connectives gradually increases as they grow older with elicited connective use developing faster than their independent use (compare findings by Braunwald, 1997; Diessel, 2004). This finding can be related to the relative complexity of independent connective clauses. First, elicited usage only requires the production of the consequent clause. This is less complex than independent usage, which often requires the child to produce both the antecedent and the consequent clause. Second, elicited usage is easier because *why/warum*-questions help the child to identify those situations in which it is appropriate to provide causal information, and hence can be said to have a beneficial effect on children's later causal competence. A similar effect has previously been reported by Peterson, Jesso, and McCabe (1999), who encouraged mothers to elicit narratives by asking open-ended and context-eliciting questions, and found that, as a result, their children showed an overall improvement in narrative skill.

Our analyses also revealed that children very rarely imitate their caretakers (10 out of 2,559 instances). This finding suggests that imitations do not play a role in children's connective acquisition. Moreover, it means that we must rule out Van Veen *et al.*'s (2009) suggestion that imitations are (partially) responsible for effects of short-term parental input (within one recording). An alternative explanation for short-term input effects is that of interactive alignment (Garrod & Pickering, 2004: 10), which predicts that 'speakers reuse the structures that they have just interpreted as listeners when formulating their response'. This would predict that when adults use a particular structure with *because*, children start to use this structure too. Interactive alignment is appealing in that it is not restricted to an occurrence in the next turn, or to the exact words making up the structure. Moreover, it does not only account for short-term effects, but also for long-term effects: since the same speakers usually surround the child, any alignment effects will continue across and between recordings. Only an extensive analysis of each recording – annotating every linguistic structure – could reveal whether interactive alignment is indeed at work here.

The growth curve analyses allowed us to generalize across children, but also made us aware of individual differences. The growth curves in Figures 3 and 4 show that children differ with respect to the speed and the extent to which they develop their connective use. This may be related to the amount

of input they receive. For example, the highest-achieving child, Abe, receives the highest amount of input (see Figure 4). However, as the growth curve for Leo illustrates, this relationship is not straightforward: although Leo is the highest-achieving German child, his caretakers provide the least input (when controlled for word count). Thus, there must be additional factors that determine the children's speed of development. For example, children may differ with respect to their cognitive development: children who are unable to cope with causal reasoning will also have difficulty using a causal connective. Apart from adult and conceptual influences, the connective's optionality may also play a role. Since it is possible to express causality without *because* or *weil* (i.e., unmarked causality), it is possible that individual preference or style – and not linguistic competence per se – leads to a higher or lower probability of occurrence.

Even though our study was based on a small number of children, which means that (the lack of) cross-linguistic effects may be due to sampling issues, the English and German children showed similar developmental patterns. Thus, the cross-linguistic differences between *because* and *weil* with respect to the relative complexity of word order and lexical competition were not found to cause any developmental differences. Rather, it would seem that their conceptual compatibility (Pit, 2003; Sanders, 2005) – the fact that *because* and *weil* are the most frequent (default) connectives for denoting backward causal relations in their language – plays a more important role than word order. In order to find true effects of word order, we must turn to a language that has two backward causal connectives that are equally frequent but require a different word order, such as Dutch. In studies of Dutch connective acquisition, the coordinator *want* 'because' is acquired before the subordinator connective *omdat* 'because' (Evers-Vermeul & Sanders, 2009, 2011). Thus, it would seem that a connective's speed of acquisition is only determined by the complexity of the word order when there is an equally frequent lexical competitor.

The adults' connective use

Although previous studies have already shown that input influences connective acquisition (Diessel, 2004; Van Veen *et al.*, 2009), they did not show whether input is constant or whether adults scaffold children's connective use. Our growth curves showed that the amount of the adult use of *because* and *weil* was by no means constant. The overall growth curves in Figure 2 seem to point towards an effect of scaffolding: adults increase their connective use in anticipation of the children's increasing abilities to use the connective (compare Figure 1c). However, the growth curves per connective type in Figures 5 and 6 suggest otherwise. First, the adult

independent use remained constant (compare Figure 1a). This means that the adults' independent use seems to function independently of the children's connective development, and hence cannot be regarded as a scaffold. Second, the adult elicited use was found to develop after the children's elicited use. Hence, adult elicited use does not fit any of the profiles in Figure 1: it is not constant, nor does it resemble audience design or scaffolding. Instead, the development of adult elicited use is controlled by the children asking *why/warum*-questions. In other words, adults' elicited use follows the children's development instead of preceding it, and therefore this type of connective use cannot be seen as a scaffold for children's connective use either.

However, we would like to claim that elicited use is the result of a different (non-connective) scaffold, namely adult *why/warum*-questions. As the children's growth curves for elicited use show (Figures 5 and 6), these adult questions ensure that children develop their elicited use at a faster rate than their independent use: each child's elicited connective use represents an accompanying adult *why/warum*-question. However, Figures 5 and 6 at best give an indication of the minimum amount of adult *why/warum*-questions, because we only present those cases in which the children answered the question by using a connective. Of course, adults may also ask questions that are left unanswered by the child. As a result, the amount of *why/warum*-questions is likely to be underestimated in the current growth curves, in particular during the children's earlier development. This would be in line with research by McCabe and Peterson (1997), who show that the number of adult *why*-questions increases in the five or six months prior to their children's first spontaneous causal productions (see also Garvey, 1984). If this also holds for our data, we would expect the growth curves of adult *why/warum*-questions in relation to children's answers to look like the scaffolding in Figure 1c, with the number of questions increasing ahead of the children's answers. A recently completed study of adult *why*-questions as scaffolds for connective acquisition supports this prediction (Van Veen, Evers-Vermeul, Van den Bergh & Sanders, unpublished observations). This study of English and Dutch parental *why*-questions shows that the probability of parents asking a *why*-question increases as children grow older, suggesting that they are timing their questions to their child's increasing ability to answer them. In line with McCabe and Peterson (1997), this study also finds that the first parental *why*-question always precedes the child's first production of a causal statement with *because*. Moreover, this study shows that parents do not only fine-tune the frequency, but also the complexity of their *why*-questions to the cognitive and linguistic abilities of the child: they ask questions that contain highly familiar propositions, trying to adapt the content to what the child already 'knows' from the previous conversation.

Not only the adults, but also the children were found to ask *why/warum*-questions. This result is in line with the study by McCabe and Peterson (1997), who find that from age 2;2 children start to ask *why*-questions increasingly often. The children's interest in causality may be related to their own ability to answer *why/warum*-questions, as the increase in adult elicited use directly follows the increase in children's elicited use. Thus, adult *why/warum*-questions are scaffolds of their children's connective use, as well as of their children's ability to explore causality by asking *why/warum*-questions themselves. The growth curve analysis shows that the adult elicited use increases at a faster rate for German adults than for English adults. This means that German children ask increasingly more *warum*-questions than English children ask *why*-questions. The growth curves in Figures 5 and 6 show that this difference is subtle, and so far the reason for this difference is unclear.

A consequence of the children's *why/warum*-questions is that children are, in effect, responsible for the increase in their own input. The rise in elicited adult examples of *weil/because* use may also contribute to the children's increasing ability to produce these connectives independently. Additional adult input also occurred as a result of adult imitations or reformulations. Although these did not occur very often (between 1 and 25% of total connective use), their function as reinforcement is important as they communicate to the children that they are on the right track.

Overall, we find that it is not the adults' use of the connectives themselves that scaffolds connective acquisition, but the adults' introduction of contexts (*why/warum*-questions) in which children may show and elicit causal knowledge that is key to the acquisition of causal connectives. Thus, by distinguishing different types of connective input, the current study disentangled parts of an intricate system in which the role of input is clearly important, but hard to discern.

Merits and limits of growth curve analyses

The current study showed that growth curve analysis, that is multilevel regression analyses, produces interpretable results. Importantly, we were able to take every datapoint into account by using Age as a continuous variable. The growth curves visualized the speed of development and helped us to identify the overall interaction between children's and adults' language use. This is an improvement in comparison to analyses in which various aspects of children's linguistic development are related to overall amounts of parental input. Moreover, the multilevel approach ensured a longitudinal analysis in which children with varying ages and number of recordings could be assessed without worrying about complications caused by missing data or lack of fixed occasions.

In our study we opted for a multilevel LOGISTIC regression analysis, which means that we characterized our dependent variable as dichotomous: absence versus presence of the connective. Hence, the analysis does not provide any details about the number of connectives that occur in a recording. A logistic regression analysis is particularly suitable for linguistic phenomena that are optional, like connectives, because it circumvents the question as to how often a connective should ‘normally’ be used.

Although growth curve analysis is useful, it also has its limits. It can be seen as a tool that can help us to identify phases we should investigate with more in-depth analyses. In order to really address some of the issues raised, we have shown that one would have to carry out additional qualitative analyses on the function and pragmatics of connective use. For example, the corpora were compiled in different situations, and the likelihood of occurrence of *because/weil* may also be determined by the topic of the conversation.

CONCLUSION

The current study set out to use growth curve analysis to study the role of scaffolding during connective acquisition. We added a cross-linguistic perspective by investigating the longitudinal corpora of English *because* and German *weil*. By looking at the development of the various uses of *because* and *weil* – imitated use, elicited use, and independent use – we were able to present a detailed view of the nature of the connectives adults and children produce.

The children’s growth curves show that connective acquisition is a long-term process during which the children’s production of elicited use develops faster than their production of independent use. We account for this finding by the fact that elicited use involves a higher level of scaffolding, as well as an overall lower level of complexity than independent use. Furthermore, we showed that children’s imitations do not play a role in the acquisition of *because* and *weil*.

The analysis of the adult connective use revealed that adults do not scaffold children’s connective use by adapting their independent use to the abilities of the children. Rather, it appears that *why/warum*-questions play an important role in the acquisition of *because* and *weil*. Although the exact nature of the interaction still needs to be determined, our findings show that we are dealing with a two-way system in which adult–child interaction is central.

In summary, we can construct a timeline along which *because* and *weil* develop. First, adults use connectives independently. Although adult independent use does not appear to function as a scaffold, children are likely to be sensitive to the input in this phase. Later, adults start to ask

why/warum-questions and thereby scaffold children's connective acquisition. Subsequently, children start to answer these questions by using connectives and begin to produce the connectives independently, although the children's elicited use increases at a faster rate than their independent use. Finally, children start to ask *why/warum*-questions themselves, which in turn leads to an increase in adult elicited use.

Cross-linguistic differences between German and English were only detected for adult elicited and independent use, but these differences were not large enough to affect the children's connective development (which was the same for German and English). This means that we could be dealing with an input threshold: children need a certain amount of input in order to develop a sensitivity for discourse connectives, and any additional input does not affect their speed of development. Future research will have to reveal whether these findings hold for additional languages and a larger number of children.

From a methodological point of view, it is important to stress the possibilities of the growth curve analyses we have used. We did not only employ growth curve analysis to determine the children's connective development, but also to reveal any changes in the connective input during children's acquisition process. In fact, we trust that growth curve analysis will be a successful method to study many aspects of children's language development, not just of discourse connectives, but of various other linguistic elements as well.

REFERENCES

- Behrens, H. (2006). The input–output relationship in first language acquisition. *Language and Cognitive Processes* 21, 2–24.
- 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, 235–61.
- Braunwald, S. (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*. New Jersey: Lawrence Erlbaum.
- Brown, R. (1973). *A first language: the early stages*. Cambridge, MA: Harvard University Press.
- Chouinard, M. M. & Clark, E. V. (2003). Adult reformulations of child errors as negative evidence. *Journal of Child Language* 30, 637–69.
- Clark, E. V. (1978). Discovering what words can do. In D. Farkas, W. M. Jacobsen & K. W. Todrys (eds.), *Papers from the parasession on the lexicon*. Chicago, IL: Chicago Linguistic Society.
- Clark, H. H. & Murphy, G. L. (1982). Audience design in meaning and reference. In J. F. LeNy & W. Kintsch (eds.), *Language and comprehension*. Amsterdam: North-Holland Publishing Company.
- De Villiers, J. & De Villiers, P. (1973). A cross-sectional study of the acquisition of grammatical morphemes in child speech. *Journal of Psycholinguistics Research* 2, 267–78.
- Diessel, H. (2004). *The acquisition of complex sentences*. Cambridge: C.U.P.

- Evers-Vermeul, J. & Sanders, T. (2009). The emergence of Dutch connectives: how cumulative cognitive complexity explains the order of acquisition. *Journal of Child Language* **36**, 829–54.
- Evers-Vermeul, J. & Sanders, T. (2011). Discovering domains – on the acquisition of causal connectives. *Journal of Pragmatics* **43**, 1645–62.
- Farrar, M. J. (1992). Negative evidence and grammatical morpheme acquisition. *Developmental Psychology* **28**, 90–98.
- Garrod, S. & Pickering, M. J. (2004) Why is conversation so easy? *Trends in Cognitive Sciences* **8**, 8–11.
- Garvey, C. (1984). Children's talk. Oxford: O.U.P.
- Goldstein, H. (1979). The design and analysis of longitudinal studies: their role in the measurement of change. London: Academic Press.
- Goldstein, H. (1999). Multilevel statistical models. London: Edward Arnold.
- Günthner, S. (1993). '...weil – man kann es ja wissenschaftlich untersuchen' – Diskurspragmatische Aspekte der Wortstellung in WEIL-Sätzen. *Linguistische Berichte* **143**, 37–59.
- Hood, L., Bloom, L. & Brainerd, C. J. (1979). What, when, and how about why: a longitudinal study of early expressions of causality. *Monographs of the Society for Research in Child Development* **44**, 1–47.
- Keller, R. (1995). The epistemic *weil*. In D. Stein & S. Wright (eds.), *Subjectivity and subjectivisation: linguistic perspectives*. Cambridge: C.U.P.
- Kirjavainen, M., Theakston, A. & Lieven, E. (2009). Can input explain children's me-for-I errors? *Journal of Child Language* **36**, 1091–114.
- Klee, T., Schaffer, M., May, S., Membrino, I. & Mougey, K. (1989). A comparison of the age-MLU relation in normal and specifically language-impaired preschool children. *Journal of Speech and Hearing Disorders* **54**, 226–33.
- Kuczaj, S. (1976). *-ing, -s and -ed: a study of the acquisition of certain verb inflections*. Unpublished doctoral dissertation, University of Minnesota.
- MacWhinney, B. (2000). The CHILDES project: tools for analyzing talk (3rd ed.). Mahwah, NJ: 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*. Mahwah, NJ: Erlbaum.
- Menard, S. (2002). *Applied logistic regression analysis* (2nd ed.) (Quantitative Applications in the Social Sciences, 106). Thousand Oaks, CA: Sage.
- Miller, J. F. & Chapman, R. S. (1981). The relation between age and mean length of utterance in morphemes. *Journal of Speech and Hearing Research* **24**, 154–61.
- Miller, M. (1979). *The logic of language development in early childhood*. Berlin: Springer-Verlag.
- 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.
- Peterson, C., Jesso, B. & McCabe, A. (1999). Encouraging narratives in preschoolers: an intervention study. *Journal of Child Language* **26**, 49–67.
- Pit, M. (2003). How to express yourself with a causal connective: subjectivity and causal connectives in Dutch, German and French. Amsterdam: Rodopi.
- Quené, H. & Van den Bergh, H. (2004). On multi-level modeling of data from repeated measures designs: a tutorial. *Speech Communication* **43**, 103–21.
- Quené, H. & Van den Bergh, H. (2008). Examples of mixed-effects modeling with crossed random effects and with binomial data. *Journal of Memory and Language* **59**, 413–25.
- Sachs, J. (1983). Talking about the there and then: the emergence of displaced reference in parent-child discourse. In K. E. Nelson (ed.), *Children's language*, Vol. 4. Hillsdale, NJ: Erlbaum.

- Sanders, T. (2005). Coherence, causality and cognitive complexity in discourse. In M. Aurnague, M. Bras, A. Le Draoulec & L. Vieu (eds.), *Proceedings/Actes SEM-05, First International Symposium on the exploration and modelling of meaning*: Biarritz, France.
- Sanders, T., Spooren, W. & Noordman, L. (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, 561–85.
- Saxton, M. (2000). Negative evidence and negative feedback: immediate effects on the grammaticality of child speech. *First Language* 20, 221–52.
- Stoll, S. & Gries, S. T. (2009). How to measure development in corpora? An association strength approach. *Journal of Child Language* 36, 1075–90.
- Suppes, P. (1974). The semantics of children's language. *American Psychologist* 29, 103–14.
- Theakston, A. L. & Lieven, E. V. M. (2008). The influence of discourse context on children's provision of auxiliary BE. *Journal of Child Language* 35, 129–58.
- Van den Bergh, H. & Rijlaarsdam, G. (1996). The analysis of writing process data: a mini-longitudinal study. In M. C. Levy & S. Ransdell (eds.), *The science of writing*. New York: Lawrence Erlbaum.
- Van Veen, R., Evers-Vermeul, J., Sanders, T. & Van den Bergh, H. (2009). Parental input and connective acquisition in German: a growth-curve analysis. *First Language* 29, 267–89.
- Vygotsky, L. (1978). Interaction between learning and development. In M. Cole (Trans.), *Mind in society*. Cambridge, MA: Harvard University Press.
- Wijnen, F., Kempen, M. & Gillis, S. (2001). Root infinitives in Dutch early child language: an effect of input? *Journal of Child Language* 28, 629–60.
- Wood, D., Bruner, J. S. & Ross, G. (1976). The role of tutoring in problem solving. *Journal of Child Psychology and Child Psychiatry* 17, 89–100.