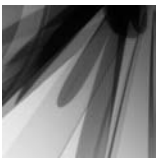

ARTICLE

FIRST LANGUAGE

www.sagepublications.com Vol 24(1): 033–070 (200402)
DO1: 10.1177/0142723704040704

Early acquisition of verb grammar and lexical development: evidence from periphrastic constructions in French and Austrian German

Dominique Bassano, *CNRS & Université Paris 5*
Sabine Laaha, *Österreichische Akademie der Wissenchaften*
Isabelle Maillochon, *CNRS & Université Paris 5*
Wolfgang U. Dressler, *Österreichische Akademie der Wissenchaften*

ABSTRACT
This paper takes a functionalist approach to the acquisition of verb morphology in French and Austrian German. The development of periphrastic constructions using auxiliaries and modal verbs (compound past, modal constructions and analytic future) was examined in two French and two Austrian children's spontaneous speech samples from the onset of production until 3;0. Cross-linguistic comparisons showed both similarities and differences in the development of periphrastic constructions (e.g., compound past was the first structure to emerge in French but not in Austrian German, analytic future was the last in both languages), suggesting an interplay between general cognitive and language-specific factors. In the four children, developmental analyses showed precursors, such as bare infinitives and past participles and preverbal fillers, which denote a gradual and continuous acquisition process. They also showed temporal relations between grammaticization and lexical production of verbs, which is an argument for the 'critical lexical mass' hypothesis and for interdependencies between lexical and grammatical developments. These analyses support a constructivist and interactive view of language acquisition.

KEYWORDS
Compound verbal constructions; continuity/discontinuity; cross-linguistic study; early language development; functionalist approach; naturalistic production; verb grammaticization

INTRODUCTION

In the last few decades, great interest has been paid to the acquisition of verb grammar, and a number of recent studies have conducted fine-grained analyses of verb grammar acquisition in natural production of young children learning various languages, e.g., Tomasello (1992) for English; Behrens (1993) for German; Pizzuto & Caselli (1994) for Italian; Choi (1998) for Korean; Aksu-Koç (1998) for Turkish. However, few studies have directly compared different languages using a common approach and method. This is the purpose of our present study. It is part of a larger research programme which takes a functionalist approach to comparing the development of verb morphology in two languages, French and Austrian German, from the onset of speech production until three years of age (cf. Bassano, Maillochon, Klampfer & Dressler, 2001*a*, 2001*b*; Klampfer, Maillochon, Bassano & Dressler, 1999). Here we focus on a central aspect of verb grammar: the emergence of periphrastic constructions consisting of a grammatical verb and a main verb. In the two languages, relevant periphrastic constructions for early language consist of three structures: 'auxiliary plus past participle', which mainly corresponds to the compound past tense; 'modal verb plus infinitive', which refers to modal constructions; and 'auxiliary plus infinitive', which mainly corresponds to the analytic future. Our present analysis of the development of these structures focuses on the question of when and how children become able to produce adult-like complete forms composed of both the grammatical and the main verb. This question is at the heart of theoretical issues such as the acquisition of grammatical morphemes, and the nature of the verb grammaticization process, e.g., innate or constructed, sudden or gradual, discontinuous or continuous.

A functionalist approach to the development of verb grammar

Like our larger research on the development of verb forms in French and Austrian German and like a number of our other recent works, the present study on the development of periphrastic constructions has been conducted in the general framework of the functionalist approach to language acquisition (cf. Bates & MacWhinney, 1987, 1989; Elman, Bates, Johnson, Karmiloff-Smith, Parisi & Plunkett, 1996; Karmiloff-Smith, 1992; MacWhinney, 1999*a*; Slobin, 1997*b*; Tomasello, 1995, 1998). It is in line with the linguistic theory of Natural Morphology (Dressler, 1997, 2000; Dressler, Mayerthaler, Panagl & Wurzel, 1987), which shares the functionalist perspective.

The functionalist approach to the acquisition of grammar is directly inspired by principles of Cognitive and Functional Linguistics, in which there is widespread agreement 'that language is not an autonomous "mental organ", but rather that it is a complex mosaic of cognitive and social communicative activities closely integrated to the rest of human psychology' (Tomasello, 1998: ix). The main idea of the functionalist

conception is that the linguistic forms of natural languages are created, governed, constrained, acquired and used for purposes of communicative function in its two basic aspects, pragmatics and semantics. Functions are embodied in structures in conventional and multi-relational ways, the structures being all composed of some combination of four types of symbolic devices: words, markers on words, word (and morpheme) order, and prosody (Bates & MacWhinney, 1987, 1989). From a developmental point of view, the more general assumption in the functionalist approach is that linguistic categories are not given to children innately, but rather they are gradually constructed by the child on the basis of an interplay between general predispositions and input stimuli. Input stimuli allow children to make generalizations on the basis of their categorization skills working on the language they hear (Tomasello, 1998). In this perspective, children could show different developmental patterns of acquisition as a function of the language learned, due to properties of individual languages, that is, to formal differences in linguistic input. Thus, fine-grained analyses and cross-linguistic comparisons of the ways in which particular languages are acquired are crucial for articulating universals and particulars in the process of language acquisition (Bates & MacWhinney, 1987; Bittner, Dressler & Kilani-Schoch, 2000; Slobin, 1985, 1997*a*).

A second – and related – main assumption of the functionalist approach concerning acquisition is that linguistic categories emerge and develop in relation with the development of cognitive and other language skills, in particular lexical skills. This conception calls into question the traditional generativist view that syntax is independent of all other levels of linguistic description including semantics. By contrast, the integrative conception of language processing has largely developed the view that listeners use at the same time and integrate different sources and levels of information, such as lexical, syntactic or pragmatic levels. Similarly, a number of recent works on early language conducted in the integrative perspective have shown that there exist strong interdependencies between the different dimensions of language in the course of the acquisition process, in particular between lexical and grammatical development. Various longitudinal and cross-sectional studies using the MacArthur CDI parental reports have found tight non-linear correlations between vocabulary size and sentence complexity across the 16- to 30-month age range, for children acquiring English as well as other languages such as Italian (Bates, Dale & Thal, 1995; Bates & Goodman, 1999; Caselli, Casadio & Bates, 1999). These data indicate that there is a powerful relationship between grammar and lexical growth across the second and third year of life, and that this relation might be a universal property of language development.

A further question is whether the link between lexical and grammatical development exists when looking at specific aspects of grammar. Marchman & Bates (1994) addressed this issue, using the MacArthur data to investigate the relation between the number of verbs children use and their progress on the verb morphology subscales on the CDI. They reported a powerful non-linear relation between the number of verbs in the child's vocabulary and three forms of past tense marking – zero stem, correct irregulars, incorrect overgeneralizations – thus showing a link between specific grammatical structures and their required critical mass of lexical items. These findings were in accordance with those observed in connectionist simulations of past tense learning (e.g., Plunkett & Juola, 1999). Similar results were found in naturalistic studies

on French children's spontaneous speech samples. A longitudinal study on the development of nouns and verbs in one child from 1;2 to 2;6 (Bassano, 2000) showed temporal relations between the quantitative development of noun and verb lexicons and the grammatical development of the noun and verb classes. For both word classes, explosions in the grammaticization processes (assessed through determiner use for nouns, and through inflection and auxiliary use for verbs) occurred slightly after increases in lexical production. In a further study focusing on the noun grammaticization process in 10 French children at ages 1;8 and 2;6 (Bassano & Eme, 2001), strong correlations were found between the number of noun types produced by the children and the level of the grammaticization process. These results are in line with the 'critical mass hypothesis' (Bates & Goodman, 1999; Marchman & Bates, 1994), which claims that developments within morphosyntax are triggered by an increase in the size of the lexicon beyond a given level, thus providing supporting evidence for the inter-dependence of lexical and morphosyntactic developments.

A comparative perspective on the French and Austrian German verbal systems

In regard to verb morphology, the comparison of French and Austrian German is interesting because their oral verbal systems present strong similarities as well as certain differences.

Concerning similarities, the Austrian system is closer to French than Northern-German in certain respects. Both verbal systems have three persons in the singular and plural, but in spoken French the impersonal form (e.g., *on parle* 'one speaks') largely replaces the first plural (*nous parlons* 'we speak'), which is not the case in German. Both systems have non-finite forms of infinitive and past participle, finite modes of indicative, imperative, subjunctive and conditional. Both languages have two subjunctives, both little used in spoken language. As for voice, they both have an analytic passive. They both express the opposition between present, past and future by means of compound past and analytic future. Whereas the analytic category of compound past is widely used in both languages, the synthetic counterparts are largely limited to literary use, i.e., the French simple past (*passé simple*) and the German preterite (widely used in Northern Germany). As in English, both French and German compound past uses two auxiliaries, F. *avoir* = G. *haben* 'to have' and F. *être* = G. *sein* 'to be'. Both languages also have present participles, which are little used in spoken language, with the exception of the French gerund (*en* + participle, e.g., *en parlant* 'speaking'), a category which German lacks. Both languages have just one productive class of verbs into which all recent loan words are integrated, i.e., the French type *parl-er* 'to speak', past participle *parl-é* (cf. *flirt-er* 'to flirt') and the German weak verbs of the type *spiel-en* 'to play', past participle *ge-spiel-t* (cf. *flirt-en*). Both languages have several homophonous forms, e.g., in French the infinitive *parler* is homophonous with the past participle *parlé*; in German the infinitive *spielen* is homophonous with the first and third plural present forms.

However, the verbal systems of spoken French and oral Austrian German, which are relevant for the input of the children investigated, also present certain interesting differences. The category of person is expressed by clitic subject pronouns in French,

whereas German subject pronouns are rarely clitic; it is often expressed by distinctive verb endings in German, rarely so in French. French uses two different past tenses, compound past and (synthetic) imperfect, whereas Austrian German has only its compound past. French has two futures, a synthetic and an analytic one (with the auxiliary *aller* 'to go'), and German only an analytic future (with the auxiliary *werden* 'to become'). Also the conditional is synthetic in French but analytic in German (auxiliary *würde*, past subjunctive of *werden*); only in dialectal Austrian German the synthetic past subjunctive has changed into a conditional. German uses many verb particles which are separable prefixes, a pattern which French lacks.

As far as periphrastic constructions are concerned, French and Austrian German present large similarities, with some differences. Both languages use mainly an auxiliary plus the past participle of the main verb (AUX+PPP in what follows) to form compound past tenses, in particular the compound past, which is the most frequently used for referring to past events. Auxiliaries are *avoir* and *être* in French, and *haben* and *sein* in Austrian German, respectively 'have' and 'be' (e.g., *il a mangé* or *er hat gegessen* 'he has eaten', *il est parti* or *er ist gegangen* 'he has gone'). Moreover, the structures *être* 'be'+PPP and *sein* 'be'+PPP are used for the expression of the state passive (e.g., *il est caché* or *er ist versteckt* 'he is hidden'). In German there exists also the structure *werden* 'become'+PPP for the expression of the event passive, e.g., *das Buch wird gelesen* 'the book is being read'. Both languages use a modal verb plus the infinitive of the main verb (MOD+INF) to form modal constructions, which refer to frequently expressed basic states of mind such as desires, capacities and permission. French uses the verbs *vouloir* 'want', *pouvoir* 'can/may', *devoir* 'must', *savoir* 'know (how)' and *falloir* 'must', which is highly defective, (e.g., *je veux manger* 'I want to eat', *tu peux regarder* 'you can look'). Austrian German uses *wollen* 'want', *mögen/möchten* 'want', *können* 'can/may', *dürfen* 'can/may', *sollen* 'must', *müssen* 'must', (e.g., *ich will essen* 'I want to eat', *du kannst schauen* 'you can look'). Finally, both languages mainly use an auxiliary plus the infinitive of the main verb (AUX+INF) to form the analytic future. French uses the verb *aller* 'go' and German uses *werden* 'become' as the auxiliary (e.g., *je vais manger* 'I am going to eat', *ich werde essen* 'I will eat'). In colloquial Austrian German, there exists also the structure AUX *tun* 'do'+INF for the expression of a present, ongoing event, e.g., *ich tue essen* 'I am eating'.

While periphrastic structures present strong overall similarities, some differences are noticeable. Analytic future is relatively frequent in oral French, which tends to replace the synthetic future by the analytic one, and less frequent in Austrian German, which often uses the present form to express future events. Concerning compound past, morphological complexity in the formation of the past participle is greater in German, which takes in general (except under certain prosodic conditions) a prefix *ge-* in addition to a suffix. Moreover, for past participles of prefixed verbs, there is a further complication: if the prefix is separable, the participle takes, in addition, the prefix *ge-*, and if it is not, it does not take it: for example, *über-setz-en* 'to ferry', 1sg.pres. *ich setze über* (separable prefix), past participle *über-ge-setzt* vs. *über-setz-en* 'to translate', 1sg.pres. *ich über-setz-e* (non-separable prefix), past participle *über-setz-t*. Also, productive types of participles are easier to construct in French than in Austrian German. Finally, the compound past in Austrian German is frequently affected by a syntactic separability and invertibility which separates the auxiliary and past participle (e.g., *er* ***hat***

gestern noch einen Apfel ***gegessen***, literally 'he has yesterday still an apple eaten'). Overall, the German properties of separability and invertibility are a main difference between French and all German periphrastic constructions: in French, the auxiliary or modal verb is in general immediately followed by the participle or infinitive whereas in German they may be separated by one or more syntactic constituents and their order may be inverted.

Hypotheses on the development of periphrastic structures in French and Austrian German

Within this framework, our present study on the emergence and development of periphrastic structures in French and Austrian German was conducted on the basis of three main sets of assumptions or hypotheses.

A first set of hypotheses, derived from the idea that language acquisition relies on the interplay between children's general predispositions and input stimuli, is that children's acquisition of the periphrastic structures depends upon both general cognitive factors and linguistic factors related to the typological or language-specific properties of the verbal system (cf. Bassano *et al.*, 2001*a*, 2001*b*; Dressler, 1997; Slobin, 1997*a*). General constraints should be reflected in similar cross-linguistic developmental patterns, whereas language-specific effects should be reflected in differences in developmental patterns. According to general cognitive factors, analytic future should emerge after compound past in both languages, for reasons of conceptual and semantic complexity (Comrie, 1985; MacWhinney, 1999*b*). In general, verbal forms referring to present events are likely to be more accessible to children than those referring to past and future events, and verbal forms referring to past events more accessible than those referring to future events, which involve a projection in mental representation. According to language-specific effects, the above-mentioned differences between the two languages, most of all concerning the frequency of use of the periphrastic future and the morphological and syntactic complexity of the compound past, should favour these structures in French rather than in Austrian German.

A second set of hypotheses, derived from the idea that linguistic categories are gradually constructed by the child, is that, in both languages, children's acquisition of the periphrastic structures is a gradual process which involves the existence of precursors denoting a continuity in the developmental process (cf. Bassano, 2000; Tomasello, 1995). From this perspective, a likely expectation is that, before producing the adult-like complete periphrastic structures, children should produce incomplete forms, resulting in 'bare past participles' and 'bare infinitives' (without auxiliary or modal preceding the main verb). Bare infinitives have been reported to be produced by young children in a number of languages including German, Dutch and French (but they seem to be rare in Italian or Spanish). However, the question remains controversial whether they reflect the omission of a modal or auxiliary (as we argue here), or whether they have a reduced syntactic representation in which tense is lacking (e.g., the optional infinitive stage hypothesis, Wexler, 1994, 1998). From the perspective of continuity, children could also employ transitional procedures in the course of the verb grammaticization process. A number of recent studies have reported how children incorporate unglossable syllables – the so-called fillers – into their early utterances in quite an array of languages, stressing the role of these elements in emerging grammar

(Kilani-Schoch & Dressler, 2001; Peters, 1997, 2001; Peters & Menn, 1993; Veneziano & Sinclair, 2000). Children could thus be expected to use fillers instead of auxiliaries and modals in front of past participles and infinitives.

Finally, a last set of hypotheses is derived from the functionalist integrative assumption that there exist interdependencies between lexical and grammatical development (cf. Bassano, 2000; Bassano & Eme, 2001; Bates & Goodman, 1999; Marchman & Bates, 1994). We expect relations between the time course of the verb grammaticization process and of the lexical development of verbs. Of course we consider here only one aspect of the verb grammaticization process, namely the aspect related to the acquisition of periphrastic structures. Following Bassano (2000), we use, for assessing this aspect of verb grammaticization, a synthetic index which measures the child's ability to produce correctly complete periphrastic constructions, that is, the ability to use a mandatory grammatical verb before past participles or infinitives. We focus on quantitative relations between the grammatical and the lexical development of verbs, looking particularly at whether there is evidence for the 'critical mass hypothesis', e.g., that a grammatical spurt is preceded by a lexical spurt.

METHOD

Participants

The four participants were two French and two Austrian children (one girl and one boy in each language group). The French girl, Pauline, who was the youngest of four children in a middle-class family living in Rouen, was recorded from the age of 1;2 to 3;0. The French boy, Benjamin, the younger of two children in a middle-class family living in Paris, was recorded from 1;11 to 3;0. The Austrian girl, Katharina, the second of three children in a middle-class family living in Vienna, was recorded from 1;6 to 3;0. The Austrian boy, Jan, the younger of two children in a middle-class family living in Vienna, was recorded from 1;3 to 3;0.

Data collection

Each child was audio- or video-recorded once or twice a month at home, during everyday activities such as eating, playing, washing, dressing, etc., in unstructured interactive sessions with his/her family. Long uninterrupted parts of each recorded session were selected for transcription in order to obtain a variety of situations and a sufficient and representative number of productions. They were transcribed in CHAT format (MacWhinney, 2000), with indications about the situations, contexts and gestures, and then stored on computer.

Data sampling

The analyses were conducted on monthly samples (combining the two sessions per month), each consisting of a constant number of 120 utterances selected from the transcribed sessions. To qualify as an utterance, a production had to be a prosodic and

meaningful unit that included at least one element resembling a word in form and meaning. Babbling, vocalizations and completely incomprehensible strings were part of the child's productions, but were not considered utterances and therefore were not analysed in the study. Working on restricted samples consisting of a constant number of utterances, whatever the session, is a methodological choice that we had adopted elsewhere (e.g., Bassano, 2000; Bassano, Maillochon & Eme, 1998) because it makes it possible to carry out longitudinal studies of relatively long duration, and because it allows maximal (also quantitative) comparability across children and across ages.

Coding and analysis

For each of the four children, all 120 utterances in each monthly sample were submitted to various types of coding. Analyses below were based on a specific coding of verbs (we considered as verbs in child productions those terms which are verbs in adult language). For the Austrian children, morphological coding was done according to the norms of CHILDES (MacWhinney, 2000). For the French children, a specific coding elaborated in previous studies and providing information on semantic, as well as morphological and structural properties of verbs, was applied. A main methodological difficulty in coding verb forms – in particular for French data – was to distinguish between bare infinitives and past participles, which are homophonous for the verbs in the first group in spoken French (e.g., infinitive *jouer* 'to play' and past participle *joué* 'played' are homophonous: /zue/). These ambiguous forms were classified according to functional analysis guided by contextual and situational information provided in transcriptions. For example, in Pauline's data, the utterance /pase/ (1;7) was interpreted as an infinitive with a modal value, i.e., *passer* 'to pass', since the child was clearly expressing her desire to pass through a busy corridor. In contrast, the utterance /uve/ (1;8) was interpreted as a past participle with a resultative value, i.e., *trouvé* 'found', since she was waving a piece of puzzle she had been looking for.

Quantitative analyses of the frequencies of verbs and verb forms relied on the classical distinction between types and tokens. For verbs, the computation of types contained the lexically different verbs (lemmas), whatever their form, whereas the computation of tokens contained all the verbs produced, as many times as they were repeated. For specific verb forms (e.g., past participle, infinitive), types refer to the different lemmas used with this verb form. For the complex verb forms under study, that is periphrastic structures consisting of a grammatical and a main verb, types referred to the different forms of the structure, excluding only strict repetitions (e.g., aux. *avoir*1sg + past participle *mangé* 'eaten' and aux. *avoir*3sg + past participle *mangé* 'eaten' were two types), while tokens included strict repetitions.

In qualitative analyses, two indicators of appearance for a structure were used: time of emergence, corresponding to the age where the first production of a given structure was found in the child's samples, and time of acquisition, corresponding to the age where the structure started to be used productively by the child. Two criteria of acquisition can be used (cf. Pizzuto & Caselli, 1994). The weak criterion requires the production of the same grammatical verb in the same form with at least two distinct main verbs in the same sample (e.g. aux. *avoir*3sg + past participle *mangé* 'eaten' and aux. *avoir*3sg + past participle *cassé* 'broken'). The strong criterion requires the

production of the same main verb in the same form with at least two distinct grammatical verbs in the same sample (e.g., 'modal *pouvoir*1sg + infinitive *manger'* and 'modal *vouloir*1sg + infinitive *manger').*

RESULTS

This section is divided into four main parts. First, a series of initial analyses is designed to place the investigation of periphrastic constructions in children in a more general context, such as the development of MLU and verb production. Then, two series of core analyses focus on (a) when and how periphrastic constructions emerge and develop in both languages, and (b) where they come from, looking in particular at whether they have precursors. These analyses are mainly devoted to examining the two first sets of hypotheses, i.e., that the development of periphrastic constructions is dependent on both general and language-specific factors, and that it is a gradual and continuous process. Finally, a last series of analyses, focusing on the verb grammaticization process drawn from the core analyses, examines the third set of hypotheses, i.e., that the grammaticization process is dependent on the lexical development of verbs.

MLU development and the lexical production of verbs

To compare the four children on the basis of general indicators of language development, we first conducted an initial series of analyses which examined MLU development and the production of verbs. Despite its obvious limitations, MLU remains the most classical index of early language development in general and grammatical development in particular. Indices of the production of verbs are given by numbers of verb types and verb tokens, which are a measure of the children's tendencies to produce this class of words from the lexical point of view.

In regard to MLU (here calculated in words, i.e., free morphemes, and in the raw version, i.e, taking into account all tentative word tokens produced including repetitions), Fig. 1 shows that the index increased in very similar ways in the two French children, Pauline and Benjamin. In both of them, MLU reached a value of around 2 at 2;0, of around 4 at 2;6 and of around 5 at 3;0. The two Austrian children presented more differences. Jan's MLU developed very much like the French children's until 2;6, but did not increase any more thereafter: it reached a value of 2 at 2;0, of around 3.5 at 2;6 and 3.2 at 3;0. Compared with the other three children, Katharina's MLU showed a delay of about four months (value of around 1.3 at 2;0, of around 2.5 at 2;6, and around 2.7 at 3;0), which reflected a delay in the emergence of language. It can be seen, however, that during the last months of the study, from about 2;6 on, the two Austrian children's scores were close to each other, and lower than those of the French children.

With regard to verb production, the developmental tendencies in numbers of verb types (Fig. 2a), as well as in numbers of verb tokens (Fig. 2b), corresponded to those in MLU for the four children. Until about 2;6, the development of verb production was similar for Pauline, Benjamin and Jan, and clearly delayed for Katharina. Thereafter, the two Austrian children's scores were lower than the two French children's.

In addition, a finer-grained analysis of verb-type production (Fig. 2a), which is the

index that most closely measures the verb lexicon size, showed that all four children presented a non-linear increase in production, that is a 'lexical verb spurt'. The timing of this spurt varied somewhat across children, with boys being slightly more precocious than girls: from 1;9 on for Jan, 2;0 for Benjamin, 2;2 for Pauline, and 2;3 for Katharina. Before this point, the number of verb types produced by each of the four children did not change markedly from one monthly sample to the following, while it increased strikingly after this point. It can be noticed that the increase appeared to be particularly sharp for the two children (the girls) who showed their verb spurt a little later in time. In Jan's samples, the number of verb types was 14 at 1;9, 21 at 1;10, 26 at 1;11, and in Benjamin's samples, it was 14 at 2;0, 24 at 2;1, 30 at 2;3. In Pauline's samples, the number of verb types jumped from 13 at 2;2 to 31 at 2;3, and in Katharina's samples, from 2 at 2;3 to 10 at 2;4 and 20 at 2;5.

In summary, two noticeable results emerged from these initial analyses. First, although the four children's scores concerning MLU and verb production increased along time – with individual differences in rate – French children had higher performances than Austrian children in the second part of the third year, in particular for MLU and verb-token production. This contrast is likely to reflect language differences, a hypothesis that will be examined in the discussion. Second, all the four children presented something like a 'lexical verb spurt' in their type-production increase. Variations among children in the moment of occurrence of this spurt do not seem to be related to language differences, but rather reflect individual differences. Further analyses will examine to what extent they could be related to variations concerning other variables, namely the grammatical development of verbs.

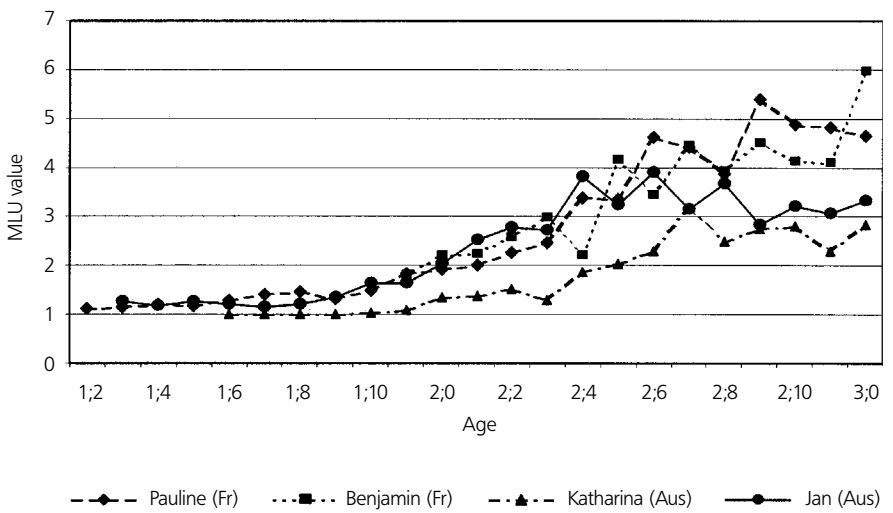

Figure 1 MLU development in the four children

Emergence and development of adult-like periphrastic structures

Turning then to the three periphrastic structures under study, we analysed when and how these structures appeared and developed in the four children's productions. Here we examined the production of the adult-like structures, that is the structures in their complete forms, with both the grammatical and the main verb: auxiliary+past participle (AUX+PPP) mostly for the compound past, modal+infinitive (MOD+INF) for the modal constructions, auxiliary+infinitive (AUX+INF) mostly for the periphrastic future. Appendix 1 provides analyses of type and token frequencies of the three structures in

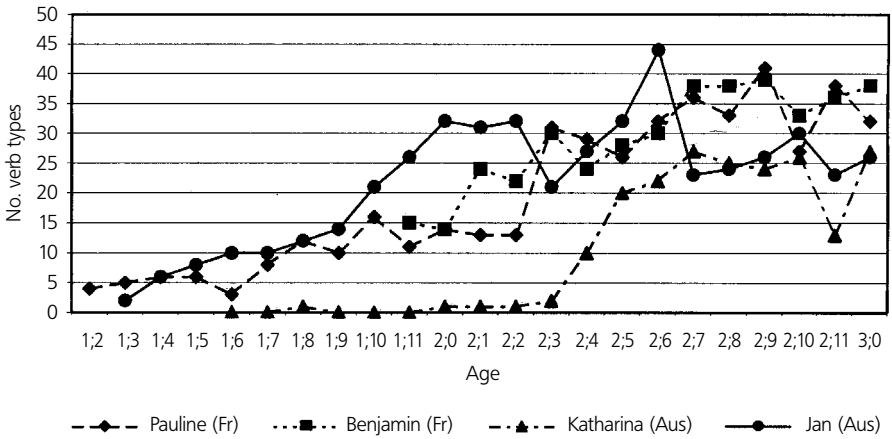

Figure 2a Development of verb-type production in the four children

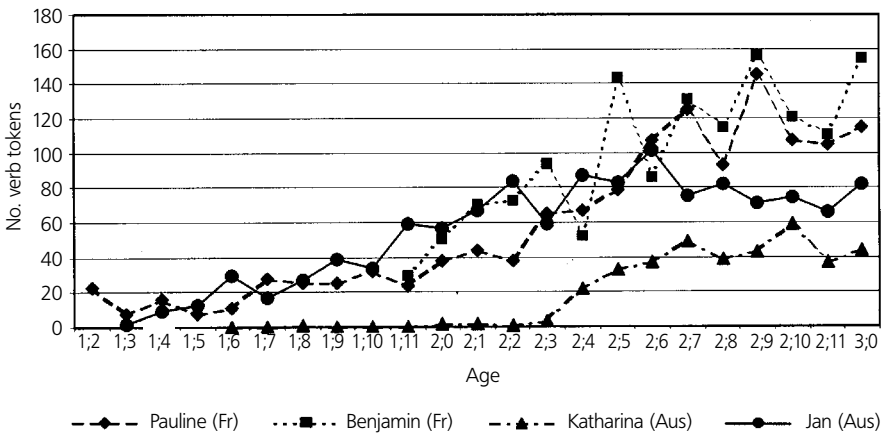

Figure 2b Development of verb-token production in the four children

each of the four children, additionally distinguishing how the different auxiliaries and modal verbs as well as their different tense forms were used by the child.

Periphrastic structures in the French children

In both French children, Pauline and Benjamin (see Fig. 3), the complete periphrastic structures started to be produced from around 2;0 on, and all three emerged in a relatively short range of time, i.e., three or four months. However, we noted differences in time of emergence of the three structures, identical for both children: AUX+PPP appeared first, followed by MOD+INF, and last by AUX+INF. Benjamin produced his first AUX+PPP at 1;11 (e.g., *est fini* 'is finished'), his first MOD+INF at 2;0 (*peux le fai(re)* 'can do it'), and his first AUX+INF at 2;1 (*va venir Elsa* 'is going to come Elsa'). Pauline started a little later, producing her first AUX+PPP at 2;0 (*/e/ porte est cassée* '/fill/ door is broken'), her first MOD+INF at 2;1 (*tu peux me tailler?* 'you can sharpen for me?), and her first AUX+INF at 2;3 (*on va se pro(me)ner* 'we are going to walk')

As can also be seen in Fig. 3, frequency analyses showed similar hierarchies in both children. The most frequent structure in type numbers as well as in token numbers was AUX+PPP, followed by MOD+INF, and then by AUX+INF. During the whole period, Benjamin produced 100 types and 141 tokens of AUX+PPP, 41 types and 57 tokens of MOD+INF, 35 types and 42 tokens of AUX+INF. Pauline produced 74 types and 95 tokens of AUX+PPP, 50 types and 64 tokens of MOD+INF, 33 types and 39 tokens of AUX+INF. As can be seen from these numbers, differences in frequencies go hand in hand with differences in time of emergence: the earlier a structure appeared, the more frequent it was in terms of type and token production. Although the general developmental patterns were strikingly similar in the two children, individual differences could be noted: in particular, the three structures emerged a little later in Pauline than in Benjamin, and Benjamin produced more AUX+PPP structures than Pauline did.

With regard to time of acquisition, differences between the three structures and between the two children seemed to be neutralized, although AUX+PPP was acquired clearly earlier in one of the children. If we consider the weak criterion of acquisition, that is, the production of the same grammatical verb with two distinct main verbs in the same sample, this criterion was surely reached by Benjamin at 2;0 for AUX+PPP (*a cassé* 'has broken', *a coulé* 'has poured'), and at 2;5 for MOD+INF (e.g., *on peut prendre* 'we can take', *on peut manger* 'we can eat') as well as for AUX+INF (*vais faire* 'am going to do', *vais te montrer* 'am going to show you'). It was reached by Pauline at 2;5 for AUX+PPP (*j'ai vidé* 'I have emptied', *j'ai renversé* 'I have spilled'), as well as for MOD+INF (*il faut couper* 'we must cut', *il faut ranger* 'we must range') and for AUX+INF (*on va mettre la couche* 'we are going to put the nappy', *on va piquer* 'we are going to sting'). Thus, except for the earlier AUX+PPP in Benjamin's data, time of acquisition was around 2;5 for all periphrastic structures in both children.

Qualitative analyses (see Appendixes 1a and 1b) show how a diversification of form use occurs in grammatical verbs. During a first period in both children, the typical auxiliary in the AUX+PPP structure was *être*3sg, a trend which fits to the idea that children's very first uses of past tense generally have a resultative value, associated with movement or change-of-state. After the acquisition criterion was reached, the auxiliary *avoir* started to be produced and progressively became more frequent than the

auxiliary *être*, showing the highest overall token frequency in both children (in Pauline 77%, in Benjamin 57% of all AUX+PPP tokens). This development is in conformity with the frequency patterns of these two auxiliaries in French adult language, in which the large majority of the verbs have their compound tenses formed with *avoir*. The range of the different forms used by the children in the auxiliaries increased: although the 1sg and 3sg remained the most frequent, the 2sg and 3pl were also found, as well as some imperfect and future forms. The first modal verbs in the MOD+INF structure were *vouloir*1sg and *pouvoir*1/2/3sg, and the auxiliary in the AUX+INF structure was *aller*1/3sg. Afterwards, the modal verbs *vouloir* and *pouvoir* remained frequent in the MOD+INF structure, but *falloir*3sg (*il faut mettre* 'we must put') and *savoir*1sg (*je sais faire* 'I know how to do') were also often produced.

The diversification process was also marked by the production of formal contrasts in the same monthly sample, that is, the production of different forms of the auxiliary or modal with the same main verb. In Benjamin's samples, the first contrast was found at 2;5 in the AUX+PPP structure (*c'est caché* 'it is hidden', *je suis caché* 'I am hidden') and in the MOD+INF structure (*peux pas mettre* 'cannot put', *on peut mettre* 'we can put', *je veux mettre* 'I want to put'), and at 2;9 in the AUX+INF structure (*je vais chercher* 'I am going to look for', *papa va chercher* 'daddy is going to look for'). In Pauline's samples, the first contrast was found at 2;5 in the AUX+INF structure (*vais manger* 'am going to eat', *on va la manger* 'we are going to eat it'), at 2;6 in the AUX+PPP structure (*j'ai compté* 'I have counted', *tu as compté* 'you have counted'), and at 2;7 in the MOD+INF structure (*il faut faire* 'we must do', *je sais pas faire* 'I don't know how to do'). A total of about 10 and 15 formal contrasts concerning periphrastic structures were found in Benjamin's and Pauline's samples, respectively.

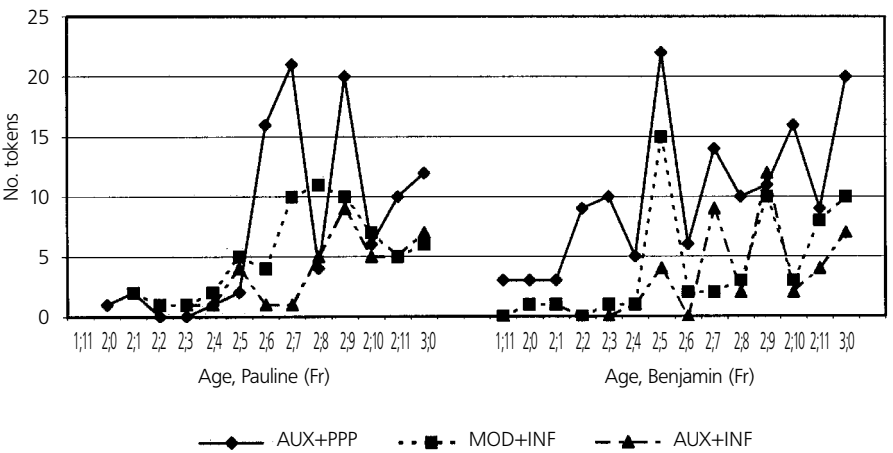

Figure 3 Development of the three periphrastic structures (token production) in the two French children

Periphrastic structures in the Austrian children

In the Austrian boy Jan (see Fig. 4), as in the French children, periphrastic structures emerged from 2;0 on; in the Austrian girl Katharina, they emerged from 2;4 on, in correspondence with her general delay. In both Austrian children, the range of time for the emergence of the three structures was also three months. In both Austrian children, the first structure to appear was not the AUX+PPP, but the MOD+INF structure. Jan produced his first MOD+INF at 2;0 (*kann s(ch)auk(e)ln* 'can swing'), his first AUX+INF at 2;1 (*tun s(ch)laf(e)n* 'are sleeping'), and his first AUX+PPP at 2;2 (*is(t) versteckt* 'is hidden'). Katharina started four months later, producing her first MOD+INF at 2;4 (*kann a(n)zieh(e)n* 'can put on'), her first AUX+PPP at 2;5 (*hab(e) ich (g)e(s)pritzt* 'I have sprinkled'), and her first AUX+INF at 2;6 (*wer(de)n wir seh(e)n* 'we will see').

Frequency analyses (Fig. 4) showed a considerably higher number of types and tokens for the three structures in Jan than in Katharina, in correspondence with her general delay. In Jan, AUX+PPP and MOD+INF were at about the same frequency level (to the advantage of MOD+INF until 2;8, however); in Katharina, AUX+PPP was the most frequent structure. AUX+INF was the least frequent structure in both children. During the whole period, Jan produced 59 types and 75 tokens of MOD+INF, 52 types and 62 tokens of AUX+PPP, 7 types and 13 tokens of AUX+INF. Katharina produced 34 types and 44 tokens of AUX+PPP, 20 types and 23 tokens of MOD+INF, 2 types and 4 tokens of AUX+INF.

With regard to time of acquisition, the reverse order of time of emergence was observed for AUX+PPP and MOD+INF: our weak criterion of acquisition (see above) was fulfilled earlier for AUX+PPP than for MOD+INF in both Austrian children. It was reached by Jan at 2;2 for AUX+PPP (*is(t) versteckt* 'is hidden', *is(t) anges(ch)nallt* 'is strapped on'), and at 2;5 for MOD+INF (*darf spiel(e)n* 'may play', *darf hau(e)n* 'may hit'); by Katharina at 2;7 for AUX+PPP (*ich hab(e) gesagt* 'I have said', *ich hab(e) zugedeckt* 'I have covered), and at 2;9 for MOD+INF (*darf ich kommen* 'I may come', *darf ich anrufen?* 'may I call?'). For AUX+INF, the acquisition criterion was not fulfilled in either of the two Austrian children. Thus, there was a clear difference in time of acquisition between the three structures: AUX+PPP was acquired first, followed by MOD+INF; there was no clear evidence for the acquisition of AUX+INF until the end of the observed period.

Qualitative analyses (see Appendixes 1c and 1d) show that, in AUX+PPP, both auxiliaries *sein* and *haben* were used by the two Austrian children from early on, with *haben* showing the highest overall token frequency in both children (in Jan 71%, in Katharina 75% of all AUX+PPP tokens). The structure *werden* + PPP (i.e., event passive) was limited to two single (but surprisingly early) occurrences in Jan, and was not found at all in Katharina. At the beginning, the typical form of auxiliaries in AUX+PPP was 1sg and 3sg for both children; from 2;6 onwards in Jan and from 2;10 onwards in Katharina, a diversification of form use (2sg, 1/3pl and pret3sg) was observed. The first modal verb in MOD+INF for both children was *können,* but soon also other modals were used (e.g., Jan: *wollen, mögen, sollen, müssen*, Katharina: *müssen, mögen*); as to overall token frequency, *können* was most frequent in both children (in Jan 41% and in Katharina 48% of all MOD+INF tokens). The typical initial form of modals in MOD+INF was 3sg in Jan and 1/3sg in Katharina; from 2;2 in Jan and from 2;9 in Katharina, an

extension of form use (2sg, 1/3pl, pret3sg/pl) took place. In AUX+INF, Jan started (at 2;1) with the structure *tun*+INF which denotes an ongoing present event; *werden*+INF (i.e., future tense) emerged much later. In Katharina, AUX+INF was limited to the single formulaic structure *werden*+INF *sehen* 'will see'.

The first examples of formal contrasts (see above) were found in Jan's samples at 2;2 in the MOD+INF structure (*muss steh(e)n bleib(e)n* 'must stop', *können steh(e)n bleib(e)n* 'can stop'), and at 2;4 in the AUX+PPP structure (*ich bin gefahr(e)n* 'I have driven', **hat (= ist) gefahr(e)n* 'has driven'). The first and only example in the AUX+INF structure was found at 3;0 (*ich werd(e) da wohnen* 'I will live here', *ich würd(e) da wohnen* 'I would live here'). In Katharina's samples, the first contrast was found at 2;7 in the AUX+PPP structure (*hab(e) ich *gehaltet (= gehalten)* 'I have held', *hat gehalten* 'has held'), and at 2;9 in the MOD+INF structure (*darf ich anrufen?* 'may I call?', *musst du anrufen* 'you must call'). The first and only example in the AUX+INF structure (**wir (= wirst) du seh(e)n* 'you will see', *werma seh(e)n* 'we will see') was found at 2;9. A total of 13 and 5 formal contrasts concerning periphrastic structures were found in Jan's and Katharina's samples, respectively.

In summary, the emergence and development of adult-like periphrastic structures showed strong cross-linguistic similarities. In both languages, these structures emerged from around 2;0 onwards, all three emerged in a range of time of about three months (with little variation between children), and the last to appear was the AUX+INF structure. They developed following a diversification process marked by the use of more and more main verb types and grammatical verb forms. However, cross-linguistic differences were also found. Obvious quantitative discrepancies concerned the frequencies of the structures, which were in general less frequent in Austrian than in French children. Another more specific difference concerned the AUX+PPP structure,

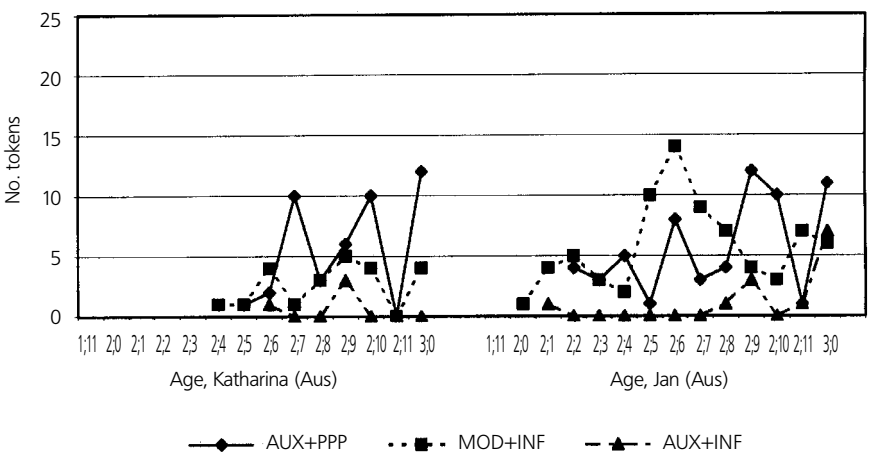

Figure 4 Development of the three periphrastic structures (token production) in the two Austrian children

that, contrary to French children, Austrian children did not produce as the first periphrastic structure, although it was the first to fulfil the criterion of acquisition. We will examine in the discussion the extent to which these similarities and differences can be related to general and language-specific factors.

Looking for precursors of periphrastic structures

Continuing with the analysis of periphrastic structures, we then examined whether these structures emerged *ex nihilo* or whether we could find precursors denoting a continuity in the developmental process. From this perspective, we identified all the forms where past participle and infinitive were employed – in addition to the strict AUX+PPP, MOD+INF and AUX+INF analysed above – in the four children's productions, and we examined which of these forms could be candidates for being precursors of the periphrastic structures.

The different configurations including past participle and infinitive are presented in Appendix 2, with numbers of tokens produced by each of the four children in the monthly samples. As can be seen from these tables, in both languages, children frequently produced incorrect 'bare past participles' (*0+PPP) and 'bare infinitives' (*0+INF), that is past participles and infinitives alone, without any preceding auxiliary, modal or other verb. Incorrect uses referred to productions that were not grammatical, since an auxiliary, modal or other verb would be required in adult language. Examples of incorrect bare past participles were, in French, *fini* 'finished', *oh cassé* 'broken', *moi déjà mangé* 'me already eaten', and in Austrian German, *zugangen (= zugegangen)* 'closed', *die Windel angepatzt* 'the diaper stained', *Pauli Kaninchen auch aufgewacht* 'Pauli (the) rabbit also awaked'. Examples of incorrect bare infinitives were, in French, *regarder* 'look', *habiller bébé* 'dress baby', *mette (= mettre) là* 'put there', *c'est moi qui ranger* 'it's me who to range', and in Austrian German, *asaun (= anschauen)* 'look at', *brrm machen* 'make brrm', *selber saukln (= schaukeln)* 'swing (my)self', *Auto Darase (= Garage) naustommen (= hinaus/rauskommen)* 'car garage come out'.

Some correct or admissible uses of bare past participles (0+PPP correct) and infinitives (0+INF correct) had to be distinguished from truly incorrect uses. Correct uses corresponded to certain specific structures where these forms can be employed alone in adult language. For example, in French, the bare past participle is correct in some adjectival structures (*une de tombée* 'one fallen'). Bare infinitive is correct when it is used with a subject function (e.g., *courir est interdit* 'to run is forbidden'), but these uses are not found in child language and are quite rare in adult language. The same holds for German bare infinitives. In addition, orders can be given in the bare infinitive, e.g., *bitte einsteigen!* 'please get in!', *nicht fallenlassen!* 'don't drop it!'. Other uses considered as correct corresponded to those particular cases of speech situation when the child is invited to answer or to continue an utterance where the auxiliary or the modal verb (or a 'light' verb, e.g., German *machen* 'make') has been produced by the interlocutor. For example,

1. Mot: *Tu voulais pas quoi?* 'what did you not want?'
   Paul: *pas manger* 'not to eat',
2. Mot: *Was machen denn die Kinder da auf dem Spielplatz?* 'What are the kids doing here on the playground?'
   Jan: *spiel(e)n* 'playing'.

Another configuration of non-canonical past participle or infinitive use was found in the present French children's data, but not in the Austrian data. It consisted of filler syllables preceding past participle (FILL+PPP) or infinitives (FILL+INF), the filler being likely to prefigure the grammatical verb required before the main verb. Examples of 'filler+past participle' were */da/ touvé (= j'ai trouvé)* '/filler/ found', */lé/ fini (= j'ai fini)* '/filler/ finished'. Examples of 'filler+infinitive' were */eum/ pas percer (= je veux pas percer)* '/filler/ not got a hole', */a/ sauter* '/filler/ jump', */è/ monter* '/filler/ climb'. In fact, the 'filler+past participle' structures were infrequent, and the 'filler+infinitive' structures were not very frequent. Two comments have to be added with respect to the sequences counted in these configurations. First, it was difficult to distinguish preverbal fillers in place of the grammatical verb from preverbal fillers in place of the subject pronoun clitic. Particularly in the 'filler+past participle' structures, the filler is likely to be a placeholder for both subject clitic and auxiliary (see above, */da/ trouvé* or */lé/ fini).* But note that the auxiliary always has a syllabic nucleus whereas the clitic pronoun often has not. Thus, the filler is a prosodic placeholder for the auxiliary rather than for the clitic. Only fillers immediately preceding a past participle or infinitive were counted. We considered that the filler was likely to be in place of the grammatical verb in productions where the grammatical verb was not expressed in any other way, and, *a fortiori*, in productions where the subject was expressed in another way (e.g., *moi /eu/ faire comme (je) veux moi* 'me /filler/ to do as (I) want, me'). Second, incomprehensible syllables (signalled by xxx in the transcription, mainly found in Benjamin's corpus) immediately preceding past participles or infinitives were considered as fillers and included in the count of the FILL+PPP or FILL+INF categories.

Finally, a last group of configurations using the infinitive (OTHER+INF) was found in both French and Austrian children. We included in this category the correct constructions produced with the verb in the infinitive following another tensed verb (distinct from the auxiliaries and modal verbs mentioned above), with or without a preposition. Examples of such structures were, in French: *a un mouchoir pour essuyer les larmes* 'has a handkerchief to wipe tears', *j'ai arrêté de pleurer* 'I stopped crying', *tu as appris à baigner, papa?* 'you learned to bath, daddy', *viens manger* 'come to eat', *maman, elle a fait tomber l'assiette* 'mummy, she made the plate fall', *il est en train de téléphoner* 'he is phoning'. Examples in Austrian German were: *alle geh(e)n s(ch)laf(e)n* 'all go to sleep', *reden *macht (= tut) er* 'he is speaking', *der bleibt steh(e)n* 'he stops', *ich lass(e) es wirklich da steh(e)n* 'I really leave it here'. For French children, the OTHER+INF category also included infinitives preceded by a preposition only, as we considered them not to be bare infinitives because of the preposition. These cases concerned some well-admitted expressions like *à boire* '(I want) to drink', produced early by children, and some incomplete sentences where the governing verb was lacking (e.g., *avant de dire non* 'before saying no').

Among these configurations, some appeared as good candidates for being precursors of the periphrastic structures. Both incorrect 'bare past participle' in the two languages and 'filler+past participle' in French were likely to be precursors of the AUX+PPP structure. Similarly, incorrect 'bare infinitive' in the two languages and 'filler+infinitive' in French were likely to be precursors of MOD+INF or AUX+INF structures. Strong arguments for considering these forms to be precursors of adult-like periphrastic structures were provided by the developmental patterns presented in the

children's productions. For obvious reasons, these patterns were particularly apparent in the two children with early data, Pauline for French and Jan for Austrian German, whose data are used in the following figures.

Figure 5 shows the developmental patterns for the AUX+PPP structure in Pauline and Jan, respectively, presenting evolutions in token frequencies for incorrect bare past participle, 'filler+past participle', and complete 'auxiliary+past participle'. As can be seen in Pauline's data, bare past participles were produced much earlier than complete forms and decreased thereafter, disappearing when complete forms increased strongly. There were few 'filler+past participles', but they also disappeared when the complete forms increased. However, for some months the three forms coexisted in the child's productions. Although Jan did not produce 'filler+past participle' configurations, similar developmental patterns appeared in his data, in which the competitive decrease of early bare past participles and increase of complete 'auxiliary+past participle' structures were obvious.

Figure 6 shows the developmental patterns for the MOD+INF and AUX+INF structures, in Pauline and Jan, respectively. Evolutions in token frequencies are presented for incorrect bare infinitive, 'filler+infinitive', and complete 'modal+infinitive' and 'auxiliary+infinitive' (cumulating the frequencies of the two structures). Again we find that in Pauline's data bare infinitives and 'filler+infinitives' were produced earlier than adult-like periphrastic structures, and that they decreased and disappeared when the complete structures increased markedly. The competitive decrease of early bare infinitives and increase of adult-like complete structures were striking also in Jan's corpus.

In summary, these developmental analyses suggest that the emergence of the periphrastic structures around 2 years of age in child language was prepared by

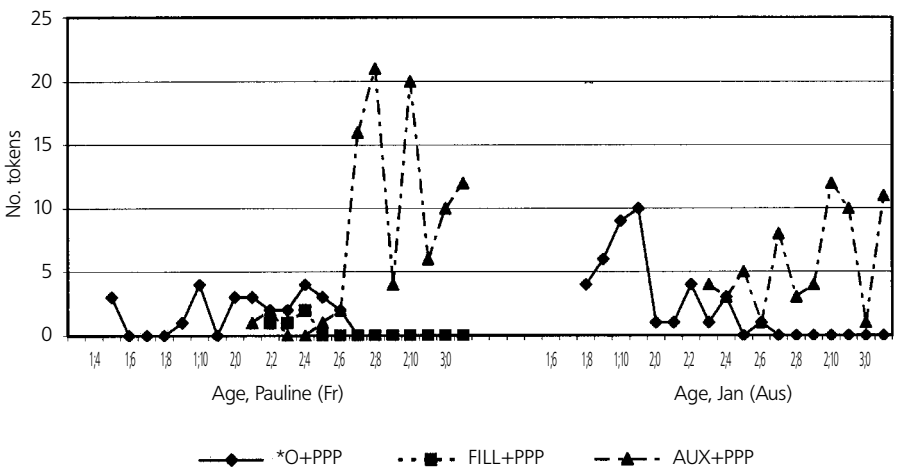

Figure 5 The 'AUX+PPP' structure and its precursors in Pauline's and Jan's data

precursors which consist in the early production of incorrect bare past participles and bare infinitives by both the French and the Austrian children, and in the early production of transitional forms using fillers preceding main verbs by French children. The implications of our present interpretation of these forms (incorrect bare past participles and infinitives, in particular) for early language will be discussed in more detail below, as well as the extent to which they support the view that the acquisition process is gradual in both languages under study.

Indices of verb grammaticization and their relations to the lexical production of verbs

Using the above results, a final series of analyses was designed for calculating indices of verb grammaticization concerning the acquisition of periphrastic structures, and for examining the relations between the grammaticization process and the lexical development of verbs.

Three indices were calculated for each of the four children and are presented in Appendix 3a for the French children, and Appendix 3b for the Austrian children. The 'AUX+PPP' index measures the child's ability to use a mandatory auxiliary before past participles (AUX+PPP monthly value = number of AUX+PPP divided by the number of PPP requiring an auxiliary, that is *0+PPP plus FILL+PPP plus AUX+PPP). The 'MOD/AUX+INF' index measures the child's ability to use a mandatory auxiliary or modal before infinitives (MOD/AUX+INF monthly value = number of MOD/AUX+INF divided by the number of INF requiring a modal or auxiliary, that is *0+INF plus FILL+INF plus MOD/AUX+INF). Finally, the 'AUX/MOD+PPP/INF' is a global index which integrates the two previous ones: it measures the child's ability to produce complete periphrastic

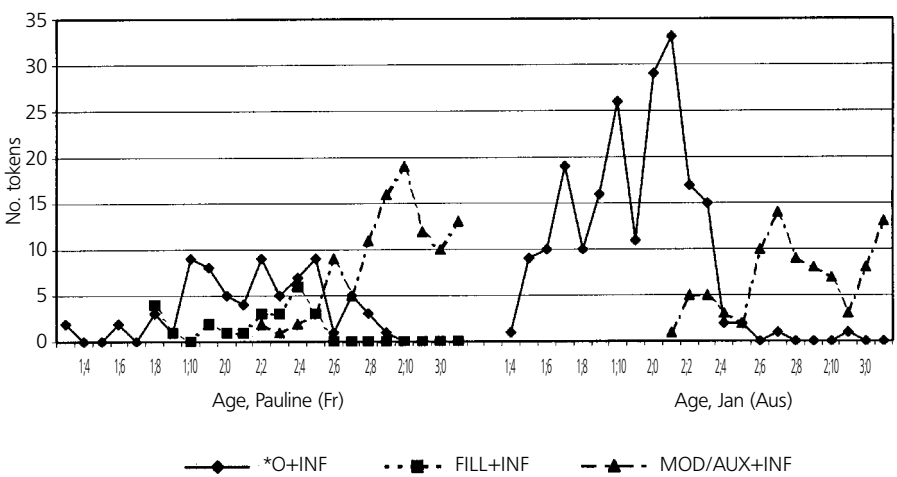

Figure 6 The 'MOD/AUX+INF' structure and its precursors in Pauline's and Jan's data

constructions correctly, that is the ability to use a mandatory preverbal auxiliary or modal verb before past participles or infinitives (AUX/MOD+PPP/INF monthly value = number of AUX+PPP plus MOD/AUX+INF divided by the number of *0+PPP plus FILL+PPP plus AUX+PPP plus *0+INF plus FILL+INF plus MOD/AUX+INF). It should be noted that: (1) given that the indices measure the child's ability to use auxiliaries and modal verbs in contexts where they are required, correct bare past participles and infinitives, as well as correct other uses of infinitives, were excluded from computation; and (2) that strict versions of the indices were calculated here, taking into account true auxiliaries and modal verbs only (and not fillers) for complete structures.

Analyses reported in Appendixes 3a and 3b make it possible to examine separately the development of the 'AUX+PPP' and 'MOD/AUX+INF' indices. In the French children, the two indices developed relatively close to each other, particularly after the age of 2;6. However, the 'AUX+PPP' index was generally higher than the 'MOD/AUX+INF' index. In Pauline's data, the 'AUX+PPP' index was overtaken by the 'MOD/AUX+INF' in three age points only, at 2;2, 2;3 and 2;5. It steadily reached the 1.00 value at 2;6, while the 'MOD/AUX+INF' index reached the 1.00 value at 2;9. The 'AUX+PPP' index advantage was even more clear-cut in Benjamin's data, in which it was overtaken by the 'mod/aux-INF' index in two age points only, at 2;5 and 2;10. Similar tendencies were found in the Austrian children, in which the two indices also developed relatively close to each other, with advantage to the 'AUX+PPP' index. In Katharina's data, the 'AUX+PPP' index was overtaken by the 'MOD/AUX+INF' index in one age point only, at 2;4, and reached the 1.00 value at 2;6. In Jan's data, it also reached the 1.00 value at 2;6; before this point, it was overtaken by the 'MOD/AUX+INF' at four age points, 2;0, 2.1, 2;3 and 2;5. Thus, the relation between the two indices appeared more competitive in Jan's data before 2;6 than with the other children.

The development of the global index 'AUX/MOD+PPP/INF', shown in Fig. 7 for each of the four children, gives a synthetic view of the verb grammaticization process. The

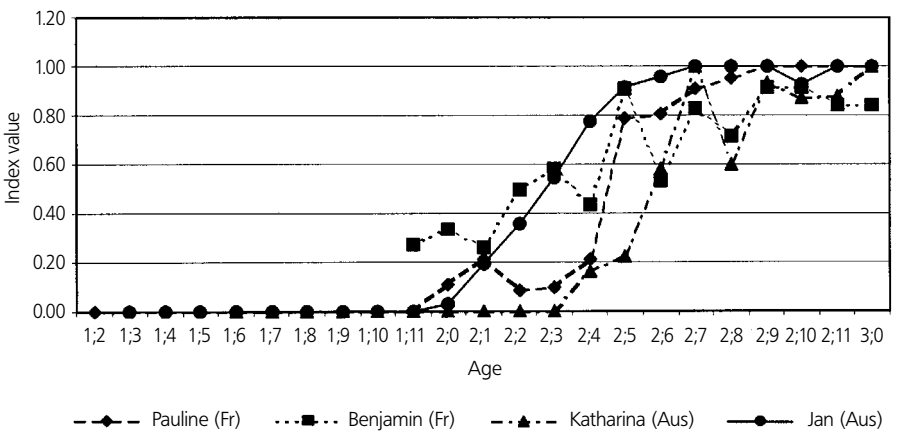

Figure 7 Development of the global verb gramaticization index in the four children

most striking result was that the index showed an explosion in verb grammaticization for the four children to various extents. The explosion was particularly clear-cut with the girls. Pauline's global index had a low value until the age of 2;4, then increased sharply between 2;4 and 2;5 (from 0.21 to 0.79) and reached the maximal value of 1.00 at 2;9. Katharina's index had a low value until the age of 2;5, and increased sharply between 2;5 and 2;7 (from 0.22 to 1.00, although with an up and down evolution afterwards). The boys' grammatical spurt started somewhat earlier, and seemed to take more time. For both of them, the index increase occurred between 2;1 and 2;5, evolving from 0.26 to 0.91 in Benjamin, and from 0.19 to 0.92 in Jan.

A further result of particular interest emerged when we interrelated the lexical and the grammatical development of verbs across time. In the initial analyses, we noticed that all four children presented a lexical verb spurt. Now, we see that in each of the four children the grammatical explosion followed the lexical spurt with some delay. Jan had the lexical verb increase from 1;9 on, and the grammatical explosion from 2;1 on. With Benjamin, the two 'spurt points' were at 2;0 and 2;1 respectively, with Pauline they were at 2;2 and 2;4, and with Katharina they were at 2;3 and 2;5. It can be seen from these temporal relations that the earlier the verb-lexicon increase occurred in a child, the earlier the verb-grammar explosion occurred, following the lexical increase with a delay of about three months. In addition, similar to what we noticed in regard to the verb-lexicon increase, the verb-grammar explosion appeared to be particularly sharp for two children – the girls – who showed their explosion later, as if they made up for lost time.

In summary, the following results emerged from the analyses of indices assessing the children's ability to use a mandatory auxiliary or modal verb in periphrastic structures. First, an explosion in this domain of the verb grammaticization process was evidenced in all four children, with some individual variations in timing and rhythm. Second, in each of the four children, the grammatical explosion followed the lexical spurt with a delay of a few months, and the two processes showed similar developmental rhythms. From these relations it appears that the verb grammaticization process was related to – and, more precisely, dependent on – the development of the lexical verb production in all four children.

DISCUSSION AND CONCLUSION

In this study, we investigated the emergence and development of periphrastic constructions in two French- and two Austrian German-speaking children. In both languages, relevant periphrastic constructions consisted of the 'auxiliary + past participle', 'modal + infinitive' and 'auxiliary + infinitive' structures, mainly corresponding to the compound past tenses, the modal constructions and the analytic future, respectively. Although the four children's corpora started to be collected at different ages (ranging from 1;2 for Pauline and 1;3 for Jan to 1;6 for Katharina and 1;11 for Benjamin) for reasons that we could not control, we tried to maximize the data's comparability. The analyses were conducted until 3;0 for all the children, and they were carried out on monthly samples equalized in number of utterances. From these analyses, it appeared that among the four children the Austrian girl Katharina, although she stayed in the normal range of language acquisition, presented a delay of some

months in the emergence of language, as compared with the other three children (the delay was apparent in MLU, as well as in the production of verbs and of periphrastic structures). Our investigation was designed to shed light on the acquisition of verb grammar by analysing when and how children become able to produce complex structures composed of both the preceding grammatical verb and the past participle or infinitive of the main verb. The discussion below focuses on three main questions resuming our hypotheses about this aspect of the verb grammaticization process: (1) the impact of general (cognitive or linguistic) factors and language-specific effects; (2) the hypothesis of a gradual and continuous acquisition process; (3) the hypothesis of inter-relations between the grammaticization process and the lexical development of verbs.

General and language-specific factors in acquisition

The question of the impact of general – cognitive or linguistic – factors and of language-specific effects was examined by looking for the most striking similarities and differences between the two languages, with the hypothesis that clear similarities are likely to reflect relatively general constraints on acquisition, while clear differences are likely to be related to language-specific effects. We considered as cross-linguistic similarities the similarities appearing in the four children, taking into account the Austrian girl's delay, and we considered as cross-linguistic differences the differences between the two French vs. the two Austrian children, tending not to confound inter-individual and inter-linguistic variations.

Striking cross-linguistic similarities were found in the development of periphrastic constructions in French and Austrian German. This was in line with the results of our previous studies restricted to analyses of the two girls' data (Bassano *et al.*, 2001*a*, 2001*b*; Klampfer *et al.*, 1999), in which strong similarities in order of acquisition of verb forms were observed, despite Katharina's delay. In both children, indicative and imperative present were the first tenses to be produced, appearing earlier than infinitive, past participle, imperfect or preterite and future; the singular forms emerged earlier than the plural ones in all tenses, and the 1st and 3rd persons emerged earlier than the 2nd in indicative tenses. The anteriority of present, singular, and 1st and 3rd persons, which were found in other French and German-speaking children (e.g., Kilani-Schoch, De Marco, Christofidou, Vassilakou, Vollmann & Dressler, 1997) as well as in other languages (Kilani-Schoch *et al.*, 1997; Pizzuto & Caselli, 1994; Tomasello, 1992), are likely to reflect general cognitive or linguistic constraints, such as conceptual complexity and semantic markedness.

In the present study, the four children's data also showed similarities between French and Austrian German in regard to the development of periphrastic structures. In both languages, these structures emerged from around 2;0 onwards, and all three emerged in a range of time of about three months, with little variation between children. In the four children, all three developed following a diversification process which combined two dimensions: the use of a wider and wider variety of types in main verbs, and of types and forms in grammatical verbs. The 'weak criterion of acquisition' (the production of the same form of a grammatical verb with two distinct main verbs in the same session), derived from the first dimension of the diversification process, denoted that

the child used the structures in a productive way and not as frozen expressions. In the four children, this criterion was generally fulfilled a few months after the point of emergence for each structure (although with variations according to the structure, and most often around 2;5 in the French children). However, it can be thought that this criterion was relatively constraining as it was used here, given that it was applied to a limited sample of utterances per session for each child, and that it could be reached earlier in the case of larger samples. Finally, a striking and more specific similarity emerging from our present data concerned the analytic future, which was the last to appear and the least frequent in both languages. This result is consistent with the hypothesis of a general greater cognitive complexity of the notion of future in comparison to past. Due to such different degrees of complexity, on the one hand languages generally have greater inflectional richness for expressing past than future, and on the other hand future tenses generally emerged later than past tenses in all languages. Moreover, our data indicate that also the analytic future appeared later than modal constructions, thus suggesting a greater complexity of the notion of future in comparison to modality. This result, which was not particularly expected, can be explained by the fact that the first modal constructions produced in both languages expressed very basic states of mind, such as desires and abilities.

In regard to cross-linguistic differences, a series of quantitative discrepancies between French and Austrian children were observed. French children presented higher performances in MLU and in the production of verbs (verb tokens, in particular) during the second part of the third year, as well as higher performances in the production of all three periphrastic structures. Differences in MLU can be explained in various ways. The advantage of French children during this period could be due to the systematic production of some specific and very frequent constructions, namely dislocations, which involve additional pronouns (e.g., ***ça c'****est un chien* 'that, this is a dog', ***moi je*** *veux des fraises* 'me, I want strawberries') and often additional verbs (e.g., *c'****est*** *moi qui* ***fais*** 'It is me who does'). Moreover, Austrian children produce a fair number of nominal compounds. Such a compound counts as one word, whereas the usual French counterpart is a multi-word construction (e.g., German *Eisenbahn* = French *chemin de fer)*. Both the language-specific morphological properties of German vs. French and the way of calculating MLU (in words, i.e., free morphemes) would be thus responsible for differences in MLU. Differences in number of verb tokens may be due to the production of dislocations involving additional verbs by French children. In addition, it may be noted that we found more tokens of the very frequent verb 'be' in the French corpus than in the Austrian corpus, although this could be an artefact of Benjamin's production (Pauline: 246 and Benjamin: 397 tokens of *être* vs. Jan: 227 and Katharina: 94 tokens of *sein*). Finally, the quantitative differences between French and Austrian children in the production of periphrastic structures, which were in line with the difference in verb token production, are not easy to account for, except for the AUX+INF structure. This difference clearly reflects the much more frequent use of the periphrastic future in French than in German, where quite often the present is used for expressing future.

A more specific and subtle difference observed in the children's production of the periphrastic structures concerned the compound past, which was the first to emerge in the French children, as well as to fulfil the criterion of acquisition. In contrast, the Austrian children did not produce the compound past as the first periphrastic structure,

although it was the first to fulfil the criterion of acquisition. The difference in emergence fits the hypothesis that the compound past is more complex in Austrian German than in French for various interactive reasons: difficulties in forming the past participle, syntactic separability and invertibility, and lack of perceptual salience of auxiliaries, as compared with modals. These factors are more important than what Wittek & Tomasello (2002) call the greater consistency of the German compound past than of the modal constructions: there are only two auxiliary verbs in compound past constructions but a much greater choice among modal verbs in modal constructions, and this should favour the acquisition of the compound past. The same holds for French, but in both languages modal verbs have a consistency of their own: they are semantically consistent and have their typical syntactic properties (similar to English); finally their paradigms exhibit many morphological similarities. What makes the main difference in German between compound past and modal constructions is less saliency of the auxiliary and the complexity of past participle formation.

In general, these data show that periphrastic structures are verb forms of relatively late acquisition in both languages. This is not surprising since they are complex structures, from the conceptual-semantic as well as from the morphological point of view. From the conceptual point of view, they typically involve references to past, modal or future events that are generally less easy to grasp by children than simple present events (the most difficult being future). From the morphological point of view, they automatically involve the ability to use a free grammatical morpheme before the main verb, which implies a steady two or more word-utterance production. It can be noticed, however, that these compound forms are not produced later than bound-inflected forms denoting past or future, such as imperfect or simple future in French (Bassano, 2000; Bassano *et al.*, 2001*b*). In the Natural Morphology framework distinction between pre-, proto- and modularized morphology phases, all these forms are likely to develop from the proto-morphological phase onwards. While the pre-morphological phase is defined as the phase with no system of grammatical morphology dissociated from a general cognitive system, the proto-morphological phase is the period where the system of morphological grammar starts to develop and where the child sets out to construct creatively morphological patterns (Dressler, 1997; Dressler *et al.*, 1987; Kilani-Schoch & Dressler, 2001; Kilani-Schoch *et al.*, 1997). Language-specific effects found to be connected with general developmental constraints were mostly related to the relative morphological or syntactic complexity of particular verb forms (e.g., compound past). Further research will have to examine more directly the hypothesis that language-specific input factors may determine cross-linguistic differences in acquisition.

Continuity of developmental processes

The second main question concerns how children proceed in the acquisition of periphrastic structures: we now examine to what extent our present data provide evidence for a gradual and continuous process. We argue that the answer depends on the level of fineness and connexity in analyses, and that sufficiently fine-grained and inter-related analyses should reveal gradual and continuous developments. A first level in our data analyses indicates that all the three structures emerged in a range of about

three months in both languages, a range of time that could be seen as a relatively short one, denoting a rapid acquisition of the core of verb grammar. Moreover, the global index of verb grammaticization, which was calculated in order to assess children's ability to produce complete periphrastic constructions correctly, showed that an 'explosion' in the verb grammaticization process occurred in each of the four children to various extents. However, we claim that thorough analyses of how the periphrastic structures developed and of how they originated lead to a modulation of this first-level view, supporting, on the contrary, the conception of a gradual and continuous acquisition process.

First, analyses of how the periphrastic structures developed showed that there exist differences in times of emergence and acquisition between the three structures, in both languages. In French children, pertinent differences in time of emergence were found, showing the order: 'compound past' first, then 'modal constructions', and afterwards 'periphrastic future'. In Austrian children, differences were found in time of emergence and in time of acquisition, showing the orders: 'modal constructions', 'compound past' and afterwards 'periphrastic future' for emergence, and 'compound past', then 'modal constructions' for acquisition. In addition, each of the three structures developed gradually in both languages, following the diversification process described above. These findings are in accordance with a view of language acquisition as a gradual process, which may differ in different linguistic substructures, and which develops progressively within a structure.

Second, analyses of how the periphrastic structures originated were conducted with the idea that these structures should not emerge *ex nihilo* at around 2;0 and that precursors should be found. Two kinds of configurations with past participle or infinitive appeared as good candidates for being precursors of the periphrastic structures: ungrammatical bare past participles or infinitives, and fillers preceding past participles or infinitives. The conception that configurations with fillers preceding past participles or infinitives are precursors of periphrastic structures is in line with a number of recent studies which all stress the relation between fillers and the development of grammatical morphemes (Kilani-Schoch & Dressler, 2001; Peters, 2001; Veneziano & Sinclair, 2000). Although the kind of first knowledge implied by the use of phonological or syntactic fillers is disputed, there is agreement that fillers are 'on the way to grammatical morphemes' (Veneziano & Sinclair, 2000: 461), with preverbal fillers being placeholders for auxiliaries and modals (or clitic pronouns), similar to prenominal fillers as placeholders for determiners. A difficulty in our present data is that precursors consisting in the use of preverbal fillers were found in French children, but not in Austrian children. This does not mean, of course, that fillers are not produced at all by German-speaking children (see Vollmann, 1997). Variability is a puzzling characteristic of fillers: not all children produce them, and it may be that some languages are more susceptible to fillerization than others (Peters, 1997, 2001). Actually, fillers are likely to be more frequent in a language such as French, which has many more proclitic elements preceding nouns and main verbs than German.

The conception that bare past participles or infinitives are precursors of periphrastic structures is more controversial. In particular, generative theories of acquisition generally consider bare infinitives as involving missing or underspecified functional projections (e.g., Wexler, 1994, 1998). Our present view, proposed elsewhere in regard to French particularly (e.g., Bassano, 2000), argues for a conception where bare infinitives are

mostly considered as the expression of incomplete complex constructions in which the modal or auxiliary is missing. In our present data, arguments in favour of this conception first come from the parallel between bare infinitives and bare past participles: in roughly the same period when they produced bare infinitives, all four children also produced bare past participles, which can hardly be explained otherwise than by omission of an auxiliary. Arguments also come from the relations between bare nonfinite forms and filler+nonfinite forms, the latter being viewed as possible transitional procedures, rooted in both phonology and syntax and making a link between ungrammatical bare forms and complete grammatical forms. Finally, the strongest arguments come from the competitive developmental patterns shown for the periphrastic structures and their assumed precursors of both kinds: bare nonfinite forms and filler+nonfinite forms were produced earlier than periphrastic structures and they decreased and disappeared as the production of periphrastic structures strongly increased. All these inter-related patterns are in accordance with the conception that bare infinitives and past participles, as well as fillers preceding infinitives and past participles, are steps in the emergence of periphrastic structures, thus supporting a continuist view of language acquisition.

Temporal relations between lexical and grammatical spurts

The third and last question, closely related to the preceding one, is that of the relations between the grammatical and the lexical development of verbs. Evidence for inter-relations between lexical and grammatical development was provided by the parallelism found between the lexical production of verbs and the development of the synthetic index of verb grammaticization. In each of the four children, we found a spurt in the lexical production of verbs as well as a spurt in the verb grammaticization index, and the grammatical explosion followed the lexical increase with a delay of a few months. As proposed by an integrative view of language acquisition (Bassano, 2000; Bassano & Eme, 2001; Bates & Goodman, 1999; Marchman & Bates, 1994), these relations suggest, first, that there exist interdependencies between the lexical and grammatical development of verbs, and, second, that verb grammaticization might need a certain 'critical lexical mass' to develop. As a result, grammatical development appears as dependent on lexical development. Further research will have to investigate more precise relations between semantic properties and the grammaticization process of verbs.

* * *

To conclude, we hope that these analyses on the development of periphrastic constructions in two languages have shed some light on crucial issues in language acquisition, such as the factors at play in early verb grammar (both general cognitive and linguistic factors and language-specific effects), and the nature of early grammatical development (gradual, continuous, interdependent with lexical development). More research is needed to go further in this direction, for instance by providing more details on the respective role of cognitive and linguistic determinisms in verb grammar, by exploring the possible role of language input factors such as frequency or positional

salience, and by looking for qualitative relations between verb grammar and lexical development.

ACKNOWLEDGEMENTS

This research was supported by a grant from the *Cognitique* Programme (French Ministry for Research). Part of it was presented as a communication at the IX International Congress for the Study of Child Language, July 16–21, 2002, Madison, USA. We are grateful to the children – Pauline, Katharina, Benjamin and Jan – and parents who took part in our study. We would also like to thank the students and collaborators who contributed to the data collection and transcription.

REFERENCES

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.

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

Bassano, D. & Eme, P. E. (2001). Development of noun determiner use in French children: Lexical and grammatical bases. In M. Almgren, A. Barreña, M. J. Ezeizabarrena, I. Idiazabal & B. MacWhinney (Eds), *Research on child language acquisition* (pp. 1207–1220). Sommerville, MA: Cascadilla Press.

Bassano, D., Maillochon, I. & Eme, E. (1998). Developmental changes and variability in early lexicon: A study of French children's naturalistic productions. *Journal of Child Language*, *25*, 493–531.

Bassano, D., Maillochon, I., Klampfer, S. & Dressler, W. U. (2001*a*). L'acquisition de la morphologie verbale à travers les langues (I): Les fondements théoriques. *Enfance*, 1/2001, 81–99.

Bassano, D., Maillochon, I., Klampfer, S. & Dressler, W. U. (2001*b*). L'acquisition de la morphologie verbale en français et en allemand autrichien (II): L'épreuve des faits. *Enfance* 2/2001, 117–148.

Bates, E., Dale, P. & Thal, D. (1995). Individual differences and their implications for theories of language development. In P. Fletcher & B. MacWhinney (Eds), *The handbook of child language*. Oxford: Blackwell.

Bates, E. & Goodman, J. C. (1999). On the emergence of grammar from the lexicon. In B. MacWhinney (Ed.), *The emergence of language* (pp. 29–79). Mahwah, NJ: Erlbaum.

Bates, E. & MacWhinney, B. (1987) Competition, variation, and language learning. In B. MacWhinney (Ed.), *Mechanisms of language acquisition* (pp. 157–193). Hillsdale, NJ: Erlbaum.

Bates, E. & MacWhinney, B. (1989). Functionalism and the Competition Model. In B. MacWhinney & E. Bates (Eds), *The crosslinguistic study of language processing* (pp. 3–73). Cambridge MA: Cambridge University Press.

Behrens, H. (1993). Temporal reference in German child language: Form and function of early verb use. PhD dissertation, University of Amsterdam.

Bittner, D., Dressler, W. U. & Kilani-Schoch, M. (Eds) (2000). *First verbs: On the way to mini-paradigms*. ZAS Papers in Linguistics, *18*.

Caselli, M. C., Casadio, P. & Bates, E. (1999). A comparison of the transition from first words to grammar in English and Italian. *Journal of Child Language*, *26*, 69–111.

Choi, S. (1998). Verbs in early lexical and syntactic development in Korean. *Linguistics*, *36*, 755–780.
Comrie, B. (1985). *Tense*. Cambridge: Cambridge University Press.
Dressler, W. U. (1997). Universals, typology, and modularity in Natural Morphology. In R. Hickey & S. Puppel (Eds), *Language history and language modellings* (pp. 1399–1421). Berlin: de Gruyter.
Dressler, W. U. (2000). Naturalness. In Booij *et al.* (Eds), *Morphologie: ein internationales Handbuch zur Flexion und Wortbildung*, I (pp. 288–296). Berlin: de Gruyter.
Dressler, W. U., Mayerthaler, W., Panagl, O. & Wurzel, W. U. (1987). *Leitmotifs in natural morphology*. Amsterdam: Benjamins.
Elman, J. L., Bates, E., Johnson, M. H., Karmiloff-Smith, A., Parisi, D. & Plunkett, K. (1996). *Rethinking innateness: A connectionist perspective on development*. Cambridge, MA: MIT Press.
Karmiloff-Smith A. (1992). *Beyond modularity: A developmental perspective on cognitive science*. Cambridge, MA: MIT Press.
Kilani-Schoch, M. & Dressler, W. U. (2001). Filler + infinitive and pre- and protomorphology demarcation in a French acquisition corpus. *Journal of Psycholinguistic Research*, *30*, 653–685.
Kilani-Schoch, M., De Marco, A., Christofidou, A., Vassilakou, M., Vollmann, R. & Dressler, W. U. (1997). On the demarcation of phases in early morphology acquisition in four languages. In K. Dziubalska-Kolaczyk (Ed.), *Pre- and protomorphology in language acquisition. Papers and Studies in Contrastive Linguistics, 33*, 15–32.
Klampfer, S., Maillochon, I., Bassano, D. & Dressler, W. U. (1999). On early acquisition of verb inflection in Austrian German and French: The case of person and number marking. *Wiener Linguistische Gazette*, *64–65*, 1–29.
MacWhinney, B. (1999*a*). *The emergence of language*. Mahwah, NJ: Erlbaum.
MacWhinney B. (1999*b*). The emergence of language from embodiment. In B. MacWhinney, *The emergence of language* (pp. 213–256). Mahwah, NJ: Erlbaum.
MacWhinney, B. (2000). *The CHILDES project: Tools for analyzing talk*. Mahwah, NJ: Erlbaum.
Marchman, V. & Bates, E. (1994). Continuity in lexical and morphological development: A test of the critical mass hypothesis. *Journal of Child Language*, *21*, 339–366.
Peters, A. (1997). Language typology, individual differences and the acquisition of grammatical morphemes. In D. I. Slobin (Ed.), *The crosslinguistic study of language acquisition*, Vol. 5 (pp. 135–197). Hillsdale, NJ: Erlbaum.
Peters, A. (2001). Filler syllables: What is their status in emerging grammar? *Journal of Child Language*, *28*, 229–242.
Peters, A. M. & Menn, L. (1993). False starts and filler-syllables: ways to learn grammatical morphemes. *Language*, *69*, 742–777.
Pizzuto, E. & Caselli, M. C. (1994). The acquisition of Italian verb morphology in a cross-linguistic perspective. In Y. Levy (Ed.), *Other children, other languages* (pp. 137–187). Hillsdale, NJ: Erlbaum.
Plunkett, K. & Juola, P. (1999). A connectionist model of English past tense and plural morphology. *Cognitive Science*, *23*, 463–490.
Slobin, D. I. (1985). Why study acquisition crosslinguistically? In D. I. Slobin (Ed.), *The crosslinguistic study of language acquisition*, Vol. 1 (pp. 3–24). Hillsdale, NJ: Erlbaum.
Slobin, D. I. (1997*a*). The Universal, the typological, and the particular in acquisition. In D. I. Slobin (Ed.), *The crosslinguistic study of language acquisition*, Vol. 5 (pp. 1–39). Hillsdale, NJ: Erlbaum.
Slobin, D. I. (1997*b*). The origins of grammaticizable notions: Beyond the individual mind. In D. I. Slobin (Ed.), *The crosslinguistic study of language acquisition*, Vol. 5 (pp. 265–323). Hillsdale, NJ: Erlbaum.

Tomasello, M. (1992). *First verbs: A case study of early grammatical development*. Cambridge, MA: Cambridge University Press.
Tomasello, M. (1995). Language is not an instinct. *Cognitive Development*, *10*, 131–156.
Tomasello, M. (Ed.) (1998). *The new psychology of language. Cognitive and functional approaches to language structure*. Mahwah, NJ: Erlbaum.
Veneziano, E. & Sinclair, H. (2000). The changing status of 'filler syllables' on the way to grammatical morphemes. *Journal of Child Language*, *27*, 461–500.
Vollmann, R. (1997). Über die Kategorie 'fillers' im Erstspracherwerb. *Grazer Linguistische Studien*, *48*, 103–132.
Wexler, K. (1994). Optional infinitives, head movement, and economy of derivation. In N. Hornstein & D. Lightfoot (Eds), *Verb movement* (pp. 305–350). Cambridge: Cambridge University Press.
Wexler, K. (1998). Very early parameter setting and the unique checking constraint: A new explanation of the Optional Infinitive Stage. *Lingua*, *106*, 23–79.
Wittek, A. & Tomasello, M. (2002). German children's productivity with tense morphology: The Perfekt (present perfect). *Journal of Child Language*, *29* (3), 567–589.

APPENDIXES

Appendix 1a Type and token (T/t) frequencies of the three periphrastic structures in Pauline's samples

Age session	*Structures: total types/tokens*	*Types/tokens for each auxiliary or modal form*
2;0	AUX+PPP: 1/1	1/1 E3sg
	MOD+INF: 0/0	
	AUX+INF: 0/0	
2;1	AUX+PPP: 1/2	1/2 E3sg
	MOD+INF: 1/2	1/2 P2sg
	AUX+INF: 0/0	
2;2	AUX+PPP: 0/0	
	MOD+INF: 1/1	1/1 V1sg
	AUX+INF: 0/0	
2;3	AUX+PPP: 0/0	
	MOD+INF: 1/1	1/1 V1ppp
	AUX+INF: 1/1	1/1 AL1sg
2;4	AUX+PPP: 1/1	1/1 E3sg
	MOD+INF: 2/2	1/1 V1sg (*L); 1/1 P3sg
	AUX+INF: 1/1	1/1 AL1sg
2;5	AUX+PPP: 2/2	2/2 A1sg
	MOD+INF: 4/5	1/1 V1sg; 3/4 F3sg

continued overleaf

Appendix 1a continued

Age session	*Structures: total types/tokens*	*Types/tokens for each auxiliary or modal form*
	AUX+INF: 4/4	1/1 AL1sg, 3/3 AL3sg
2;6	AUX+PPP: 10/16	1/1 E3sg, 1/1 E3pl, 1/1 Eimp1sg, 1/2 Eimp3sg; 1/2 A1sg, 2/6 A2sg, 3/4 A3sg
	MOD+INF: 3/4	1/2 V1sg; 1/1 F3sg; 1/1 S1sg
	AUX+INF: 1/1	1/1 AL3sg
2;7	AUX+PPP: 15/21	2/4 E3sg; 9/12 A1sg (1*AUX), 3/4 A3sg, 1/1 Aimp1sg
	MOD+INF: 7/10	1/1 V1sg; 5/8 F3sg; 1/1 S1sg
	AUX+INF: 1/1	1/1 AL3sg
2;8	AUX+PPP: 3/4	1/1 E3sg; 2/3 A1sg
	MOD+INF: 9/11	4/5 P3sg; 5/6 F3sg
	AUX+INF: 5/5	2/2 AL1sg, 3/3 AL3sg
2;9	AUX+PPP: 16/20	1/1 E3sg, 1/1 Eimp3pl; 6/9 A1sg, 1/1 A2sg, 2/2 A3sg, 1/1 Aimp1sg, 3/3 Aimp3sg, 1/2 Afut1sg
	MOD+INF: 7/10	1/1 P1sg, 1/1 P3sg; 2/3 F3sg; 3/5 S1sg
	AUX+INF: 5/9	3/7 AL1sg, 1/1 AL2sg, 1/1 AL3sg
2;10	AUX+PPP: 5/6	1/1 E3sg; 4/5 A1sg (1*AUX)
	MOD+INF: 5/7	3/4 V1sg; 1/1 P3sg; 1/2 F3sg
	AUX+INF: 5/5	4/4 AL1sg, 1/1 AL3sg
2;11	AUX+PPP: 9/10	1/1 E3sg, 1/1 Eimp3sg; 2/2 A1sg, 1/2 A2sg, 1/1 A3sg, 1/1 Aimp1sg, 1/1 Aimp2sg, 1/1 Aimp3sg
	MOD+INF: 4/5	2/2 P3sg; 2/3 F3sg
	AUX+INF: 4/5	1/1 AL1sg, 3/4 AL3sg
3;0	AUX+PPP: 11/12	2/3 E3sg; 3/3 A1sg (1*AUX), 1/1 A2sg, 3/3 A3sg, 2/2 Aimp3sg
	MOD+INF: 6/6	1/1 V1sg; 1/1 P3sg; 4/4 F3sg
	AUX+INF: 6/7	2/2 AL1sg, 4/5 AL3sg (1/2*G)

Note. E = *être* 'be'; A = *avoir* 'have'; V = *vouloir* 'want'; P = *pouvoir* 'can/may'; S = *savoir* 'know how'; F = *falloir* 'must'; AL = *aller* 'go'; inf = infinitive; imp = imperfect; fut = future; ppp = past participle. The symbol (*G) refers to a form error on the grammatical verb, and (*L) to a form error on the lexical verb of the periphrastic structure. The symbol (*AUX) refers to the use of a wrong auxiliary.

Appendix 1b Type and token (T/t) frequencies of the three periphrastic structures in Benjamin's samples

Age session	*Structures: total types/tokens*	*Types/tokens for each auxiliary or modal form*
1;11	AUX+PPP: 2/3	2/3 E3sg
	MOD+INF: 0/0	
	AUX+INF: 0/0	
2;0	AUX+PPP: 3/3	1/1 E3sg; 2/2 A3sg
	MOD+INF: 1/1	1/1 P1sg
	AUX+INF: 0/0	
2;1	AUX+PPP: 2/3	1/2 E3sg; 1/1 A2sg
	MOD+INF: 1/1	1/1 V1sg
	AUX+INF: 1/1	1/1 AL3sg
2;2	AUX+PPP: 6/9	1/1 E3sg; 4/6 A1sg, 1/2 A3sg
	MOD+INF: 0/0	
	AUX+INF: 0/0	
2;3	AUX+PPP: 7/10	2/3 E3sg, 1/1 E3pl; 1/1 A2sg, 3/5 A3sg
	MOD+INF: 1/1	1/1 V1sg
	AUX+INF: 0/0	
2;4	AUX+PPP: 4/5	3/4 E3sg; 1/1 A1sg (1*AUX)
	MOD+INF: 1/1	1/1 V1sg
	AUX+INF: 1/1	1/1 AL3sg
2;5	AUX+PPP: 10/22	1/1 E3sg, 1/1 E1sg, 1/6 E3pl; 3/5 A1sg, 1/2 A2sg, 3/7 A3sg
	MOD+INF: 8/15	1/1 P1sg, 5/12 P3sg; 2/2 V1sg
	AUX+INF: 3/4	2/2 AL1sg, 1/2 AL3sg
2;6	AUX+PPP: 5/6	3/4 E3sg; 2/2 A3sg
	MOD+INF: 2/2	2/2 V1sg
	AUX+INF: 0/0	
2;7	AUX+PPP: 12/14	5/7 E3sg; 2/2 A1sg, 3/3 A2sg (1*AUX), 2/2 A3sg
	MOD+INF: 1/2	1/2 S1sg
	AUX+INF: 7/9	3/3 AL1sg, 1/3 AL2sg, 3/3 AL3sg
2;8	AUX+PPP: 7/10	1/1 E3sg; 2/2 A1sg, 3/6 A2sg, 1/1 A3sg
	MOD+INF: 2/3	1/2 V2sg; 1/1 F3sg
	AUX+INF: 2/2	2/2 AL1sg
2;9	AUX+PPP: 9/11	3/5 E3sg; 4/4 A1sg, 2/2 A3sg
	MOD+INF: 6/10	3/5 V1sg; 3/5 P2sg
	AUX+INF: 11/12	8/9 AL1sg, 1/1 AL2sg, 2/2 AL3sg
2;10	AUX+PPP: 13/16	3/6 E3sg, 2/2 E2sg; 1/1 A1sg, 1/1 A2sg, 4/4 A3sg, 2/2 A3pl
	MOD+INF: 3/3	1/1 V1sg; 1/1 P1sg, 1/1 P2sg
	AUX+INF: 2/2	2/2 AL1sg

continued overleaf

Appendix 1b continued

Age session	*Structures: total types/tokens*	*Types/tokens for each auxiliary or modal form*
2;11	AUX+PPP: 8/9	1/2 E3sg; 2/2 A1sg, 5/5 A3sg
	MOD+INF: 7/8	3/4 V1sg, 1/1 V2sg; 3/3 F3sg
	AUX+INF: 3/4	1/1 AL1sg, 1/1 AL2sg, 1/1 AL3sg
3;0	AUX+PPP: 12/20	1/1 E1sg, 2/4 E3sg, 2/3 E3pl, 1/1 Einf, 1/1 Eimp3sg; 2/4 A1sg, 2/4 A2sg, 1/2 Aimp3sg (*AUX)
	MOD+INF: 8/10	1/1 V1sg; 2/2 P1sg,1/2 P2sg, 1/1 Pimp3sg; 3/4 F3sg
	AUX+INF: 5/7	2/3 AL1sg, 1/1 AL2sg, 2/3 AL3sg

For an explanatory note, see Appendix 1a.

Appendix 1c Type and token (T/t) frequencies of the three periphrastic structures in Katharina's samples

Age session	*Structures: total types/tokens*	*Types/tokens for each auxiliary or modal form*
2;4	AUX+PPP: 0/0	
	MOD+INF: 1/1	1/1 K1/3sg
	AUX+INF: 0/0	
2;5	AUX+PPP: 1/1	1/1 H1sg
	MOD+INF: 1/1	1/1 K1sg
	AUX+INF: 0/0	
2;6	AUX+PPP: 1/2	1/2 H3sg
	MOD+INF: 2/4	1/2 K3sg; 1/2 MÜ1sg
	AUX+INF: 1/1	1/1 W1pl
2;7	AUX+PPP: 8/10	1/1 S3sg; 5/6 H1sg (1/1 *L), 2/3 H3sg
	MOD+INF: 1/1	1/1 MÖ1/3sg
	AUX+INF: 0/0	
2;8	AUX+PPP: 3/3	1/1 H1sg, 2/2 H3sg
	MOD+INF: 3/3	1/1 MÖ1sg; 1/1 K3sg; 1/1 MÜ3sg
	AUX+INF: 0/0	
2;9	AUX+PPP: 4/6	4/6 H1sg (1/1 *L)
	MOD+INF: 5/5	1/1 K1sg; 2/2 D1sg; 1/1 MÜ1sg, 1/1 MÜ2sg
	AUX+INF: 2/3	1/2 W2sg (*G), 1/1 W1pl

continued

Appendix 1c continued

Age session	*Structures: total types/tokens*	*Types/tokens for each auxiliary or modal form*
2;10	AUX+PPP: 8/10	3/3 S3sg, 1/1 S3pl; 1/2 H1sg, 2/2 H3sg, 1/2 H1pl
	MOD+INF: 3/4	3/4 K3sg
	AUX+INF: 0/0	
2;11	AUX+PPP: 0/0	
	MOD+INF: 0/0	
	AUX+INF: 0/0	
3;0	AUX+PPP: 10/12	4/5 S3sg (1/1 *L), 1/1 Spret3s (*L); 3/3 H1sg, 1/2 H3sg, 1/1 H1pl (*L)
	MOD+INF: 4/4	1/1 K3sg; 1/1 D3sg; 1/1 MÜ3sg, 1/1 MÜ1pl
	AUX+INF: 0/0	

Note. S = *sein* 'be'; H = *haben* 'have'; T = *tun* 'do'; W = *werden* 'will'; WO = *wollen* 'want'; MÖ = *mögen/möchten* 'want'; K = *können* 'can/may'; D = *dürfen* 'can/may'; SO = *sollen* 'must'; MÜ = *müssen* 'must'; inf = infinitive; pret = preterite; con = conjunctive
The symbol (*G) refers to a form error on the grammatical verb, and (*L) to a form error on the lexical verb of the periphrastic structure. The symbol (*AUX) refers to the use of a wrong auxiliary.

Appendix 1d Type and token (T/t) frequencies of the three periphrastic structures in Jan's samples

Age session	*Structures: total types/tokens*	*Types/tokens for each auxiliary or modal form*
2;0	AUX+PPP: 0/0	
	MOD+INF: 1/1	1/1 K1/3sg
	AUX+INF: 0/0	
2;1	AUX+PPP: 0/0	
	MOD+INF: 4/4	1/1 WO3sg; 1/1 MÖ3sg; 1/1 K3sg; 1/1 SO3sg
	AUX+INF: 1/1	1/1 T3pl
2;2	AUX+PPP: 4/4	2/2 S3sg; 1/1 W3sg, 1/1 Winf
	MOD+INF: 5/5	1/1 WOpret3sg; 1/1 K1pl , 1/1 K3pl; 1/1 MÜ2sg, 1/1 MÜ3sg
	AUX+INF: 0/0	
2;3	AUX+PPP: 2/3	2/3 H1sg
	MOD+INF: 3/3	1/1 WO1sg; 1/1 K3sg; 1/1 SO3sg
	AUX+INF: 0/0	

continued overleaf

Appendix 1d continued

Age session	*Structures: total types/tokens*	*Types/tokens for each auxiliary or modal form*
2;4	AUX+PPP: 3/5	1/1 S1sg; 2/4 H3sg (1/1 *AUX)
	MOD+INF: 1/2	1/2 K3sg
	AUX+INF: 0/0	
2;5	AUX+PPP: 1/1	1/1 S3sg
	MOD+INF: 8/10	1/2 WOpret3sg; 2/2 K3sg, 1/2 K1pl, 4/4 D3sg
	AUX+INF: 0/0	
2;6	AUX+PPP: 8/8	1/1 S3sg, 1/1 S3pl; 4/4 H1sg, 1/1 H2sg, 1/1 H3sg (*L)
	MOD+INF: 11/14	1/1 WOpret3sg, 1/1 WOpret3pl; 4/5 K3pl; 1/1 MÜ3sg, 4/6 MÜ3pl (4/6 *G, 1/1 *L)
	AUX+INF: 0/0	
2;7	AUX+PPP: 3/3	1/1 S3sg; 2/2 H1sg
	MOD+INF: 5/9	1/1 WO1sg, 1/1 WO2sg; 1/2 K3sg; 1/1 SO1sg, 1/4 SO1pl
	AUX+INF: 0/0	
2;8	AUX+PPP: 4/4	2/2 S3g, 1/1 S1pl; 1/1 H1sg
	MOD+INF: 5/7	1/1 WO2sg; 2/4 K3sg, 1/1 K1pl; 1/1 MÜ3sg
	AUX+INF: 1/1	1/1 T1pl
2;9	AUX+PPP: 9/12	1/1 S3sg; 2/2 H1sg, 4/7 H2sg, 1/1 H1/3pl, 1/1 H3pl
	MOD+INF: 4/4	2/2 WO1sg, 1/1 WO1/3sg; 1/1 MÜ3sg
	AUX+INF: 2/3	1/2 W1sg (*L), 1/1 W2sg (*L)
2;10	AUX+PPP: 10/10	2/2 S1sg; 2/2 H1sg, 4/4 H2sg (1/1 *L), 1/1 H3sg, 1/1 H1pl
	MOD+INF: 3/3	1/1 K2sg, 1/1 K3sg, 1/1 K1pl
	AUX+INF: 0/0	
2;11	AUX+PPP: 1/1	1/1 S3sg
	MOD+INF: 5/7	1/2 WO2sg; 2/3 MÖ1sg; 1/1 K1pl; 1/1 D1sg
	AUX+INF: 1/1	1/1 W3sg (*G)
3;0	AUX+PPP: 8/11	2/2 S3sg; 3/4 H1sg, 1/2 H2sg, 2/3 H3pl
	MOD+INF: 5/6	1/2 WO2sg; 2/2 K1sg, 1/1 K3sg, 1/1 K3pl
	AUX+INF: 2/7	1/5 W1sg, 1/2 Wcon:pret1sg

For an explanatory note, see Appendix 1c.

Appendix 2a Configurations with past participle and infinitive in the corpus of each of the French children: net numbers by monthly sample

Age	**0+PPP*	*0+PPP correct*	*FILL+ PPP*	*AUX+ PPP*	**0+INF*	*0+INF correct*	*FILL+ INF*	*OTHER +INF*	*MOD/AUX +INF*
Pauline									
1;2	0	0	0	0	2	0	0	0	0
1;3	0	0	0	0	0	0	0	0	0
1;4	3	0	0	0	0	0	0	0	0
1;5	0	0	0	0	2	0	0	0	0
1;6	0	0	0	0	0	0	0	0	0
1;7	0	0	0	0	3	0	4	0	0
1;8	1	0	0	0	1	1	1	0	0
1;9	4	0	0	0	9	0	0	0	0
1;10	0	0	0	0	8	1	2	0	0
1;11	3	0	0	0	5	0	1	0	0
2;0	3	0	0	1	4	0	1	0	0
2;1	2	0	1	2	9	0	3	0	2
2;2	2	0	1	0	5	0	3	0	1
2;3	4	0	2	0	7	0	6	0	2
2;4	3	0	0	1	9	0	3	3	3
2;5	2	1	0	3	1	0	0	0	9
2;6	0	0	0	16	5	0	0	1	5
2;7	0	0	0	21	3	1	0	6	11
2;8	0	0	0	4	1	0	0	0	16
2;9	0	0	0	20	0	1	0	8	19
2;10	0	0	0	6	0	0	0	0	12
2;11	0	0	0	10	0	0	0	3	10
3;0	0	0	0	12	0	0	0	1	13
Benjamin									
1;11	3	0	0	3	1	0	4	2	0
2;0	5	0	2	3	1	0	1	3	1
2;1	2	0	4	3	4	1	4	0	2
2;2	4	0	0	9	5	0	0	0	0
2;3	0	0	0	10	5	2	3	2	1
2;4	3	0	1	5	1	1	4	2	2
2;5	2	0	1	22	1	0	0	9	19
2;6	2	0	0	6	5	0	0	5	2
2;7	0	0	1	14	2	0	2	5	11
2;8	2	1	0	10	4	0	0	3	5
2;9	0	1	0	11	1	0	2	5	22
2;10	0	0	2	16	0	1	0	3	5
2;11	1	0	0	9	1	1	2	1	12
3;0	0	0	1	20	1	0	5	3	17

Appendix 2b Configurations with past participle and infinitive in the corpus of each of the Austrian children: net numbers by monthly sample

	**0+PPP*	*0+PPP correct*	*AUX+PPP*	**0+INF*	*0+INF correct*	*OTHER +INF*	*MOD/AUX +INF*
Katharina							
2;4	1	0	0	4	2	0	1
2;5	3	0	1	4	3	0	1
2;6	0	0	2	5	1	0	5
2;7	0	0	10	0	1	1	1
2;8	0	0	3	4	3	1	3
2;9	0	0	6	1	1	0	8
2;10	1	0	10	1	3	2	4
2;11	1	0	0	0	2	0	0
3;0	0	0	12	0	0	2	4
Jan							
1;3	0	0	0	1	0	0	0
1;4	0	0	0	9	0	0	0
1;5	0	0	0	10	0	0	0
1;6	0	0	0	19	0	0	0
1;7	4	0	0	10	0	0	0
1;8	6	0	0	16	0	0	0
1;9	9	0	0	26	0	0	0
1;10	10	0	0	11	0	0	0
1;11	1	0	0	29	1	0	0
2;0	1	1	0	33	0	2	1
2;1	4	0	0	17	0	2	5
2;2	1	0	4	15	0	4	5
2;3	3	0	3	2	0	1	3
2;4	0	1	5	2	3	3	2
2;5	1	0	1	0	1	2	10
2;6	0	0	8	1	5	4	14
2;7	0	0	3	0	0	1	9
2;8	0	0	4	0	0	0	8
2;9	0	0	12	0	0	1	7
2;10	0	1	10	1	3	1	3
2;11	0	0	1	0	2	0	8
3;0	0	0	11	0	0	0	13

Appendix 3a Indices of verb grammaticization in Pauline's and Benjamin's data

	Pauline			*Benjamin*		
Age	*AUX +PPP index*	*MOD/AUX +INF index*	*AUX/MOD +PPP/INF global index*	*AUX+PPP index*	*MOD/AUX +INF index*	*AUX/MOD PPP/INF global index*
1;11	0.00	0.00	0.00	0.50	0.00	0.27
2;0	0.25	0.00	0.11	0.33	0.33	0.33
2;1	0.40	0.14	0.21	0.33	0.20	0.26
2;2	0.00	0.11	0.08	0.69	0.00	0.50
2;3	0.00	0.13	0.10	1.00	0.11	0.58
2;4	0.25	0.20	0.21	0.56	0.29	0.44
2;5	0.50	0.90	0.79	0.88	0.95	0.91
2;6	1.00	0.50	0.81	0.75	0.29	0.53
2;7	1.00	0.79	0.91	0.93	0.73	0.83
2;8	1.00	0.94	0.95	0.83	0.56	0.71
2;9	1.00	1.00	1.00	1.00	0.88	0.92
2;10	1.00	1.00	1.00	0.89	1.00	0.91
2;11	1.00	1.00	1.00	0.90	0.80	0.84
3;0	1.00	1.00	1.00	0.95	0.74	0.84

Appendix 3b Indices of verb grammaticization in Katharina's and Jan's data[†]

	Katharina			*Jan*		
Age	*AUX +PPP index*	*MOD/AUX +INF index*	*AUX/MOD +PPP/INF global index*	*AUX+PPP index*	*MOD/AUX +INF index*	*AUX/MOD PPP/INF global index*
1;11	0.00	0.00	0.00	0.00	0.00	0.00
2;0	0.00	0.00	0.00	0.00	0.03	0.03
2;1	0.00	0.00	0.00	0.00	0.23	0.19
2;2	0.00	0.00	0.00	0.80	0.25	0.36
2;3	0.00	0.00	0.00	0.50	0.60	0.55
2;4	0.00	0.20	0.17	1.00	0.50	0.78
2;5	0.25	0.20	0.22	0.50	1.00	0.92

continued overleaf

Appendix 3b continued

	Katharina			*Jan*		
Age	*AUX +PPP index*	*MOD/AUX +INF index*	*AUX/MOD +PPP/INF global index*	*AUX+PPP index*	*MOD/AUX +INF index*	*AUX/MOD PPP/INF global index*
2;6	1.00	0.50	0.58	1.00	0.93	0.96
2;7	1.00	1.00	1.00	1.00	1.00	1.00
2;8	1.00	0.43	0.60	1.00	1.00	1.00
2;9	1.00	0.89	0.93	1.00	1.00	1.00
2;10	0.91	0.80	0.88	1.00	0.75	0.93
2;11	0.00	0.00	0.00	1.00	1.00	1.00
3;0	1.00	1.00	1.00	1.00	1.00	1.00

† The 0.00 value of Katharina's indices at 2;11 is due to the low number of past participles and infinitives in the sample (1 *0+PPP and 2 0+INF correct, cf. Appendix 2b). In Fig. 7, it is replaced by the same value as at 2;10 (0.88).

ADDRESS FOR CORRESPONDENCE

Dominique Bassano, Laboratoire Cognition et Développement,
Centre Henri Piéron, 71 avenue Edouard Vaillant,
92774 Boulogne-Billancourt Cedex, France
E: bassano@psycho.univ-paris5.fr