

**Early development of nouns and verbs in French:
exploring the interface between lexicon and
grammar***

DOMINIQUE BASSANO

Laboratoire Cognition et Développement, Université Paris V – CNRS

(Received 8 March 1999. Revised 8 February 2000)

ABSTRACT

Early acquisition of nouns and verbs across languages is a key issue for a number of recent studies that question the reality of the ‘noun-bias’ and wonder about the reasons why it exists as they explore the role of cognitive vs. more language-specific input factors. Addressing this issue, the present study investigates how the noun and verb word classes develop in the free speech of a French child between the ages of 1;2 and 2;6, from the perspective of semantic and grammatical development. The analyses indicate that, in French acquisition, nouns clearly predominate over verbs until age 1;8 at least, but that verbs are produced in the early stages. Concrete object names among nouns and concrete action verbs among verbs were found to be the most prevalent categories, but they were not the earliest to appear and their distribution revealed an asymmetry in the conceptual packaging of nouns and verbs. Verb grammaticalization, assessed through inflection and auxiliary use, lagged somewhat behind noun grammaticalization, assessed through determiner use. This result supports the hypothesized noun–verb grammatical asynchrony. Verb grammaticalization seems to be related to the production of concrete action verbs, and noun grammaticalization to that of concrete object nouns, indicating interactions between semantic and grammatical development. These findings, discussed in a cross-linguistic perspective, suggest that both conceptual and grammatical packaging are important and interacting factors in noun and verb development, and argue in favour of a constructivist approach to language acquisition.

[*] We are grateful to Pauline and her family. We acknowledge Isabelle Maillochon who collected and transcribed the child’s corpus and Madeleine Léveillé for assistance in data processing. Address for correspondence: Dominique Bassano, Laboratoire Cognition et Développement (UMR 8605), 71 avenue Edouard Vaillant, 92774 Boulogne-Billancourt Cedex, France. e-mail: dominique.bassano@psycho.univ-paris5.fr

INTRODUCTION

In recent developmental psycholinguistic research, an increasing interest has been taken in the issue of how young children acquire nouns and verbs, with particular attention being paid to crosslinguistic variations in development. One of the reasons for this interest is that noun and verb word classes are good candidates for investigating crucial questions, such as the respective roles of universal and language-specific processes during the initial stages of language acquisition, and the relationship between lexical-semantic and grammatical development. The present study addresses these questions. Using analyses of one child's spontaneous free speech, it investigates the emergence and early development of noun and verb word classes in French speech production, in a crosslinguistic perspective.

Noun and verb word classes

The idea that words could be distinguished and classified comes from the work of Greek and Latin grammarians. The noun-verb distinction is the oldest found in those scientists' attempts to classify parts of speech (see Plato's *Cratylus*), and it remains the central one in current classifications, which usually involve about ten word classes (i.e. noun, verb, adjective, adverb, pronoun, determiner, conjunction, preposition, interjection). In these classifications, word class definition is based on various complementary types of criteria, mainly semantics and grammar. The semantic criteria are often seen as the most intuitive and should allow definition of noun and verb categories on a universal basis. Traditionally, nouns are devoted to denoting entities, especially concrete objects, and have a referential function that MacNamara (1982) described by saying that nouns refer to 'things that you can pick up or bump into'. However, the semantic class of nouns is much broader, including abstractions such as 'friend', 'family', 'name', as well as proper names. Lyons (1977) postulates a focal semantic concept of the noun which states that all language encodes, namely words for persons, animals and things, and proposes that more abstract categories denoted by nouns are generalizations from this core concept. In contrast with nouns, verbs, used to assert something about entities that nouns denote, have a predicative function. They denote events, especially actions, but the semantic category of verbs also include references to states, mental activities or attitudes, and so on. Moreover, verbs are semantically close to another class of word, namely adjectives, which denote qualities and are somewhat intermediate between nouns and verbs.

Although ontological semantic properties appear to form the basis of these word classes, noun and verb categories cannot be completely defined without using grammatical criteria, which involve the distributional and morphological properties of words. Grammatical criteria emphasize that the

words of a class occur in particular positions in sentences, are marked in particular morphophonemic ways, or agree with particular morphemes. For instance, noun phrases have a grammatical role in sentences, such as that of subject or object of the verb, and, in English, among other languages, nouns are marked for the plural in a characteristic way. Verbs have syntactic properties, such as being the main component of the predicate and agreeing with the subject, and English verbs have characteristic ways of taking present and past tensing whereas nouns and adjectives do not. Laying stress on such formal properties, Maratsos (1982) defended the view that word classes are basically formed and shaped by the set of grammatical operations in which their members participate rather than by their semantic denotations *per se*. However, there seems to be a large agreement within linguistics that a purely grammatical approach cannot be sufficient because distributional and morpho-syntactic properties generally reflect language-specific options and vary across languages, whereas word classes (noun, at least) appear to be linguistic universals.

Thus, the categories of noun and verb are both semantic and grammatical concepts, and, as noted by Lyons (1977), there are good correlations between semantic and syntactic classifications in a large number of languages. The central issue approached in the present paper, that is, when and how noun and verb word classes start to appear in child language, is shaped by the intrinsically double dimension of word classes. From the semantic point of view, children are likely to have some notion of nouns and verbs early as words are linked conceptually and are used with appropriate overall referents, although, as we know, children's words do not always have the same delimitations as adults' words. The current debate on noun bias and noun-verb developmental asynchrony, presented below, takes this conceptual approach, with nouns and verbs defined in terms of adult notional categories. However, grammatical adult like word classes of noun and verb cannot be allotted to the child as far as nouns and verbs are not employed in conformity with the grammatical constraints which characterize these categories in the language being learned. To be sure, the sign that adultlike word classes are acquired is that their specific grammatical properties are acquired and produced. This conception implies that analysis of noun and verb semantic development, which gives the initial pace, is complemented by analysis of their grammaticalization, which provides the strict criteria for word class acquisition. Moreover, in the light of the relationship existing between semantic and grammatical properties of word classes, it can be hypothesized that the grammatical categories of noun and verb do not develop independently of their semantic bases. This hypothesis, which agrees with an interactive approach to language development that claims interdependencies between lexical and grammatical development (cf. Bates, Bretherton & Snyder, 1988; Marchman & Bates, 1994; Bates & Goodman,

1997), requires investigation of the interactions between lexical-semantic and grammatical properties within noun and verb classes.

Conceptual development of nouns and verbs : the debate on the noun-verb asynchrony

A main issue regarding noun and verb development is to determine whether regularities in their order of acquisition can be found. Until recently, common observations in studies of children's first spoken words have shown that content words are acquired before closed-class words, and that within content words, nouns, mainly words denoting objects, appear earlier and with greater frequency than verbs and other predicative terms, which develop later. The initial predominance of nouns in early vocabularies – the so-called noun bias – seems to be one of the best established phenomena in studies of English-speaking children, clearly appearing in natural free speech as well as in parents' diaries or checklists until at least the child reaches the age of 1;6–1;8 (Nelson, 1973; Gentner, 1982; MacNamara, 1982; Bates *et al.*, 1988; Bates, Marchman, Thal, Fenson, Dale, Reznick, Reilly & Hartung, 1994). For example, in Nelson's (1973) well-known study, nouns formed more than half of the first 50 words produced by children. In Bates *et al.*'s (1994) large-scale, cross-sectional study of 1800 children from age 0;8 to 2;6 conducted using parental reports from the MacArthur Communicative Development Inventories (CDI, Fenson, Dale, Reznick, Thal, Bates, Hartung, Pethick & Reilly, 1993), nouns were first largely predominant, occupying an average of 55% of the lexicon of children with vocabularies between 100 and 200 words, while predicates (verbs and adjectives) formed less than 15%. The emphasis in vocabulary content was found to shift from nouns to predicates and finally to closed-class words. Some doubt was cast on the generality of the noun-bias in English-speaking children by analyses of stylistic interindividual variability (e.g. Nelson, 1973; Bates *et al.* 1994; Bates, Dale & Thal, 1995), which suggested that some children, referred to as expressive children, learned fewer nouns than the prototypical referential children. The proportion of nouns in early lexicons seems to partially depend on sociocultural factors such as birth order and the parents' socioeconomic class. Yet, even expressive children appear to possess a fair number of nouns (33% in Nelson's study), and the noun-bias is the predominant pattern for learners of English. Criticisms of the claim for a noun-bias in English-speaking children have also been based on questions regarding the definition and/or extension of lexical categories. It has been stressed that children's early vocabularies involve not only nominals, but also a number of non-nominal elements such as social or cognitive relational words with a predicative use, which may not fall into reliable word categories (Bloom, Tinker & Margulis 1993), and that the nouns acquired by young children are far from referring solely to concrete objects (Nelson, Hampson & Kessler-

Shaw, 1993). However, these considerations do not directly challenge the predominance of lexical adult-like nouns over lexical adult-like verbs in the first stages of English development.

The noun–verb developmental asynchrony was in fact questioned in studies focusing on the acquisition of languages other than English. Gentner's (1982) analyses of children's early words in six languages (mainly English, but also German, Japanese, Mandarin Chinese, Turkish, and Kaluli) did not provide evidence of any crosslinguistic differences, and suggested that the initial predominance of nouns was universal and language-independent. Nevertheless, the data that formed the bases of her empirical analyses were of limited reliability, coming from indirect sources and heterogeneous collecting and sampling methods. Reliable data were obtained from some large-scale studies using parental reports based on adapted versions of the CDI to examine early vocabulary development of Spanish-speaking (Jackson-Maldonado, Thal, Marchman, Bates & Gutierrez-Clellen, 1993) and Italian-speaking (Caselli, Bates, Casadio, Fenson, Fenson, Sanderl & Weir, 1995; Caselli, Casadio & Bates, 1999) children. The developmental patterns found therein were very similar to those of English-speaking children. In particular, the initial predominance of nouns and the late emergence of verbs was observed for both Italian- and English-speaking children in Caselli *et al.*'s (1995; 1999) comparison. The noun–verb asynchrony was challenged by studies on the natural productions of children learning Japanese, Korean, and Mandarin Chinese, i.e. languages that differ substantially from English in a number of structural and pragmatic respects. Studies on Korean (Choi & Gopnik, 1995; Gopnik, Choi & Baumberger, 1996) report that Korean children use verbs earlier than English-speaking children, but use fewer and less varied nouns. Thus, both verbs and nouns in Korean seem to be dominant categories at the single-word stage, and an early verb spurt often precedes the naming spurt. An even more conclusive criticism of early noun dominance can be found in analyses of natural speech of children learning Mandarin Chinese (Tardif, 1996; Tardif, Shatz & Naigles, 1997; Gelman & Tardif, 1998), who appeared to produce more verbs than nouns at the age of about 1;8–1;10.

These findings raise a number of interesting issues about the nature of the underlying acquisitional processes and factors involved in noun and verb development. Supporters of the universality of the noun–verb developmental asynchrony emphasize the importance of cognitive and perceptual constraints in vocabulary acquisition. Gentner's (1982) NATURAL PARTITIONS hypothesis states that the difference between nouns and verbs is primarily 'based on a preexisting perceptual–conceptual distinction between concrete concepts such as persons or things and predicative concepts of activity, change-of-state, or causal relations' (1982: 301). Nouns are conceptually simpler, more basic, more tangible, and easier to grasp by children than verbs and other

predicates because nouns correspond to particularly cohesive perceptual entities, while ‘the perceptual elements that are packaged into verb referents are distributed more sparsely through the perceptual field’ (1982: 324). Differences in the conceptual-cognitive bases of noun and verb learning have been underlined in a number of subsequent studies on verbs (e.g. Smith & Sachs, 1990; Tomasello, 1992; Smiley & Huttenlocher, 1995). Verbs are thought to typically refer to transient events which are not perceptually available to the child when the word is uttered, i.e. they are learned in what Tomasello called non-ostensive contexts. Verbs are typically ambiguous because they do not refer to a clear conceptual packaging whose situational aspects are relevant to their meaning.

Gentner (1982) claimed that the conceptual-perceptual basis is the key factor in noun and verb acquisition. Not having found any clear differences in her crosslinguistic data analyses, she concluded that none of the language-based factors she examined (frequency, word order, morphological transparency, cultural patterns of language teaching) are strong enough to change the conceptual predisposition effect, although she recognized that language variation certainly affects acquisition. A similar position is shared by Caselli *et al.* (1995; 1999), who found no difference between Italian- and English-speaking children and concluded that the structural differences between these languages – including the greater salience of verbs in Italian than in English – do not affect the emergence and growth of lexical categories. These authors suggested that methodological effects might account for the differences found in other children, e.g. the Korean-speaking children in Choi and Gopnik’s studies. It is now well-known, indeed, that the two main methods for collecting data in studies of lexical development, i.e. parental reports vs. free speech samples, probably do not chart the exact same language abilities. According to Caselli *et al.* (1995; see also Bates *et al.* 1988), free speech samples show the forms that children prefer to use in various contexts, while parental reports should reveal the words they are able to use.

However, although it is generally agreed that methodological effects are indeed at play, methodological factors cannot fully account for the crosslinguistic differences found in acquisition data, and attention has been drawn to the possible role of linguistic input factors. The impact of language input can be analysed at various levels. From a theoretical linguistic point of view, the formal and structural properties of the language itself are examined, while in the more empiricist approach found in recent studies, the way these features are actually reflected in adult-to-child speech are investigated (e.g. Choi & Gopnik, 1995; Gopnik *et al.*, 1996; Tardif *et al.*, 1997; Naigles & Hoff-Ginsberg, 1998). Four main factors have been regarded as likely to favour nouns vs. verbs differently in input across languages: frequency, salience/utterance position, morphological transparency, and pragmatic salience (for a review, see Tardif *et al.*, 1997). All these language input factors

certainly play a role in noun and verb development. However, their effects on child language have not been directly demonstrated, since what has generally been shown is simply a correspondence between children's early vocabularies and adult models. In addition, input factor effects are probably the result of certain specific combinations in which not all factors have the same weight. For example, Italian is a pro-drop language with a relatively flexible SVO word order, a set of features which could favour verbs; on the other hand, it has a rich verbal morphology which jeopardizes verbs for reasons of morphological complexity. Moreover, an interesting idea that has arisen from some analyses is that linguistic input might play a different role according to word class. As suggested by Gentner, children may need more linguistic input information for verbs than for nouns because 'the assignment of relational terms is more variable crosslinguistically than that of nominal terms' (1982: 323). In a similar perspective, Maratsos (1991) highlighted the intrinsic asymmetry between nouns and verbs regarding the respective roles of conceptual and formal properties, and assumed that noun and verb acquisition differ in that noun acquisition is mostly based on the perception of the conceptual-semantic core, while verb acquisition is mostly based on the analysis of formal-grammatical properties.

Grammatical development of nouns and verbs

A complementary way to approach noun-verb learning asynchrony is to investigate not only how conceptual but also how formal noun and verb categories develop in the child's output. Surprisingly, although a number of studies have been concerned with the development of grammar in young children's speech, few have examined noun and verb grammatical development in the perspective of the noun-verb asynchrony issue. A main contribution can be found in Tomasello's recent work. In his (1992) case study of the early grammatical development of an English-speaking child between the ages of 1;0 and 2;0, Tomasello provided a detailed account of verb learning as compared to noun learning. Claiming that the major organizing element in children's early grammar is the semantics of individual verbs rather than abstract, syntactic categories (what he called the Verb-Island Hypothesis), he argued that, although the child under study showed evidence of having a paradigmatic noun category, there was very little evidence of a paradigmatic verb category at 2;0. Evidence of the asymmetry in the child's construction of the two grammatical categories was found in analyses of her use of argument structure and morphology. Tomasello noted that the child seemed to use her nouns in all kinds of new argument roles fairly soon after she learned them, but she used her verbs only in ways that she had heard them used. Regarding noun and verb morphology, he argued that the child used plural -s and possessive -s markers on nouns starting at around 1;5-1;6, while her first contrastive marking of verbs with the

progressive *-ing* and past tense *-ed* endings appeared at 1;7–1;8 and 1;8–1;9, respectively, and were used sporadically thereafter. Further evidence that children learning English have word classes of nouns quite early, while not forming verb categories until later was provided in a series of experimental studies where the linguistic context within which children learned novel nonce nouns and verbs was controlled (Olguin & Tomasello, 1993; Tomasello & Olguin, 1993; Tomasello, Akhtar, Dodson & Rekau, 1997). Children under 2;0 (Tomasello *et al.*, 1997) were found to combine the novel nouns productively much more often than they did the novel verbs, and to pluralize a newly learned noun but not to form a past tense with a newly learned verb.

Present study on French

The present study addresses the question of noun–verb asynchrony in French. It is directly in line with previous research (Bassano, 1998a; Bassano, Maillochon & Eme, 1998) which investigated early lexical development in French in longitudinal (from 1;2 to 2;6) and cross-sectional (focusing on ages 1;8 and 2;6) studies of children’s natural productions. These studies indicated that, among the four categories of words identified in children’s speech, nouns and the so-called para-lexical elements (including interjections, fillers, social-interactive particles and conventional expressions) were predominant until 1;8 and decreased over time, while predicates (including verbs and adjectives) and grammatical words were less frequent at first but strongly increased later. The present longitudinal study focuses on the development of nouns and verbs. The question of noun–verb asynchrony is addressed with the idea that finding out how these word classes are acquired involves investigation of both the semantic and grammatical developmental processes, and of the relation and interdependencies between these processes. In this perspective, we conducted analyses of nouns and verbs across development, how their semantic subclasses developed, and the acquisition of noun and verb grammar, while looking for possible effects of lexical-semantic properties on grammaticalization.

Our analyses of the semantic development of nouns and verbs were more particularly designed to determine to what extent concrete object names and concrete action verbs form the core part of the notional classes of nouns and verbs, respectively. In this view, we adopted a traditional conception of the noun semantic category, calling nouns those terms that refer to entities and that are nouns in French adult language (including proper names, but excluding pronouns which are sometimes included in the nominal category). We analysed nouns produced by the child using a three-level classification which takes into account this semantically well-structured domain. This classification mainly distinguished animate vs. inanimate names, as well as proper vs. common names and concrete vs. abstract names, making six possible subclasses of nouns, among which is the subclass of concrete object

names. To define the verb category, we adopted a similar conception, considering as verbs those terms that are verbs in French adult language and that denote events in a broad sense (here including auxiliaries and verbs of being, but excluding adjectives). As noted above, the semantics of verbs is much more complex than the semantics of nouns, and a number of dimensions have been proposed across studies. To give the simplest account of all the verbs used by the child and to facilitate the comparison with nouns, we proposed a two-level classification. We first used a classical distinction that opposes action verbs to other verbs. These other (non-action) verbs include verbs of being and modals or auxiliaries, i.e. verbs which have a grammatical function and form a kind of closed class of verbs. Then we introduced some more data-driven distinctions among action verbs, namely, we distinguished concrete action verbs, which denote perceptible actions or results of actions, from abstract action verbs, which denote actions with no perceptual appearance, and also from highly contextualized situational verbs, which are somewhat marginal action verbs.

Turning now to grammatical properties in oral French, the noun class is mainly characterized by the fact that nouns are usually preceded by determiners, which are marked for gender and number. Only some nouns themselves are marked for gender, and the written *-s* plural marking is not pronounced. Thus, the preceding determiner serves as the grammatical constraint on nouns. This constraint, however, is not general, and determiners are not used in certain contexts, such as before proper names or names used as proper names, within verbal-nominal expressions (e.g. *avoir faim* 'to be hungry'), or after some prepositions (e.g. *sans sucre* 'without sugar'). In contrast to the relative simplicity of the grammatical noun class, the grammatical verb class in French is characterized by a number of distributional and morphological properties, mainly mood, tense, and person markings. Person is indicated primarily by the person pronoun that must be used in the absence of a noun phrase subject. Here we focus on the verb tense system, considered the main aspect of verb grammar. Traditionally, French verbs are assigned to one of three conjugations. The first group, with verb infinitives that end in *-er* (like *jouer* 'to play'), is the most regular, the largest (about 4000 verbs, or 90% of all verbs) and the productive conjugation. The second group is also regular, with verb infinitives that end in *-ir* (like *finir* 'to finish'), but small, consisting of some 300 verbs. The third group contains the remaining verbs, some 350, all irregular in form, most of them very common in everyday use, such as modals and the verb *être* 'to be'. The complete written French inflectional verb system includes 23 simple and compound, finite and nonfinite forms (cf. Table 1, from Béchade, 1992: 248). However, some ten or twelve forms only are commonly used in standard oral French, moreover, a number of them are homophones, i.e. identical in spoken language. Commonly used simple forms are the present indicative,

TABLE 1. *Mood and tense forms in the French verb system* (from Béchade, 1992: 248)*

Moods	Simple tense forms	Compound tense forms
Indicative	present (je joue [zu]) imperfect (je jouais [zue]) future (je jouerai [zure]) simple past (je jouai)	compound past (j'ai joué [e zue]) pluperfect (j'avais joué [ave zue]) future perfect (j'aurai joué [ore zue]) past anterior (j'eus joué)
Imperative	present (joue [zu])	past (aie joué)
Conditional	present (je jouerais [zure])	past (j'aurais joué)
Subjunctive	present (que je joue) imperfect (que je jouasse)	past (que j'aie joué) pluperfect (que j'eusse joué)
Infinitive	present (jouer [zue])	past (avoir joué)
Participle	present (jouant) past (joué [zue])	past (ayant joué)
Gerund	present (en jouant)	past (en ayant joué)

* Commonly used forms are in bold; examples are given for a verb of the 1st conjugation (*jouer* 'to play'), in the 1st person singular.

used for actual time events, and the present imperative, used for requests, which are both unmarked forms in the singular (sounding as bare stems), are homophonous for almost all verbs, and differ only in the presence of an obligatory subject in the indicative form (e.g. indicative *je/tu/il joue* and imperative *joue* are homophonous). Other relatively commonly used simple forms are the more marked forms of the indicative imperfect and simple future, used for past and future events, and the conditional present, used for hypothetical events. The future and the conditional present are homophonous in the singular for all verbs. Nonfinite forms of the present infinitive and the past participle, which are generally barely marked and are homophonous for a large majority of the verbs (verbs in the first group: e.g. infinitive *jouer* and past participle *joué* are homophonous), are often used, mostly in compound and periphrastic forms. Compound forms require the auxiliaries *être* 'to be' and *avoir* 'to have' plus the past participle of the verb. A very commonly used compound form in spoken French is the compound past, used for past events with completive aspect. Two other compound forms are the pluperfect and the future perfect. Other forms very commonly used are periphrastic constructions (not presented in table 1), namely the periphrastic future, requiring the auxiliary *aller* 'to go' plus the infinitive of the verb (e.g. *je vais jouer*), and modal constructions, requiring a modal verb plus the infinitive of the verb (e.g. *je veux jouer*).

Some information about noun and verb grammatical development in French is available in both older and recent studies. Grégoire's (1947) study of his two sons mentioned phenomena such as the initial omission of articles, person pronouns and auxiliaries, as well as the early appearance of indicative present verb forms, and of past participles and infinitives, used alone to refer

to past and to present or future events, respectively. More recent studies taking a constructivist approach to language acquisition and generally based on longitudinal analyses of one French child (e.g. Veneziano, Sinclair & Berthoud, 1990; Bassano, 1996; Kilani-Schoch, de Marco, Christofidou, A., Vassilakou, Volmann & Dressler, 1997; Veneziano & Sinclair, 1997) provided detailed accounts of the emergence of grammatical morphemes. In particular, the emergence of morphemes such as articles, pronouns, auxiliaries and modals was shown to be linked to a phenomenon which appears very early in language development, namely the production of short elements at the beginning of words, referred to as 'filler syllables' following Peters & Menn (1993). In the present study, early grammatical development of nouns and verbs is documented from the perspective of a comparison of the developmental course of these two word classes, and with a focus on the relation between grammatical and semantic development. From the above descriptions, it can be expected that, overall, grammaticalization is likely to be more complex and to be achieved later for verbs than for nouns. According to the dependency hypothesis, grammaticalization processes could be expected to depend on lexical-semantic properties of nouns and verbs, to some extent.

METHOD

Subject

The girl, Pauline (the same child studied in Bassano, Maillochon & Eme, 1998), was the youngest of four children in a middle-class family living in Rouen. At the time of the study she was attending a nursery school.

Data collection

The child was audio- and video-recorded twice a month at home, during everyday activities, such as eating, playing, washing, dressing, etc., in unstructured interactive sessions (of about two hours each) with her family. Long uninterrupted parts of each recorded sessions were selected for transcription so as to obtain a variety of situations and a sufficient and representative number of productions. They were fully transcribed, with indications about the situations, contexts, and gestures, and then stored on computer in CHAT format (MacWhinney, 1995). Appendix 1 gives detailed information about the transcribed corpus, using months as the unit of analysis (combining the two sessions per month). For each month's data, we calculated the total number of vocal productions, and the number of utterances, defined as the number of productions that were linguistic. To qualify as an utterance, a production had to be a prosodic and meaningful unit that included at least one element resembling a French 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. In addition, two main indices of the child's general language development across time were calculated (see Appendix 1): the utterance/minute rate, which is an index of the child's linguistic productivity, and the mean length of utterance (MLU in words), the most classical index of young children's language level.

Data sampling

The analyses were conducted on monthly samples, each consisting of a constant number of 120 utterances selected from the transcribed session. 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 *et al.*, 1998) and that is justified for two main reasons. First, this method allows studies which would not otherwise be possible if working on extensive transcriptions, namely studies using a relatively large number of subjects, or longitudinal case-studies of relatively long duration (such as the present one, for which a total of 1920 utterances to be analysed was yielded). Second, it allows a maximal comparability across subjects for studies involving different children, and across ages for longitudinal studies. An apparent disadvantage is that the linguistic material is reduced. However, it can be noted that even extensive sessions are restricted samples of the child's production at a point in time, and that samples of 120 utterances provide good amounts of materials as they are. Moreover, the reduction inconvenience was neutralized for a large part by the way the selection was done. Indeed, samples were the more naturally and representatively constituted, since they were formed by all the utterances of several long and uninterrupted discursive sequences (excluding incomprehensible productions when necessary, as explained above), with the sequences being chosen among the most interesting and the richest parts of the transcription.

Coding and analysis

The analyses of noun and verb production were based on a main coding consisting of a sequential description of all the units in each utterance, with indications of their grammatical class (e.g. *determiner*, or *noun*, or *verb*), their morphological properties if necessary (e.g. for a verb: *present*, *indicative*, *third person singular*), and their lexical content (e.g. *lapin* 'rabbit', or *manger* 'to eat'). The semantic subclass of each noun or verb was indicated by means of a specific coding that distinguished the categories mentioned above, e.g. common concrete animate, common concrete inanimate, concrete action verb, abstract action verb, etc. Since some distinctions cannot be made independently of the way words are used, codings were made using linguistic and extra-linguistic contextual information provided by the method of detailed transcription. Thus, the same word could be classified into different

semantic subclasses depending on the context. For example, the verb *regarder* 'to look' was classified as a situational verb when it was used in highly contextualized imperatives, and as a concrete action verb in other cases. Moreover, two semantically close words could be classified in different subclasses; for instance, in contrast with the verb *regarder*, the verb *voir* 'to see' was usually considered as an abstract action verb. A similar use of contextual information was required for grammatical codings, especially for homophonous verb forms. For example, 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. On the other hand, 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. In the same way, an utterance like *papier maman* 'paper mummy' (1;4) could be clearly interpreted as a request from the context, with *maman* being a vocative that did not require any determiner. Codings were made independently by two people including the author (initial agreement rate of about 90%) and discussed until complete agreement was reached.

The frequencies of nouns and verbs and their subclasses were calculated by analysing types and tokens, respectively. As customary for these analyses, the different phonological forms that could be produced for a word (e.g. *apin* for *lapin*), as well as the differently inflected forms of a verb (e.g. *mange* and *manger*) were considered as the same word type.

RESULTS

Overall development of nouns and verbs: type and token frequencies

The child's production of nouns and verbs between 1;2 and 2;6 is shown across five consecutive age periods (P1: 1;2-1;5, P2: 1;6-1;8, P3: 1;9-1;11, P4: 2;0-2;2, P5: 2;3-2;5) in Figure 1, for types and tokens, respectively.

These analyses indicate that nouns and verbs have quite different developmental patterns, as shown in Figure 1. Noun-type production was relatively important as early as period 1, but it did not increase over the first four periods (except for a slight peak in period 2) and only presented substantial growth at the onset of the last period. Although verbs were found as early as the first period, verb-type production was initially lower than noun-type production but showed a regular increase over time and shot up at the onset of the last period, so that verbs gradually overtook nouns (monthly mean for each period, in chronological order: 13.5, 17.7, 14.7, 14.7, 27.7 for nouns, and 5.3, 7.7, 12, 13.3, 28.7 for verbs). A 2-way ANOVA on noun and verb types (2 word-classes, 5 age-periods) yielded significant main effects for word class ($F = 12.36$, $p < 0.01$) in favour of nouns (monthly mean: 17.4 for noun types vs. 13 for verb types), and for age period ($F =$

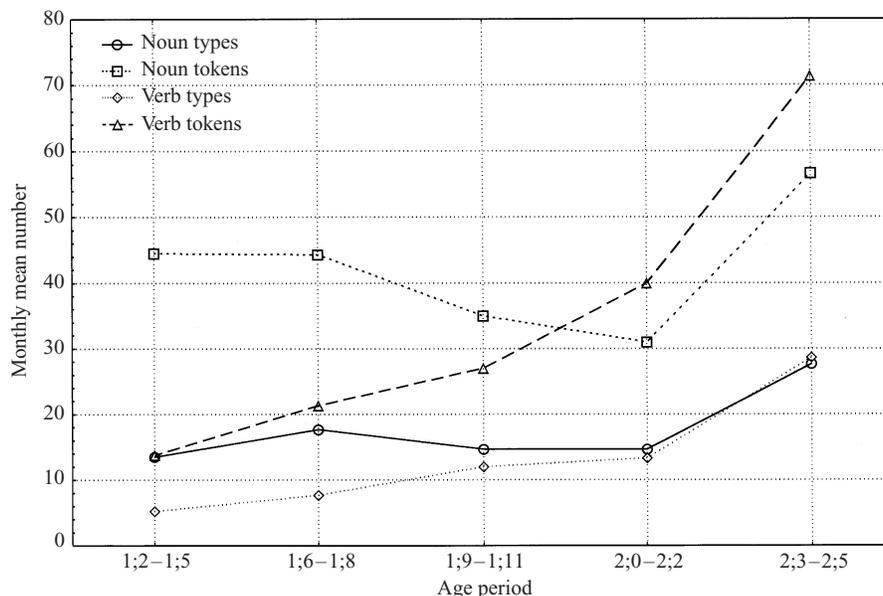


Fig. 1. Changes in noun and verb type and token frequencies in Pauline's corpus from age 1;2 to 2;6.

25.09, $P < 0.001$). The effect of age was significant both for noun types ($F = 6.46$, $p < 0.01$), and for verb types ($F = 33.12$, $p < 0.001$). However, when only the first four age periods were considered, the effect of age was significant for verb types ($F = 5.43$, $p < 0.05$) but not for noun types. The difference between the noun and verb frequencies was significant only in periods 1 ($F = 12.90$, $p < 0.01$) and 2 ($F = 14.21$, $p < 0.01$), with nouns outnumbering verbs.

The contrast in the developmental patterns of nouns and verbs was even more apparent in the token analyses. Noun-token frequencies, at first largely predominant, decreased from period 1 to period 4, while verb-token frequencies increased, so that verb frequency exceeded noun frequency from 2;0 on. Both nouns and verbs exhibited a sharp increase in token production in period 5 (monthly means: 44.5, 44.3, 35, 31, 56.7 for nouns, and 13.8, 21.3, 27, 40, 71.3 for verbs). The ANOVA on tokens yielded a significant main effect of word class ($F = 6.19$, $p < 0.05$) in favour of nouns (monthly mean: 42.4 for noun tokens *vs* 33.4 for verb tokens), a significant main effect of age period ($F = 24.17$, $p < 0.001$), and a significant word-class by period interaction ($F = 9.04$, $p < 0.01$). Significant effects of age period were found for both noun tokens ($F = 4.05$, $p < 0.05$) and verb tokens ($F = 33.09$, $p < 0.001$), but again, the effect of age on the first four periods was significant for verbs only ($F = 8.17$, $p < 0.01$). Significant differences in noun and verb

token frequencies – with nouns exceeding verbs – were found only in periods 1 ($F = 27.24, p < 0.001$) and 2 ($F = 11.43, p < 0.01$).

In sum, although verbs were noted as early as the first recording, nouns significantly predominated over verbs in this child's early productions both in number of types and in number of tokens until at least 1;8. But from around 2;0 on, verbs became equal with nouns in type frequencies, and surpassed them in token frequencies. Marked increases in type and token frequencies occurred from 2;3 on in nouns and in verbs. A look at the type/token ratios indicated at first glance that as a whole, nouns and verbs did not differ strongly in lexical diversity (mean type/token ratio: 0.42 for nouns and 0.38 for verbs). However, the evolution across age periods was quite different for nouns and verbs. The noun type/token ratio increased constantly (monthly means: 0.30, 0.40, 0.42, 0.47, 0.49) while the verb type/token ratio, initially higher than the noun ratio, exhibited an up-and-down pattern without a developmental progression (monthly means: 0.38, 0.36, 0.44, 0.33, 0.40). These numbers suggest that nouns clearly develop in lexical diversity from 1;2 to 2;6, while verbs tend to develop in both lexical diversity and in frequency of use.

Semantic development of nouns

The theoretical three-level classification along the animate/inanimate, proper/common, and concrete/abstract dimensions resulted in fact in five subclasses of nouns: among nouns referring to animates: proper names, concrete common nouns, and abstract common nouns; and among nouns referring to inanimates: concrete and abstract common nouns. Proper names referring to inanimates were not found. The developmental picture provided in Figures 2a and 2b shows how the five different semantic noun subclasses are distributed across the age periods in Pauline's corpus, for types (2a) and for tokens (2b).

Qualitative analyses indicated that, among nouns denoting animates, a notable frequency of proper names was found as early as the first period, during which they were used to refer to other members of the family, exclusively and repeatedly. After a time of disappearance, they started up again but were more diversified at around 2;0, when the child began to refer to herself as well as to symbolic persons and more distant individuals. Common nouns referring to concrete animates were frequent in both types and tokens, particularly during the first two periods, but tended to decline afterwards. They first designated close human beings, often being used like proper names (*maman* 'mommy', *bébé* 'baby'), and then later animals, usually referred to in play situations such as doing a puzzle or looking at a book. More abstract nouns referring to animates (*bête* 'animal', *copain* 'friend') were late and very infrequent. Regarding nouns denoting inanimates, all were common names, largely concrete object names. As early

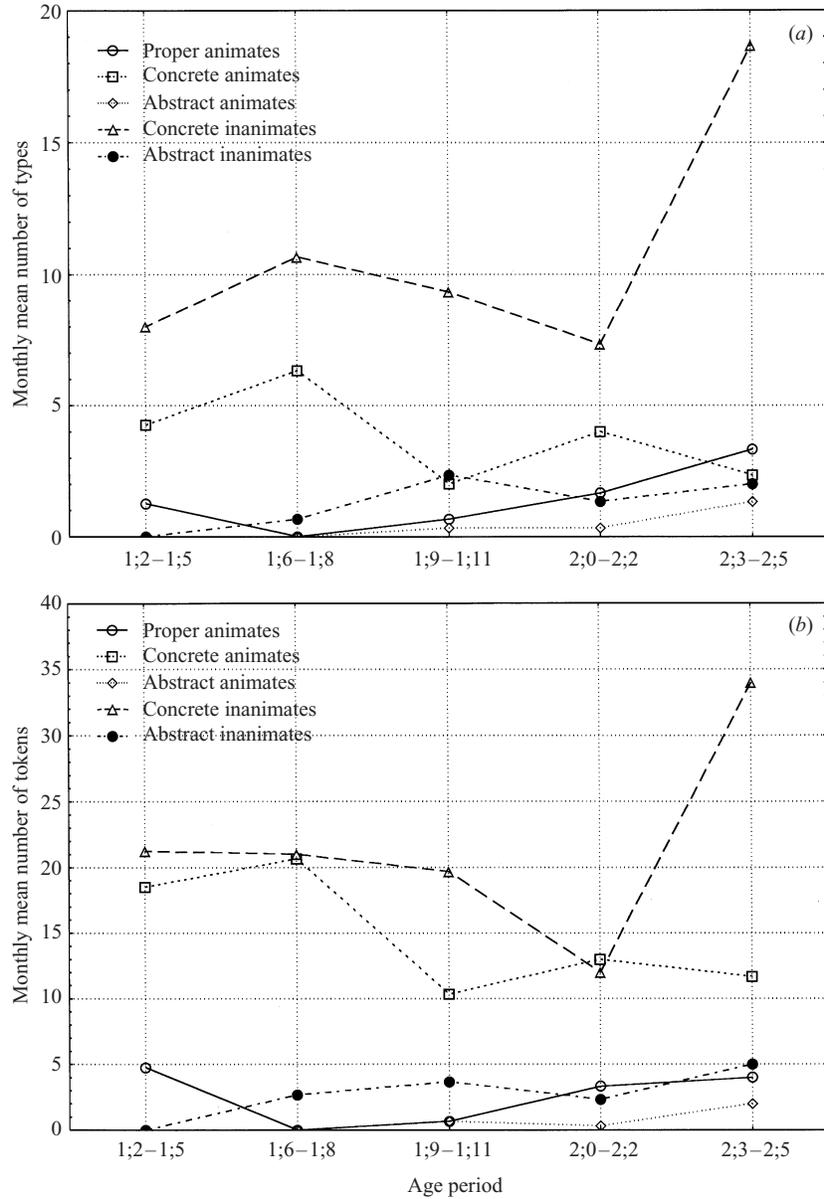


Fig. 2a, b. Changes in frequencies for the five subclasses of nouns.

as the first two periods, the child entered semantic fields such as body parts, food, clothes, furniture, and other objects and places related to everyday life (*pied* 'foot', *pomme* 'apple', *chambre* 'room'), and progressively enriched all

these fields as she extended her experience and activity, so that she produced this kind of nouns in very large quantities during period 5. More abstract nouns referring to inanimates, which started to be produced in period 2, clearly had an experiential basis, referring to sequences of everyday actions or situations (*histoire* 'story') or evocating emotions (*peur* 'afraid') often in predicative constructions. Although abstract nouns for inanimates were more frequent than abstract nouns for animates, and increased with age, they remained scarce.

Several series of statistical analyses were conducted on these data. First, two 2-way ANOVAs on noun types and noun tokens, respectively (2 semantic categories, 5 age periods) were designed to examine the role of the main animate/inanimate distinction. Both for type and token frequencies, they showed significant effects of period ($F = 6.5$, $p < 0.01$ for types, and $F = 4.1$, $p < 0.05$ for tokens), significant effects of semantic category ($F = 33.2$, $p < 0.001$ for types, and $F = 8.5$, $p < 0.05$ for tokens) with inanimates prevailing, and significant category by period interactions ($F = 3.5$, $p < 0.05$ for types, and $F = 4.3$, $p < 0.05$ for tokens). Further analyses were designed to test the subclass distributions. The 2-way ANOVAs on noun types and noun tokens (5 semantic classes, 5 age periods) both yielded significant effects of semantic class ($F = 103.5$, $p < 0.001$ for types, and $F = 81.9$, $p < 0.001$ for tokens) and age period ($F = 6.5$, $p < 0.01$ for types, and $F = 4.1$, $p < 0.05$ for tokens), and significant class by period interactions ($F = 5.6$, $p < 0.001$ for types, and $F = 3.3$, $p < 0.001$ for tokens). Specific effects of age period upon each semantic class were examined through analyses of the period 1 vs. period 5 contrast. They indicated significant developmental changes in type frequencies for concrete inanimate nouns ($F = 27.3$, $p < 0.001$) and abstract animate nouns ($F = 5.6$, $p < 0.05$) only, and significant changes in token frequencies for concrete inanimate nouns ($F = 6.5$, $p < 0.05$), concrete animate nouns ($F = 5.3$, $p < 0.05$), abstract animate nouns ($F = 8.1$, $p < 0.05$), and abstract inanimate nouns ($F = 5.1$, $p < 0.05$).

What emerges from these data is that nouns referring to inanimates were more frequent overall than nouns referring to animates, in type (68%) and token (57%) production. Inanimate concrete common nouns, i.e. object names, largely predominated overall (61% of all noun types and 51% of all noun tokens), while animate concrete common nouns had lower frequencies (22% of noun types and 35% of noun tokens), and the other three subclasses occurred at very moderate rates. However, as suggested by the period by semantic class interactions noted above, this overall result must be qualified by developmental analyses. In the first two age periods, the differences between animates and inanimates in type and token production was not statistically significant. This means that, at the emergence of language, noun production is divided into object names and animate names (mostly proper and almost-proper names). The prevalence of inanimates became statistically

significant in period 5 only, both for types ($F = 27.9$, $p < 0.001$) and for tokens ($F = 18.3$, $p < 0.01$). Yet animates remained frequent in the tokens. This finding is related to the generally high frequency of use of animates (the type/token ratio was 0.49 for inanimate concrete nouns and 0.25 for animate concrete nouns), due in particular to the repeated use in early language of some nouns, such as *maman* ‘mommy’, *papa* ‘daddy’, and *bébé* ‘baby’.

Semantic development of verbs

Five semantic subclasses of verbs were identified in Pauline’s corpus: among non-action verbs, modals or auxiliaries, and verbs of being, and among the broad class of action verbs, situational verbs, concrete action verbs, and abstract action verbs. The developmental distribution of the five subclasses of verbs is shown in Figure 3a for types and 3b for tokens.

Among non-action verbs, four modal types were found in the child’s corpus: *vouloir* ‘want’ appeared as early as 1;3 and was used frequently (41 tokens), *pouvoir* ‘can/may’ and *savoir* ‘to know’ appeared at about 1;8, and the impersonal *falloir* ‘must’, at 2;5. Auxiliaries, used in compound tense forms, were produced late and infrequently: *être* ‘to be’ from 2;0 on (4 tokens), *aller* ‘to go’ from 2;3 on (6 tokens), and *avoir* ‘to have’ at 2;5 (2 tokens). The verbs *être* ‘to be’ and *avoir* ‘to have’, used as main verbs, were the only verbs of being strictly speaking found in the child’s productions. The verb *être* was produced as early as 1;2 and had a remarkable and consistently increasing frequency of use (136 tokens), whereas *avoir* was first produced at 1;11 and was employed much less (14 tokens). Both were used first in formulaic expressions, such as *ça y est* ‘all done’ or *il y a* ‘there is/are’, before they entered more diversified phrases. Among the action verbs, situational verbs characterized by a strongly contextualized use are at the boundary of action verbs. This category included stereotyped imperatives, found as early as 1;2–1;3 and used solely in exchange situations to express attention request, i.e. *donne* ‘give’ (14 tokens), *tiens* ‘take’ (34 tokens), *regarde* ‘look’, *fais voir* ‘let’s see’, as well as some verbs in expressions that often have a modal value, such as *j’arrive pas* ‘I can’t’, *j’aime pas* ‘I don’t like’. Situational verbs, produced early and with the largest frequency during period 1, tended to decline over time. By contrast, concrete action verbs had relatively low frequencies at first and increased with time. Although some could be found in period 1 (*chanter* ‘to sing’, *cacher* ‘to hide’), these core action verbs started to develop in period 2, where they were used to refer to everyday life activities related to eating, playing, dressing, moving, etc. (e.g. *manger* ‘to eat’, *colorier* ‘to colour’, *habiller* ‘to dress’, *tomber* ‘to fall’, *courir* ‘to run’), and reached a peak in period 5. It is worth noting that concrete action verbs could have a more or less specific meaning: for example, *coiffer* ‘to hairdress’, *manger* ‘to eat’, and *courir* ‘to run’ are highly specific activities, while *aider* ‘to help’, *mettre* ‘to put’ and *poser* ‘to put down’ are not related

EARLY NOUN AND VERB DEVELOPMENT IN FRENCH

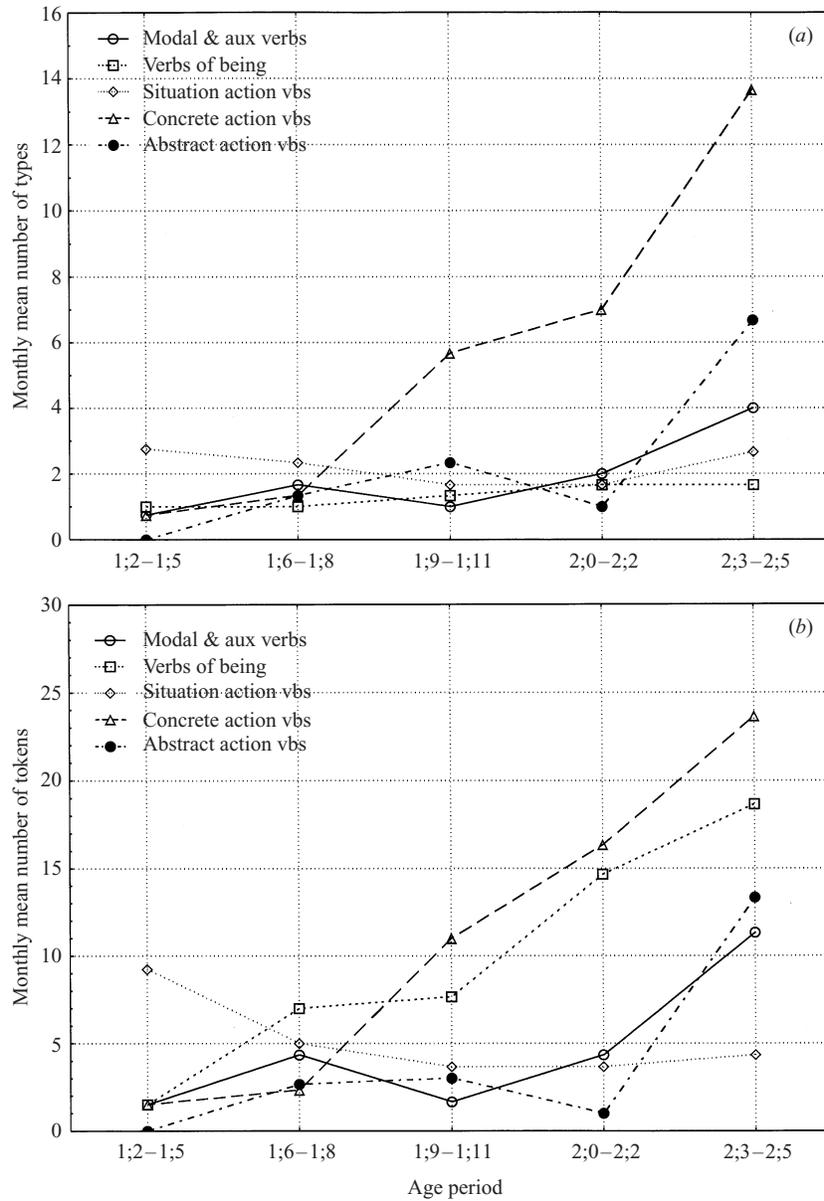


Fig. 3a, b. Changes in frequencies for the five subclasses of verbs.

to specific objects, and can therefore be seen as having a certain degree of abstractness. Moreover, some action verbs (5 %) are used in contexts where they actually refer to states, as is the case in verbs employed in perfective

forms such as *tombé* 'fallen' and *cassé* 'broken'. The subclass of abstract action verbs was defined as containing verbs denoting actions with no perceptual appearance. Abstract action verbs started in period 2, and like concrete verbs, increased over time and rose sharply in period 5. They first consisted of perception verbs, such as *entendre* 'to hear' or *voir* 'to see'. Mental and opinion verbs such as *trouver* 'to find', *essayer* 'to try', *réfléchir* 'to think', and *connaître* 'to know' were used later.

To analyse the role of the main action/non-action distinction in verbs, two 2-way ANOVAs on types and tokens (2 semantic categories, 5 age periods) were first conducted. For both type and token frequencies, the analyses yielded significant effects of period ($F = 35.4$, $p < 0.001$ for types, and $F = 37.2$, $p < 0.001$ for tokens) and semantic category ($F = 103.9$, $p < 0.001$ for types, and $F = 5.3$, $p < 0.05$ for tokens), with action verbs predominating. A significant category by period interaction was obtained for types ($F = 16.9$, $p < 0.001$) but not for tokens. In addition, as for nouns, two 2-way ANOVAs (5 semantic classes, 5 age periods) were conducted on verb types and verb tokens to test the subclass distributions. Significant main effects were found for semantic class ($F = 38.8$, $p < 0.001$ on types; $F = 12.6$, $p < 0.001$ on tokens), and age period ($F = 35.4$, $p < 0.001$ on types; $F = 37.2$, $p < 0.001$ on tokens). The interactions were significant ($F = 12.8$, $p < 0.001$ on types; $F = 5.2$, $p < 0.001$ on tokens).

These results can be summarized as follows. Overall, action verbs as a broad class predominated over non-action verbs in type frequencies (76%) and to a lesser extent, in token frequencies (58%). However, the developmental analyses showed that the difference between action verbs and non-action verbs was not statistically significant in periods 1 and 2, either in the production of types or in the production of tokens. It became significant in period 3 ($F = 21.98$, $p < 0.001$), period 4 ($F = 14.71$, $p < 0.01$), and period 5 ($F = 122.8$, $p < 0.0001$) for type production only. This means that action verbs are supplemented by other verbs, particularly at the onset of verb production. This conclusion is even more obvious when only the subclass of concrete action verbs, that is, core action verbs, is considered. Although the concrete action verb subclass was the most frequent of the five verb subclasses in both types and tokens (42% of all verb types and 32% of all verb tokens), there was a relatively strong impact of various other subclasses in the course of verb development. Far from dominating the verb lexicon at first, concrete action verbs had low frequencies until the age of about 1;8. During the first two periods of the child's production, the earliest and most frequent verbs were situational verbs such as attention requests, along with certain verbs of being like *être* 'to be' and modal verbs like *vouloir* 'want'. On the other hand, abstract verbs played a non-negligible role, particularly during the last period. Moreover, verbs of being competed with concrete action verbs in token frequencies throughout the development period studied

here, as reflected by their high overall token frequency (28%) and their particularly low type/token ratio (0.14, compared to 0.52 for concrete action verbs), mostly due to the notable frequency of use of the verb *être* 'to be' in early language.

Grammatical development of nouns

Development of noun determiner use. The question raised now is whether, when, and how, terms that have the semantic properties of nouns were employed by the child in conformity with the perceivable grammatical constraint of the noun class, i.e. being preceded by a determiner. Four noun use patterns were found in the child's corpus, examples of which are provided in the Appendix 2: correct determinerless nouns, incorrect determinerless nouns, nouns preceded by fillers, and nouns preceded by true determiners. Correct determinerless nouns correspond to cases in which a determiner is not required in adult language (examples (1) to (4)). The other three patterns of using nouns were in contexts where a determiner is required in adult language. Nouns were classed as incorrect determinerless when they were produced without the required determiner and thus denoted determiner omission (examples (5) to (8)). Another pattern consisted of producing the noun with a preceding filler syllable (Peters & Menn, 1993), likely to be a prefiguration or approximation of a determiner not yet clearly defined in nature or in form (examples (9) to (12)). The last pattern was the production of an adultlike determiner, clearly and distinctly specified (examples (13) to (16)).

The frequencies by time period of the four noun use patterns are shown in Figure 4. Correct determinerless nouns were particularly frequent in period 1 (44% of all nouns), when proper names and quasi-proper names were produced the most, declined suddenly, and then gradually increased again after that. Incorrect determinerless nouns (determiner omission) were highly prevalent before 1;8, especially during period 2 (52% of all nouns), and decreased drastically after that. Prenominal filler use was found as early as period 1, and presented an inverted-U profile with a maximum in period 2 (24% of nouns), followed by a gradual and almost complete disappearance with age. Determiners were not produced in period 1 and remained very infrequent until period 4. But they increased sharply in period 5, when they were used before 51% of all nouns. As could be expected, developmental changes assessed by means of period 1 vs. period 5 contrasts were not significant for correct bare nouns or filler use. In contrast, the decrease in determiner omission was significant ($F = 5.9$, $p < 0.05$), as well as the increase in true determiner use ($F = 22.8$, $p < 0.001$).

Index of noun grammaticalization. A main index of noun-class grammaticalization was derived from this analysis. This index provides an assessment of the child's ability to use a noun determiner in mandatory

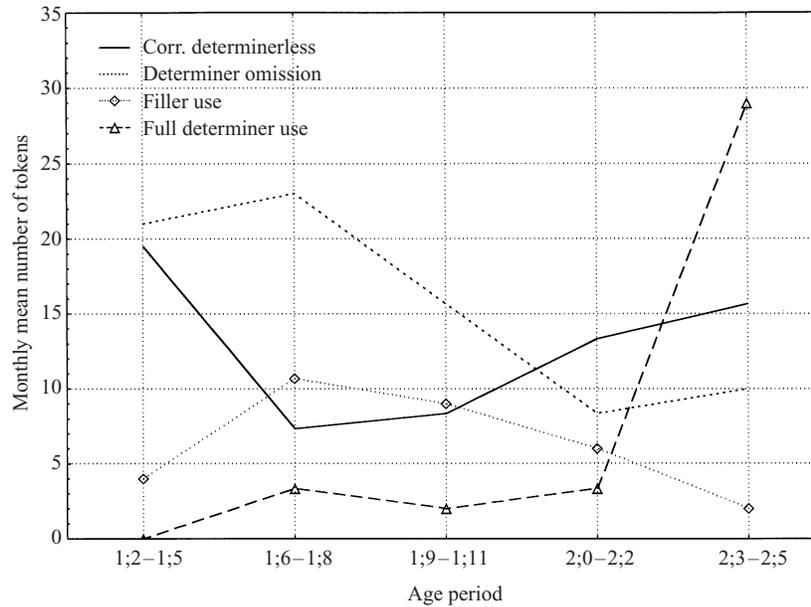


Fig. 4. Temporal distribution of the different noun use patterns observed in Pauline's corpus.

contexts. Figure 5 presents two possible versions of the index: a strict measure, calculated by taking into account only true determiners (proportion of nouns used with a determiner relative to nouns requiring a determiner in adult language), and a more generous measure that takes fillers and determiners into account (proportion of nouns used with a filler or a determiner relative to nouns requiring a determiner in adult language). The strict version of the index shows that the child did not use determiners before 1;6, used determiners at a low frequency between 1;6 and 2;0 (ratio around 0.10), at a moderate frequency until 2;3 (ratio around 0.30), and then showed a sudden sharp rise in the last few months of the study, with a ratio of 0.75 at 2;4 and 0.95 at 2;5. The 95% correct use rate in the strict version supports the conclusion that this child had almost completed the process of noun-class grammaticalization by the age of 2;6. The more generous version of the index suggests an earlier and more gradual development of noun grammaticalization, with a ratio of more than 0.10 at 1;4 and more than 0.40 at 2;0. This measure suggests that determiner use started early with fillers, a transitional pattern in which the child's first sporadic awareness of the syntactic constraint can be seen. It also indicates that syntactic constraint implementation is not a once-and-for-all development mechanism. The long-lasting coexistence of these three patterns (omission, filler, and determiner), sometimes even within the same discursive sequence, over ten

EARLY NOUN AND VERB DEVELOPMENT IN FRENCH

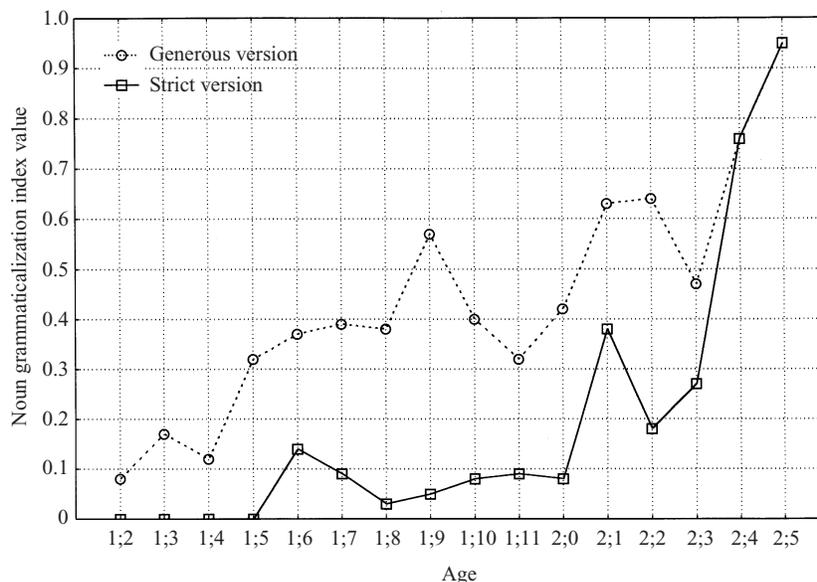


Fig. 5. Changes in the noun grammaticalization index, i.e. determiner use: strict (determiner only) and generous (determiner or filler) versions.

TABLE 2. *Noun use patterns as a function of noun semantic subclasses: tokens, for each subclass and for the animate vs. inanimate categories*

	ANIMATES			INANIMATES		
	Proper	Common	All ANI	Concrete	Abstract	All INA
Correct determinerless	42	140	182	18	12	30
Incorrect determinerless	0	70	70	172	13	185
Filler/determiner use	1	40	41	155	16	171
Filler	0	20	20	69	10	79
Determiner	1	20	21	86	6	92

months of the child's production, suggests that noun class grammaticalization is a gradual process.

Role of semantic properties in noun grammaticalization. In order to investigate the relation between semantic and grammatical noun development, we asked whether the grammaticalization process may have been influenced by certain lexical or semantic properties of nouns reflected by the subclass distinctions, such as animacy. In this perspective, we analysed (Table 2) how the different noun use patterns were distributed across the semantic subclasses of nouns. It may be noted that all proper names followed the 'correct determinerless' pattern, except for one occurrence in period 5. To evaluate

the possible role of animacy in the child's use of determiner in obligatory context, the frequencies of 'incorrect determinerless' (i.e. omission) and 'filler/determiner use' patterns were analysed as a function of the animate vs. inanimate overall categories. This interaction analysis yielded a slight but significant effect of non-animacy on filler/determiner use ($\chi^2 = 4.62$, $p < 0.05$). Inanimates tended to be produced with a filler or determiner (171, i.e. 48% of inanimates) proportionately more frequently than did animates (41, i.e. 36% of animates). This advantage of inanimates in taking a filler or determiner might be related to developmental processes mentioned above, since animates were found early while inanimates increased later. A finer-grained analysis, taking into account the distinction between determiners and fillers, showed that the highest proportion of determiners was found in concrete inanimates (26%), while abstract inanimates tended to have the highest proportion of fillers (34%). These results suggest that inanimates, especially concrete object nouns, may be the better candidates for triggering determiner use, and as such, form the basis for noun-class grammaticalization.

Grammatical development of verbs

Development of tense markings. We now examine the question of when and how, terms that have the semantic properties of verbs started to function as a grammatical class in the child's production. As noted in the introduction, verbs are characterized by a complex set of grammatical properties, including the use of subject clitics, auxiliaries and morphological markings. Here the analysis was restricted to the development of tense forms, which involve suffix and auxiliary use. The investigation of whether and how the child produced differentiated tense verbal forms posed some problems, due to the homophonous realizations of forms likely to emerge early in child language (e.g. present indicative and imperative; infinitive and past participle). As already noted in the method section, contextual information provided in the detailed transcription of natural speech enabled us to identify the function of each verb token and distinguish between forms that are formally identical in speech. On these bases, the following main verb forms were found to be produced by the child in the whole corpus before the age of 2;6 (see examples in Appendix 3): present imperative forms (examples (17) to (19)); present indicative forms (examples (20) to (22)); some imperfect and conditional forms (examples (23) to (25)); infinitive forms (used alone as simple forms, or in compound constructions: examples (26) to (34)); past participle forms (used alone, or in compound constructions: examples (35) to (40)).

Developmental analyses, presented in Figure 6, for tokens, show that present indicatives and imperatives were the two verb forms produced with some frequency before 1;6, as early as period 1. However, they had extremely different evolutions. Verbs in the imperative form, which were the

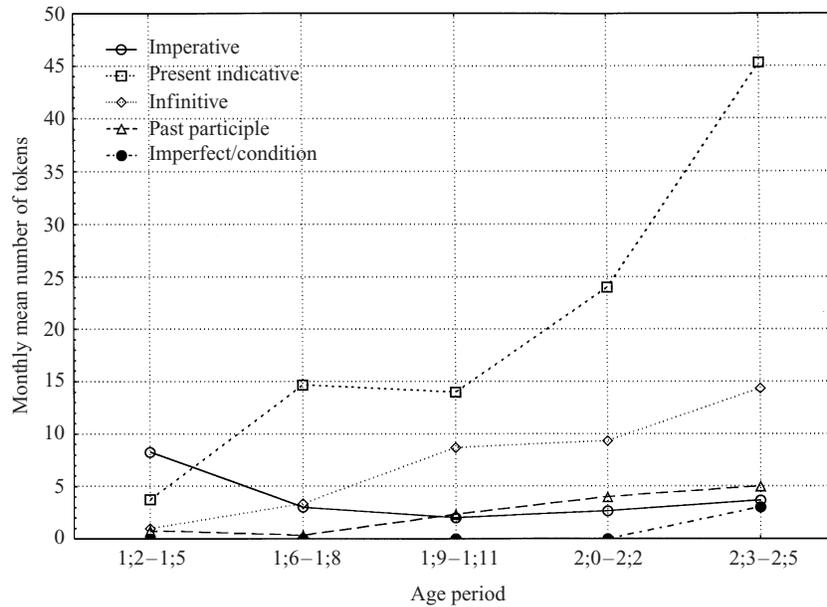


Fig. 6. Temporal distribution of the different verbal forms observed in Pauline's corpus.

most frequent in period 1, tended to diminish over time, while verbs in the present indicative strongly increased over time, showing a sharp rise in period 5. Infinitive and past participle forms, found in very small numbers in period 1, were actually produced in period 2, and increased with age regularly, with infinitives being more frequent. Finally, verbs in the conditional or imperfect did not appear until period 5, when they were still scarce. Developmental changes assessed by means of the period 2 vs. period 5 contrast in token production were significant on all of these forms, except the imperative (present indicative: $F = 80.2$, $p < 0.001$; infinitive: $F = 19.9$, $p < 0.001$; past participle: $F = 17.9$, $p < 0.01$; imperfect: $F = 5.7$, $p < 0.05$). A three step developmental pattern arises from these data: emergence of present indicative and imperative, then infinitive and past participle, and finally imperfect and conditional. It should be noted that the earliest forms to be produced were the morphologically unmarked forms (present indicative and imperative, which sound like bare stem), followed by barely marked forms (infinitive and past participle), and finally, some more marked forms (imperfect and conditional), a pattern which suggests that morphological complexity could be a determining factor in verb form appearance.

An interesting result concerning the infinitive and past participle is that these forms were used in various configurations or structures (see Appendix 3) which, in our view, reflect the emergence of compound forms. Infinitives

appeared in two main types of configurations. Most often they were used alone, i.e. they were bare infinitives (62 % of all infinitives, examples (26) to (28)) which would be incorrect in adult language and are typical of early child productions. But infinitives were also found in complete compound constructions (38 % of all infinitives): mostly, they were modal constructions (examples (29) to (31)) and periphrastic futures (examples (32) to (33)), as well as a very few prepositional constructions (example (34)). In these compound constructions, a filler could sometimes take the place of the modal. Infinitives preceded by fillers appeared first, in period 2, then infinitives preceded by modals, in period 4, and finally infinitives preceded by auxiliaries, in period 5. In a similar way, past participles were also used in two main types of configurations. The most frequent were past participles used alone, i.e. bare past participles (71 % of all past participles, examples (35) to (37)) which are incorrect uses typical of early child speech. But past participles were also found used with a preceding auxiliary or filler, from period 4 (29 % of all past participles, examples (38) to (40)). These configurations were clear compound forms, namely the compound past. A logical hypothesis that might be derived from these observations is that bare past participles are incomplete forms of the compound past in which the auxiliary is omitted. In a similar way, bare infinitives can be largely considered as incomplete modal constructions or periphrastic futures, and can be explained by the omission of the modal or auxiliary rather than by a confusion between the present and infinitive forms. This interpretation is based on the developmental patterns observed in the appearance of the different configurations: bare infinitives and bare past participles were produced as early as period 1 or, more surely, period 2, whereas compound forms emerged later – in period 2 for filler plus infinitive forms and in period 4 for clear modal constructions and compound pasts – and increased in frequency with age. The analysis of Pauline's verb forms production from 2;6 to 3;0, currently in progress, shows that bare past participles and bare infinitives decreased over time, and completely disappeared from 2;7 and 2;10, respectively, a result which appears to support the omission interpretation.

Indices of verb grammaticalization. Two main indices, shown in Figure 7, were derived from these analyses to assess the process of verb class grammaticalization. The first index measures the occurrence of marked verbal forms, i.e. infinitive, past participle, conditional and imperfect, in contrast to the earliest unmarked forms, i.e. present indicative and imperative (index 1 = proportion of marked form tokens relative to all verb tokens). This index has an up-and-down profile, but nevertheless indicates an overall increase with age, from an initial value of about 0.10 to a final value of about 0.30. The zigzag shape of this curve along with its moderate increase, due to the prominent and constant role of the present indicative, can be explained

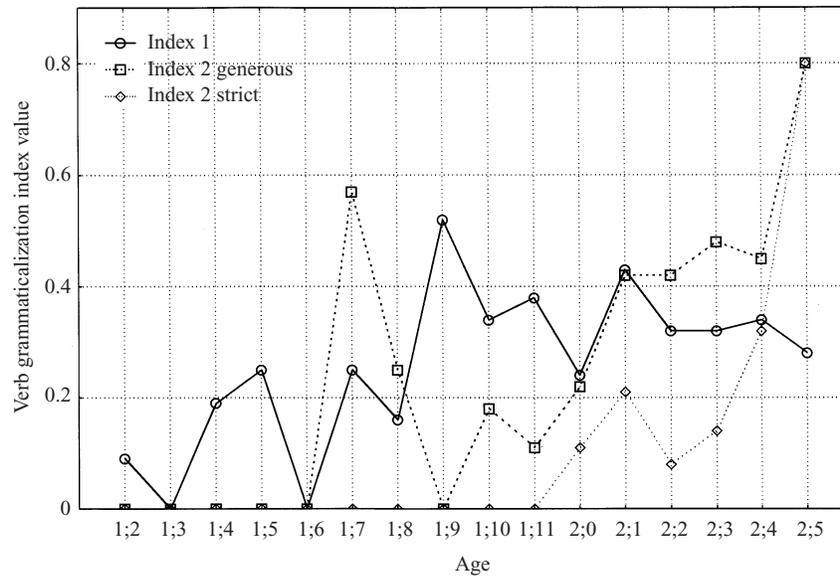


Fig. 7. Changes in the verb grammaticalization indices, i.e. marked form use (index 1) and auxiliary/modal use (index 2), strict (auxiliary/modal only) and generous (auxiliary/modal or filler) versions.

by the fact that this index is only a partial reflection of developmental changes in the child's linguistic ability. The capacity to use marked forms is a developmental step, but their use largely depends on contextual and pragmatic constraints.

The situation is different for the second index, which measures the child's ability to correctly produce complete compound verbal constructions, i.e. the ability to use a mandatory preverbal element before infinitives or past participles (index 2 = proportion of correct complete compound verbal constructions relative to all infinitive and past participle tokens). This index assesses the child's acquisition of a linguistic constraint in verb grammar, and to that extent, is comparable to the index previously calculated for assessing noun grammaticalization. Like the noun index, this verb index has two versions. The strict version only takes into account the use of true modal or auxiliary verbs, while the more generous version also takes into account the use of preverbal fillers in compound-like constructions. It should be noted that the strict version is a very strict one here, since all modals that were not phonologically perfect, as in *(v)eux boire* '(w)ant to drink', were counted as fillers although they were counted as modals in semantic analyses. The values of the strict version of the index indicates that the child did not use phonologically perfect modals or auxiliaries until the age of 2;0, and that she

used them with a low, variable frequency between 2;0 and 2;3 (mean ratio of around 0.15) and massively increased their use in the last two months of the study, with a ratio of 0.32 at 2;4 and 0.80 at 2;5. The rate of correct compound form use (80%) suggests that the child's assimilation of the grammatical constraints was growing, but was not yet completely achieved at the age of 2;6. The more generous version of the index indicates that this child's use of preverbal fillers started earlier, but not until the age of 1;6, had an up-and-down profile (mean around 0.20) between 1;6 and 2;0, and finally became more regular at a value of about 0.45 between 2;0 and 2;4.

Role of semantic properties in verb grammaticalization. Let us now look at whether the verb grammaticalization process seems to be affected by the verb lexicon. The question is whether certain semantic subclasses of verbs were more often associated with the emergence of verbal form differentiation. Table 3 shows how the different grammatical verb forms were distributed

TABLE 3. *Verbal forms as a function of verb semantic subclasses: tokens, for each subclass and for the non-action vs action categories*

	NON-ACTION VERBS			ACTION VERBS			
	Modal & auxiliaries	Verbs of being	All NON-ACTION	Situational verbs	Concrete verbs	Abstract verbs	All ACTION
Unmarked	70	144	214	84	47	31	162
Imperative	0	0	0	57	6	4	67
Present ind.	70	144	214	27	41	27	95
Marked	1	6	7	3	123	26	152
Infinitive	0	0	0	0	96	16	112
Past participle	1	0	1	1	27	10	38
Imperfect & conditional	0	6	6	2	0	0	2

across the semantic subclasses of verbs. First, the overall frequencies of unmarked forms (indicative and imperative present cumulated) and marked forms (infinitive, past participle, and imperfect and conditional cumulated) were analysed as a function of action verb vs. non-action verb overall categories. This analysis yielded a highly significant effect of activity ($\chi^2 = 142.9$, $p < 0.0001$) on the production of marked verbal forms. Non-action verbs were produced massively in unmarked forms (214, i.e. 96%) and almost never in marked forms (7, i.e. 4%), whereas action verbs were produced in near equal proportions of unmarked forms (162, i.e. 52%) and marked forms (152, i.e. 48%).

A finer-grained analysis shows how the subclasses of verbs were associated with different grammatical forms. Non-action verbs, whether modals, auxiliaries, or verbs of being, were produced almost exclusively in the present indicative (96%), except for a few imperfect forms (4%) associated with the

verb *être*. Situational verbs were mostly produced in the imperative form (66 %) and in the present indicative (30 %), except for some conditionals in formulaic expressions (3 %). Diversification of verbal forms involving the production of marked forms of the infinitive and past participle was found for concrete and abstract action verbs. In both of these verb subclasses, all four main forms were represented: imperatives (4 % and 7 %, respectively), present indicative (24 % and 47 %, respectively), infinitives (56 % and 28 %, respectively), and past participles (16 % and 18 %, respectively). It is worth noting that concrete action verbs had a particularly high proportion of infinitives reflecting modal use. These analyses suggest that concrete and abstract action verbs, and more particularly concrete action verbs, are the verbs that allow for the emergence of the infinitive and the past participle, and thus, the diversification of verbal forms. Verbs of this kind can therefore be seen as the basis for verb-class grammaticalization.

DISCUSSION AND CONCLUSION

In this study, we analysed noun and verb development in free speech samples of a French child during her second and third years of life. The first main result is that nouns largely predominated over verbs in Pauline's early productions, both in the number of types and in the number of tokens, until at least the age of 1;8. From around 2;0 on, verbs equalized nouns in types and outnumbered them in tokens. The early predominance of nouns and the later development of verbs observed for this child is in line with the results of a cross-sectional study of two groups of 12 children aged 1;8 and 2;6, respectively, in which nouns were on average about twice as frequent as verbs in the children's speech at 1;8, while verbs were almost as frequent as nouns for types, and outnumbered them for tokens at 2;6 (Bassano, 1998b). Overall, the noun-verb developmental asynchrony found here in French acquisition is in keeping with the patterns obtained for English-, Italian- and Spanish-speaking children in studies using similar or different data collection methods (e.g. Jackson-Maldonado *et al.*, 1993; Bates *et al.* 1994; Caselli *et al.*, 1995; 1999). Such cross-language consistency suggests that this pattern may extend to a larger number of Indo-European languages. Moreover, there is some evidence that the noun-verb asynchrony that appears in speech production for these languages also applies in comprehension. Gentner (1982) reported comparative data of words understood and produced by English-speaking children which indicated that production lags behind comprehension for both word classes, and that verbs lag behind nouns in both production and comprehension. Recent studies focusing on verb comprehension in English-learning children have shown that verb comprehension precedes verb production (Golinkoff, Hirsh-Pasek, Cauley & Gordon, 1987; Smith & Sachs, 1990), and that children seem to be sensitive to verbs and to their different syntactic frames at an early age (Gleitman &

Gillette, 1995; Naigles, 1996). However, a verb comprehension study using a preferential looking paradigm (Naigles, 1997) showed that significant verb learning in English does not seem to take place until the second half of the second year.

Although French production data clearly support the noun–verb asynchrony, they also indicate that some verbs are produced as early as 1;2–1;3 and that verbs are produced with a noticeable frequency by 1;8. Compared to English, for example (e.g. Bates *et al.* 1994, with the limitations involved in such an indirect comparison), nouns seem to be somewhat less frequent and verbs more frequent in French than in English in the early stages of language emergence. This and other specific tendencies (e.g. the early onset and remarkable expansion of grammatical words) were already noticed in our previous study of lexical composition in French children (Bassano, 1998a; Bassano *et al.* 1998), where nouns formed 28% and predicates 18% of children’s type vocabularies at 1;8. Such differences suggest that a common general developmental process found in the acquisition of a set of languages does not rule out some particular language-specific patterns. The question now addressed is how to account for the general noun–verb asynchrony found in French as in other languages like English, Italian, or Spanish, as well as for the relatively early verb production specific to French acquisition. We could argue that conceptual and grammatical packaging are both determining factors which interact in the acquisition of nouns and verbs. These points will be more precisely summarized in the conclusion, after a discussion of how semantic and grammatical development of noun and verb classes occurred in Pauline’s speech.

Our analyses of the semantic development of nouns and verbs were primarily designed to assess the role of concrete object names and concrete action verbs in the more general acquisition of nouns and verbs, respectively. They showed that overall, concrete object names formed the largest part of the child’s noun production between the ages of 1;2 and 2;6 (around 60% of all types and 50% of all tokens), a finding which is consistent with the most classical results in the field (see MacNamara, 1982; Maratsos, 1991; Nelson *et al.*, 1993), and which is not surprising since the child was surrounded by everyday objects that she was prompted to identify. However, what also emerges from our analyses is the important role of animate names, in particular proper and quasi-proper names, in noun production until the age of 1;8. This observation leads us to qualify the idea that concrete object names form the basis of the noun class, a restriction also stated in other ways in Nelson *et al.*’s (1993) study. This finding suggests that the referential function of language originates in two main sources, the identification of animate individuals on the one hand, and the identification of objects and object classes on the other (two semantic fields related in some ways, however, since proper nouns can be seen like very specific concrete nouns).

Similarly, concrete action verbs were the most frequent of the five verb subclasses (around 40% of types and 30% of tokens), and seemed to play for verbs the role that concrete object names play for nouns. Again, like concrete object names, concrete action verbs are not the sole – and maybe not the main – basis for the conceptual development of verbs, since they were not the earliest to be produced, nor the most frequent before 1;8. At the beginning, verb production seems to have been initiated by two kinds of verbs: situational verbs consisting of stereotyped imperatives that encode attention requests or verbs that are part of formulaic expressions, and general verbs such as modals or *être* ‘to be’. This suggests that the emergence of verbs in early speech is driven by verb elements of this kind, more than by concrete action references.

Despite the similarities just highlighted between concrete object names and concrete action verbs, do these categories differ in the role they play in noun and verb conceptual development? Maratsos (1991) proposed that noun–verb asymmetry can be seen in the fact that early noun uses ‘stick much more closely to their concrete object core than do early verblike uses’ (1991: 79). Reinterpreting MacNamara’s (1982) analysis of one of Roger Brown’s three subjects, in which it was claimed that young children avoid semantically noncharacteristic nouns and verbs, Maratsos argued that the proportions of nonaction verbs in the child’s production were relatively high and that the data supported the idea of active avoidance of nonconcrete object names, but did not support the idea of avoidance of nonaction verb uses. Our own analyses of nouns and verbs in French suggest that neither nonconcrete object names nor nonconcrete action verbs were avoided in the child’s initial speech, and we insist on the role of both of these types of words in the early stages of language acquisition. However, quantitative and qualitative comparisons of concrete object nouns and concrete action verbs argue for the conceptual noun–verb asymmetry. As the proportions mentioned above show, concrete action verbs are not as prominent within verbs as concrete object nouns are within nouns, and they are in greater competition with other verbs than concrete object nouns are with other nouns. Moreover, the domain of concrete action verbs appears more complex and diversified than the domain of concrete object nouns, for instance, it involves a number of semantic distinctions, including ones that we did not specifically analyse here such as the causative, durative, and resultative dimensions, etc. These differences are indicative of an asymmetry in the respective roles and natures of core concrete object names and action verbs, and argue in favour of differences in conceptual packaging for nouns and verbs.

Our analyses of the grammatical development of nouns and verbs primarily addressed the questions of whether the child had acquired the syntactic categories of nouns and verbs, and whether there is grammatical noun–verb asynchrony. As was highlighted in these analyses, in French, noun gram-

matalization relies on a simple main process consisting of using preceding determiners. In contrast, verb grammaticalization, even restricted to the acquisition of the tense system, involves two combined processes consisting of using inflectional suffixes on the verb to make simple forms such as the infinitive, past participle, or imperfect tense, and using auxiliaries or modals preceding the verb to make compound forms. Accurate information about how these various processes emerged and evolved in the present child was obtained through the computation of grammaticalization indices that took into account the relative frequencies of the markings used. The child produced her first adult like determiners at the age of 1;6, with four occurrences in the sample and with an example of contrastive articles on the same noun, i.e. *le chien* 'the dog' and *un chien* 'a dog'. Determiner use in mandatory contexts increased very slowly until 2;0 (10% at that time), and only one other contrast (*un pied* 'a foot' and *au pied* 'at the foot') noted at 1;11 was found before that age. A sharp rise occurred at around 2;4, and determiner use reached a rate of 95% by the end of the study. The developmental patterns here found in Pauline's determiner use seem to agree with other naturalistic observations of French children. For example, in Kilani-Schoch *et al.*'s (1997) study, Sophie was noted to have a first advance in article use by 1;11–2;0, in the so-called protomorphological phase, and to develop correct use of articles from 2;4 on, in the modularized morphology phase.

As for verbs, although some isolated examples of the infinitive and the past participle were noted earlier in formulaic expressions, Pauline produced her first reliable marked forms at 1;7, with three different verbs in the infinitive (*passer*, *caler*, *voir*) in the sample. Inflectional marking on verbs was used rather frequently between 1;7 and 2;0, but the first contrastive use on a same verb was noted at 1;10, in present indicative *met*s and infinitive of *mettre* 'to put', and it was the only contrast of this type observed before 2;0. Inflectional verb marking did not increase in relative frequencies after that point, partly for natural pragmatic reasons, and it lagged behind determiner use (but inflectional contrasts on a same verb developed, e.g. *taille/tailler*, *essaie/essayer*, *dort/dormir*, *fait/faire*, *donne/donner*, etc.). Finally, the child produced her first adult-like auxiliaries or modals at 2;0. The progression of auxiliary/modal use with verbs was remarkably similar to that of noun determiner use, showing in particular a similar explosion from 2;4 on, but always lagging behind determiner use, with a rate of only 80% at the end of the study. Some concordant results can be found in Kilani-Schoch *et al.*'s (1997) analyses of Sophie's data, who was said to have a rich premorphology on verbs, with categories such as bare past participle in addition to infinitive, imperative and present indicative in the 1;6–1;10 period, but slowly developed some paradigmatic activity together with compound past only from the protomorphological phase (1;11–2;2).

From these analyses of Pauline's data, it can be concluded that noun and

verb grammaticalization emerged side by side for noun determiners and verb inflections (with a slight advantage for the former), but that verb grammaticalization was clearly behind when it came to auxiliary and modal verb use. Regarding developmental tendencies, overall verb grammaticalization lagged somewhat behind noun grammaticalization, a tendency that became more apparent after 2;0. The verb grammaticalization lag exhibited by this French child is largely in line with Tomasello's (1992; Tomasello *et al.*, 1997) observations of English-speaking children, and seems to support the hypothesis that the formation of the verb category occurs later than that of the noun category, at least for the two languages in question here. However, Tomasello's argument mainly relied on analyses of children's speech before 2;0, whereas the French child we studied did not actually show strong evidence of noun grammaticalization until that age, and she had relatively frequent inflectional markings on verbs between 1;7 and 2;0, although these markings were not necessarily used productively. These discrepancies may be due to differences in the grammaticalization indicators available in the languages considered here, and in particular to the fact that determiner use is the only indicator for noun grammaticalization in spoken French, which has no perceivable inflectional marking on nouns. As far as determiner and modal/auxiliary use are concerned (two similar processes), the noun-verb grammatical asynchrony was confirmed here across the entire period. A likely explanation for this asynchrony is the obvious difference in complexity of noun and verb grammatical packaging. Verb grammaticalization is more complex than noun grammaticalization, if for no other reason than it requires a combination of two distinct processes, inflection and modal/auxiliary use. When a child has to produce a compound form, he/she must think about both the appropriate inflection on the verb and the appropriate modal or auxiliary. The implementation of the latter constraint, in particular, may be delayed because the child's cognitive resources are being allotted to the former, which is a process that emerges early and is particularly active in the period before 2;0 due to the relative morphological simplicity of infinitive and past participle markings in French.

Another related issue raised here concerns the nature of these grammatical processes and their relationship to conceptual and lexical development. In contrast to nativist approaches which claim that syntactic categories are innately given and initially present in the child's speech, constructivist approaches assume that children 'begin more conservatively by repeating what they have heard' (Tomasello *et al.*, 1997: 374), and gradually construct linguistic categories through their language experiences in the semantic and lexical domains (cf. Tomasello, 1992; Pine & Martindale, 1996; Lieven, Pine & Baldwin, 1997). Arguments supporting a constructivist approach to grammar acquisition can be found in the results presented here, which emphasize that noun and verb grammaticalization processes are both basi-

cally gradual, and that throughout the period under study, construction of grammatical categories can be rather precisely tracked. Progressiveness in noun-class grammaticalization was shown in particular in the role of prenominal fillers that emerged earlier and were gradually replaced by true adultlike determiners (see also Peters & Menn, 1993, and Veneziano & Sinclair, 1997, for discussion of this phenomenon). Progressiveness in verb-class grammaticalization was shown in the similar gradual appearance of auxiliaries and modals following preverbal fillers. More generally, the step-by-step development of the French inflectional verb system is apparent in the steps that could be identified in the period under study. From 1;2 to 2;6, simple forms were produced in the order: present indicative and imperative, then infinitive and past participle, and finally imperfect and conditional, while compound forms, constructed with auxiliaries or modals preceding verbs in the infinitive or past participle, started to be produced by 2;0, thus appearing before imperfect forms. From these data, we propose (see Bassano, 1998b, for more details) that the French verb tense system in the early stages develops according to the following pattern: 1) emergence of the contrast between the simple unmarked forms of the present indicative and imperative (used for the expression of actual present events), and the barely-marked nonfinite forms of the infinitive and past participle (used for the expression of non-actual events); 2) emergence of compound forms (allowing a clear distinction between past, future, and modal uses among non-actual events); and finally, 3) emergence of some more-marked simple forms, such as the imperfect, serving the expression of non-actual events. The gradual development of the verb tense system certainly depends on a series of diversified and interactive factors, including cognitive and morphological complexity, as well as discursive and pragmatic aspects (cf. Pizutto & Caselli, 1994).

Further arguments in favour of a constructivist approach can be found in the interactions between lexical and grammatical development in the noun and verb classes alike, which suggest that certain aspects of the acquisition of grammar are related to lexical development. First, a relation appears between the noun and verb grammaticalization processes and the quantitative development of the child's noun and verb lexicons. The explosions in the grammaticalization indices noted at the age of 2;4 occurred slightly after the lexical increases noted in our analyses of overall noun and verb frequencies, which occurred at around 2;3. This temporal relation is in line with the 'critical mass' hypothesis, which claims that developments within morpho-syntax are triggered by an increase in the size of the lexicon beyond a given level, and which provides supporting evidence for the interdependency of lexical and morphosyntactic development (Marchman & Bates, 1994; Bates *et al.*, 1995; Bates & Goodman, 1997). Moreover, the results presented here suggest the additional hypothesis that progress in grammatical development is related not only to quantitative lexical development, but also to lexical

structure and content. Particular semantic subclasses of nouns and verbs appear to be more likely to form the bases for grammaticalization than others. Among nouns, concrete object nouns received determiners the most frequently and can thus be regarded as the privileged type of nouns that triggers grammaticalization. Among verbs, concrete action verbs were the subclass of verbs in which marked forms were first produced the most, and they may be the privileged type of verbs that trigger diversification in verbal forms, that is grammaticalization. These effects suggest that, although concrete object nouns and concrete action verbs were not the only or earliest types of nouns and verbs in semantic development, both appeared to play an important role in grammaticalization processes and may be the semantic primers in the emergence of noun and verb grammar. They also suggest another explanation for the noun–verb grammatical asynchrony, which appears to be partly related to differences between concrete object nouns and concrete action verbs in the child’s lexical composition.

To conclude, let us come back to the idea that both conceptual and grammatical packaging are crucial interactive factors in the general and specific patterns of noun and verb development. As noted above, in French and probably in other languages like English or Italian, verbs develop later than nouns for conceptual AND grammatical reasons. In those languages, not only do verbs generally encode notions and percepts that are complex and variable across languages, as emphasized by Gentner (1982) and others, they also intrinsically encode a number of mood and tense distinctions resulting in complex morphological systems that the child has to learn when learning verbs. The relative frequency of early verbs found in French could also be related to more specific semantic and grammatical characteristics. The semantic domains of particular significance for the child, such as requesting actions or attention, locating, and indicating that a process is over, are encoded by verbs, the very same verbs that make up the child’s earliest verb productions. Regarding morphosyntax, the complexity of the inflectional system is reduced by the large number of unmarked and barely-marked forms as well as the many homophones which strongly attenuate morphological variation and are likely to facilitate a child’s verb detection. It is worth noting that similar reasons of both kinds have been set forth to account for the very high verb frequencies in extremely different languages. For example, Gopnik *et al.* (1996) argued that in Korean, concepts that are salient to young children, such as success and failure, recurrence and location, are expressed by verbs rather than by interjections or particles like in English. And although verb morphology is very rich, verbs are in a salient final position, a structural property that facilitates detection. All these analyses suggest that conceptual and grammatical packaging interact in early lexical development. A challenge for future research is to more precisely determine the place, timing, and nature of such interactions.

REFERENCES

- Bassano, D. (1996). Functional and formal constraints on the emergence of epistemic modality: a longitudinal study on French. *First Language* **16**, 77–113.
- (1998a). L'élaboration du lexique précoce chez l'enfant français : structure et variabilité. *Enfance*, **4/1998**, 123–53.
- (1998b). Sémantique et syntaxe dans l'acquisition des classes de mots: l'exemple des noms et des verbes en français. *Langue Française* **118**, 26–48.
- 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.
- Bates E., Bretherton I. & Snyder L. (1988). *From first words to grammar : individual differences and dissociable mechanisms*. New York: Cambridge University Press.
- 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*, Blackwell.
- Bates, E. & Goodman, J. C. (1997). On the inseparability of grammar and the lexicon: evidence from acquisition, aphasia, and real-time processing. *Language and Cognitive Processes*, **12**, 507–86.
- Bates, E., Marchman, V., Thal, D., Fenson, L., Dale, P., Reznick, J. S., Reilly, J. & Hartung, J. (1994). Developmental and stylistic variation in the composition of early vocabulary. *Journal of Child Language* **21**, 85–124.
- Béchade, H. (1992). *Phonétique et morphologie du français moderne et contemporain*. Paris: Presses Universitaires de France.
- Bloom, L., Tinker, E. & Margulis, C. (1993). The words children learn: evidence against a noun bias in early vocabularies. *Cognitive Development* **8**, 431–50.
- Caselli, M. C., Bates, C., Casadio, P., Fenson, L., Fenson, J., Sanderl, L. & Weir, J. (1995). A cross-linguistic study of early lexical development. *Cognitive Development* **10**, 159–99.
- 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. & Gopnik, A. (1995). Early acquisition of verbs in Korean: A cross-linguistic study. *Journal of Child Language* **22**, 497–529.
- Fenson, L., Dale, P., Reznick, J. S., Thal, D., Bates, E., Hartung, J., Pethick, S. & Reilly, J. (1993). *The Mac Arthur Communicative Development Inventories : user's guide and technical manual*. San Diego: Singular Publishing Group.
- Gelman, S. & Tardif, T. (1998). Acquisition of nouns and verbs in Mandarin and English. In E. Clark (Ed). *The proceedings of the twenty-ninth annual child language research forum*. Stanford, CA, USA.
- Gentner, D. (1982). Why nouns are learned before verbs: Linguistic relativity versus natural partitioning. In S. Kuczaj (Ed). *Language development, Volume 2*. Hillsdale, NJ: Erlbaum.
- Gleitman, L. & Gillette, J. (1995). The role of syntax in verb learning. In P. Fletcher & B. MacWhinney (eds), *The Handbook of Child Language*, Blackwell.
- Golinkoff, R., Hirsh-Pasek, K., Cauley, K. & Gordon, L. (1987). The eyes have it: lexical and syntactic comprehension in a new paradigm. *Journal of Child Language* **14**, 23–45.
- Gopnik, A., Choi, S. & Baumberger, T. (1996). Cross-linguistic differences in early semantic and cognitive development. *Cognitive Development*, **11**, 197–227.
- Grégoire, A. (1947). *L'apprentissage du langage, II*. Faculté de philosophie et Lettres de l'Université de Liège.
- Jackson-Maldonado, D., Thal, D., Marchman, V., Bates, E., & Gutierrez-Clellen, V. (1993). Early lexical development in Spanish-speaking infants and toddlers. *Journal of Child Language* **20**, 523–49.
- Kilani-Schoch, M., De Marco, A., Christofidou, A., Vassilakou, M., Volmann, R. & Dressler, U. (1997). On the demarcation on phases in early morphology acquisition in four languages. *Papers and Studies in Contrastive Linguistics* **33**, 15–32.

- Lieven, E., Pine, J. M. & Baldwin, G. (1997). Lexically-based learning and early grammatical development. *Journal of Child Language* **24**, 187–219.
- Lyons, J. (1977). *Semantics*. Cambridge: C. U. P.
- MacNamara, J. (1982). *Names for things: a study of child language*. Cambridge, MA: Bradford Books/MIT Press.
- MacWhinney, B. (1995). *The CHILDES project: tools for analysing talk*. Hillsdale, NJ: Erlbaum.
- Maratsos, M. (1982). The child's construction of grammatical categories. In E. Wanner & L. Gleitman (eds), *Language acquisition: the state of the art*. New York: C.U.P.
- Maratsos, M. (1991). How the acquisition of nouns may be different of that of verbs. In N. Krasnegor, D. Rumbaugh, R. Schiefelbusch & M. Studdert-Kennedy (eds), *Biological and behavioral determinants of language development*. Hillsdale, NJ: Erlbaum.
- Marchman, V. A. & Bates, E. (1994). Continuity in lexical and morphological development: a test of the critical mass hypothesis. *Journal of Child Language* **21**, 339–66.
- Naigles, L. (1996). The use of multiple frames in verb learning via syntactic bootstrapping. *Cognition* **58**, 221–51.
- (1997). Are English-speaking one-year-olds verb learners, too? In E. Clark (ed.), *The proceedings of the twenty-eighth annual child language research forum*. Stanford, CA, USA.
- Naigles, L. & Hoff-Ginsberg, E. (1998). Why are some verbs learned before other verbs? Effects of input frequency and structure on children's early verb use. *Journal of Child Language* **25**, 95–120.
- Nelson, K. (1973). Structure and strategy in learning to talk. *Monographs of the Society for Research in Child Development*, **38** (1–2), serial n° 149.
- Nelson, K., Hampson, J. & Kessler-Shaw, L. (1993). Nouns in early lexicons: evidence, explanations, and implications. *Journal of Child Language* **20**, 61–84.
- Olguin, R. & Tomasello, M. (1993). Twenty-five-month-old children do not have a grammatical category of verbs. *Cognitive Development* **8**, 245–72.
- Peters, A. M. & Menn, L. (1993). False starts and filler-syllables: ways to learn grammatical morphemes. *Language* **69**, 742–77.
- Pine, J. M. & Martindale, H. (1996). Syntactic categories in the speech of young children: the case of the determiner. *Journal of Child Language* **23**, 369–95.
- 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*. Hillsdale, NJ: Erlbaum.
- Smiley, P. & Huttenlocher, J. (1995). Conceptual development and the child's early words for events, objects and persons. In M. Tomasello & W. Merriman (eds), *Beyond names for things: young children's acquisition of verbs*. Hillsdale, NJ: Erlbaum.
- Smith, C. A. & Sachs, J. (1990). Cognition and the verb lexicon in early lexical development. *Applied Psycholinguistics* **11**, 409–24.
- Tardif, T. (1996). Nouns are not always learned before verbs: evidence from Mandarin speakers' early vocabularies. *Developmental Psychology* **32**(3), 492–504.
- Tardif, T., Shatz, M. & Naigles, L. (1997). Caregiver speech and children's use of nouns versus verbs: a comparison of English, Italian and Mandarin. *Journal of Child Language* **24**, 535–65.
- Tomasello, M. (1992). *First verbs: a case study of early grammatical development*. Cambridge: C.U.P.
- Tomasello, M., Akhtar, N., Dodson, K. & Rekau, L. (1997). Differential productivity in young children's use of nouns and verbs. *Journal of Child Language* **24**, 373–87.
- Tomasello, M. & Olguin, R. (1993). Twenty-three-month-old children have a grammatical category of noun. *Cognitive Development* **8**, 451–64.
- Veneziano, E., Sinclair, H. & Berthoud, I. (1990). From one word to two words: repetition patterns on the way to structured speech. *Journal of Child Language* **17**, 633–50.
- Veneziano, E. & Sinclair, H. (1997). From the surface inward: a discontinuous continuity in the emergence of grammatical morphemes. *Archives de Psychologie* **65**, 107–16.

APPENDIX 1

PAULINE'S LONGITUDINAL CORPUS: CHARACTERISTICS OF
MONTHLY TRANSCRIBED SESSIONS FROM AGE 1;2 TO 2;6

Age (yrs; mos)	Duration of transcribed session (min)	No. of productions	No. of utterances	Frequency (utter./min)	MLU (words)
1;2	46	195	137	2'98	1'12
1;3	56	183	134	2'39	1'15
1;4	45	148	129	2'87	1'21
1;5	60	212	186	3'10	1'20
1;6	29	196	166	5'72	1'29
1;7	40	306	229	5'73	1'42
1;8	33	266	235	7'12	1'46
1;9	32	300	257	8'03	1'33
1;10	67	428	386	5'76	1'50
1;11	40	374	339	8'48	1'83
2;0	30	408	383	12'77	1'93
2;1	48	413	386	8'04	2'00
2;2	43	408	390	9'07	2'27
2;3	41	437	424	10'34	2'47
2;4	58	410	398	6'86	3'39
2;5	50	415	409	8'18	3'36
Mean	45	319	287	6'71	1'74

APPENDIX 2

EXAMPLES OF THE DIFFERENT NOUN USE PATTERNS FOUND IN
PAULINE'S CORPUS*Correct determinerless nouns*

- (1) *est Claire* 'is Claire' (1;3)
- (2) *papier maman* 'paper mommy' (1;4)
- (3) *pas envie* 'don't want' (1;11)
- (4) *non, sans beurre* 'no, without butter' (2;4)

Incorrect determinerless nouns

- (5) *est chat* 'is cat' (1;6)
- (6) *c'est pied* 'this is foot' (1;6)
- (7) *veux biberon* 'want bottle' (1;6)
- (8) *moi goûter fraises* 'me taste strawberries' (2;3)

Prenominal filler use

- (9) */e/ nez* '/filler/ nose' (1;2)
- (10) */a/ lapin* '/filler/ rabbit' (1;4)
- (11) *veux /m/ biberon* 'want / filler/ bottle' (1;6)
- (12) */eum/ boîte* '/filler/ box' (2;0)

Prenominal adultlike determiner use

- (13) *les chaussures* 'the shoes' (1;6)
- (14) *mon chausson* 'my slipper' (1;8)
- (15) *c'est la culotte* 'it's the shorts' (1;11)
- (16) *cherche une serviette* 'look for a towel' (2;3)

APPENDIX 3

EXAMPLES OF THE DIFFERENT VERB FORMS FOUND IN PAULINE'S
CORPUS*Present imperative forms*

- (17) *tiens* 'take' (1;3)
 (18) *attends* 'wait' (1;10)
 (19) *raconte une histoire* 'tell a story' (2;5)

Present indicative forms

- (20) *est Claire* 'is Claire' (1;3)
 (21) *veux pas* 'don't want' (1;3)
 (22) *tu peux me tailler* 'you can point me' (2;1)

Imperfect indicative forms

- (23) *mais j'étais pas contente* 'I was n't happy' (2;5)
 (24) *il criait tout le temps* 'he was crying all the time' (2;5)

Present conditional forms

- (25) *on dirait une fille* 'it looks like a girl' (2;4)

Infinitive forms

Simple : bare infinitive

- (26) *chanter* 'to sing' (1;2)
 (27) *maman ranger* 'mommy to arrange' (1;9)
 (28) *pas mettre ça* 'not to put this' (2;4)

Compound constructions

- (29) *veux cacher* 'want to hide' (1;7)
 (30) *peux pas mettre chaussure* 'cannot put on shoe' (1;8)
 (31) *veux manger* 'want to eat' (2;0)
 (32) *vais faire ça* 'going to do this' (2;4),
 (33) *on va piquer* 'it going to sting' (2;5)
 (34) *a un mouchoir pour essuyer les larmes* 'has a handkerchief to wipe tears' (2;4)

Past participle forms

Simple : bare past participle

- (35) *caché* 'hidden' (1;4)
 (36) *non, pas fini* 'no, not done' (2;0)
 (37) *moi déjà mangé* 'me already eaten' (2;5)

Compound : compound past

- (38) */e/ porte est cassée* '/filler/ door is broken' (2;0)
 (39) *j'ai renversé* 'I have spilled' (2;5)
 (40) */m/ déjà dit ça* '/filler/ already said that' (2;3)