1. Introduction

The general background for our study is the hypothesis of Lexical Learning, i.e., the question in what ways the emergence of syntactic structure in child language is determined through the acquisition of properties of lexical (and morphological) items. In the Universal Grammar (UG) framework it has recently been argued that parameters of UG should only refer to categories of lexical items or to properties of lexical items, for example canonical government. Rizzi (1989) proposed, for example, that only heads (=X⁰) or properties of heads may enter into a UG parameter. Chomsky (1989), based on previous work by Borer (1984), explored the possibility of allowing only functional categories to be parameterized. These attempts to constrain the class of UG parameters have lead to the hypothesis that in child language development the syntax of a particular language could be determined by the acquisition of lexical and/or functional categories (X⁰).

The role subject verb agreement morphology might have for the acquisition of syntactic phenomena such as V2 and the use of empty subjects is an interesting test case for the hypothesis of Lexical Learning. In the acquisition literature on German child language, there is an ongoing debate as to the potential developmental connections between subject verb agreement and the syntax. Some authors (Weissenborn, 1990; Felix, 1987) consider subject verb agreement to be simply irrelevant for syntactic development. As opposed to this, we will try to show that the acquisition of subject agreement serves as a lexical trigger for parameter setting in German child language.

The approach we will advocate here involves the assumption that children construct syntactic representations in a gradual fashion, based on
UG principles and Lexical Learning. Our view is that specific syntactic positions gradually emerge in children's grammars. From the perspective of the Lexical Learning hypothesis, we will try to show that the emergence of syntactic properties, i.e. verb-second and empty subjects in German child language, is determined through the acquisition of properties of lexical and morphological items, namely subject verb agreement affixes.

One problem in the debate on the role of subject verb agreement in German child language is that the amount of analyzed child language data is rather limited. In order to determine when the acquisition of subject verb agreement takes place it is necessary to carry out distributional analyses and to quantify the results, because from isolated examples it is impossible to decide whether or not a morphological paradigm has been acquired. With respect to the development of subject verb agreement in (monolingual unimpaired) German children, distributional analyses of large longitudinal corpora do not exist. The quantitative results available from previous studies are based on rather small corpora (Clahsen, 1986a, 1986b) and should be tested with more data. This is what will be done in the present paper. We have been able to analyze the development of subject verb agreement, verb-second and empty subjects in the extensive longitudinal corpus collected by Max Miller, on the acquisition of German by his daughter Simone (cf. Miller, 1976). In addition to the linguistic analyses presented in an earlier version of this paper (cf. Clahsen and Penke, 1991), we carried out a statistical analysis to test our results. Simone's speech was recorded between the age of 1;7 and 3;5 at regular intervals of 1 to 14 days; the recording sessions lasted for about 2 to 4 hours. This resulted in a rather large corpus of about 1,000 to more than 2,000 utterances per month.

The paper is structured as follows. In the next section, we will briefly describe what the three phenomena (subject verb agreement, verb-second, empty subjects) that will be studied in the acquisition data look like in the adult language. Section 3 will briefly summarize results from previous acquisition studies. In section 4, the results of the Simone-Corpus will be presented in detail. In section 5, a grammatical analysis for these data will be proposed. An alternative proposal from the literature will be discussed in section 6.