## *The benefit of doubt*

We address the following problem: How should ambiguous examples in natural language production be treated, in particular when the data set is small? This question arose during the study of two children (age 3;10-6;01) who are acquiring Swiss German. The focus of the analysis is on verb placement in embedded clauses introduced by a complementizer. In contrast to the target grammar, where embedded clauses generally show the verb-final pattern, these children opt for verb movement for a considerable period of time, which results in many verb-placement errors. The verb-final pattern appears to be acquired late. As well as examples that clearly involve verb movement and those that do not, there are many examples which are ambiguous, i.e. in which one cannot determine, based on the surface string, whether verb movement has applied or not. Moreover, while the data set for one of the children (Moira) is large, that of the other child (Eliza) is relatively small.

If one omits all structurally ambiguous examples in clauses introduced by a complementizer (Approach A) produced by Moira (Fig. 1a) then examples produced before age 4;11 usually involve verb movement and those produced after age 5;1 usually involve the verb-final pattern. *Our hypothesis* is that embedded clauses produced before age 5;0 involve verb movement, whereas those produced after age 5;0 do not.

For our statistical analysis we also leave out the intermittent period (age 4;11 and 5;1) and then group count data for the early and late period. We assume that both periods are homogeneous, with a high proportion of verb movement in the former and a low proportion of verb movement in the latter. Our hypothesis stipulates that there is a considerable difference between these two proportions, while the *null hypothesis* is that the same proportion of verb movement is found in both periods, i.e. any observed differences are due to chance. We apply a proportion test. This test rejects the null hypothesis with high significance, thus strongly supporting our hypothesis.

But, although the analysis seems to provide clear evidence for our hypothesis, we have ignored the fact that there is a substantial number of ambiguous examples, which may ultimately jeopardise our hypothesis (Fig. 1b). If all structurally ambiguous examples are included in these counts–again leaving out the intermittent period (age 4;11 and 5;1)–and counting all of these ambiguous examples *against* our hypothesis (Approach B)–i.e. adding all ambiguous examples produced before age 4;11 to the verb-final examples and those after age 5;1 to the verb-movement examples–the difference between verb movement (before age 4;11) vs. non-verb movement (after age 5;1) is less striking. However, even adopting this approach, a clear difference remains between examples produced before age 4;11 and those produced after age 5;1. But now the 99% confidence interval for the difference between the proportions obtained in Approach A gives a minimum difference of 47.8 percent points, which is just enough to support our hypothesis.

In the case of Eliza there is little data (Fig. 2a/2b). Thus if ambiguous examples are analysed *against* our hypothesis, then the experimental hypothesis is difficult to support. From Fig. 2a, we define the early period as before age 5;01 and the late period as after age 5;01, leaving out 5;01. If ambiguous examples are omitted (Fig. 2a), the proportion test still rejects the null hypothesis. However, the difference between the early and late proportions is only 22.3 percent points, which is not enough to support our hypothesis. On the other hand, if the ambiguous examples were interpreted in favour of our hypothesis (Approach C), we would obtain a highly significant result. The difference between the early and late proportions is at least 48.7 percent points at the 99% confidence level (which is more than would be the case in Moira if this approach were adopted!). Clearly, interpreting ambiguous examples in this way is suspect.

We conclude that for sparse data the correct treatment of ambiguous cases is crucial. Depending on which approach we adopt (B vs. C), we can obtain anything between clear support for our hypothesis and no rejection of the null hypothesis. The "true" value, i.e. the cases where there was actually verb movement, lies somewhere between these two extremes. Thus ambiguous examples in Eliza's case lead to a high degree of uncertainty. In the light of this problem we suggest either using additional linguistic assumptions to classify ambiguous examples or increasing the data set, for instance by including another type of embedded clause which never shows verb movement in the target grammar. Thus we could add Eliza's relative clauses to those with a complementizer to boost the number of embedded clauses.

**Fig. 1a: Clauses introduced by complementizer (Moira)**

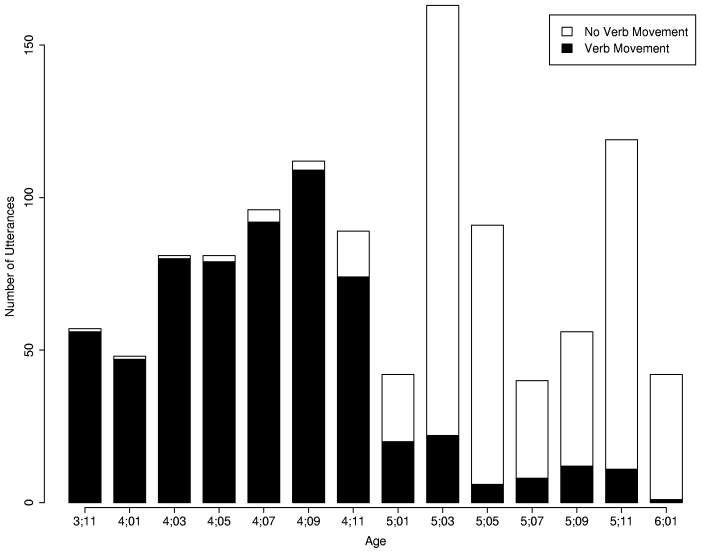


**Fig. 1b: Clauses introduced by complementizer (Moira)**

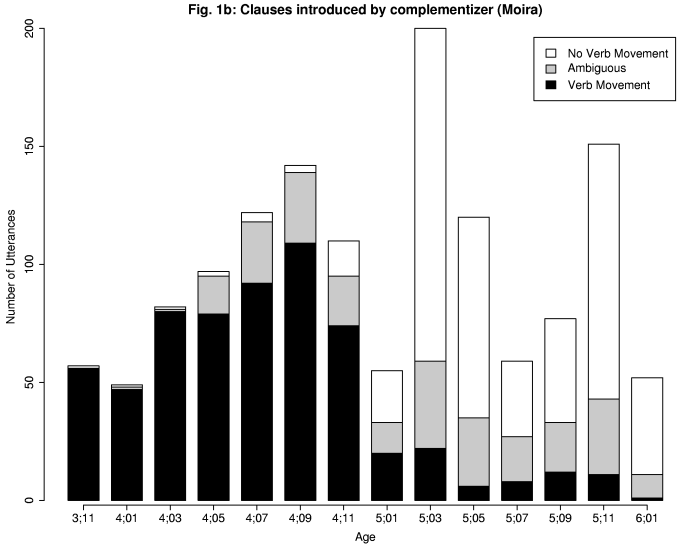


**Fig. 2a: Clauses introduced by complementizer (Eliza)**

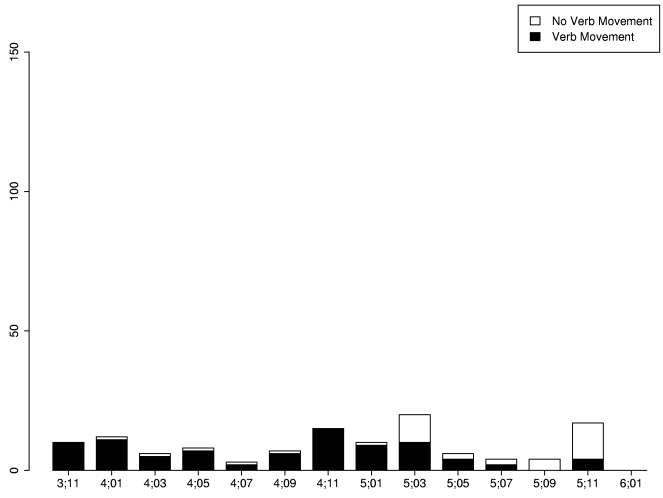


**Fig. 2b: Clauses introduced by complementizer (Eliza)**