Structure of MLU: prediction from tense and non-tense morphology
Filip Smolík (amolik@ku.edu)
Child Language Doctoral Program, University of Kansas

Summary
The structure of MLU and its growth in early language development is analyzed in this study. Using random-effects (multilevel) modeling of the relationships between MLU and measures of grammatical morphology usage, the study investigates the hypothesis that growth of MLU is primarily driven by morphosyntactic development. The results demonstrate that productive use of inflectional morphology improves prediction of MLU compared to prediction based on the age only. Although original analyses suggested that tense-related inflection is a better predictor or MLU than inflection unrelated to tense, the final results are not consistent in this respect. The findings confirm the MLU is a valid indicator of syntactic development, and that development of sentential syntax is the primary correlate of MLU growth in typically developing children.

Introduction
Mean length of utterance is widely used in child language research as an approximate measure of language developmental level. Many researchers expressed concerns about validity and reliability of the measure (e.g., Keen & Fitzgerald, 1985; Eisenberg, Fersko, & Lundgren, 2001). Some concerns about MLU include:

• it is not clear what aspects of language development MLU reflects
• utterance length is determined by the context to a large extent
• the observed variability in MLU values is large, and the relationship between MLU and age is not always clear

One problem is that most data on MLU development comes from cross-sectional studies. The observed range of MLU values in children at a particular age is due both to the measurement error and to the true differences in developmental level. This study attempts to address these problems:

• if MLU reflects morphosyntactic development, indicators of morphol-ogy use should improve prediction of MLU compared to predictions made from age alone

Data and Method
Two longitudinal corpora of spontaneous language transcripts, available from CHILDES (MacWhinney, 2000), were used:

• Manchester corpus (Theakston, Lieven, Pine, & Rowland, 2001) – a total 798 transcripts from 12 children aged 20–36 mo. used for this study
• Wells corpus (Wells, 1981) – total of 279 transcripts from 32 children were used here; age range 17–43 mo.

The data was analyzed using mixed-effects modeling (Bryk & Raudenbush, 1992; Pinheiro & Bates, 2000). All models were fit separately for each corpus:

• first step was finding the best model for prediction of MLU from MLU-referenced age (reported Smolík, 2004, this meeting)
• number of tense or non-tense inflections in a transcript was added as time-changing (level 1) covariate to the above models
• resulting models compared by overall fit indices and by effect sizes of the inflection variables (using value of $t$-statistics as an estimate of effect size)

Results
Adding the inflection measures resulted in clear improvement in fit for both corpora, both for the tense and non-tense inflection measures. The effects of inflectional variables were always significant with $p < 0.001$.

Wells corpus
• model with referenced age only as predictor: $\text{BIC} = 272.11$, $\text{logLik} = −118.71$
• model with tense inflection added: $\text{BIC} = 244.61$, $\text{logLik} = −94.15$, comparison with age-only model $\chi^2(1) = 33.14$, $p < 0.001$;
• non-tense inflection as predictor: $\text{BIC} = 225.64$, $\text{logLik} = −73.16$, comparison with age-only model $\chi^2(1) = 77.10$, $p < 0.001$
• adding tense inflection results in worse model than adding non-tense inflection (cf. the BIC values)

Conclusions
MLU can be predicted more precisely if the prediction takes into account indicators of inflection mastery. The consistent growth and sensitivity to inflection variables indicate that MLU reflects some latent characteristic of language growth.

References