

Child language documentation: The sketch acquisition project

Birgit Hellwig

University of Cologne

bhellwig@uni-koeln.de

Mittwoch,
04.03.2020
14:45–15:15
VMP5 2101/2105

Language documentation is concerned with “the observable linguistic behavior, manifest in everyday interaction between members of the speech community” (Himmelmann 1998: 166) – with a clear focus on the linguistic behavior of adults. The last two decades have seen the construction of numerous adult corpora, enabling and stimulating cross-corpus typological research. But the “observable linguistic behavior” also includes the language used by and with children, and there is a growing interest within language documentation in constructing corpora of child language and child-directed language and utilizing them in cross-linguistic research (e.g., Kelly et al. 2015; Moran et al. 2016). This interest is set against the background of a severe bias within language acquisition research: much more so than in typological research, theories of linguistic and cognitive development are grounded in empirical data that is strongly biased towards the major European languages, with acquisition data being available for only 1–2% of the world’s languages (Lieven & Stoll 2010: 144).

Cross-linguistic research on language acquisition made major progress in the 1980s/ 1990s, largely thanks to a number of large-scale initiatives such as Dan Slobin’s (1985–1997) “cross-linguistic study of language acquisition.” But despite this early promise, 20 years on we are not markedly further. Very likely, the reasons are to be found in the forbidding methodological challenges of constructing child language corpora under fieldwork conditions with their strict requirements on participant selection, sampling intervals and amounts of data (e.g., Behrens 2008; Parisse 2019; Tomasello & Stahl 2004).

Building up on Dan Slobin’s original idea, this talk reports on an ongoing joint project (Defina et al., in prep.) that combines the construction of manageable corpora of natural interaction with children (of 5 hours of analyzed data) with a sketch description of the corpus data – resulting in a database of comparable corpora and sketches across languages, which can form a basis for cross-linguistic comparisons.

References: Behrens, H. (ed.). 2008. Corpora in language acquisition research. Amsterdam: Benjamins. Defina, R., B. Hellwig et al. (in prep.). The sketch acquisition project. Himmelmann, N. P. (1998). Documentary and descriptive linguistics. *Linguistics* 36(1), 161–195. Kelly, B., W. Forshaw, R. Nordlinger & G. Wiggelsworth (2015). Linguistic diversity in first language acquisition research: Moving beyond the challenges. *First Language* 35(4–5), 286–304. Lieven, E. & S. Stoll (2010). Language. In M. H. Bornstein (ed.), *The Handbook of Cross-Cultural Developmental Science*. Mahwah, NJ: Erlbaum, 143–165. Moran, S., R. Schikowski, D. Pajović, C. Hysi & S. Stoll (2016). The ACQDIV Database: Min(d)ing the Ambient Language. In N. Calzolari et al. (eds.), *Proceedings of the Tenth International Conference on Language Resources and Evaluation (LREC 2016)*. Paris: European Language Resources Association (ELRA), 4423–4429. Parisse, C. (2019). Representativeness in (target-specific) corpus construction. *CogniTextes* 19. Slobin, D. I. (1985–1997). The crosslinguistic study of language acquisition, Vol. 1–5. Mahwah, NJ: Erlbaum. Tomasello, N. & D. Stahl (2004). Sampling children’s spontaneous speech: How much is enough? *Journal of Child Language* 31(1), 101–121.

AG 10