

Ongoing change and where to track it: Studying incipient grammaticalisation on the example of the German *am*-progressive

Arne Zeschel, Thilo Weber & Felix Bildhauer
Leibniz-Institut für Deutsche Sprache, Mannheim
{zeschel, weber, bildhauer}@ids-mannheim.de

Keywords: grammaticalisation, aspect, subjectification, spoken language, web corpora

Linguistic innovation is commonly assumed to originate in informal spoken language with low normative pressure on the speaker. Corpora of authentic private conversations are therefore of great interest to studies that seek to identify current shifts and ongoing changes within the usage patterns of grammatical constructions. Unfortunately, such corpora are costly to compile and therefore usually rather small, so that it is often not possible to draw robust inferences from the observations they afford. However, research in Interactional Linguistics (IL, cf. Couper-Kuhlen & Selting 2018) has proposed that certain kinds of computer-mediated communication (CMC) are highly similar to spoken conversational data in relevant respects, so that they are profitably analysed with IL methods and provide a rich source of comparably 'cheap' new data for the field (Dürscheid & Brommer 2009; Imo 2013). The present paper explores whether a similar argument can be made for studies of language change and grammaticalisation. In order to assess this question, we consider a fairly well-studied construction that is widely assumed to be grammaticalising in present-day German, the so-called *am*-progressive (cf. e.g. Flick & Kuhmichel 2013, Anthonissen et al. 2016) as illustrated in (1):

- (1) bin da am Umbauen und gleichzeitig bin ich auch noch eine meine Wohnung am
am there at:the rebuild:INF and simultaneously am I also also a my apartment at:the
Renovieren und dahin am Ziehen
renovate:INF and there at:the move:INF
'I am rebuilding there and simultaneously I am also renovating my apartment and moving there.'

FOLK_E_00182_SE_01_T_01_DF_01, c442

Previous studies have identified a number of properties that we take as indications of the degree of grammaticalisation that the construction has attained so far: (i) the proportion of usages with non-(simple-) intransitive argument structures (as an indication of its spread to novel syntactic contexts; cf. Flick & Kuhmichel 2013), (ii) the proportion of usages with non-activity verbs (spread to novel semantic contexts; cf. Flick & Kuhmichel 2013), and (iii) the proportion of usages with additional subjective or intersubjective discourse functions (spread to novel pragmatic contexts as a result of subjectification; cf. Anthonissen et al. 2016). We compare these properties across uses of the *am*-progressive in the reference corpus of written German, DeReKo (Kupietz et al., 2010), the reference corpus of spoken German, FOLK (Schmidt, 2016) and in discussion forums from the DECOW16B web corpus (Schäfer & Bildhauer, 2012). The forum data are both informal and interactive and, owing to corpus size, they contain substantially more attestations of the construction (approximately 70,000) than FOLK (less than 150 attestations in the current release 2.19). For each parameter, we assess whether the informal written usage in DECOW is more similar to the spoken usage of the construction in FOLK than to its more formal written usage in DeReKo. We predict that usage in DECOW is on the whole more grammaticalised than in DeReKo, but not necessarily less grammaticalised than in FOLK. We close with a discussion of whether (certain kinds of) informal written data can be used as a proxy for informal spoken data in studies of synchronic variation and incipient grammaticalisation, or whether the former should receive greater attention in their own right.

References

- Anthonissen, Lynn, Astrid De Wit & Tanja Mortelmans. 2016. Aspect Meets Modality: A Semantic Analysis of the German *am*-Progressive. *Journal of Germanic Linguistics* 28(1). 1–30.
- Couper-Kuhlen, Elizabeth & Margret Selting. 2018. *Interactional linguistics: Studying language in social interaction*. Cambridge: Cambridge University Press. doi:10.1007/9781139507318.
- Dürscheid, Christa & Sarah Brommer. 2009. Getippte Dialoge in neuen Medien. Sprachkritische Aspekte und linguistische Analysen. *Linguistik Online* 37(1). doi:10.13092/lo.37.511.
- Flick, Johanna & Katrin Kuhmichel. 2013. Der *am*-Progressiv in Dialekt und Standardsprache. *Jahrbuch für germanistische Sprachgeschichte* 4(1). 52–76. doi:10.1515/jbgsg-2013-0005.

- Imo, Wolfgang. 2013. *Sprache in Interaktion: Analysemethoden und Untersuchungsfelder*. De Gruyter. doi:10.1515/9783110306323.
- Kupietz, Marc, Cyril Belica, Holger Keibel & Andreas Witt. 2010. The German reference corpus DeReKo: A primordial sample for linguistic research. In Nicoletta Calzolari, Khalid Choukri, Bente Maegaard, Joseph Mariani, Jan Odijk, Stelios Piperidis, Mike Rosner & Daniel Tapias (eds.), *Proceedings of the seventh international conference on language resources and evaluation (LREC '10)*, 1848–1854. Valletta, Malta: European Language Resources Association (ELRA).
- Schäfer, Roland & Felix Bildhauer. 2012. Building large corpora from the web using a new efficient tool chain. In Nicoletta Calzolari, Khalid Choukri, Thierry Declerck, Mehmet Ugur Dogan, Bente Maegaard, Joseph Mariani, Jan Odijk & Stelios Piperidis (eds.), *Proceedings of the Eighth International Conference on Language Resources and Evaluation (LREC'12)*, 486–493. Istanbul: ELRA.
- Schmidt, Thomas. 2016. Good practices in the compilation of folk, the research and teaching corpus of spoken German. *International Journal of Corpus Linguistics* 21(3). 396–418.