Morphological Productivity and Linguistic Creativity:

Measuring Linguistic Innovation and Change in Diachronic Corpora

The concept of productivity has been widely discussed in the morphological literature throughout the past few decades in both synchronic and diachronic perspectives (cf. e.g. Plag, 1999; Bauer, 2001; Scherer, 2005). Although considerable progress has been made since Rainer's (1987: 187) verdict that the notion of productivity was often used, but seldom explicitly discussed, some key questions are still subject to debate. Concerning quantitative approaches to productivity, for example, the question what exactly the number of hapax legomena, i.e. words occurring only once in a given corpus, of a certain morphological category reveals about the morphological productivity of the respective morphological process has not been conclusively answered yet. The significance of hapax-based measures of productivity such as Baayen's (e.g. 2009) 'potential productivity' and 'expanding productivity' is hotly debated (cf. e.g. Bauer, 2001: 155 for a critique of such approaches and Hilpert, 2013: 132f. for a defence). In addition, there is no consensus to what extent corpus-based measures of productivity can be seen as clues to the cognitive processes underlying the capacity to productively coin new words as well as to the linguistic knowledge constraining the productivity of existing patterns. Yet another, more methodologically oriented open question concerns the size and structure of the corpora to be used. While there are fairly large and well-balanced corpora of English as well as highly sophisticated methods that have been proposed for their diachronic analysis (cf. e.g. Hilpert & Gries, 2009), we have to rely on rather small corpora for many other languages, which provides a further challenge for assessing diachronic changes in productivity.

This paper discusses such open questions with the example of a corpus-based study of German nominalization patterns. Drawing on data from an as yet unpublished corpus of Early New High German (1350-1650) as well as from the GerManC corpus (cf. Durrell et al., 2007), which covers the time span of 1650-1800, the diachronic development of two highly productive word-formation patterns, namely nominalization in the suffix *-ung* and Infinitival Nominalization, is discussed. It is shown that *ung*-nominalization is subject to a growing amount of constraints: For example, words like *Murmelung* 'muttering' and *Schweigung* 'silence', which are felicitous in the Early New High German period, are ungrammatical in Present Day German. The quantitative corpus data provide a valuable means to track the emergence of these constraints. However, some problems arise due to the rather small size of the Mainz Early New High German corpus (c. 280,000 words) and the coarse-grained

periodization of the GerManC corpus, which is only balanced for 50-year periods. I will discuss how such problems can be addressed and argue that even from these less-than-ideal data, important conclusions concerning the cognitive representation of morphological constructions can be drawn.

References

Baayen, R. Harald (2009). Corpus Linguistics in Morphology. Morphological Productivity. In A. Lüdeling & M. Kytö (Eds.), *Corpus Linguistics. An International Handbook. Vol. 2* (pp. 899-919). Berlin, New York: De Gruyter (HSK, 29.2).

Bauer, L. (2001). *Morphological Productivity*. Cambridge: Cambridge University Press.

Durrell, M., Ensslin, A. & Bennett, P. (2007). The GerManC Project. In: *Sprache und Datenverarbeitung,* 31, 71-80.

Hilpert, M. (2013). *Constructional Change in English. Developments in Allomorphy, Word Formation, and Syntax*. Cambridge: Cambridge University Press.

Hilpert, M. & Gries, S. Th (2009). Assessing Frequency Changes in Multistage Diachronic Corpora. Applications for Historical Corpus Linguistics and the Study of Language Acquisition. In: *Literary and Linguistic Computing,* 24, 385-401.

Plag, I. (1999). *Morphological Productivity. Structural Constraints in English Derivation*. Berlin, New York: De Gruyter (Topics in English Linguistics, 28).

Rainer, F. (1987). Produktivitätsbegriffe in der Wortbildungstheorie. In H. Geckeler, H.-M. Gauger & W. Dietrich (Eds.), *Grammatik und Wortbildung romanischer Sprachen. Beiträge zum Deutschen Romanistentag in Siegen, 30.9-3.10.1985* (pp. 187-202). Tübingen: Narr.

Scherer, C. (2005). *Wortbildungswandel und Produktivität. Eine empirische Studie zur nominalen er-Derivation im Deutschen*. Tübingen: Max Niemeyer (Linguistische Arbeiten, 497).