Morpho-syntactic perspectives on the development of the German nominalized infinitive in the Austrian Media Corpus and the Austrian Baroque Corpus.

Martina Werner, Hannes Pirker and Claudia Resch (Austrian Center for Digital Humanities, Austrian Academy of Sciences)

The nominalized infinitive (*NI*, also *conversion, gerund* such as Italian *(il) mangiare carote* '(the) carrot eating') is a very frequent phenomenon in many languages cf., Alexiadou 2011 et al. on variation between Germanic vs. Romance NIs, theoretically cf. Spencer 2013: 340ff). This holds also for present-day-German (PDG). Aside from its usage as an abstract noun (such as PDG *das Laufen* 'the running', *das Herumstehen* 'the standing about'), the NI also appears in light-verb constructions (such as *ins Rollen kommen, zum Verschwinden bringen*), in the progressive (such as *Er ist am Arbeiten*, cf. Pottelberge 2004), and potentially (cf. e.g. Engelberg 2004) in the so-called 'absentive' (such as *Er ist Arbeiten/arbeiten*). In German, NIs exhibit particular morphosyntactic properties, especially as regards the avoidance of plurals (on synchrony cf. e.g. Ehrich 1991; on the diachrony cf. Behaghel 1906/07, Kloocke 1974). Historically, the diachronically increasing productivity (in the sense of serialness) of the NI is simultaneous to the decrease of the *ung*-nominalization (cf. PDG **Lachung, *Hustung;* cf. e.g. ten Cate 1985: 227, Vogel 1996: 250ff, Demske 1999, Pottelberge 2004: 98f, Werner 2010).

Apart from a questionnaire-study on PDG (see Blume 2004), more complex forms of the NI such as the nominalization of verbal complexes and VPs including its formal restrictions (such as *?das Besucht-worden-sein-können, ?das Weggegessen-Haben-Sollen*) and the linkages of the NI to the nominalization of other verbal constructions (such as predicatives as in *Lisa ist schön à ihr Schön-Sein*) have never been corpuslinguistically investigated, neither synchronically, nor diachronically. This is crucial since some formal particularities of the NI such as the selection of auxiliaries are triggered by verbal semantics (cf. e.g. Leiss 1992, cf. *?das Gegessen-Haben, ?das Besuchtwordensein*), i.e. they are not directly attributable to (a transposition of) the verbal infinitive (*gegessen zu haben, besucht worden zu sein*). Starting from this observation, it is unclear when verbal complexes have been nominalized at first in diachrony and how this refers to possible syntactic (cf. Haider 2011) or semantic restrictions (Bücking 2012). Since NIs are a rare and very restricted phenomenon in MHG (cf. Koning 1933), we raise the question of the morphosyntactic developments of the NIs from ENHG onwards.

For this purpose, we present a comparison between two corpus-based studies on the NI; one on Austrian Standard German, with results on PDG from the *Austrian Media corpus* (*AMC*, see https://acdh.oeaw.ac.at/acdh/en/amc), with examples such as *(das) Drehen* '(the) turning' or *(das) Gfrast-Sein* '(the) being a rascal'; the second study is on ENHG with results from the *Austrian Baroque Corpus* (*ABaC:us*, see https://acdh.oeaw.ac.at/abacus/).

AMC is a digital collection of journalistic prose, comprising 35 million articles from newspapers and magazine, providing almost complete coverage of Austria's print-media production of the last two decades. The whole corpus (ca. 9 billion token) has been automatically provided with part-of-speech (PoS) tags and morphological information. *ABaC:us is* a digital thematic collection of Baroque texts of Austrian German with textual data in full length of early and, if possible, the first known editions and rare specimens from different monastic and public libraries. A large part of the corpus, ca. 180,000 tokens, has been annotated with the STTS-TagSet (cf. www.sfs.uni-tuebingen.de/resources/stts-1999.pdf) which was adopted for ENHG, including manual control for a gold-standard corpus. Both corpora are located at the Austrian Center for Digital Humanities (ÖAW) and are accessible via SketchEngine (cf. Kilgarriff et al. 2014), a query tool for morphosyntactic information.

Our empirical findings of both studies result in descriptive generalizations with respect to the historical development of the NIs and its semantics from ENHG to PDG, including a cross-linguistic perspective. The conclusions from category-specific, morphosyntactic distributions with still attestable restrictions in PDG suggest an ongoing language change in the nominal domain at the interface between syntax and morphology.

Selected references:

Alexiadou, A., G. Iordăchioaia & F. Schäfer (2011): Scaling the variation in Romance and Germanic nominalizations. In Antonia Petronella Sleeman & Harry Perridon (eds.), The Noun Phrase in Romance and Germanic: Structure, Variation, and Change, 25–40. Amsterdam: Benjamins.

Behaghel, O. (1906/07): Zum substantivierten Infinitiv. Zeitschrift für Wortforschung 8, 329–332.

Blume, K. (2004): Nominalisierte Infinitive: Eine empirisch basierte Studie zum Deutschen (Linguistische Arbeiten, Band 487). Tübingen: Niemeyer.

Bücking, S. (2012): Müdigkeit und Müde-Sein: Zur Semantik adjektivbasierter Zustandsnominalisierungen im Deutschen. Linguistische Berichte 232, 361–397.

Demske, U. (1999): Nominalisierungen im Deutschen und Englischen: Überlegungen zu einer Theorie des sprachlichen Wandels. In: Kanngießer, S. / Vogel, P. (eds.): Elemente des Sprachwandels. Opladen, 98–138.

Ehrich, V. 1991. Nominalisierungen. In Stechow, A. von / Wunderlich, D. (eds). Semantik/Semantics. Ein internationales Handbuch der zeitgenössischen Forschung. Berlin/ New York, 441–458.

Engelberg, S. (2004): The Structural Ambiguity of PPs Containing Nominalized Infinitives. In: Moskovsky, C. (ed.): Proceedings of the 2003 Conference of the Australian Linguistics Society. Permanent URL: http://www.als.asn.au/.

Haider, H. (2011): Infinitivnominalisierungen als Testgelände. PPT-Presentation of a lecture held on 07/14/11, Bochum University. Permanent URL: *https://www.uni-salzburg.at/fileadmin/oracle_file_imports/1537188.PDF*, 2016/06/10.

Kilgarriff, A. et. al. (2014): The Sketch Engine: ten years on. Lexicography 1, 7–35.

Koning, W. (1933): Substantivierte infinite Verbformen mit verbaler Rektion. Die neueren Sprachen 41, 90–98.

Kloocke, H. (1974): Der Gebrauch des substantivierten Infinitivs im Mittelhochdeutschen. Göppingen.

Leiss, E. (1992): Die Verbalkategorien des Deutschen: Ein Beitrag zur Theorie der Sprachlichen Kategorisierung. Berlin, New York: de Gruyter.

Pottelberge, J. v. (2004): Der am-Progressiv. Struktur und parallele Entwicklung in den kontinentalwestgermanischen Sprachen. Tübingen: Narr.

Spencer, A. (2013): Lexical relatedness. Oxford: OUP.

Ten Cate, A. (1985): Aspektualität und Nominalisierung. Zur Bedeutung satzsemantischer Beziehungen für die Beschreibung der Nominalisierung im Deutschen und im Niederländischen. Frankfurt: Lang.

Vogel, P. (1996): Wortart und Wortartwechsel. Zur Konversion und verwandten Erscheinungen im Deutschen und in anderen Sprachen. Berlin, New York: de Gruyter.

Werner, M. (2010): Substantivierter Infinitiv statt Derivation. Ein ‚echter' Genuswechsel und ein Wechsel der Kodierungstechnik innerhalb der deutschen Verbalabstraktbildung. In: Bittner, Dagmar / Gaeta, Livio (eds.): Kodierungstechniken im Wandel. Das Zusammenspiel von Analytik und Synthese im Gegenwartsdeutschen. Berlin, New York: de Gruyter, 159–178.