etymological counterparts in the ancient IE languages. Interestingly, most of the properties not inherited from Proto-IE in Baltic and Slavic – as far as I can judge from the data available – are not inherited in Finnic either. This means that these properties were created relatively recently, and that the major part of these properties is the result of common developments in Baltic, Finnic and East Slavic.

References

Productivity of selected dejectival suffixes in Czech.*

Sevcikova, Magda
(Charles University, Prague)

Productivity has become one of the central issues of research into word-formation in recent decades; it has mostly been studied on the basis of corpus data. In the respected literature on word-formation in Czech, however, the productivity of word-formation means has been indicated only generally, based usually on existing dictionaries and using scales like “very productive / less productive / unproductive” (cf. Dokulil 1962, Daneš et al. 1967, Šmilauer 1971, Hauser 1980, Dokulil et al. 1986, Karlík et al. 2000).

In our paper, we present a pilot corpus-based study of productivity of selected suffixes used for dejectival derivation of nouns in Czech. After an introductory comparison of how productivity is defined across approaches (cf. Baayen 1992 and 2001, Bauer 2001, Štekauer – Lieber 2005), and a brief survey of selected approaches to measuring productivity in languages other than English (e.g. Bolozky 1999, Evert – Lüdeling 2001, Gaeta – Ricca 2006, Pustylnikov – Schneider-Wiejowski 2009), we focus on the suffixes that compete in deriving names of qualities from adjectives in Czech: the suffixes -ost/-nost, -ota and -da are used to derive feminine names (e.g. lehkost ‘lightness’, kritičnost ‘criticality’, samota ‘loneliness’, pravda ‘truth’), -ství and -i occur in neuter derivatives (e.g. slaboství ‘weakness’, pokrytí ‘hypocrisy’, zdraví ‘health’). The analysis is based on recent language data stored in representative sub-corpora of the Czech National Corpus (SYN2000, SYMN2005 and SYN2010). The productivity of the suffixes is determined according to their token frequency, type frequency and to the count of their hapax legomena in the corpora (Baayen 1992). The results obtained will be compared with results from several corpora of journalistic texts (Prague Dependency Treebank 2.0, SYN2006PUB, SYN2009PUB) on the one hand, and with the frequency data from the recent dictionary of Czech (Filipec et al. 1998) on the other.

The analysis of productivity is completed with an analysis of the lexical meaning of the derivates, which enables us to detect the differences among competing suffixes, to describe the polysemy of the derivates (and the suffixes), to speculate on reasons why some adjectives are compatible with more than one of the suffixes while with other adjectives only one suffix is used, and to ask further related questions. For instance, the suffix -ost is expected to be the most frequent (and most productive) one among suffixes used for derivation of names

338
of qualities (type frequency), nevertheless, if there are several nouns derived from the same base adjective, the noun with the suffix -ost is often less frequent than the others (token frequency; e.g. mladí vs. mladosť ‘youth’ with 3853 vs. 101 tokens in the SYN2010 corpus, veselí vs. veselost ‘merriment’ with 987 vs. 185 tokens in the same corpus).

The presented analysis might be a considerable, real data-based contribution to the theoretical discussion on which component of the linguistic description productive derives belong to in Czech, i.e. whether they should be considered a part of the lexicon or of the grammatical component. Our paper concludes with a discussion of dis/advantages of both possibilities.

* Acknowledgement: This work has been supported by the project GA ČR P406/12/P175.

References

Examining the role of structural features upon that/zero variation in verbs of locution: a diachronic corpus-based multivariate analysis of say, tell, assert, declare, and state.

Shank, Christopher and Plevoets, Koen
(Bangor University and University of Ghent)

Most of the attention following Rissanen (1991) and Finnegan and Biber’s (1995) seminal research on the rise and predominance of the zero-complementizer form as an object-clause link in PDE has focused on the mental state verbs (i.e. think and know) while considerably less has been directed at the equivalent claims and conclusions made regarding the remaining 50% of their data set namely say and tell aka verbs of locution