# NetWordS short visit

Final report

by

Stela Manova

*Austrian Academy of Sciences & University of Vienna*

stela.manova@univie.ac.at

http://homepage.univie.ac.at/stela.manova/

I used my NetWodS grant to spend two weeks, 6 through 19 May 2013, as a visiting scholar at the Laboratorio di Linguistica 'Giovanni Nencioni' of Scuola Normale Superiore in Pisa, Italy. Website of the host institution: http://linguistica.sns.it/.

## 1. Purpose of the visit

The major goal of my visit was a comparative, corpus-based study of the Italian and Russian suffix combinations in derivation. The work was done in collaboration with Luigi Talamo who had carried out research on parsability and selectional restrictions of Italian derivational suffixes (Talamo 2011). The idea was to combine his research with my own research on affix ordering (Manova 2010, 2011, 2012; Manova & Aronoff 2010) and also to use the new Corpus e Lessico di Frequenza dell'Italiano Scritto (CoLFIS) and the Russian National Corpus (RNC) for the purposes of our research. CoLFIS was prepared at the Linguistic laboratory of SNS (Talamo & Celata 2011). The RNC is the largest electronic resource on a Slavic language. In addition, CoLFIS is annotated for derivational morphology while such an annotation is currently under development for RNC but there is a Morpheme Dictionary of Russian (Kuznecova and Efremova 1986) that can be used for research on affix combinations in derivation and that serves as basis for the annotation of the corpus.

Information about CoLFIS is available online at: http://linguistica.sns.it/CoLFIS/Home.htm. The Russian National Corpus can be accessed at: http://www.ruscorpora.ru/.

As Prof. Pirrelli's CNR DyLan Lab is also located in Pisa, the second goal of my visit was to see how their research and findings relate to my own results and those of the research on affix ordering in Italian and Russian derivational morphology I did in collaboration with Luigi Talamo.

**2. Description of the work carried out during the visit**

During the first week of my visit, the colleague Talamo and I checked my usage-based approach to Bulgarian and English suffix combinations (Manova 2011) against word-formation data from Italian and Russian. My approach is based on two major assumptions:

1) Suffix combinations should be tackled in terms of binary structures of the type SUFF1-SUFF2, SUFF1 and SUFF2 being any two neighboring suffixes; and

2) All SUFF2 suffixes that can follow a particular SUFF1 in a language should not be analyzed together, as done in the literature so far, but, according to their lexical-category specifications, they should be distributed into three groups: $SUFF2_N$, $SUFF2_{ADJ}$, $SUFF2_V$.

Significantly, if SUFF2 suffixes are distributed according to their lexical-category specifications, the majority of the SUFF1 suffixes combine with a single $SUFF2_N$, $SUFF2_{ADJ}$ and $SUFF2_V$, i.e. participate in up to three combinations. In cases where SUFF1 is followed by more than one SUFF2 of the same lexical category, the principle of default (the majority of types are derived by a particular SUFF2) and the principle of blocking (the different SUFF2 suffixes derive different semantics) differentiate among the available options (Manova 2011).

Such a treatment of suffix combinations is consonant with research in cognitive neuroscience revealing that different word classes have their representations in different areas in the brain (Mestres-Missé et al. 2010); as well as with research on first language acquisition - the acquisition of nouns differs from the acquisition of verbs (cf. research in Professor Dressler's lab at the University of Vienna).

For Italian, we looked for suffix combinations first in the four-million ColFIS and then rechecked the results obtained in the larger La Repubblica corpus that contains 380

million tokens. La Repubblica is available at: http://dev.sslmit.unibo.it/corpora/corpus.php?path=&name=Repubblica. The results of both searches always converged, i.e. CoLFIS turned out to be a very well-balanced and reliable resource for investigation of the Italian word-formation.

As already mentioned, we checked the combinability of the Russian derivational suffixes with the help of the RNC and the morpheme dictionary of Russian (Kuznecova & Efremova 1986).

In addition to the investigation of the role of the lexical-category specification of a suffix in suffix ordering, we also did some research on the role of suffix-particular semantics. Following research in neural science (Kandel et al. 2012), we expected SUFF1 semantics to impact further suffixation, i.e. SUFF1 suffixes deriving the same basic cognitive category to select the same or similar SUFF2 suffixes. We compared the combinability of nine suffixes deriving persons in Russian with the set of all nine Italian suffixes for derivation of personal nouns and found support for the correctness of our hypothesis for both Italian and Russian.

During the second week of my visit, L. Talamo and I started working on the preparation of an article about our collaborative research on affix ordering in Italian and Russian and the results obtained. I also met Prof. Pirelli, discussed my research with him and learned about the research at his lab. Additionally, during this week I presented my research and the collaborative work done so far with a talk at the seminar of the Linguistic Laboratory of the Scuola Normale Superiore. Moreover, I used the visit to make progress and continue another collaboration with the SNS Linguistic Laboratory - a project that investigates the psycholinguistic status of the basic form of an inflectional paradigm. In this investigation, we analyze data from Italian and Bulgarian verbal morphology and run experiments at the SNS and New Bulgarian University in Sofia. This is work in progress and closely related to the goals of NetWordS.

## 3. Description of the main results obtained

The research I carried out during my stay at Scuola Normale Superiore reveals important aspects of the organization of the lexicon. In most SUFF1-SUFF2 combinations, a

SUFF1 combines with only one nominalizing, verbalizing and adjectivizing SUFF2, the majority of the combinations being word-class changing, i.e. SUFF1 and SUFF2 have different lexical-category specifications. For Italian, we differentiate between two types of suffix combinations:

1) *fixed* SUFF1-SUFF2 combinations: SUFF1 combines with a single SUFF2 of a particular lexical category.

2) *predictable* SUFF1-SUFF2 combinations: SUFF1 combines by more than one SUFF2 of a particular lexical category but one of the SUFF2 suffixes applies by default (the majority of the types are derived by that SUFF2) or intentional semantics (SUFF2 particular semantics), i.e. what the speaker intends to say, differentiates between the possible options. Due to the well-known blocking effect in morphology, if more than one SUFF2 is available, those SUFF2 suffixes never have exactly the same semantics.

We also established that, in Italian, in cases of predictable suffix combinations in which the SUFF2 applies by default, the other available SUFF2 suffixes always derive up to ten types. We checked the correctness of this observation in La Repubblica corpus, too; and it holds even for such a large corpus.

Similar results were obtained for Russian, though in a very few cases the Russian suffix combinations challenge the principle of morphological blocking.

In sum, all investigated SUFF1-SUFF2 combinations in Italian and most such combinations in Russian are either fixed or predictable. The number of the fixed combinations is greater than the number of the predictable ones. If lexical-category specification and suffix-particular semantics are considered, most SUFF1-SUFF2 combinations appear fixed and predictable, i.e. they are most probably processed and produced as rote-learned ready-made structures by speakers. Thus, SUFF1-SUFF2 combinations appear good examples of constructions in morphology. Since lexical-category specification and suffix-particular semantics explain the way suffix combinations are formed, our results suggest that, in the lexicon, suffixes should be specified for lexical category and suffix-particular semantics in terms of cognitive categories.

Finally, if a SUFF1 combines with a single SUFF2$_N$, SUFF2$_{ADJ}$ and SUFF2$_V$, i.e. participates in up to three combinations, that SUFF1 can be identified in an electronic corpus statistically - on the basis of its further combinability. Thus, the results of our research can be used for development of strategies for automatic corpus annotation at a morpheme level as well as for creating tools for search for derivational suffixes in non-annotated corpora.

### 4.  Future collaboration with host institution (if applicable)

My near-future collaboration with the Linguistic Laboratory of SNS is on inflectional morphology. As mentioned in section 2 above, we investigate the psycholinguistic status of particular forms of the Italian and Bulgarian verb paradigms and the next series of experiments is planned for October 2013.

### 5.  Projected publications/articles resulting or to result from your grant

An article reporting the results of the collaborative research on Italian and Russian word-formation is currently in preparation. The research done during my stay in Italy will also be presented at the third annual workshop of NetWordS, to be held in Dubrovnik (Croatia) on September 19-20, 2013.

### 6.  Other comments

I would like to thank for the grant. The short visit gave me the chance to carry out research at a renowned foreign institution and acquaint myself with the scientific practices there. I could attend the seminar series of the Linguistic Laboratory and listened to a talk by Prof. Geert Booij who presented on construction morphology, a topic highly relevant to what Luigi Talamo and I investigated during my short visit. I myself presented at the same seminar series the next week and received input on my own research and the collaborative research L. Talamo and I carried out on the combinability of the Italian and Russian derivational suffixes. I had an inspiring time, could compare

the scientific practices in Italy and Austria, worked with CoLFIS, the newest corpus of Italian, and learned a great deal. Moreover, I will continue collaborating with the Italian partner in the future.

**References**

Kandel, E. R.; J. H. Schwartz; T. M. Jessell; S. A. Siegelbaum; & A. J. Hudspeth (2012). *Principles of Neural Science*, 5th Edition. McGraw-Hill Professional.

Kuznecova, A. I. and T. F. Efremova. 1986. *Slovar' morfem russkogo jazyk / A Morpheme Dictionary of Russian*. Moskva: Russkij jazyk.

Manova, S. (2010). Suffix Combinations in Bulgarian: Parsability and Hierarchy-Based Ordering. *Morphology* 20(1), 267-296.

Manova, S. & M. Aronoff (2010). Modeling Affix Order. *Morphology* 20(1): 109-131.

Manova, S. (2011). A cognitive approach to SUFF1-SUFF2 combinations: A tribute to Carl Friedrich Gauss. *Word Structure* 4(2): 272–300.

Manova, S. (2012). Closing suffixes. To appear in Müller, P.; I. Ohnheiser; S. Olsen & F. Rainer (eds.). *Word Formation in the European Languages*. Handbooks of Linguistics and Communication Science. Berlin: de Gruyter, 15 pp.

Mestres-Missé, A.; A. Rodriguez-Fornells & T. F. Münte (2010) Neural differences in the mapping of verb and noun concepts onto novel words. *NeuroImage 49(3): 2826.*

Talamo, L. (2011). Suffix combinations in Italian: selectional restrictions and processing constraints. To appear in Manova, S. *Affix ordering accross languages and frameworks*. Oxford University Press.

Talamo, L. & C. Celata (2011). Toward a morphological analysis of the Italian lexicon: developing tools for a corpus-based approach. *Quaderni del Laboratorio di Linguistica* 1(10).