DR. MONICA-MIHAELA RIZEA
INDEPENDENT RESEARCHER, ROMANIA
EMAIL: MONICAMIHAELARIZEA@GMAIL.COM

STSM REPORT
ANALYSIS OF LEXICAL AND SYNTACTIC PROPERTIES OF ROMANIAN MWEs
COST ACTION IC 1207 PARSEME

HOST INSTITUTION

Institut für England- und Amerikastudien, Goethe-Universität Frankfurt am Main

TIMEFRAME

18 April 2016 – 29 April 2016

PURPOSE OF THE STSM

The purpose of this STSM implied:
- Working on the classification of Romanian Multiword Expressions within the multilingual template set up by the PARSEME Working Group 1;
- Contributing to the collection of Romanian MWEs that are Negative Polarity Items (NPIs), both in terms of extending the current list with new items and of enriching the existing description with corpus examples for every type of context in which NPIs expressions are licensed.

WORK CARRIED OUT

The first part of the STSM implied enriching the classification of the Romanian MWEs (according to the requirements of the multilingual template designed by PARSEME WG 1) in terms of 1. prominent syntactic patterns 2. flexibility of MWE parts (with special emphasis on the syntactic flexibility).

1. A rich classification of the Romanian MWEs according to syntactic patterns was proposed and documented with numerous glossed examples; real usage examples were also provided (especially from the Romanian Web Corpus, and Opus2 parallel corpora, accessible via the the.sketchengine.co.uk online platform).
For example, in the case of verbal MWEs I was particularly interested in structures with clitics as obligatory constituents attested in patterns such as **V NP, V N N_dat, V PP, V NP PP, V ADVP, V AP** (for details, see http://wiki.studiumdigitale.uni-frankfurt.de/FB10_Parseme/index.php/MWE_Template:_Romanian#Verbal_MWEs).
As a general comment, many verbal MWEs represent structures with obligatory ACC/DAT reflexive clitics ('se', '-şi'), with personal ACC/DAT clitics (e.g.: 'o', '(-)i', '-l'), or with both personal and reflexive clitics ('i se'). When they are obligatory constituents of verbal MWEs, personal clitics such as the non-anaphoric feminine clitic 'o' and the non-referential '(-)i' have an expletive function. I will only provide examples for the **V NP** and the **V PP** structures (more examples are available on the Romanian MWE wiki template):

V NP

The idiomatic combinations of the type V NP can contain Nominative, Accusative, and (rarely) Dative nouns; they can also include Accusative or Dative (personal or reflexive) clitics:

- ***a-l duce (pe cineva) capul (la ceva) - by convention, the obligatory ACC personal clitic is represented as a singular masculine form ('-l')***
  'understand / know how to do something/be smart'
  'Nu crezi că mă duce capul ca să iau asta în serios, nu? Nu te prosti. ' (You don' t think I' m smart enough to be serious about. Don' t be ridiculous.)
  (Opus 2 Romanian / English - the.sketchengine.co.uk)

V PP

- ***a i se urca (cuiva ceva) la cap (with a personal Dative clitic and an Accusative reflexive clitic)***
  (of alcohol) 'go to someone's head'
  'Nu mă grăbi! Înainte, vinul mi se urca la cap și mă înveselea. Acum îmi coboară în picioare și mă îmbătrânește.' (Don' t hurry me. Wine used to go to my head and make me gay ... and now it goes to my legs and makes me old.)
  (Opus 2 Romanian / English - the.sketchengine.co.uk)

2. In terms of syntactic flexibility of the MWE parts, the most prominent syntactic operations undergone by the Romanian MWEs seem to be: Passivization, Raising, Topicalization, Dative/PP alternation.

Examples can be found in the Romanian MWE template (accessible at http://wiki.studiumdigitale.uni-frankfurt.de/FB10_Parseme/index.php/MWE_Template:_Romanian).

The second part of the STSM program was dedicated to strategies for enriching the Romanian database of Negative Polarity Items (NPIs) (available via https://www.english-linguistics.de/codii/) with new Multiword Expressions and to completing the existing description with corpora examples. Discussions with Prof. Dr. Manfred Sailer, PD Dr. Frank Richter, and with the initial author of the Romanian NPI collection, Dr. Gianina Iordăchioaia were very helpful, leading to a number of important decisions on modifying and extending the current NPI collection. During this second stage of the STSM, I have contributed a number of modifications on the original XML files, and I have also collected a list of new Negative Polarity Items MWE candidates. However, this is work-in-progress that will be fulfilled during future collaborations.

MAIN RESULTS OBTAINED

- An enriched classification of Romanian MWEs within the multilingual template set up by PARSEME WG 1. It now contains a description of the most prominent syntactic patterns, of the Romanian-specific aspects related to fixedness/flexibility of MWE parts, and a classification of Romanian MWEs by levels of idiomaticity (especially lexical, syntactic, and semantic idiomaticity). The Romanian MWE template contains multiple glossed examples as well as real usage (corpus) examples.
- Important decisions regarding the completion of the collection of Romanian NPIs (in terms of adding more syntactic information and more corpus examples, defining new Romanian-specific licensers, and enriching the current list with new multiword expression items).
- Incipient modifications of the RO XML files (Extensible Markup Language representing the encoding format of the entire CODII collection).

- A list of new Negative Polarity Multiword candidates collected from the most important Romanian dictionaries.
  I also benefited from my stay in Frankfurt beyond what has been planned in my STSM proposal:
- I was invited to give a talk on the *Particularities of Romanian as a Negative Concord Language. The Idiosyncratic Behaviour of Nici and an HPSG Account of Long Distance Negative Concord* and I received precious feedback on my previous work on Romanian negation. Discussions with Prof. Dr. Manfred Sailer and with PD Dr. Frank Richter led to an improved HPSG analysis and to a deeper understanding of the LRS approach.
- I have also received useful information about the strategies of automatic extraction of MWE candidates and on corpus annotation for German; I intend to apply this knowledge in my future work.
- I also had the chance to discuss with Dr. Irene Franco from the Institute for Romance Languages and Literature, and I received important details about the historical development of quantifiers, and about negative indefinites and negative concord in Old Italian (http://www.quantification.eu/). I intend to use this information in my future research on the historical development of negation from a comparative perspective.

FUTURE COLLABORATION WITH THE HOST INSTITUTION

Professor Manfred Sailer and I intend to continue our collaboration both in terms of obtaining a rich and improved database of Romanian Negative Polarity idiomatic expressions and in terms of further theoretical research on the classification of Negative Polarity Items.

FORSEEN PUBLICATIONS TO RESULT FROM THE STSM

Starting from the Romanian MWE template, Professor Manfred Sailer and I intend to develop an overview paper on the phraseology of Romanian.

We also intend to collaborate on a joint paper about strategies of *Enriching the Collection of Romanian Negative Polarity Items (CoDII-NPI.ro) with Multiword Expressions.*

CONFIRMATION BY THE HOST INSTITUTION OF THE SUCCESSFUL EXECUTION OF THE STSM

Prof. Dr. Manfred Sailer writes:

I confirm the successful execution of Monica-Mihaela Rizea's STSM at the Institute for English and American Studies at Goethe-University Frankfurt a.M. in April 2016. We worked together on the Romanian MWE template of WG1 and on the Romanian collection of Negative Polarity Items (NPI), available via https://www.english-linguistics.de/codii/. Monica presented her work in our colloquium and had the chance to meet with several colleagues from Frankfurt. Our work on the MWE templates for Romanian also led to serious enrichments in the English and German templates. We put an emphasis on the syntactic flexibility, which is a dimension that has received less attention in many of the MWE templates.
Dr. Gianina Iordăchioaia (Stuttgart), the original author of the Romanian NPI collection, spent one day with us in Frankfurt and we discussed and decided on extensions and improvements in the encoding of the collection. We also managed to arrive at a number of important decisions concerning the encoding in the MWE template and the NPI collection. All of these decisions could only be made in the direct, personal interaction that was made possible by the STSM grant.

In addition to the work on these two resources of data, we laid the basis for future joint work on idiomatic NPIs in Romanian, where we could find empirical justification for a finer distinction within the classes of NPIs than has previously been assumed.

Manfred Sailer
Insitut für England- und Amerikastudien
Goethe-Universität Frankfurt a.M.
Norbert-Wollheim-Platz 1
D-60623 Frankfurt a.M.
sailer@em.uni-frankfurt.de