SHORT TERM SCIENTIFIC MISSION (STSM)
SCIENTIFIC REPORT

This report is submitted for approval by the STSM applicant to the STSM coordinator

Action number: CA18209
STSM title: Corpus Analysis of Covid and health-related metaphors
STSM start and end date: 15/08/2021 to 30/08/2021
Grantee name: Liudmila Mockiene

PURPOSE OF THE STSM:

(max.200 words)

The aim of the STSM was the extension of multilingual resources, i.e. carrying out research on COVID-19 and health-related metaphors based on a multilingual corpus ParlaMint and creation of a corpus covering the COVID-19 pandemic situation as presented in the Lithuanian news and social media. This STSM falls within the research carried out by a group of researchers Kristina Štrkalj Despot, Ana Ostroški Anić and Petya Osenova, who work on a Use Case of Public Health under Task 4.4 Life Sciences in WG4, NexusLinguarum COST Action. The research is focusing on Parliamentary Data from *ParlaMint project* in *NoSketch concordancer* for Bulgarian, Croatian and Slovenian, thus aiming to analyse varieties of metaphors related to the current pandemic situation. The result from this STSM will contribute to the multilingual set of uniformly annotated metaphor corpora as part of the parliamentary COVID corpora with the analysis of the Lithuanian data, as a less-resourced language, in ParlaMint-LT 2.0 (Lithuanian parliamentary corpus).

Our work will make possible to compare the reactions in the above-mentioned parliaments and thus contribute to observing the societal map in a pandemic.

DESCRIPTION OF WORK CARRIED OUT DURING THE STSMS

(max.500 words)

During the STSM the following work was carried out: discussion of a common research design including issues related to the work with parliamentary data from ParlaMint multilingual corpus via NoSketch Engine (public, https://www.clarin.si/noske/index-en.html), working with NoSketch concordancer for Lithuanian, extraction of data relevant for the research, solving the ongoing technical problems together with the reserachers of the host institution (such as issues with lemmatization, representation of the extracted concordance lines in a readable format, etc.), discussing methodological issues of identifying the types of Covid and health-related metaphors,

classification of source frames of conceptual Covid and health-related metaphors, analysis of the corpus data in Lithuanian and, together with the researchers from the host institution, representation of the analysed data as an initial ontology/hierarchy formalization of the frames obtained together with the related entries.

DESCRIPTION OF THE MAIN RESULTS OBTAINED

The first step while working with ParlaMint-LT 2.0 was to create a subcorpus for analysing COVID related metaphors. It was analysed further to establish the keyness (relevance) of the keywords in comparison with the rest of the ParlaMint-LT 2.0 corpus.
Each keyword related to COVID (*pandemija* (pandemic), *epidemija* (epidemic), *COVID*, *COVID-19*, *virusas* (virus), *koronavirusas* (coronavirus), *korona* (corona), *karantinas* (quarantine)) was further analysed individually to extract the concordance lines which include specific context the keyword is used in.
All in all 1,305 concordance lines were extracted for the qualitative analysis.
Each concordance line was analysed to establish whether the keyword is used metaphorically or not.
In total, 301 cases were marked as metaphorical, which is 23% of all cases (see Table 1).
The respective metaphorical phrase was highlighted.
All metaphoric phrases were 'lemmatised' and translated into English.
Next, the frame of the metaphor was established.
The frames found in the data are PERSONIFICATION (22%), EVENT STRUCTURE > CAUSES ARE FORCES (20%), COMBAT > WAR (17%), REIFICATION (13%), CONTROL (10%), DISASTER > FIRE (4%), DISASTER > CRISIS (3.3%), DISASTER > THREAT (2.4%), DISASTER > WATER (2.3%), EVENT STRUCTURE > ACTION IS MOTION (2%), DISASTER (1.6%), DANGER (0.6%), PARTNERSHIP (0.3%).

Keyword	Freq. / concordance lines	Metaphoric, %
1. pandemija + epidemija	**336 (275 + 61)**	**114 (34%)**
2. koronavirusas + korona	**197 (182 + 15)**	**70 (36%)**
3. COVID + COVID-19	**171 (59 / 112)**	**42 (25%)**
4. virusas	**153**	**57 (37%)**
5. karantinas	**448**	**18 (4%)**
Total	**1305**	**301 (23%)**

Table 1. Freuqency of metaphoric uses of the analysed keywords

It should be noted that the keyword *karantinas* (quarantine), although being among most relevant (having high keyness) was among the least used in metaphoric sense in the data – only 4%. Thus, it is recommended to exclude it from further research in either other domains or other languages. If we exclude the keyword *karantinas* (quarantine) from the summary of the results, the average percentage of the metaphorical use of other keywords would be significantly higher, i. e. 33% (283 cases out of 857).

The analysed and processed data was represented as an initial hierarchy/ontology of the frames obtained together with the related entries and next converted into an interoperable format as LOD ready to be incorporated into other resources. Now entries include two languages – Lithuanian and English, however, data in any other language can be easily added to the hierarchy/ontology. This format enables further multilingual linking of the COVID and health-related metaphors.

Sample entry:

%%%%%%% Lexical Entries

nexusi:LexicalEntry1 a nexuso:LexicalEntry ;
nexuso:lexicalForm [nexuso:writtenRep "pandemijos įveikimas"@lt ; nexuso:writtenRep "overcoming/ beating the pandemic"@en] ;
nexuso:frame nexuso:Combat ;
nexuso:frame nexuso:War .

For the encoding of the hierarchy/ontology of the COVID related metaphorical frames see Appendix 1.

Within the STSM, a lecture was delivered in the host institution to present the preliminary research results on the 25 August, 2021. Title: ***Corpus Analysis of Covid and health-related metaphors in ParlaMint-LT 2.0. Initial results***

FUTURE COLLABORATIONS (if applicable)

Due to the limited time of the STMS in the context of all empirial work that was performed, the initial ontology version was only partially aligned with the Lemon model (https://lemon-model.net/). As a next step we envisage to convert the methaphor ontology to Lemon completely, and also to validate it before starting alighnments with other resources/models.

The STSM also targeted discussions on the possibilities of further joint analysis of tendencies of the use and spread of COVID-19 and health-related metaphors in other languages and domains (news media and social media in a cross-lingual setting), focusing on the Lithuanian data, which will enable both comparison of the research results with the data from ParlaMint-LT 2.0 and the data in other languages.

Appendix 1. Encoding of the hierarchy/ontology of the COVID related metaphorical frames

Pref nexuso = http::/www.nexus-ling.eu/metaphorOnto.

%%%%%%%%%
nexuso:CovidFrame rdf:type owl:Class .

nexuso:Disaster rdf:type owl:Class .

nexuso:Disaster rdf:subClass owl:CovidFrame .

nexuso:Water rdf:subClass nexuso:Disaster .

nexuso:WaterWave rdf:subClass nexuso:Water .

nexuso:Tsunami rdf:subClass nexuso:WaterWave .

nexuso:Fire rdf:subClass nexuso:Disaster .

nexuso:Tragedy rdf:subClass nexuso:Disaster .

nexuso:Fire rdf:subClass nexuso:Disaster .

nexuso:Crisis rdf:subClass nexuso:Disaster .

nexuso:Plague rdf:subClass nexuso:Disaster .

nexuso:Explosion rdf:subClass nexuso:Disaster .

nexuso:Cataclysm rdf:subClass nexuso:Disaster .

nexuso:Chernobyl rdf:subClass nexuso:Disaster .

nexuso:Reification rdf:type owl:Class .

nexuso:Reification rdf:subClass owl:CovidFrame .

%%%%%%%%%
nexuso:Personification rdf:type owl:Class .

nexuso:Personification rdf:subClass owl:CovidFrame .

nexuso:Killer rdf:subClass nexuso:Personification .

nexuso:Ruler rdf:subClass nexuso:Personification .

nexuso:Teacher rdf:subClass nexuso:Personification .

nexuso:Partner rdf:subClass nexuso:Personification .

%%%%%%%%%
nexuso:ImageSchemas rdf:type owl:Class .

nexuso:ImageSchemas rdf:subClass owl:CovidFrame .

nexuso:Posession rdf:subClass nexuso:ImageSchemas .

nexuso:Container rdf:subClass nexuso:ImageSchemas .

nexuso:DestructiveForce rdf:subClass nexuso:ImageSchemas .

nexuso:Control rdf:subClass nexuso:ImageSchemas .

%%%%%%%%%
nexuso:EventStructure rdf:type owl:Class .

nexuso:EventStructure rdf:subClass owl:CovidFrame .

nexuso:States rdf:subClass nexuso:EventStructure .

nexuso:AreLocations rdf:subClass nexuso:EventStructure .

nexuso:CausesAreForces rdf:subClass nexuso:EventStructure .

nexuso:ActionIsMotion rdf:subClass nexuso:EventStructure .

nexuso:Purposes rdf:subClass nexuso:EventStructure .

nexuso:AreDestinations rdf:subClass nexuso:EventStructure .

%%%%%%%%%
nexuso:Danger rdf:type owl:Class .

nexuso:Danger rdf:subClass owl:CovidFrame .

nexuso:Evil rdf:subClass nexuso:Danger .
nexuso:Beast rdf:subClass nexuso:Danger .

%%%%%%%%%
nexuso:LiteraryForm rdf:type owl:Class .

nexuso:LiteraryForm rdf:subClass owl:CovidFrame .

nexuso:Story rdf:subClass nexuso:LiteraryForm .

nexuso:Drama rdf:subClass nexuso:LiteraryForm .

%%%%%%%%%
nexuso:Sports rdf:type owl:Class .

nexuso:Sports rdf:subClass owl:CovidFrame .

nexuso:Football rdf:subClass nexuso:Sports .

nexuso:Race rdf:subClass nexuso:Sports .

nexuso:Box rdf:subClass nexuso:Sports .

nexuso:Dance rdf:subClass nexuso:Sports .

```
%%%%%%%%%
nexuso:RareOrNovelMetaphors rdf:type owl:Class .

nexuso:RareOrNovelMetaphors rdf:subClass owl:CovidFrame .


nexuso:Religion rdf:subClass nexuso:Sports .

nexuso:Cult rdf:subClass nexuso:Sports .

nexuso:Plant rdf:subClass nexuso:Sports .

nexuso:DentalCavity rdf:subClass nexuso:Sports .

nexuso:MovieSeries rdf:subClass nexuso:Sports .


%%%%%%%%%
nexuso:War rdf:type owl:Class .

nexuso:War rdf:subClass owl:CovidFrame .


%%%%%%%%%
nexuso:Combat rdf:type owl:Class .

nexuso:Combat rdf:subClass owl:CovidFrame .


%%% THREAT
nexuso:Threat rdf:type owl:Class .

nexuso:Threat rdf:subClass owl:CovidFrame .


%%%% The Following are copied from Lemon, but modified

nexuso:LexicalEntry
   a rdfs:Class, owl:Class ;
   rdfs:comment "An entry in the lexicon. This may be any morpheme, word, compound, phrase
or clause that is included in the lexicon"@en ;
   rdfs:comment "It is copied from Lemon, but modified"@en ;
   rdfs:label "Entrada léxica"@es, "Entrée lexicale"@fr, "Lexical entry"@en, "Lexikaal
item"@nl, "Lexikoneintrag"@de ;
   rdfs:subClassOf :HasLanguage, :HasPattern, :LemonElement, [
      a owl:Restriction ;
      owl:minCardinality "1"^^xsd:nonNegativeInteger ;
      owl:onProperty nexuso:lexicalForm
   ] .
```

```
nexuso:lexicalForm
   a rdf:Property, owl:InverseFunctionalProperty, owl:ObjectProperty ;
   rdfs:comment "Denotes a written representation of a lexical entry"@en ;
   rdfs:domain nexuso:LexicalEntry ;
   rdfs:label "Forma léxica"@es, "Forme lexicale"@fr, "Lexical form"@en, "Lexikaal vorm"@nl, "Lexikalische Form"@de ;
   rdfs:range nexuso:Form .

nexuso:Form
   a rdfs:Class, owl:Class ;
   rdfs:comment "A given written or spoken realisation of a lexical entry"@en ;
   rdfs:label "Form"@de, "Form"@en, "Forma"@es, "Forme"@fr, "Vorm"@nl ;
   rdfs:subClassOf nexuso:LemonElement, [
      a owl:Restriction ;
      owl:minCardinality "1"^^xsd:nonNegativeInteger ;
      owl:onProperty nexuso:representation
   ] .

nexuso:representation
   a rdf:Property, owl:DatatypeProperty ;
   rdfs:comment "A realisation of a given form"@en ;
   rdfs:domain nexuso:Form ;
   rdfs:label "Darstellung"@de, "Representación"@es, "Representation"@en, "Représentation"@fr, "Voorstelling"@nl ;
   rdfs:range xsd:string .

nexuso:writtenRep
   a rdf:Property, owl:DatatypeProperty ;
   rdfs:comment "Gives the written representation of a given form"@en ;
   rdfs:domain nexuso:Form ;
   rdfs:label "Representación escrita"@es, "Représentation écrite"@fr, "Schriftelijke voorstelling"@nl, "Schriftliche Darstellung"@de, "Written representation"@en ;
   rdfs:range xsd:string ;
   rdfs:subPropertyOf nexuso:representation .

%%%%

nexuso:frame
   a rdf:Property, owl:InverseFunctionalProperty, owl:ObjectProperty ;
   rdfs:comment "Link a lexical entry with a methafor frame"@en ;
   rdfs:domain nexuso:LexicalEntry ;
   rdfs:label "Forma léxica"@es, "Forme lexicale"@fr, "Lexical form"@en, "Lexikaal vorm"@nl, "Lexikalische Form"@de ;
   rdfs:range nexuso:CovidFrame .
```