

SHORT TERM SCIENTIFIC MISSION (STSM) – SCIENTIFIC REPORT

The STSM applicant submits this report for approval to the STSM coordinator

Action number: CA16105

STSM title: Crowdsourcing collocations for language learning purposes

STSM start and end date: 03/06/2018 to 09/06/2018

Grantee name: Elena Volodina

PURPOSE OF THE STSM/

(max.500 words)

The main aim of this STSM was to prepare a crowdsourcing experiment that we will run together with the members of the WG1. We started from the idea of turning a (collocation) dictionary into a language learning resource based on the project run in Slovenia (Kosem et al. Forthcoming; <https://www.cjvt.si/kssj/en/>).

In particular, we needed to decide on:

- the data,
- research question(s),
- task type and scope in relevance to the research question(s),
- ranking approach,
- platform and time scale.
- involvement of the specialists that would make the whole experiment successful,
- the “type of” crowd and its size, as well as
- possible extensions of the initial experiment.

Besides the experiment planning, we needed to discuss the contents of the Teacher survey that is under preparation within WG1, feasibility to launch the in the near future, as well as the status and next steps in Literature review deliverable.

DESCRIPTION OF WORK CARRIED OUT DURING THE STSMS

(max.500 words)

The STSM was a successful event, where meeting researchers from Slovenia (Iztok Kosem, Simon Krek, Jaka Cibej, Spela Arhar Holdt) helped to brainstorm the above-mentioned questions, design the experiment, define expert/crowd involvement and set a timeline. As it looks now, the experiment (crowdsourcing task) will be launched in September-October 2018, results will be analysed during November and presented at the WG1 meeting in December 2018.

We had multiple intensive meetings daily. On June, 7, I gave a presentation to researchers from the University of Ljubljana and Trojina Institute, where a good portion of audience represented the highly needed group of L2 researchers, teachers and assessors. LINK HERE TO THE LECTURE INVITATION + DISSEMINATION – SPELA?

DESCRIPTION OF THE MAIN RESULTS OBTAINED

(max. 500 words)

Main results:

- **Call for participation in the hands-on workshop for WG1 has been prepared, where we defined the**
 1. **Program of the workshop, including invited speaker(s)**
 2. **Reimbursement criteria**
 3. **Involvement criteria**
 4. **Set up the workshop webpage (<https://spraakbanken.gu.se/eng/wg1-sept18-gbg>)**
 5. **Designed the registration form and the type of information we would need to know (<https://spraakbanken.gu.se/eng/WG1-Dec2018-registration>)**

- **Design of the pre-workshop crowdsourcing experiment for multi-word expressions and collocations**
 1. **Data: 50% Dante database + 50% English vocabulary profile**
 2. **Data size: 60 headwords containing at least two MWEs**
 3. **Approach: best-worst scaling**
 4. **Platform: pyBossa, Slovenian installation**

- **Task distribution and involvement (tasks 1 and 2 below will potentially be carried out during a coming STSM)**
 1. **Data preparation: Sweden, SB**
 2. **Data import and task design: Slovenia**
 3. **Dissemination of the task: Slovenia, Sweden → WG1 members coming to the hands-on WG1 workshop**
 4. **Data export: Slovenia**
 5. **Data post-processing: Sweden**
 6. **Experiment analysis: Sweden and Slovenia**

- **Teacher survey**
 1. **Status: ready to launch with slight modifications**
 2. **Period: launch - late June, open till end of September; analysis before the WG1 workshop in December 2018**
 3. **Responsible per country: MC representatives, information mail will be sent in June**

FUTURE COLLABORATIONS (if applicable)

(max.500 words)

As described above, there are plenty of tasks to address, and there is a collaboration plan spanning June-December 2018 already sketched there.

On top of this experiment, we plan to write an article

Ideally, the hands-on WG1 workshop will lead to a cross-linguistic crowdsourcing experiment on MWEs and collocations in language learning and teaching context.