Deliverable Number: 3.5

Deliverable Title: First SERISS Symposium on synergies between survey translation and developments in translation sciences. Minutes of the expert meeting

Work Package: 3

Deliverable type: Other

Dissemination status: Public

Submitted by: ESS ERIC, Universitat Pompeu Fabra (UPF)

Authors: Diana Zavala-Rojas

Date Submitted: August, 2017

www.seriss.eu   @SERISS_EU

SERISS (Synergies for Europe's Research Infrastructures in the Social Sciences) aims to exploit synergies, foster collaboration and develop shared standards between Europe's social science infrastructures in order to better equip these infrastructures to play a major role in addressing Europe's grand societal challenges and ensure that European policymaking is built on a solid base of the highest-quality socio-economic evidence.

The four year project (2015-19) is a collaboration between the three leading European Research Infrastructures in the social sciences – the European Social Survey (ESS ERIC), the Survey of Health Ageing and Retirement in Europe (SHARE ERIC) and the Consortium of European Social Science Data Archives (CESSDA AS) – and organisations representing the Generations and Gender Programme (GGP), European Values Study (EVS) and the WageIndicator Survey.

Work focuses on three key areas: Addressing key challenges for cross-national data collection, breaking down barriers between social science infrastructures and embracing the future of the social sciences.

Please cite this deliverable as: Zavala-Rojas, D. (2017) *First SERISS Symposium on synergies between survey translation and developments in translation sciences. Minutes of the expert meeting*. Deliverable 3.5 of the SERISS project funded under the *European Union's Horizon 2020 research and innovation programme* GA No: 654221. Available at: www.seriss.eu/resources/deliverables

**1st SERISS Symposium on synergies between survey translation
and developments in translation sciences**
Universitat Pompeu Fabra, Barcelona
June 1-2, 2017

Minutes of the expert meeting
Diana Zavala-Rojas
(ESS ERIC, UPF)

Summary

This report summarises the discussion of the first expert meeting organized as part of the "1st SERISS Symposium on synergies between survey translation and developments in translation sciences" which took place on June 1-2, 2017 at Universitat Pompeu Fabra (UPF) in Barcelona. SERISS Work Package 3, which is dedicated to maximising the equivalence of survey items through translation. In particular, Task 3.2 addresses the "feasibility of applying computational linguistic methods to survey translation.

The expert meeting, with a round table format, brought together 15 experts from the public and private sectors in the fields of comparative survey methodology, computational linguistics, and translation and language sciences. One main theme and three related subtopics shaped the round table: 1) how survey research - including survey translation, its assessment, and the management of the translation process - could learn from the past with respect to machine or automated translation (MT) and computer-assisted translation (CAT) tools. 2) Methods developed in computational linguistics that can be potentially used for translation assessment. 3) How the translation process could be facilitated based on the information that is already available (e.g., formerly validated translations of survey items). 4) how to gather evidence (i.e., how to design experimental research) to assess language or translation effects in survey questions. After an overview describing the task, this deliverable presents the summary of the discussion.

Overview

Synergies for Europe's Research Infrastructures in the Social Sciences (SERISS) is a collaboration between the European Social Survey (ESS ERIC), the Survey of Health, Ageing and Retirement in Europe (SHARE ERIC), the Consortium of European Social Science Data Archives (CESSDA AS), the European Values Study (EVS), the Generations and Gender Programme (GGP), and the Wageindicator Survey. The ultimate goal of the project is to both strengthen and harmonize social science research in Europe.

The "1st SERISS Symposium on synergies between survey translation and developments in translation sciences" took place on June 1-2, 2017 at Universitat Pompeu Fabra (UPF) in Barcelona. It was organised by the European Social Survey ERIC, and chaired by Diana Zavala-Rojas, SERISS

Task 3.2 leader.

The symposium was organized as part of SERISS Work Package 3, which is dedicated to maximising the equivalence of survey items through translation. In particular, Task 3.2 addresses the "feasibility of applying computational linguistic methods to survey translation." Computational linguistics is an interdisciplinary area of research which applies statistical modelling and computational developments to language processing.

The general objective of the symposium was to bring together a group of specialists from survey methodology and translation and language sciences to initiate a roadmap for the potential synergies between survey translation and assessment and the technological developments in language and translation sciences. A second goal was to form a multidisciplinary network of researchers that would enhance the possibilities for collaboration. The symposium brought together 15 experts from the public and private sectors in the fields of comparative survey methodology, computational linguistics and translation, and language sciences.

According to the SERISS work programme, the mandate of Task 3.2 consists in organising an expert meeting. At the stage of initial contacts with experts, the conveners realised that it would be very difficult to bring together experts from three disciplines, i.e. survey methodology, translation, and computational linguistics, to discuss openly areas for potential collaboration in the field of survey translation, if they did not get familiar with each other's work first. Therefore, the convener changed the meeting's format to a closed symposium i.e. experts were invited to present their work to the group without an open audience before starting the 'meeting' in the round table format.

It was agreed with the experts that the presentations will not be made public, but that they will be used indirectly in Deliverable 3.6 that reports the state of the art in computational linguistic methods and survey translation, because they cover key areas that bridge survey research and translation and language sciences.

The first part of the symposium, after the general welcome and *tour de table,* consisted of participants presenting their areas of expertise, grouped into five themes, and a key note presentation. The first theme was "**Translation process in comparative surveys".** There were presentations of the translation process in the ESS ERIC, the Survey of Health, Ageing and Retirement in Europe (SHARE), and in the European Values Study (EVS). The second topic was "**Technological advancements in translation studies - what do they offer to survey translation?"** The third theme centred around "**Making use of past experience: managing the translation process in large scale surveys".** The fourth topic was **"Survey translation and assessment: possibilities for incorporating automated procedures"**, finally the fifth topic was **"Research to assess translation effects".** Toni Badia, professor of computational linguistics and director of the PhD program in that area at UPF, gave a key note presentation on the history of the use of technology in translation studies and translation assessment. The content of the presentations from the symposium will be included in SERISS Deliverable 3.6, a report in the form of an academic article that presents to survey methodologists' key terms from translation and language sciences, and summarises the state of the art regarding advancements from computational linguistics that can be applied to survey translation.

On the second day, an expert meeting in the format of a round table took place. Participants openly discussed how to establish a roadmap for the potential synergies between survey translation and assessment, and for technological developments in language and translation sciences. One main theme and three related subtopics shaped the round table. The first theme centred around how survey research - including survey translation, its assessment, and the management of the translation process - could learn from the past with respect to machine or automated translation (MT) and computer-assisted translation (CAT) tools. The second topic focused on the methods developed in computational linguistics that can be potentially used for translation assessment. The third topic was about how the translation process could be facilitated based on the information that is already

available (e.g., formerly validated translations of survey items). The final theme dealt with the question of how to gather evidence (i.e., how to design experimental research) to assess language or translation effects in survey questions.

Appendix 1 contains the detailed agenda of the symposium, Appendix 2 the list of participants, and Appendix 3 the short biographies of the participants. A glossary of survey translation and computational linguistics terms is provided at the end of the document to assist readers.

To keep the privacy of participants and respect the open environment for discussion, the minutes of this roundtable have been summarized and refer to actors in terms of the larger group of which they form part: experts from the survey infrastructures, academic experts in computational linguistics and experts from the private sector.

The chair of the meeting had prepared, in advance of the meeting, some questions in each of the themes to structure the discussion. Participants expressed their spontaneous opinions to answer questions according to their area of expertise. The summary that follows this introduction is organised by themes, with the first one being the main topic that was discussed for a longer time than the other topics. For each theme, the summary presents the open questions that were presented to participants and the summary of the answers. When there were clear conclusions, they are presented in a short conclusions section.

Summary of the expert meeting (round table)

Theme 1: learning from the past

The participants discussed advancements related to Computer Assisted Translation (CAT) and Machine Translation (MT) and the extent to which it was possible to automate existing translation procedures. Although the expert meeting was not about the Translation Management Tool (TMT), this CAT tool originally developed for SHARE and being extended for use on ESS, EVS and GGP as part of the SERISS project, was, spontaneously, focused on during the round table because during the first part on the first day, when the survey infrastructures presented their translation processes, this tool was mentioned and the TMT is the most imminent application of CAT or MT technologies for the surveys participating in this roundtable.

Q1: **Why is survey translation assessment complicated?**.

This discussion allowed academic experts from computational linguistics to get familiar with the challenges survey researchers face in the evaluation of translated questionnaires.

Survey experts listed several reasons why this is difficult:
1) In the surveys participating in this meeting, the translations are done on the national level, and only guidelines are given by the central coordination teams, but they are not involved in the translation processes.
2) Translation assessment is very complicated because there are many different languages in a comparative survey and survey experts need to ensure that the quality of the translations of all languages is good and compares well across languages, also of those they do not speak by themselves.
3) A seemingly tiny translation mistake can have big consequences as it may be that the data is not comparable anymore and this is expensive and can lead to wrong substantive conclusions.
4) There are many nuances in languages which can easily change the meaning of a survey question when it is translated. These nuances can in many cases not be captured by the existing CAT tools in the market yet.

5) Translations are not always carried out by trained translators but, for instance, by PhD students or survey researchers, therefore assessment is particularly important to monitor translation quality.

Experts from the survey infrastructures recognised that a big limitation of survey translation is that in many cases translators are not professionals. Besides that, in some surveys the translation process is largely decentralized at the country level but translation assessment is centralized.

Consequences of this are that national teams require a user friendly tool: For instance it was mentioned during the roundtable that, in some cases, translators have done their translations outside of CAT tools and then copied their translation into it, because the tools were not easy to use for them. This way the documentation function of the tools (the TMT, as an example) – that is, documenting the whole translation process including all steps from the very first drafts to the final versions – is not used which does not correspond to the ideas behind introducing these tools. So, tools need to be user friendly for all kinds of people who translate. If a CAT tool is used for survey translation, this may in the end contribute to minimizing the effect of the lack of professional translators, but highly experienced and trained translator will never be replaceable by CAT tools or MT. The second consequence is that translation assessment becomes a fundamental aspect where technology can aid the survey infrastructures.

Academic experts from computational linguistics and the private sector recommended that survey infrastructures conduct a detailed review of the processes and techniques that are commonly used at each step in the translation of their survey questionnaires, to analyse which steps should be improved, modified or possibly even left out. For instance, translation scientists strongly recommended against using back translation because, from a translation science point of view, evidence has shown it does not work – but nevertheless, there are still surveys using it. In at least one of the surveys present at this meeting (the ESS), the whole translation and assessment process is transparent and documentation is easily available on its website or other publications, but not so from other surveys participating in this meeting.

The revision of the procedures should be done by translators, linguists and technical experts jointly. Analyses should be carried out to try and determine why some steps are working better than others. In the case of translation errors, it should be analyzed why they happened and how they can be avoided. For example, quite often the same instructions to the respondents are repeated throughout a survey questionnaire. If these expressions are translated each time, and written down by human subjects instead of being automatically called by a program, they can result in inconsistent translations or simply in typographical errors. One possibility for preventing inconsistencies when texts that were already translated are used could be using a CAT tool and translation memories, because they can help automating the use and format of repeated text.

Q2. Is survey translation ready to incorporate a more intense use of language technologies?

Pre-editing the source and translated texts is a precondition for the use of technology, both for translation and assessment purposes. Some basic rules: Machine translation (MT) uses machine learning techniques, while 'Machine learning' requires formatted datasets. If translation and translation assessment is to be carried out or assisted by technological means, there is a need to build corpora that can then be analysed afterwards by machines: if there are 17 languages to assess, you need 17 corpora. Commercial automated tools, already available on the market, provide functionalities to automatically format databases. So, for instance, the TMT should also ideally store the source and translated texts in the standard form of commercial tools that use corpora, because this would allow using this data with machine learning techniques at a later point in time.

Experts pointed out the possibilities that Machine Translation (MT) already offers nowadays: with the current technology (commercial MT programs used in the translation business) it is indeed possible

(and already a standard practice) to quickly obtain translations after entering source text ('Google translate' is an example with a similar functionality that may be familiar to many). As there are many survey questions that have already been translated, a CAT tool – into which existing survey translations have been fed and that includes Translation Memory (TM) functionalities – would recognize new source text that is similar to source text that already exists in the database. If MT is used, the tool could provide the user with a quick and dirty translation draft. Human translators would have to post-edit, that is, fine-tune this quick and dirty translation before it could be used as one of the initial translations in the team approach, for instance. For computational linguists, programming such a process is already a standard procedure.

Although survey experts pointed out that the team approach has certainly improved the quality of translation, for academic experts in computational linguistics, it does not seem to make sense to have two translators translating, one reviewer, one adjudicator, and then a large assessment procedure. It seems more efficient to try to prevent errors from the beginning, drafting the source text in multilingual teams, avoiding terms which are problematic in the different languages, and starting the translation process early during the questionnaire design, and making use of automated tools for regular expressions and translation memories. A suggestion was to get a quick, automated translation of a new survey question in many target languages, with the aid of technology, to foresee problems.

Survey experts added that some of these recommendations are indeed already practiced in some of the major cross-national surveys, such as the ESS: the ESS has been involving as many languages as possible in its questionnaire development since its beginning, and has been using 'advance translation' since its 5th round for this purpose (albeit this is carried out by human translators and not machines); and there are other forms of 'Translatability Assessment', carried out by experienced – human – survey translators during the questionnaire translation in order to detect and iron out potential translation problems in the source text. So adding a quick automated translation in several languages may be worth a trial indeed in order to foresee potential problems from as many languages as possible. However, the quality of this technology and in particular the database will of course be crucial for detecting problems also at a very nuanced level.

Conclusion:
1) For the inclusion of technology in the survey translation processes, a thorough revision of the complete set of translation and translation assessment procedures by survey methodologists, linguists, translators, and linguistic technologies experts seems necessary.

2) To incorporate a more intense use of language technologies in survey translation, pre-editing the source texts and the target ones is a necessary precondition. Every language included needs a validated corpus to be analysed afterwards.

Q3. Does technology allow survey researchers to simplify the translation process and, if so, at what steps?

Academic experts in computational linguistics concur that it is essential to do a thorough identification of all minor and major steps involved in questionnaire translation according to the 'team' or 'committee approach' (Harkness, 2003). There is currently no commercial tool or technological advancement ready for all the steps (altogether) involved in survey translation under the committee approach (the TMT does provide the functionalities to support translating questionnaires following the committee approach but is not a pre-existing commercial tool). Therefore, in order to assess whether language technologies can simplify the process of questionnaire translation, academic experts in computational linguistics suggested revising the committee or team approach itself.

One possibility could be that at each step of the translation process, each expert involved should summarize their own contribution, and discuss as a group how their work is interrelated. Translation memories, machine translation, etc. may be added within this process at a later point in time, but

right now as a starting point, all actors should decide in which steps this will be most useful. For doing so, the process should be split into specific tasks, and then it should be analyzed where the existing technology is ready to be incorporated.

Conclusion: In order to assess whether the use of language technologies can simplify survey translation processes, it is necessary to analyse in detail each step to find potential tasks that can be done with the aid of technology.

The example of the Translation Management Tool (TMT)

The TMT was mentioned in the first part of the symposium in some of the presentations of the survey experts, therefore it was discussed during the experts meeting, as an example of a CAT tool currently in use by some of the surveys participating in this roundtable (SHARE, ESS, EVS).

Participants discussed specifically the opportunities and challenges of this tool. Experts from the private sector who work with several other tools such as programs for computer assisted translation (CAT), translation memories, machine translation, and thesauri, concluded that the TMT needs to be improved in order to meet commercial CAT standards. In its current form it mainly focuses on the management of the translation process but does not integrate the translation tool functionalities that are available on the CAT market, for instance, translation memories or thesauri. However, it must be stated here that the TMT is currently being upgraded and, for instance, a translation memory functionality is being added. It was stressed how important it is that the TMT needs to remain flexible, adjust over time and include all the different actors in its further development and if possible, it should, for instance, be enhanced to create more user-friendly and formatted printable versions of all intermediate and final versions of translated questionnaires, e.g. in a .pdf or .docx file (this resulted from a criticism that, at the moment of the symposium, printable versions of questionnaires created out of the TMT are not formatted and therefore not useful for proofreading purposes).
More work should be done to balance human and automatic intervention, and improvements regarding automated procedures are being worked on.

Participants concluded that the TMT needs to be further enhanced, in particular the translation tasks, the communication between the people involved, and the use of translation memories. This enhancement of the TMT is going on under different tasks within the SERISS project, so part of this symposium's recommendations are scheduled to be resolved under SERISS.

Representatives from the survey infrastructures agreed that the TMT in its current version is too much oriented as a translation process documenting tool rather than a (survey) translation tool. More survey methodologists and linguists, possibly also computational linguists, should be included in the further TMT development than is currently the case. Within SERISS WP4, the main objective was to test the TMT with regard to its fitness to be used for the 'committee approach to survey translation' (see Harkness, 2003) and to document the translation process. TMT has been successful as a documenting tool, however feedback from the three test countries in the ESS ERIC shows that TMT needs to be improved with regard to its user-friendliness and usefulness as a classical CAT tool and from an end-user perspective.

Academic experts in computational linguistics pointed out that a possibility to make economies of scale in the further development of the TMT is that survey infrastructures get together to identify the needs which are common and those that are specific. Then, a catalogue of functionalities already available in common commercial tools can be investigated to see which improvements need to be prioritized. Furthermore, to create corpora, survey infrastructures could share their source and validated translations.

Conclusion: More collaboration between the infrastructures (beyond what can be achieved under the SERISS workprogramme) is certainly necessary to share experiences, data, know-how and needs regarding the further development of the TMT. The TMT needs to evolve into a tool that also

incorporates automated procedures, such as automated spell-check and programming of regular expressions and translation memories.

Q4. Tools for survey translation: Does it make sense to explore the commercial or open source solutions rather than developing something from scratch?

Academic experts in computational linguistics suggested comparing the TMT to other tools. Maybe a tool like DéjàVu could be customized. OmegaT, an open source platform, was also suggested as it is easy to create add-ons or additional functionalities. They acknowledged that for the survey infrastructures, documenting the process is key. Therefore, although the TMT is limited as a translation tool it has the advantage that it documents the steps in a conscientious way, something that is limited in most of the classical or commercial CAT tools. Therefore, different tools could also be used for different purposes while being interrelated. For instance, translators could work on a common CAT tool, but behind that another tool could keep the documentation by keeping the history of the translation.

Q5. How to incorporate machine translation or CAT technologies in survey translation?

Academic experts in computational linguistics suggested several possibilities for how to incorporate machine translation and other technologies into survey translation processes. All participants agreed that a follow up from SERISS Task 3.2 could be to carry out empirical research implementing pilots of several possibilities to incorporate technology in the translation process, among them:
1) Once a source question is drafted, linguists translate in a multilingual team and report back to the substantive team, similar to the 'advance translation' technique or to 'Translatability Assessment' (see Dorer, 2015). This would contribute to the questionnaire design process.
2) Substantive experts draft the question, linguists revise it, substantive experts conclude and then machine translation is used in combination with professional human translators for the translations of the final source questionnaire. Similar to the 'team approach', one version would be translated with the machine, and the second one by a trained/professional translator, the other steps in the committee approach can be revised.
3) In the case of only non-professional translators being used, an experiment could consist in producing two translations: one translated by a machine and the other one by a non-professional translator, and then assess the differences between both translations with a multilingual expert team.
4) A possible experiment could consist in one version translated by a professional translator and one version machine translated and then a review could be carried out of both translations, according to the review and adjudication steps in the 'team' or 'committee' approach or TRAPD.

In any case, the experts concluded that the design of such pilot testing should depend on the target populations and the topic of the questionnaire, e.g. the different approaches could be tested with specific, sophisticated questionnaires and compared to surveys to the general population.

Is the existing body of survey questions enough to build a corpus? Can we use other corpora? What are the technical requirements and possibilities to prepare for MT, translation memories and CAT tools (in addition to the preparation of databases)?

Academic experts in computational linguistics suggested that survey experts should assess to what extent there is an existent corpus that can be used, because for machine learning, having a corpus built up with survey data is a requirement: a corpus specialized on other text types than survey questions would not be as relevant to train machines as using survey questions right from the beginning. They mentioned that MOSES works with an alignment tool developed in Stanford University. So the survey questions are the data that is prepared in a raw format, which then is aligned into MOSES. This requires an extensive work with parallel corpora including various domains

(that is, topics of questions to be included in these parallel corpora).

Different ideas were mentioned and discussed how to build such multilingual corpora to be fed into survey translation tools: As most surveys cover topics such as political issues, sociology, health topics, etc., one interesting possibility was mentioned: using the corpus built from the written transcripts of the sessions of the European Parliament, which cover most participating languages in the survey projects. Another possibility is to take the questionnaires of one of the survey infrastructures or that of the Eurobarometer. This should yield reasonable data to start checking with MOSES, a machine learning algorithm. If this works – which would mean that the translations are of reasonable quality – then a neural model can be explored. This should be combined with the use of translation memories and a separate programming of regular expressions, for instance, full sentences that are instructions for the respondent, such as "Please use this card to answer." Ideally, to produce a translated version of a questionnaire by using only language technologies and no human involvement, only less than half will be done by machine translation e.g. using prediction algorithms for new texts based on an existing database of translations, and the rest would come from translation memories and programming regular expressions. That makes it very important to have large databases of translation memories of high quality. All participants mentioned the importance of working in collaboration with CESSDA ERIC experts who are working on a thesaurus in English and French (see SERISS Task 3.3) to assess whether corpora do exist from their side that could be used to be fed into machine translation or CAT tools.

Building corpora from pre-existing (and pre-edited) survey questions may become a possible follow-up project from SERISS Task 3.2. The task leader of SERISS Task 3.3 was also present during this meeting and confirmed this would be a worthwhile way to inquire.

Experts from the private sector concluded that from their experience using CAT tools, thesauri and translation memories, there is a gap between corpora, that is, the data needed for the machine learning processes, and the way survey questions are commonly layouted and stored by survey designers. Preparing the data in the format needed for machine learning (the method behind machine translation) is a key step prior to implementing technological translation tools into survey translation processes.

Conclusion: In order to incorporate machine translation (MT) into survey translation tools, annotated databases should be prepared (corpora), this requires planning such databases carefully (preparing texts already translated) and sharing databases across the infrastructures and possibly with survey practitioners in the private sector.

Q6. Which are the challenges of the validation of the translated questions to prepare a corpus?

Academic experts in computational linguistics pointed out that in a translation memory, one only can include validated translations, if these are included ex-post. A Translation Memory also 'learns' while translation is ongoing because it runs behind the translation process; that means, it will always contain, per definition, validated AND draft – that is, non-validated and potentially wrong – translations. Human validation is a requirement for the domain data, i.e. questionnaires. Ideally, when searching in a TM, the source text needs to be seen, together with information of when and by whom and in what form the translation was carried out and, if applicable, validated (this is the case with the common CAT tools, such as SDL-Trados, where this information is shown automatically). If the translation memory is carefully designed, it is possible to take into account meta-data to filtering the questions, e.g. only translations after 2010 or validated according to a specific requirement of by a specific person or institution and so on.

Experts pointed out that in the case of a transition to use technology during the survey translation process, working with CAT or translation management tools will always require a process of validating the translations. The current practice in the translation business is that the translation

memories or MT provide suggestions and then the human translator needs to work from that (translate or post-edit), this practice can also be extended to survey translation.

It is also possible to introduce an automated validation step for existing translations – that is, all existing translations are automatically validated according to certain criteria by a machine. Currently, some tools exist that provide flexible solutions for these needs.

Translation validation may be an additional challenge and something that needs to be revised when analyzing the different steps of the translation process, as for instance the translation process in the ESS ERIC is not centralized, translations are not validated, but the process is signed off. None of the surveys present during the symposium has a proper process of validating translated questions.

Theme 2: Methods in computation linguistics for translation assessment

Academic experts in computational linguistics pointed out that translation consistency is strongly improved and facilitated when using translation memories (TM). As the TM will help in finding and, possibly, filling gaps out of its database for those terms, expressions or sentences that are missing in a translation produced by a human or a machine. For this to happen, a careful design and programming of the TM is a requirement. Resources that are saved by the TM detecting and helping to fill these inconsistencies can be applied elsewhere.

Survey experts pointed out that a problem with thesauri and existing corpora is that many times, survey questions do not have enough contextual information to select proper terms. Once the database of questions (that is, the corpus) is prepared, it is possible to work with a multilingual team to flag and foresee the complicated terms.

Experts suggested several possibilities for solving contextual challenges:
1) Validate a translation, then build corpus linguistics and deep machine learning techniques, and create experimental conditions in which these corpus linguistics and deep machine learning techniques will be applied by different actors on different texts. Then the results should be compared and analysed.
2) It is relatively easy to get web data to build up corpora. For instance, in the 'Language effects in surveys' project (Gothenburg University, see agenda and bios below), the focus is on few concepts, e.g. democracy or corruption; and these researchers managed to create huge corpora from the web around these concepts. Borrowing from this idea, survey researchers could take English master questionnaires, select specific terms in the source text that are known to be problematic, build up corpora by web research, and then check the neighbouring words of these terms both in the source and in the translated questionnaires. In this way it could be assessed whether the translated words are used in similar contexts as the words in the source language, which may help assessing whether the right words were used for the context required by the translation. Of course these corpora need to be built with great caution because on the web, text quality is in many cases not checked, one doesn't, for example, know whether texts are written by native speakers or people from the right country (in the case of 'shared languages', that is, languages spoken in more than one country, for instance), and how much text will be available in the smaller languages, for instance. So the reliability of this approach needs to be considered very carefully.

Theme 3: Facilitating tools for translators

Q7. How can survey translation make use of the evidence collected during the translation and fieldwork processes to facilitate future translation?

In surveys, information is gathered during the fieldwork from the interviewers that may be useful for survey translation purposes: in a detailed follow up, they could share their observations, then this feedback is analysed with the purpose of improving questionnaire design elements of future questionnaires. For instance, if a problem is detected in many or even all national teams of a survey, it is very likely that this is a general design problem in the source instrument. If, however, problems are detected in only one or very few national teams, then this may be a translation problem.

Experts from the private sector suggested that it is important – and possible with current technology – to collect information on the problematic cases during the fieldwork phase and then incorporate them in the CAT tools; this relates to both translators' and interviewers' feedback. This information could be collected, for instance, by translators, and then programmed to form glossaries pointing out "black" lists of terms that have been found to be problematic both in the translation and in the fieldwork phases. This would help the translation and the translation assessment teams to identify terms which should merit special attention when translating.

Survey experts mentioned that in addition to the TMT, other survey software is being developed in SERISS, in particular a 'questionnaire design documentation tool' (QDDT) and a 'question and variables database' (QVDB). In a follow-up pilot project, one idea may be to explore how to use such databases of survey questions to build bilingual or multilingual corpora from these databases for CAT tools and machine learning.

Theme 4: Gathering evidence to assess language effects

This last topic centered around how the effectiveness of the different specific steps involved in questionnaire translation processes could be tested and how computational linguistics might help with this. Amongst other questions, it was discussed how different translation options can be compared.

Q8. How can different translation options be compared?

So far, one of the ways to compare different translation options is to use cognitive interviewing in questionnaire pretests. The question is whether computational linguistics could help to evaluate the meaning of translated questions or terms.

Academic experts in computational linguistics concluded that corpus linguistics could be helpful if a certain term is always used along with other specific words – that is, to verify that words in the source and target languages are used in comparable contexts, but there is no method yet to verify the content of translated texts.

Next steps

A second expert meeting will take place at the end of year three of the SERISS project (by June 2018). The task leader is preparing the meeting based on the outcomes of this first symposium.

In preparation of the meeting, the SERISS WP3.2 task leader will conduct desk work to study which of the many possibilities for further collaboration outlined in this first expert meeting may be realistic for a follow up project. This will be done in close collaboration with SERISS survey infrastructures and the SERISS WP3 leader. At the same time, the WP3.2 task leader will keep in frequent contact with the members of the group to finish the roadmap for potential synergies between survey translation and developments in translation and languages sciences. Once realistic possibilities for further collaboration will be established, experts on those areas will be invited to participate in the second meeting, alongside with representatives of the SERISS infrastructures and the experts of this first meeting.

Glossary

Advance translation: A source questionnaire is translated during the questionnaire development process with the aim of detecting problems in the source text, the insights are used to improve the source questionnaire before it is finalized (see Dorer 2015).

Computational linguistics: An interdisciplinary area of research which applies statistical modelling and computational developments to language processing.

Computer-assisted translation (CAT): Computer-assisted translation, computer-aided translation or CAT is a form of language translation in which a human translator uses computer software to support and manage the translation process.

Corpus, pl. corpora: Text corpus, pl. text corpora is in linguistics a comprehensive and structured set of texts. It is used to do statistical analysis, linguistic analysis or to validate linguistic rules within a specific language.

> →**Parallel corpus/corpora:** A corpus that contains original texts in a source language matched with the translations of those source texts into one or more target languages.
> →**Bilingual corpora** (consisting of a source and a target language) or parallel corpora can be used to support translation and also translation assessment activities.

Machine learning: It is a subfield of computer science that explores algorithms that learn features from samples of data ('trained' data) to make predictions about those features when new data is incorporated. In the context of machine translation, machine learning algorithms are used based on validated translations to predict the translation of new source text.

Machine translation (MT): A sub-field of computational linguistics that investigates the use of software to automatically translate text or speech from one language to another (not to be confused with computer-aided translation). 'Google translate' is a well-known example of a machine translation tool that is available free of charge on the internet.

MOSES: A statistical machine translation system that allows automatically training translation models (application of machine learning). It requires a parallel corpus that is used to train a model. It uses a search algorithm that finds the highest probability translation among the exponential number of choices.

Post-editing: A process in which human translators correct and improve a machine translation.

Regular expression / repeated text: A sequence of characters that define a search pattern. In the context of this text it is used to indicate expressions that are used more than once in a questionnaire and that can always be translated in the same way, therefore a CAT tool can call them back when a text is being translated and block them for new translations. Examples are instructions such as 'please use this card' or routing instructions.

Source language: The language being translated from.

Source text/questionnaire: It is the original text/questionnaire that is to be translated into another language.

Survey translation also referred as questionnaire translation, is the communication of the meaning of a source-language questionnaire by means of an equivalent target-language questionnaire.

Target language: The language into which the source text/questionnaire is being translated.

Target text/questionnaire: The text/questionnaire being translated.

Team/committee translation approach: In the context of questionnaire translation, a 'committee / team translation approach' brings together a multidisciplinary group, e.g. professional translators and survey researchers, to produce an optimal translation version in a multi step procedure. Each stage of the team translation process builds on the foregoing steps and uses the documentation required for the previous step to inform the next. The translation, review, adjudication, pretesting and documentation (TRAPD) model proposed by Janet Harkness (2003) is an example of a committee approach.

Thesaurus pl., thesauri: It is a controlled vocabulary in which concepts are represented by terms organised so that relationships are made explicit. It is a type of Knowledge Organization System (KOS) that is used to support subject access to information.

Translation memory: A database within a CAT tool that stores 'units', that is, typically sentences or sentence-like units (headings, titles or elements in a list) that have previously been translated, to aid human translators to translate new texts.

Translation Management Tool (TMT): It is a web-based tool specially designed to allow translators to translate questionnaires without the burden of understanding complex routing and programming codes for large multi-lingual questionnaires. It has been used in several large international studies. It had originally been developed by CentERdata (University of Tilburg / Netherlands) for the SHARE project; an earlier version of this software was called 'Language Management Utility (LMU)'. Under SERISS the TMT has been adapted to support (a) the team or committee approach, and (b) other large scale studies that need to translate survey questionnaires including ESS, EVS and GGP. The program has been modularized for easier adaptation to other research infrastructures. The TMT is developed in PHP with an underlying MySQL database.

TRAPD: The Translation, Review, Adjudication, Pretesting and Documentation (TRAPD) model proposed by Janet Harkness (2003) is an example of a committee approach.

References

Dorer, B. (2015). "Carrying out 'advance translations' to detect comprehensibility problems in a source questionnaire of a cross-national survey." In *Translation and Comprehensibility*, edited by K. Maksymski, S. Gutermuth, and S. Hansen-Schirra, TransÜD 72, 77-112. Berlin: Frank & Timme.

Harkness, J. A. (2003). "Questionnaire translation". In *Cross-cultural survey methods,* edited by J. A. Harkness, F. van de Vijver, & P. Ph. Mohler, 35–56. Hoboken, NJ: John Wiley & Sons.

Appendix 1: Final agenda of the 1st SERISS Symposium on synergies between survey translation and developments in translation sciences

European Social Survey

seriss

1st SERISS Symposium on synergies between survey translation and developments in translation sciences

Thursday, June 1, 2017 [13.00-19.00]

Room 24.009, Mercè Rodoreda Building, UPF*

Friday, June 2, 2017 [9.30-15.00]

Room 20.283, Jaume I Building, UPF*

Thursday, June 1, 2017 [13-19:00] Chair: Diana Zavala-Rojas
LUNCH [13:00-14:00]
WELCOME AND *TOUR DE TABLE* **[14.00-14.10]**
1. **Translation process in comparative surveys [14.10-15.00]**   Brita Dorer, European Social Survey (ESS)   Yuri Pettinicchi, Survey of Health, Ageing and Retirement in Europe (SHARE)   Ruud Luijkx, European Values Study (EVS)
2. **Technological advancements in translation studies - what do they offer to survey translation?**   Dorothée Behr, GESIS [15.00-15.30]
3. **Making use of past experience: managing the translation process in large scale surveys**   Multi-language tool for the translation of questionnaires: NeferTT   Oscar Riviere, Kantar Public [15.30-16.00]   Asking MOSES to help with translation verification   Yuri Pettinicchi, SHARE [16.00-16.30]
COFFEE BREAK [16.30-17.00]
4. **Survey translation and assessment I: possibilities for incorporating automated procedures**   Manuel Souto, cApStAn [17.30-18.00]   Lorna Balkan, University of Essex [18.00-18.30]
5. **Technology in translation studies and translation assessment**   Antoni Badia, UPF [18.30]
DINNER [20.00][1]

Appendix 1. (Cont) Final agenda

1st SERISS Symposium on synergies between survey translation and developments in translation sciences
Thursday, June 1, 2017 [13.00-19.00]
Room 24.009, Mercè Rodoreda, Building, UPF*
Friday, June 2, 2017 [9.30-15.00]
Room 20.283, Jaume I Building, UPF*

Friday, June 2, 2017 [9.30-15.00] Chair: Diana Zavala-Rojas
1. Survey translation and assessment II: possibilities for incorporating automated procedures Danina Lupsa, cApStAn [09.30-10.00]
2. Research to assess translation effects Comparing Public Opinion - Language Effects in Comparative Survey Research Stefan Dahlberg, University of Gothenburg [10.00-10.30] Striving for evidence - research on effects of different translation variants on data Dorothée Behr, GESIS [10.30-11.00]
COFFEE BREAK [11.00-11.30]
ROUND TABLE Synergies between survey translation and advancements in translation studies All participants [11.30-14:00]
CLOSURE DAY 2 [14.00-14.10]
LUNCH [14.00-15.00]

Appendix 2: List of participants of the 1st SERISS Symposium on synergies between survey translation and developments in translation sciences

1st SERISS Symposium on synergies between survey translation and developments in language and translation sciences
List of participants

1. **Badia, Toni**, Computational Linguistics Research Group (GLICOM), Universitat Pompeu Fabra
2. **Balkan, Lorna**, University of Essex
3. **Behr, Dorothée**, GESIS
4. **Brunori, Claudia** European Values Survey (EVS), Tilburg University
5. **Colominas, Carme**, GLICOM, Universitat Pompeu Fabra
6. **Dahlberg, Stefan**, University of Gothenburg
7. **Dorer, Brita**, ESS ERIC, GESIS
8. **Luijkx, Ruud**, EVS
9. **Lupsa, Danina**, cApStAn
10. **Pettinicchi, Yuri**, Survey of Health, Ageing and Retirement in Europe (SHARE)
11. **Riviere, Oscar**, Kantar Public
12. **Souto, Manuel**, cApStAn
13. **Villar, Ana**, ESS ERIC, City University
14. **Weber, Wiebke**, ESS ERIC, Universitat Pompeu Fabra
15. **Zavala-Rojas, Diana**, ESS ERIC, Universitat Pompeu Fabra

Apologies: Pascal Chelala (Kantar Public), Susana Cabaço (GGP), Anna DeCastellarnau (ESS ERIC)

Appendix 3: Short biographical information of participants

1ˢᵗ SERISS Symposium on synergies between survey translation and developments in language and translation sciences

Short biographies of participants

Badia, Toni. Full professor, Computational Linguistics Research Group (GLiCom), Universitat Pompeu Fabra (UPF)

Toni Badia has been an active teacher, researcher and manager for more than 20 years at Universitat Pompeu Fabra. As teacher he has taught at undergraduate and graduate levels courses about computation applied to language processing: computational linguistics and machine translation, technology for translators and basic tools for natural language processing and language analysis (syntactic typology, analysis based on corpus...). As researcher he has been involved in a number of publicly funded projects at Catalan, Spanish and European levels. He coordinates the GLiCom group at UPF, cooperating with the Barcelona Media Technology Centre as scientific coordinator of the language research area. He was the vice dean of the Faculty of Translation and the head of the Department of Translation and Language Sciences. Currently, he coordinates the PhD programme.

Balkan, **Lorna.** Consortium of European Social Science Data Archives (CESSDA), Thesaurus Co-ordination Officer, University of Essex

Lorna Balkan works within the Access team and runs the multilingual-related activities required to maintain, quality assure and extend the European Language Social Science Thesaurus (ELSST), the multilingual thesaurus, including: content management; stakeholder relations; and IP review groups. She is the first point of contact for organizations, such as CESSDA members, with an interest in the multilingual thesaurus.

Behr, **Dorothée.** Cross-cultural survey methodologist, GESIS

Dorothée Behr is a senior researcher at GESIS – Leibniz Institute for the Social Sciences in Mannheim (Germany). She holds a degree in translation studies and a doctorate on applied translation studies (questionnaire translation). Her research and services focus on translation, translation assessment and cross-national web probing. Previous projects include the European Social Survey or the Programme for the International Assessment of Adult Competencies. Between April 2015 and March 2016 she served as interim professor for Applied Translation Studies at the Magdeburg-Stendal University of Applied Sciences. She is a member of the Translation Expert Panel of the European Social Survey European Research Infrastructure Consortium (ESS ERIC).

Brunori, Claudia. Intern for EVS/SERISS at Tilburg University

Claudia Brunori is Master student in Sociology and Social Research at the Universities of Trento and Tilburg. She is currently working as an intern for the Synergies for Europe's Research Infrastructures in the Social Sciences (SERISS) project and for the European Values Survey (EVS). Her research interests focus on the socioeconomic integration of immigrants.

Colominas, **Carme.** Associate professor, GLiCom, UPF

Carme Colominas completed two undergraduate degrees in philosophy and German in Spain and Germany respectively. She pursued her doctoral studies on translation and formal linguistics at the *Universitat Autònoma in Barcelona* (Spain) (Doctor in Translation 2001, UAB, Diss: Semantics Representation of light verbs, Adviser: Toni Badia). Since 1992 she has held a teaching post in translation from German into Spanish at the UPF. Since 1997 she takes part on a Postgraduate Course

teaching computer-assisted translation (CAT), machine translation (MT) and postediting courses. Besides these academic activities she has been working as a translator. Her areas of work and research include multilingual corpora, translation memories, machine translation and postediting.

Dahlberg, **Stefan.** Associate professor, University of Gothenburg

Stefan Dahlberg is Associate Professor in Political Science and holds the position as Database manager at the Quality of Government Institute. He is also connected to the Swedish National Election Study programme and the Multidisciplinary Opinion and Democracy Research Group (MOD) at the University of Gothenburg. He teaches quantitative methods at Ph.D and Masters' level at the Faculty of Social Sciences, University of Gothenburg. His research focuses on representative democracy, democratic legitimacy, quality of government, survey methodology and voting behavior. He is the project manager of the Language Effects in Surveys project founded by the Swedish Research Council.

Dorer, **Brita**. Head of the Translation Team, ESS ERIC CST, GESIS

Brita Dorer is a researcher at GESIS, specialized in the field of questionnaire translation and adaptation. She is heading the translation team of the ESS. Her scientific interests include the quality of questionnaire translations and adaptations, translation and assessment methods, translatability of source questionnaires/advance translations, intercultural aspects of questionnaire translation and translation process research. She is currently preparing a PhD on advance translations carried out for improving the translatability of survey questionnaires in the ESS ERIC. She holds a degree in English, French and Italian translation studies from Johannes-Gutenberg-University Mainz, FTSK Germersheim, where she also worked as a freelance lecturer for English-to-German and French-to-German translation, and a degree in 'Interfacultative Studies of France' from University of Freiburg. She has been involved in translating survey questionnaires into German, such as the ESS ERIC, the ISSP, the PIAAC and the Survey of Health, Ageing and Retirement in Europe (SHARE).

Luijkx, Ruud. Chair of the Methodology Group, EVS

Ruud Luijkx is an associate professor at the department of Sociology of Tilburg University (the Netherlands), an associate member of Nuffield College, Oxford, and a research fellow at the department of Sociology and Social Research of Trento University (Italy). He is the chair of the Methodology Group of EVS.

Lupsa, **Danina**. Translation technologist, cApStAn

Danina Lupsa studied multidisciplinary translation in Brussels and Saarbrücken. She combined her passion for science and languages and specialized in scientific translation. Her passion for science also sparked an interest in translation technology and she started learning how to use CAT tools in her free time. She speaks French, English, Romanian and German. As a student, she started working for cApStAn and BranTra to offer first-hand technical support for PISA 2018 and was gradually involved in other projects before joining the BranTra Team.

Pettinicchi, Yuri. Survey specialist, SHARE

Yuri Pettinicchi is a survey specialist at the SHARE Central (MEA – Max Planck Institute) since February 2016. He is part of "Operations" team where he supervises the translation procedures in 28 Country Teams. Previously he was a Post-Doc fellow (GO-IN EU program – 2 years) at the Chair of Macroeconomics of Prof. Michalis Haliassos. He contributed to develop a questionnaire in the domain of Household Finance. He holds a Ph.D in Economics from the University of Venice (2012). He worked in the SHARE project (Italian country team) from 2010 to 2013. His research interests are experimental economics, applied economics and household finance. In particular, he is interested in financial education and belief formations.

Riviere, **Oscar**. Head of translation, Kantar Public Brussels

Oscar Riviere, as head of translation, is in charge of the translation department at Kantar Public (formerly TNS opinion). He is responsible for the translation production of questionnaires and analytical reports, vendor management and quality assurance. He joined Kantar Public in July 2005 to set up the translation department from scratch. He has nearly 20 years' experience in the language industry which includes a traineeship with the European Commission's terminology unit and three years as language expert for the Official Journal of the European Union. He also worked as a freelance translator and translation manager in several translation or communication companies.

Souto, **Manuel**. Linguist and translation technologist, cApStAn

Manuel Souto is a Galician translator and linguist by education and a self-taught language engineer and polyglot. He works as a translation technologist, specializing in terminology management, localization engineering, CAT tool consultancy and quality assurance automation. He drives cApStAn's R&D effort, dedicated to researching ways to optimize cApStAn's processes and developing verification-oriented prototypes (quality assessment checker, a MT lookup engine, etc.).

Villar, Ana. Research fellow, ESS ERIC CST, City University

Ana Villar is part of the Centre for Comparative Social Surveys as Research Fellow since January 2012. Her main focus is on the Mixed Mode Methodology Programme of the ESS and, under the Horizon 2020 SERISS project, she is coordinating the recruitment of a cross-national online survey panel (the CRONOS panel). She has published numerous articles, chapters and conference presentations on the areas of web survey data collection, mixed mode data collection, cross-cultural survey research, nonresponse bias in telephone surveys, survey translation assessment methods, the use of interpreters in survey interviews and cross-cultural differences in response styles.

Weber, Wiebke. Head of ASQME research group, ESS ERIC CST, UPF [Convener]

Wiebke Weber is the head of the Advanced Survey Quality Methods (ASQME) research group as well as head of the Core Scientific Team of the ESS ERIC at the Research and RECSM, UPF (Barcelona, Spain). Within the ESS she is mainly focused on measurement quality and cross-national measurement equivalence. She is also a researcher of the SERISS project, studying web survey panelists' responding behavior as well as strategies to improve this behavior. Finally, as a survey consultant, Wiebke regularly advices on questionnaire design and survey data analyses with correction for measurement error.

Zavala-Rojas, **Diana**. Research fellow, ESS ERIC CST, UPF [Convener and Chair]

Diana Zavala-Rojas is a survey methodologist interested in all aspects of the survey lifecycle, holding a doctorate in comparative survey methodology. She is a member of the CST and the Translation Expert Task Group of the ESS ERIC collaborating on questionnaire design, translation process, measurement quality and cross-national measurement equivalence. She is also researcher in the SERISS project, studying the feasibility of applying computational linguistic methods, CAT and MT to survey translation. Her current research lines focus on 1) quality of cross-cultural survey data, analyzing the effects of minority languages in the reliability and validity questions. 2) Language effects in surveys by studying multilingual measurement equivalence. 3) Survey translation and, 4) correction for measurement errors in survey research. Other research interests are experimental designs in survey research and structural equation modelling.