Using comparable corpora for interpreters’ terminology preparation

Ran Xu
ranxu2015@outlook.com

China Foreign Affairs University
Beijing, China
Challenges

- Interpreters often work in a wide range of domains and have limited time to prepare for and activate domain-specific terminologies before interpreting.

- The terminological resources for technical meetings are rarely specific enough for interpreters to use straight away.

- Interpreters have to spend a lot of time reading through meeting documents. Interpreters’ term lists are largely done manually.
Current situation for research

- Limited research into the use of modern term extraction and concordance tools for the task of SI.
- A few previous studies mentioned the application of corpora as potential tools for interpreters.
  

- However, no empirical study to test whether the use of term extraction and concordance tools can help interpreters increase their preparation efficiency and to what extent the use of the tools influences interpreters’ SI performance.
Research objectives

- To investigate how to integrate the use of corpus tools into interpreters’ terminology preparation
- To measure the effect of using the proposed preparation procedure and the corpus tools on simultaneous interpreting performance.
The preparation procedure
Corpus-based preparation procedure and tools used

Step 1: Corpus building
- Manual collection
- Web crawlers

Step 2: Automatic term extraction
- Term extractor

Step 3: Annotation of auto-lists
- Human annotation

Step 4: Term exploration
- Concordancer

Step 5: Term management
- Excel/ term management tool
Tools to use

- Term extraction tool: Syllabs

- Concordance tool: SketchEngine’s concordance function
This preparation procedure is expected to bring a twofold benefit to simultaneous interpreters:

- Helping to form and manage their tailor-made terminology resources in their work environments
- Boosting readiness of relevant terms for quick access and retrieval in SI.
Design of a pilot experiment
# Experiment setting and procedure

<table>
<thead>
<tr>
<th>Pilot Experiment (FR)</th>
<th>Group</th>
<th>Preparation method</th>
<th>Preparation Time</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Control</strong></td>
<td>Control</td>
<td>Traditional</td>
<td>9 days</td>
</tr>
<tr>
<td><strong>Test</strong></td>
<td>Test</td>
<td>Using auto-lists and concordancer</td>
<td>9 days</td>
</tr>
</tbody>
</table>

- **Training**
  - Pre-task briefing
  - Initial preparation
  - Group practice
  - Further preparation

- **SI tasks**
  - Focus group
  - Term quiz
Dependent variables

- SI performance scores
- Terminological accuracy scores
- Terminology error types and error numbers
- Degrees of departures for each error type
Results
The effect of using both auto-lists and the concordancer in interpreting preparation

The pilot experiment with 22 trainee interpreters shows:

- The test group had significantly higher term accuracy scores (↑7.5%). (P<0.05)
- The test group made significantly fewer term omission errors (OM) in SI tasks (↓9.3%). (P<0.05)
- The test group had significantly better post-task recall of terms than the group without using any tool (↑18%). (P<0.05)
- The test group spent significantly less preparation time than the control group. (↓17%). (P>0.05)
• An increased level of term density in the source speech and working into B language affected trainee interpreters’ performance by deteriorating term accuracy in interpretations.

• However, using both tools played a significant role in mitigating the detrimental effects of increased term density in the source speech and working into B language on the numbers of serious errors and omission errors (OM) in interpretations.
Summary of results

• The preparation procedure using both the term extractor and the concordancer within ample preparation time yielded the better preparation result than the traditional preparation procedure.

• Using both tools generally helped to improve trainee interpreters’ terminological performance during SI by increasing term accuracy scores by 7.5% and reducing the number of omission errors by 9.3%.
• On the other hand, terminology preparation (through using both tools) is not a “magical cure” for all. The data shows that the preparation procedure only helped to improve students’ holistic SI scores to a certain extent (but not yielding any statistical significance).
Pedagogical implications

- This study demonstrates that training on terminology preparation by using comparable corpora could be a useful supplement to the already existing professional interpreting training.

- It is important for both students and trainers to be aware that corpora and corpus tools when used properly can assist interpreters’ terminology preparation and achieve an enhanced performance.
Limitation

- Providing the automatically-generated term lists to the test group was to avoid overburdening them. However, this is arranged at the expense of a weaker ecological validity of the experiment, as in a real interpreting assignment, the interpreters have to get the term lists by using the tool themselves.
Future work

- Future experiment:
  - the participants use term extraction tool themselves
  - SketchEngine platform
Thank you!