UniDive COST action Progress Review - Month 24

Jose F Quesada

April 14, 2025

1 Clarifications to UniDive Core Group comments

1.1 Regarding the Level of Ambition of MoU Objectives

Original comment (Page 5):

"The level of ambition of the MoU objectives is Medium."

Clarification:

The objective of the evaluation report specifically focuses on the 24-month progress report (mid-project development), and at no point does it address the MoU proposal itself, which is clearly outside the scope and objectives of this evaluation. Personally, I find the project proposal as outlined in the MoU completely adequate. This comment refers to the level of ambition in terms of objective fulfillment. In fact, the phrase "The level of ambition of the MoU objectives" is generated directly by the system.

Without entering into an unnecessary debate about specific quantitative assessments that would require considering what percentage of importance each objective has, it can be observed, even starting from the research team's own assessment, that of the 10 objectives proposed, two are in the 0-25% range, seven in the 26-50% range, and one objective in the 76-100% range. Given that this is an evaluation corresponding to the midpoint of the project duration, it has been considered that risks may appear to complete the fulfillment of the planned objectives.

1.2 Regarding the Level of Ambition of Deliverables

Original comment (Page 7):

"The level of ambition of the deliverables is low"

Clarification:

Indeed, aligned with the previous comment, my assessment refers to the level of ambition regarding compliance with the deliverables production schedule. It is important to note that of the 10 planned deliverables, only one has been completed, which is the construction of the project website.

1.3 Regarding Publications Output

Original comment (Page 9):

"In general, a fairly high number of publications have been generated, although the topics covered are relatively limited. A large majority of publications specifically address Multi-word expressions."

Clarification:

It might have been better to state "A large number" instead of "A large majority." But ultimately, what this comment intends to capture is that a relatively high percentage of the efforts derived from the project seem to be concentrated on one topic, which is Multiword expressions. As indicated, this is a totally important and relevant aspect, but even comparing the same specification of progress beyond the state of the art in the MoU, MultiWord Expressions are one topic among a very varied and broad set of aspects to consider. Therefore, it was only intended to indicate that there is a considerable focus on one aspect, in contrast to the broad and diverse approach of the MoU itself.

1.4 Regarding Publication Quality and Review Process

Original comment (Page 9):

"Furthermore, practically all mentioned publications have been contributions to workshops held as part of the project itself, and very few publications are the result of a peer-review process."

Clarification:

The motivation for this comment is that 12 of the 24 publications indicated (9 to 20) correspond to publications made in the UniDive Workshop editions. For other publications such as 4, 5, and 24, it has not been possible to determine if there has been any peer-review process. And as an additional consideration, the publications for which there is a known peer-review process specifically address the topic of Multiword expression, such as (1, 2, 21). In any case, I would like to emphasize that there is no problem with the MWE topic; I only want to indicate that it constitutes a focal point that practically seems to monopolize the main attention of the project.

1.5 Regarding Publication Significance and Impact

Original comment (Page 9):

"Regarding the level of significance, the publications can generally be considered to be integrated within the project's general objectives, although their theoretical and practical impact is relatively limited."

Clarification:

In connection with the previous comment, in this case, I wanted to highlight that the publications and results mentioned, while forming part of the key lines of this project's proposal, at this stage represent isolated aspects of what could be the research approach proposed in the MoU. Analyzing the MoU proposal, specifically page 9 where a broad group of ideas and strategies to address the challenges of this project are specified, one gets the feeling that there is no comprehensive approach to topics such as (not exhaustively but representatively):

- Spotting similarities in unrelated or geographically remote languages
- Discovery and validation of hypothesized language universals
- Increasing the cross-lingual consistency of the already existing language resources
- Creating language resources for at least 14 not yet covered endangered/extinct languages

And again, it is important to emphasize that these evaluations aim primarily to detect potential risks regarding the fulfillment of the objectives set in the project.

1.6 Regarding Publication Relevance to Core Issues

Original comment (Page 9):

"Regarding relevance, while publications align with project objectives, they have not addressed core issues"

Clarification:

The comment actually aims to highlight what was indicated in the previous point, which is that the MoU itself indicated a series of key ideas that are not currently being addressed, without detracting from what has been developed. But the impression that emerges from some of the contributions is that very specific phenomena are being considered and that the project could be a good opportunity to address central issues to linguistic universality and diversity.

1.7 Regarding Strategic Direction for Next Phase

Original comment (Page 9):

"In the next project phase, research teams should adopt a more comprehensive vision aligned with strategic goals. Both linguistic diversity and universality are crucial aspects with increased impact following LLM consolidation."

Clarification:

I do not consider or propose that the project objectives should be changed, nor of course change the funding scheme of this project. I believe that the project proposal itself as indicated and highlighted in the MoU on page 8 is sufficiently general: "Given the above state of the art, this Action will address the following main challenge: To reconcile language diversity with rapid progress in language technology." I also do not believe that a mono-paradigmatic approach focused exclusively on LLMs should be adopted. And on the other hand, I believe that the working group is actually sensitive to the current technological ecosystem, since even one of the proposed publications includes issues such as LLM prompting or cross-lingual transfer. The comment only intended to emphasize the need and convenience for efforts to analyze topics such as universality and linguistic diversity to benefit from the latest trends in research and industrial development.

1.8 Regarding Output Significance Assessment

Original comment (Page 10):

"Significance of the output was judged as "low" by the Rapporteur for 3 achievements:"

Clarification:

I do not question the relevance of these models or modules for the project. The assessment refers to the contributions regarding these issues presented as project results up to the time of their evaluation.

1.9 Regarding Progress on Language Coverage

Original comment (Page 10):

"The results obtained at this level address important but generally very specific aspects. The proposed challenges of addressing up to more than 100 languages remain very far from being achieved at this point. Additionally, special attention to low-resourced languages has not been specifically addressed."

Clarification:

In accordance with the working group's comments, the information provided does not clearly indicate the level of contributions specifically derived from the work carried out as part of this project to these objectives.