

Envisioning a Trilingual Machine Translation System for the Language Pairs – <Tamang –English – Nepali>
Bal Krishna Bal, Amrit Yonjan Tamang, Lasang Jimba Tamang

Abstract: A Machine Translation system is a system that provides a gist or tentative translation in any target language for any given text in the source language. Such a system can be very instrumental in bridging the digital divide between any two or more language speaking communities. In the context of the degrading number of Tamang speakers among the younger generations of the Tamang community and in other contexts where the knowledge of English and/or Nepali may not necessarily be good among Tamang speakers, technology like Machine Translation system can help establish a strong stake or identity from a language empowerment or enabling perspective. The poster presents a high level plan for developing a trilingual Machine Translation System for the language pairs – <Tamang – English – Nepali>. The plan involves a detailed description of the efforts for the development of the linguistic resources, namely the trilingual parallel corpora and then the technical efforts on developing a Neural Machine Translation model for realizing a Machine Translation System. It is expected that after the development of the Machine Translation System, it can be used in the multiple sectors of the society, viz., teaching & learning in the mother tongue (for example, Tamang language in this case), tourism, promotion of local culture and traditions etc.

1) Introduction

Problem Context:

- Number of Tamang speakers in the latter generations has greatly reduced – a serious concern in terms of the identity of the Tamangs.
- There are Tamang communities who are well-versed with the Tamang language but have difficulties communicating in other languages, for example, English and Nepali.
- Need for preserving and promoting the Tamang language among Tamang speakers and non-Tamang speakers.

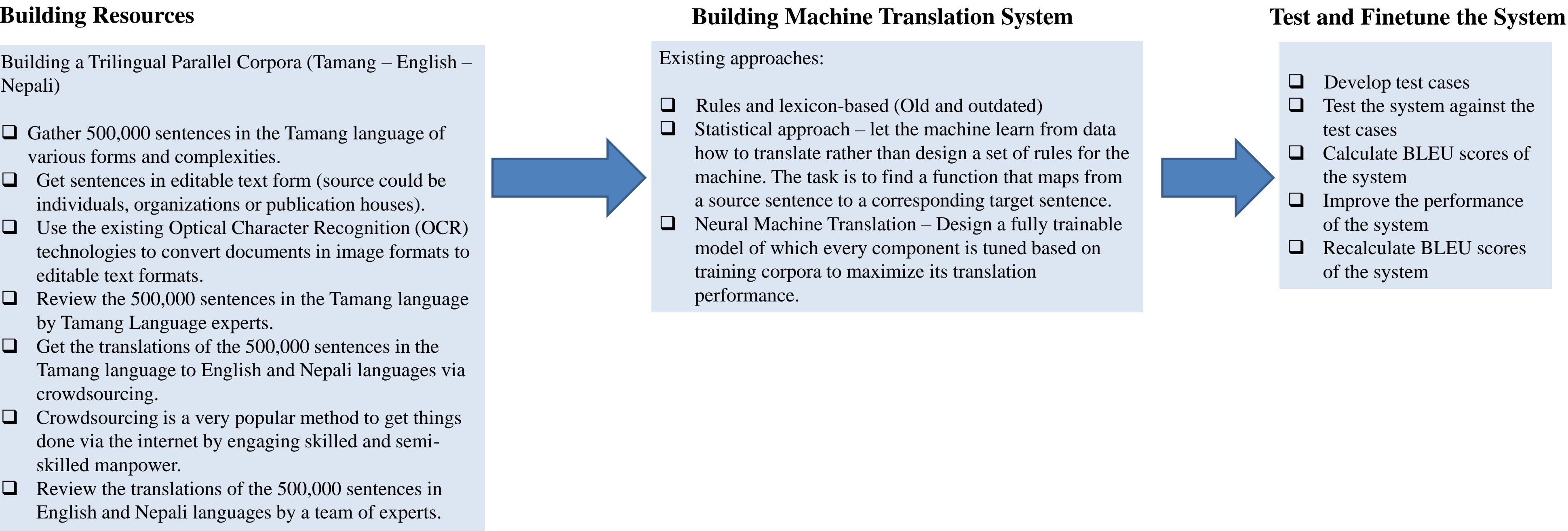
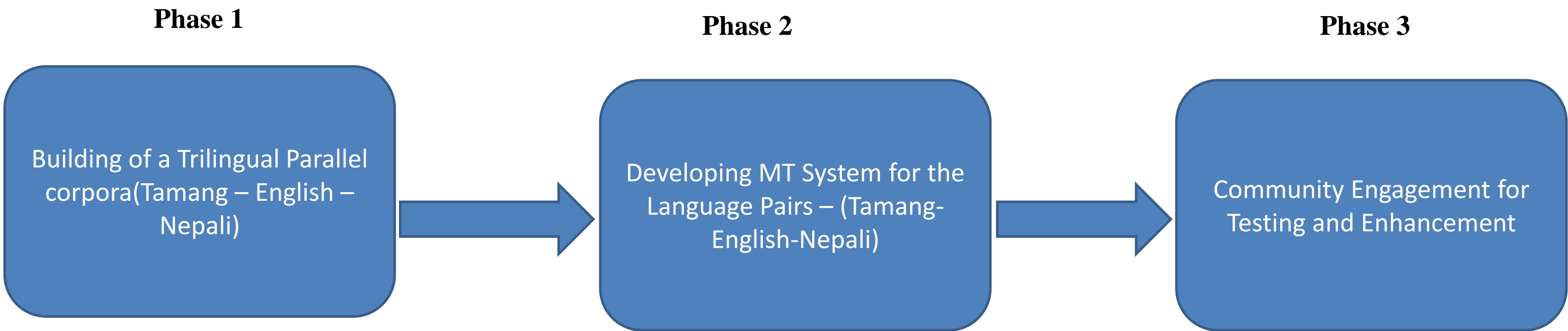
Solution Approach:

- Develop a Machine Translation System that basically provides a gist translation of any source language text into the target language text.
- The goal of a Machine Translation System is not in the least to replace human translators.

2) Methodology

- Develop linguistic resources : Trilingual Parallel Corpora (Tamang – English – Nepali)
- Develop Machine Translation System : Trilingual Machine Translation System using the latest MT Technologies
- Testing, Deployment and Enhancing the System : Community engagement in Testing, Deployment and Enhancement of the System

3) Project Design



4) Collaborators

- ☐ Mr. Amrit Yonjan Tamang – Linguist and Language Activist, Kathmandu, Nepal
- ☐ Mr. Lasang Jimba Tamang – PhD student, FedEx Institute of Technology, University of Memphis, USA
- ☐ Dr. Bal Krishna Bal – Associate Professor and Head, Information and Language Processing Research Lab, Department of Computer Science & Engineering, Kathmandu University, Nepal

5) Funding

- ☐ Total Project Period : 1 year
- ☐ Phase I : 9 K USD
- ☐ Phase II & III : 12 K USD
- ☐ Total estimated budget : 21 K USD

6) Project Funding Potential Sources:

- ☐ Tamang Community both within the country and abroad
- ☐ Government
- ☐ National and International Funding Agencies

\*Please let us know if you are interested to contribute.

7) Project Outcomes:

- ☐ Trilingual MT System for Tamang – English - Nepali
- ☐ Scientific paper publications and presentations
- ☐ Trilingual parallel corpora of 100,000 sentences in Tamang – English-Nepali

7) Way to go from here

- ☐ Will take some time before we have the funding for the Project
- ☐ Plan to start from something small – start writing Wikipedia articles in the Tamang language (1500 articles in the minimum)
- ☐ Translate these articles in English and Nepali.
- ☐ Need some small funding to start the Project, say 10 K USD.
- ☐ Having the presence of the Tamang language in Wikipedia would be the first steps for building the Trilingual Machine Translation System.
- ☐ We seek support from the Tamang community both within Nepal and abroad not just financially but also morally.

8) How to reach out to us

On Facebook: Machine Translation for Tamang, <https://www.facebook.com/groups/820029341538275/>  
Bal Krishna Bal – [balkrishna7bal@gmail.com](mailto:balkrishna7bal@gmail.com)  
Amrit Yonjan Tamang – [amrityonjant@yahoo.com](mailto:amrityonjant@yahoo.com)  
Lasang Jimba Tamang - [zimbabiru@gmail.com](mailto:zimbabiru@gmail.com)