CALL FOR PARTICIPATION
A Half Day Tutorial on
Language documentation using the Linguist’s Assistant (LA) software.
Presenter: Dr. Stephen Beale

Abstract

Consider the situation that linguists, computational linguists and Natural Language Processing (NLP) researchers in the Philippines face. With more than 170 languages and dialects being spoken, many of which have little or no documentation, the task of how to document these languages and/or compare and computationally represent each of them to enable automated translation becomes a big problem. Such documentation and representation is critical not only for the purpose of scientific research, but also because many of these languages are fast becoming extinct.

Now imagine a tool that can be used to generate descriptions of these languages, along with facilities that allow the translation of one language to another. Imagine also the researchers using the same tool to generate language resources that can be used and reused by the community.

The vision of such a tool is the motivation behind the development of a software tool known as Linguist’s Assistant (LA). The goal of this tool is to alleviate the task of documenting, representing, and translating languages. The purpose of this tutorial is to educate the concerned community about this software, and demonstrate how they can use the features incorporated into it to support their specific research needs.

About the Linguist’s Assistant

The Linguist's Assistant (LA) is a practical computational paradigm for efficiently and thoroughly describing languages. LA contains a semantic-based elicitation corpus that is the starting point and organizing principle from which a linguist describes the surface forms of a language using a visual lexicon and grammatical rule development interface. The resulting computational description can then be used in our document authoring and translation applications. LA has been used to produce extensive grammars and lexicons for English, Korean, Jula (a Niger-Congo language), Kewa (Papua New Guinea) and North Tanna (Vanuatu). The resulting computational resources have been used to produce high-quality translations in each of these languages.

Tutorial Objectives

This tutorial targets linguists, computational linguists, and Natural Language Processing researchers alike, and provides a detailed look at LA, both how LA works and how to use LA. Participants will obtain a good understanding of the LA language documentation paradigm. The goals of LA and the methodologies employed to achieve those goals will also be briefly discussed. The bulk of the class time will consist of a practical tutorial in the use of LA. Class participants will then follow a common example to set up and populate a lexicon and develop
phrase structure and surface form creation rules. As time permits, examples of more complex deep structure adjustment rules will also be covered. Finally, the tutorial will conclude with an examination of case studies based on participant’s answers to a grammar questionnaire which will be collected as part of the pre-tutorial tasks. This last activity will highlight how LA can be used to produce preliminary descriptions of the respective Philippine languages of the participants and will help them to decide whether LA can be applied to their own work/research.

**Important Dates**

Tutorial Date: March 1, 2012, 9:00 am to 12:00 pm

**Registration Information**

Participation in this tutorial requires registration and a participation fee of P350. The fee covers tutorial materials, snacks, and lunch. To register, fill-up the attached registration form and send to cpssig_nlp_secretariat@yahoo.com. Participation fees will be collected on site. Early registration is advised due to limited slots.

**Contact Information**

If you have any questions please feel free to contact any of the following:

<table>
<thead>
<tr>
<th>Name</th>
<th>Email</th>
</tr>
</thead>
<tbody>
<tr>
<td>Dr. Rachel Edita Roxas</td>
<td><a href="mailto:rachel.roxas@delasalle.ph">rachel.roxas@delasalle.ph</a></td>
</tr>
<tr>
<td>Mr. Joel Ilao</td>
<td><a href="mailto:joel.p.ilao@gmail.com">joel.p.ilao@gmail.com</a></td>
</tr>
<tr>
<td>Mr. Rodolfo Raga Jr.</td>
<td><a href="mailto:jundy.raga@gmail.com">jundy.raga@gmail.com</a></td>
</tr>
</tbody>
</table>

**Organizer**

This event is organized by the Computing Society of the Philippines Special Interest Group on Natural Language Processing (CSPSIG-NLP).

**Presenter Bio**

Dr. Stephen Beale is an Honors College Fellow at the University of Maryland. Dr. Beale received his PhD in Computer Science and Language Technologies from Carnegie Mellon University and has been involved in NLP research for twenty years. He also has considerable experience in practical linguistic fieldwork, having worked with the Summer Institute of Linguistics in Thailand and Vanuatu. He is co-developer of LA and has used LA to describe several languages and translate community development and Old Testament texts.

**Downloadable Registration Form for CSPSIG-NLP Tutorial:**

- [CPS SIG-NLP Registration Form](#)