

# CNGL Undergraduate Students as Researchers Programme PROJECT DESCRIPTION

<b><i>Institution/Track:</i></b>	TCD / ILT 3	
<b><i>Project Title:</i></b>	Development of a small hand-annotated corpus	
<b><i>Suitable for students who are studying in the following areas:</i></b>	Linguistics	
<b><i>Skills needed:</i></b>	Knowledge of syntax and semantics for analysis of text , Knowledge of foreign languages and computer science will be much appreciated.	
<b><i>Project Description:</i></b>	<p>Hand annotated corpora with semantic annotations has lead into the development of automatic semantic annotators. These automatic annotators are used in multiple applications such as: search engines, automatic text summarization, question answering systems, text paraphrasing or machine translation.</p> <p>Unfortunately, current state-of-the -art systems lose accuracy when the domain is changed. The aim of this project is to annotate an out of domain corpora for future developments of semantic role labelling systems on domain adaptation.</p> <p>The text may be annotated in FrameNet and Propbank style. Possible multiple language annotation.</p>	
<b><i>The Role of the student &amp; benefits gained from participation in this project:<sup>1</sup></i></b>	<p>The role of the students is to annotate the corpora, and coordinate any volunteers.</p> <p>The students will obtain experience in analysing and labelling text. There are publication possibilities from the project, which may be useful if the student is considering a research career.</p>	
<b><i>Who will be working with you?</i></b>	Our undergraduate students will be working closely with Hector Franco, Liliana Mamani and Gerard Lynch(all post-graduate students). The student will attend regular research meetings during his/her time with us.	
<b><i>Short description of the group:</i></b>	We have six people in our group: PIs Carl Vogel and Martin Emms, post-docs: Baoli Li, PhD students: Liliana Mamani, Gerard Lynch and Hector Franco.	
<b><i>Recommended Reading Material:</i></b>	SemEval-2010 task 10: linking events and their participants in discourse The Proposition Bank: An Annotated Corpus of Semantic Roles	
<b><i>Other information:</i></b>	This project can accept several undergraduate students.	
<b><i>For further details on this project please contact:</i></b>	<b>Name:</b> <b>Phone:</b> <b>E-Mail:</b> <b>Website:</b>	<b>Hector Franco / Gerard Lynch / Liliana Mamani</b> <b>Office: 018962017</b> <a href="mailto:francoph@cs.tcd.ie">francoph@cs.tcd.ie</a> / <a href="mailto:gplynch@cs.tcd.ie">gplynch@cs.tcd.ie</a> / <a href="mailto:mamanisl@cs.tcd.ie">mamanisl@cs.tcd.ie</a>

<sup>1</sup> *This is an initial description of the role of the student and it is liable to change following discussions between the investigators and the student.*