

# Combining Approaches to Machine Translation: the DCU Experience

Andy Way

DCU University

Until quite recently, having a 'hybrid' MT system meant enriching rules in a transfer-based system with statistics in order to constrain the processing of the system depending on different contexts.

We have conducted a number of novel pieces of research where this concept of 'hybridity' has been extended to allow sources of information other than just 'rules' and 'statistics' to be combined to good effect. These include:

- comparing EBMT and word-based SMT [Way & Gough, 2005]
- combining chunks from EBMT and PB-SMT [Groves & Way, 2005a/b, 2006]
- adding statistical language models to EBMT [Groves & Way, 2005b, 2006]
- (attempts at) combining chunks from two different EBMT systems augmenting PB-SMT with subtree pairs [Tinsley et al., 2007]
- incorporating supertags into PB-SMT [Hassan et al., 2006, 2007, 2008]
- adding source language context into PB-SMT [Stroppa et al., 2007]
- combining examples, statistics and rules in tree-based translation [Hearne & Way, 2003, 2006]

We will present the rationale behind these pieces of research, describe the various improvements made, and comment on other possible system combinations which might improve system performance further.