

# Developing an English to Afrikaans Machine Translation System

HJ Groenewald

Centre for Text Technology (CTexT™), North-West University, Potchefstroom Campus (PUK)

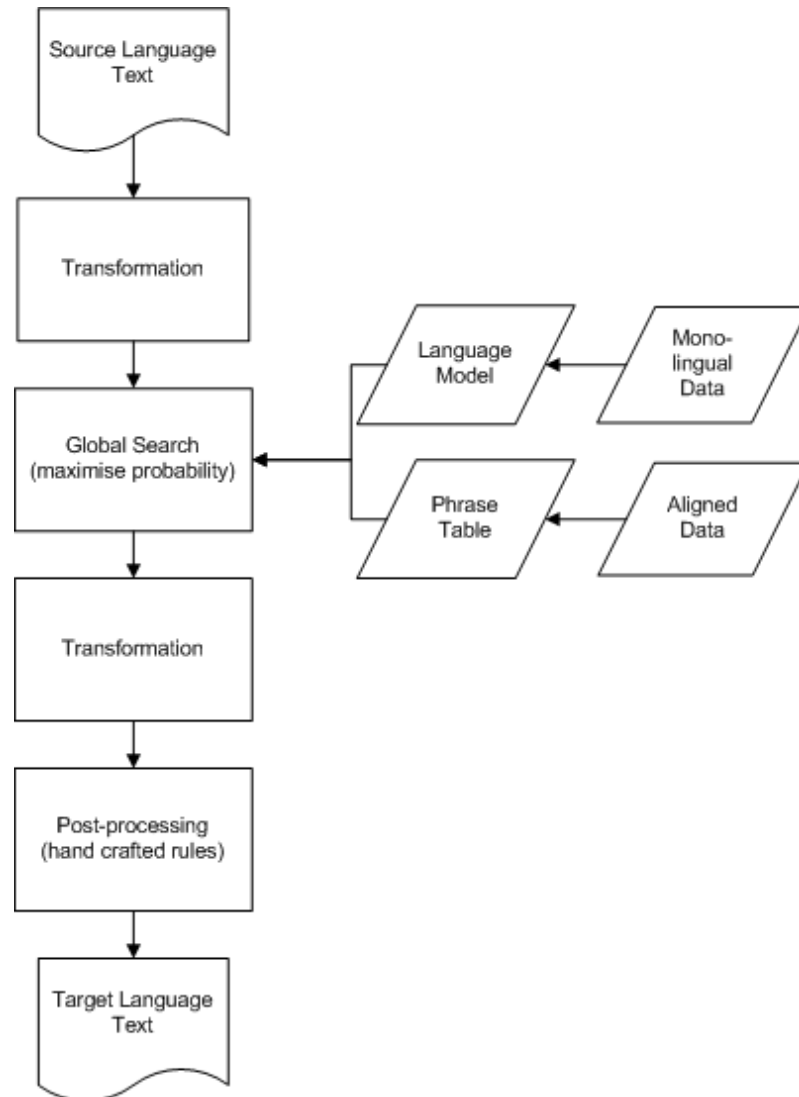
## Introduction

- Machine Translation Definition
- Machine Translation in South Africa
  - Previous projects
  - Need for Machine Translation
- Autshumato Project
  - Open source machine-aided translation tools and resources for South African Languages
    - Integrated Translation Environment
    - Terminology Management System
    - Machine Translation for 3 South African Language Pairs

# Approach

- Hybrid Approach to MT
  - Moses Statistical Machine Translation System
    - Free Software
    - Automatically train translation models for any language pair
    - Large parallel corpora required: Currently 441 124 aligned English-Afrikaans sentences
  - Hand-crafted rules
    - Based on in-depth knowledge about the languages involved
    - Various errors in the statistical machine translation output that we are trying to solve
      - Double negative not applied
      - Incorrect use of conjunctions
      - Stylistic Inconsistencies
      - Incorrect use of the attributive
      - Etc.

# Process Flow Diagram



# Results

- Metis-II evaluation set

	BLEU	NIST
Autshumato	0.5097	8.7050
Google	0.44	8.0806

- Evaluation set based on data obtained form *Gov.za*

	BLEU	NIST
Autshumato	0.5108	9.1916
Google	0.4658	8.6382