

# Introduction to Computational Linguistics

Frank Richter

[fr@sfs.uni-tuebingen.de](mailto:fr@sfs.uni-tuebingen.de)

Seminar für Sprachwissenschaft  
Eberhard-Karls-Universität Tübingen  
Germany

# The First NLP Application

## Bi-lingual Dictionaries for Word-to-Word Machine Translation

- 1947** A.D. Booth and D.H.V. Britten worked out a detailed code for realizing dictionary translation on a digital computer.
- 1948** R.H. Richens worked out a stem-affix encoding with a longest-match strategy for stem identification and translation.

# Richens' Dictionary Encoding

**An Example:** Latin *amat* = love (3rd pers. sing.)

## Dictionary Encoding:

stem dictionary: a-m

suffix dictionary: -at

## Search Strategy:

for a given input (e.g. *amat*) find longest prefix (in this case *am*) that matches the stem dictionary;

match the remainder of the string (in this case *at*) against the suffix dictionary.

**Note:** this is a direct precursor to later finite-state-encodings of computational lexica.