

Machine Translation II: Hybrid Methods and Evaluation

January 21, 2008



Andreas Eisele

UdS Computerlinguistik & DFKI

eisele@dfki.de

Language Technology I

WS 2007/8

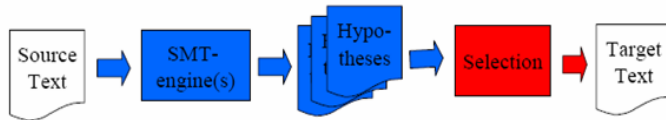
Overview

- Hybrid MT architectures in more detail
- Evaluation techniques

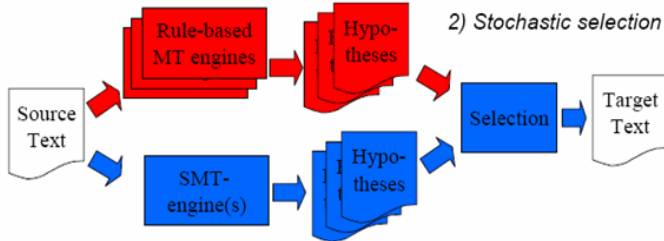
Some hybrid MT architectures

■ = SMT Module
■ = RBMT Module

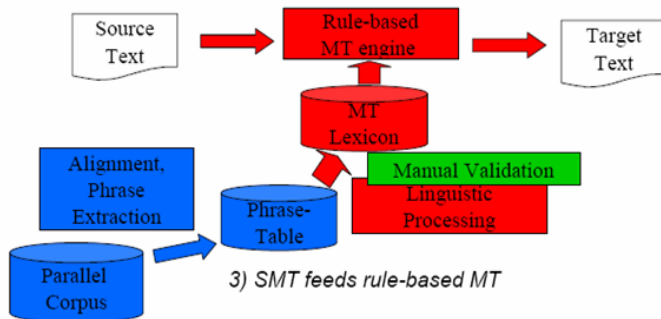
1) Syntactic selection



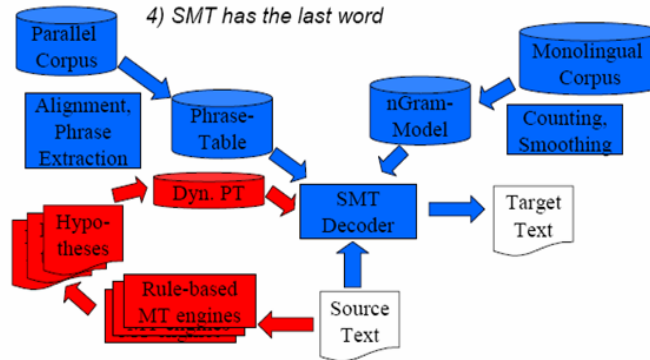
2) Stochastic selection



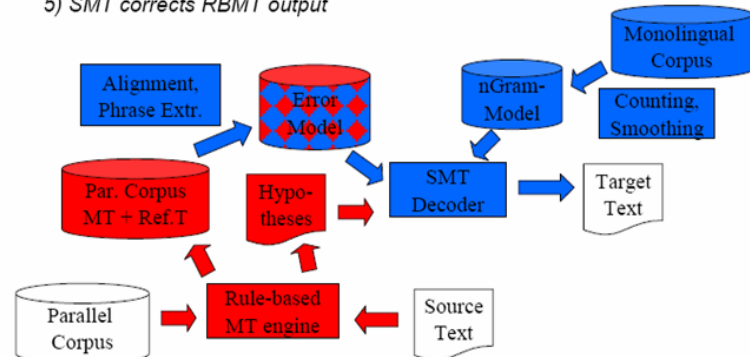
3) SMT feeds rule-based MT



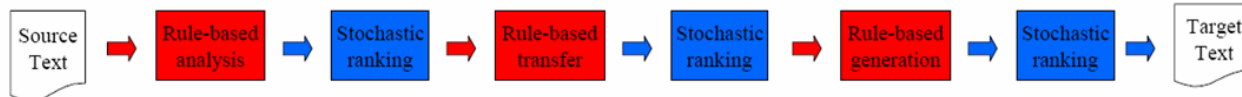
4) SMT has the last word

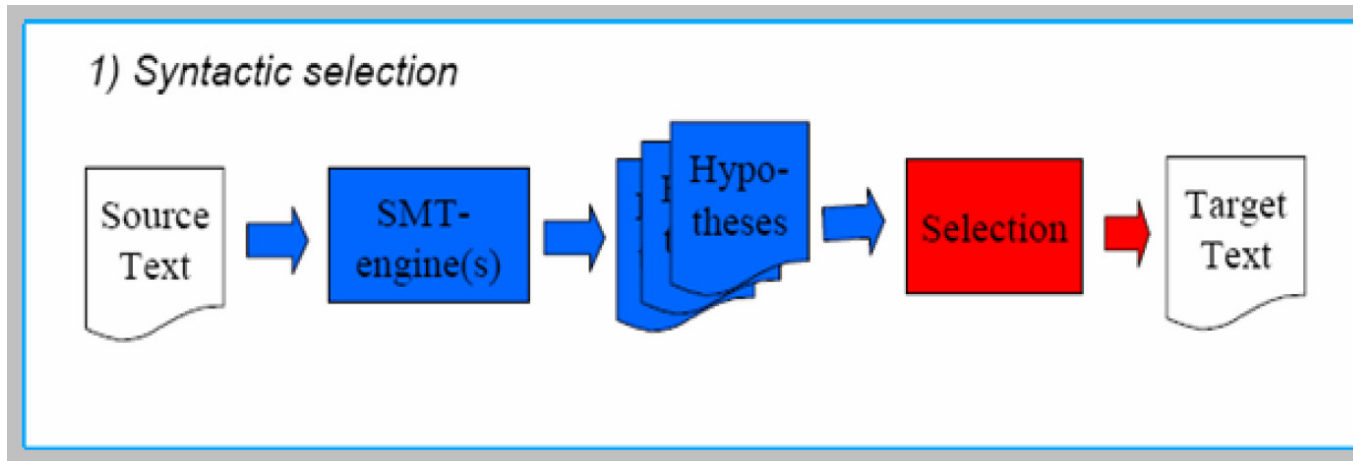


5) SMT corrects RBMT output



6) Rule-based transfer architecture interleaved with stochastic ranking



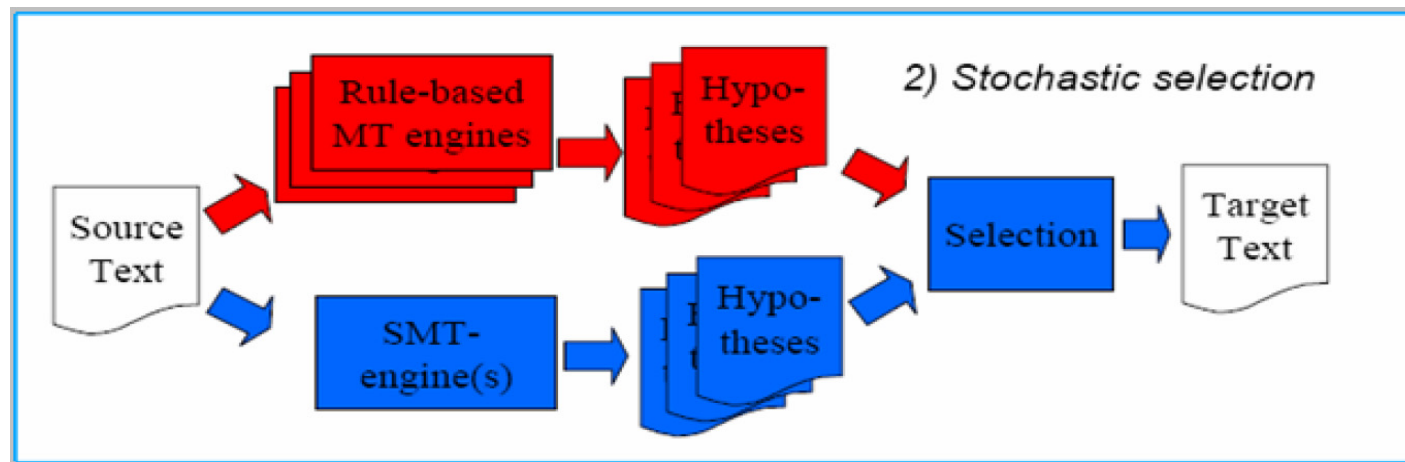


Motivation: SMT output often syntactically ill-formed

➔ Selection mechanism in SMT „generate and test“ should be enriched with syntactic knowledge

BUT:

- syntactic parsers not (yet) robust enough
- High computational cost of processing many ill-formed candidates



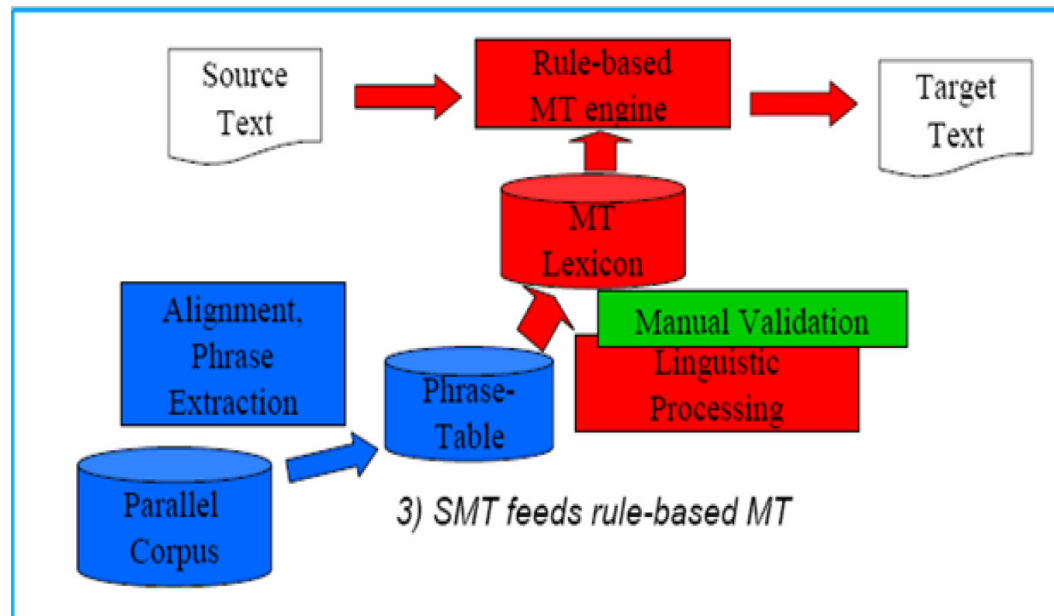
Motivation: Selection from an increased number of candidates can improve overall quality

BUT:

- Works mainly for short utterances, where one of the candidates may be good enough (VerbMobil)
- Different candidates may have problems in different parts of the sentence, granularity of decisions too coarse

Motivation:

- Adapting RBMT to new domains requires lots of new lexical entries that are difficult to write manually
- SMT techniques can help to partially automate this process



BUT:

- Not all required information can be learned from data
- Errors in examples/SMT alignment may creep in, but RBMT has no mechanism to discard implausible outcomes
- Some manual effort is required

European Patent Office (EPO):

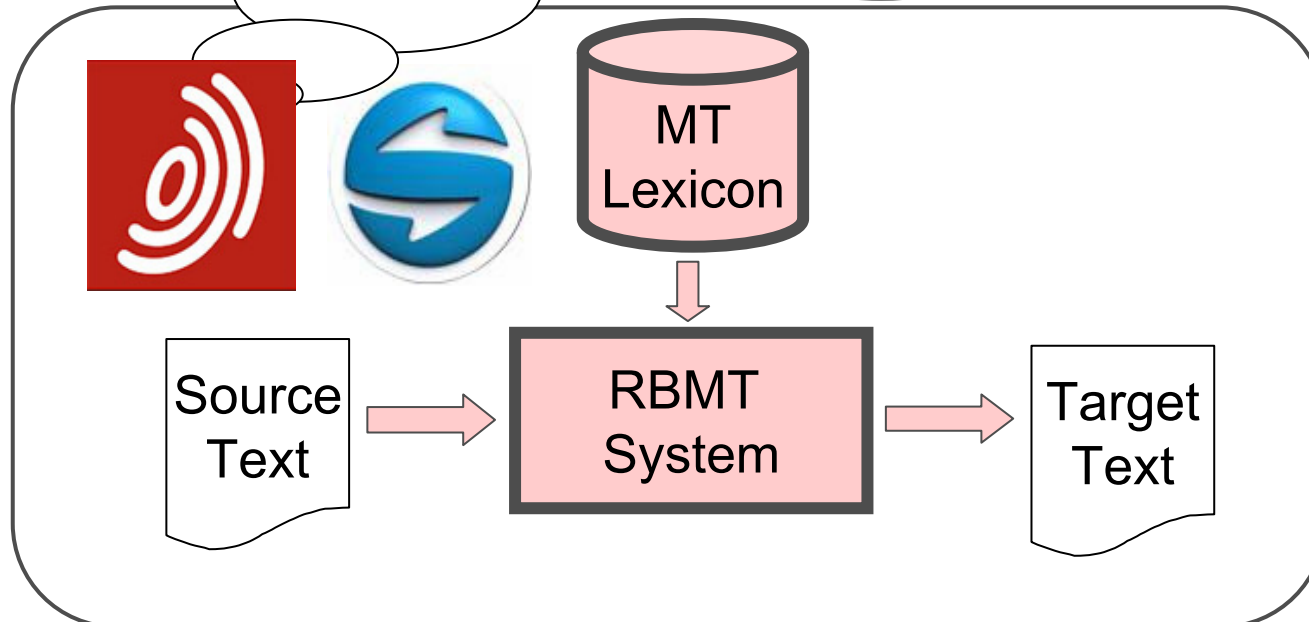
6000 employees from > 30 countries in Munich, The Hague, Berlin, Vienna, Brussels

Collection of > 60 Mio. patent documents

130000 patent applications/year (2006)

Prepares translation service for patent documents

Call for tenders & **selection test**, fall 2005



Language pairs

DE ↔ EN

ES ↔ EN

FR ↔ EN

IT ↔ EN

planned:

EL ↔ EN

PT ↔ EN

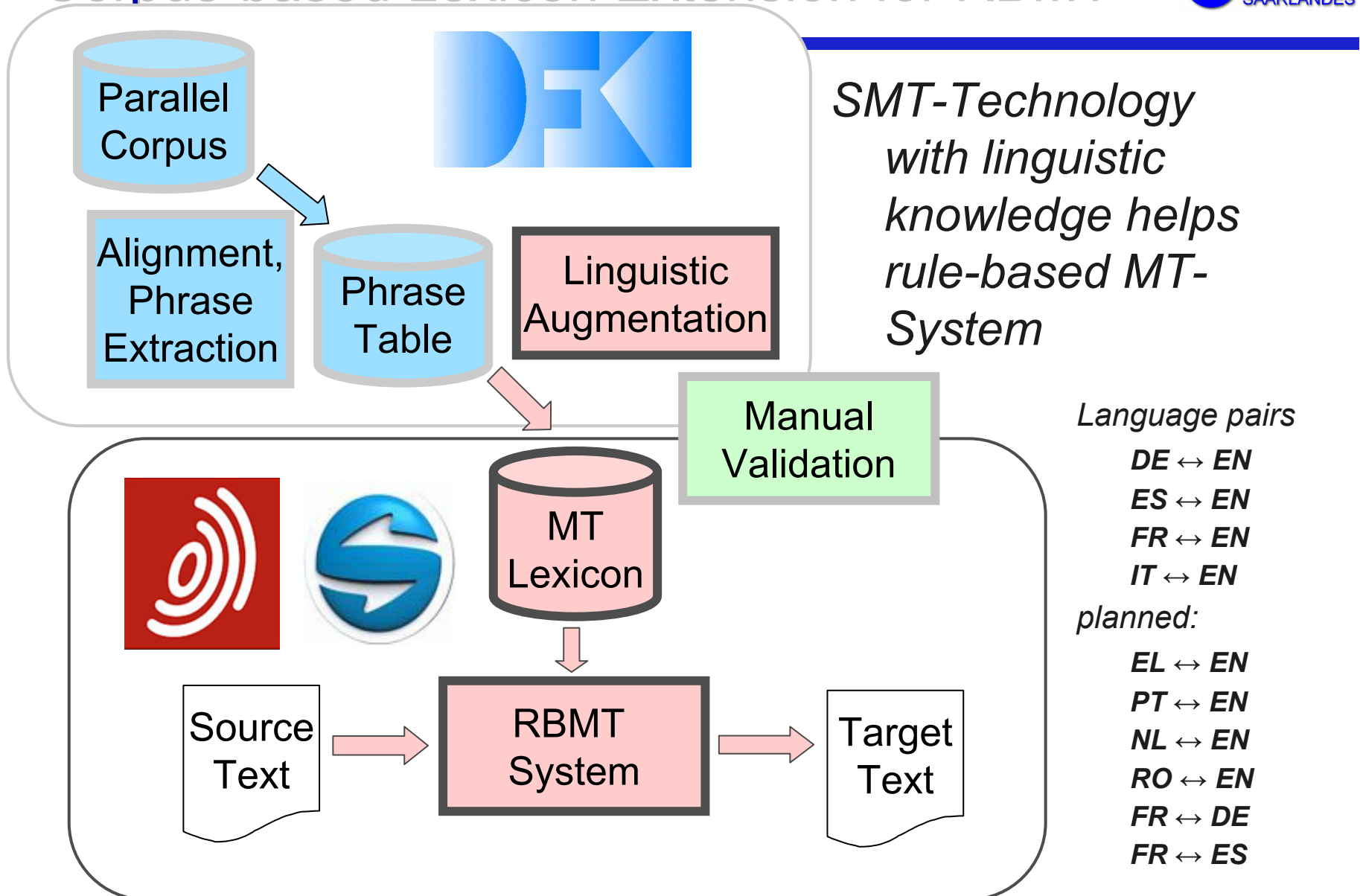
NL ↔ EN

RO ↔ EN

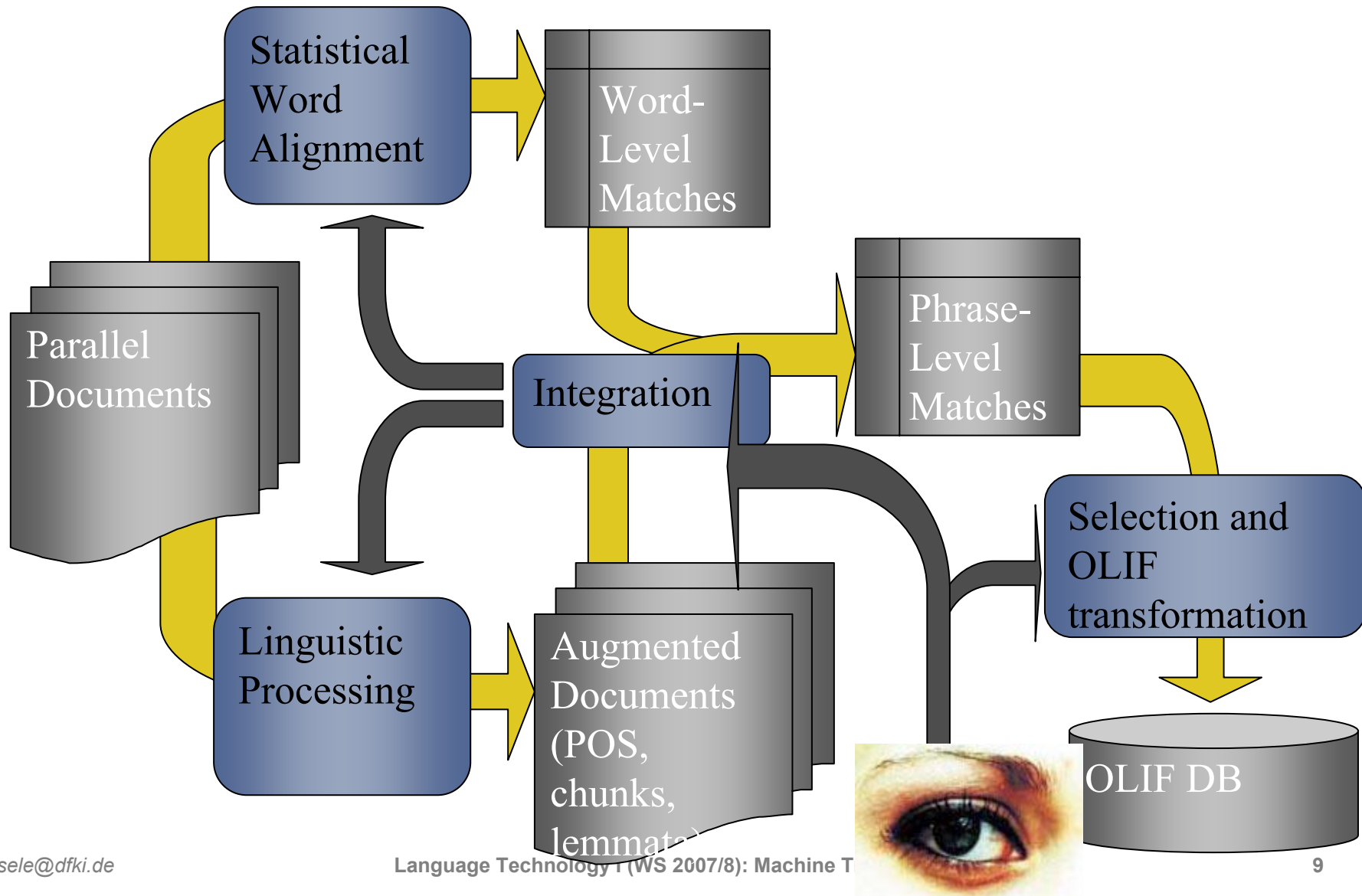
FR ↔ DE

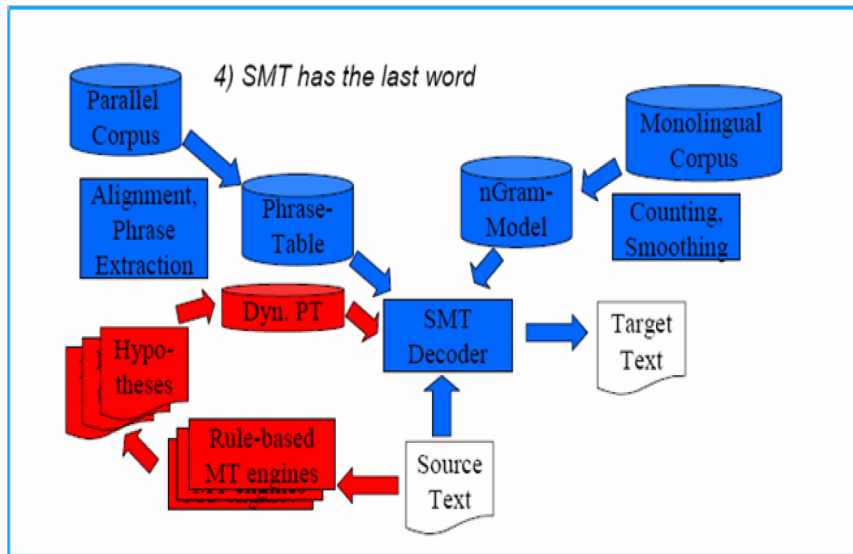
FR ↔ ES

Corpus-based Lexicon Extension for RBMT



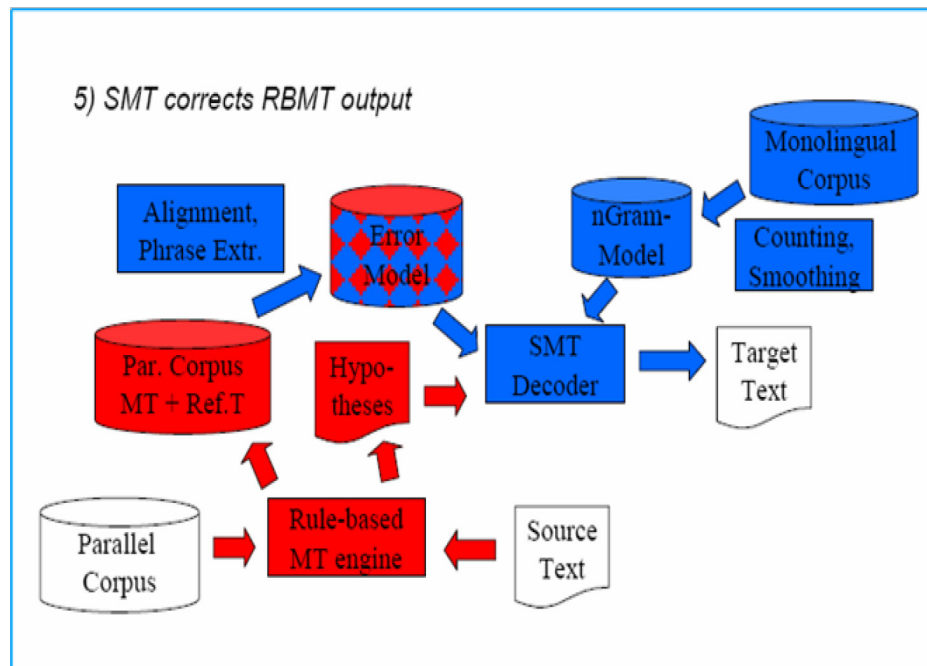
Terminology Extraction for MT: Architecture





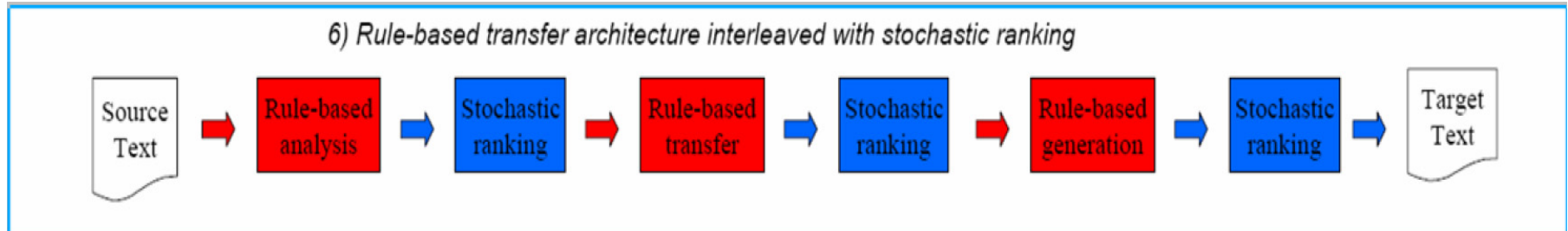
Motivation: SMT can only know what is in the training data,
RBMT systems often contain extensive lexical knowledge
BUT:

Architecture can fix lexical gaps, but will not overcome
problems with syntactically ill-formed candidates



Motivation: Errors in RBMT can be systematic/regular, may be fixed automatically. Target language model helps to find most natural wording in context

BUT: Sometimes RBMT messes a sentence completely up, no hope to repair these cases via SMT



Motivation: Fine-grained combination of statistical and linguistic evidence on all levels requires a closely coupled implementation

BUT:

- Chain can only be as good as the weakest link
- Difficult to avoid mismatches between representations when hand-crafting grammars
- Many existing processing components are designed for deterministic processing; building up forests of alternative solutions may require redesign of algorithms

Two types of MT evaluation

- Human („subjective“)
- Automatic („objective“)

The evaluation dilemma:

- Manual evaluation is meaningful, but expensive, tedious, and error-prone, not useful for regression testing
- Automatic evaluation is repeatable, objective, but not necessarily relevant; better systems may have worse scores

We need to

- ☐ lower the effort for manual evaluation,
- ☐ increase the quality of automatic evaluation,
- ☐ or do both

Highlights from Exercises

