

Anna Sgvall Hein, Institutionen fr lingvistik, Uppsala universitet
Rosenbad/2001-08-24

Automatisk versttning och versttningshjlpmedel

Klassiska problem med maskinöversättning

- **orealistiska förväntningar**
- **dåliga översättningar**
- **svårigheter att integrera maskinöversättning i arbetsflödet**

Vad kan man vinna genom automatisk versttning?

- **Snkta kostnader**
- **Snabbare produktion**
- **Konsekventare och mer exakta versttningar**

Tre exempel på etablerade maskinöversättningsprodukter på den svenska marknaden

- **Power Translator Pro (Lernout & Hauspie)**
 - engelska, franska, tyska, italienska, spanska och portugisiska
- **Systran**
 - engelska, franska, tyska, italienska, spanska, portugisiska, kinesiska, japanska, koreanska och ryska; grekiska och holländska under utveckling
- **Xplanation/LANT**
 - engelska, tyska, franska, spanska, holländska

Dock finns ingen lösning för högkvalitativ maskinöversättning till eller från svenska på marknaden idag.

Konstruktiva utgngspunkter i dag och fr framtiden

- **samspel mellan mnniskan och datorn**
- **bttre system genom utveckling och forskning**
- **sprkgranskning av klltexten**
- **kvalitetskontroll av mltexten**
- **realistiska frvntningar**
- **systematik fr evaluering**
- **beaktande av hur maskinversttning kan integreras i
arbetsfldet**

Datoriserade sprkstd fr versttare

On-line

- **textbaserade lexikon och termdatabaser**
- **versttningsminnen**
- **maskinversttningsystem**
- **verktyg fr kontroll av kllsprket**
- **verktyg fr kontroll av mlsprket**

Off-line

- **verktyg fr att bygga lexikon och minnen frn text**
- **verktyg fr att anpassa och vidareutveckla maskinversttningsystem**

Hur utveckla maskinöversättningen?

- återanvändning av översättningar
 - bättre språkliga resurser
 - lexikon, grammatiker, översättningsregler
 - nya översättningsstrategier
- kombinerade översättningsstrategier
- bättre kunskap om kontrastiva förhållanden
- beaktande av dokumentstruktur
- översättning över meningsgränserna

Vad kan man utvinna ur versatt text?

- **Rmaterial fr versttningslexikon**
 - **Ord och fraser**
- **versttningsexempel**
 - **Ord, fraser, meningar, stycken**
- **Material fr exempelbaserad maskinversttning**
 - **versttningsfragment av olika slag**
- **Material fr statistiskt baserad maskinversttning**
 - **versttningsfragment av olika slag**
- **Basresurs fr versttningsstudier**
 - **Ord, fraser, meningar**

Hur går det till?

- **Huvudsakligen statistiskt baserad länkning på olika nivåer**
 - **Text**
 - **Stycke**
 - **Mening**
 - **Fras**
 - **Ord**

Kombinerade versttningsstrategier

- **Exempelbaserad versttning**
- **Transferbaserad versttning**

En svensk satsning på automatisk översättning och översättningshjälpmedel

- **Forskning och utveckling av ett övergripande system som bygger på kombinerade översättningsstrategier**
- **Forskning om maskinöversättning över meningsgränserna**
- **Uppbyggnad av parallellkorpusar för olika ämnesområden**
 - **Fortsatt forskning om länkning och evaluering**
 - **Korpusbaserad lexikonkonstruktion**
 - **Forskning om exempelbaserad maskinöversättning**
 - **Översättningsforskning grundad på parallellkorpusar**
- **Utveckling av metoder för evaluering av maskinöversättning**
- **Utveckling av strategier för kvalitetsgranskning**
- **Forskning kring maskinöversättning och arbetsflöden**

Finns det en svensk marknad fr maskinversttning?

Ja

Se: Marknadsunderskning. En studie inom projekt MATS (Methodology and Application of a Translation System) utfrd av Explicon ab 2001.

MATS stds av VINNOVA och bedrivs i samarbete mellan Uppsala universitet, Scania CV AB och Explicon ab.