

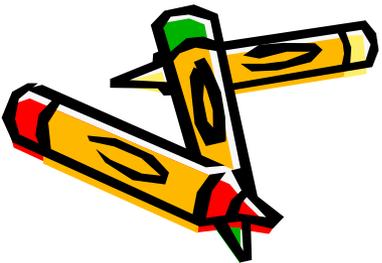
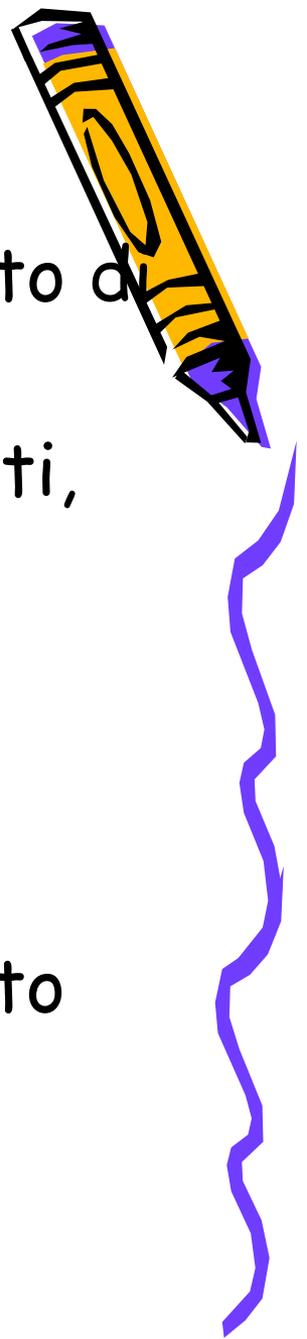
Memorie di traduzione e traduzione automatica

Fiorenza Mileto
fiorenza.mileto@unint.eu



TM e MT

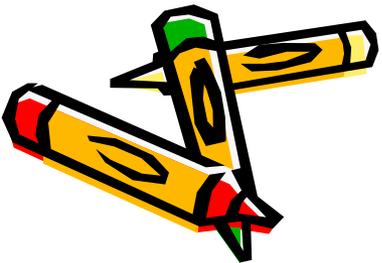
- Memorie di traduzione e Baseline come punto di partenza per la Machine Translation
 - Pulizia delle TM (segmentazione, duplicati, coerenza, terminologia, volumi)
 - Statistical MT vs Rule-based MT
 - Quando usare la machine translation?
- Test per implementazione della MT
 - Valutazione qualità/tempo/tipologia testo
- Traduzione vs post-editing



MT: obiettivi

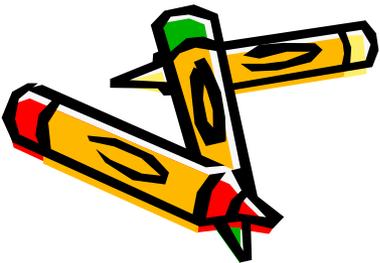
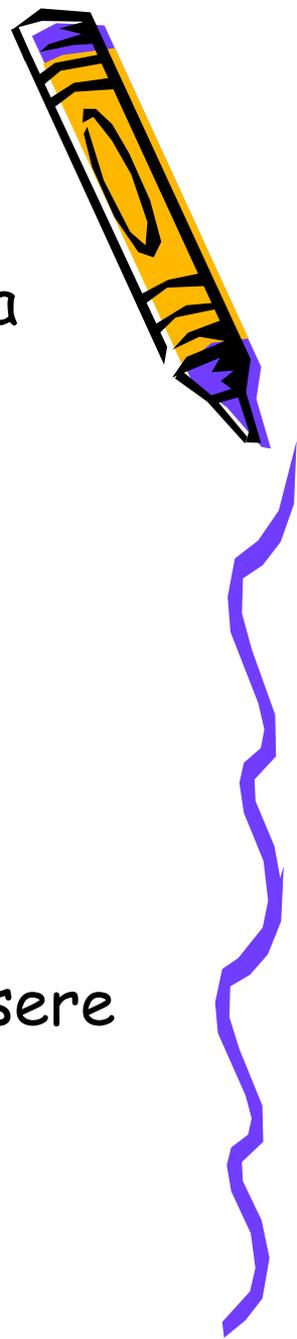


- La Machine Translation non è in grado di sostituire i traduttori
- Eseguire il post-editing delle traduzioni elaborate con machine translation è più rapido che tradurre da zero
- La Machine Translation aiuta a risparmiare tempo



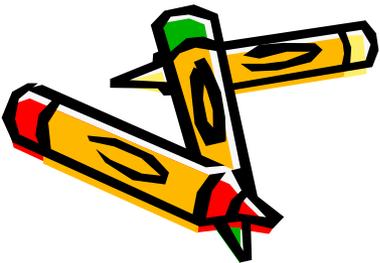
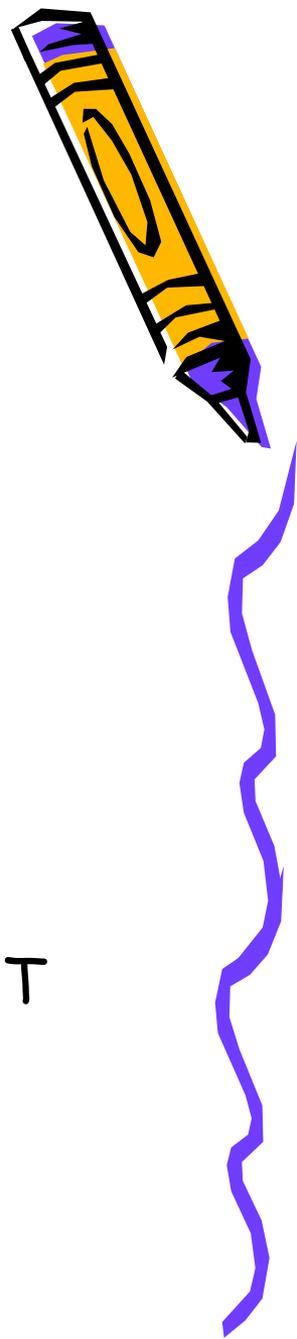
Condizioni ottimali per MT

- Materiale nuovo simile:
 - Se il nuovo materiale da tradurre somiglia a quello utilizzato per il training della MT, si otterranno risultati di MT migliori
- Qualità del source (risultati MT migliori):
 - Grammatica corretta
 - Segmentazione corretta
 - Ortografia corretta
 - Terminologia coerente
- La quantità di materiale per il training deve essere sufficiente e di buona qualità

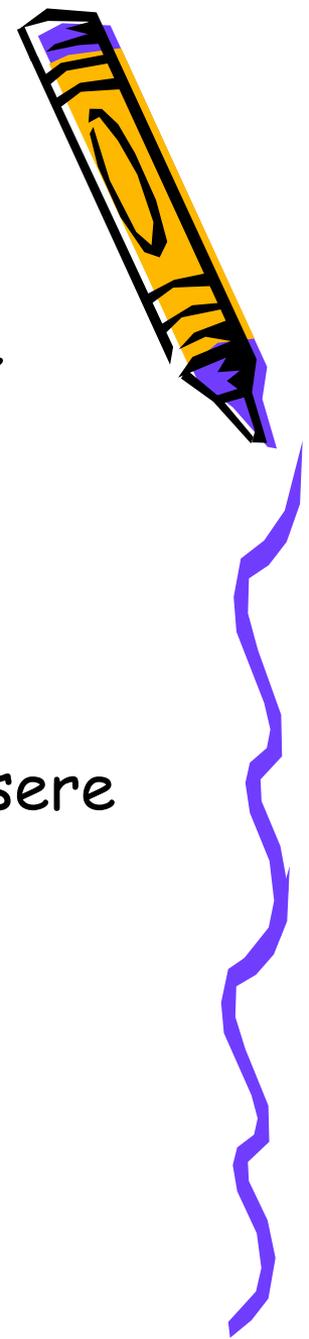


Traduzione vs post-editing

- Correggere fuzzy match e segmenti MT
- Requisiti:
 - applicare la terminologia corretta
 - seguire le istruzioni del cliente
 - la traduzione deve riflettere il source
 - correttezza grammaticale
- Qualità standard = publishable quality
- La costruzione della frase deve riflettere la MT per consentire un miglior training MT



Post-editing: alcuni consigli



- Leggere prima il source, poi iniziare a tradurre
- Non ignorare o cancellare l'output MT e non tradurre da zero
- Spesso alcune parti dell'output MT possono essere utilizzati velocizzando la traduzione
- No **over-edit** o **under-edit** l'output MT

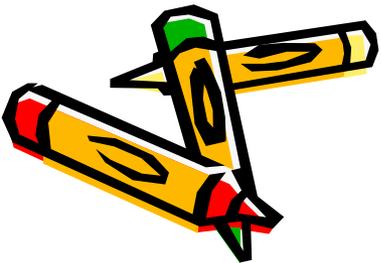


Alcuni problemi tipici: grammatica



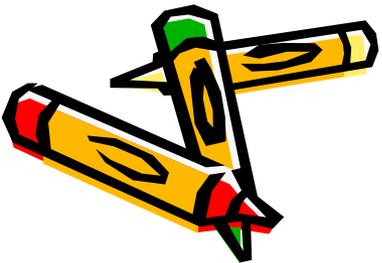
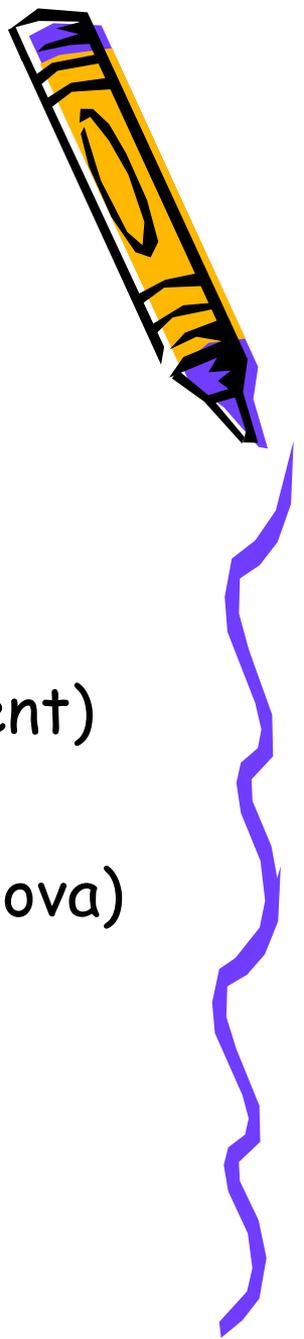
Lo stile dell'output della SMT è buono, MA presenta alcuni problemi:

- Parole extra nel target, non presenti nel source
- Parole mancanti nel target (aggettivi, verbi, prefissi)
- Mistranslation (opposti: remove/install, can/cannot, do/don't)
- Nomi composti nell'ordine errato
- Genere errato, mancata concordanza o flessione del verbo
- Verbi duplicati/mancanti/prefissi mancanti

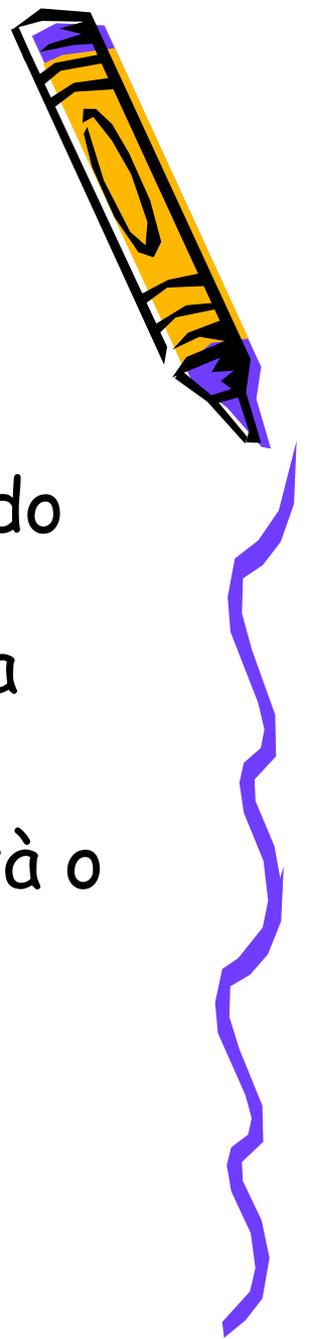


Alcuni problemi tipici: terminologia

- Terminologia che non rispetta il materiale di riferimento (non appropriata per il contesto)
- Terminologia non coerente (se la TM è inconsistente)
- Traduzione estratta da baseline (terminologia nuova)



Produttività



- Familiarizzare con post-editing e funzionamento del motore di MT: secondo gli standard di mercato la produttività giornaliera dovrebbe essere raddoppiata
- Non prendere in considerazione la qualità o l'impegno richiesto a ogni segmento: valutare la produttività nel complesso



Domande?

