

Fine-grained alignment, subsegment recall and match repair

Kevin Flanagan

1st February 2016



Match repair - simple example

Sentence to translate

=>

Connect the **green** cable to the probe.



S: Connect the **blue** cable to the probe.
T: Connectez le câble bleu à la sonde.

91%



EN	FR
blue	bleu
green	vert

<= Fragment translation source (FTS)
- termbase

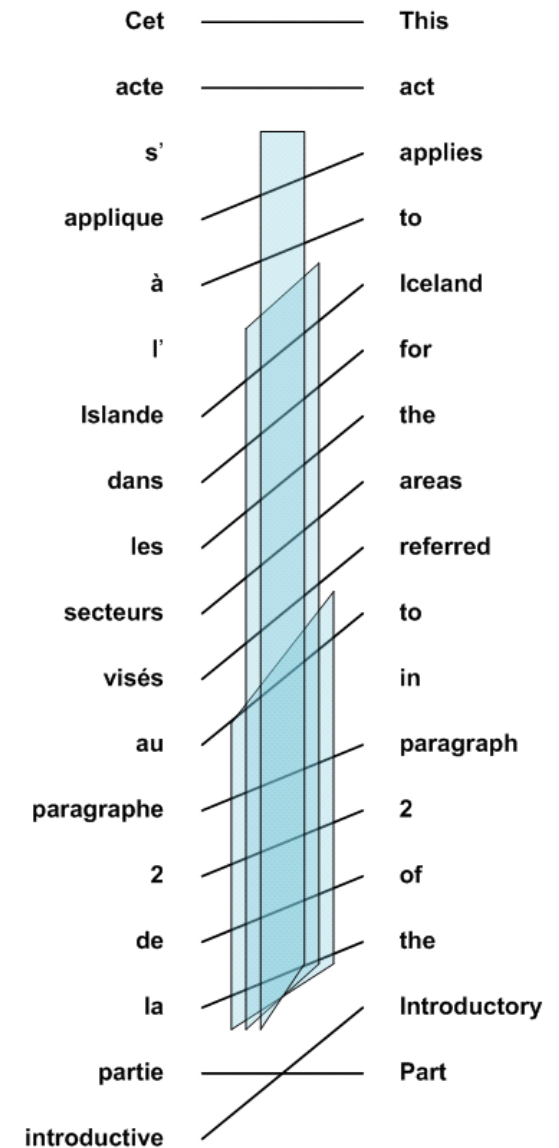
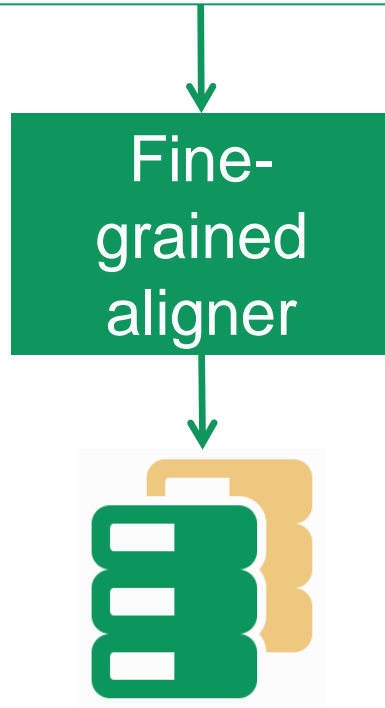
S: Connect the green cable to the probe.
T: Connectez le câble vert à la sonde.

Simple example – unpacked

- Fragment translation source (FTS) is used in two ways:
 - A. determines which word in the translation ('bleu') corresponds to the non-matching word ('blue'), i.e. creates a ***partial alignment***
 - B. provides a translation ('vert') to replace the non-matching target word ('bleu') thereby identified
- Possible FTSs: termbase, MT ... and subsegment recall
- Step A may fail using FTSs (unknown term; target differs e.g. "câble azur"; ...)
- If TM already has fine-grained alignment, FTS generally not needed for step A => more fuzzy matches can be repaired.

Fine-grained alignment¹

S: Cet acte s'applique à l'Islande dans les secteurs visés au paragraphe 2 de la partie introductive.
T: This act applies to Iceland for the areas referred to in paragraph 2 of the Introductory Part.



1. Further details and comparison with existing tools in: Flanagan, K. (2014), 'Filling in the gaps: what we need from subsegment TM recall', Proceedings of Translating and the Computer 36, London.

Types of repair

- Change (“Connect the bluegreen cable to the probe”)
 - requires (complete or partial) TU alignment
 - requires translation of new content
- Deletion (“Connect the ~~blue~~ cable to the probe”)
 - requires (complete or partial) TU alignment
- Insertion (“Connect the long blue cable to the probe”)
 - requires more-or-less complete TU alignment
 - requires translation of new content
- Movement (“Connect the ~~blue~~cable to the blue probe”)
 - requires more-or-less complete TU alignment

- Requirements can be met with FTSs (and fine-grained alignment, if available)
- Some sources possibly less reliable than others, e.g. new content generic MT translation may be less suitable than new content retrieved from termbase or TM.

Related literature

- Koehn, 2010: “Convergence of Translation Memory and Statistical Machine Translation”
 - TM word-aligned with GIZA++ (adding Europarl)
 - New/modified content translated contextually by MT.
- Ortega et al, 2014: “Using any machine translation source for fuzzy-match repair in a computer-aided translation setting”
 - TUs aligned on-the-fly using MT evidence
- N.B. differences from integrating TM data into MT decoder
 - Match repair need not require MT

Using subsegment recall in match repair

- Several existing match repair implementations use generic MT for match repair proposals (if no termbase entry etc.)
- Theory: proposals provided from subsegment recall should be more domain-specific and more suitable
- However, recall constraints (i.e. unlike MT, will sometimes return silence, as with TM in general) may inhibit repair
- So, try scoring match repair using subsegment recall vs. using MT, for different datasets.

Repair scoring

Query:

The Commission therefore concludes that the measures listed in recitals 209-232 are not compatible with the common market as rescue aid.

Reference translation:

La Commission conclut par conséquent que les mesures énumérées dans les considérants 209 à 232 sont incompatibles avec le marché commun en tant qu'aide au sauvetage.

Match:

The Commission therefore concludes that the measures ~~covered by this Decision~~ listed in recitals 209-232 are not compatible with the common market as rescue aid.

~~En conséquence,~~ la Commission conclut par conséquent que les mesures ~~couvertes par la présente décision~~ énumérées dans les considérants 209 à 232 ~~ne sont pas compatibles~~ incompatibles avec le marché commun en tant qu'aides au sauvetage.

Edit distance: 15 (stemmed) words

Repaired match:

~~En conséquence,~~ la Commission conclut par conséquent que les mesures **énumérées dans les considérants 209 à 232** ~~ne sont pas compatibles~~ incompatibles avec le marché commun en tant qu'aides au sauvetage.

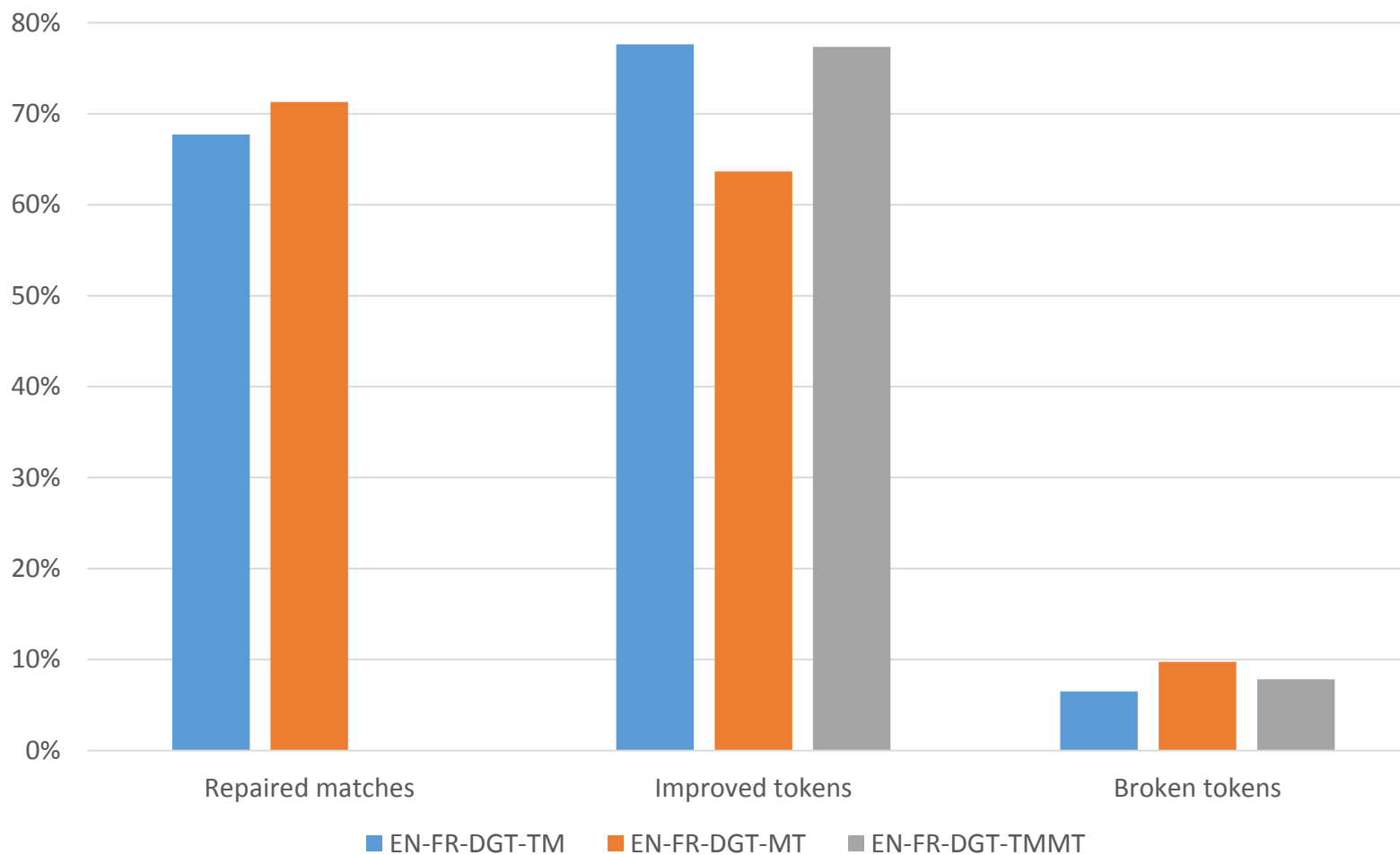
Edit distance: 8 (stemmed) words

Methodology

- Create de-duplicated test TMs with 10,000 TUs from DGT and SAP data, English-French and English-German; include fine-grained alignment
- Find all source segments with fuzzy matches scoring $\geq 70\%$
- For each, attempt match repair using subsegment recall, then generic MT, then both (favouring subsegment recall translations)
- Score repair results using target segment as reference translation

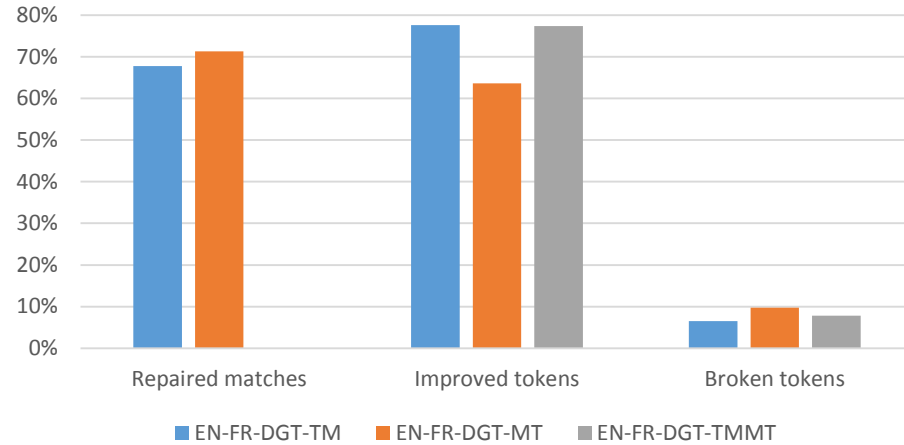
Results #1

Repair scoring results: DGT, en-fr

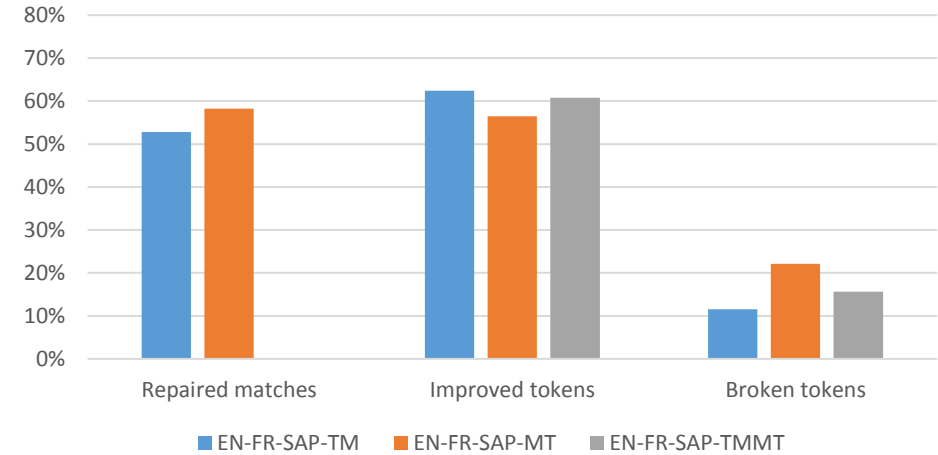


Results #2

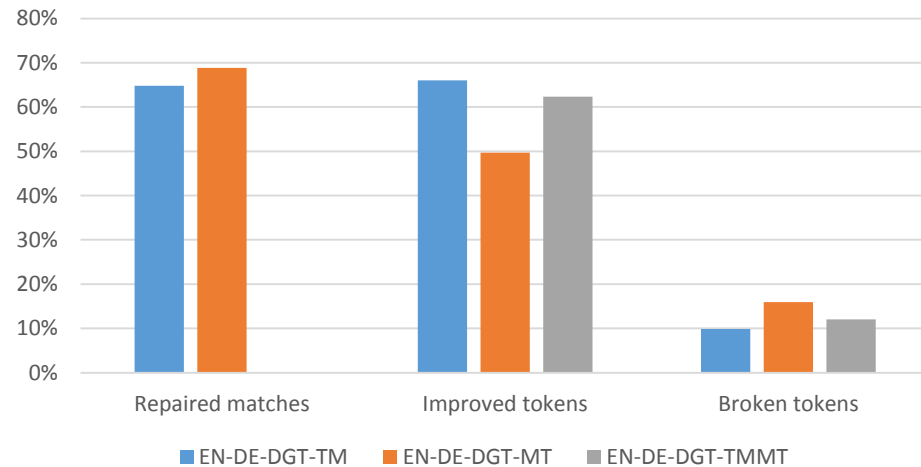
Repair scoring results: DGT, en-fr



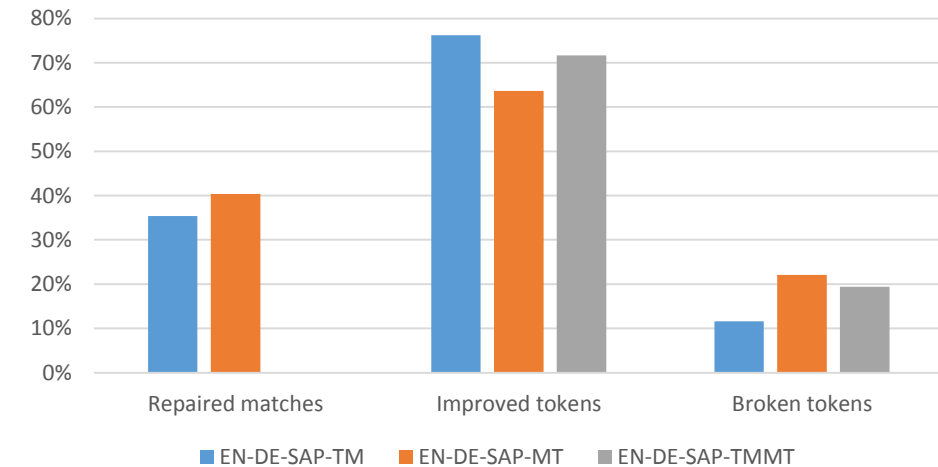
Repair scoring results: SAP, en-fr



Repair scoring results: DGT, en-de



Repair scoring results: SAP, en-de



Conclusions

- Subsegment recall alone – if comprehensive enough – may allow almost as much match repair as MT (and with better scores than generic MT)
- Introducing generic MT increases possible repairs, with relatively little reduction in scores
- Further tests required to compare with (for example) contextual MT
- UX considerations and cognitive impact
- Difficult to generate match repair performance statistics for other tools, for comparison.



Global Customer Experience Management

Copyright © 2008-2015 SDL plc. All rights reserved. All company names, brand names, trademarks, service marks, images and logos are the property of their respective owners.

This presentation and its content are SDL confidential unless otherwise specified, and may not be copied, used or distributed except as authorised by SDL.