

1.	This is a new sentence.	#Gābl·b ·t·gxp·lxgmxǵvx._\$	N/A
2.	This is a short new sentence.	#Gābl·b ·t·lahkm·gxp·lxgmxǵvx._\$	N/A
3.	This is a short wonderful sentence.	#Gābl·b ·t·lahkm·phgwxkÿnē·lxgmxǵvx.__\$	N/A
4.	This is a new sentence.	#Gābl·b ·t·gxp·lxgmxǵvx._\$	N/A
5.	Sentence 5 is similar to the next sentence.	#Fxgmxǵvx·5·bl· bfbētk·mh·māx·ǵxqm· lxgmxǵvx.__\$	N/A
6.	Sentence 6 is similar to the previous sentence.	#Fxgmxǵvx·6·bl· bfbētk·mh·māx·īkxôbhnl· lxǵmxgvx.__\$	N/A

# Pseudo-Translation

## (what will happen to my document)

Angelika Zerfaß

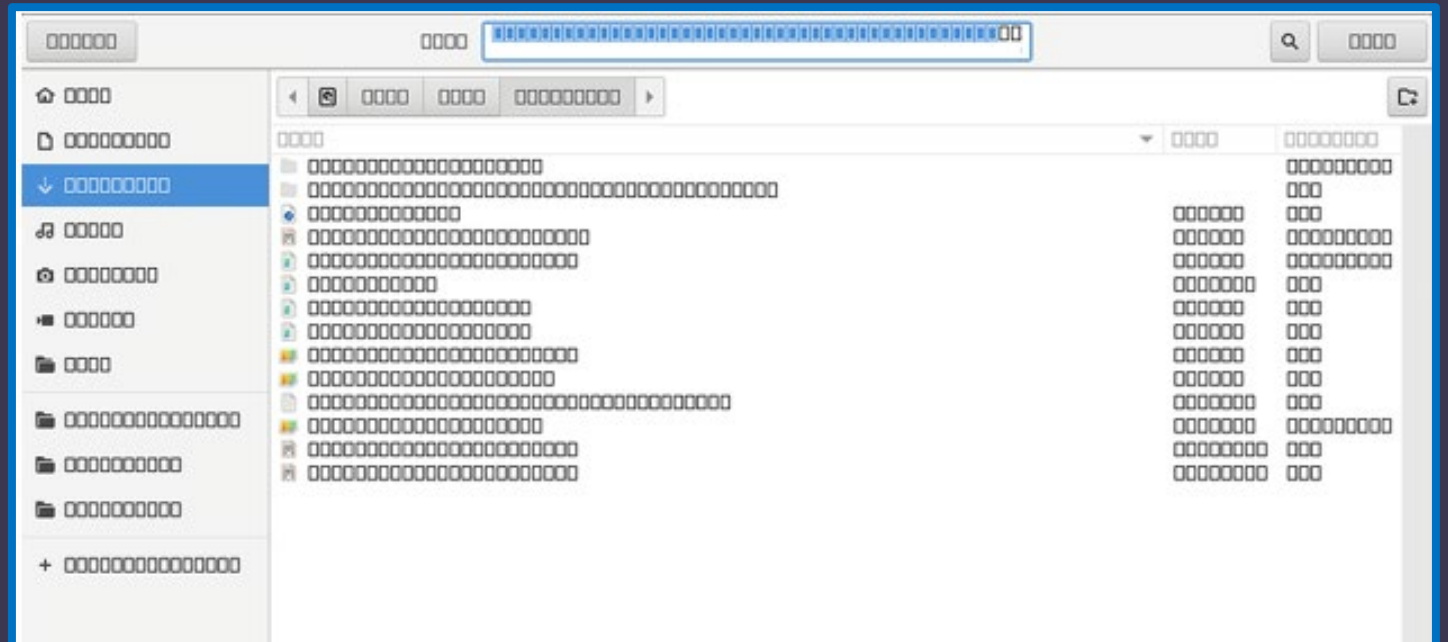
zaac

# Pseudo-Translation / Simulated Translation

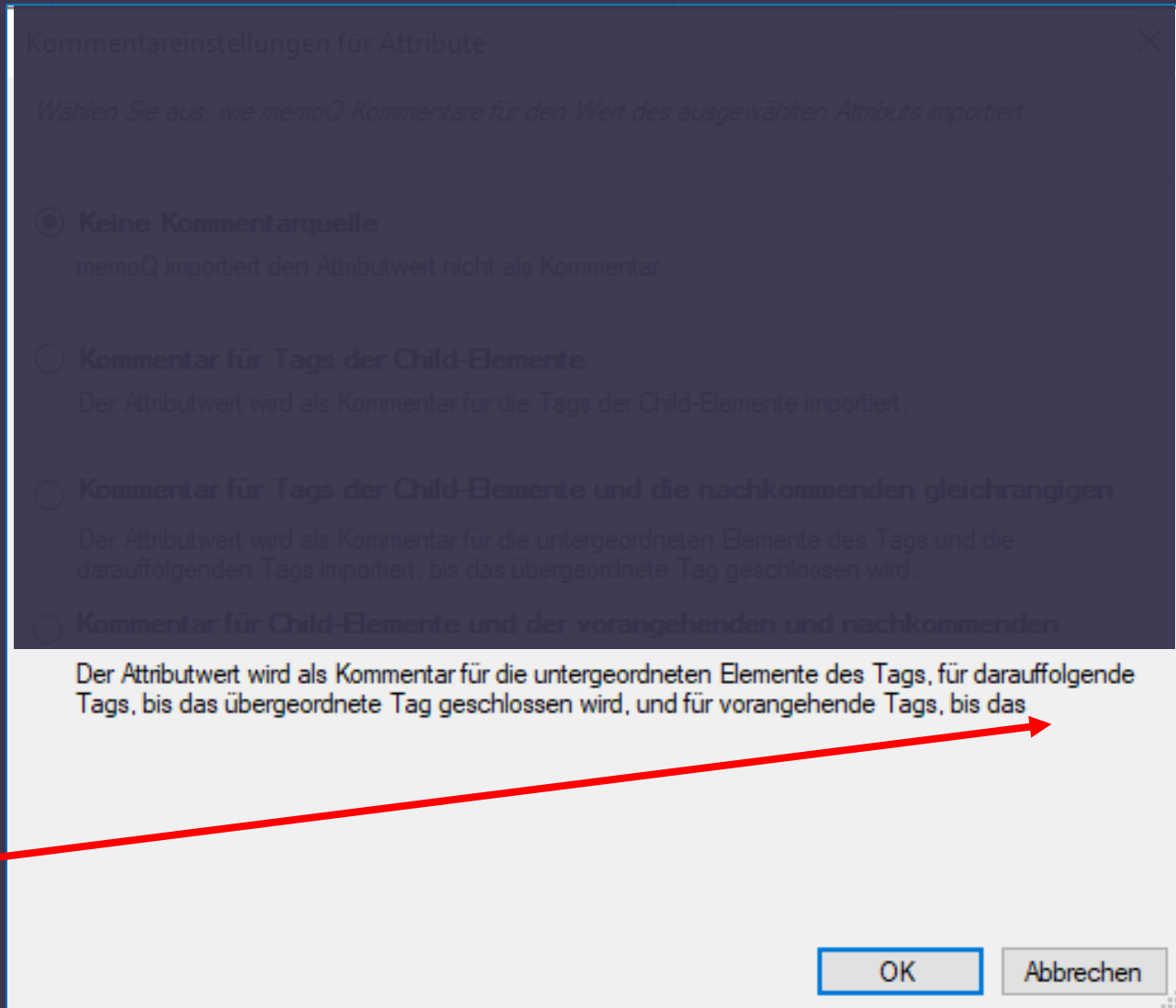
- Originally, pseudo-translation was a way to test how user interfaces of software products behave, once they get translated.
  - Can it show all the characters of the target language?
  - Will text still be visible even if it expands during translation?
  - Is the font size right for the target language?
  - Does the software crash because of the translation?
  - Is everything translated that is visible to the user?
  - Is there hard-coded text the translation cannot access?

# Examples from Software UI

Font used for the software UI cannot show the translated characters



# Examples from Software UI



Truncated text  
because the text field  
does not allow for  
more than 2 lines

# Pseudo-Translation

- Why do we need this in a tool that is not dedicated to software localization?
  - **Fonts** matter for documents, too!
  - There is not only translation, but also **layout work**.
    - Size of text boxes, text overlapping image, text moving to the next page...
    - From left-to-right to right-to left – not only text direction needs to change, but also layout of the page
  - **Default filters** might have to be adjusted
    - Keeping line spacing in InDesign files

# Pseudo-Translation

- Why do we need this in a tool that is not dedicated to software localization?
  - Custom import filters need to be checked whether they will really import **all translatable text**.
  - Documents might contain **text that cannot be extracted** for translation.
    - Text on images, text embedded in diagrams, text from documents that are only linked, but not embedded

# Character display

Text using the font Berlin Sans FB

Text using the font Comic Sans MS

Text using the font Microsoft Sans Serif

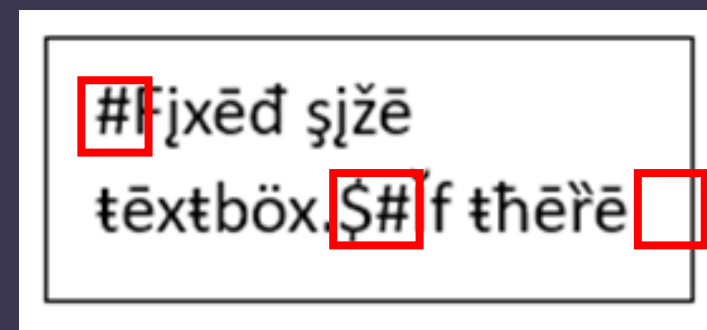
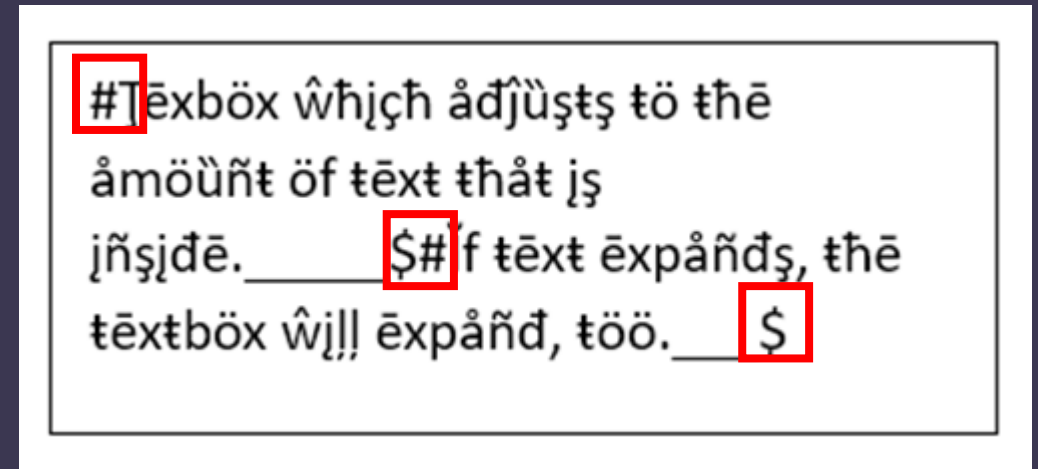
□□□ |

Ğİş |

Ğİş |

# Visible text

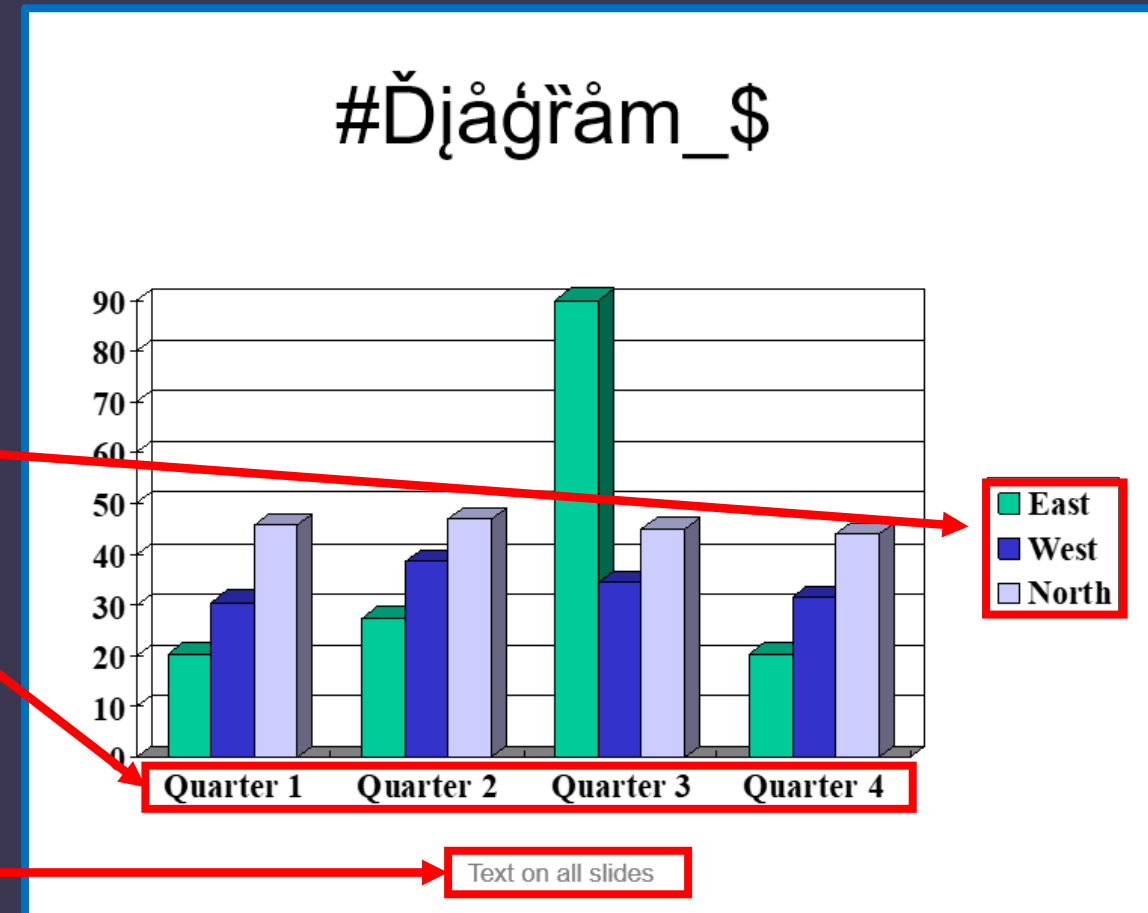
- In many editing environments (Word, InDesign...) a text box can have a fixed size or dynamically adjust to the amount of text inside.
- Pseudo-translation can add a character at the beginning and the end of each sentence.
- If the ending character of the pseudo-translation is not visible, the textbox is too small.



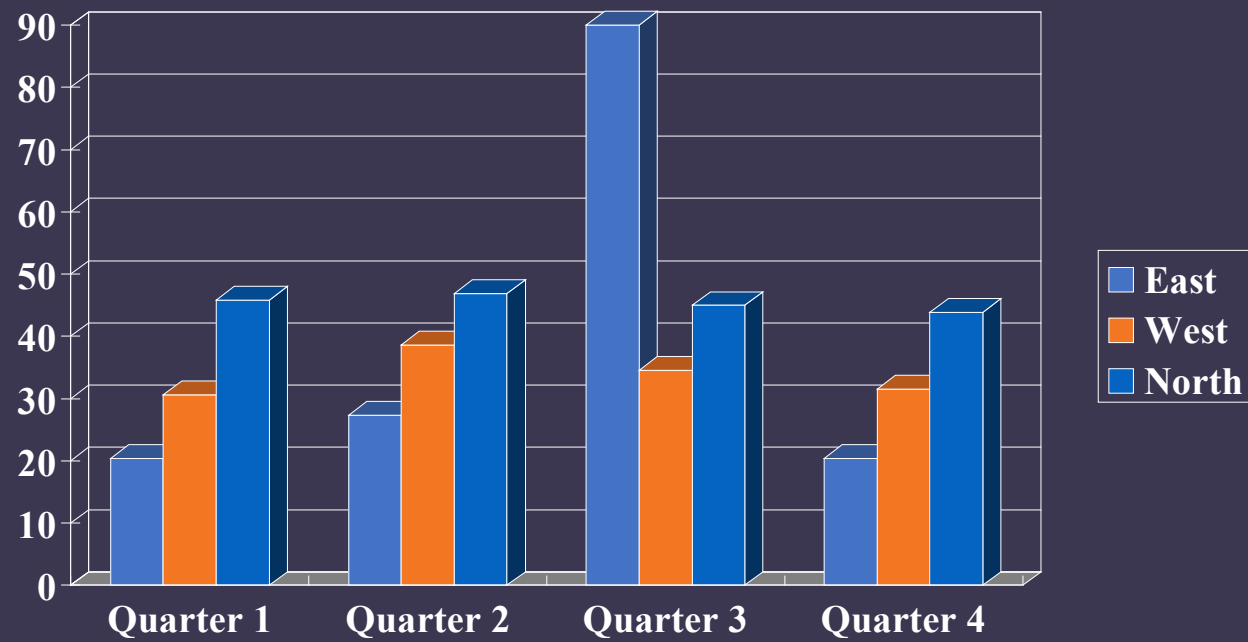


# Untranslatable Text

- Text that is not accessible when importing the file into memoQ, like text on images, text in diagrams or org charts created within the authoring tool itself (here PowerPoint).
- Text that was not imported because the respective option in the import filter was not activated (here text on master slide in the footer)



# #Djagrām\_



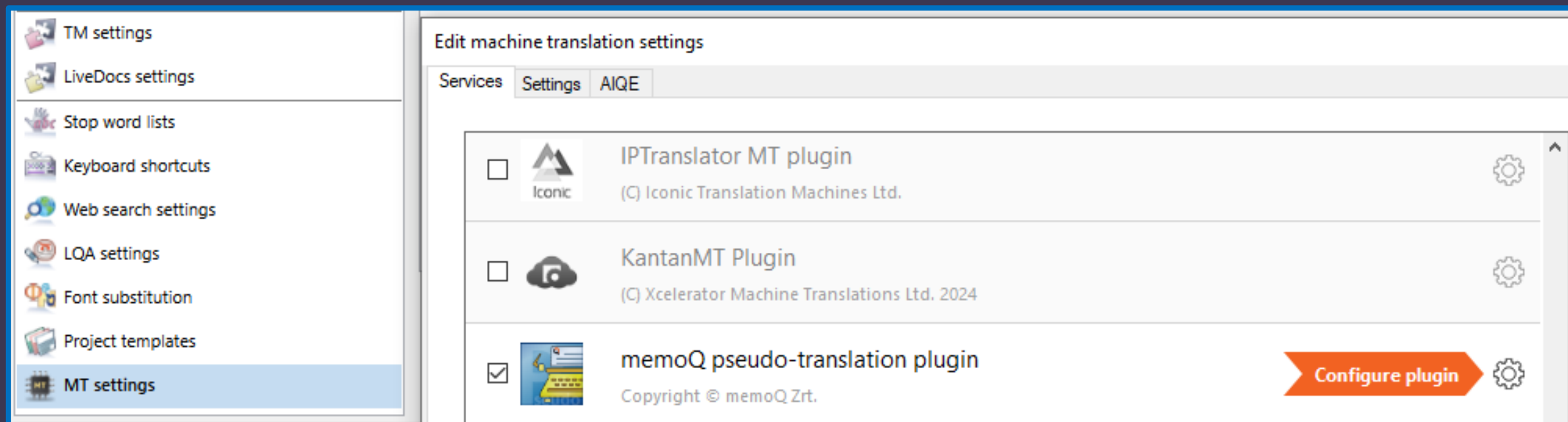


# Setting up Pseudo-Translation

---

# Pseudo-Translation Options

- MT Settings – memoQ pseudo-translation plugin

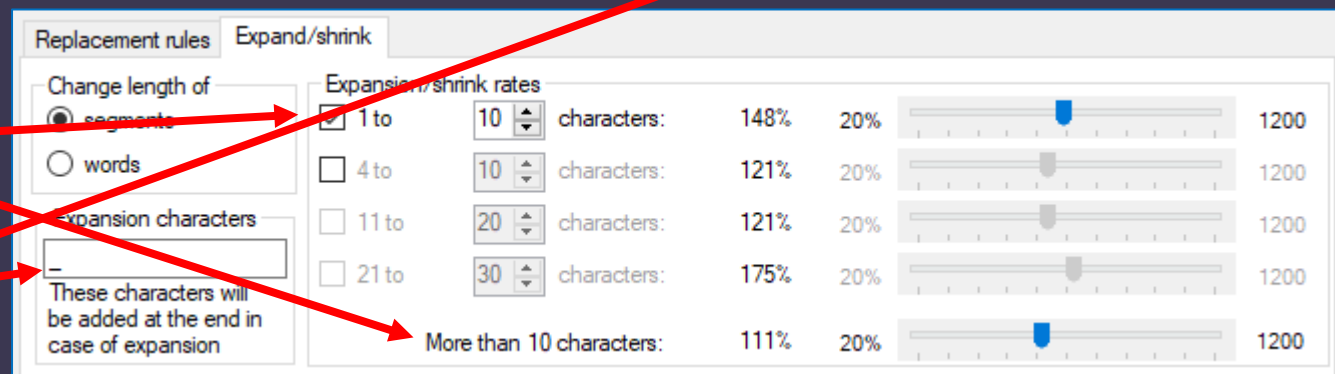
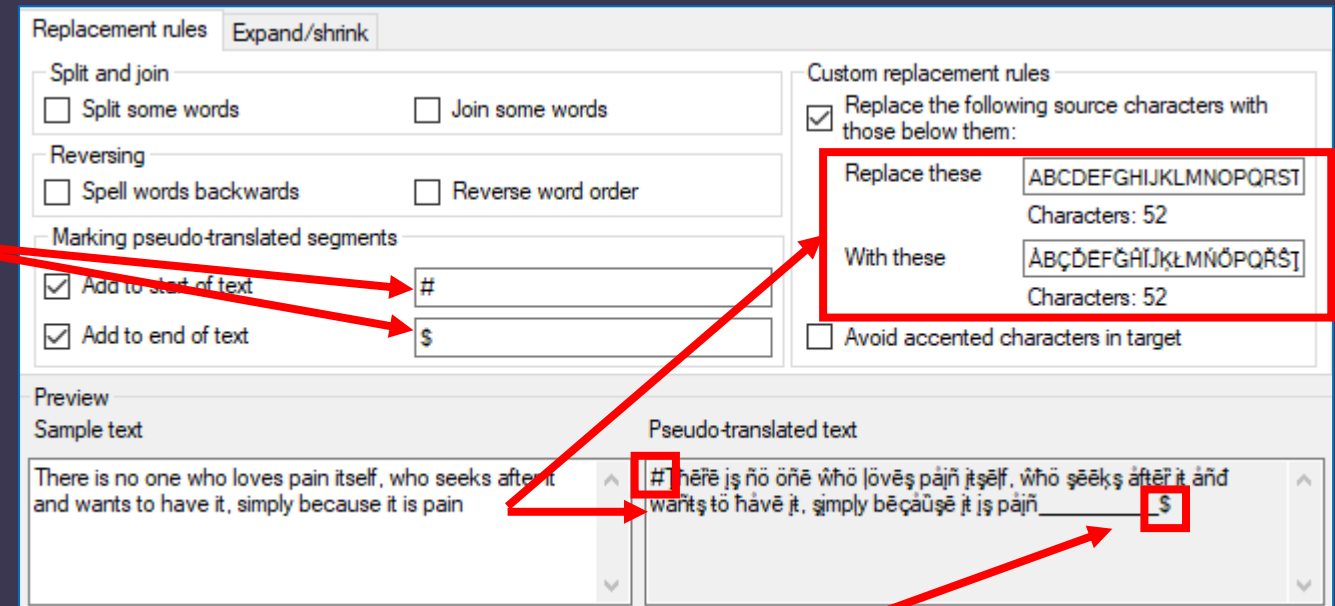


The screenshot displays the 'Edit machine translation settings' window. On the left is a sidebar menu with options: TM settings, LiveDocs settings, Stop word lists, Keyboard shortcuts, Web search settings, LQA settings, Font substitution, Project templates, and MT settings (highlighted). The main panel has tabs for 'Services', 'Settings', and 'AIQE'. Under the 'Settings' tab, three MT plugins are listed:

Plugin Name	Copyright	Status	Action
IPTranslator MT plugin	(C) Iconic Translation Machines Ltd.	<input type="checkbox"/>	Settings icon
KantanMT Plugin	(C) Xcelerator Machine Translations Ltd. 2024	<input type="checkbox"/>	Settings icon
memoQ pseudo-translation plugin	Copyright © memoQ Zrt.	<input checked="" type="checkbox"/>	Configure plugin button and Settings icon

# Default Pseudo-Translation Options

- Default Configuration
  - Add a symbol at the beginning (#) and at the end (\$) of the translation.
  - Change characters to accented characters.
  - Expand the translated segment relative to their original length.
  - Add an expansion indicator (\_) at the end.



# Text Expansion Tables

- Text expansion/shrinking depends on
  - Language pair
  - Type of text
  - Length of segments
  - ...

No. of characters in English source	Average expansion
Up to 10	200-300%
11-20	180-200%
21-30	160-180%
31-50	140-160%
51-70	151-170%
Over 70	130%

IBM, 1994 National Language Design Guide

Language	Translation	Ratio
Korean	조회	0.8
English	views	1
Chinese	次檢視	1.2
Portuguese	visualizações	2.6
French	consultations	2.6
German	-mal angesehen	2.8
Italian	visualizzazioni	3

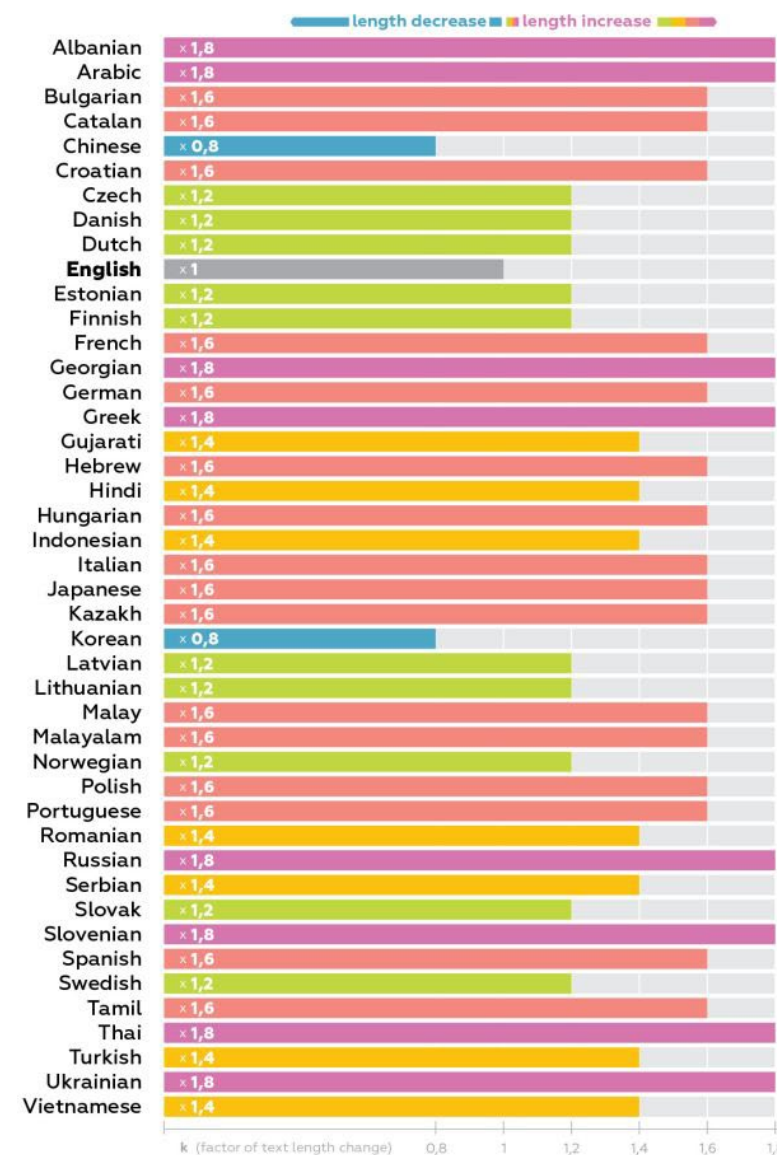
Richard Ishida, W3C

# Text Expansion Examples

Language	From English	To English
Arabic	+25%	-20% to -25%
Chinese	varies	varies
Danish	-10% to -15%	+10% to +15%
Finnish	-25% to -40%	+30% to +40%
French	+15% to +20%	-10% to -15%
German	+10% to +35%	-20% to -35%
Greek	+5% to +10%	-5% to -20%
Japanese	varies	+10% to +55%
Korean	-10% to -15%	+10% to +20%
Norwegian	-10% to -15%	+10% to +15%
Portuguese	+15% to +30%	-5% to -15%
Spanish	+20% to +25%	-10% to -20%
Swedish	-10% to -15%	+10% to +15%

## Change in text length after translation from English

Length of translated text = length of English source × k



# Tip for Multilingual Projects and different scripts

**Pseudo-translation plugin settings**

Configurations

- en-el
- en-jp

Configuration settings

Name: en-el

Use this as the fallback

Transformation

Source script: Latin

Target script: Greek

Assigned language pairs

- English (United States) to Greek

Replacement rules

Expand/shrink

- Split some words
- Join some words
- Spell words backwards
- Reverse word order
- Add to start of text (#)
- Add to end of text (\$)

Custom replacement rules

- Replace the following source characters with those below them:
- Replace these: Characters: 0
- With these: Characters: 0
- Avoid accented characters in target

Preview

Sample text: There is no one who loves pain itself, who seeks after it and wants to have it, simply because it is pain

Pseudo-translated text: #Κέρόβ ζη λμ μλβ τεμ ίμοβη νκζλ ζρηβιγ, τέμ ηββθη χγρόβο ζρ κλά τκλόη ρμ εχσβ ζρ, ηζκνίφ ψβωκρηβ ζρ ζη νκζλ \_\_\_\_\_ \$

**Pseudo-translation plugin settings**

Configurations

- en-el
- en-jp

Configuration settings

Name: en-jp

Use this as the fallback

Transformation

Source script: Latin

Target script: Japanese

Assigned language pairs

- English (United States) to Japanese

Replacement rules

Expand/shrink

- Split some words
- Join some words
- Spell words backwards
- Reverse word order
- Add to start of text (#)
- Add to end of text (\$)

Custom replacement rules

- Replace the following source characters with those below them:
- Replace these: Characters: 0
- With these: Characters: 0
- Avoid accented characters in target

Preview

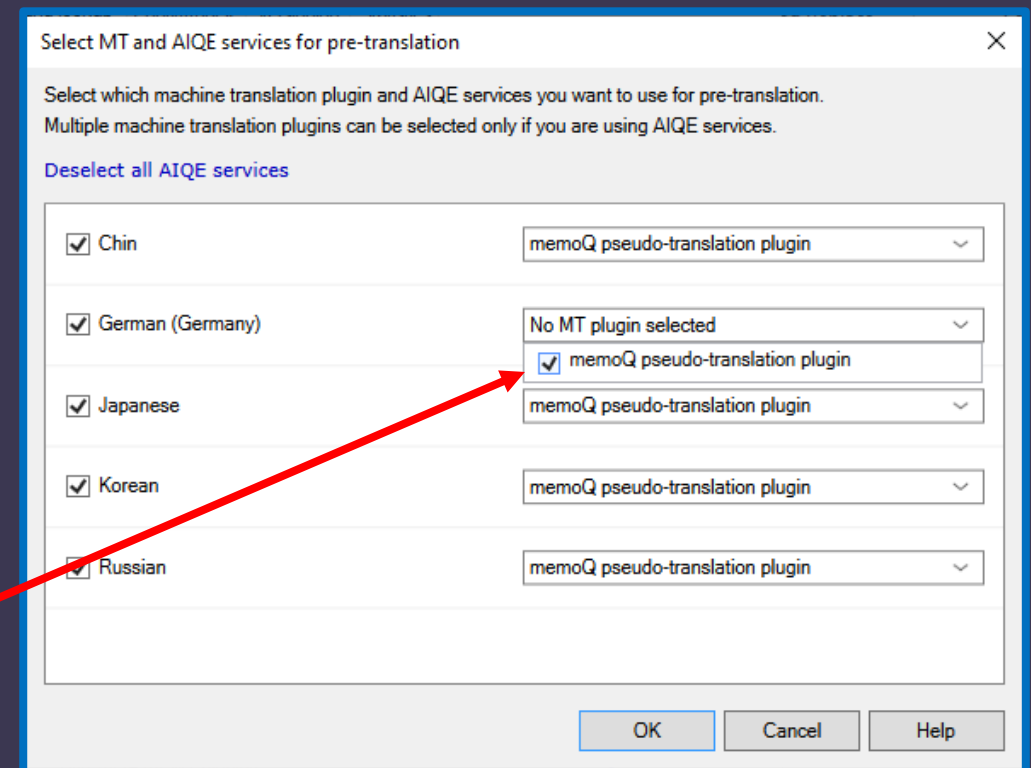
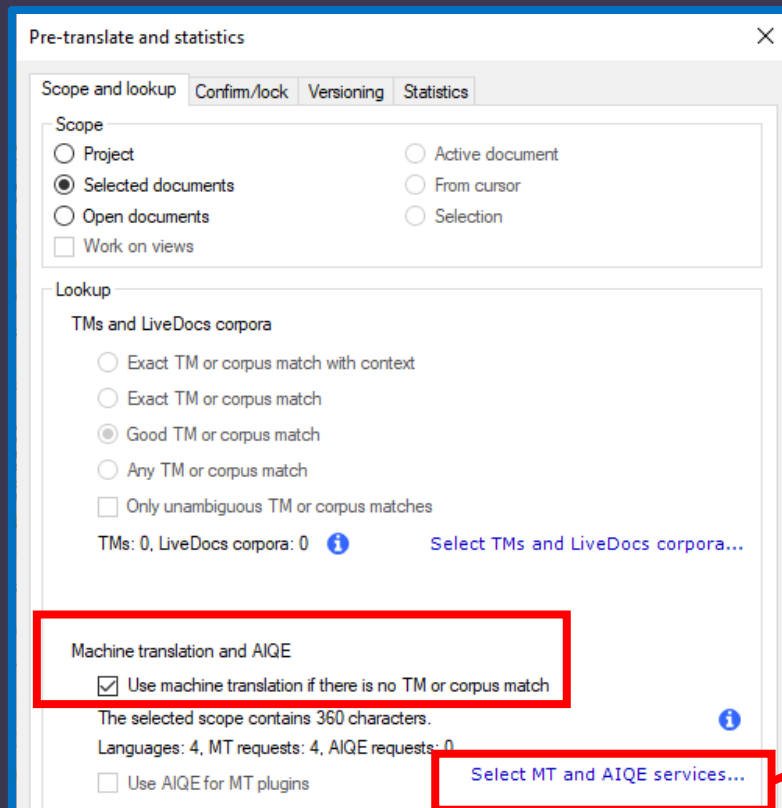
Sample text: There is no one who loves pain itself, who seeks after it and wants to have it, simply because it is pain

Pseudo-translated text: #開す間入間懸節の備備命の間等す備命屏命のつ間節の剣懸命懸(に節間屏肌等す備命節間間懸節剣肌(に間入懸(に剣懸れ等剣懸(に節に備す剣の間懸(に節懸てわ屏材妻間剣剣節間懸(に懸節の剣懸命



# Pseudo-Translation Options

- To run pseudo-translation, use the pre-translation dialog and use the MT settings to select the pseudo-translation option.



1.	This is a new sentence.	#Gābl·bl̥t·gxp·lx̥gm̥x̥ǧvx._\$	N/A
2.	This is a short new sentence.	#Gābl·bl̥t·lahkm·gxp·lx̥gm̥x̥ǧvx._\$	N/A
3.	This is a short wonderful sentence.	#Gābl·bl̥t·lahkm·ph̥gw̥x̥k̥ŷnē·lx̥gm̥x̥ǧvx.__\$	N/A
4.	This is a new sentence.	#Gābl·bl̥t·gxp·lx̥gm̥x̥ǧvx._\$	N/A
5.	Sentence 5 is similar to the next sentence.	#F̥x̥gm̥x̥ǧvx·5·bl̥l̥b̥fb̥ētk·mh̥·māx̥·ǧ̥x̥qm̥·	N/A

6.	Sentence	1. This is a new sentence.	#Κέζπ·ζπ·χ·λβτ·πβλρβλωβ._\$	N/A
		2. This is a short new sentence.	#Κέζπ·ζπ·χ·πεμόρ·λβτ·πβλρβλωβ._\$	N/A
		3. This is a short wonderful sentence.	#Κέζπ·ζπ·χ·πεμόρ·τμλάβογσι·πβλρβλωβ.__\$	N/A
		4. This is a new sentence.	#Κέζπ·ζπ·χ·λβτ·πβλρβλωβ._\$	N/A
		5. Sentence 5 is similar to the next sentence.	#Ιβλρβλωβ·5·ζπ·πζκζιχό·όμ·ρέβ·λβύρ·	N/A

6.	Sentence	1. This is a new sentence.	#閱す憩節憩節剣ぬ間等節間ぬに間ぬ劑間\$	N/A
		2. This is a short new sentence.	#閱す憩節憩節剣節す齡入にぬ間等節間ぬに間ぬ劑間_\$	N/A
		3. This is a short wonderful sentence.	#閱す憩節憩節剣節す齡入に等齡ぬれ間入肌創扉節間ぬに間ぬ劑間_\$	N/A

4.	This is a r	1. This is a new sentence.	\$_.عؤچعوطع.ؤعد·ص·كث·كث·#	N/A
5.	Sentence	2. This is a short new sentence.	\$_.عؤچعوطع.ؤعد·تقتج·ص·كث·كث·#	N/A
		3. This is a short wonderful sentence.	\$_.عؤچعوطع.ؤعد·ذبؤطعتغخن·عؤچعوطع.ؤعد·#	N/A
6.	Sentence	4. This is a new sentence.	\$_.عؤچعوطع.ؤعد·ص·كث·كث·#	N/A
		5. Sentence 5 is similar to the next sentence.	#عؤچعوطع·5·كث·كث·أئتصت·جى·أقع·وعذج·عؤچعوطع·\$_.	N/A
		6. Sentence 6 is similar to the previous sentence.	#عؤچعوطع·6·كث·كث·أئتصت·جى·أقع·إتعخ·أخت·عؤچعوطع.ؤعد·\$_.	N/A

# Special Use Cases

- Join words to simulate languages that use compound words (like German, Finnish or Dutch).
- This can be useful to see if hyphenation might be needed later.

Source	Target
1. This·is·a·new·sentence.	Gābl·b ·t·gxp·lxǵmxǵvx.
2. This·is·a·short·new·sentence.	Gābl·b ·t·lahkm·gxp·lxǵmxǵvx.
3. This·is·a·short·wonderful·sentence.	Gābl·b ·t·lahkm·phǵwxkǵnē·lxǵmxǵvx.

Source	Target
1. This·is·a·new·sentence.	Gābl·b ·t·gxp·lxǵm·xǵvx.
2. This·is·a·short·new·sentence.	Gābl·b ·t·la·hkm·gxp·lxǵm·xǵvx.
3. This·is·a·short·wonderful·sentence.	Gābl·b ·t·la·hkm·phǵw·xkǵnē·lxǵmxǵvx.

# Special Use Cases

- Checking import filters
- By changing all text characters to dots or underscores, it is easier to spot untranslated text in the exported file.

Replacement rules **Expand/shrink**

Split and join

Split some words       Join some words

Reversing

Spell words backwards       Reverse word order

Marking pseudo-translated segments

Add to start of text      #

Add to end of text      \$

**Custom replacement rules**

Replace the following source characters with those below them:

Replace these  Characters: 52

With these  Characters: 52

Avoid accented characters in target

**Preview**

Sample text      Pseudo-translated text

There is no one who loves pain itself, who seeks after it and wants to have it, simply because it is pain

-----

-----

-----

```
--<sample>
  <title>_____</title>
  -<text>
    _____
    <emphasis>_____</emphasis>
    .
  </text>
  <instruction>4f</instruction>
  -<text>
    _____
    <emphasis>_____</emphasis>
    _____
    <emphasis>_____</emphasis>
    .
  </text>
  <instruction>5r</instruction>
  -<text>
    _____
    </text>
    <instruction>6m</instruction>
  -<text>
    _____ ( _____ ) _____
    </text>
  <button value="Cancel"/>
  <button value="OK"/>
</sample>
```

# Pseudo-Translation in Templates

- Create a dedicated pseudo-translation project
  - Import of a document automatically runs the pseudo-translation.
  - The translated file is exported automatically into the stored path.
  - Check the file and prepare as necessary.
  - Import into the real project.
- Create a template for this client...
  - Import of a document automatically runs the pseudo-translation.
  - The translated file gets exported.
  - A search&replace script deletes the pseudo-translation.
  - If preparation is necessary, re-import the file (after preparation)
  - If no preparation is necessary, use the initially imported file



# Thank you!

Any questions?

---

[my-email@address.com](mailto:my-email@address.com)