Effects of Word Alignment Visualization on Post-Editing Quality & Speed

Lane Schwartz   Isabel Lacruz   Tatyana Bystrova

University of Illinois at Urbana-Champaign
lanes@illinois.edu

Kent State University
ilacruz@kent.edu   tbystro2@kent.edu

2 November 2015

All code, scripts, data & analysis files for this paper are at
`https://github.com/dowobeha/MT_Summit_2015`

Introduction
○○○○

MT for Post-Editing
○○○○

Results
○○○

MT Summit XV

Lane Schwartz, Isabel Lacruz, Tatyana Bystrova

Introduction
●○○○

MT for Post-Editing
○○○○

Results
○○○

MT Summit XV

Lane Schwartz, Isabel Lacruz, Tatyana Bystrova

- Post-editing
- Productivity
- Quality
- Effort

Introduction
○●○○

MT for Post-Editing
○○○○

Results
○○○

MT Summit XV

Lane Schwartz, Isabel Lacruz, Tatyana Bystrova

- User interface design
- Human factors
- Translation studies
- Machine translation quality

Introduction
○○●○
MT Summit XV

MT for Post-Editing
○○○○

Results
○○○
Lane Schwartz, Isabel Lacruz, Tatyana Bystrova

PBMT produces word alignments

- Do visualized word alignments affect PE quality? or speed?
- Do such effects depend on MT quality? or language pair?

Introduction
○○○●
MT Summit XV

MT for Post-Editing
○○○○

Results
○○○

Lane Schwartz, Isabel Lacruz, Tatyana Bystrova

- Russian-English
- Two news articles (WMT 2014)
- 32 segments in Doc A: 3 to 35 words (mean length 17 words)
- 33 segments in Doc B: 9 to 55 words (mean length 23 words)

Introduction
○○○○
MT Summit XV

MT for Post-Editing
●○○○

Results
○○○
Lane Schwartz, Isabel Lacruz, Tatyana Bystrova

- Russian-English
- Translated with Moses (out-of-domain TM)
- Tested 6 PE participants (4 ru L1, 2 en L1)

Introduction
0000
MT Summit XV

MT for Post-Editing
0●00

Results
000
Lane Schwartz, Isabel Lacruz, Tatyana Bystrova

File

Германия первой в Европе вводит " неопределенный пол " при регистрации детей

germany is the first in europe enters an uncertain " " floor when registering children

Germany is the first in Europe to allow " uncertain gender " when registering children

Германия станет первой страной в Европе , где будет разрешено при регистрации детей , родившихся гермафродитами , вместо мужского или женского пола указывать " неопределенный " .

germany will be the first country in europe , where will be allowed in the registration of children born гермафродитами_germafroditami instead of male or female indicate " " uncertain .

Germany will be the first country in Europe , where " uncertain " will be allowed in the registration of children born with unclear gender , instead of male or female .

У родителей ребенка будет право на время оставить графу " пол " в свидетельстве о рождении незаполненной .

the parents of the child will be the right time to leave the earl " paul " in birth certificate незаполненной_nezapolnennoy .

The parents of the child will be allowed to leave the gender " uncertain " in the birth certificate registry .

Вместо " F " (женский пол) или " M " (мужской пол) родители могут на время записать " X " (неопределенный) .

instead of " f " (female gender) or " m " (male) parents can at a time of record " x " (uncertain) .

Instead of " f " (female gender) or " m " (male gender) , parents can initially record " x " (uncertain gender) .

Это решение направлено на то , чтобы не подталкивать родителей к спешке с операцией по " уточнению " пола .

this decision is aimed at the not to push parents to hurry with the operation to " " definition of sex .
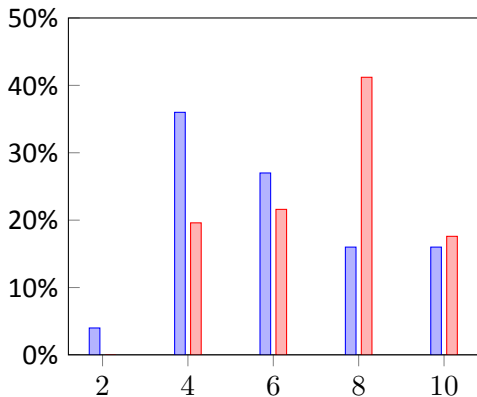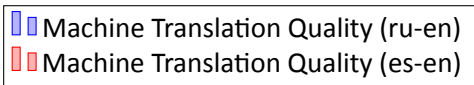
This decision is aimed to not push parents to hurry with gender clarification surgery .

Согласно статистике , один из 2 тысяч новорожденных в мире рождается гермафродитом .

according to the statistics , one in 2,000 babies in the world is born гермафродитом_germafroditom .

According to the statistics , one in 2,000 babies in the world is born with unclear gender .
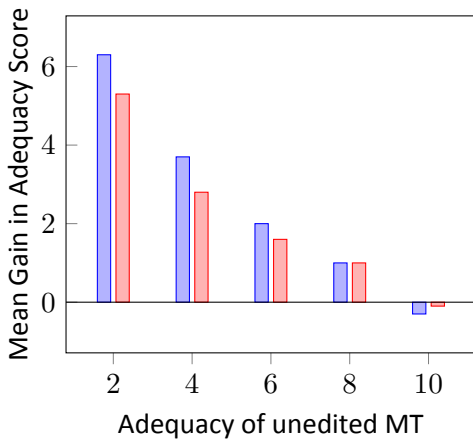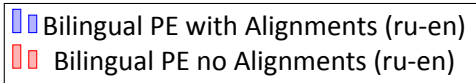
Introduction
○○○○

MT for Post-Editing
○○●○

Results
○○○

MT Summit XV   Lane Schwartz, Isabel Lacruz, Tatyana Bystrova

12	The post-edited translation is superior to the reference translation
10	The meaning of the source sentence is fully conveyed in the English translation
8	Most of the meaning of the source sentence is conveyed in the English translation
6	The English translation misunderstands the source sentence in a major way, or has many small mistakes
4	Very little information from the source sentence is conveyed in the English translation
2	The English translation makes no sense at all

Introduction
OOOO

MT for Post-Editing
OOO●

Results
OOO

MT Summit XV

Lane Schwartz, Isabel Lacruz, Tatyana Bystrova

Introduction
○○○○

MT for Post-Editing
○○○○

Results
●○○

MT Summit XV

Lane Schwartz, Isabel Lacruz, Tatyana Bystrova

El BCE dio anoche un empujón a Atenas , con la ampliación de las líneas de liquidez de emergencia .

The ECB last night gave a boost to Athens , with the extension of the emergency liquidity lines .

The ECB last night gave a boost to Athens, with the extension of the emergency liquidity lines.

Introduction
○○○○

MT for Post-Editing
○○○○

Results
○●○

MT Summit XV

Lane Schwartz, Isabel Lacruz, Tatyana Bystrova

Bilingual PE with Alignments (ru-en)
Bilingual PE no Alignments (ru-en)

Mean Gain in Adequacy Score

Adequacy of unedited MT

Introduction
○○○○
MT Summit XV

MT for Post-Editing
○○○○

Results
○○●
Lane Schwartz, Isabel Lacruz, Tatyana Bystrova

- The Russian-English aligner produced an improvement in post-editing quality

- Will there be a similar effect for other language pairs?

- Investigated this question for Spanish-English

Experiment 2 – Spanish-English

- Two newspaper articles in Spanish
- Text 1 had 26 segments (401 words) and Text 2 had 25 segments (402 words)
- Segment length: Text 1: Avg. 15, range from 4 to 24 and Text 2: Avg. 16, range from 4 to 28
- Machine translated into English
- Tested 11 participants – one had to be dropped
- All Master of Spanish Translation students
- All L1 English
- Keystroke log was recorded
- Within-subjects design

Alignment Effects: Post-Editing Mean Rating Gain for Spanish
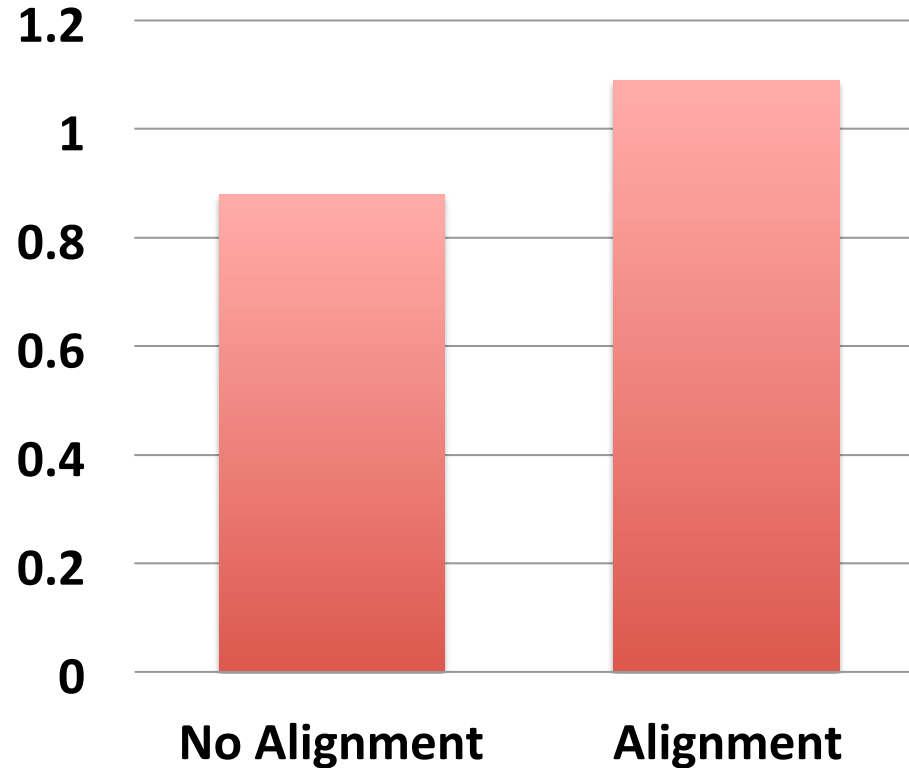
Discussion

- Overall, post-editing quality gain for Spanish-English aligned text was numerically higher than for non-aligned text.

- This difference was not statistically significant

- Does the Spanish-English aligner bring most benefit for low quality MT segments, as with Russian-English?

Mean Rating Gain as a Function of MT Quality

	MT 4	MT 6	MT 8	MT 10
■ No Alignment	1.88	2.15	0.25	-0.20
■ Alignment	3.03	1.91	0.36	-0.23
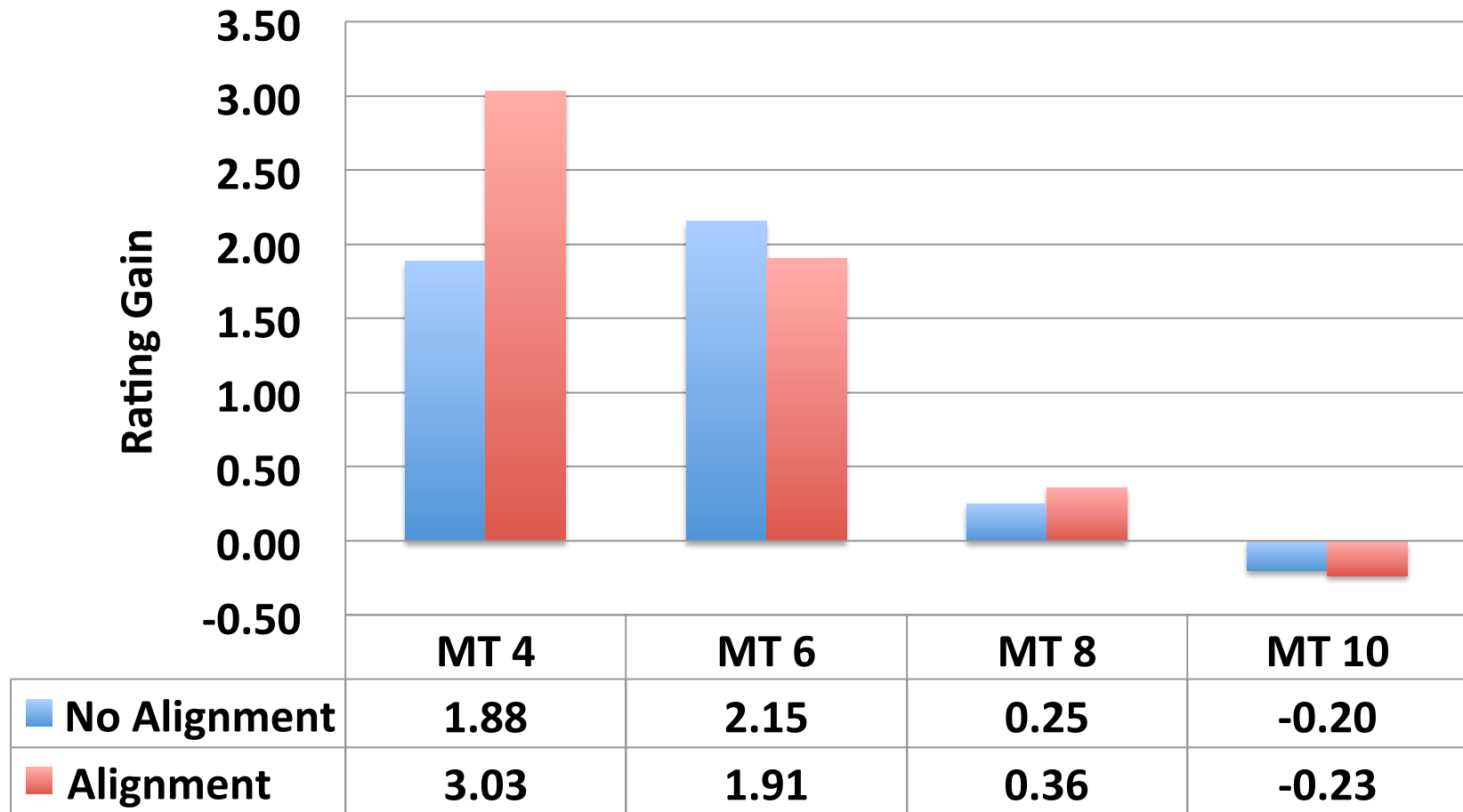
Discussion

- As with Russian-English, we find most benefit from the aligner when MT quality is low

- For Spanish-English, there is no clear benefit except for the lowest MT quality

- For Russian-English, there was a more gradual drop-off in benefit as MT quality improved

- This difference may be due to the greater similarity of English to Spanish

How about post-editing effort?

- Alignment results in increased quality of post-edited text, at least some of the time.

- Do post-editors expend more or less effort when they use the aligner?

Cognitive Effort

- The proportion of available mental resources, including memory, perception, evaluation, that are used to carry out a task (Russo and Dosher, 1983)

- The effort of thinking about actions to take
 - Can only be measured indirectly (Krings, 2001)

Pauses – Measuring Cognitive Effort

- Pauses in language production indicate cognitive effort
  - Can be measured automatically through keystroke logs (Shilperood, 2011; Krings, 2001; Shreve et al., 2011)
- Pause to Word Ratio, PWR, measures number of pauses per word
  - Correlates well with density of complete editing events (Lacruz & Shreve, 2014) and HTER

Effect of Alignment on Effort

- The keystroke logging data gathered with the Spanish-English texts allows the computation of Pause to Word Ratio (PWR)

- PWR is a measure of cognitive effort in post-editing

- Is it more or less effortful to post-edit with the aligner?

# Alignment Effects - Pause to Word Ratio
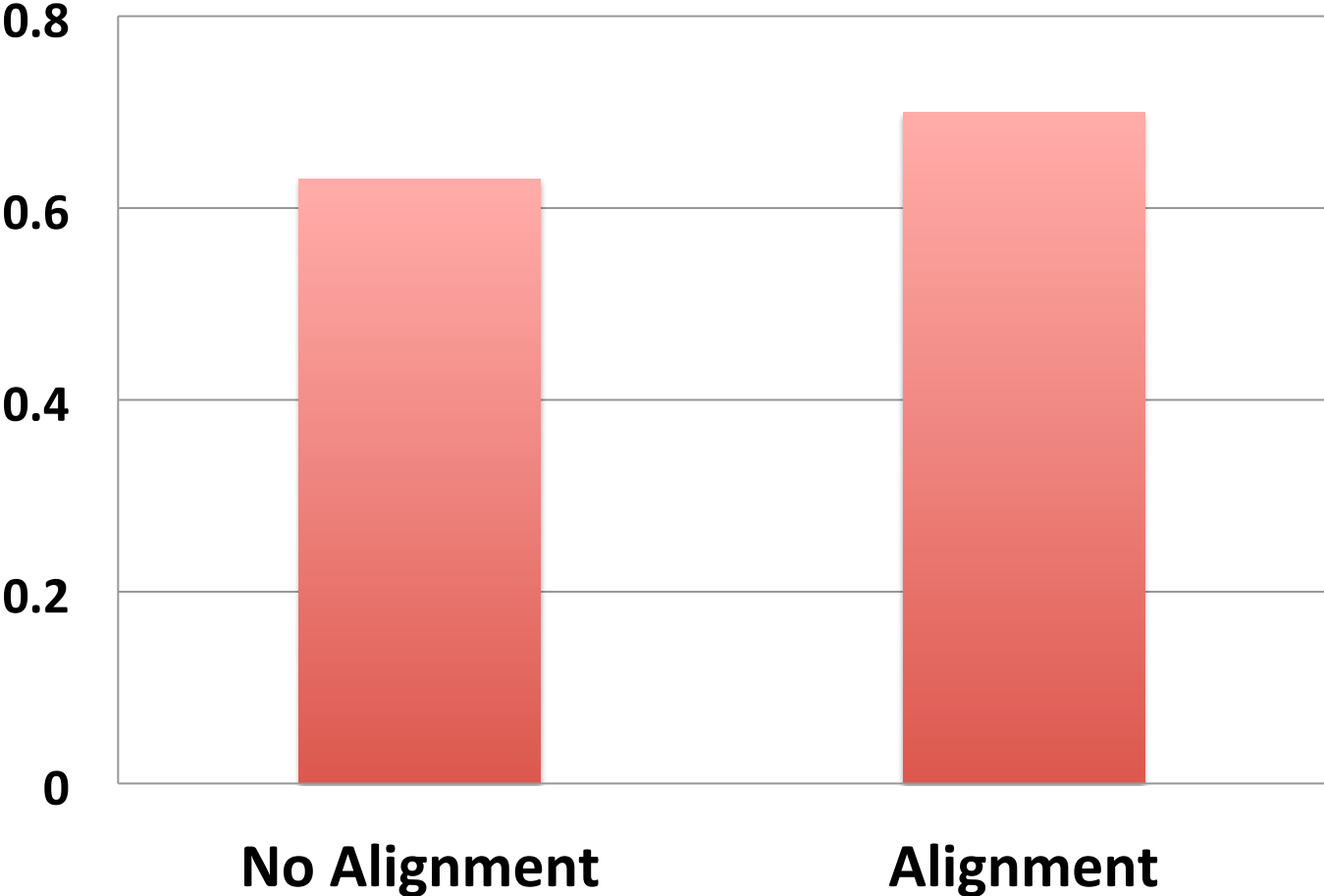
Discussion

- Overall, cognitive effort, as measured by PWR, increases slightly numerically when alignment is used

- This increase is not statistically significant

Further Discussion

- Intuitively, we expect the aligner should
  - increase post-editing quality
  - reduce post-editing effort
- If this is true, then Gain to Effort Ratio (GER)

  GER = (PE Rating – MT Rating)/PWR

  should be higher when the aligner is used

Alignment Effects - Gain to Effort Ratio
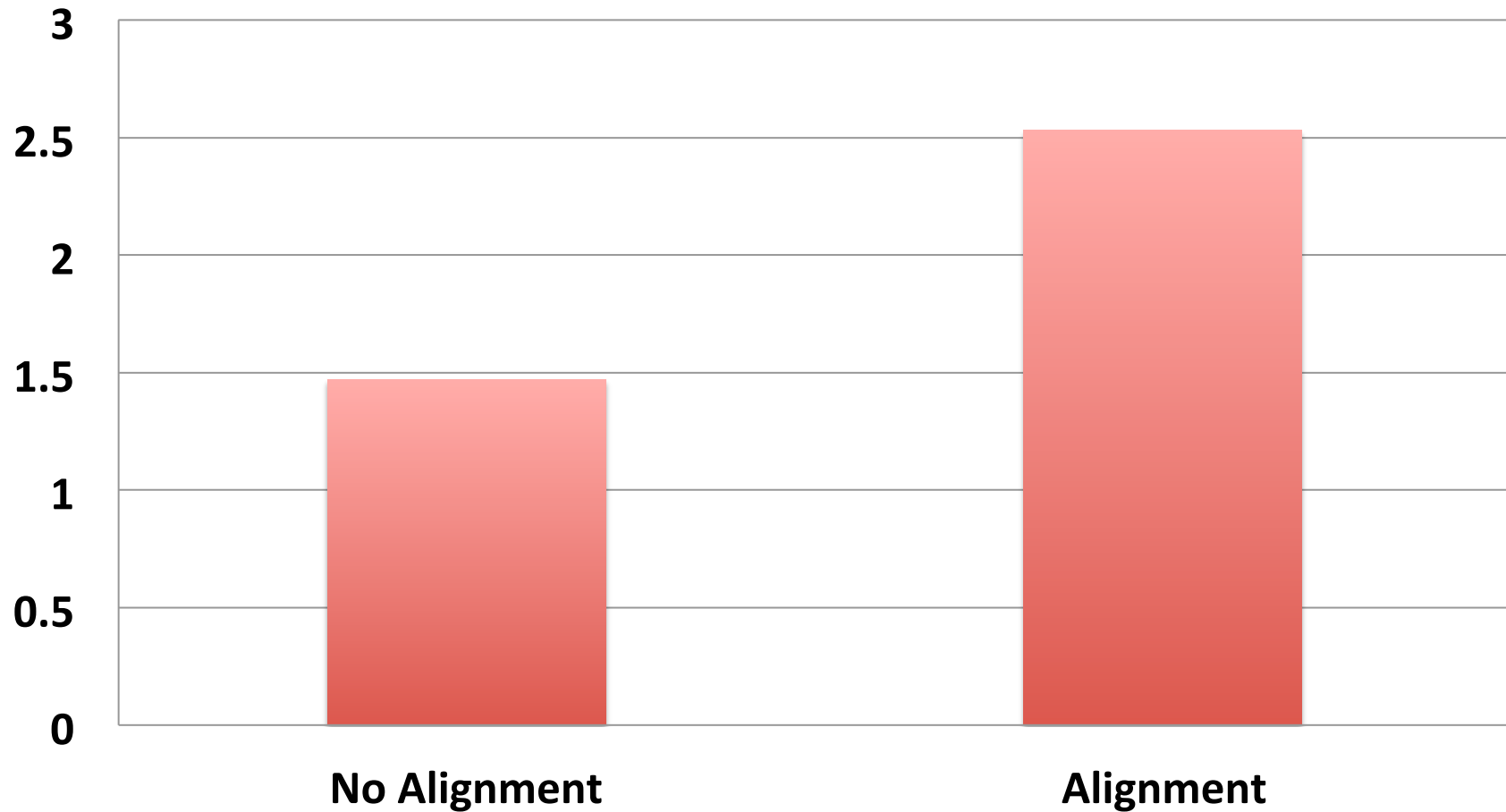
Discussion

- Overall, the Gain to Effort Ratio was significantly higher when post-editors used the aligner

- The effort exerted is more effective when the aligner is used

Summary: Spanish-English

- No significant difference in quality gain with and without the aligner - numerically there is a difference, most noticeable in the lower quality segments

- No significant difference in cognitive effort (PWR) with and without the aligner

- Significant increase in quality gain per effort when using the aligner