

# Peru is Multilingual, Its Machine Translation Should Be Too?

Spoiler:  
Yes

What is the main research question?

Do multilingual NMT systems outperform pairwise ones for four languages spoken in Peru?: **Aymara** (aym), **Ashaninka** (cni), **Quechua Chanka** (quy), **Shipibo-Konibo** (shp).

What are the challenges?

Low-resourcedness, sparse vocabulary, no relationship with Spanish, and domain mismatch between training and evaluation sets.

How were they addressed?

(i) Large pre-trained NMT model including Spanish, (ii) Multilingual segmentation model, (iii) Tagged back-translation, (iv) Sampling train data for fine-tuning.

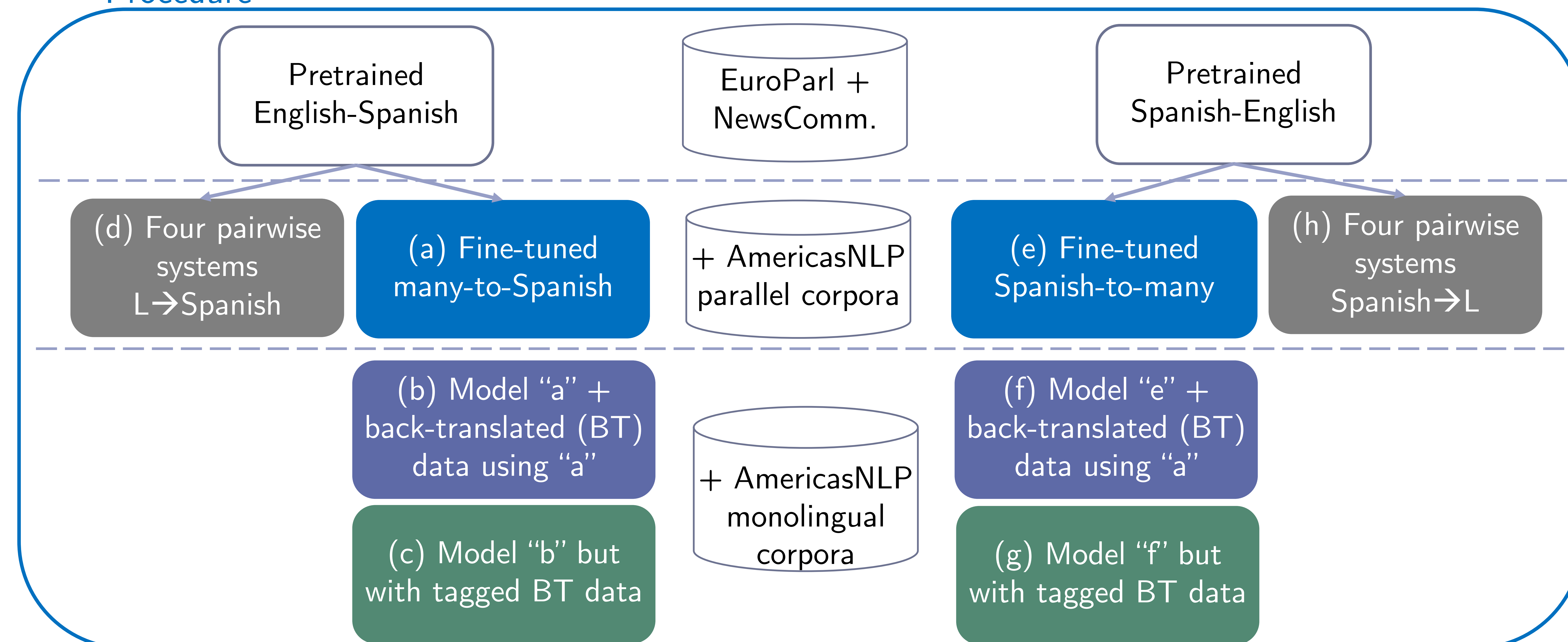
Any interesting findings?

Back-translated (BT) synthetic data did not help, a special tag was needed (tagged BT). Quechua, the highest-resourced of the four, was the only one that underperformed.

What is next to come?

Many-to-many NMT models, domain tags, advanced multilingual sampling and extra components for restoring performance in the higher-resource language.

## Procedure



## Main results

Table 1. **chrF** scores for our dev and devtest custom partitions and the official **AmericasNLP** test sets for the best multilingual setting and the pairwise baseline in each direction.

chrF	Aymara			Ashaninka			Quechua			Shipibo-Konibo		
→ <b>Spanish</b>	dev	devtest	test	dev	devtest	test	dev	devtest	test	dev	devtest	test
●(a) Multilingual	<b>31.73</b>	<b>28.82</b>	<b>22.01</b>	<b>26.78</b>	<b>26.82</b>	<b>22.27</b>	32.92	32.99	29.45	<b>31.41</b>	<b>33.49</b>	<b>31.26</b>
●(d) Pairwise	28.77	25.03	19.79	20.43	20.40	18.83	<b>36.01</b>	<b>36.06</b>	<b>30.90</b>	27.25	29.91	25.31
<b>Spanish</b> →	dev	devtest	test	dev	devtest	test	dev	devtest	test	dev	devtest	test
●(g) Multi+BT[t]	<b>37.32</b>	<b>35.17</b>	<b>26.70*</b>	<b>38.94</b>	<b>38.44</b>	<b>30.81*</b>	44.60	38.94	<b>37.80*</b>	<b>40.67</b>	<b>39.47</b>	<b>33.43*</b>
●(h) Pairwise	28.89	28.23	21.13	32.55	32.29	27.10	<b>45.77</b>	<b>39.68</b>	36.86	34.97	34.96	27.09

\*Our test results could have achieved 2<sup>nd</sup> place in the Shared Task (Track One)

## Extras

Check our paper and repository for more info. →



Out-of-domain evaluation for Quechua and Shipibo-Konibo

New test set for Shipibo-Konibo (*Kirika*) based on traditional tales

News about the release of the models!