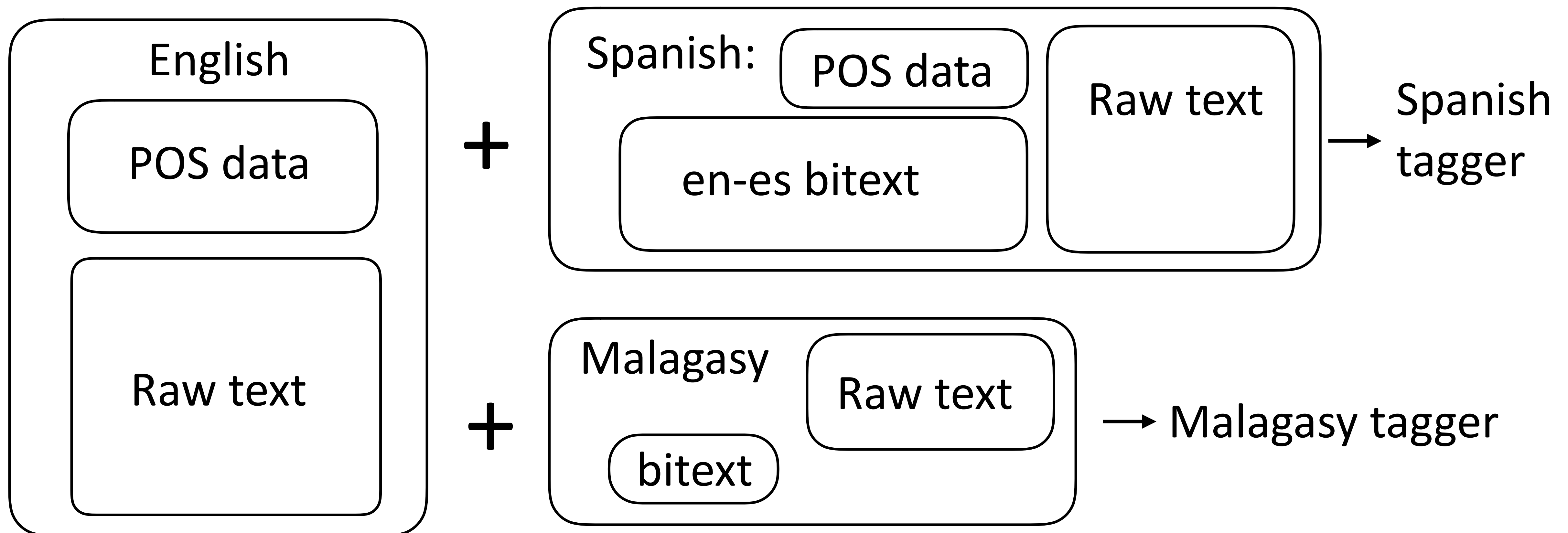


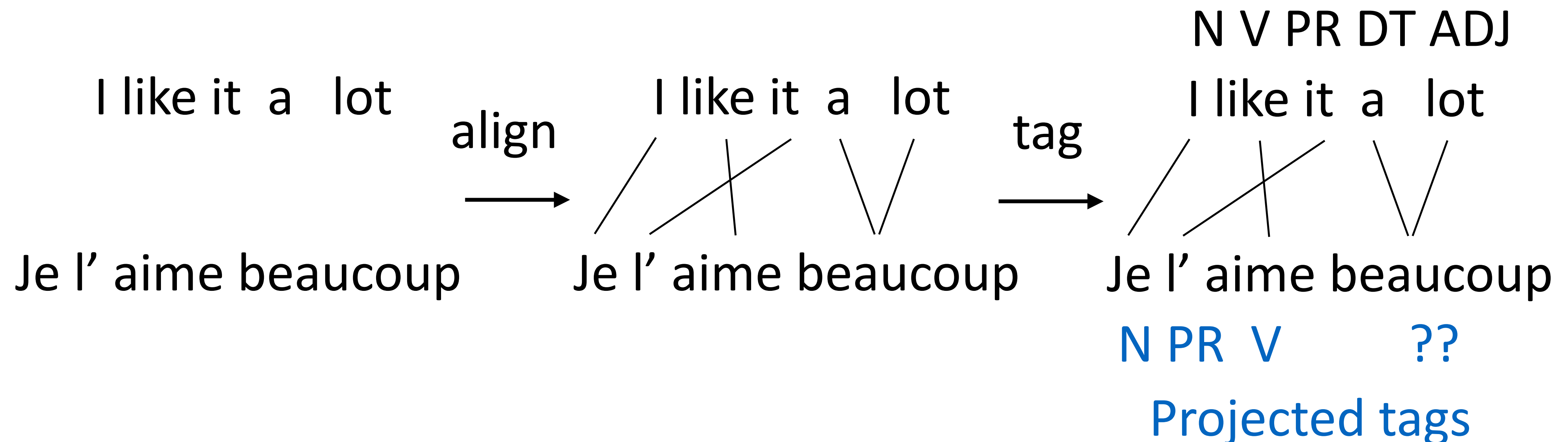
Cross-lingual Tagging

- ▶ Labeling POS datasets is expensive
- ▶ Can we transfer annotation from *high-resource* languages (English, etc.) to *low-resource* languages?



Cross-lingual Tagging

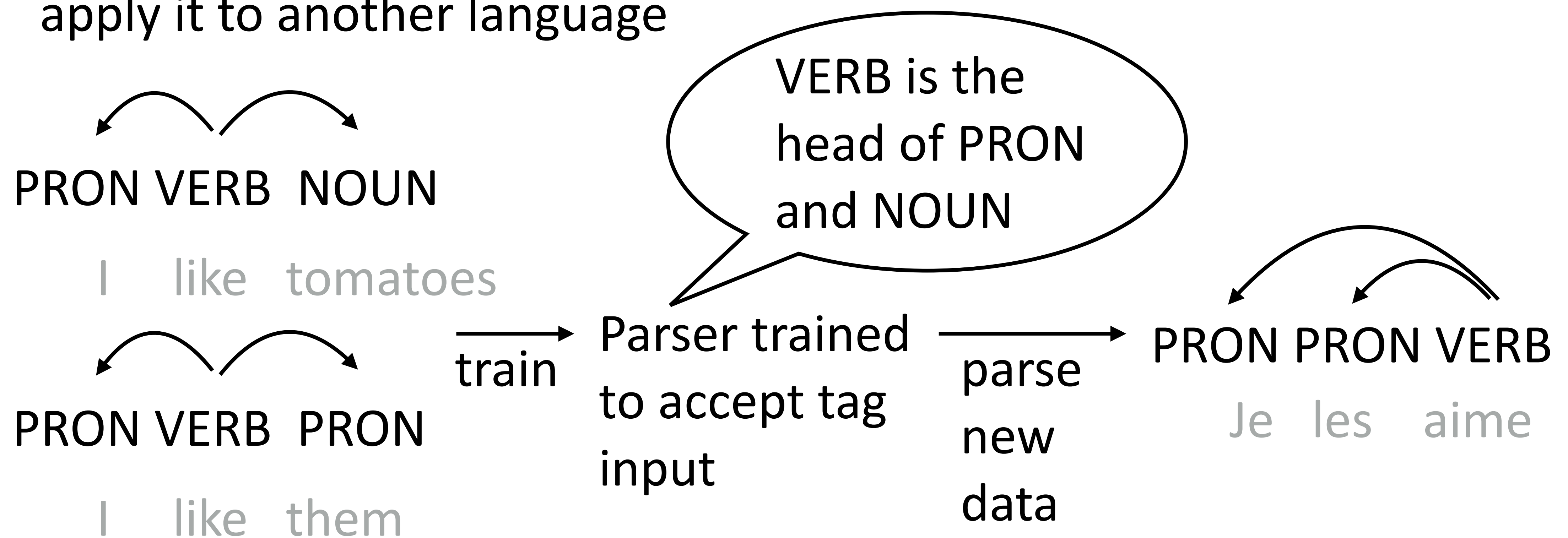
- Can we leverage word alignment here?



- Tag with English tagger, project across bitext, train French tagger?
Works pretty well

Cross-lingual Parsing

- ▶ Now that we can POS tag other languages, can we parse them too?
- ▶ Direct transfer: train a parser over POS sequences in one language, then apply it to another language



Cross-lingual Parsing

	best-source		avg-source gold-POS	gold-POS		pred-POS	
	source	gold-POS		multi-dir.	multi-proj.	multi-dir.	multi-proj.
da	it	48.6	46.3	48.9	49.5	46.2	47.5
de	nl	55.8	48.9	56.7	56.6	51.7	52.0
el	en	63.9	51.7	60.1	65.1	58.5	63.0
es	it	68.4	53.2	64.2	64.5	55.6	56.5
it	pt	69.1	58.5	64.1	65.0	56.8	58.9
nl	el	62.1	49.9	55.8	65.7	54.3	64.4
pt	it	74.8	61.6	74.0	75.6	67.7	70.3
sv	pt	66.8	54.8	65.3	68.0	58.3	62.1
avg		63.7	51.6	61.1	63.8	56.1	59.3

- ▶ Multi-dir: transfer a parser trained on several source treebanks to the target language
- ▶ Multi-proj: more complex annotation projection approach