Multilingual WSD-like Constraints for Paraphrase Extraction

Wilker Aziz[1] Lucia Specia[2]

[1]University of Wolverhampton
Wolverhampton, UK
w.aziz@wlv.ac.uk

[2]University of Sheffield
Sheffield, UK
l.specia@sheffield.ac.uk

August 9, 2013

Table of Contents I

Paraphrasing

"Re-write some text preserving its meaning"

at phrase level

para hacer frente al cambio climático de **forma** mensurable, notificable y verificable

Paraphrasing

"Re-write some text preserving its meaning"

at phrase level

para hacer frente al cambio climático de **forma** mensurable, notificable y verificable

Maybe

- the original phrase is OOV to an SMT system

Paraphrasing

"Re-write some text preserving its meaning"

at phrase level

para hacer frente al cambio climático de **forma** mensurable, notificable y verificable

Maybe

- the original phrase is OOV to an SMT system
- the paraphrase is easier to understand (targeted readers)

Paraphrasing

"Re-write some text preserving its meaning"

at phrase level

para hacer frente al cambio climático de **forma** mensurable, notificable y verificable

Maybe

- the original phrase is OOV to an SMT system
- the paraphrase is easier to understand (targeted readers)
- one needs to expand a query

Paraphrase by pivoting

Pivot through phrase-aligned parallel text [Bannard and Callison-Burch, 2005]

Paraphrase by pivoting

Pivot through phrase-aligned parallel text [Bannard and Callison-Burch, 2005]

para hacer frente al cambio climático de **forma**$_{e_1}$ mensurable, notificable y verificable	

Paraphrase by pivoting

Pivot through phrase-aligned parallel text [Bannard and Callison-Burch, 2005]

para hacer frente al cambio climático de **forma**$_{e_1}$ mensurable, notificable y verificable	to address climate change in a measurable, reportable and verifiable **manner**$_f$

Paraphrase by pivoting

Pivot through phrase-aligned parallel text [Bannard and Callison-Burch, 2005]

para hacer frente al cambio climático de **forma**$_{e_1}$ mensurable, notificable y verificable	to address climate change in a measurable, reportable and verifiable **manner**$_f$
...	...

Paraphrase by pivoting

Pivot through phrase-aligned parallel text [Bannard and Callison-Burch, 2005]

para hacer frente al cambio climático de **forma**$_{e_1}$ mensurable, notificable y verificable	to address climate change in a measurable, reportable and verifiable **manner**$_f$
...	...
	everyone can participate and contribute in an integrated **manner**$_f$.

Paraphrase by pivoting

Pivot through phrase-aligned parallel text [Bannard and Callison-Burch, 2005]

para hacer frente al cambio climático de **forma**$_{e_1}$ mensurable, notificable y verificable	to address climate change in a measurable, reportable and verifiable **manner**$_f$
...	...
cada uno pueda participar y contribuir **de manera**$_{e_2}$ integrada.	everyone can participate and contribute in an integrated **manner**$_f$.

Paraphrase by pivoting

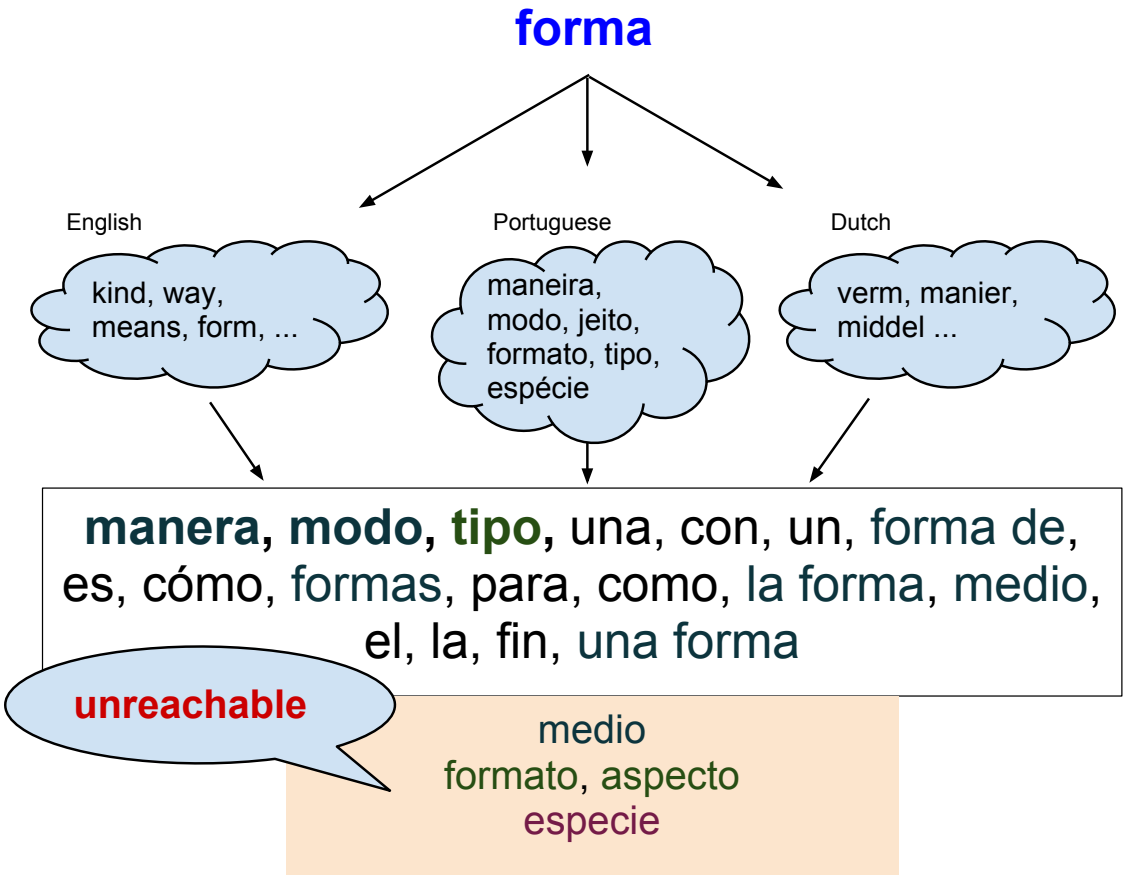
Pivot through phrase-aligned parallel text [Bannard and Callison-Burch, 2005]

para hacer frente al cambio climático de **forma**$_{e_1}$ mensurable, notificable y verificable	to address climate change in a measurable, reportable and verifiable **manner**$_f$
...	...
cada uno pueda participar y contribuir **de manera**$_{e_2}$ integrada.	everyone can participate and contribute in an integrated **manner**$_f$.

Model

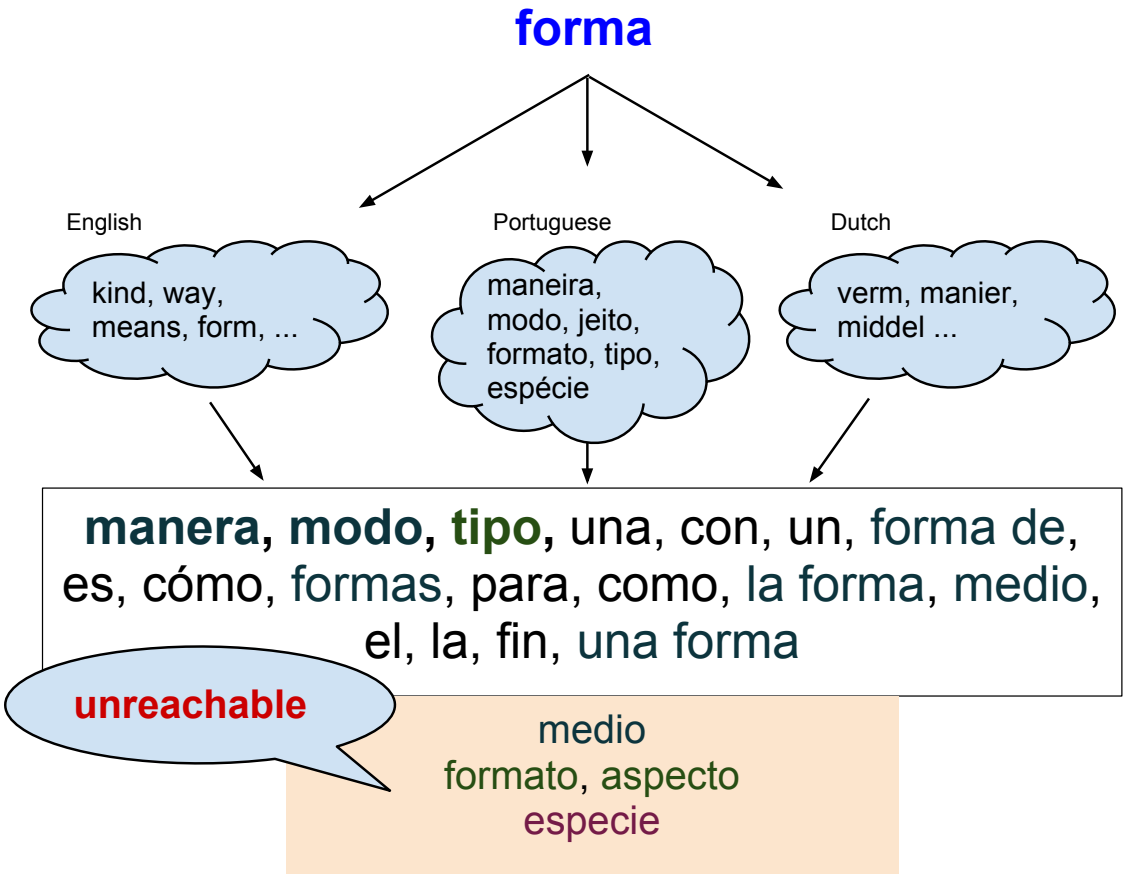
$$p(e_2|e_1) = \sum_{f \in F} p(f|e_1)p(e_e|f)$$

Illustration



- mix of valid senses of the input
- very low diversity
- unreachable paraphrases
- alignment noise

Illustration



Paraphrasing in context

- LM (important role)
- 25% F_1

WSD assumption

What if we could obtain "sense" annotation for free?

WSD assumption

What if we could obtain "sense" annotation for free?

para hacer frente al cambio climático de **forma**$_{e_1}$ mensurable, notificable y verificable	to address climate change in a measurable, reportable and verifiable **manner**$_q$

WSD assumption

What if we could obtain "sense" annotation for free?

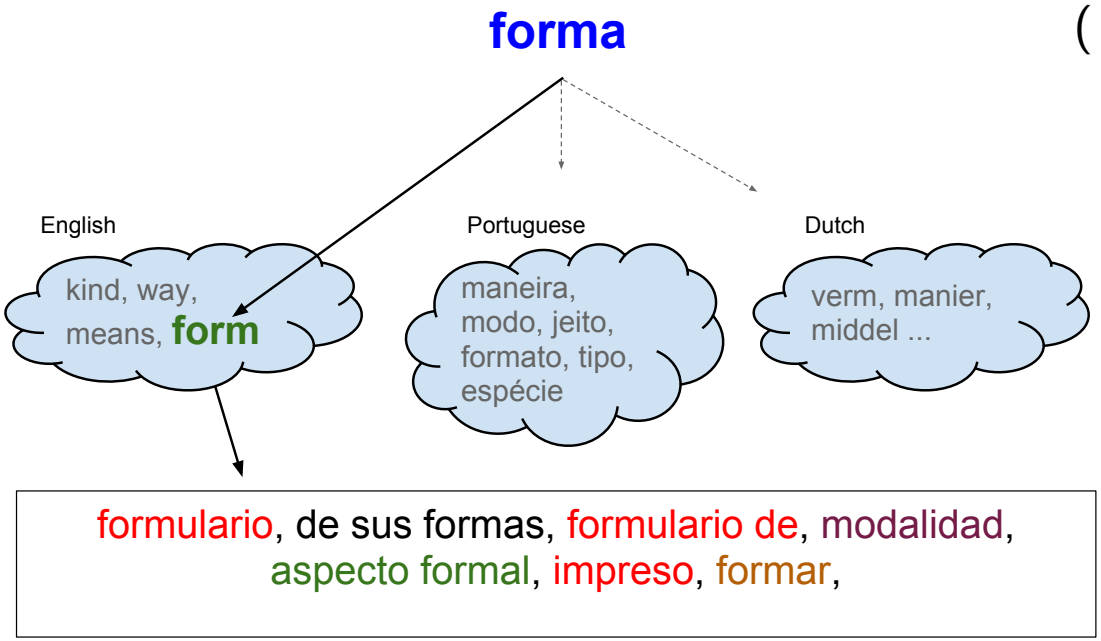
para hacer frente al cambio climático de **forma**$_{e_1}$ mensurable, notificable y verificable	to address climate change in a measurable, reportable and verifiable **manner**$_q$

Different translations of a phrase discriminate different senses of the input

[Resnik and Yarowsky, 1999, Specia et al., 2006, Carpuat and Wu, 2007]

Pivoting through one phrase

$p(e_2|e_1, f) = p(e_2|f)p(f|e_1)$ [Callison-Burch, 2007]

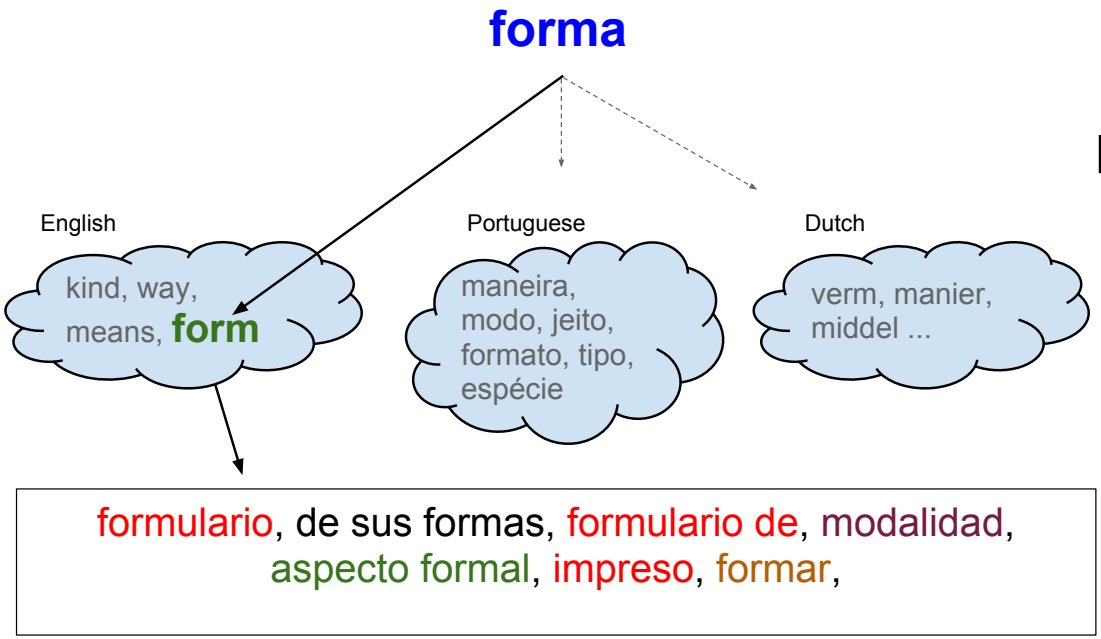


Pivot through a single phrase ("label")

- reach new options
- more diversity
- waste evidence
- sensitive to ambiguity of the "label"
- unreachable paraphrases

Pivoting through one phrase

$$p(e_2|e_1, f) = p(e_2|f)p(f|e_1)$$ [Callison-Burch, 2007]



Paraphrasing in context

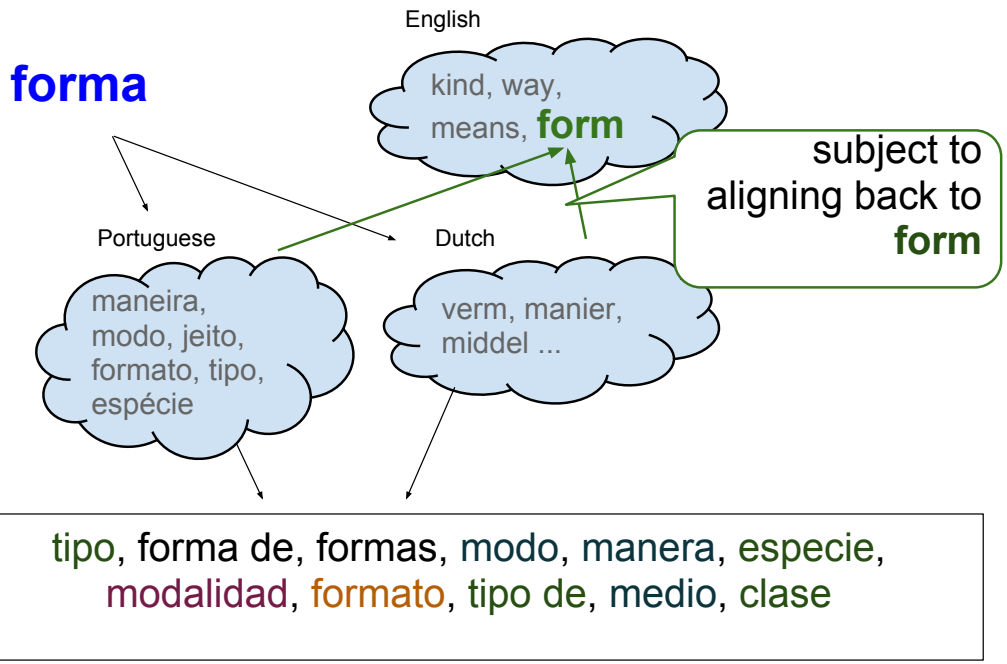
- LM does most of the job
- 30% F_1

Approach

Different translations of a phrase discriminate different senses of the input

[Resnik and Yarowsky, 1999, Specia et al., 2006, Carpuat and Wu, 2007]

$$p(e_2|e_1, q) = \frac{1}{z} \sum_{f \in F} p(e_2|f)p(q|f)p(f|e_1)$$



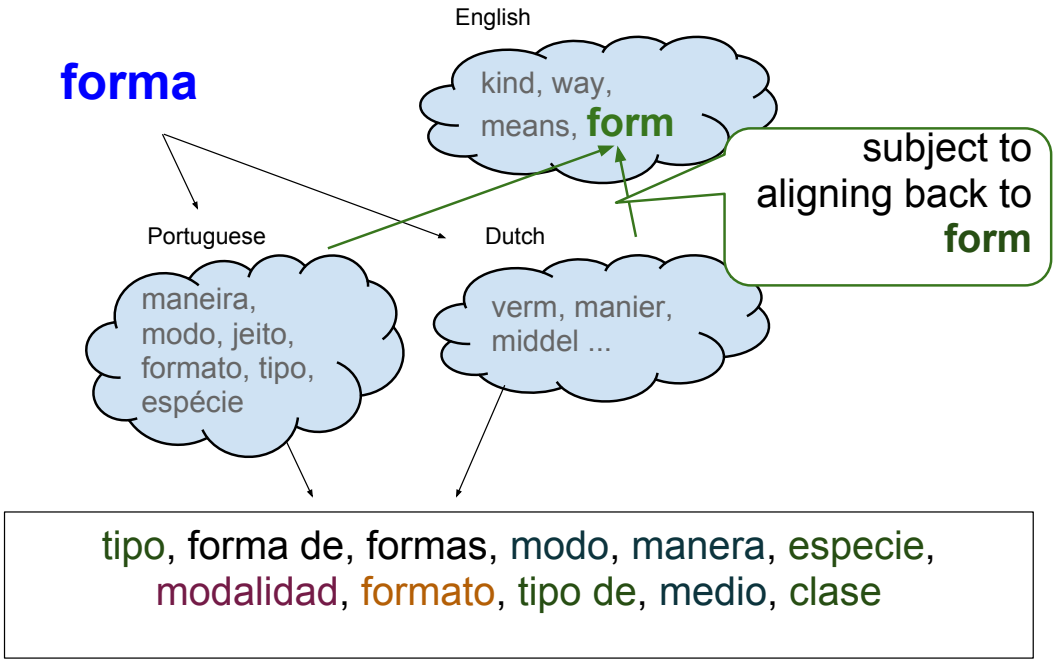
1. pivot through all available phrases
2. observing a constraint q in English

Approach

Different translations of a phrase discriminate different senses of the input

[Resnik and Yarowsky, 1999, Specia et al., 2006, Carpuat and Wu, 2007]

$$p(e_2|e_1, q) = \frac{1}{z} \sum_{f \in F} p(e_2|f)p(q|f)p(f|e_1)$$



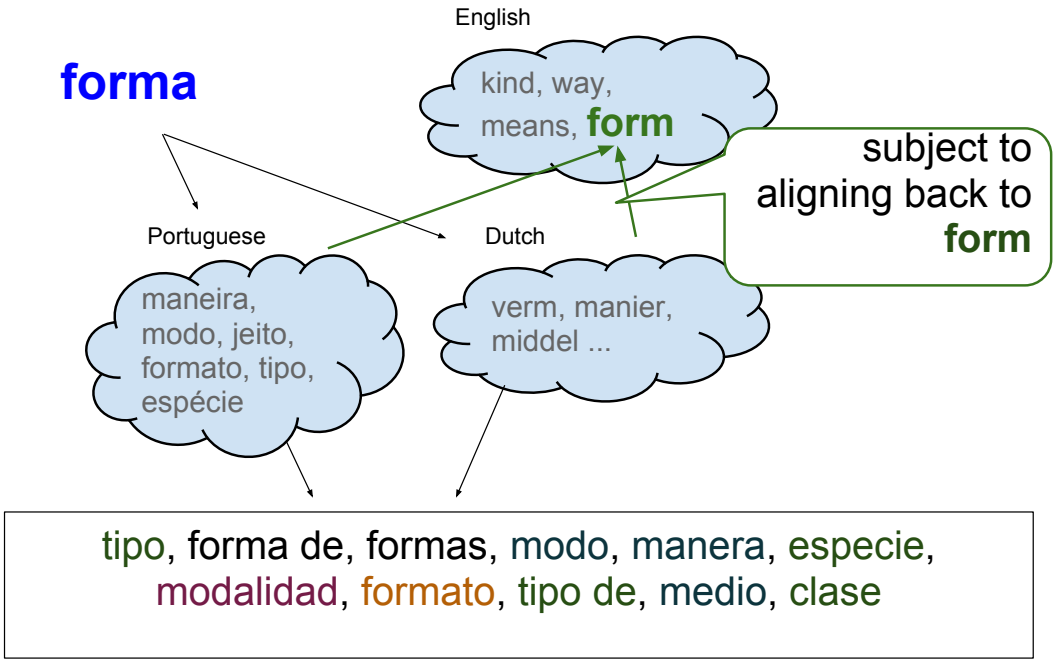
- reach many more options
- quite diverse
- less categorial ambiguity
- requires additional parallel data

Approach

Different translations of a phrase discriminate different senses of the input

[Resnik and Yarowsky, 1999, Specia et al., 2006, Carpuat and Wu, 2007]

$$p(e_2|e_1, q) = \frac{1}{z} \sum_{f \in F} p(e_2|f)p(q|f)p(f|e_1)$$



Paraphrasing in context

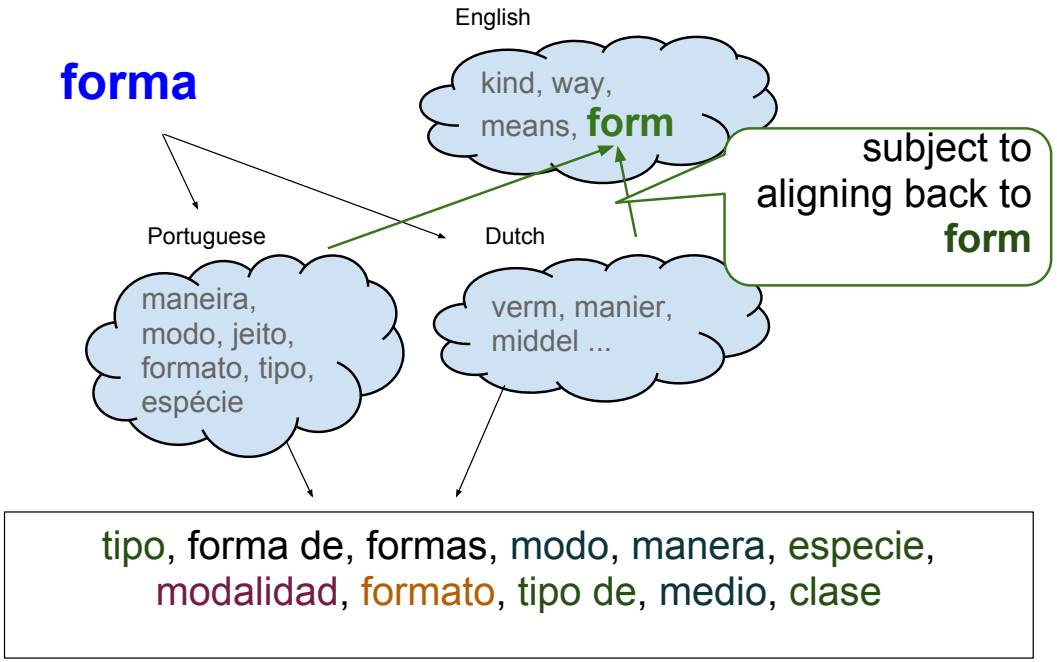
- LM is mostly responsible for fluency
- 49% F_1

Approach

Different translations of a phrase discriminate different senses of the input

[Resnik and Yarowsky, 1999, Specia et al., 2006, Carpuat and Wu, 2007]

$$p(e_2|e_1, q) = \frac{1}{z} \sum_{f \in F} p(e_2|f)p(q|f)p(f|e_1)$$



Related to SMT by triangulation [Cohn and Lapata, 2007]

Experimental setup

Parallel data from Europarl

- input: Spanish
- annotation: English
- 9 pivot languages
  de, nl, da, sv, fi, fr, it, pt, el

Phrase alignment

- Giza++
- SMT symmetrisation heuristics

Test set

- 50 phrases (80% with 2 or more senses in Spanish WordNet)
- in context (2-6 contexts per input phrase)
- 258 sentences

Comparisons

Models

- **ccb**
  original formulation by [Bannard and Callison-Burch, 2005]
- **ccb-wsd**
  pivot through one phrase [Callison-Burch, 2007]
- **multi** proposed model

5-gram LM for re-ranking: simple product

Labels

HT human translation (from parallel corpus)

MT machine translation (phrase-based SMT)
in-domain (0.49 BLEU)

Evaluation

Annotator is given the original sentence and one paraphrase

- M meaning-preserving
- G grammatical
- C correct (M and G)

κ Agreement (8 annotators)

- M 0.54 ± 0.15
- G 0.63 ± 0.16
- C 0.62 ± 0.2

Results

Method	Top	Human Translation					Machine Translation				
		M	G	Correct			M	G	Correct		
		F_1	F_1	P	R	F_1	F_1	F_1	P	R	F_1
CCB	1	32	28	25	25	25					
CCB-wsd	1	61	38	34	28	30	71	39	34	32	33
multi	1	62	55	59	42	49	69	55	50	45	48
CCB	2	41	37	33	33	33					
CCB-wsd	2	68	44	40	33	36	79	46	40	38	39
multi	2	71	64	66	47	55	82	69	63	57	60
CCB	3	46	42	37	37	37					
CCB-wsd	3	71	47	45	36	40	83	50	44	41	42
multi	3	74	67	71	50	59	85	74	69	62	65

SMT does better WSD than people?!

- most likely not, but it did offer more straightforward labels

No LM re-ranking

	Human Translation			Machine Translation		
Method	**M**	**G**	**Correct**	**M**	**G**	**Correct**
CCB	33	23	22			
CCB-wsd	19	9	8	16	8	7
multi	64	43	37	63	40	35

For our model

- options are mostly meaning-preserving
- LM helps w.r.t. grammaticality

Applications

MT evaluation

- paraphrase the MT
  (source text becomes "annotation")

OOV in SMT

- paraphrase the source phrases
  (target phrases become "annotation")

Final remarks

Summary

- paraphrase extraction with WSD-like labels
- pivoting through multiple languages
- separate senses of the input
- reach more paraphrases
- produce less noise
- less LM dependent

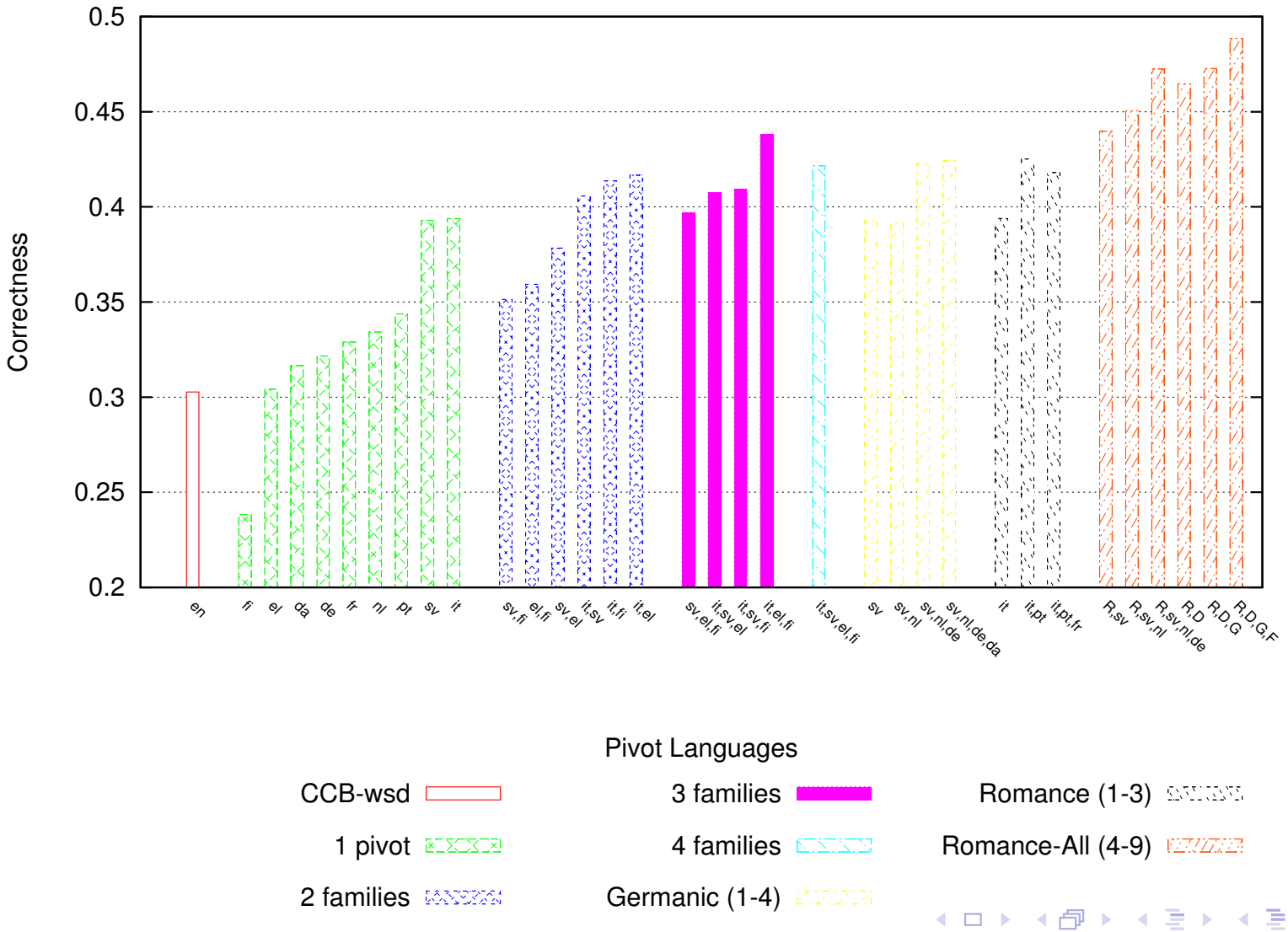
Future work

- couple related labels
  e.g. “way” and “means”, or “form” and “forms”

Thanks!

Questions/comments?

Different families



References I

Colin Bannard and Chris Callison-Burch. Paraphrasing with bilingual parallel corpora. In *Proceedings of the 43rd Annual Meeting on Association for Computational Linguistics*, pages 597–604, Ann Arbor, Michigan, 2005. doi: 10.3115/1219840.1219914.

Chris Callison-Burch. *Paraphrasing and Translation*. PhD thesis, University of Edinburgh, Edinburgh, Scotland, 2007. URL `http://cs.jhu.edu/~ccb/publications/callison-burch-thesis.pdf`.

References II

Marine Carpuat and Dekai Wu. Improving statistical machine translation using word sense disambiguation. In *The 2007 Joint Conference on Empirical Methods in Natural Language Processing and Computational Natural Language Learning*, EMNLP-CoNLL '07, pages 61–72, Prague, Czech Republic, 2007.

Trevor Cohn and Mirella Lapata. Machine translation by triangulation: Making effective use of multi-parallel corpora. In *Proceedings of the 45th Annual Meeting of the Association for Computational Linguistics*, Prague, Czech Republic, 2007.

Philip Resnik and David Yarowsky. Distinguishing systems and distinguishing senses: new evaluation methods for word sense disambiguation. *Nat. Lang. Eng.*, 5(2):113–133, 1999.

References III

Lucia Specia, Mark Stevenson, Maria das Graças Volpe Nunes, and Gabriela C.B. Ribeiro. Multilingual versus monolingual WSD. In *Proceedings of the EACL Workshop "Making Sense of Sense: Bringing Psycholinguistics and Computational Linguistics Together"*, pages 33–40, Trento, Italy, 2006.