Machine Translation Part 3
Alternatives & Evaluation

CS 585, Fall 2015
Introduction to Natural Language Processing
http://people.cs.umass.edu/~brenocon/inlp2015/

Brendan O'Connor
College of Information and Computer Sciences
University of Massachusetts Amherst

[Some slides borrowed from mt-class.org]

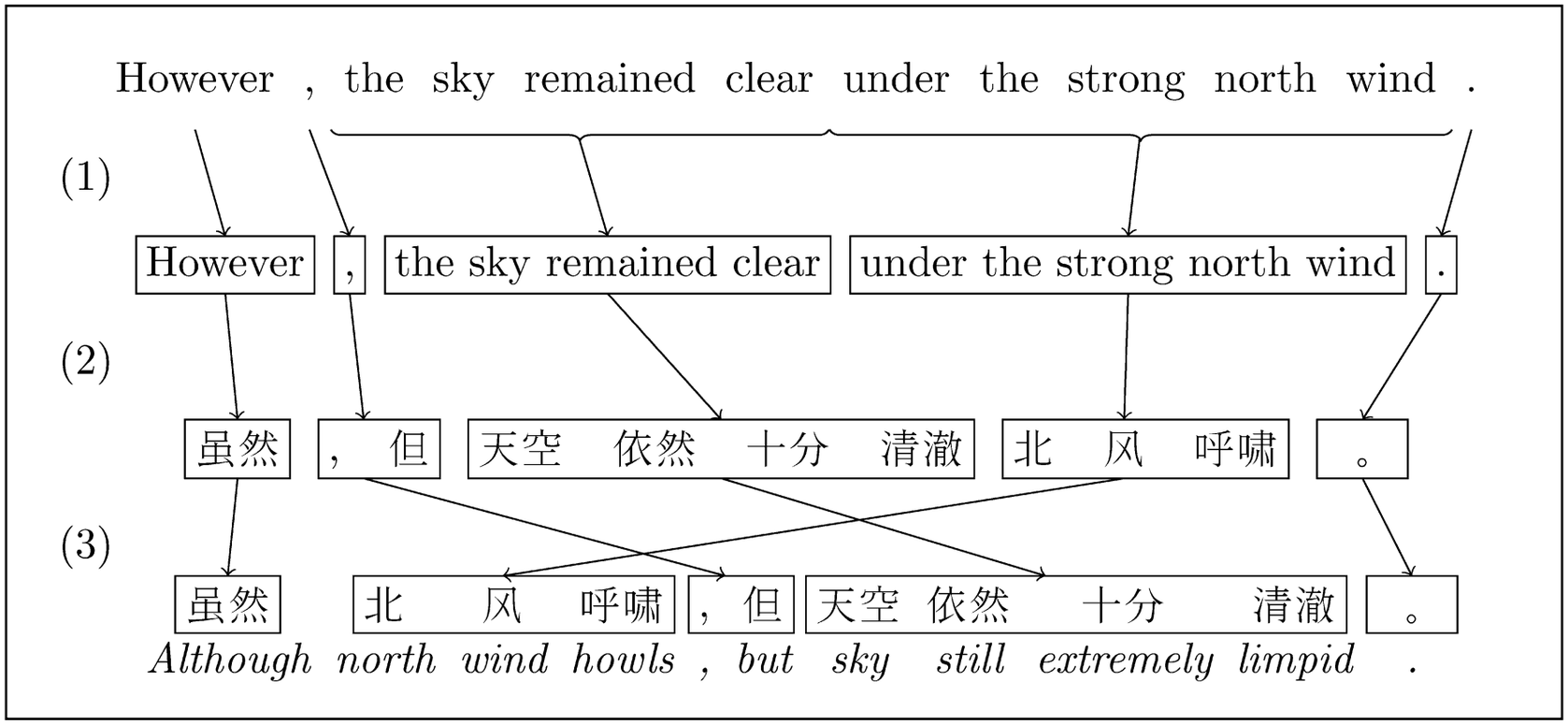
Machine Translation

- MT beyond word-based models
  - Phrase-based
  - Syntax
  - Features
- Evaluation

Phrase-based MT

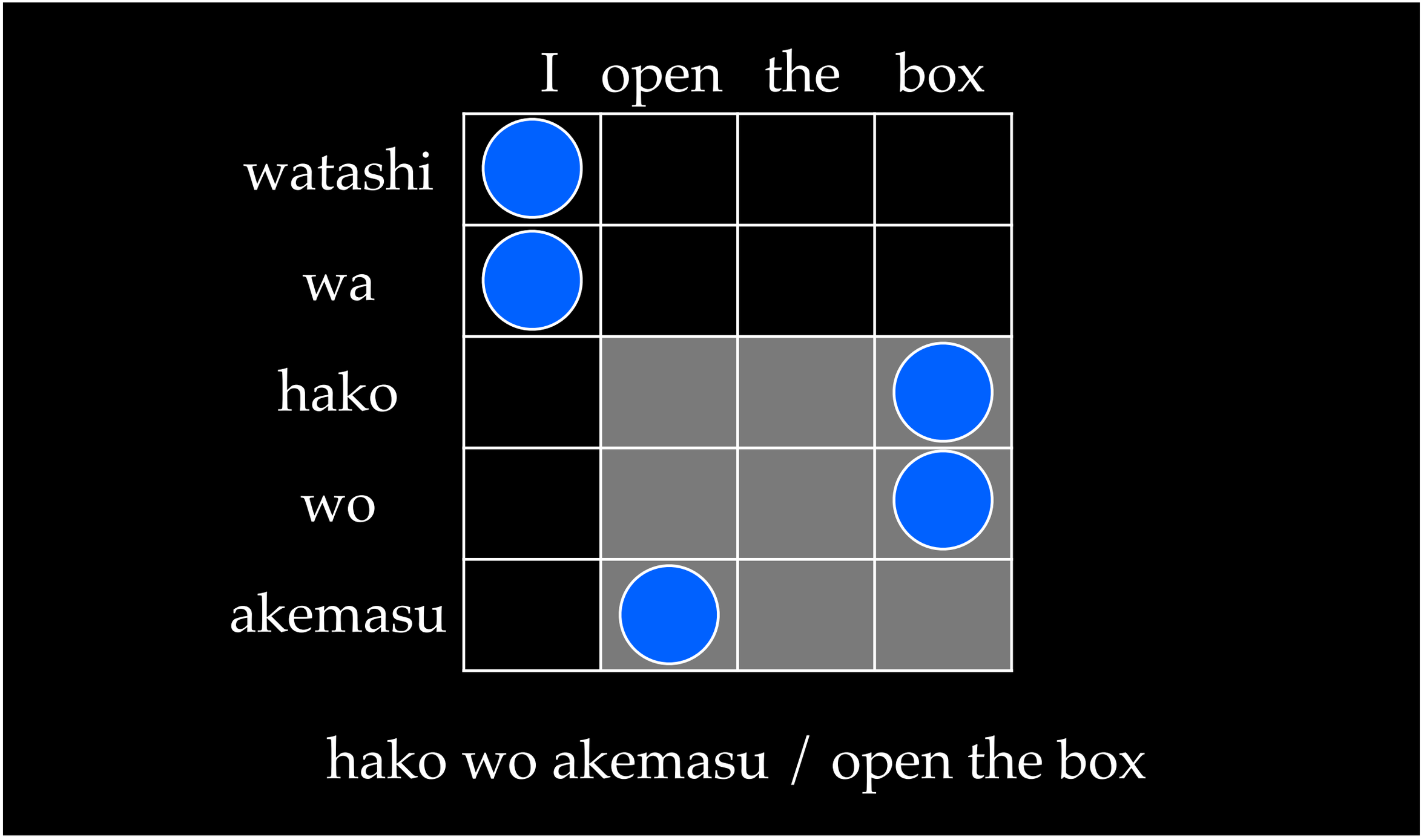
$$p(\mathbf{f}, \mathbf{a} \mid \mathbf{e}) = p(\mathbf{f} \mid \mathbf{e}, \mathbf{a})\ p(\mathbf{a} \mid \mathbf{e})$$

Phrase-to-phrase translations

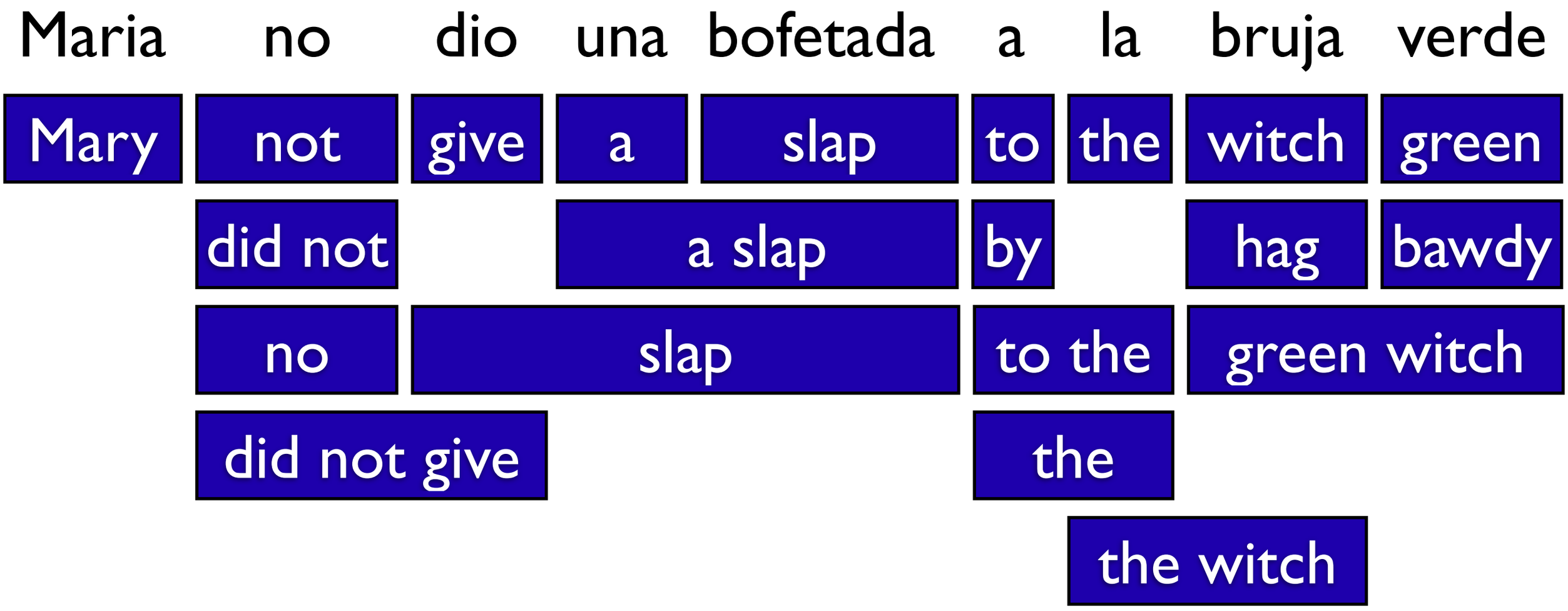


- Phrases can memorize local reorderings
- State-of-the-art (currently or very recently) in industry, e.g. Google Translate

Phrase extraction for training:
Preprocess with IBM Models to predict alignments



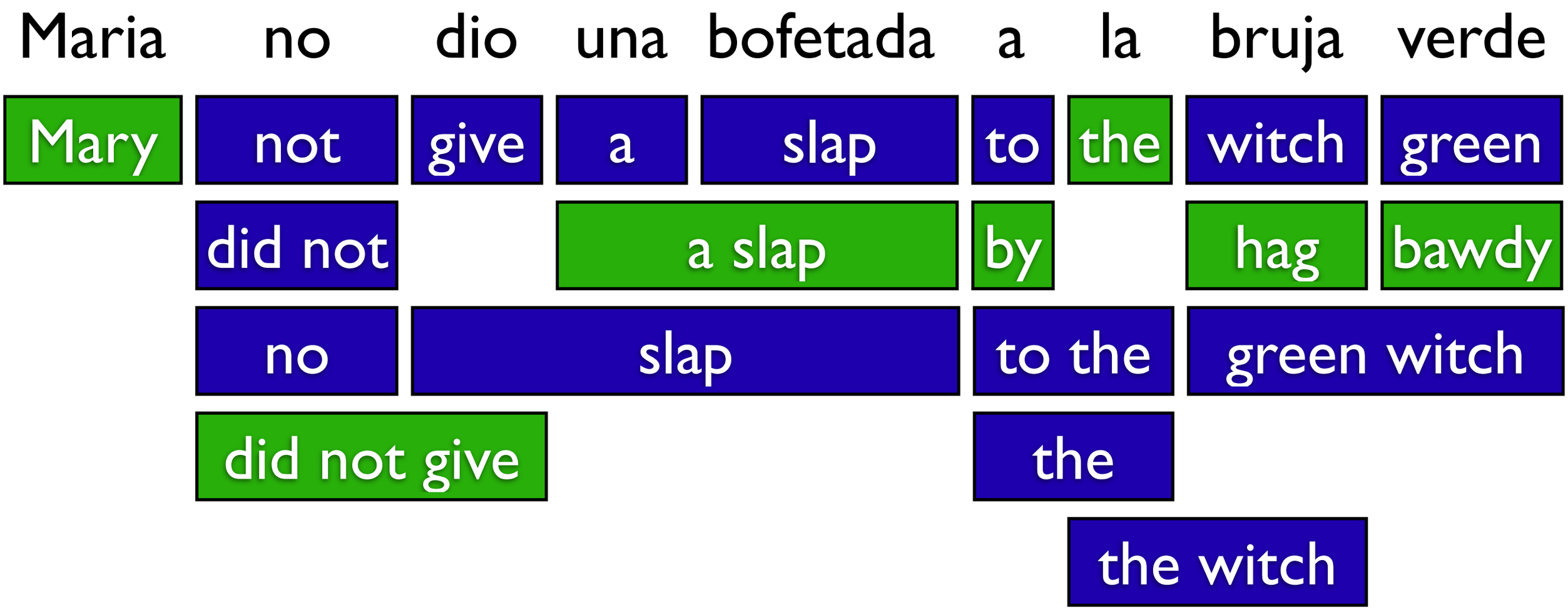
Decoding


Maria no dio una bofetada a la bruja verde
Mary not give a slap to the witch green
did not a slap by hag bawdy
no slap to the green witch
did not give the
the witch

Decoding

Maria	no	dio	una	bofetada	a	la	bruja	verde
Mary	not	give	a	slap	to	the	witch	green
did not			a slap		by		hag	bawdy
no		slap			to the		green witch	
did not give					the			
						the witch		

Decoding


Maria no dio una bofetada a la bruja verde
Mary not give a slap to the witch green
did not a slap by hag bawdy
no slap to the green witch
did not give the
the witch

More MT issues

- MT beyond word-based models
  - Phrase-based
  - Syntax
  - Features
    (noisy channel is just two log-linear features)
- Evaluation