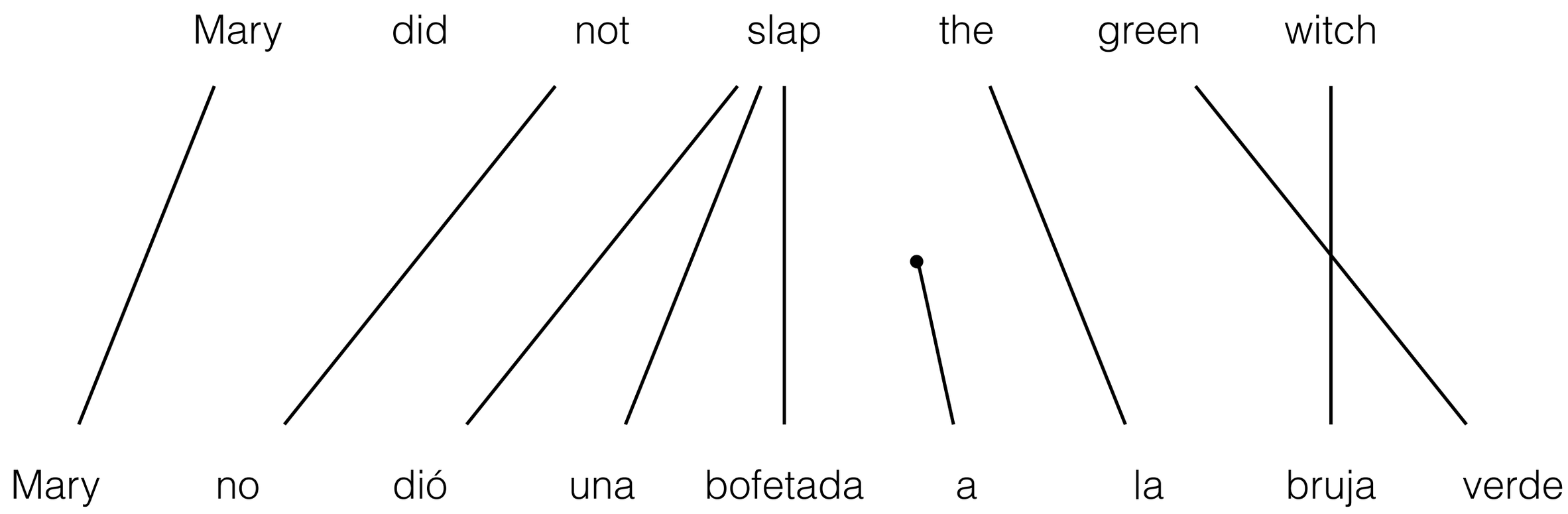
Word-Based Translation

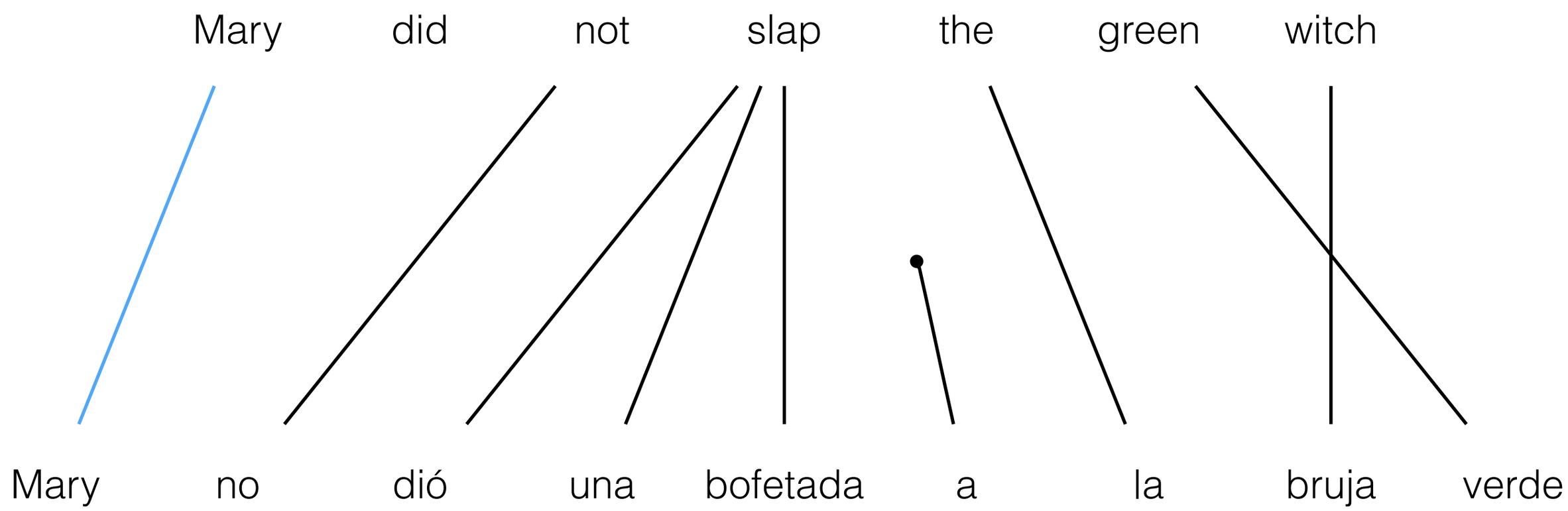
Wilker Aziz
5/4/16

Word-based Translation



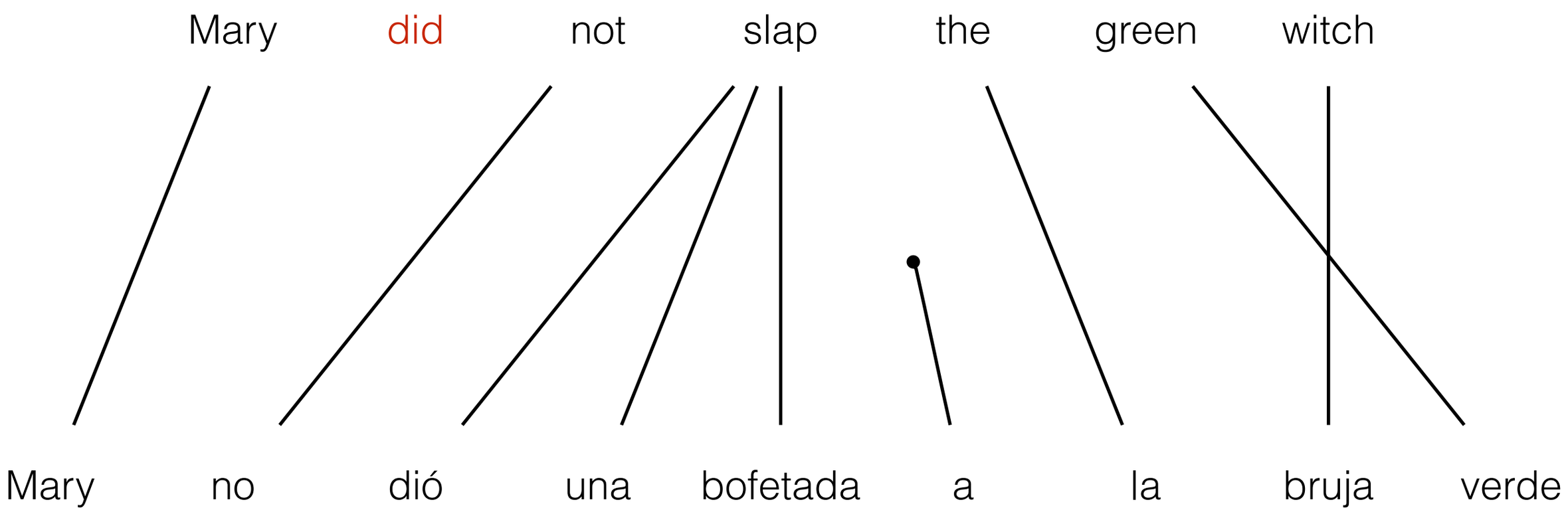
Every French word is generated by an English word (or null)

Word-based Translation



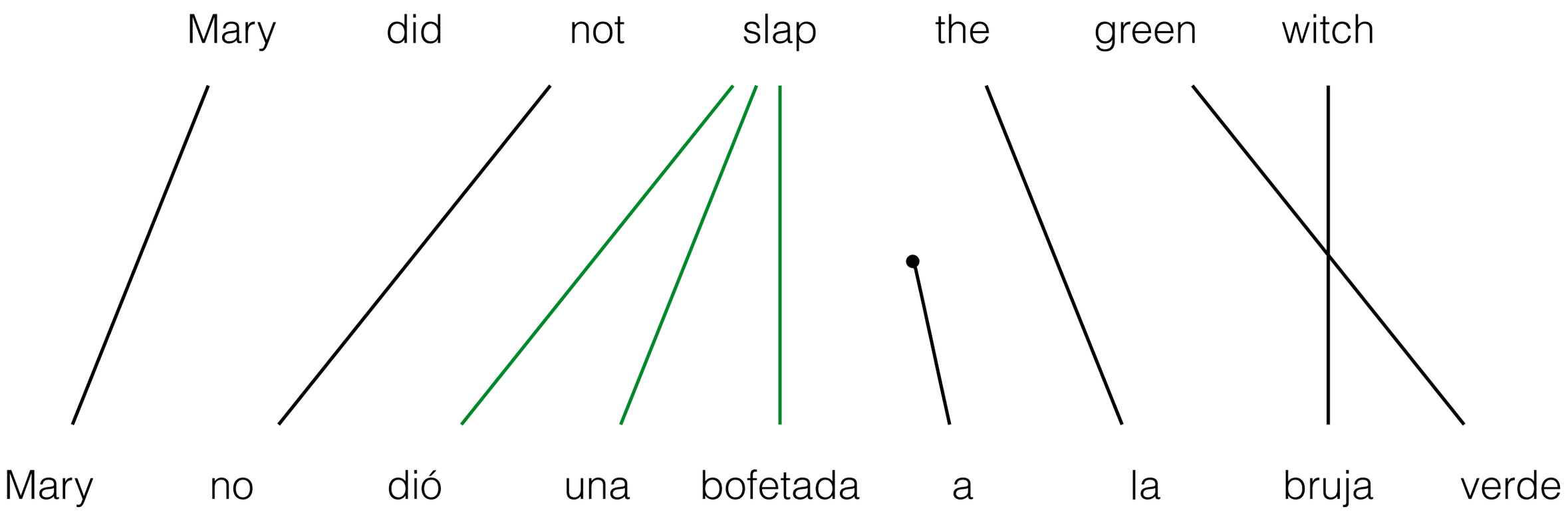
Some English words align once

Word-based Translation



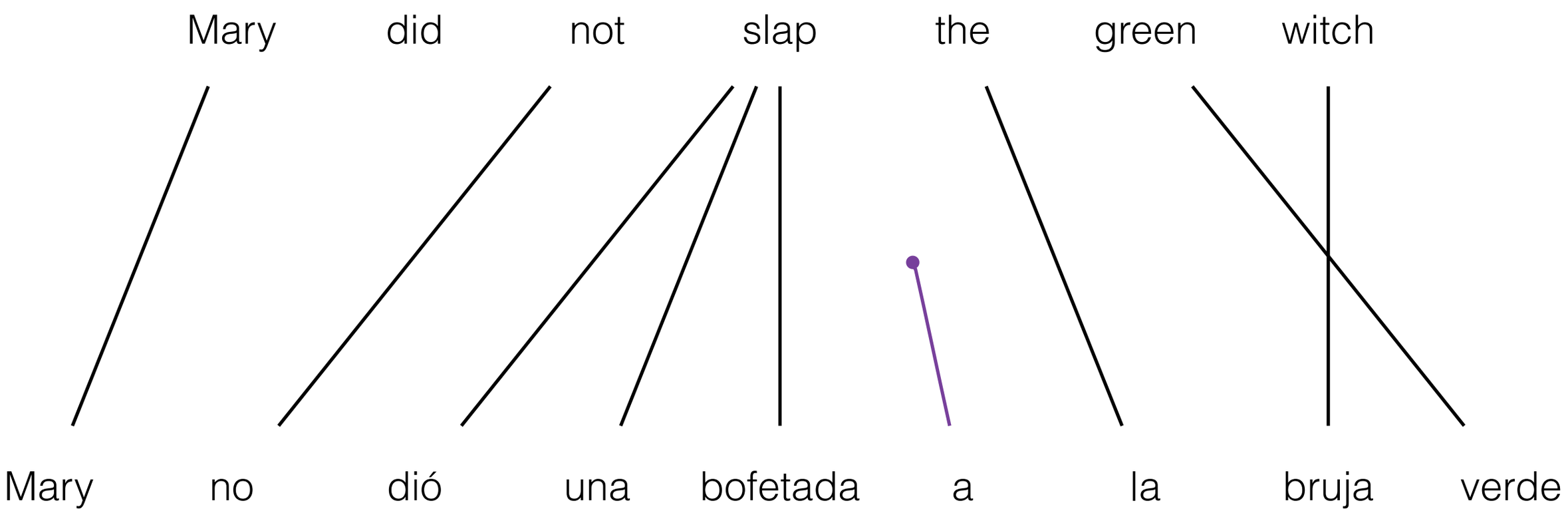
Some English words disappear

Word-based Translation



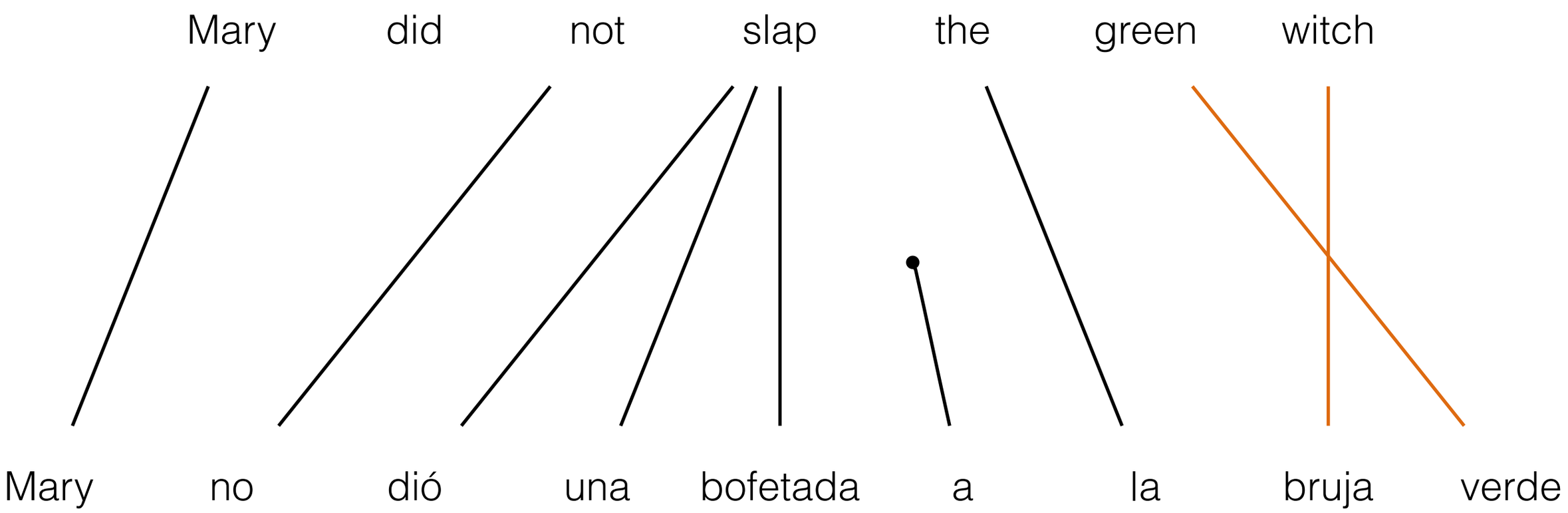
Some English words align multiple times

Word-based Translation



Some French words are null-aligned

Word-based Translation



Some French words are translated out of order

IBM models 1 and 2

Generative story

1. given an English sentence E whose length is *l*, choose length *m* of French sentence F

2. choose how to connect each position in F to a position in E

3. for each position in F, choose which word to place there

IBM1: uniform distribution over connections

IBM2: probability of a connection depends on the positions it connects and the lengths

Inference (IBM 1 and 2)

- Infer assignments for latent variables (alignment)
  - Tractable
    alignments are independent of one another
- Typical use
  - Estimation: expected alignments
  - Prediction: Viterbi alignments

Generation

When do we ever need to generate from our model?

> we do not use our translation models to predict French given English, but rather as a component of a system designed to **predict English given French** (Brown et al., 1993)

Translation

$$P(E|F) \propto P(E)P(F|E)$$

Generation

Do you see a problem in generating from IBM1 or 2?

Generation

Do you see a problem in generating from IBM1 or 2?

- given a French sentence, there are too many English candidates (perhaps infinitely many)

Generation

Do you see a problem in generating from IBM1 or 2?

- given a French sentence, there are too many English candidates (perhaps infinitely many)
- the generative story does not account for all English words

Generation

Do you see a problem in generating from IBM1 or 2?

- given a French sentence, there are too many English candidates (perhaps infinitely many)
- the generative story does not account for all English words
- we model very little about English words

Generation

Do you see a problem in generating from IBM1 or 2?

- given a French sentence, there are too many English candidates (perhaps infinitely many)
- the generative story does not account for all English words
- we model very little about English words
  - how many deletions?

Generation

Do you see a problem in generating from IBM1 or 2?

- given a French sentence, there are too many English candidates (perhaps infinitely many)
- the generative story does not account for all English words
- we model very little about English words
  - how many deletions?
  - how many insertions?

Generative Story $IBM_{\geq 3}$: Given E

Generative Story IBM$_{\geq 3}$: Fertility

Mary	did	not	slap				the	green	witch
Mary	~~did~~	not	slap	slap	slap		the	green	witch

Generative Story IBM$_{\geq 3}$: NULL insertion

Mary	did	not	slap				the	green	witch
Mary	~~did~~	not	slap	slap	slap		the	green	witch
						NULL			

Generative Story IBM$_{\geq 3}$: Translation

Mary	did	not	slap				the	green	witch
Mary	~~did~~	not	slap	slap	slap		the	green	witch
						NULL			
Mary		no	dió	una	bofetada	a	la	verde	bruja

Generative Story IBM$_{\geq 3}$: Distortion

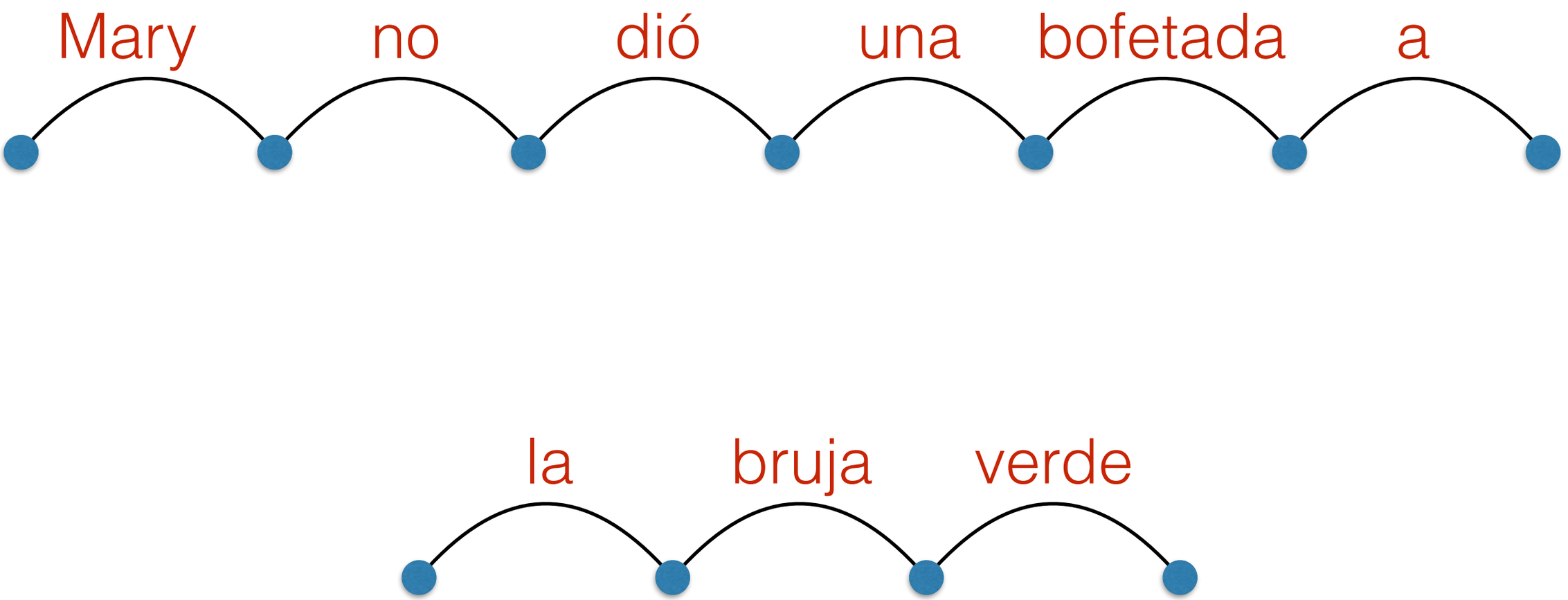
Mary	did	not	slap				the	green	witch
Mary	~~did~~	not	slap	slap	slap		the	green	witch
						NULL			
Mary		no	dió	una	bofetada	a	la	verde	bruja
Mary		no	dió	una	bofetada	a	la	bruja	verde

"Reverse" Generation Example

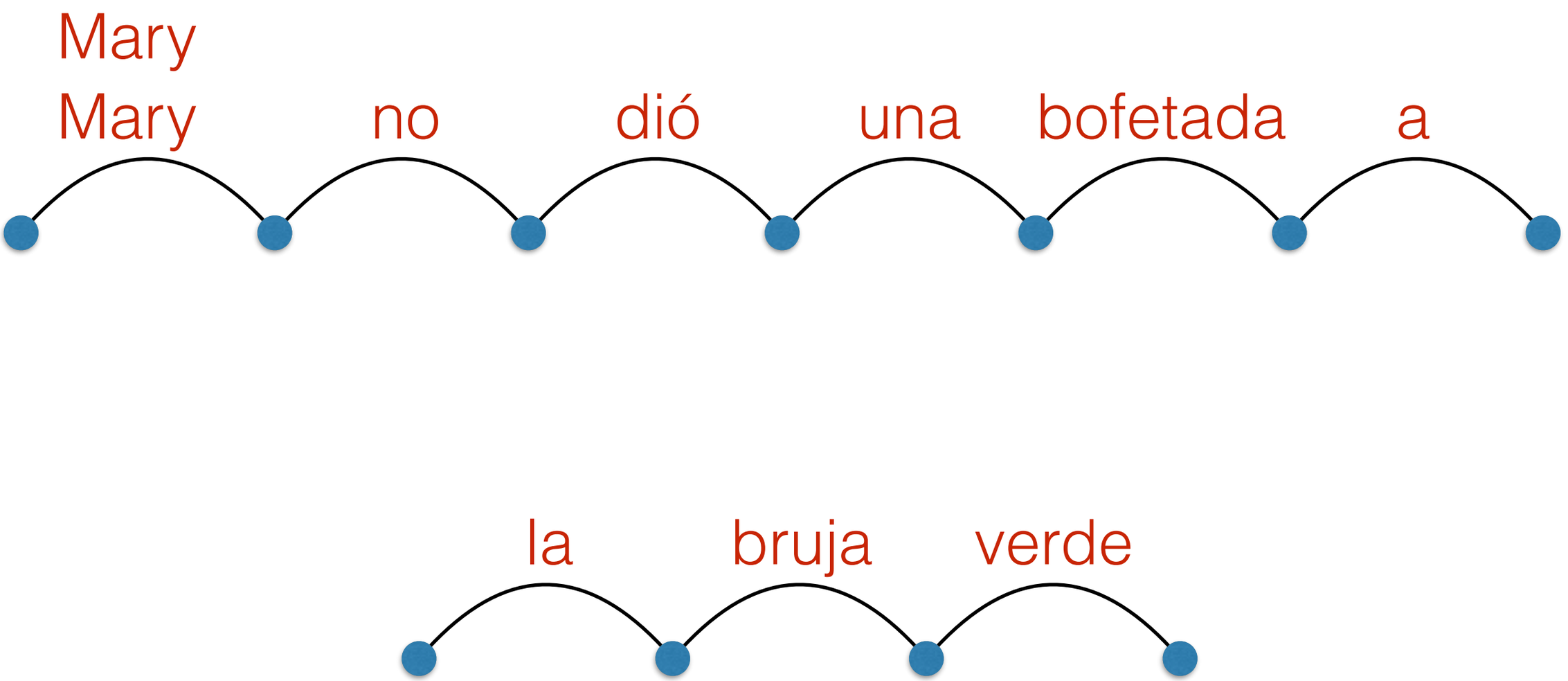
Mary no dió una bofetada a la bruja verde

Example and figures from
(Knight and Al-Onaizan, 1998) and (Knight and May, 2009).

"Reverse" Distortion



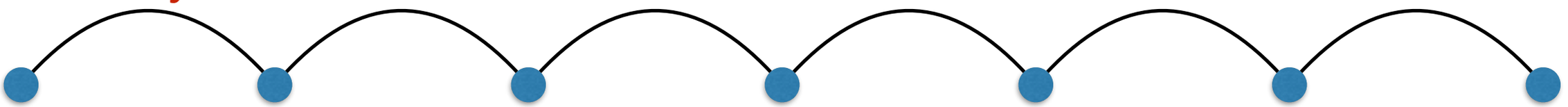
"Reverse" Distortion



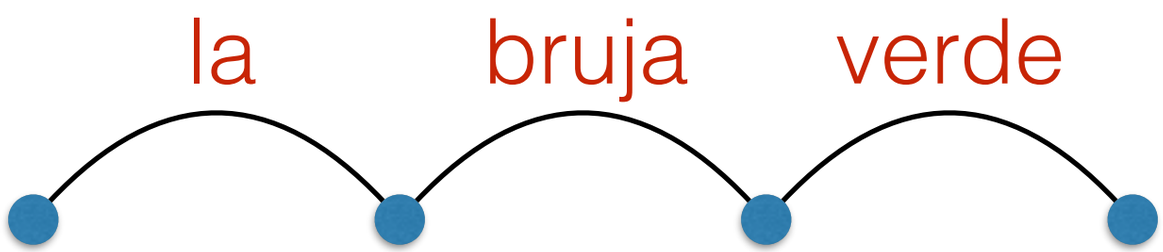
"Reverse" Distortion

Mary no

Mary no dió una bofetada a



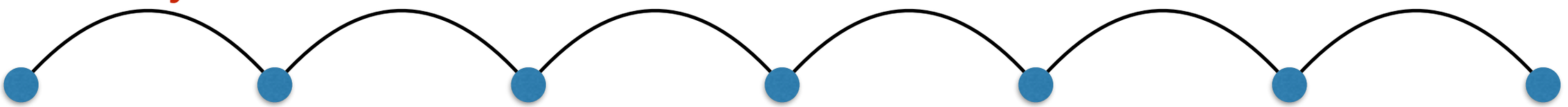
la bruja verde



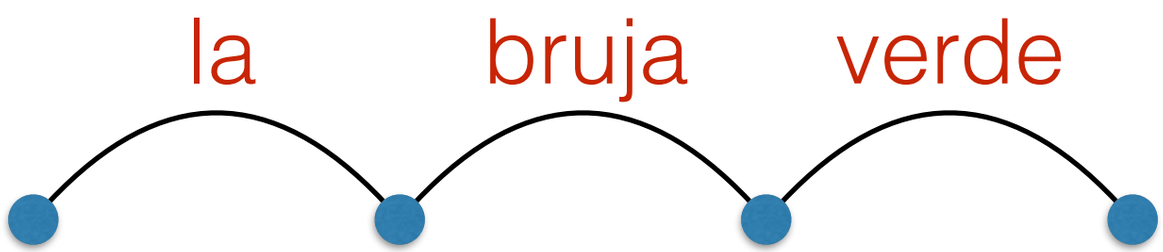
"Reverse" Distortion

Mary no dió

Mary no dió una bofetada a



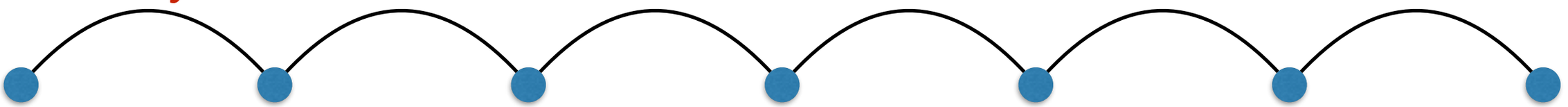
la bruja verde



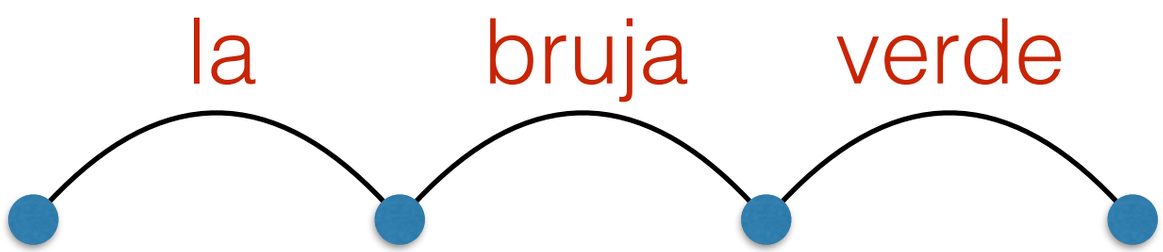
"Reverse" Distortion

Mary no dió una

Mary no dió una bofetada a



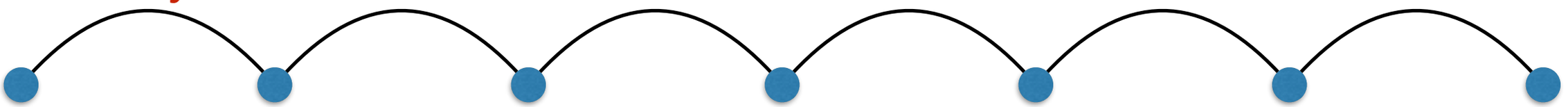
la bruja verde



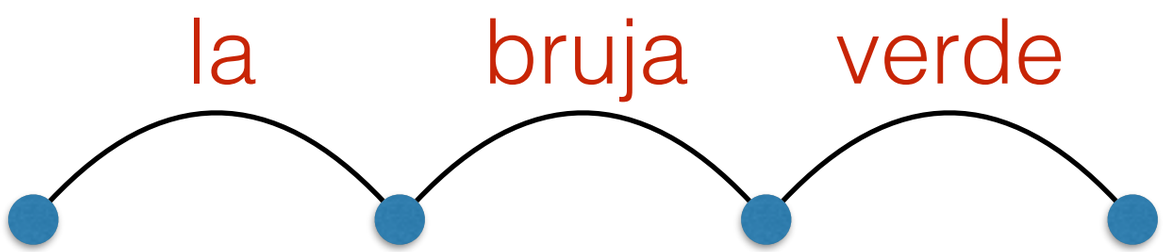
"Reverse" Distortion

Mary no dió una bofetada

Mary no dió una bofetada a

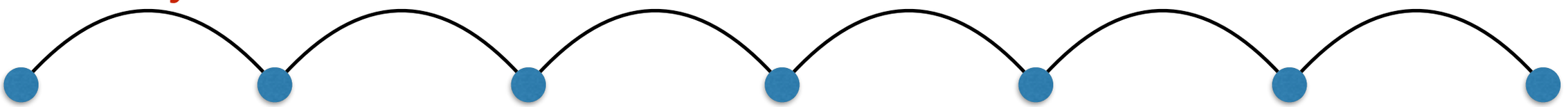


la bruja verde

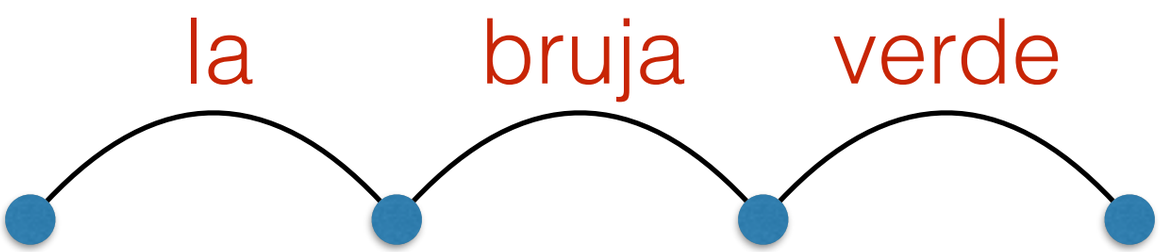


"Reverse" Distortion

Mary no dió una bofetada a

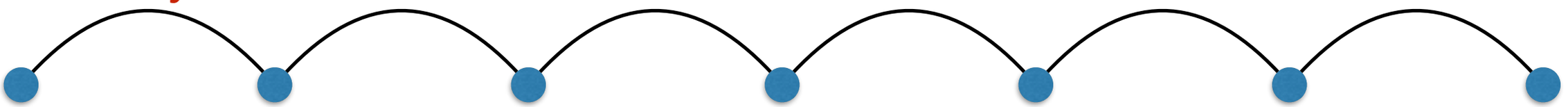


Mary no dió una bofetada a

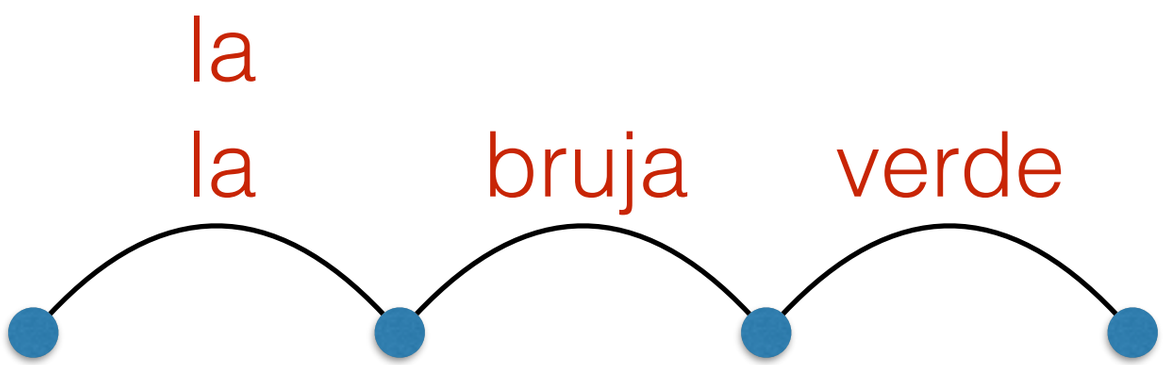


"Reverse" Distortion

Mary no dió una bofetada a

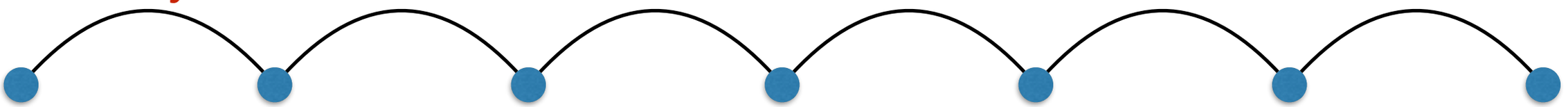


Mary no dió una bofetada a



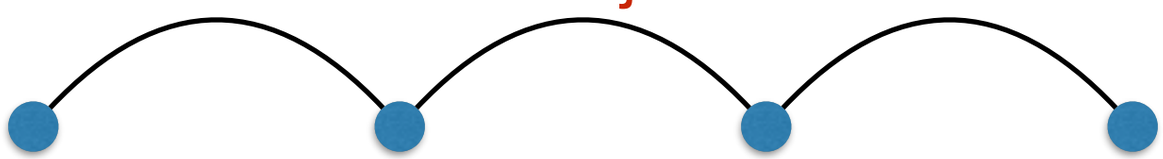
"Reverse" Distortion

Mary no dió una bofetada a



Mary no dió una bofetada a

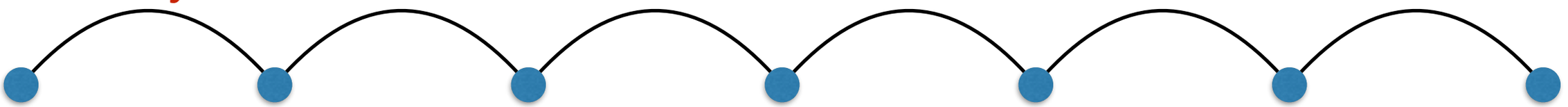
la verde



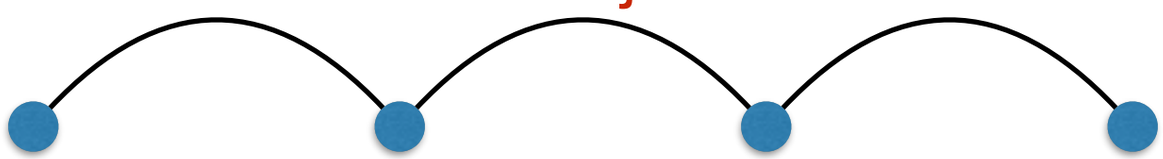
la bruja verde

"Reverse" Distortion

Mary	no	dió	una	bofetada	a
Mary	no	dió	una	bofetada	a

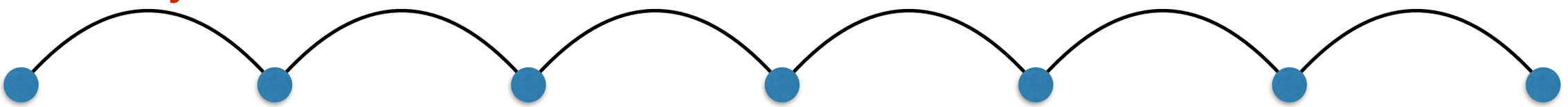


la	verde	bruja
la	bruja	verde

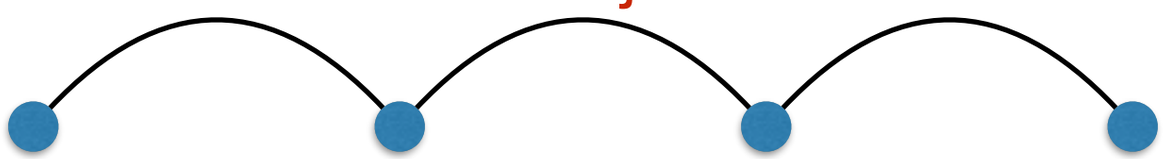


"Reverse" Translation

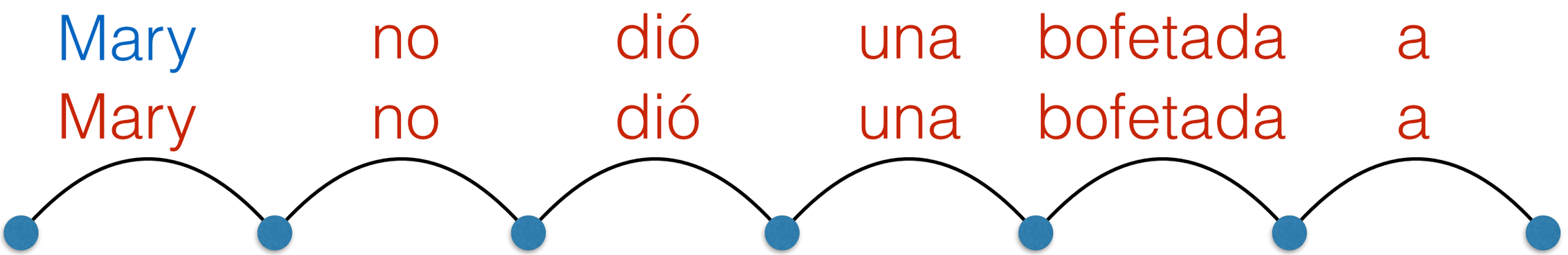
Mary	no	dió	una	bofetada	a
Mary	no	dió	una	bofetada	a

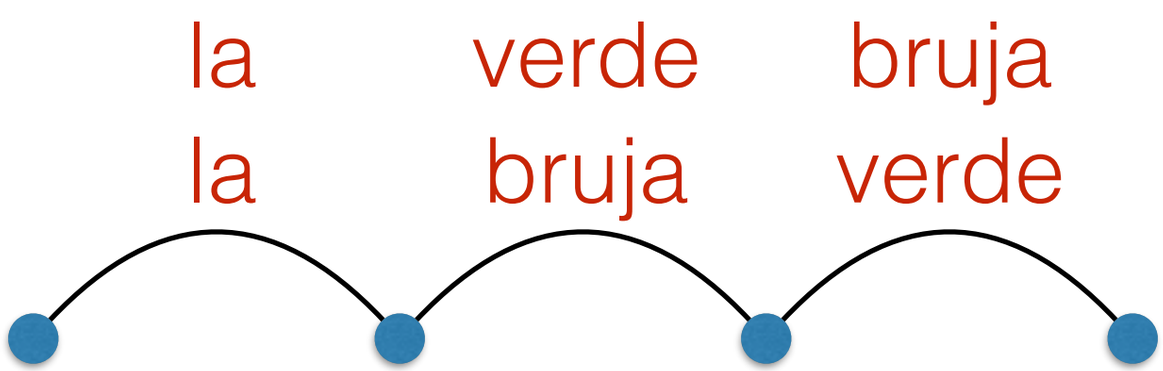


la	verde	bruja
la	bruja	verde



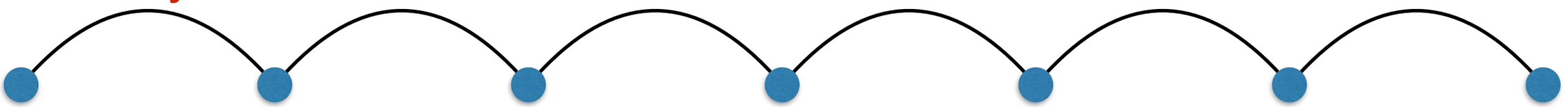
"Reverse" Translation



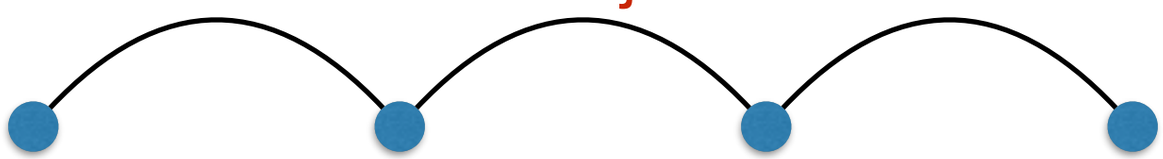


"Reverse" Translation

Mary	not	dió	una	bofetada	a
Mary	no	dió	una	bofetada	a

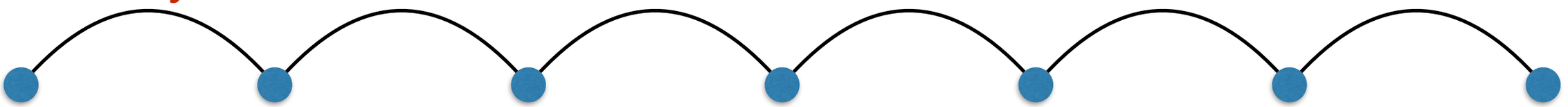


la	verde	bruja
la	bruja	verde

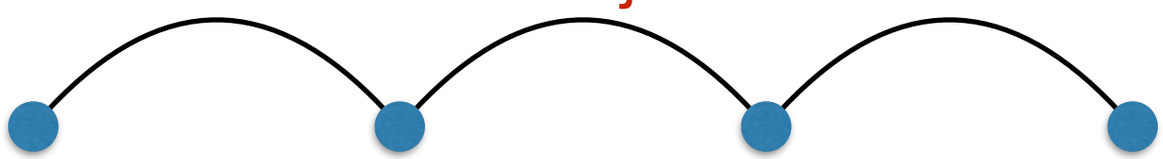


"Reverse" Translation

Mary	not	slap	una	bofetada	a
Mary	no	dió	una	bofetada	a

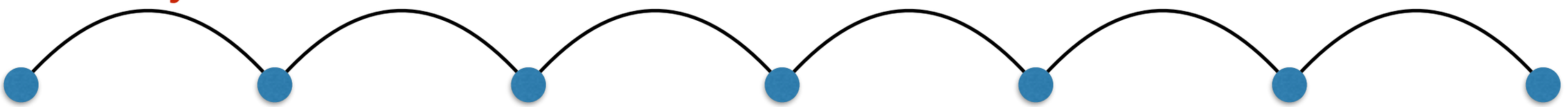


la	verde	bruja
la	bruja	verde

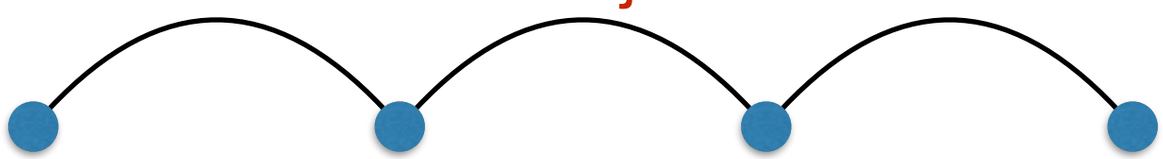


"Reverse" Translation

Mary	not	slap	slap	bofetada	a
Mary	no	dió	una	bofetada	a

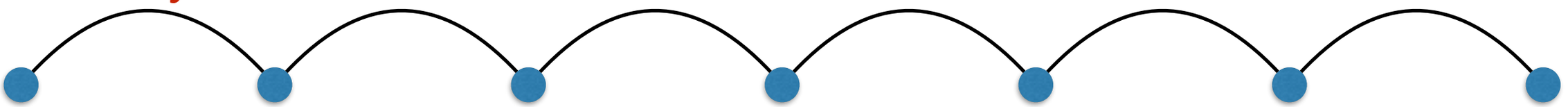


la	verde	bruja
la	bruja	verde

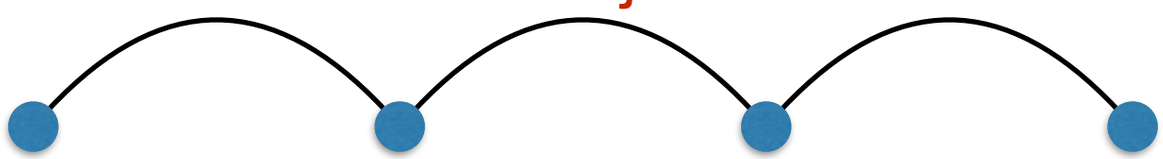


"Reverse" Translation

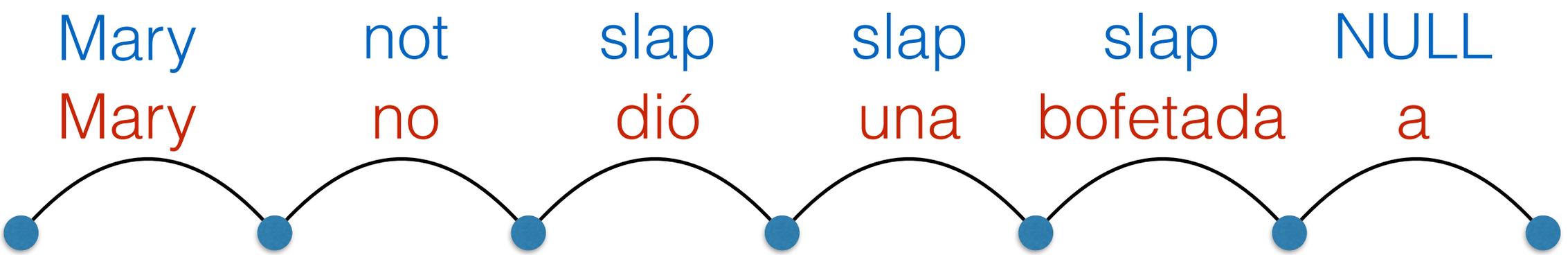
Mary	not	slap	slap	slap	a
Mary	no	dió	una	bofetada	a

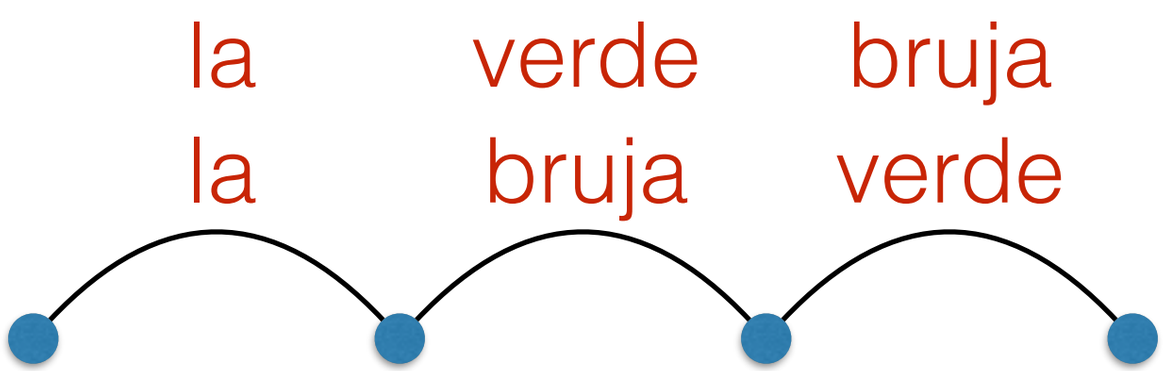


la	verde	bruja
la	bruja	verde

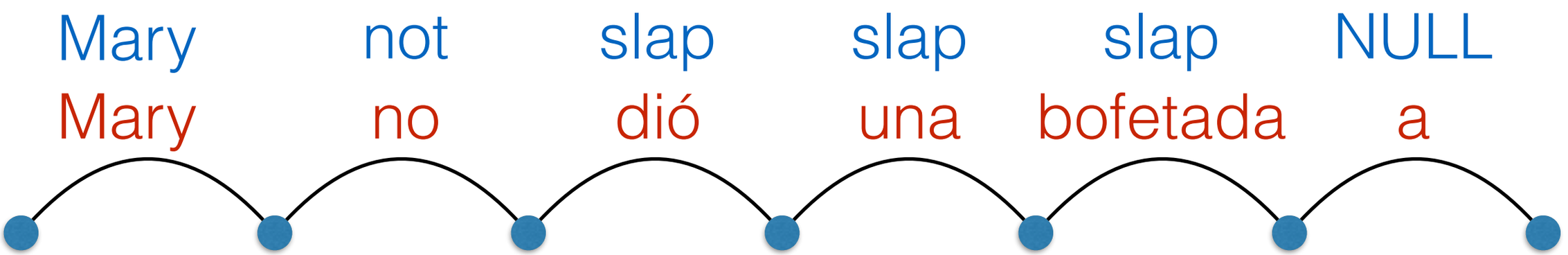


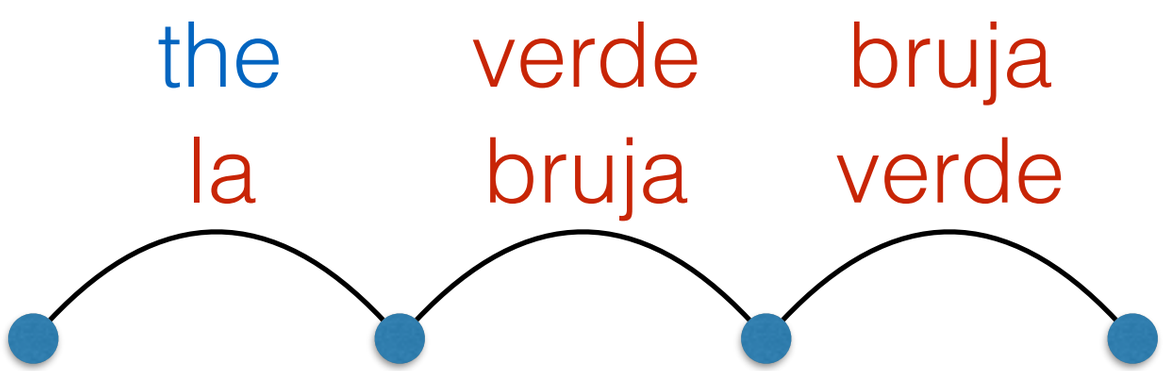
"Reverse" Translation



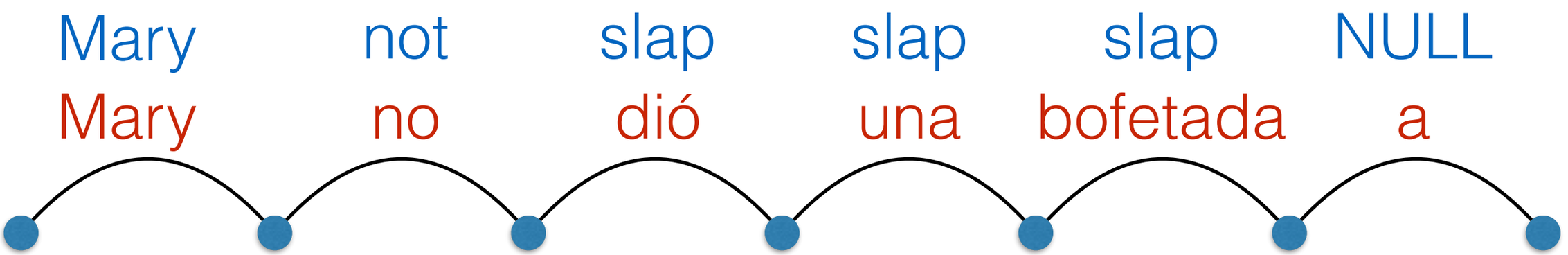


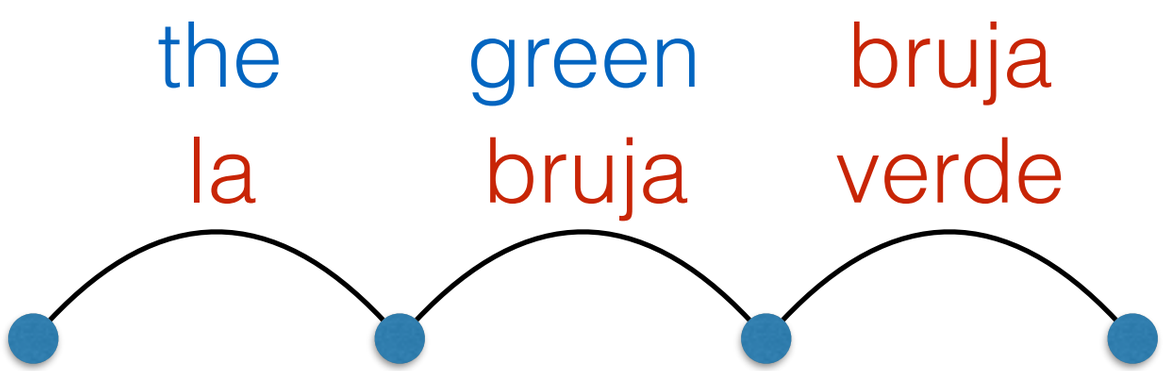
"Reverse" Translation



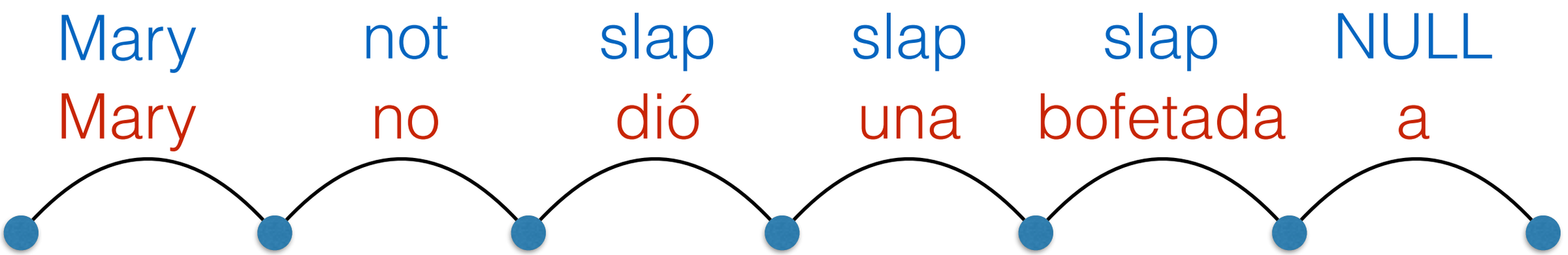


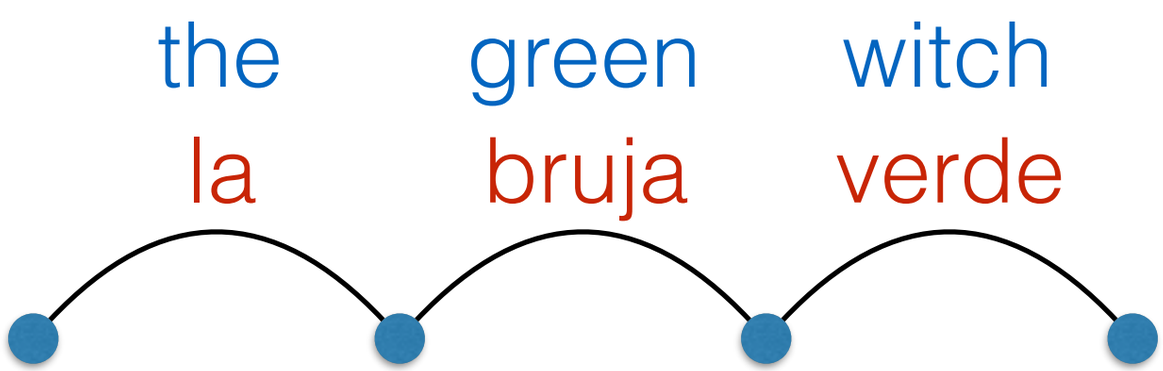
"Reverse" Translation



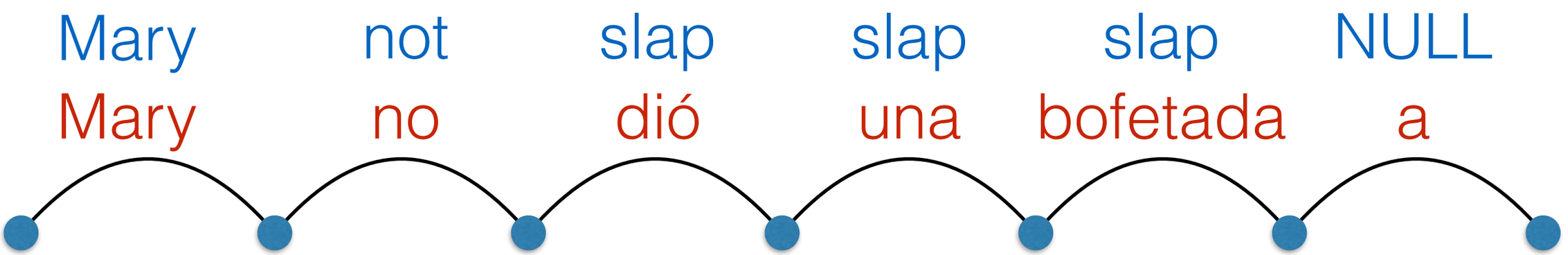


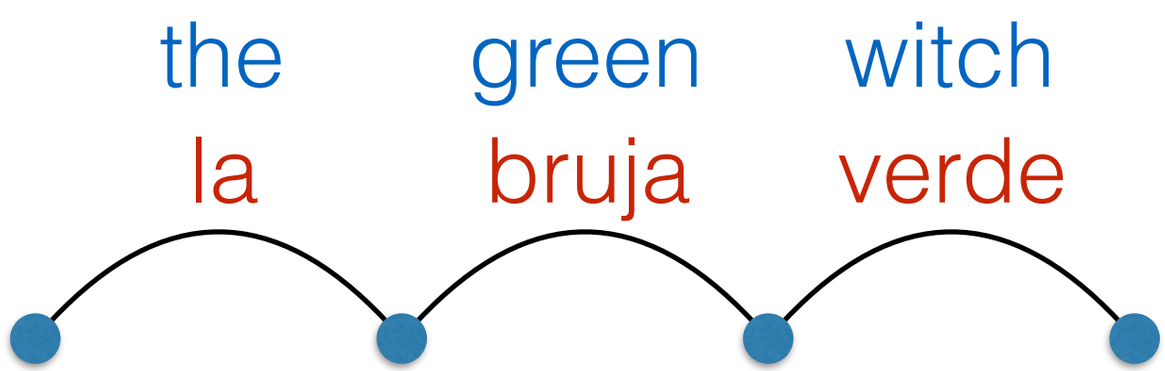
"Reverse" Translation



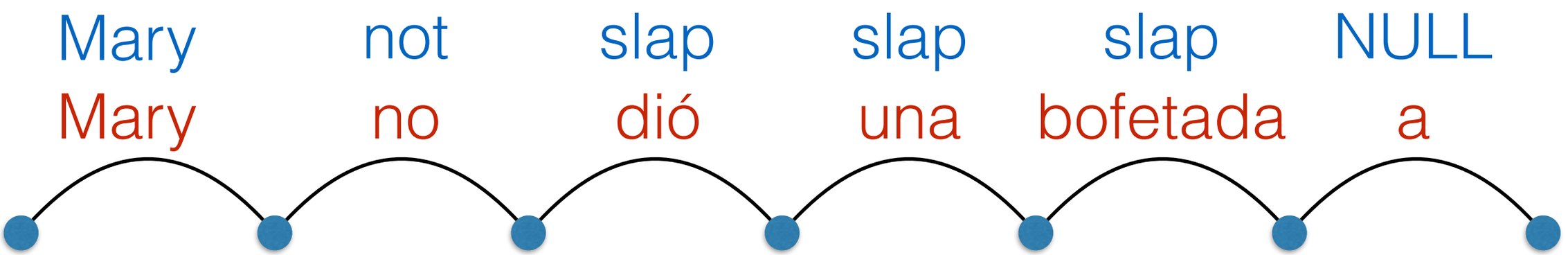


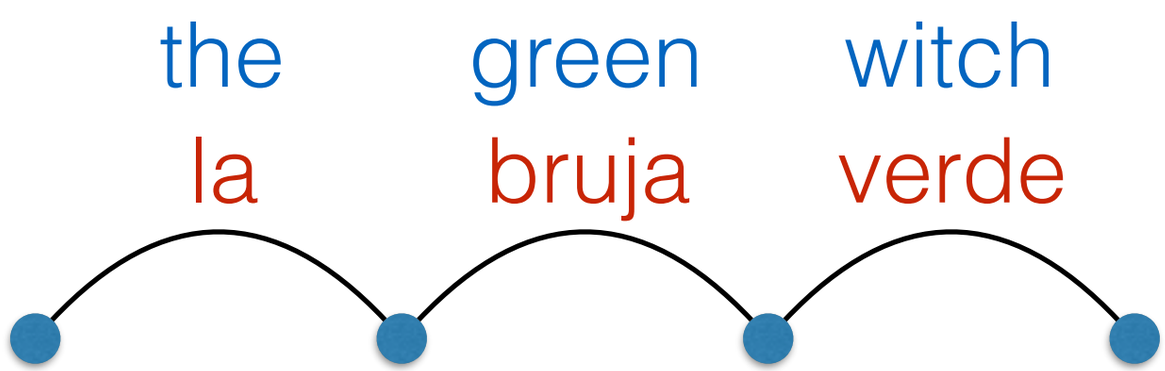
"Reverse" NULL insertion



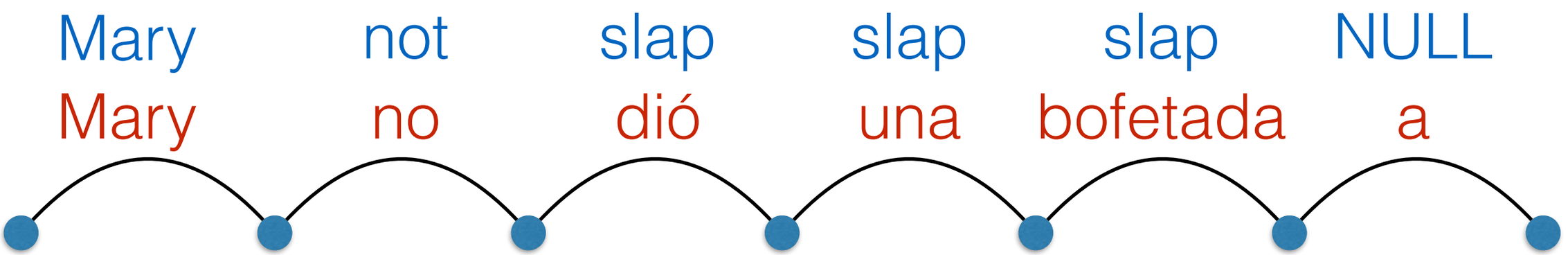


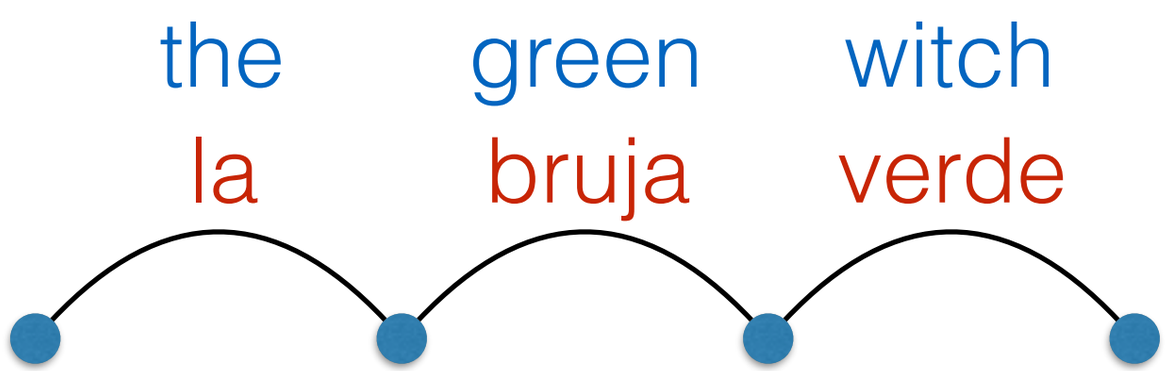
"Reverse" NULL insertion



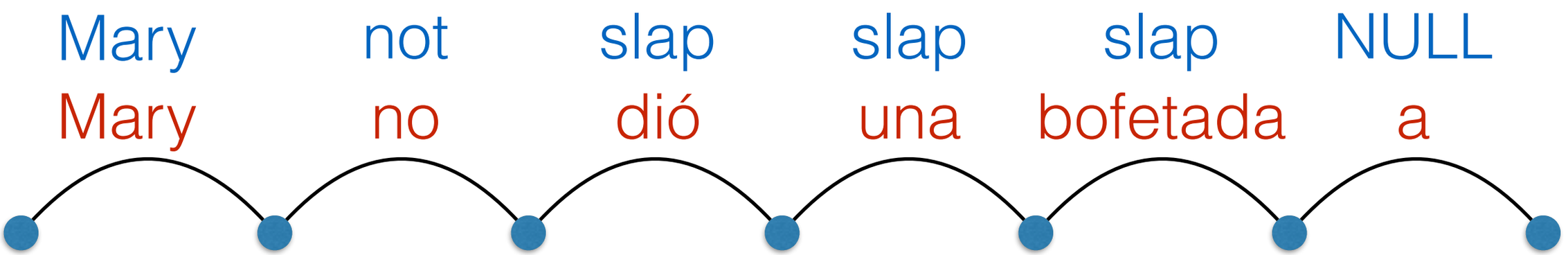


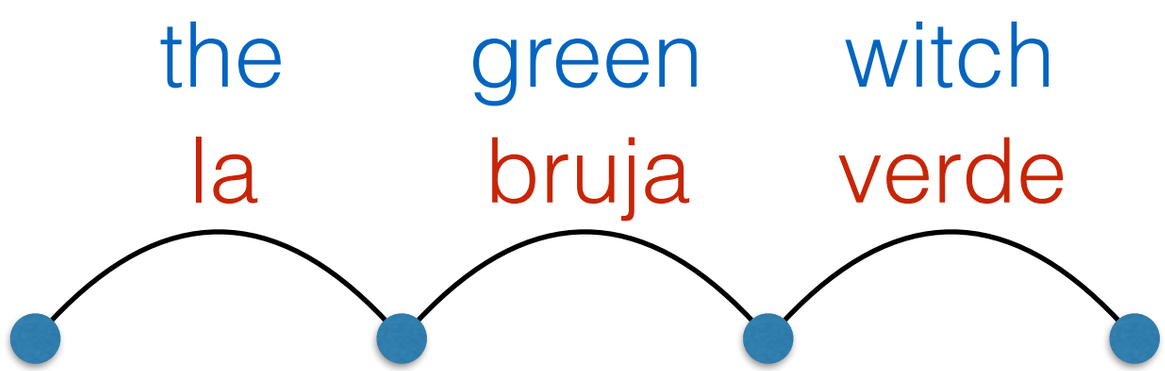
"Reverse" NULL insertion



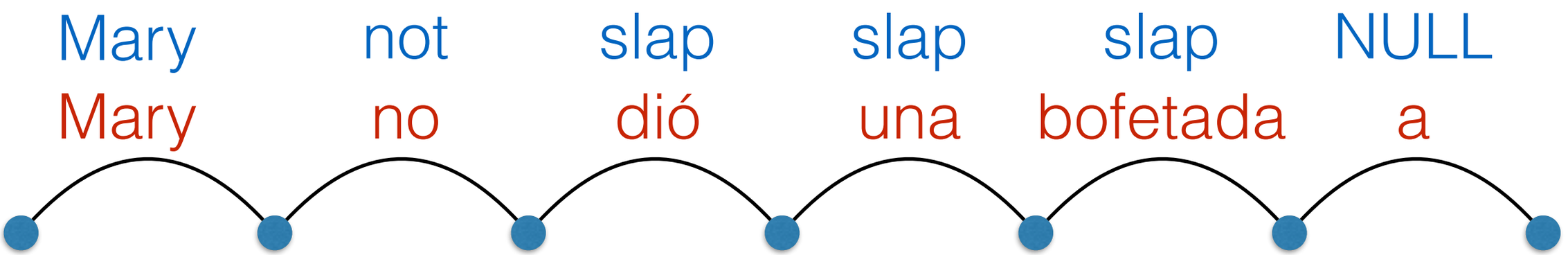


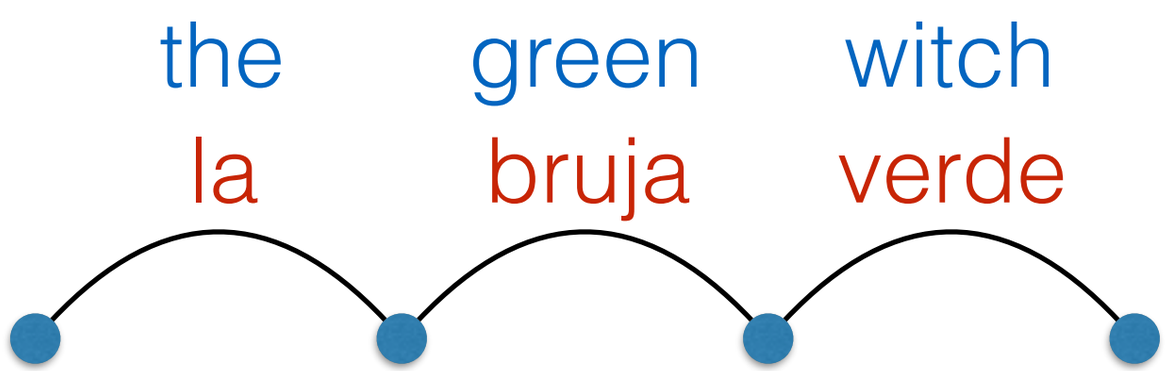
"Reverse" NULL insertion



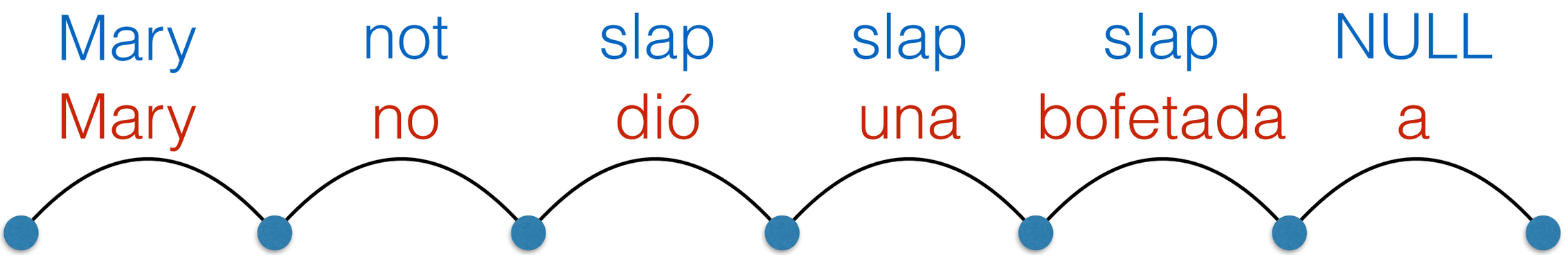


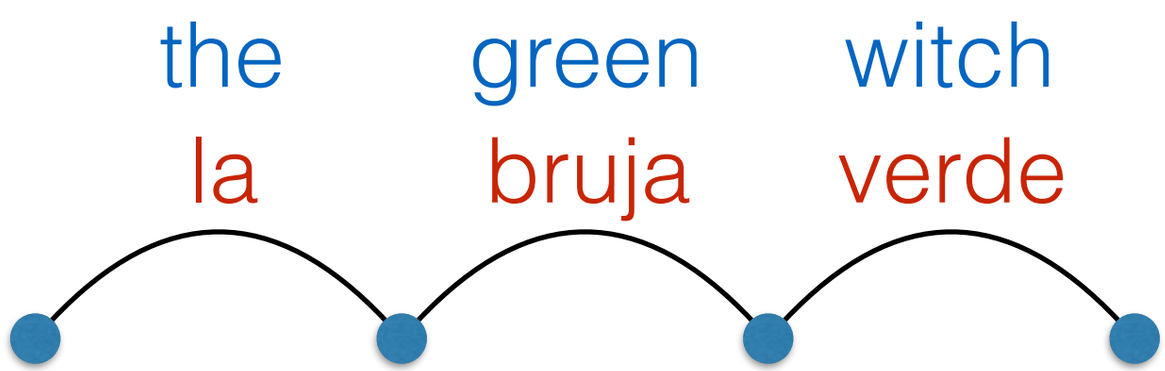
"Reverse" NULL insertion



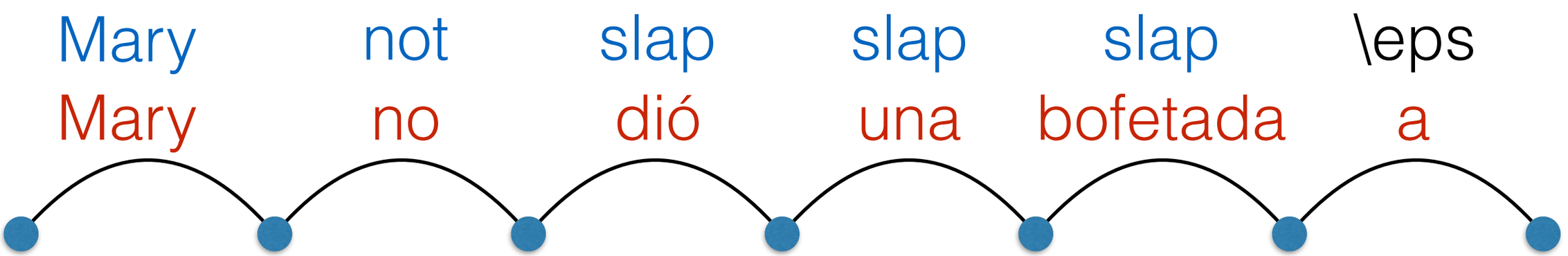


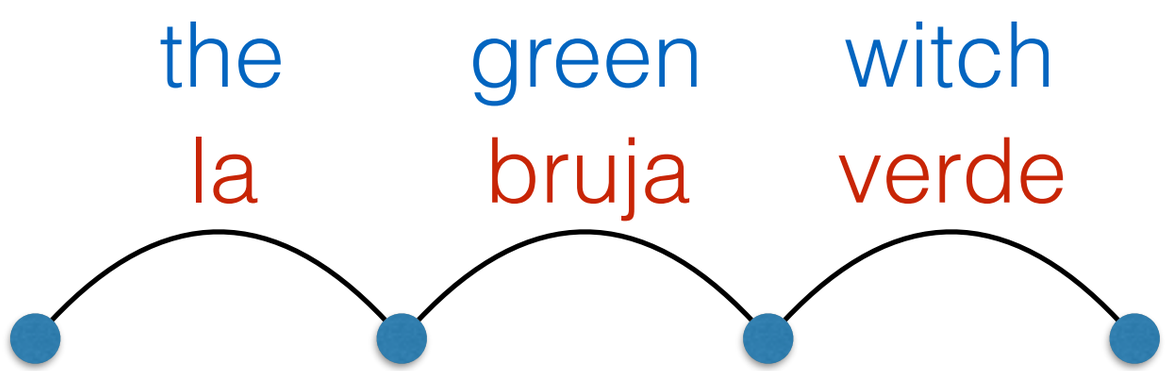
"Reverse" NULL insertion



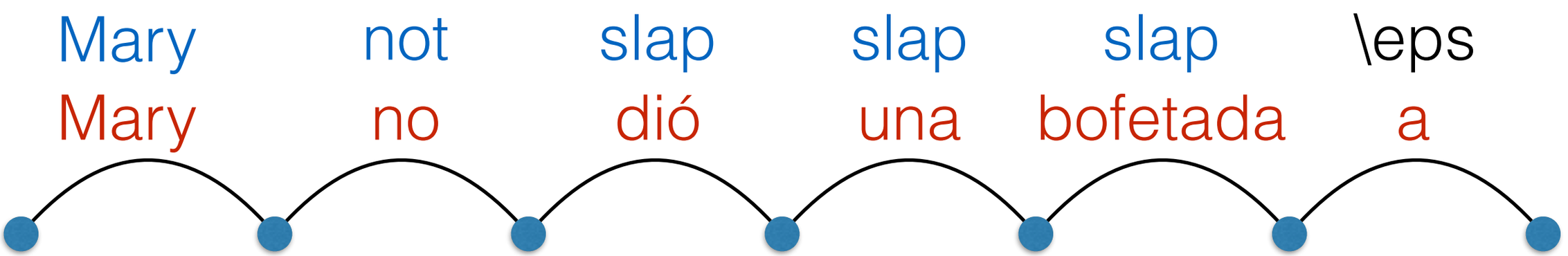


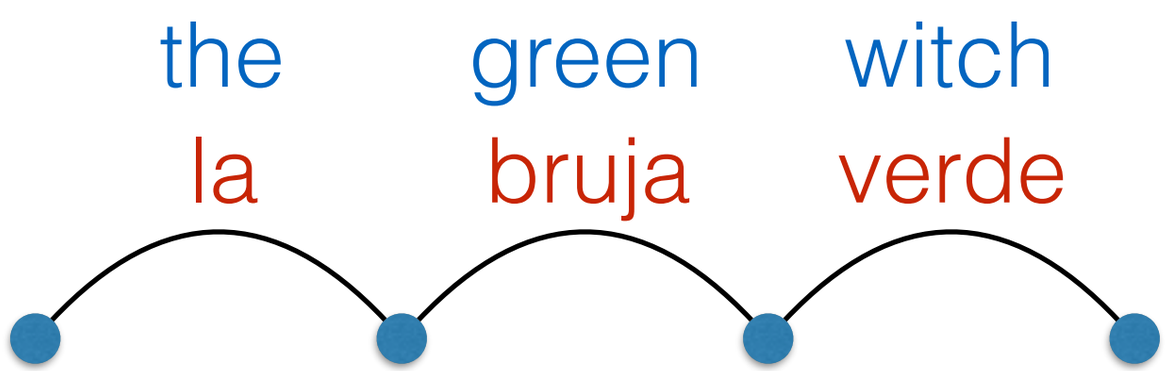
"Reverse" NULL insertion



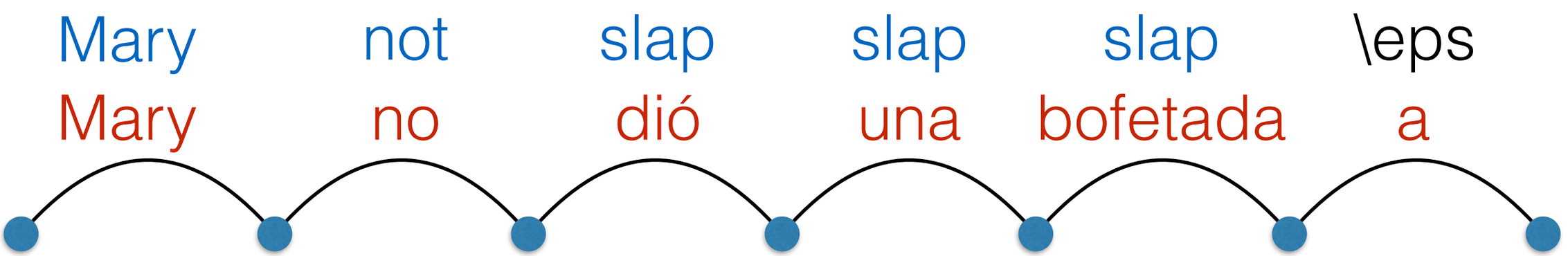


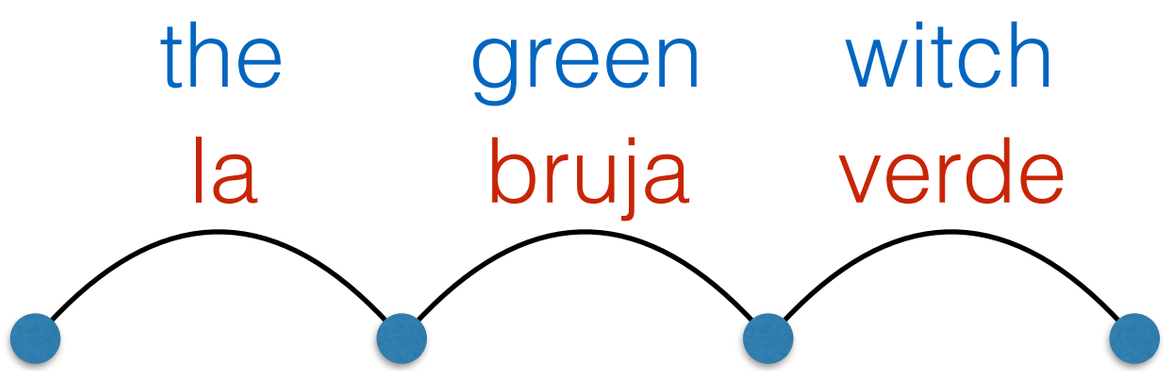
"Reverse" NULL insertion



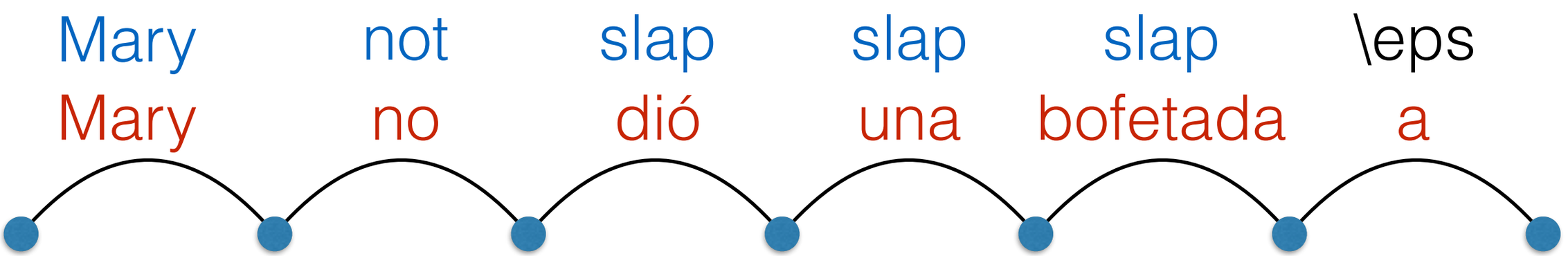


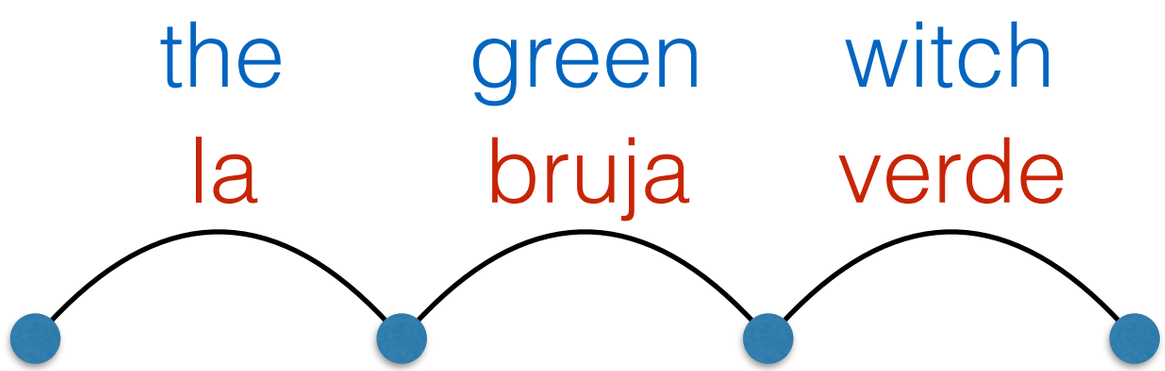
"Reverse" NULL insertion



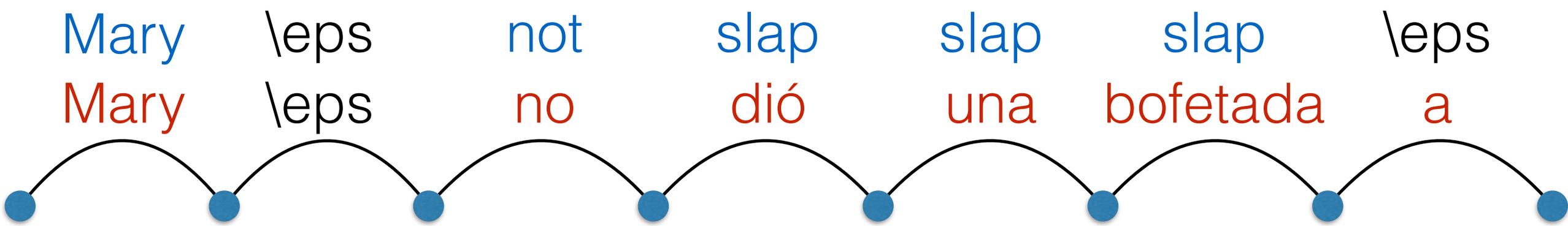


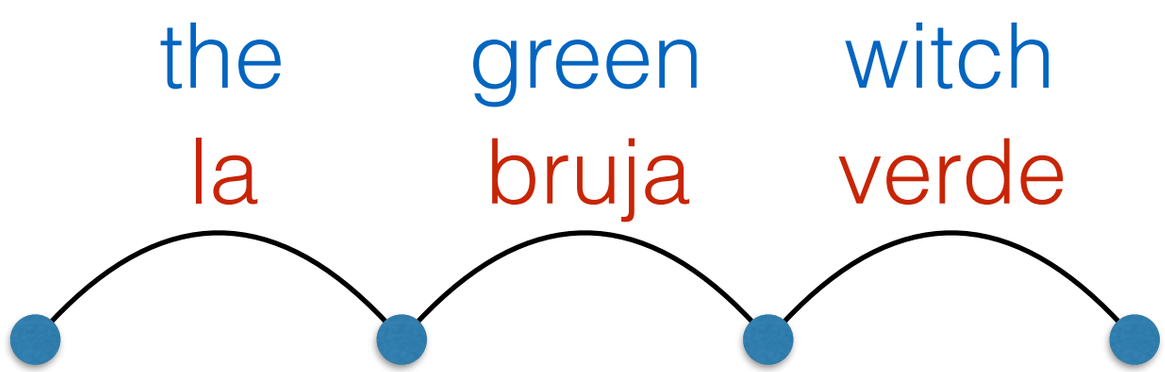
"Reverse" NULL insertion



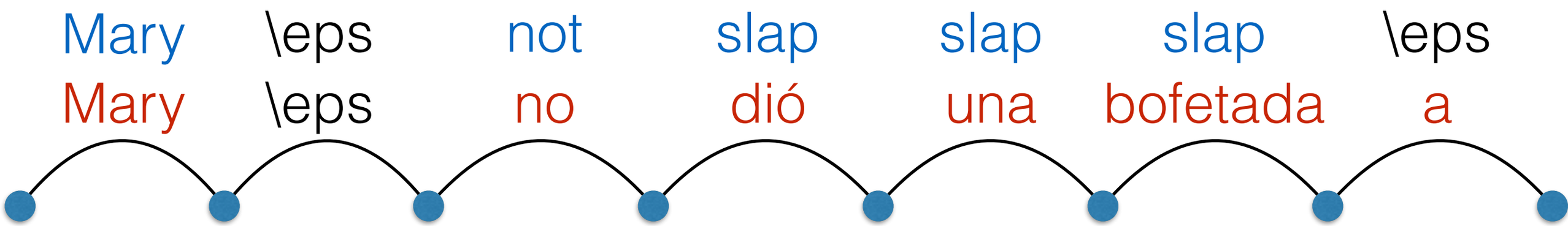


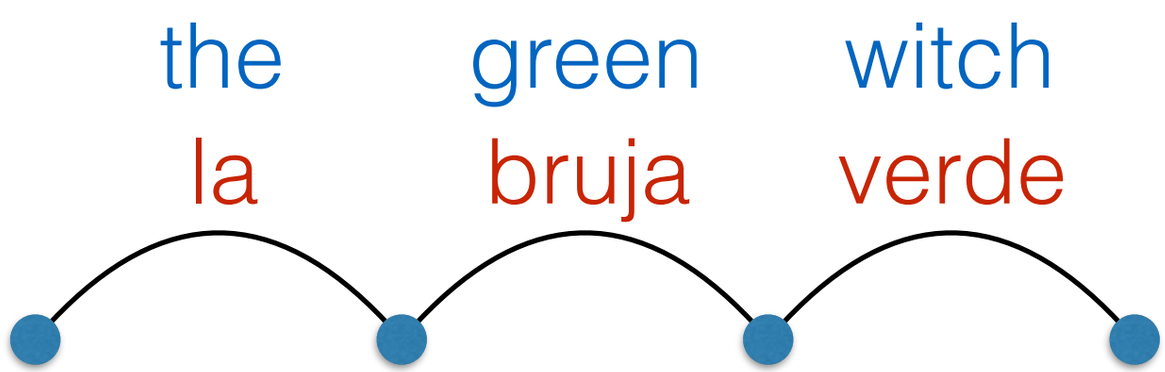
"Reverse" Fertility



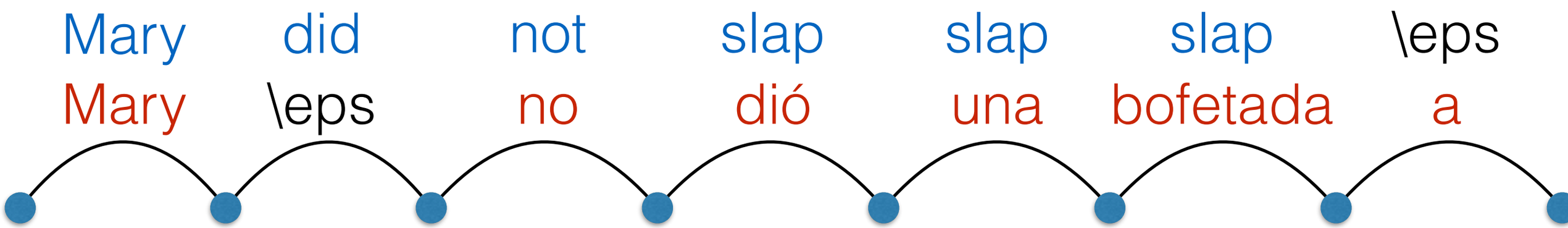


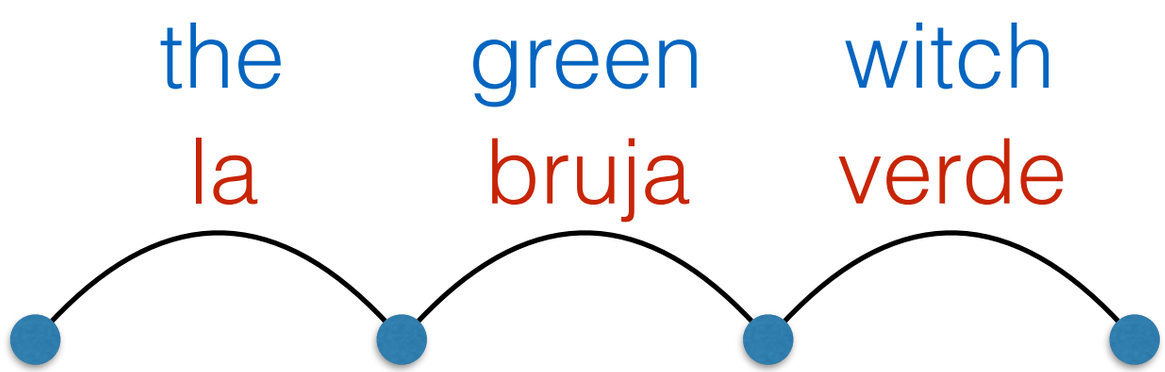
"Reverse" Fertility



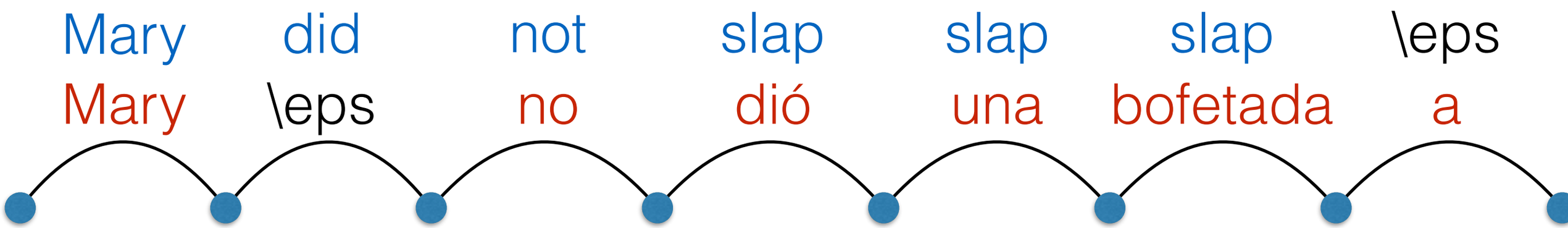


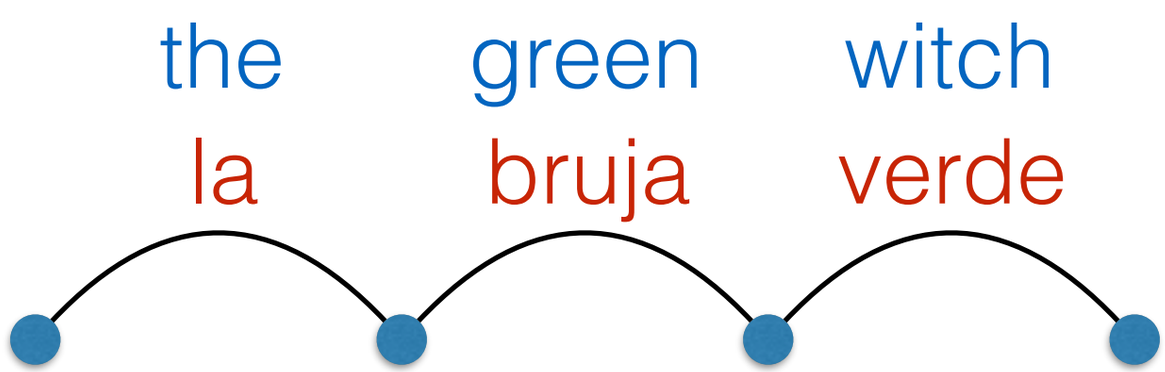
"Reverse" Fertility



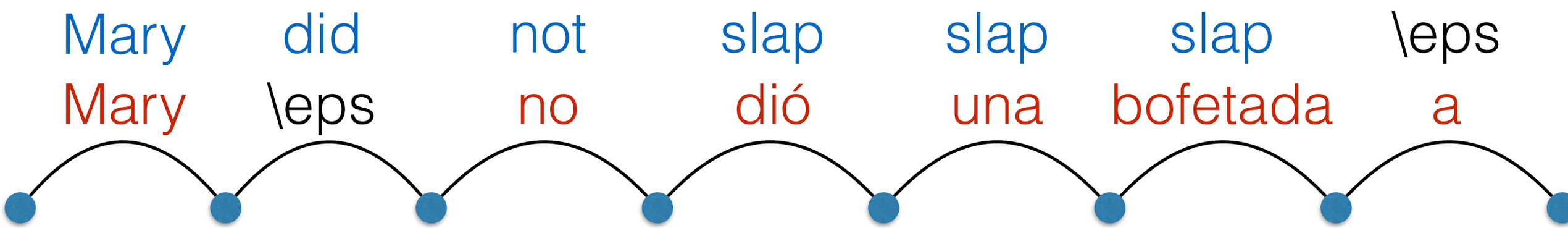


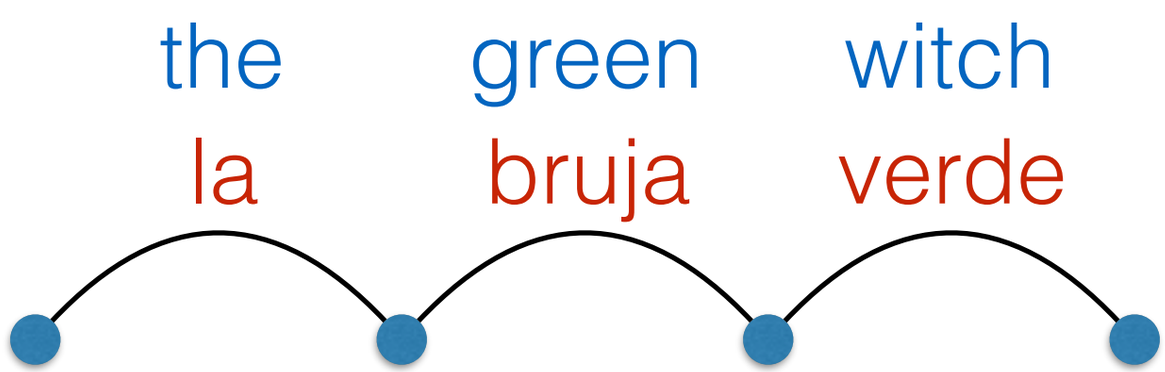
"Reverse" Fertility



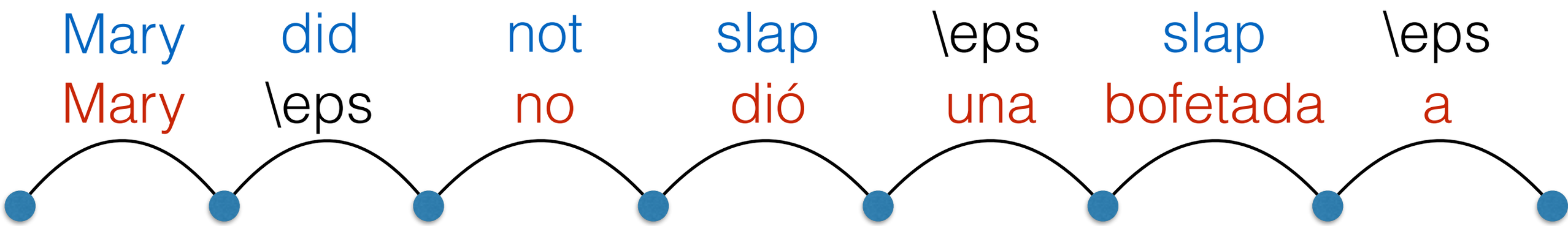


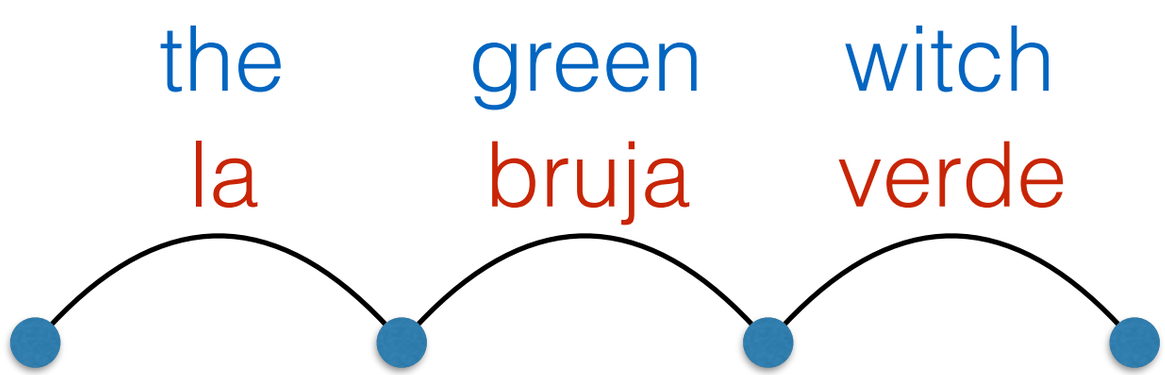
"Reverse" Fertility



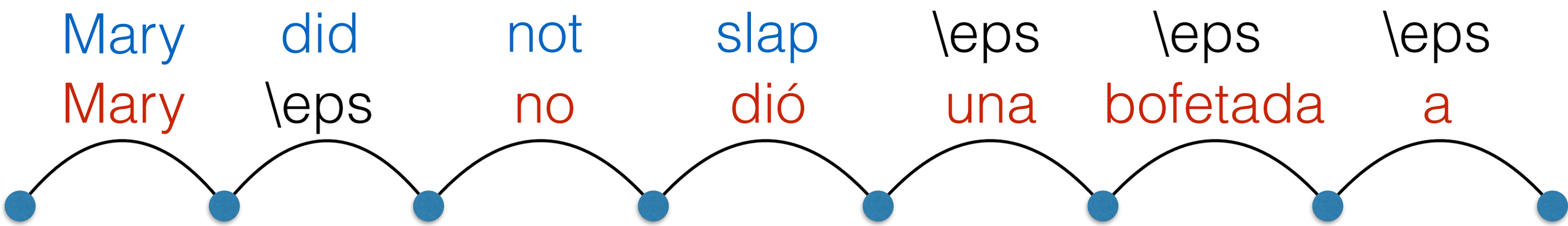


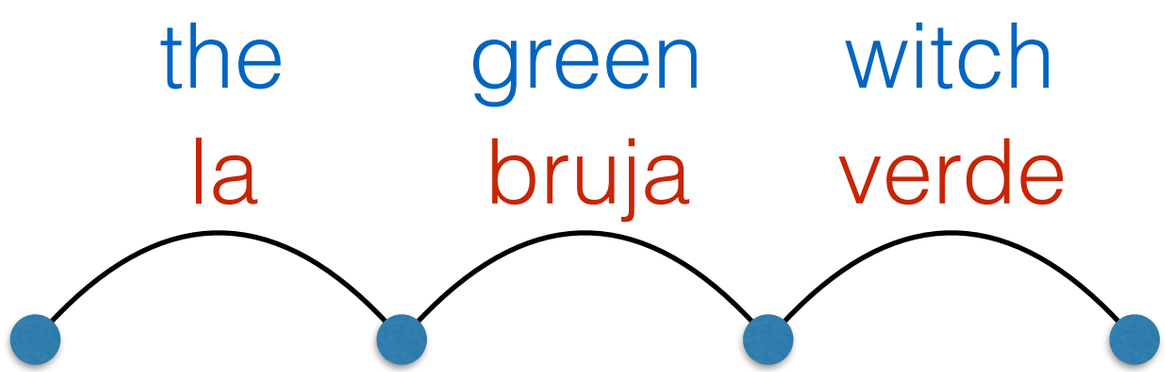
"Reverse" Fertility



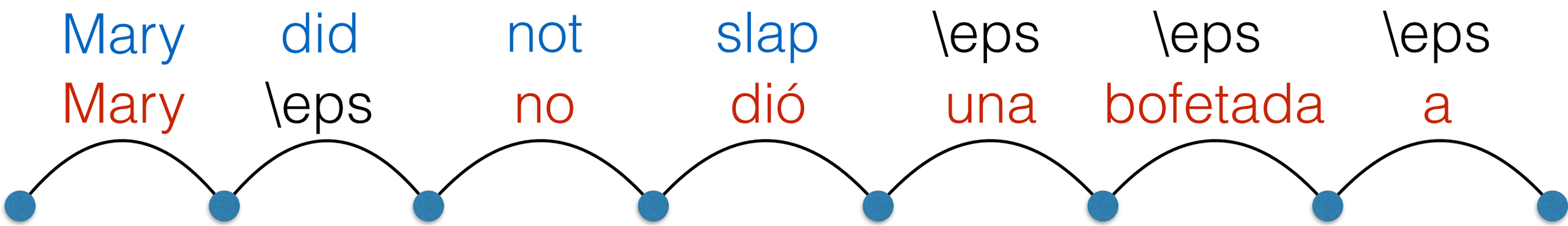


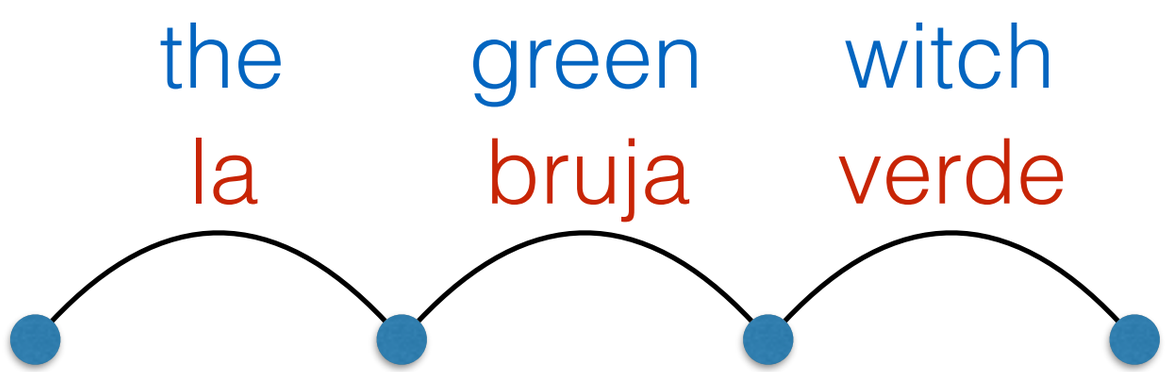
"Reverse" Fertility



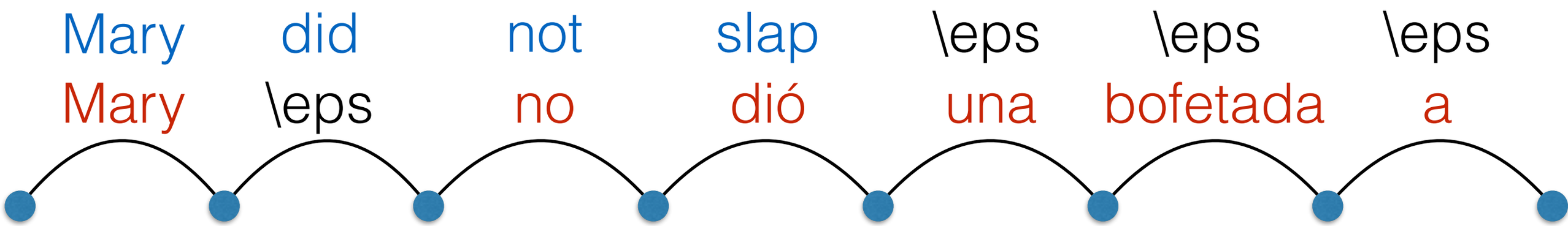


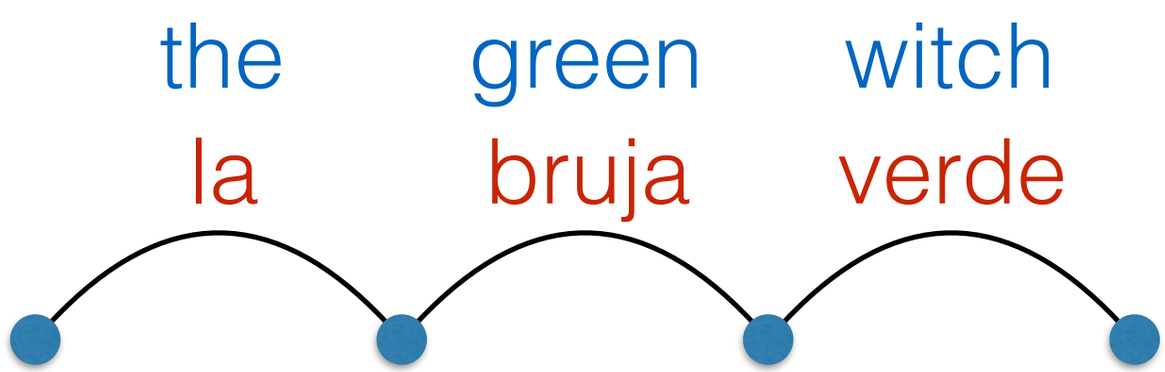
"Reverse" Fertility



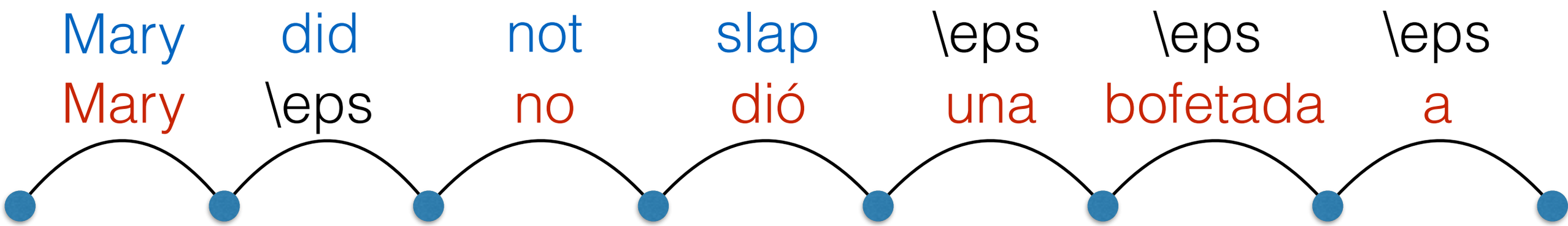


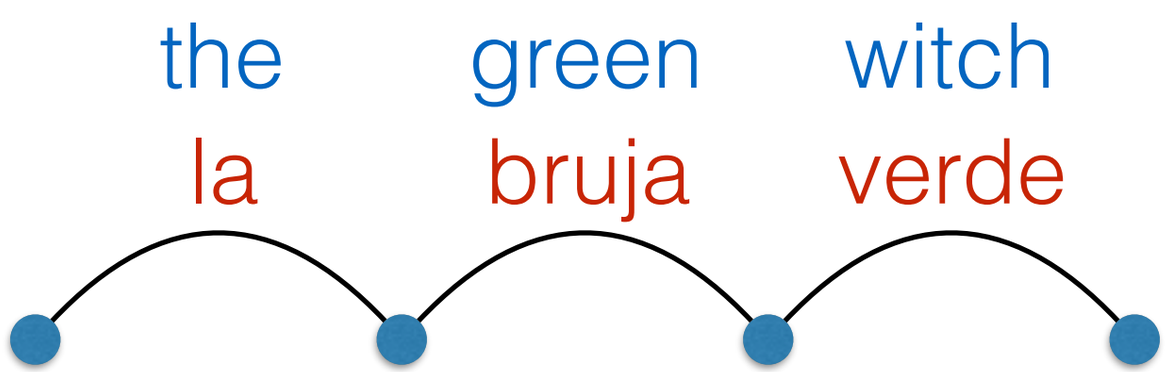
"Reverse" Fertility



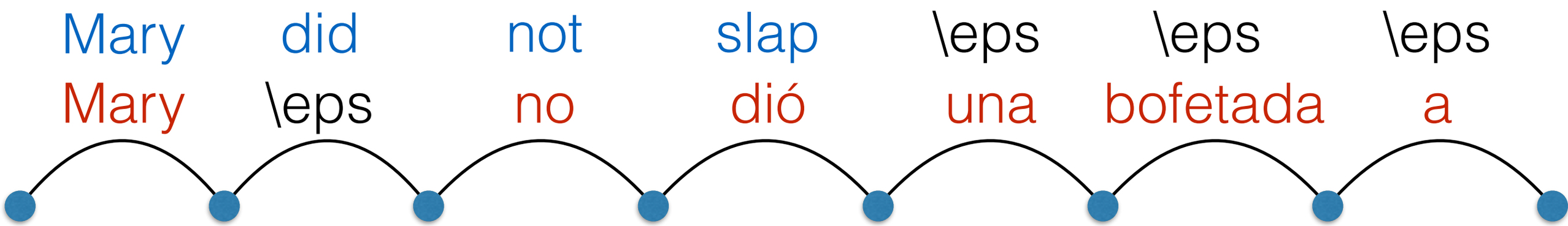


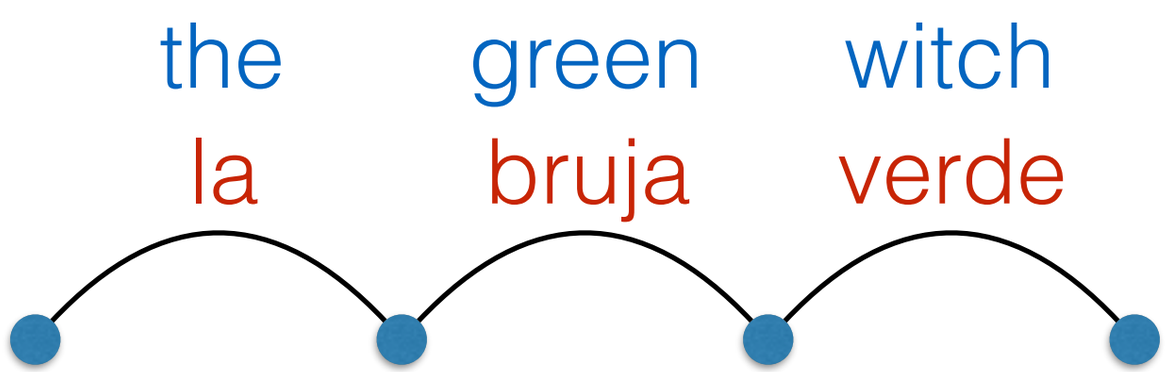
"Reverse" Fertility





"Reverse" Fertility





Word-based translation and FSTs

Language model

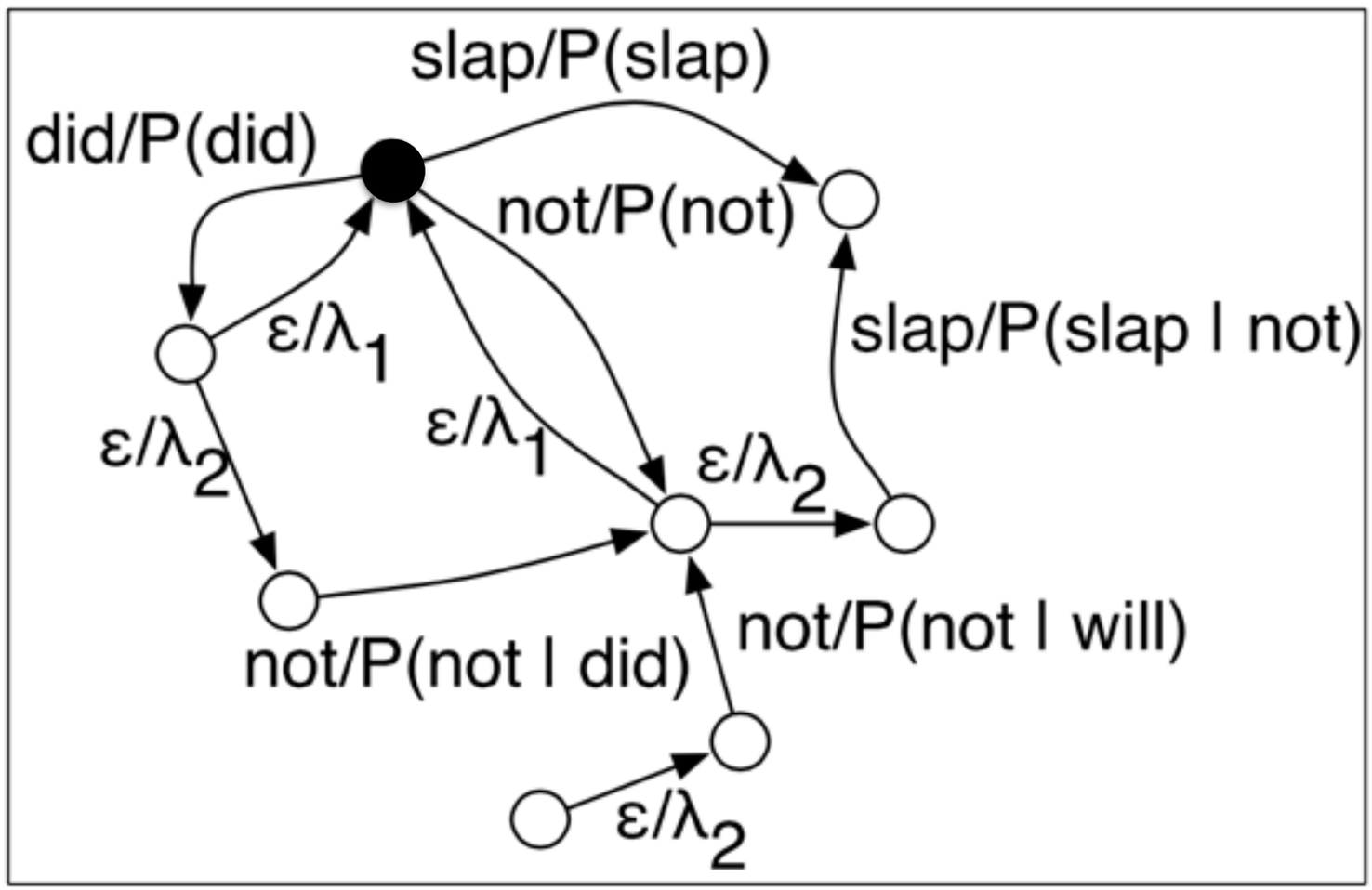
Fertility

Null insertion

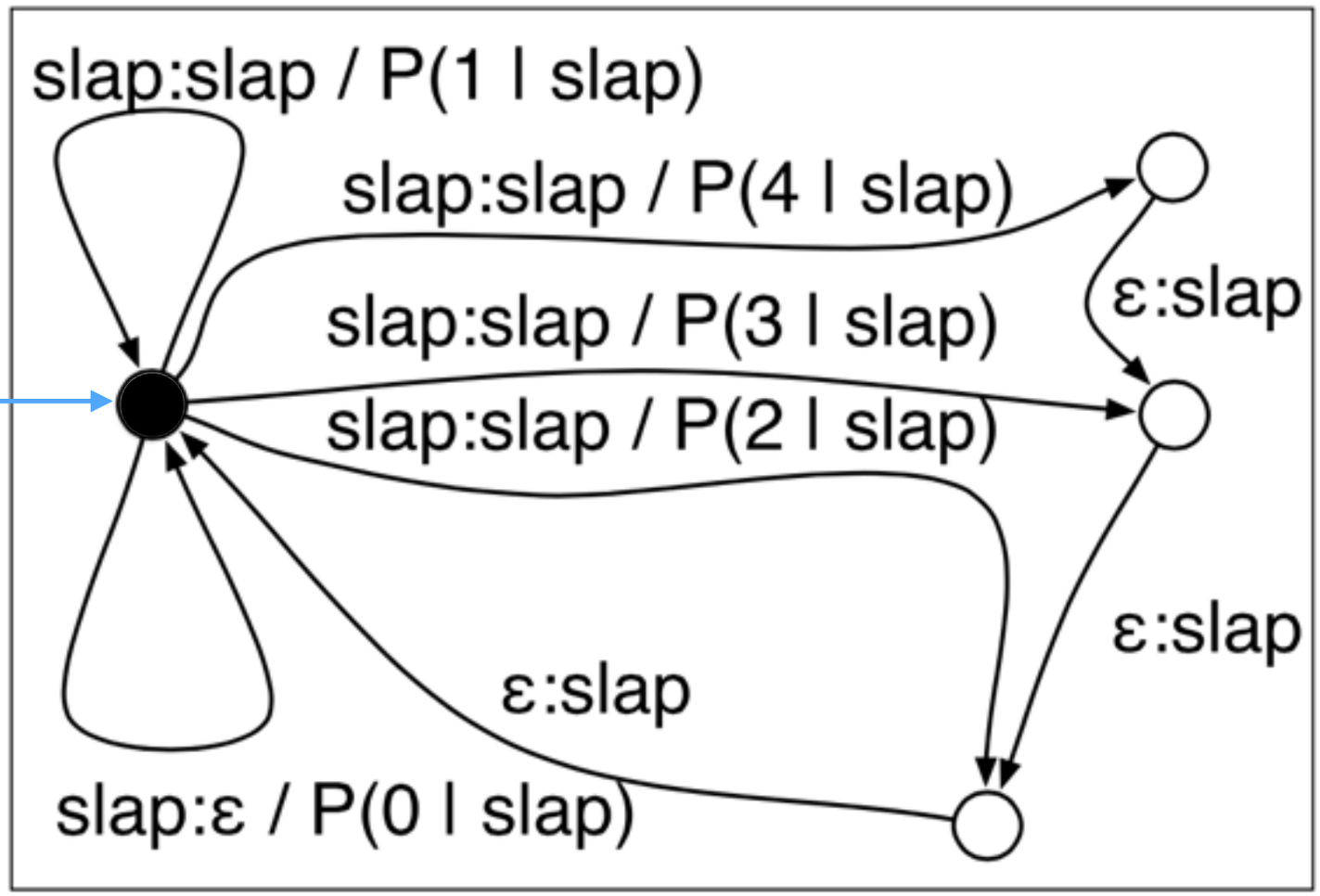
Translation

Distortion

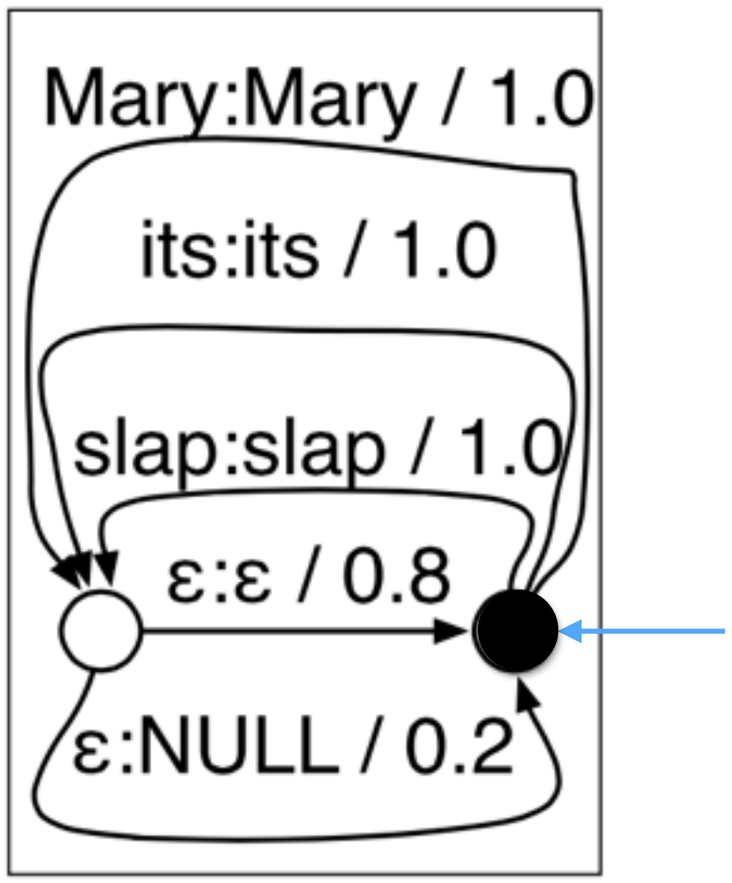
Language Model



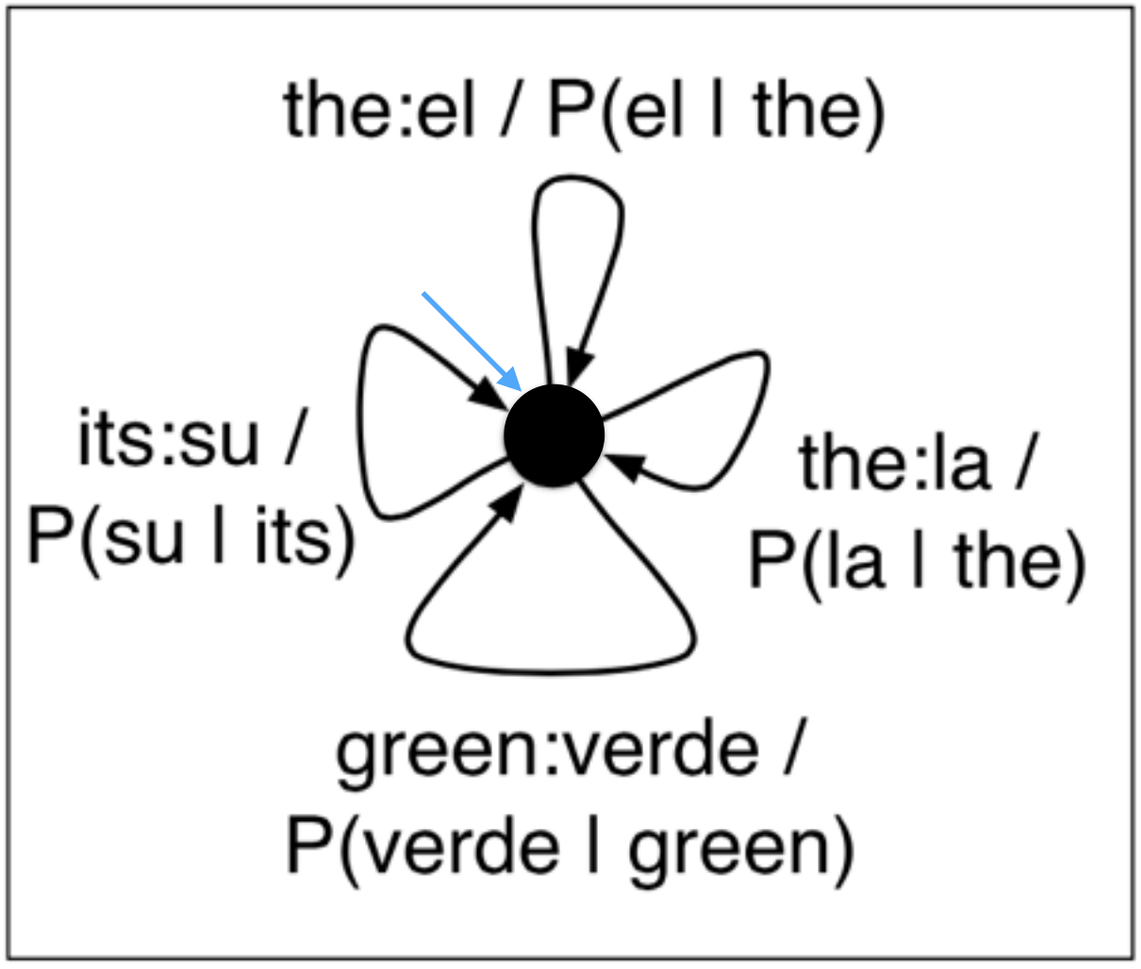
Fertility


slap:slap / P(1 | slap)
slap:slap / P(4 | slap)
slap:slap / P(3 | slap)
slap:slap / P(2 | slap)
ε:slap
ε:slap
ε:slap
slap:ε / P(0 | slap)

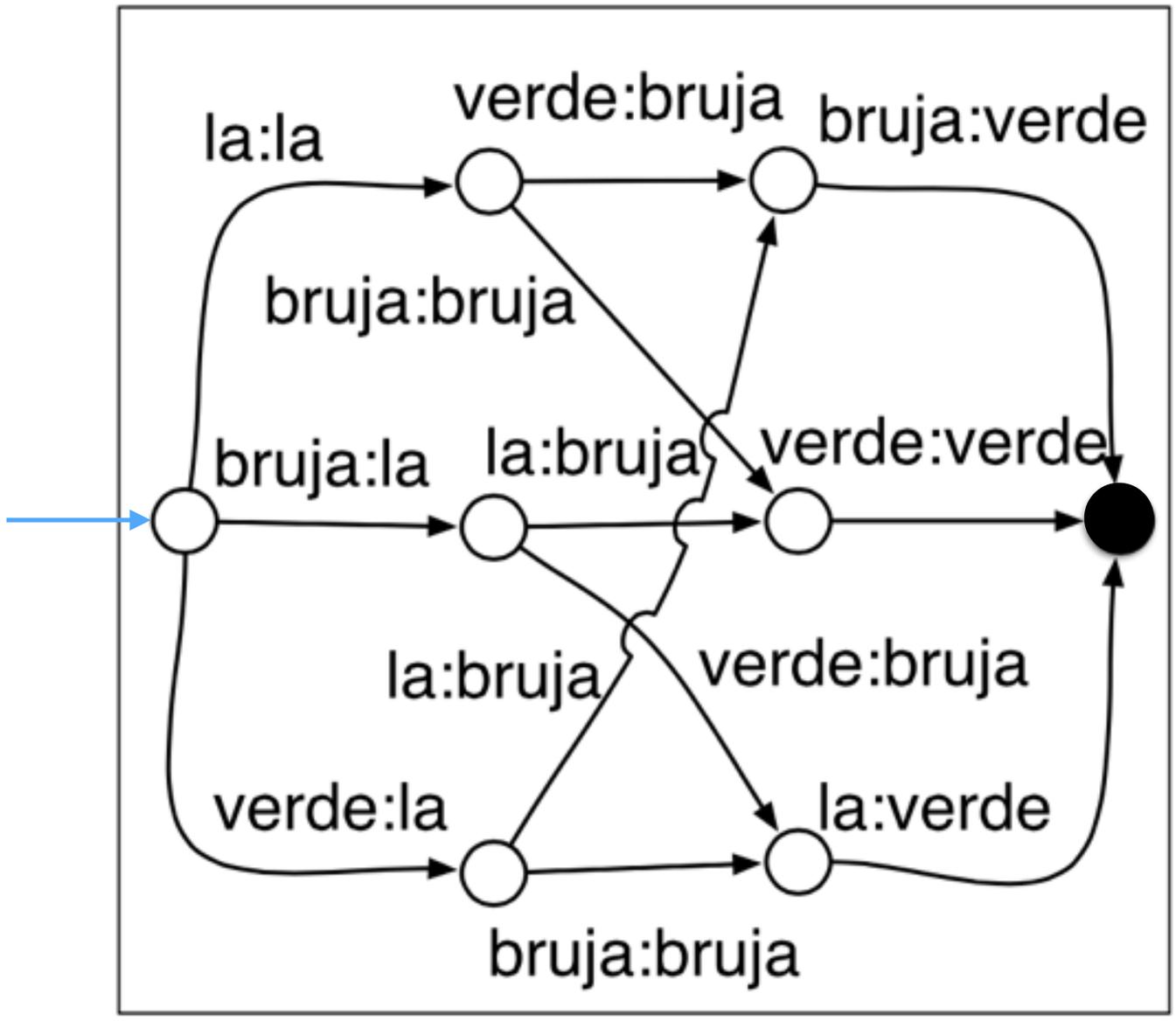
Null Insertion



Translation



Distortion


la:la
verde:bruja
bruja:verde
bruja:bruja
bruja:la
la:bruja
verde:verde
la:bruja
verde:bruja
verde:la
la:verde
bruja:bruja

Bibliography

- Brown, Peter F. and Pietra, Vincent J. Della and Pietra, Stephen A. Della and Mercer, Robert L.. 1993. The mathematics of statistical machine translation: parameter estimation. In *Computational Linguistics*. MIT Press.

- Kevin Knight and Yaser Al-Onaizan. 1998. Translation with finite-state devices. In *Proceedings of the Association for Machine Translation in the Americas*.

- Knight, Kevin and May, Jonathan. 2009. Applications of Weighted Automata in Natural Language Processing. In *Handbook of Weighted Automata*. Springer Berlin Heidelberg.