An Empirical Examination of Challenges in Chinese Parsing

THE UNIVERSITY OF SYDNEY

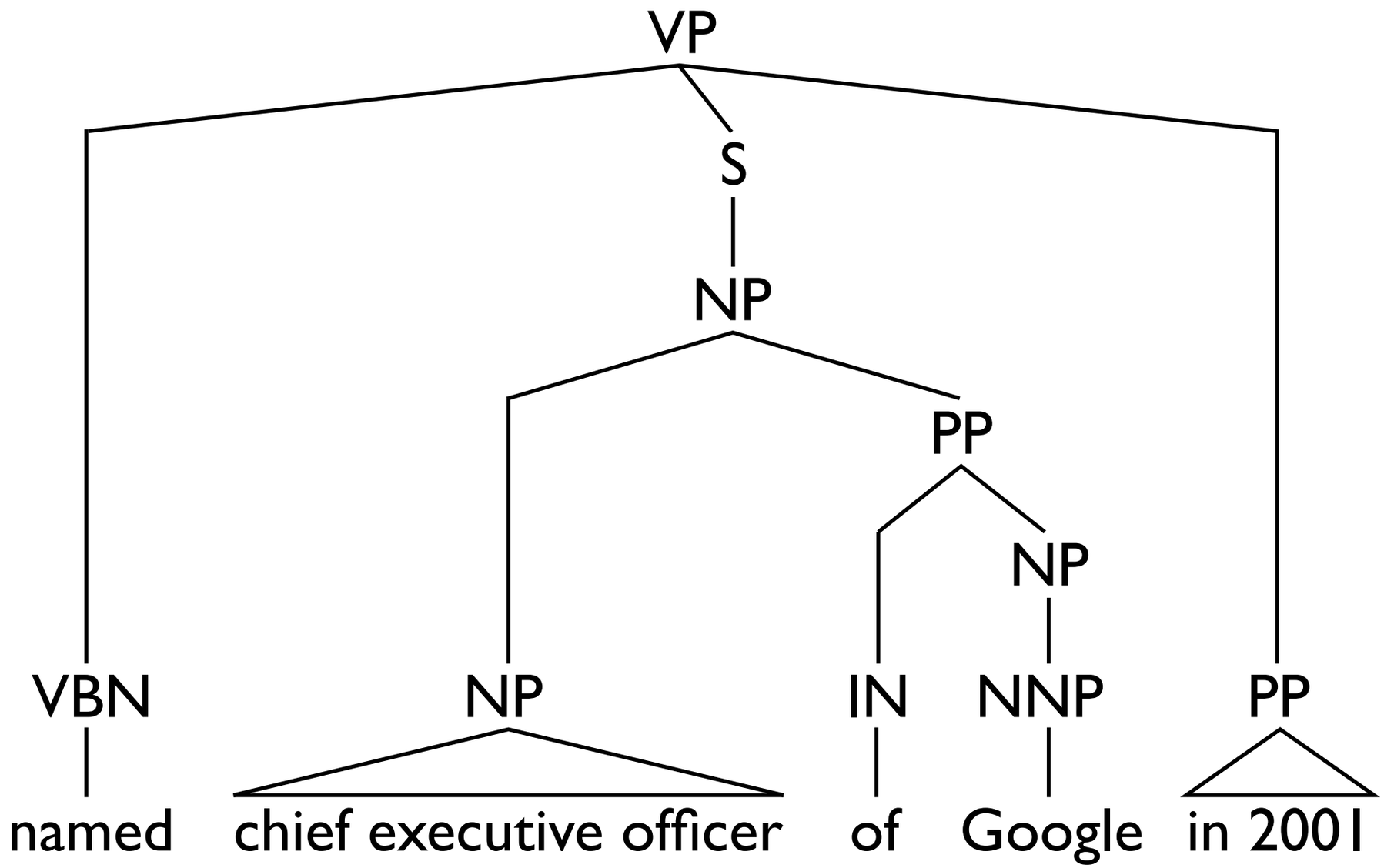
Jonathan Kummerfeld
Dan Klein
UC Berkeley

Daniel Tse
James Curran
U Sydney

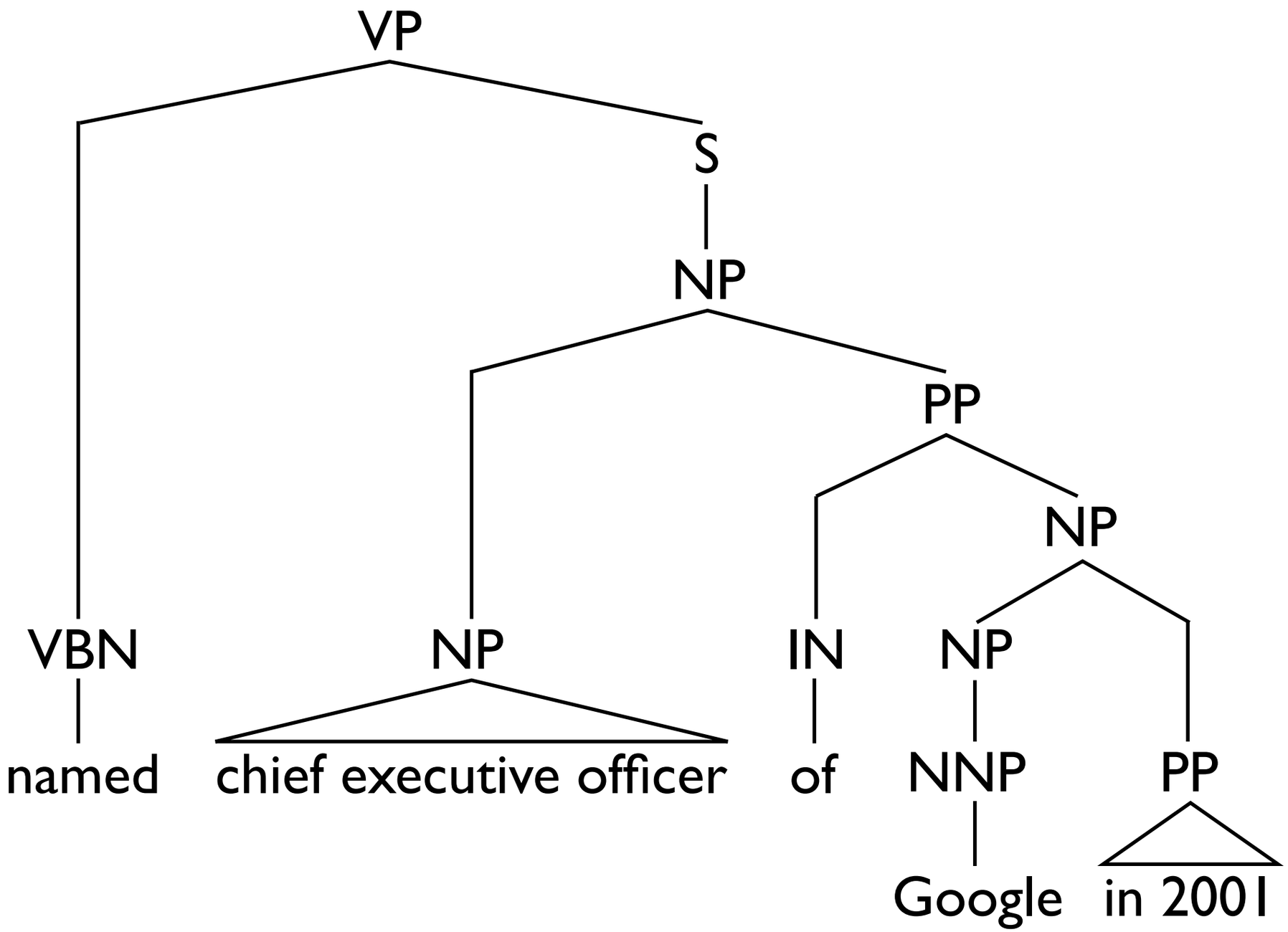
Parsing Errors

named chief executive officer of Google in 2001

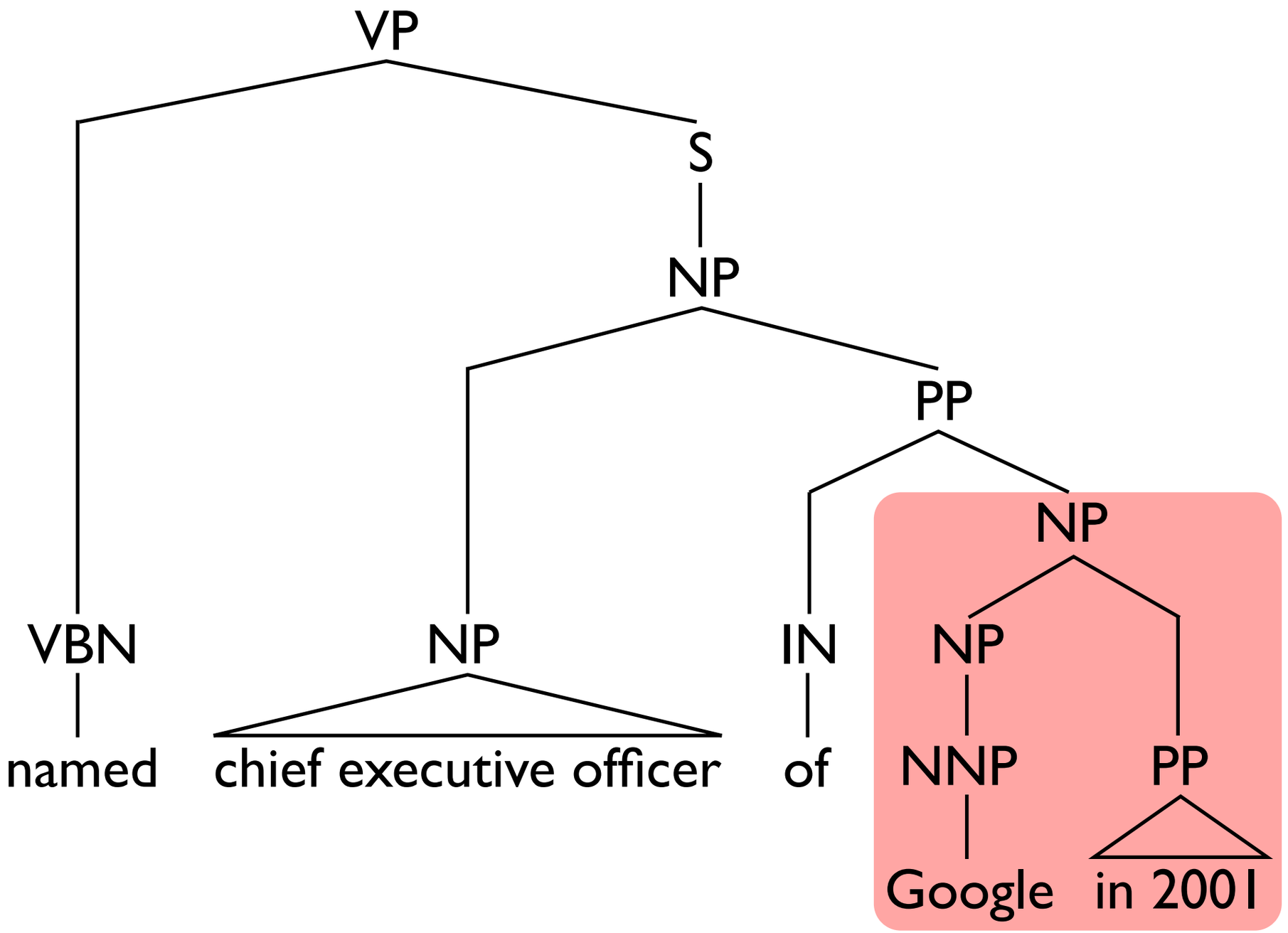
Parsing Errors



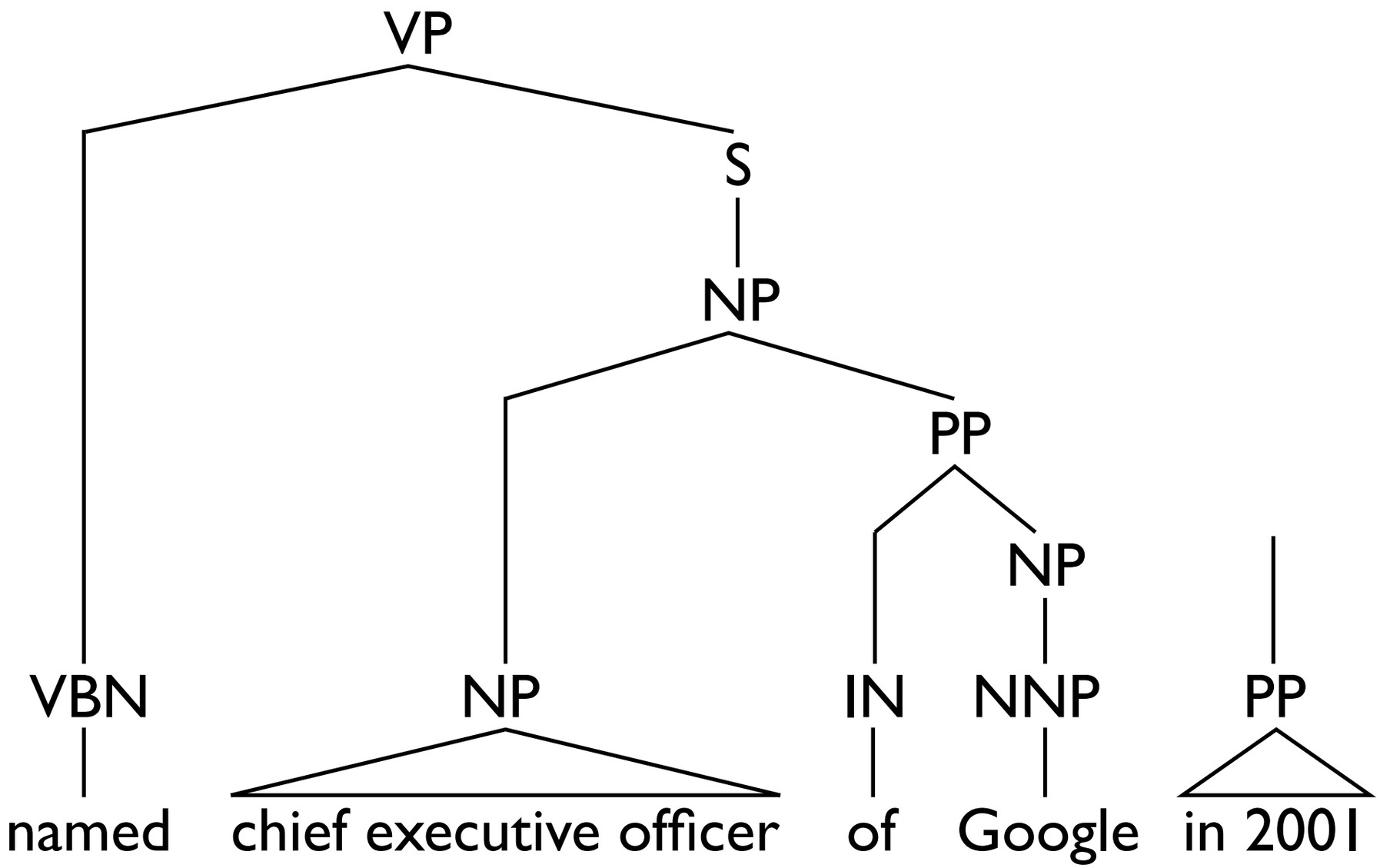
Parsing Errors



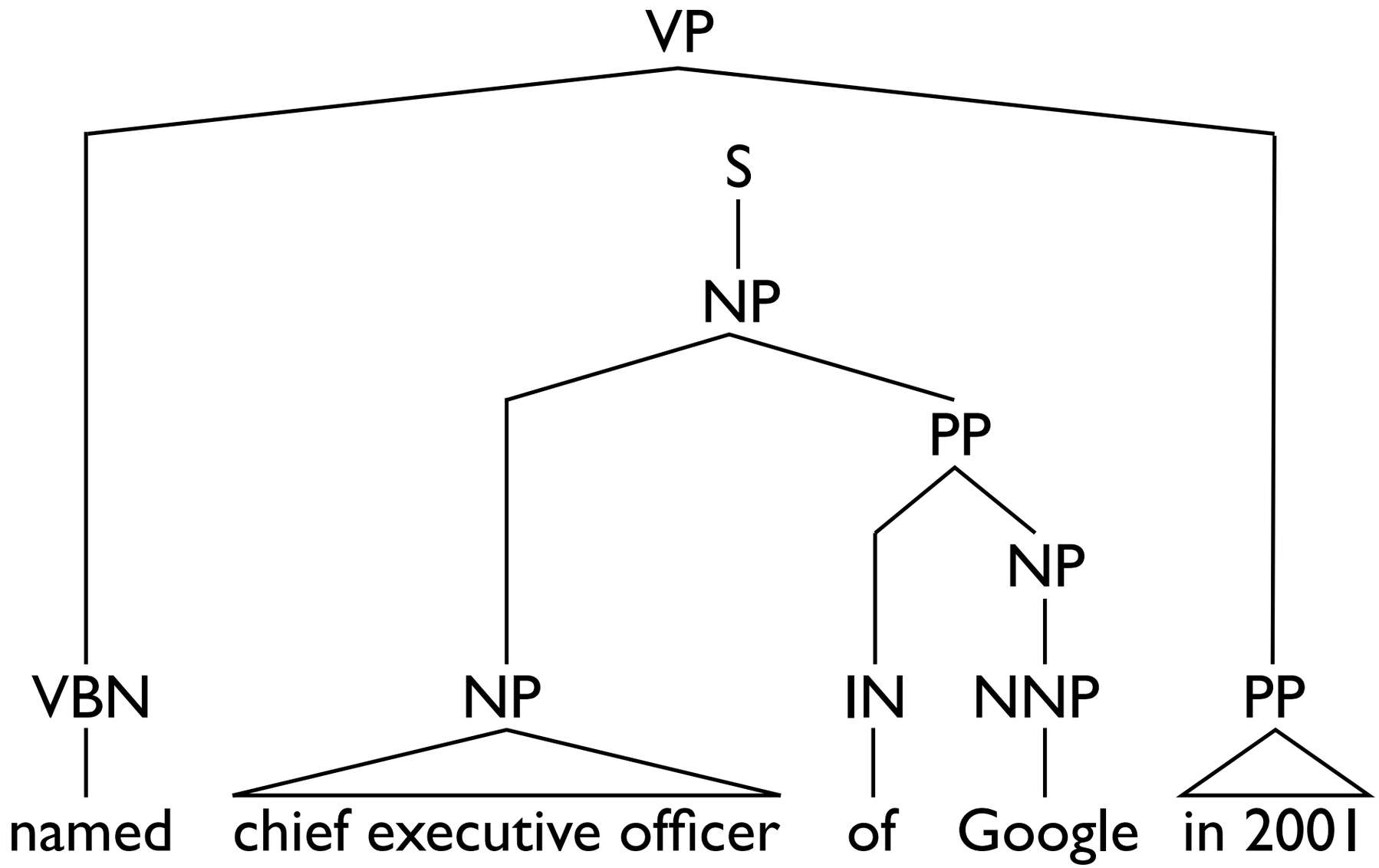
Parsing Errors



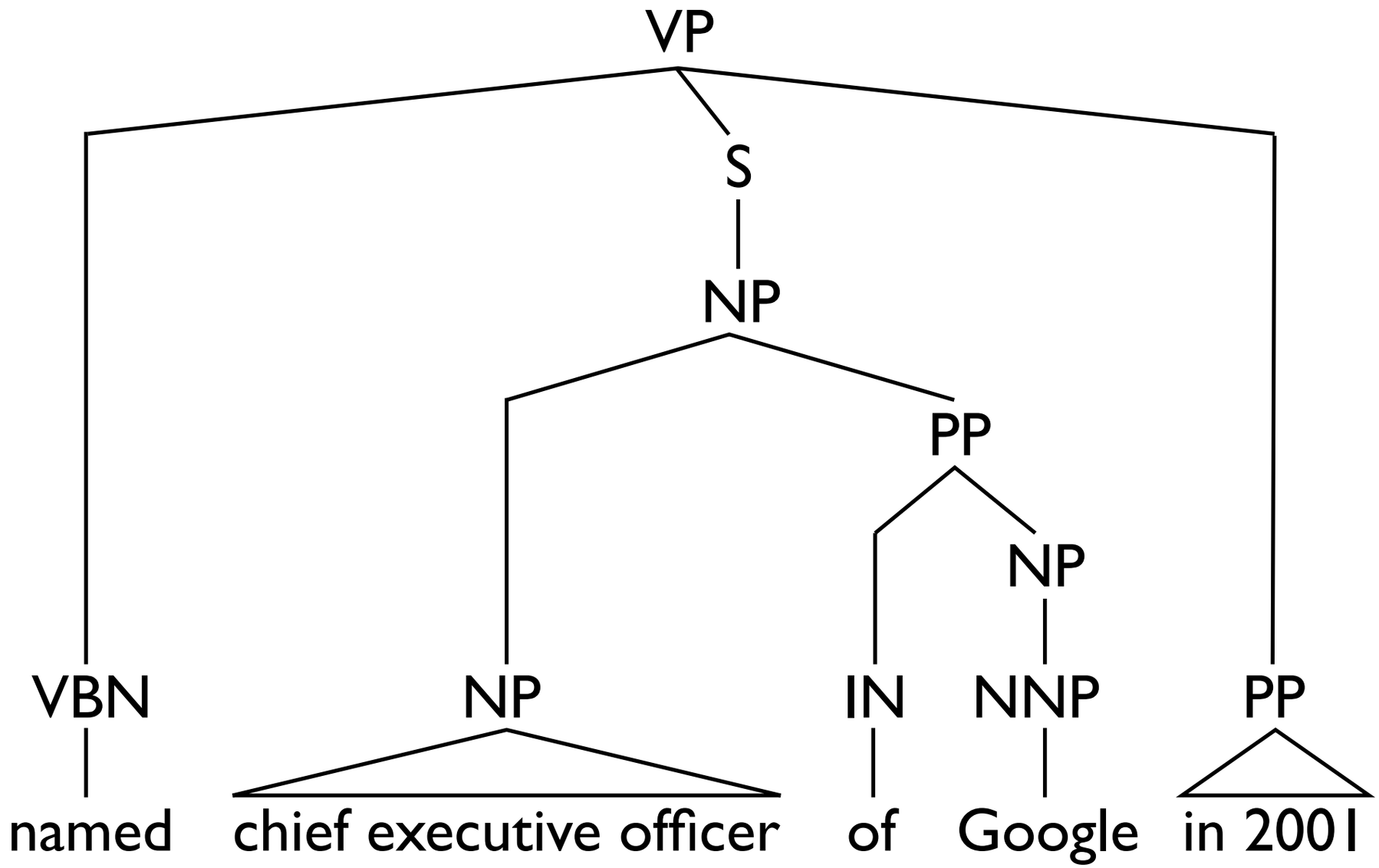
Parsing Errors

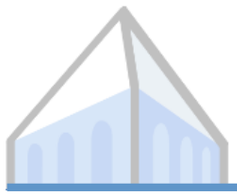


Parsing Errors



Parsing Errors





Manual Error Analysis

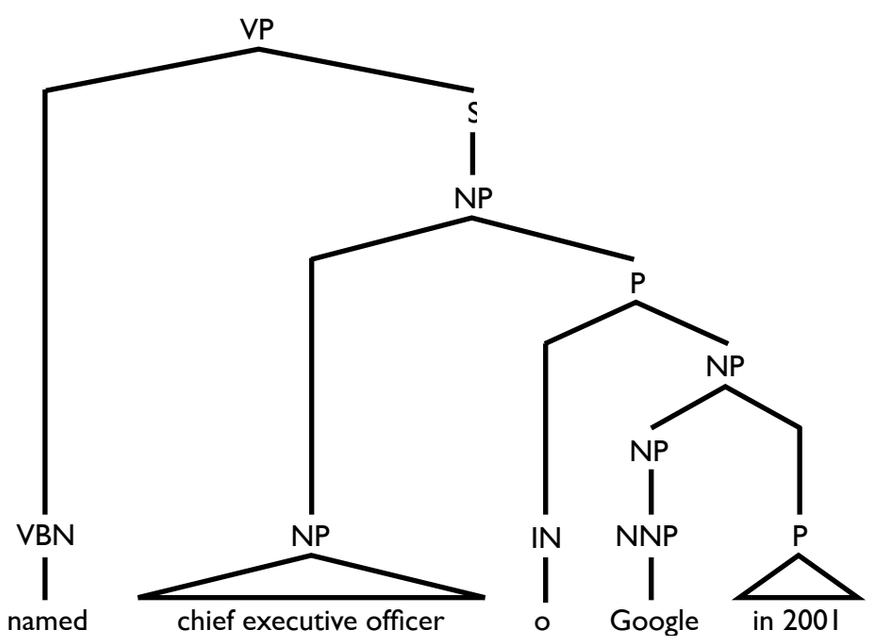
Manual Error Analysis

- Too slow!

Automatic Error Analysis

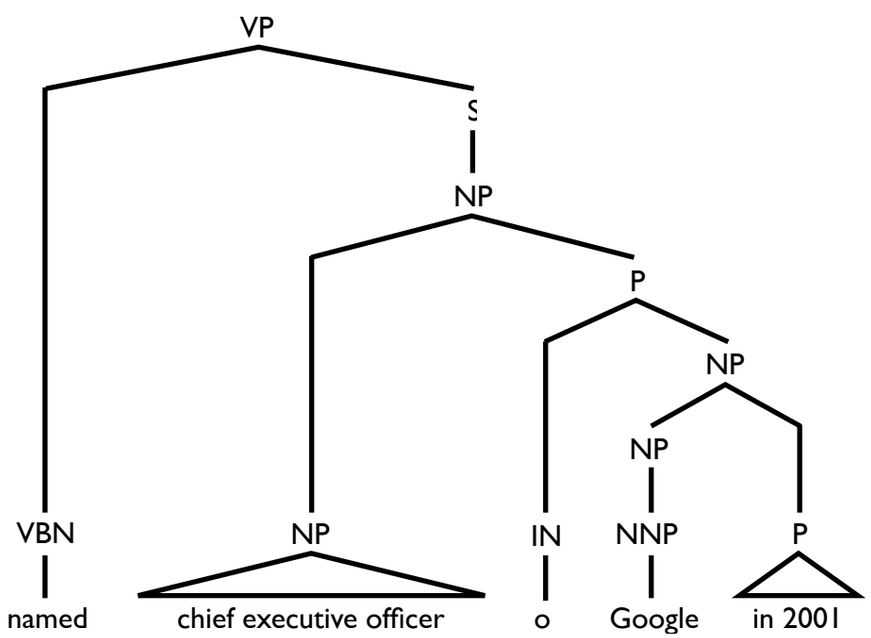
Kummerfeld et al. (2012)

Automatic Error Analysis



Kummerfeld et al. (2012)

Automatic Error Analysis

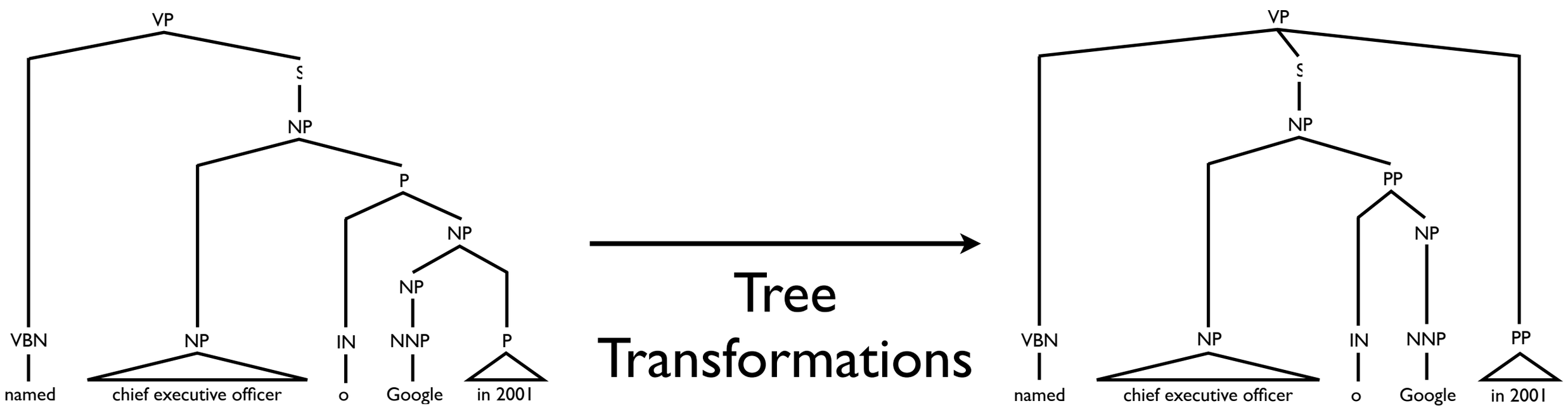


Tree

Transformations

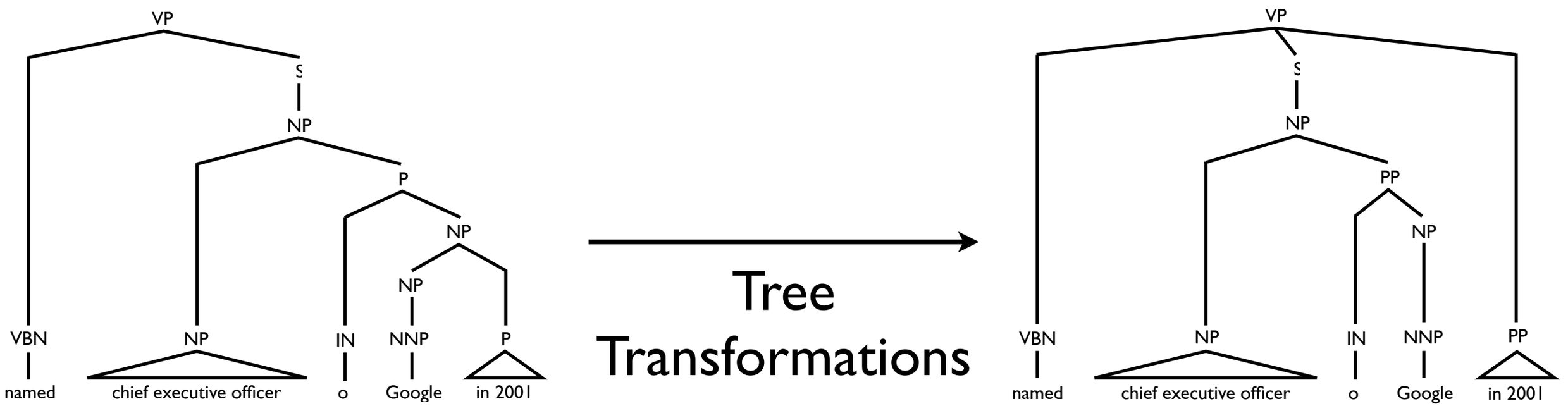
Kummerfeld et al. (2012)

Automatic Error Analysis



Kummerfeld et al. (2012)

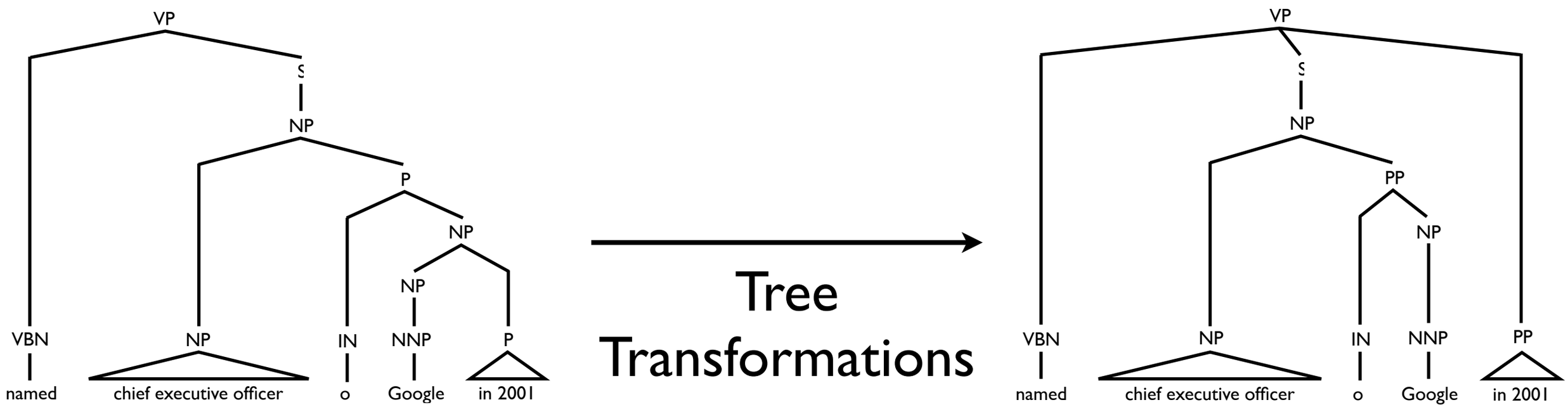
Automatic Error Analysis



List of Transformation Steps

Kummerfeld et al. (2012)

Automatic Error Analysis

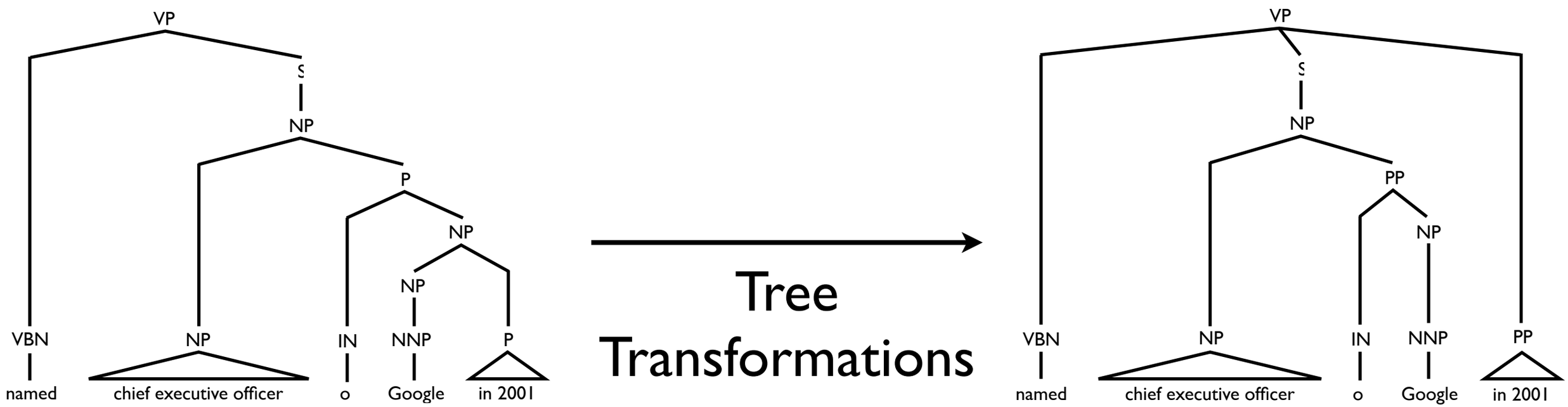


List of Transformation Steps

Categorise

Kummerfeld et al. (2012)

Automatic Error Analysis



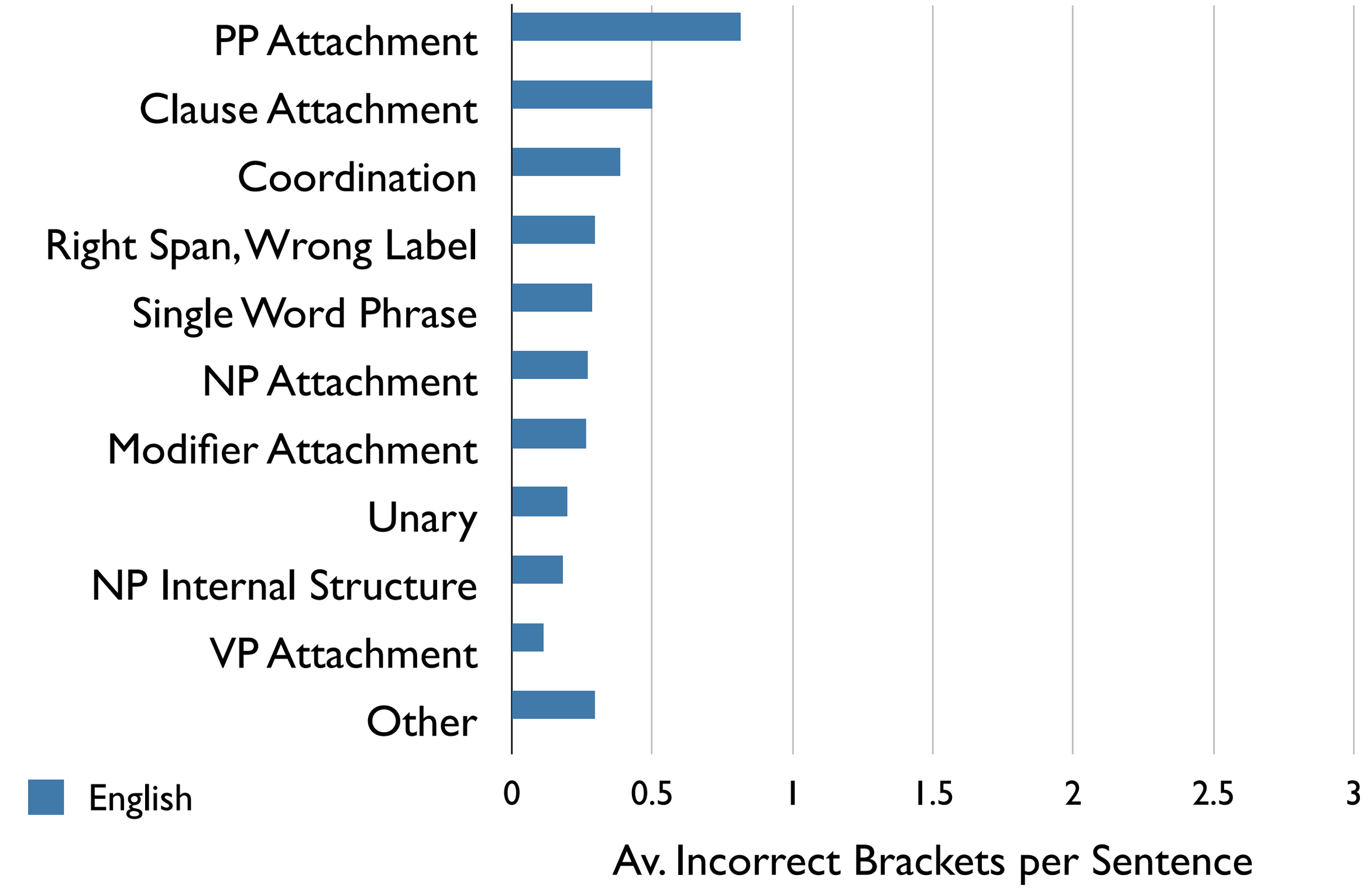
List of Transformation Steps

Categorise

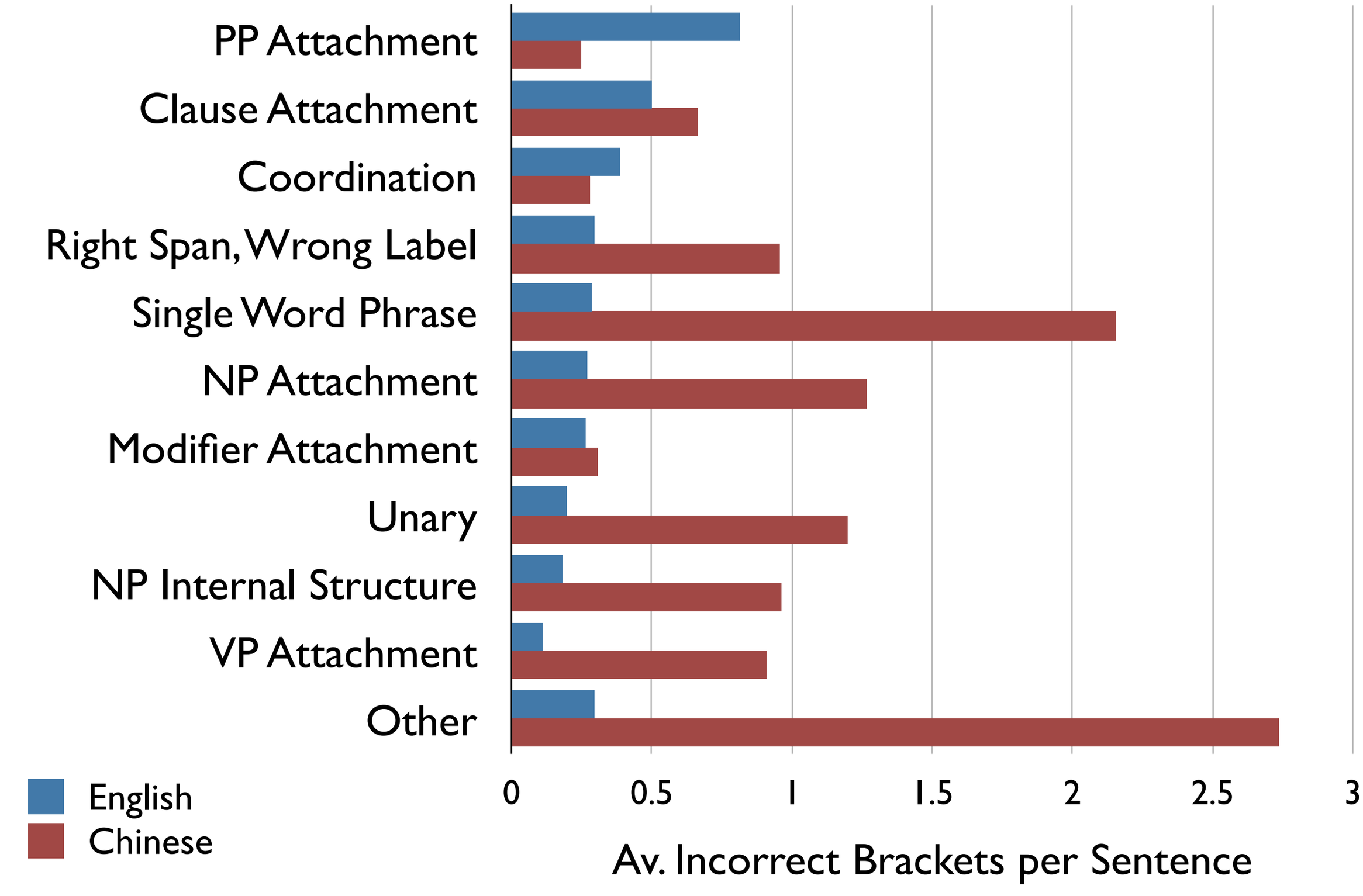
List of Errors

Kummerfeld et al. (2012)

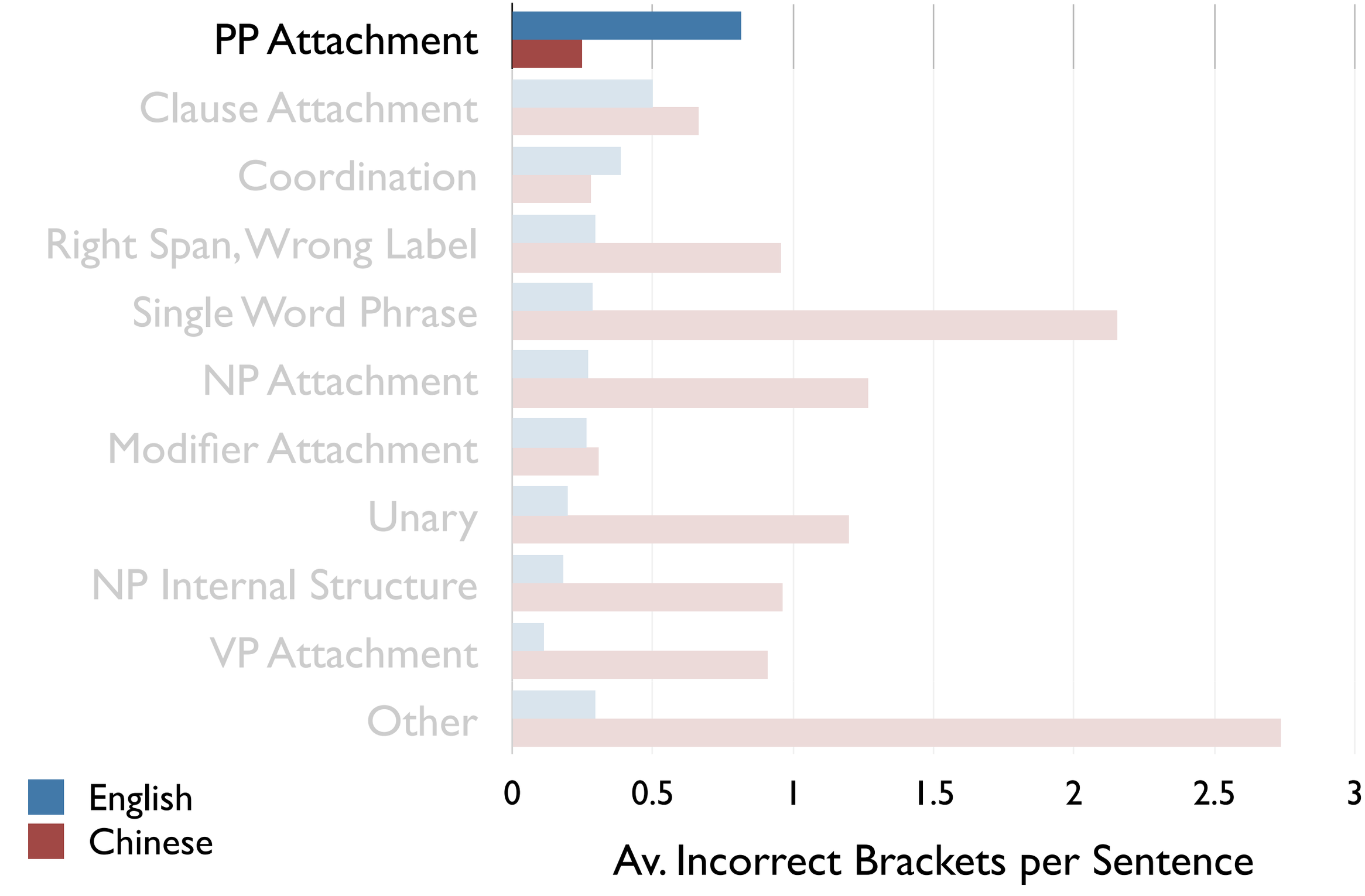
English and Chinese Errors



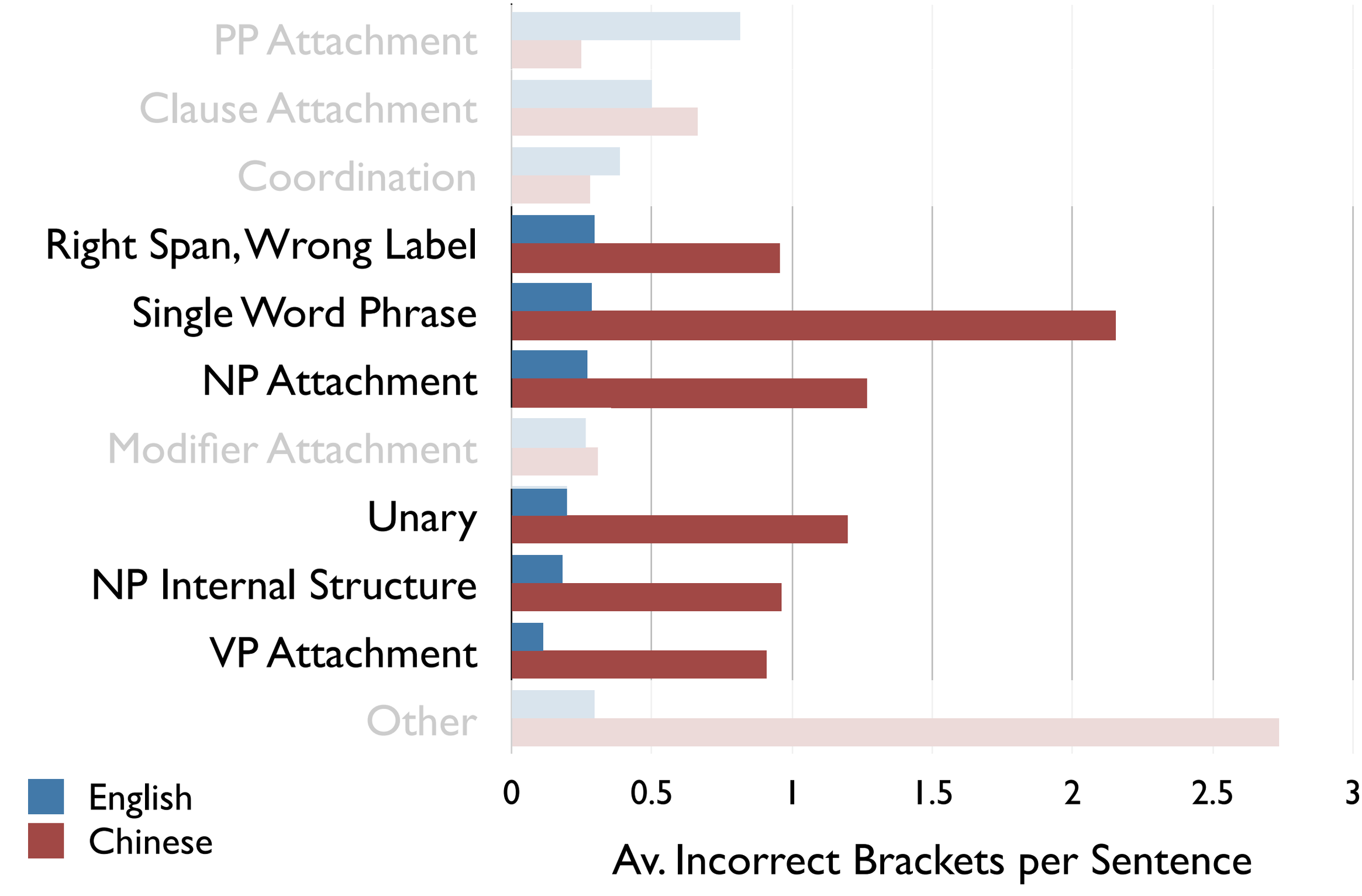
English and Chinese Errors



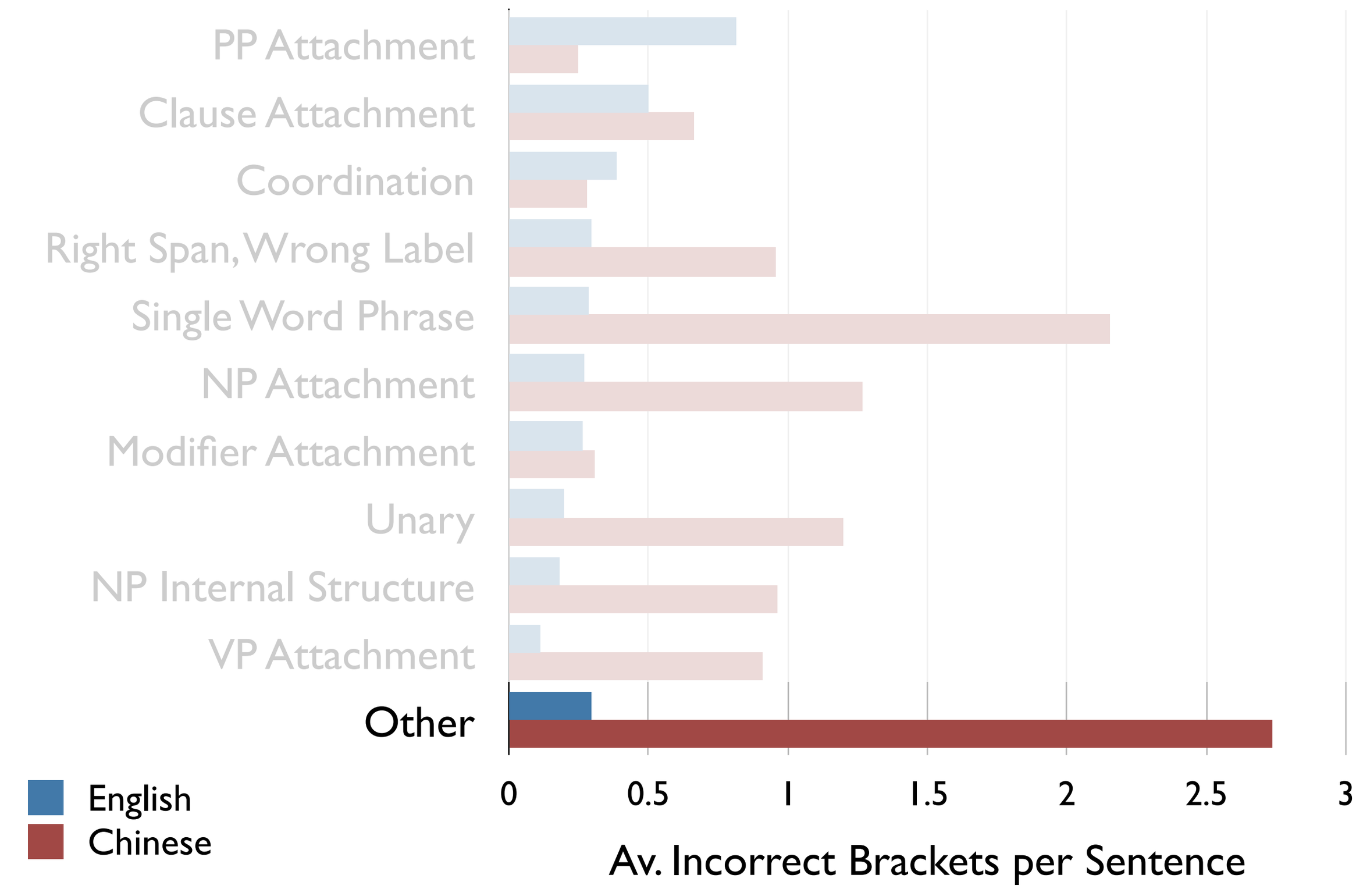
English and Chinese Errors



English and Chinese Errors



English and Chinese Errors

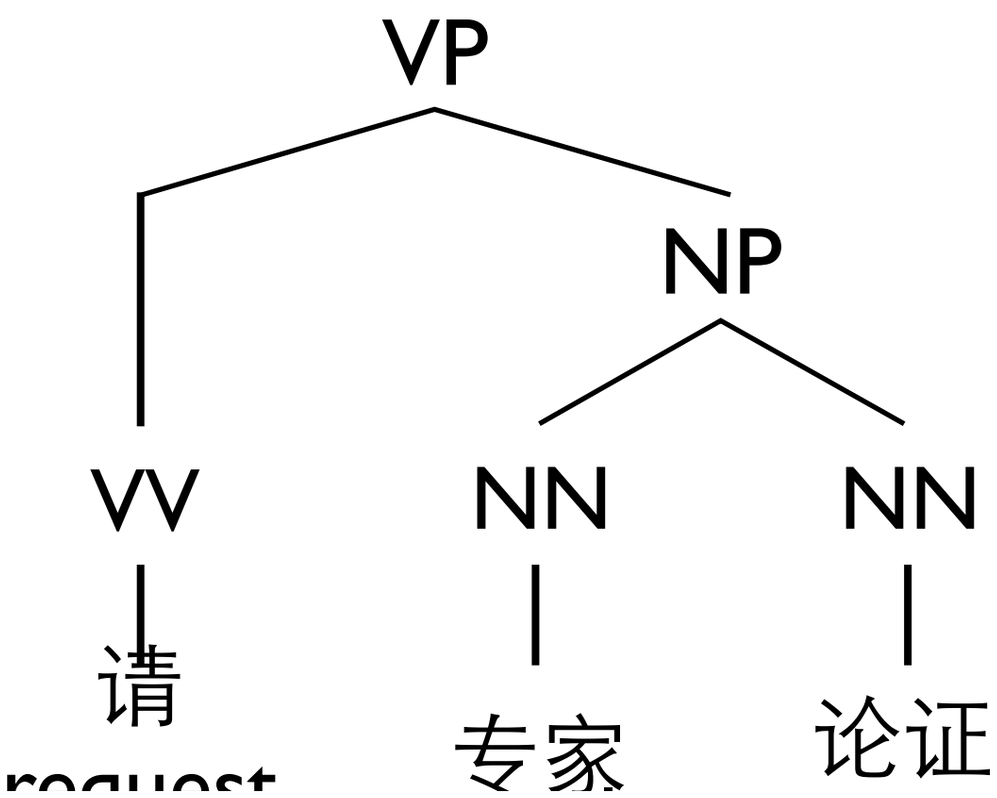


Chinese Parsing Errors

请 专家 论证

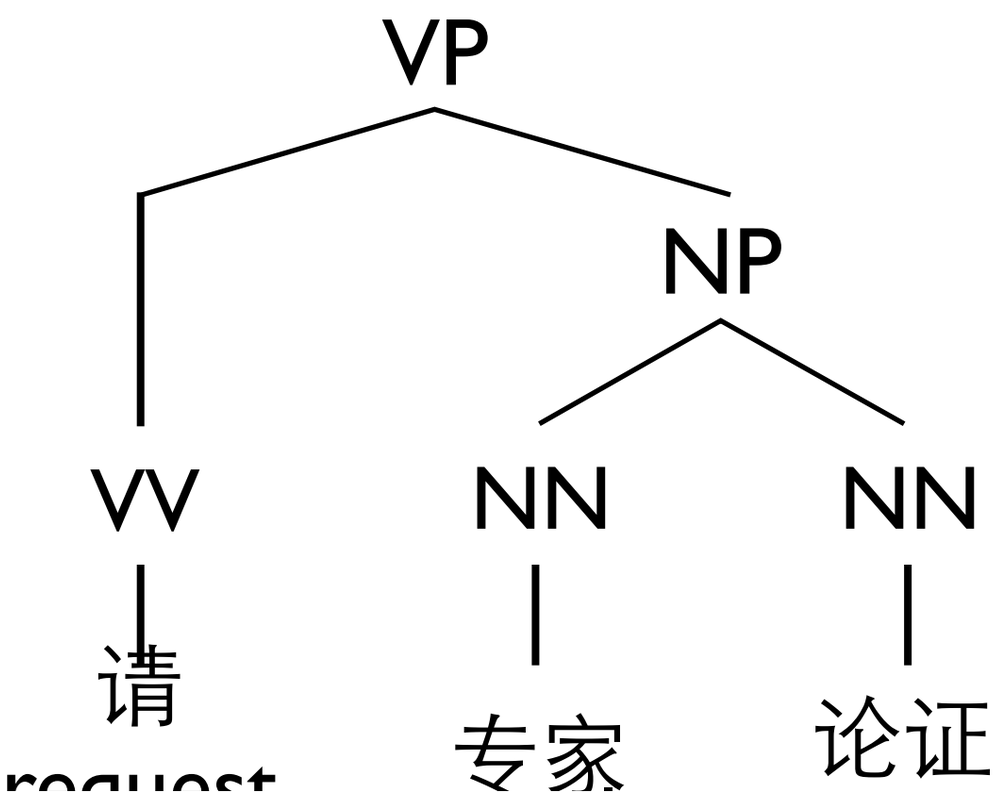
request

Chinese Parsing Errors

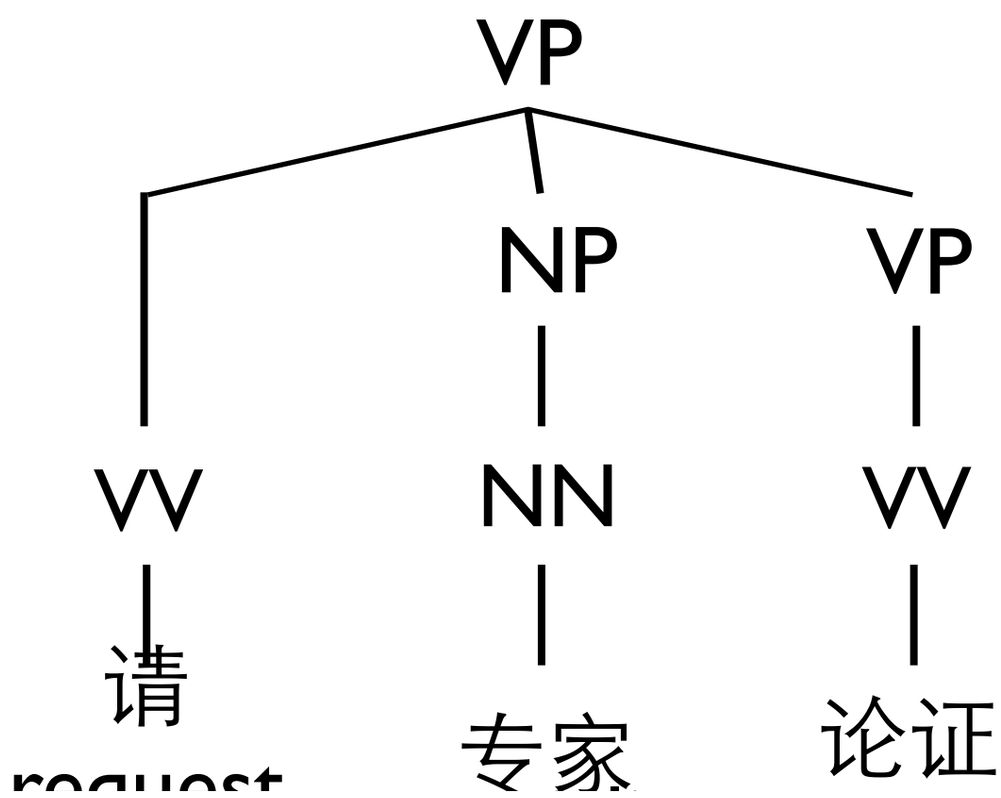


System

Chinese Parsing Errors

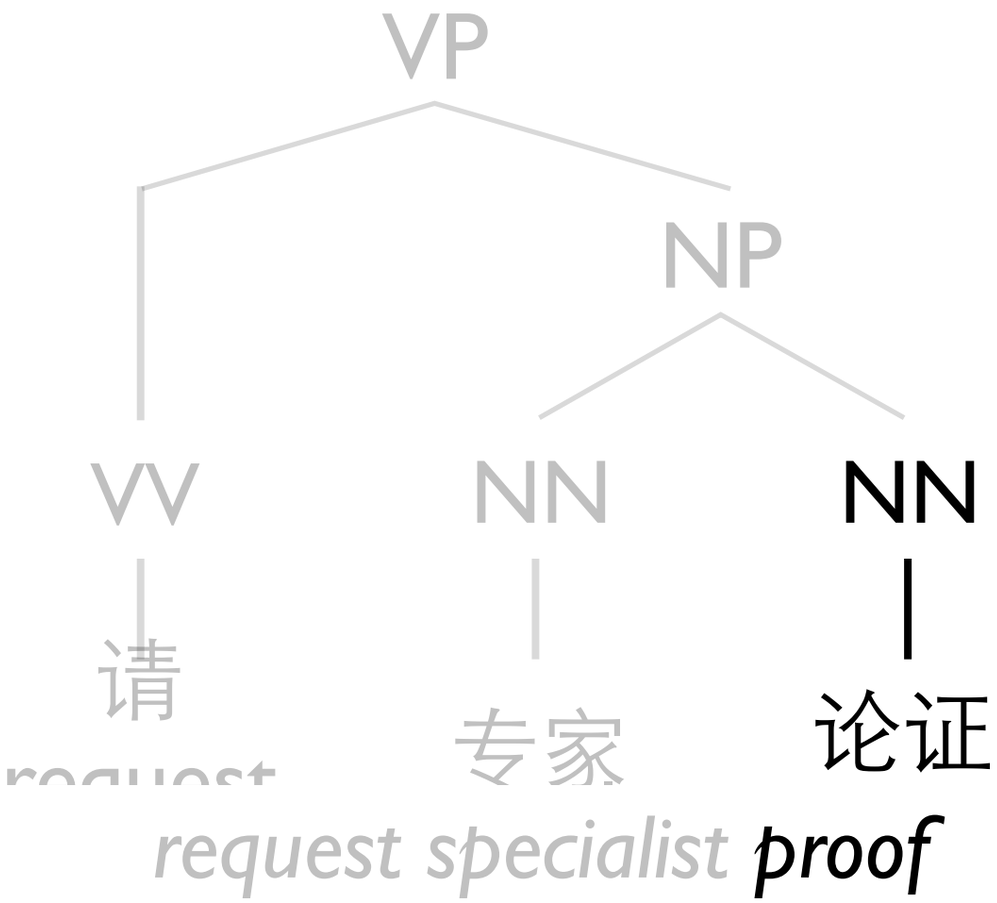


System

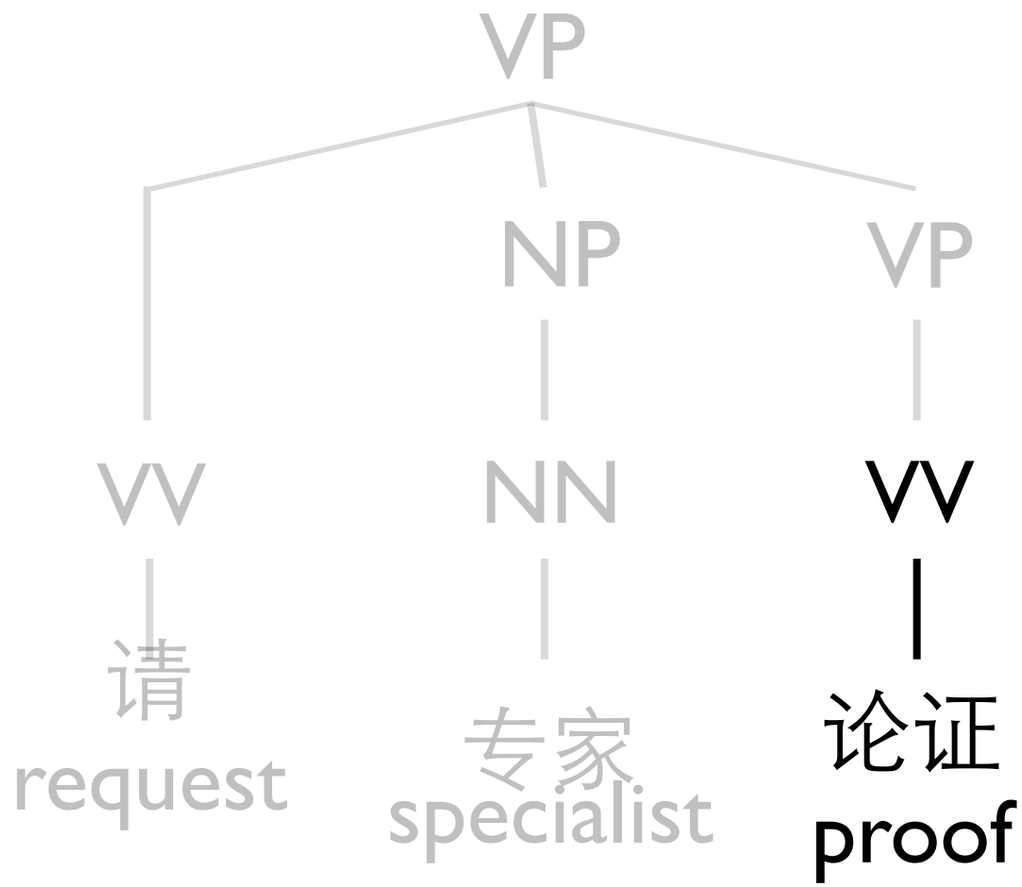


Gold

Chinese Parsing Errors

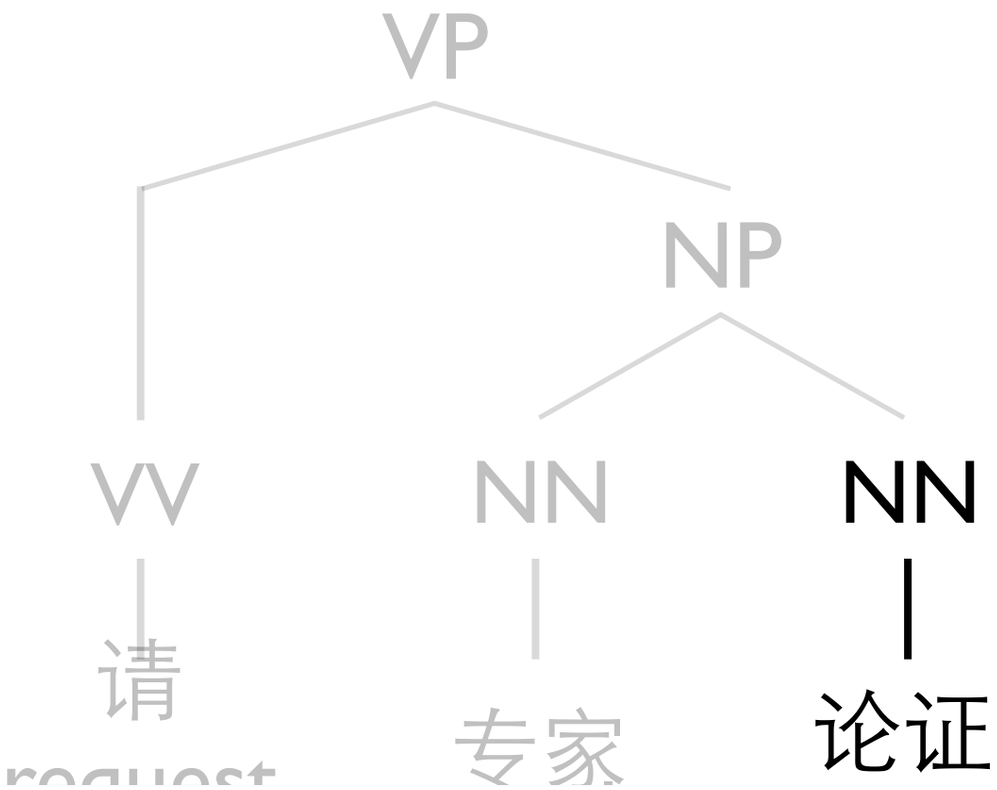


System

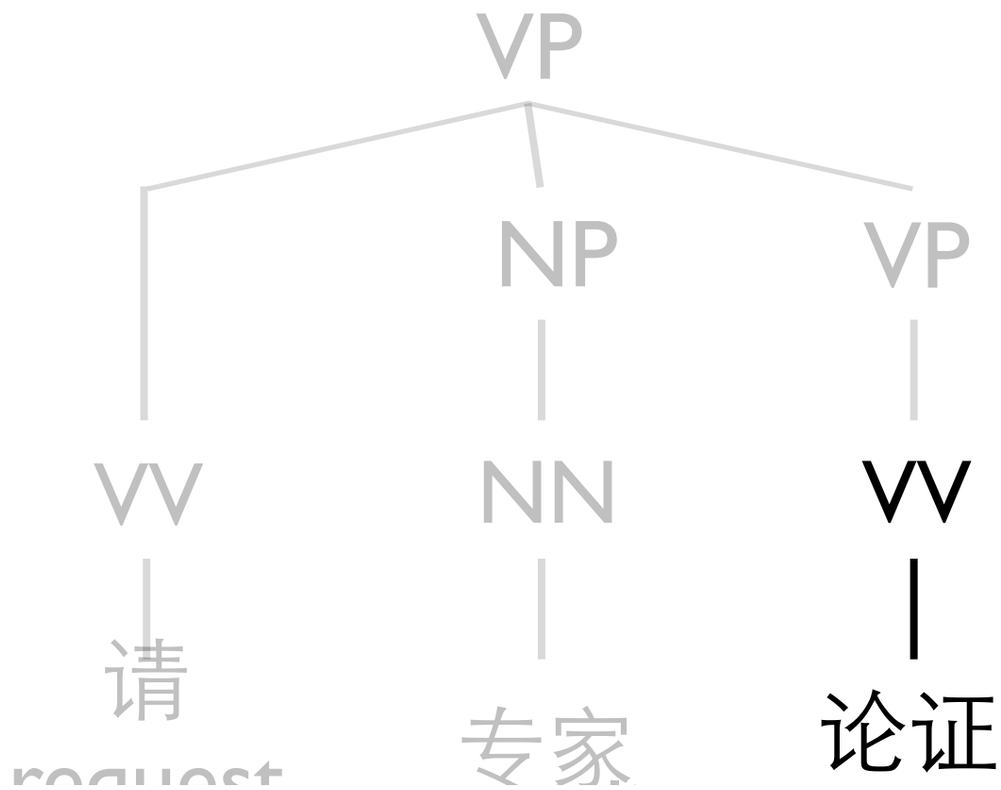


Gold

Chinese Parsing Errors

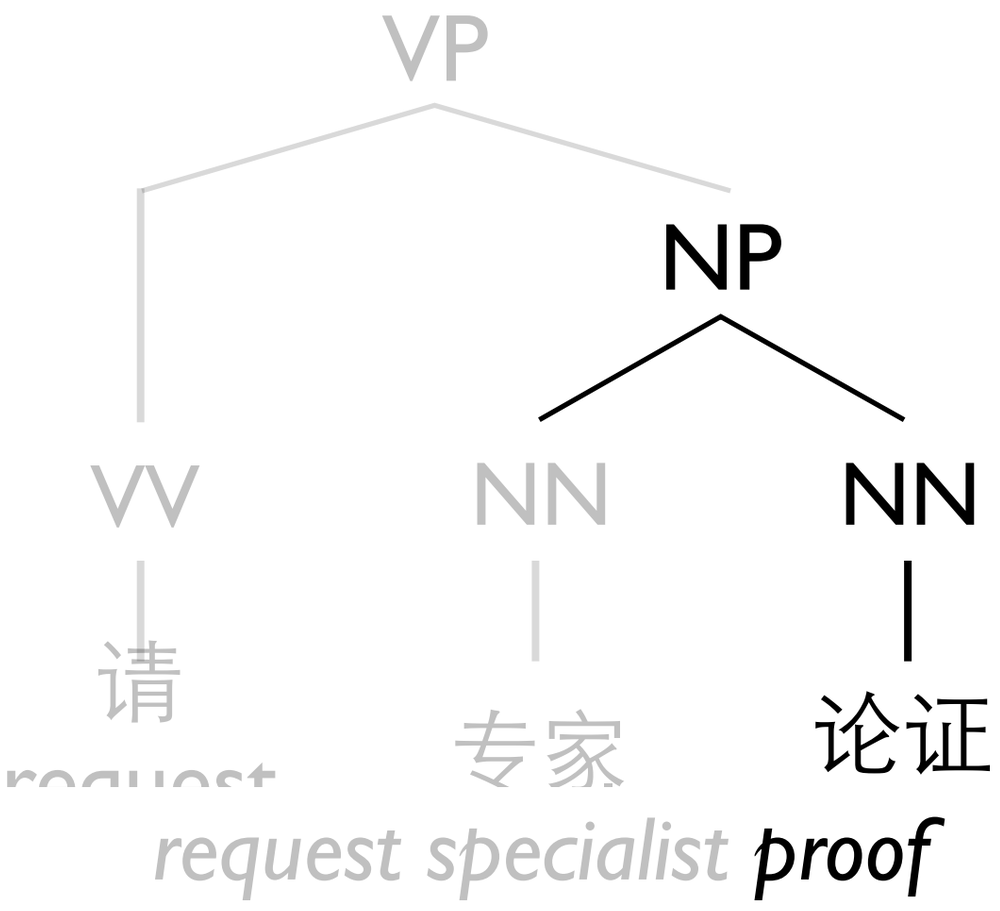


System

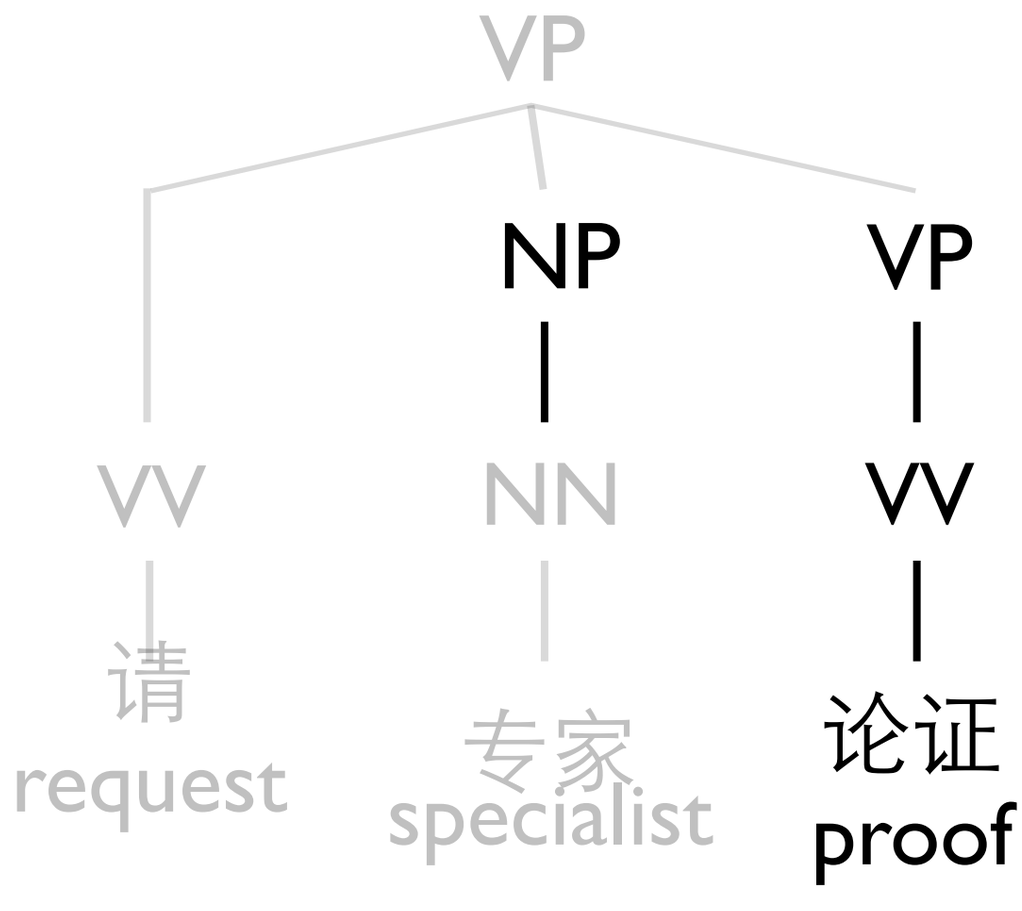


Gold

Chinese Parsing Errors

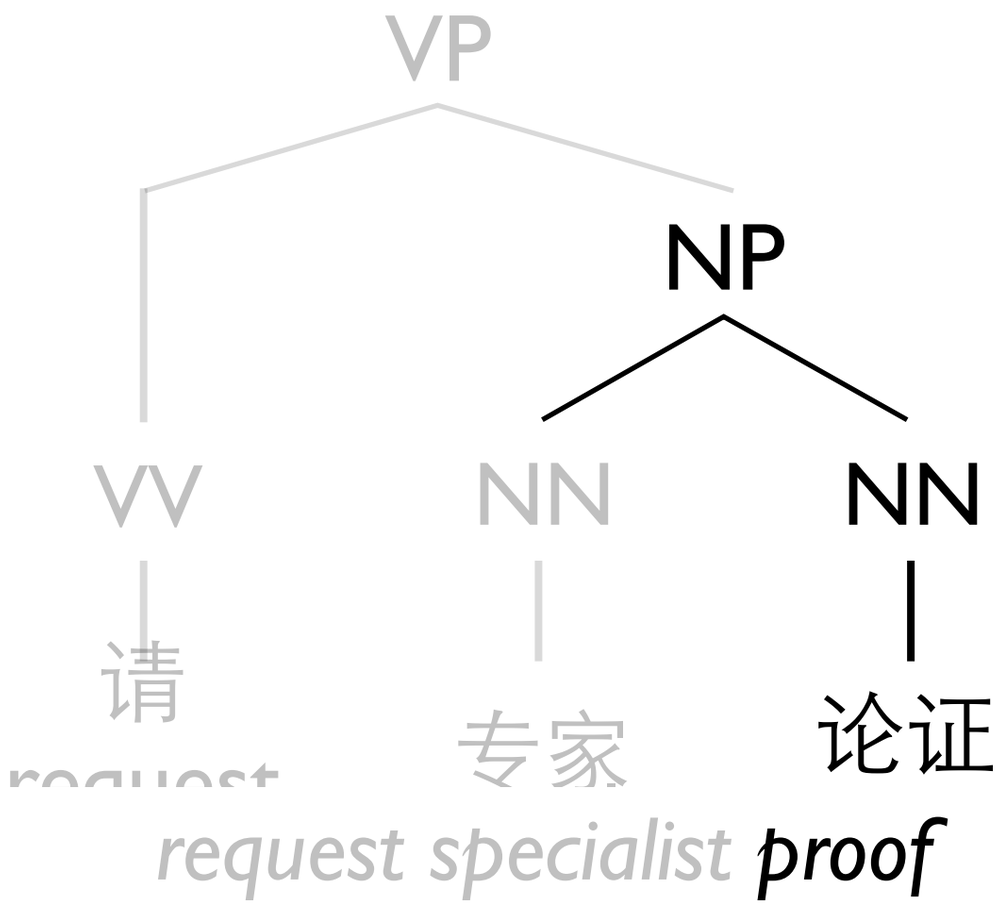


System

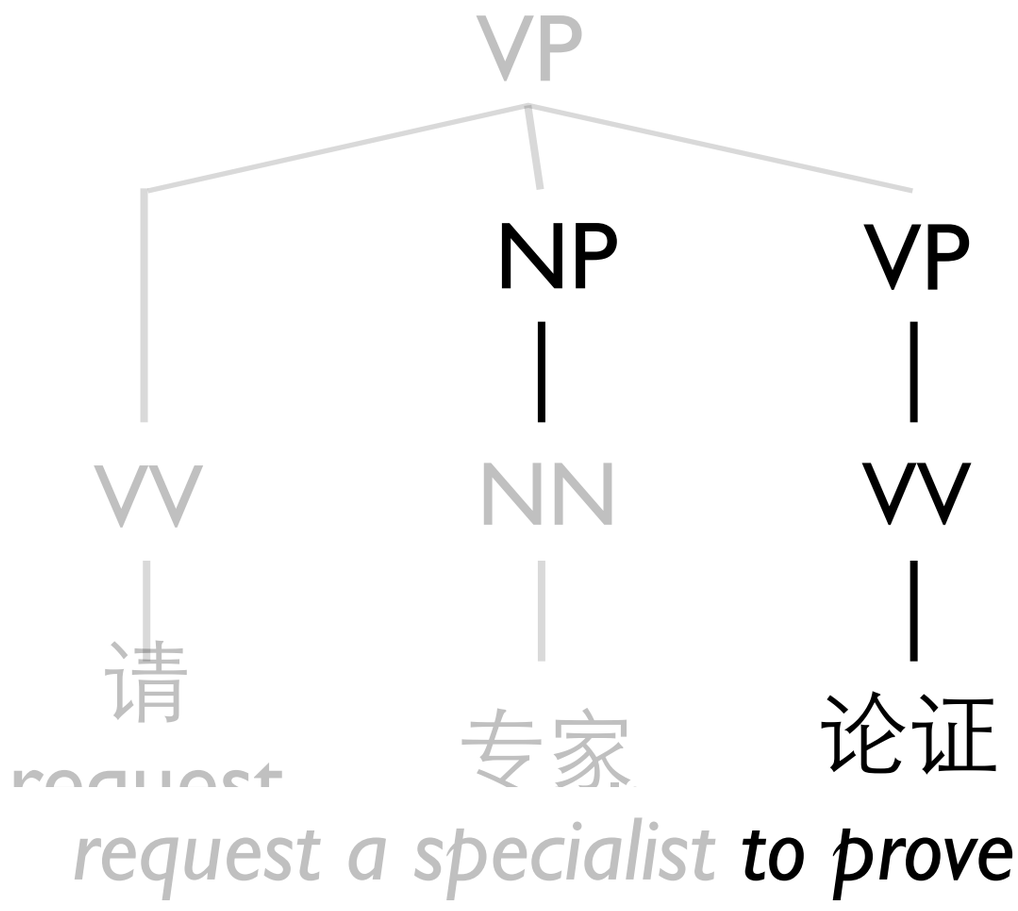


Gold

Chinese Parsing Errors



System

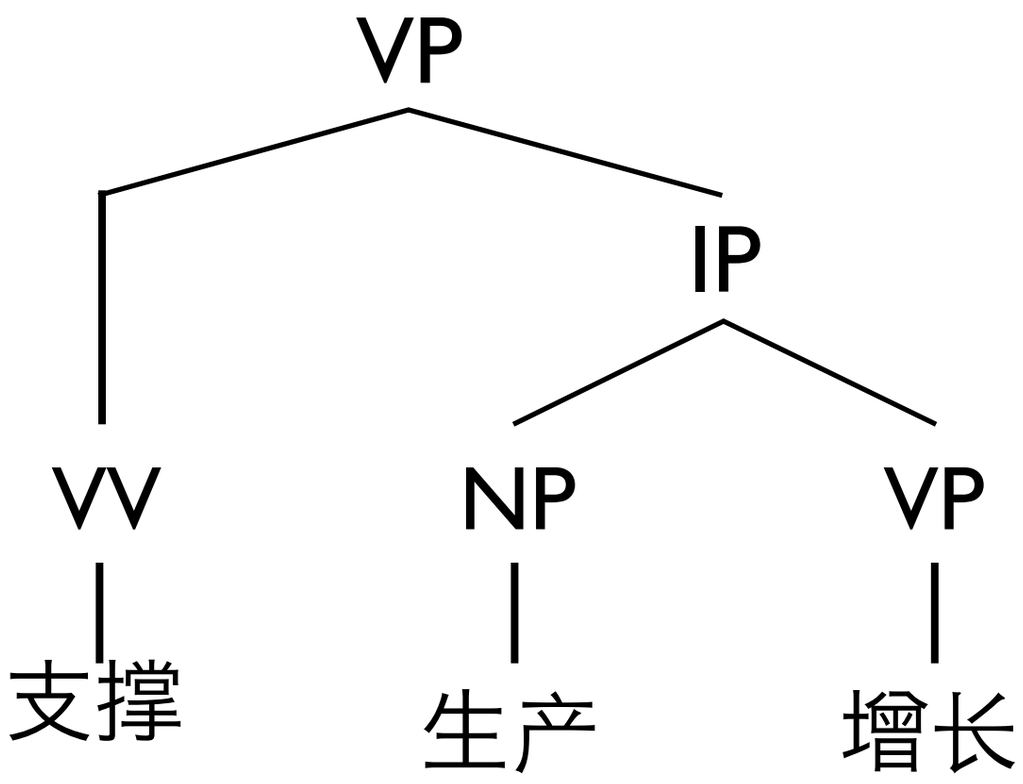


Gold

Chinese Parsing Errors

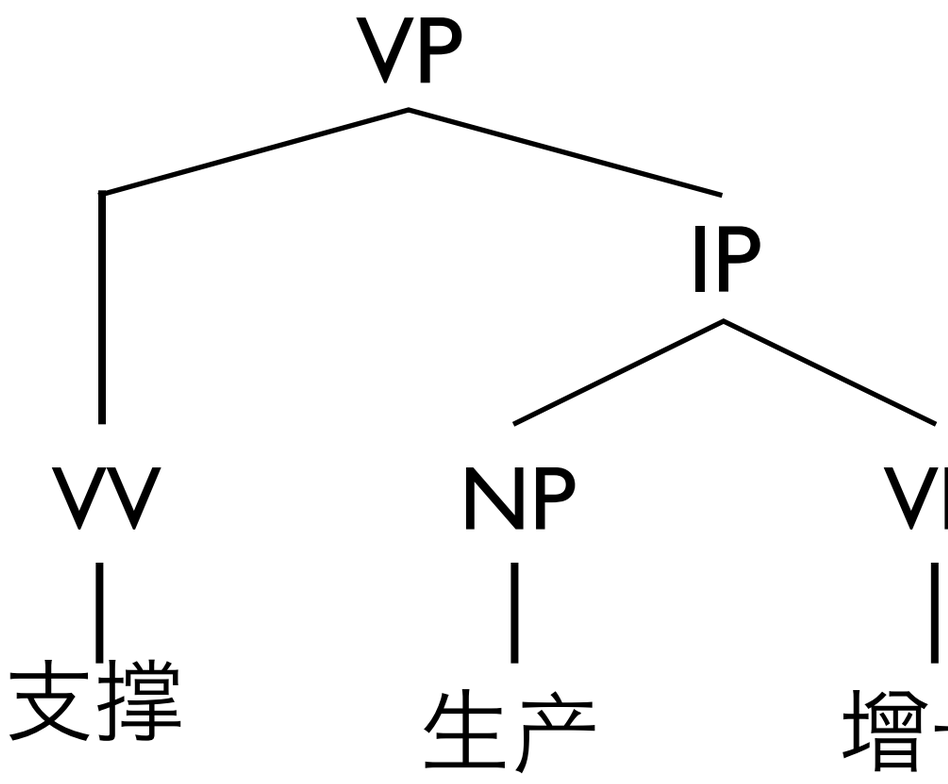
支撑　　生产　　增长

Chinese Parsing Errors



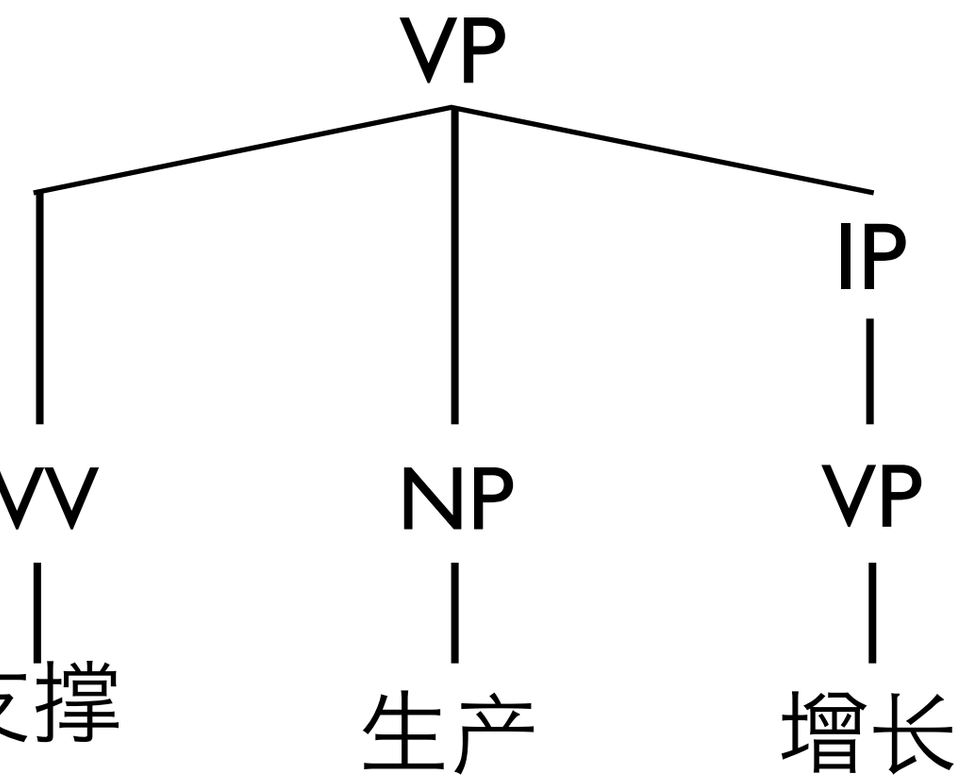
System

Chinese Parsing Errors



support increasing production

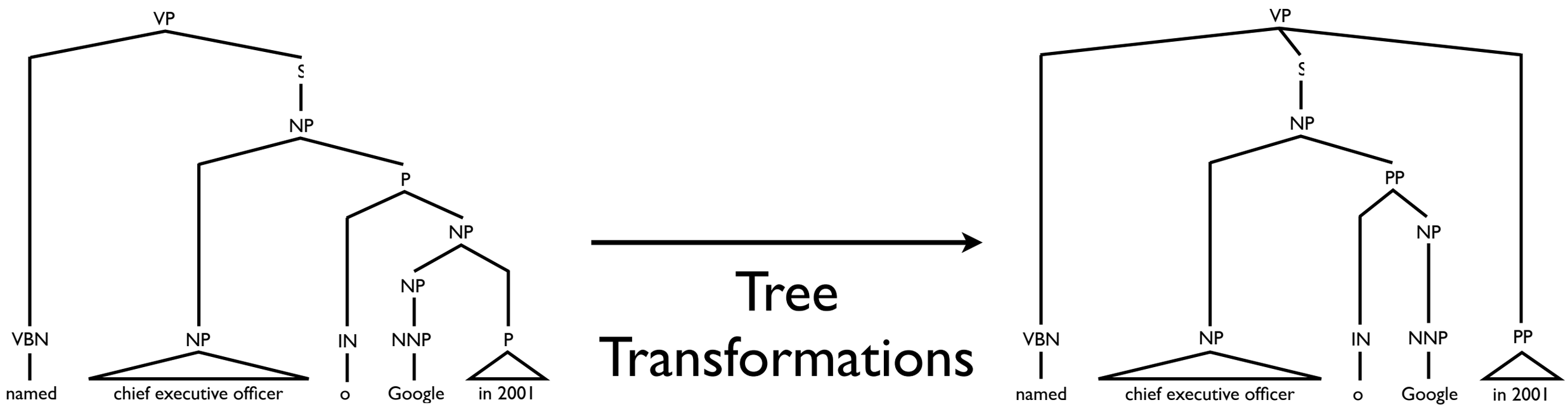
System



support production to increase

Gold

Automatic Error Analysis

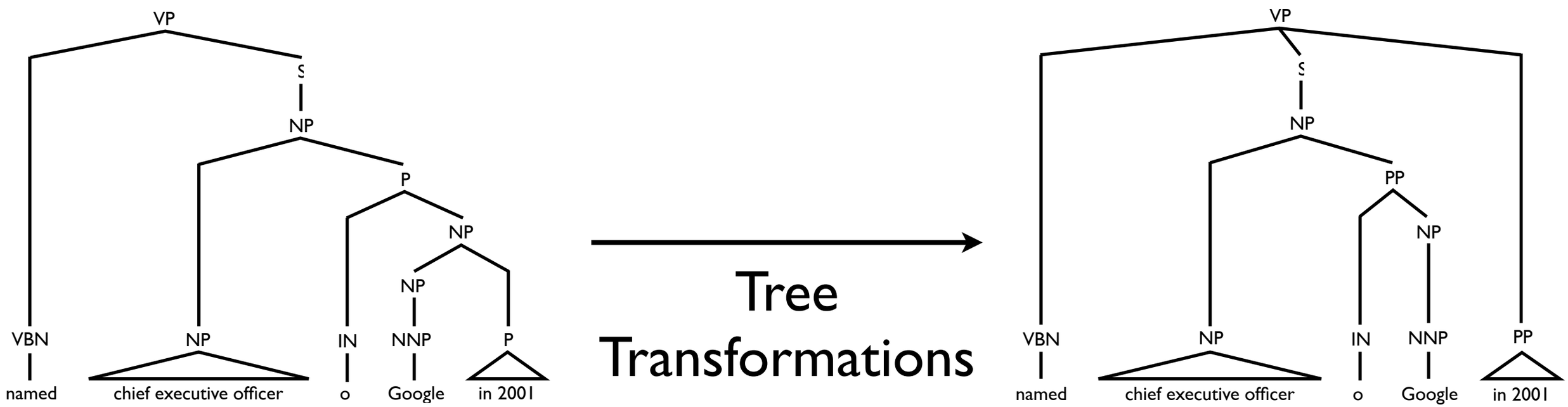


List of Transformation Steps

Categorise

List of Errors

Automatic Error Analysis

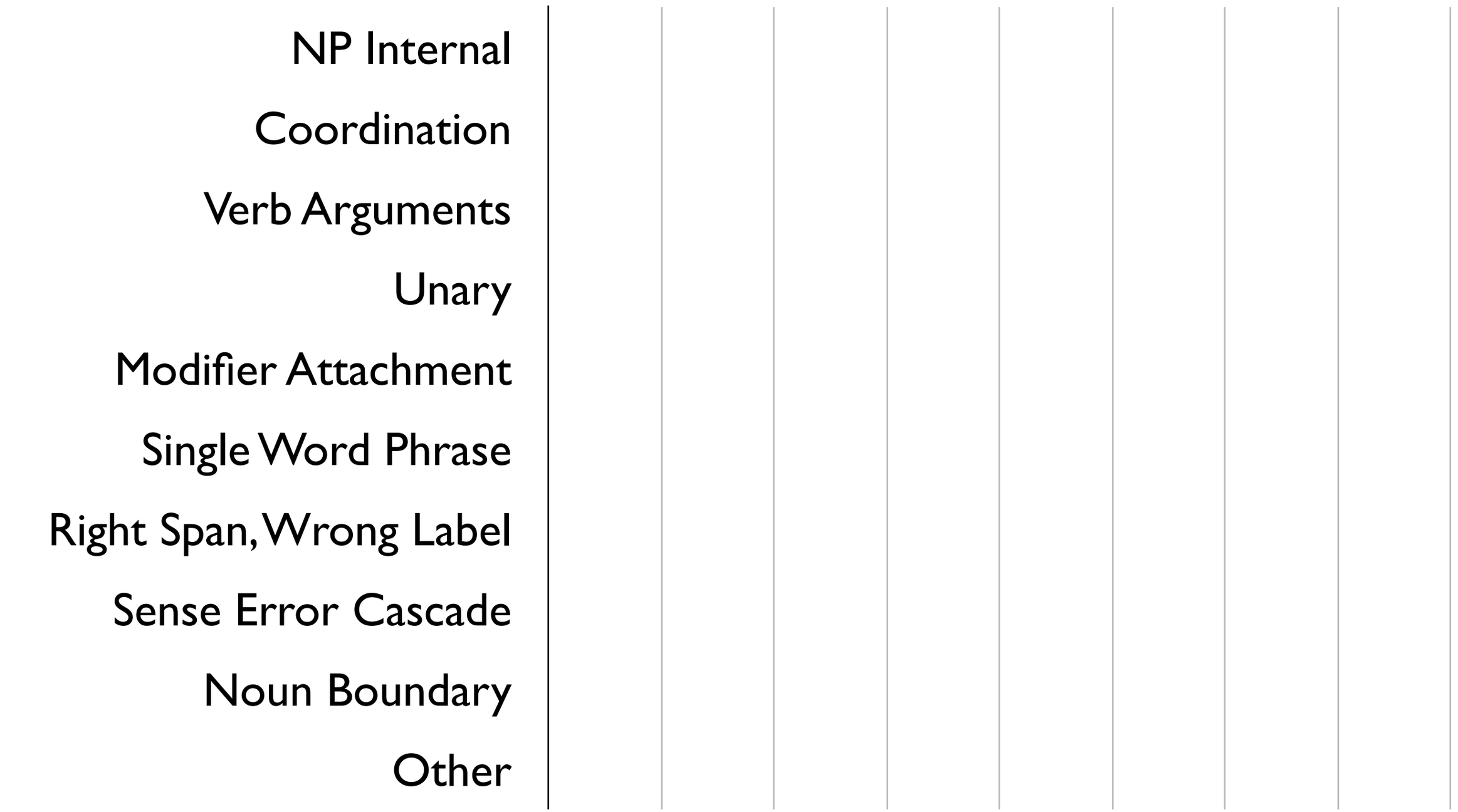


List of Transformation Steps

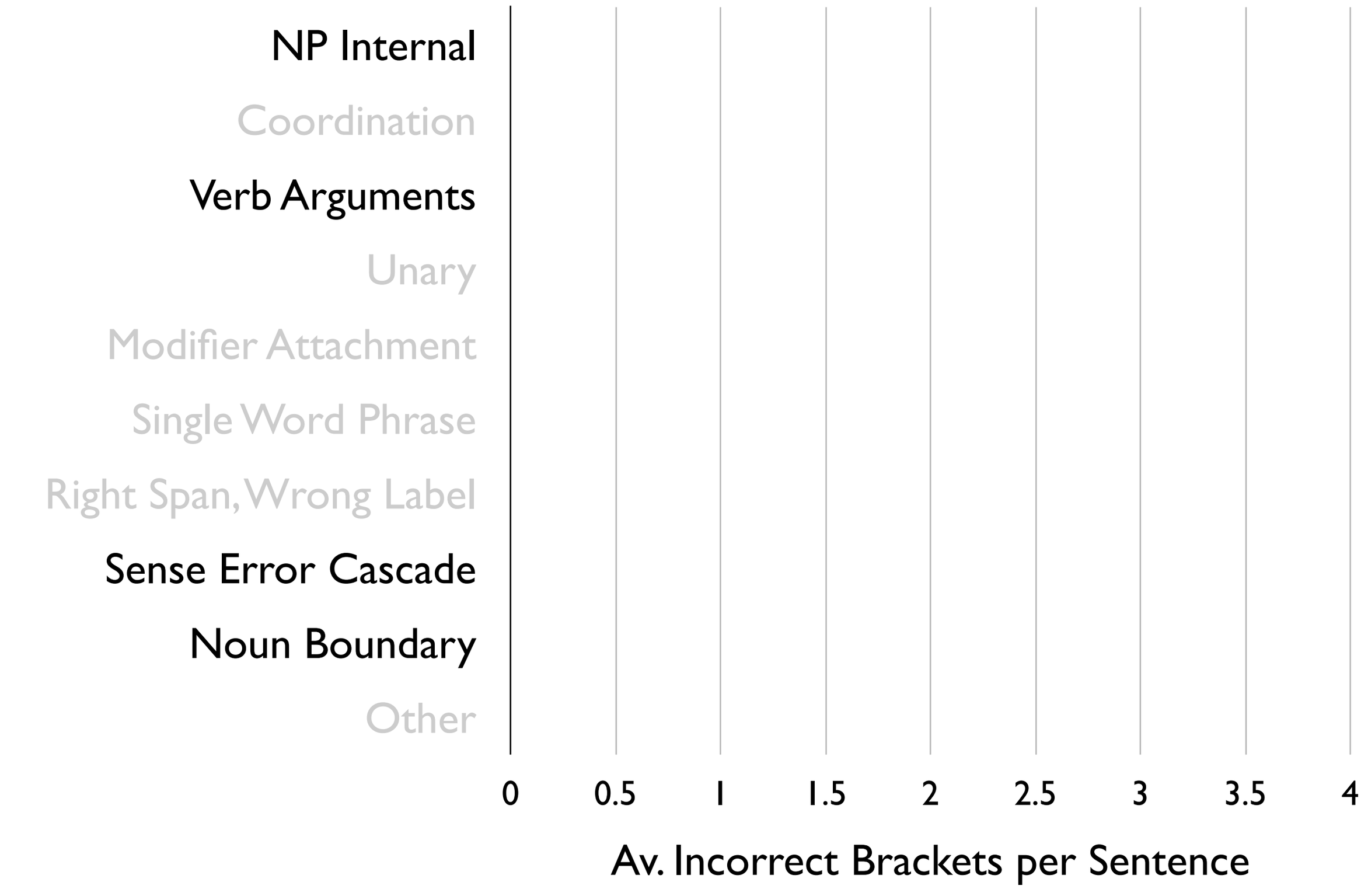
Categorise

List of Errors

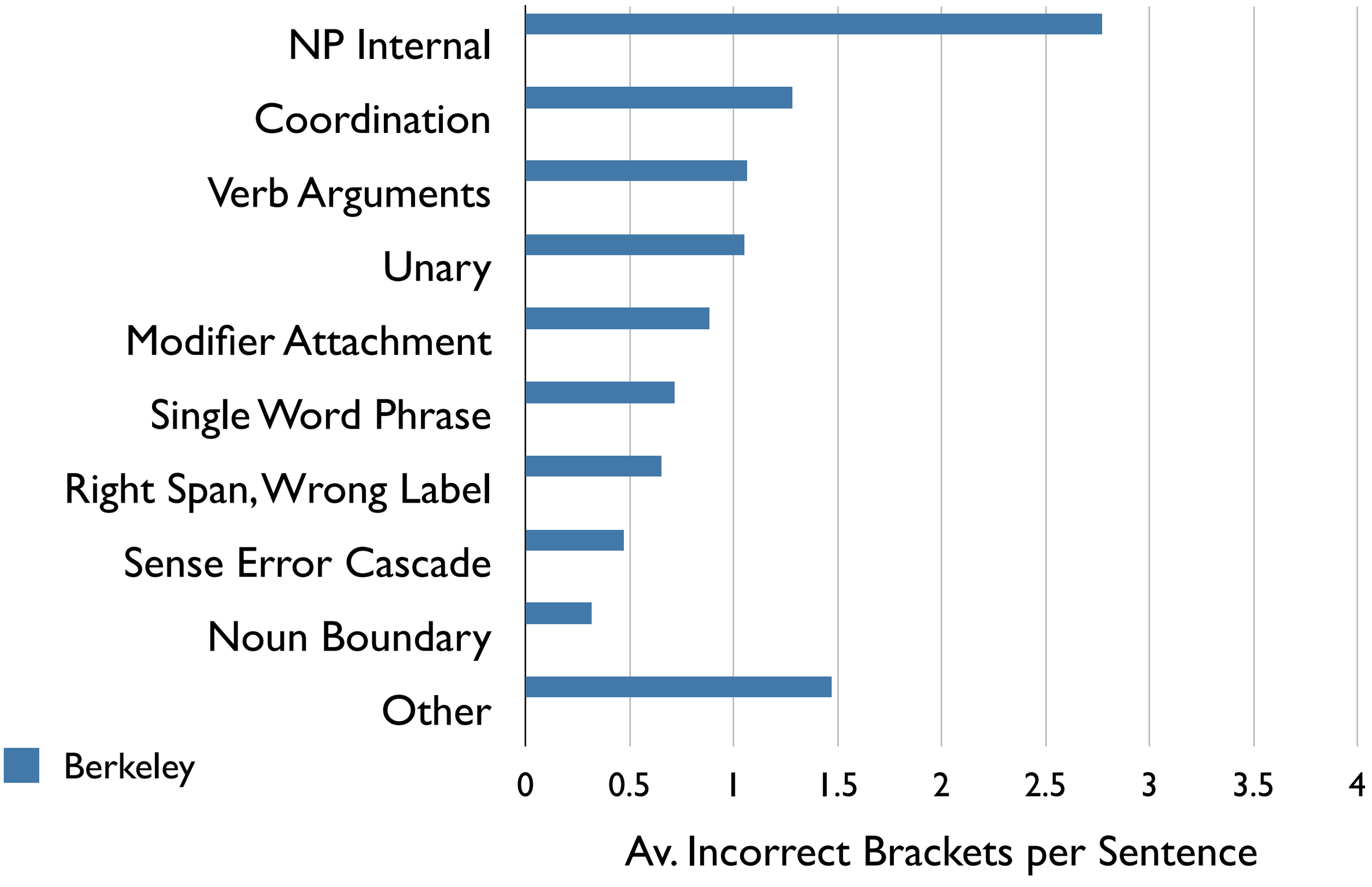
Chinese Parse Errors



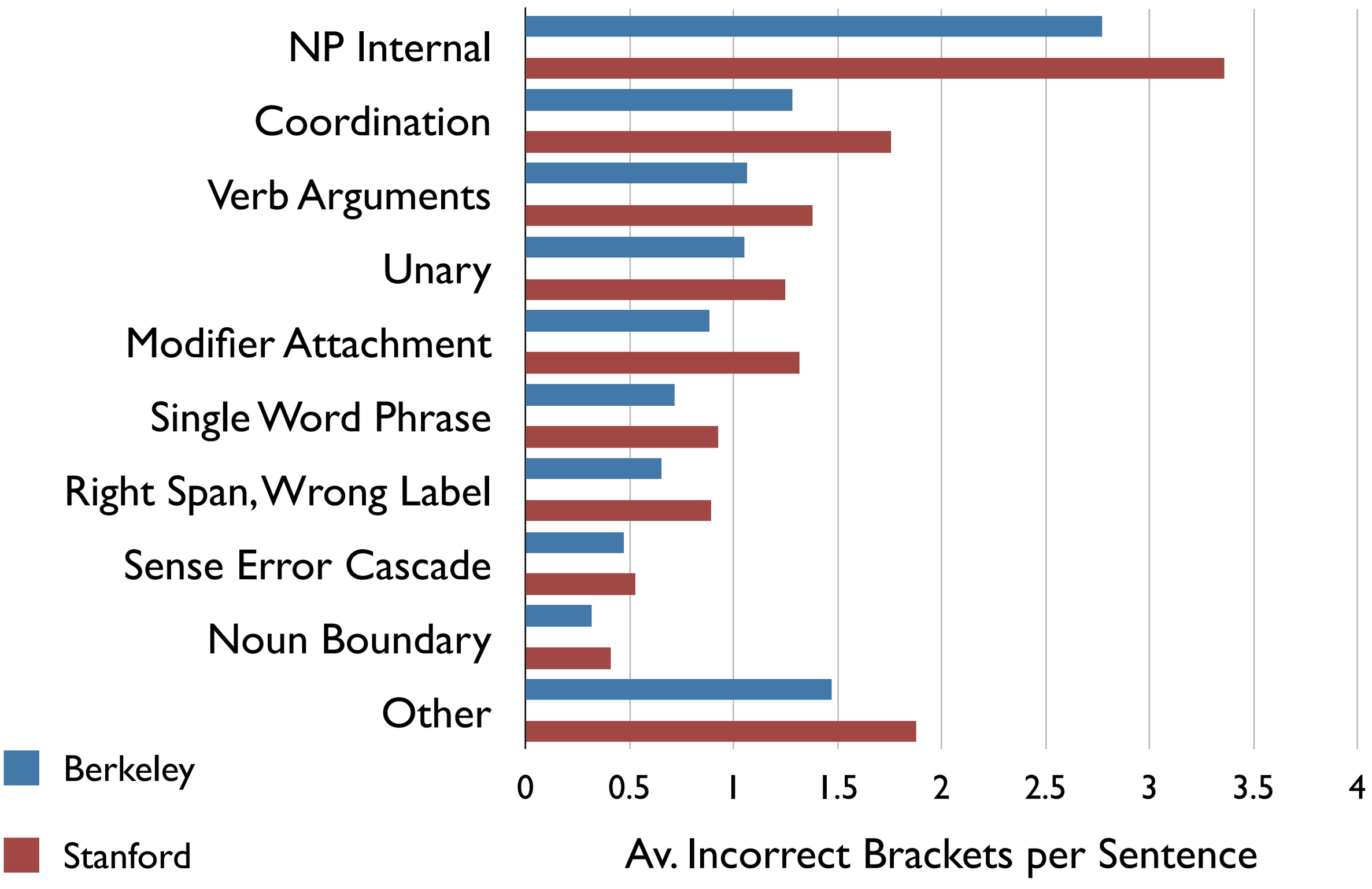
Chinese Parse Errors



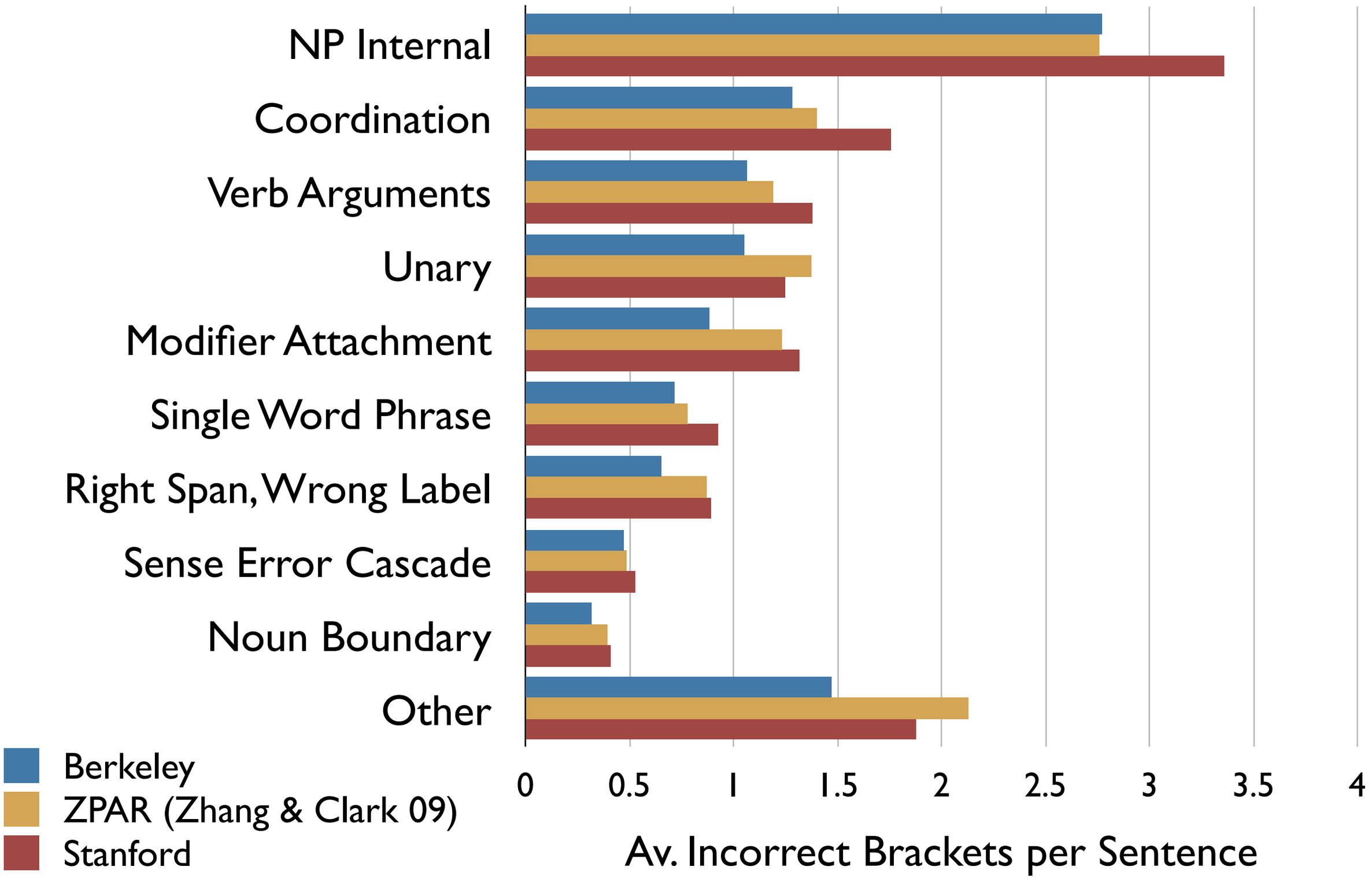
Chinese Parse Errors



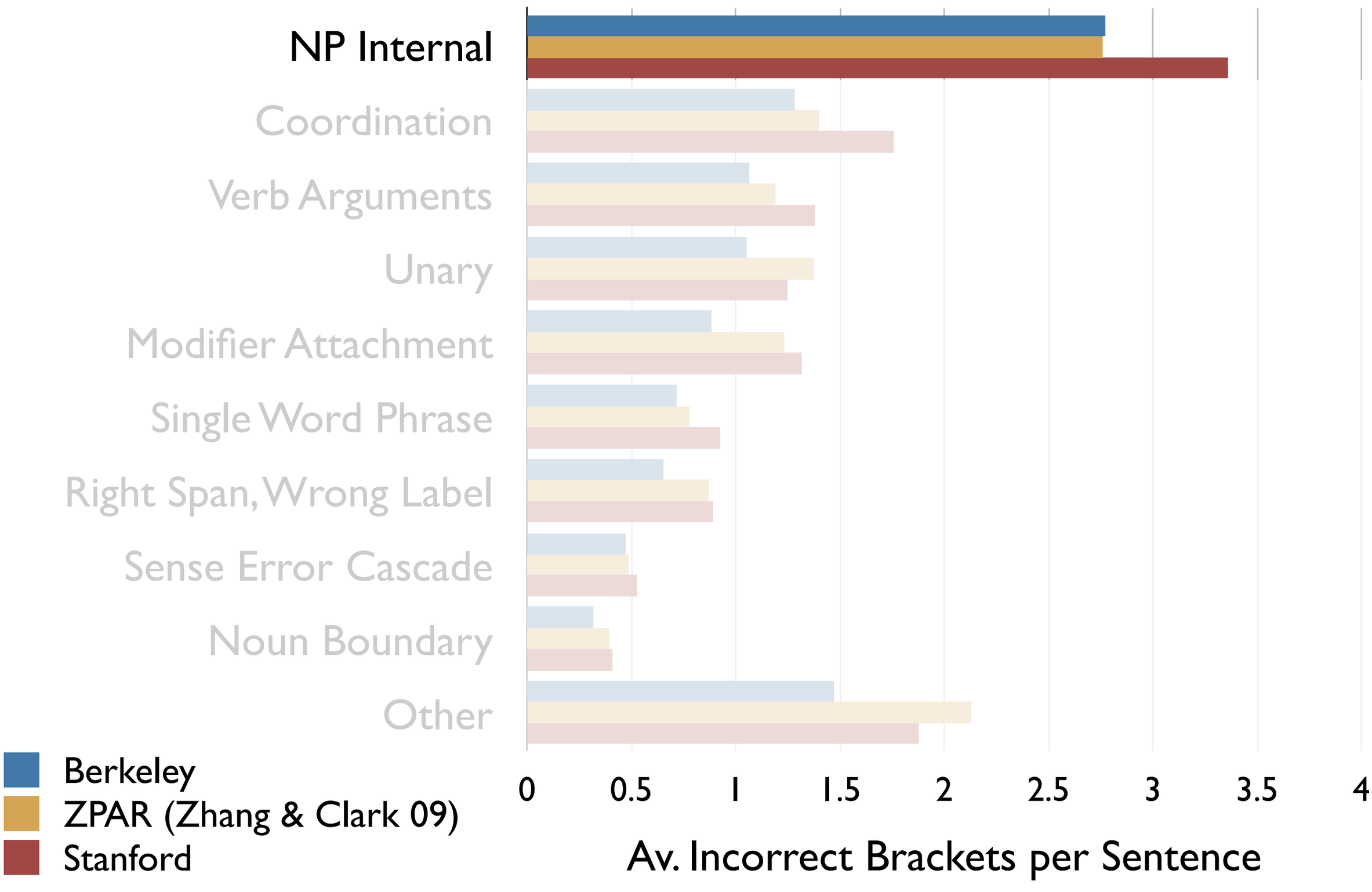
Chinese Parse Errors



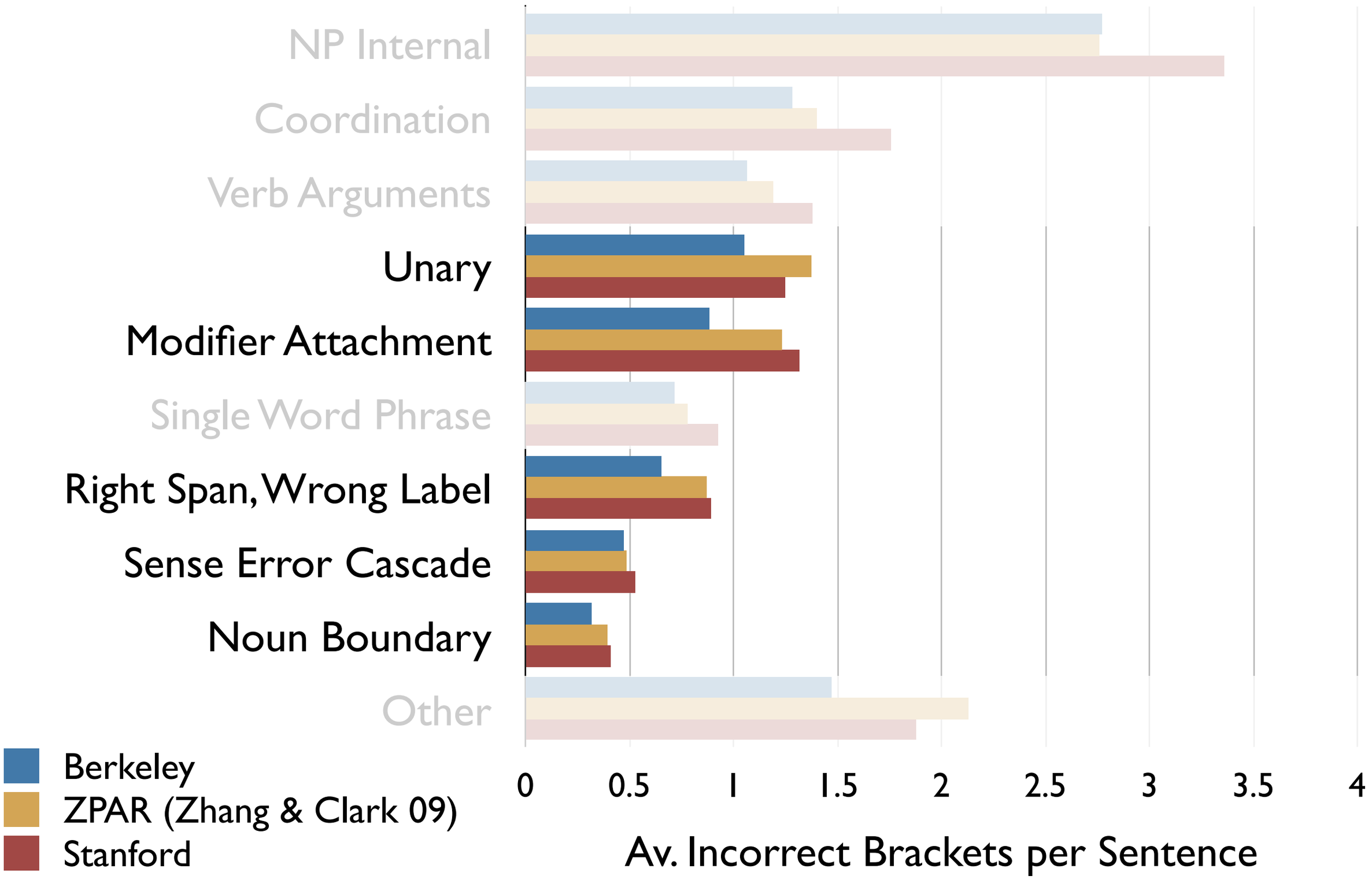
Chinese Parse Errors



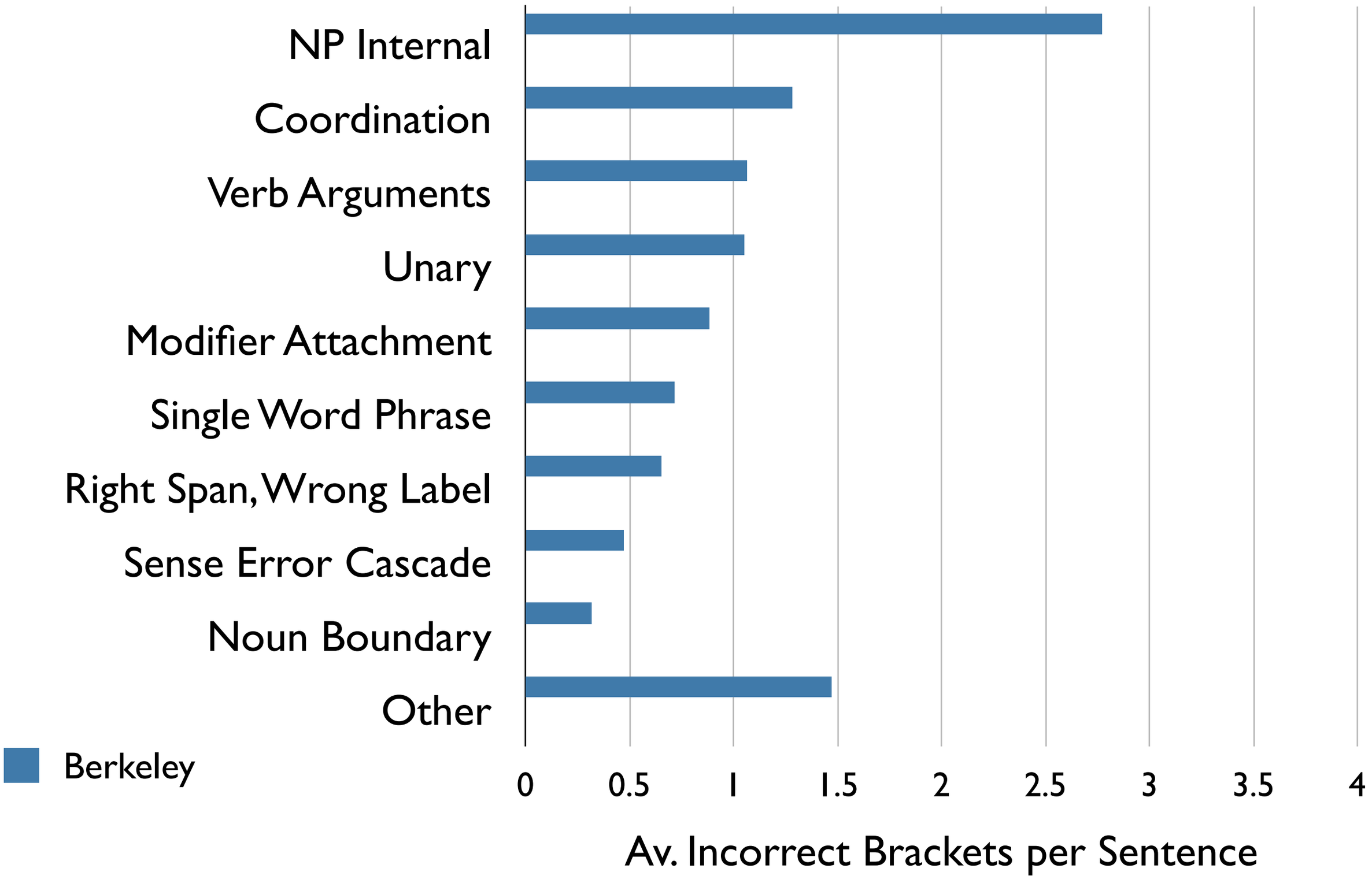
Chinese Parse Errors



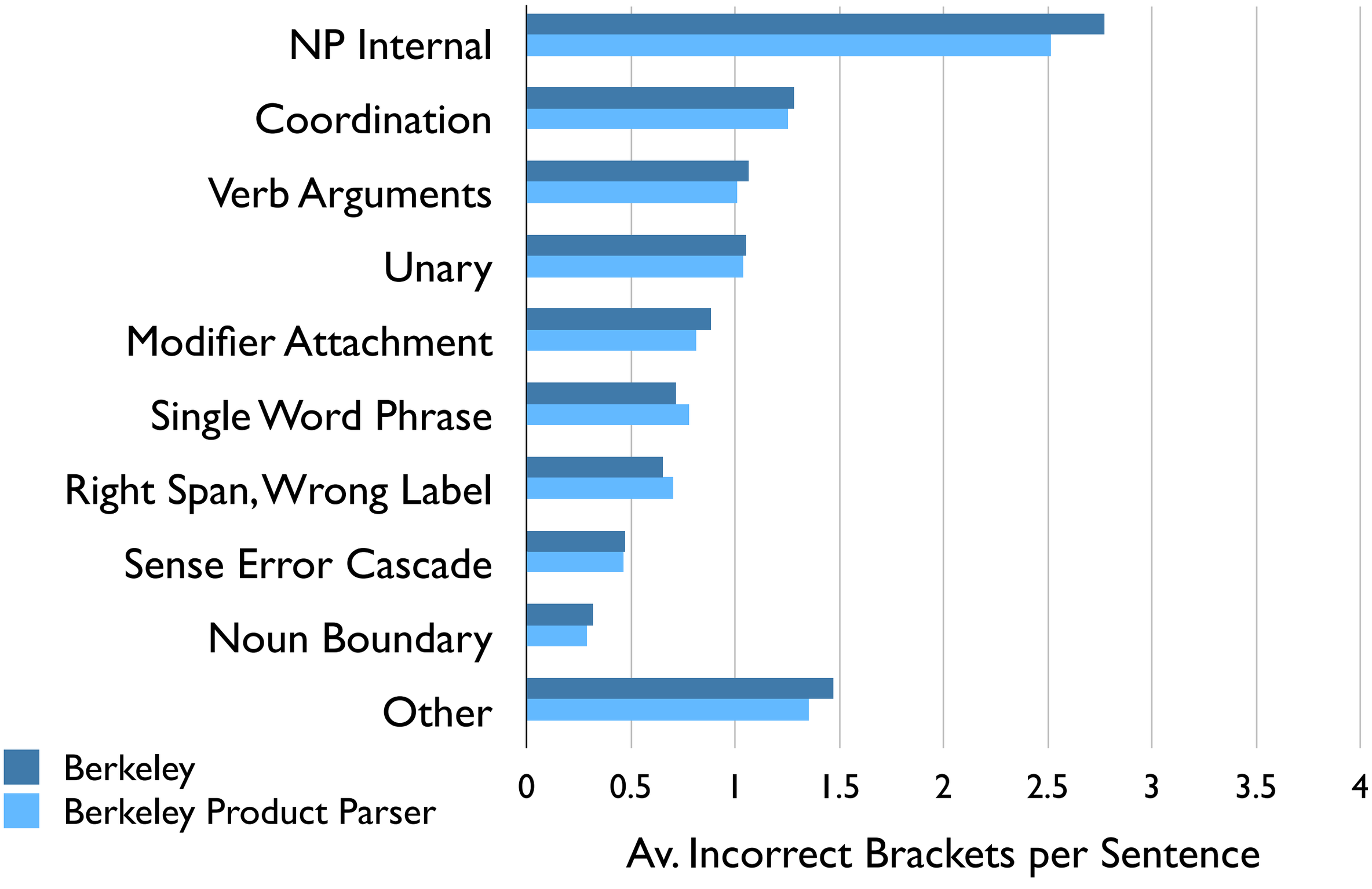
Chinese Parse Errors



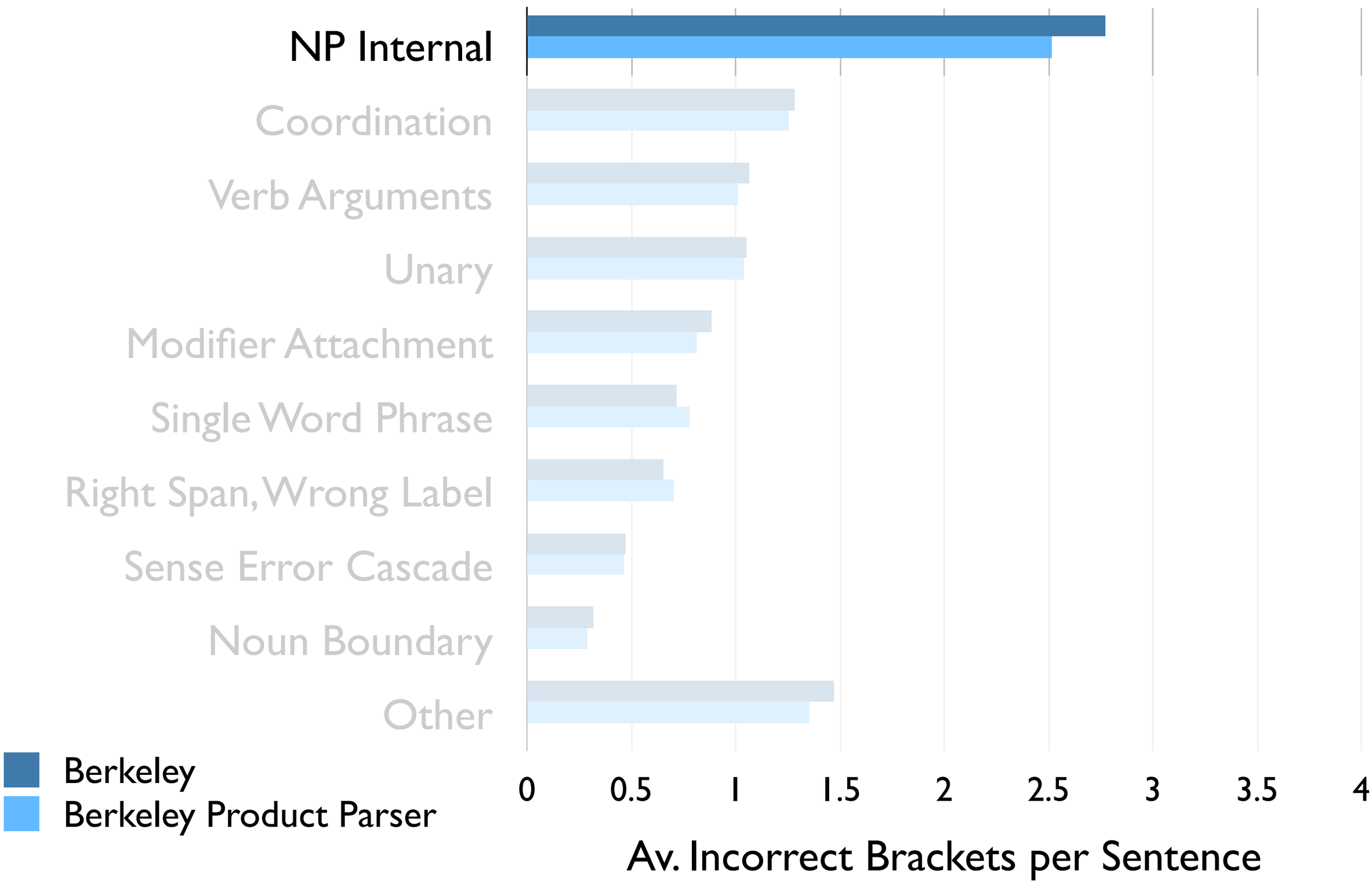
Chinese Parse Errors



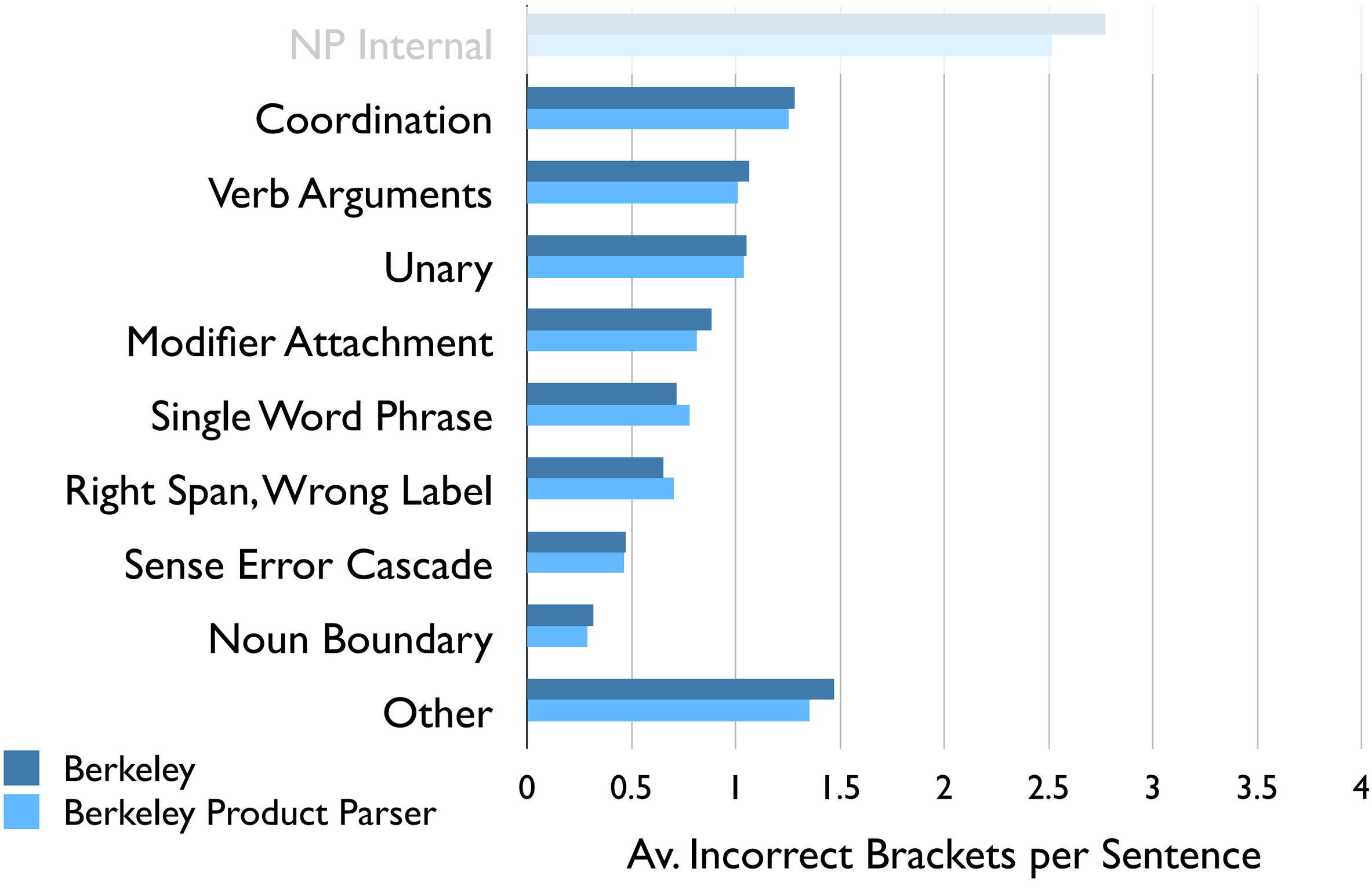
Chinese Parse Errors



Chinese Parse Errors

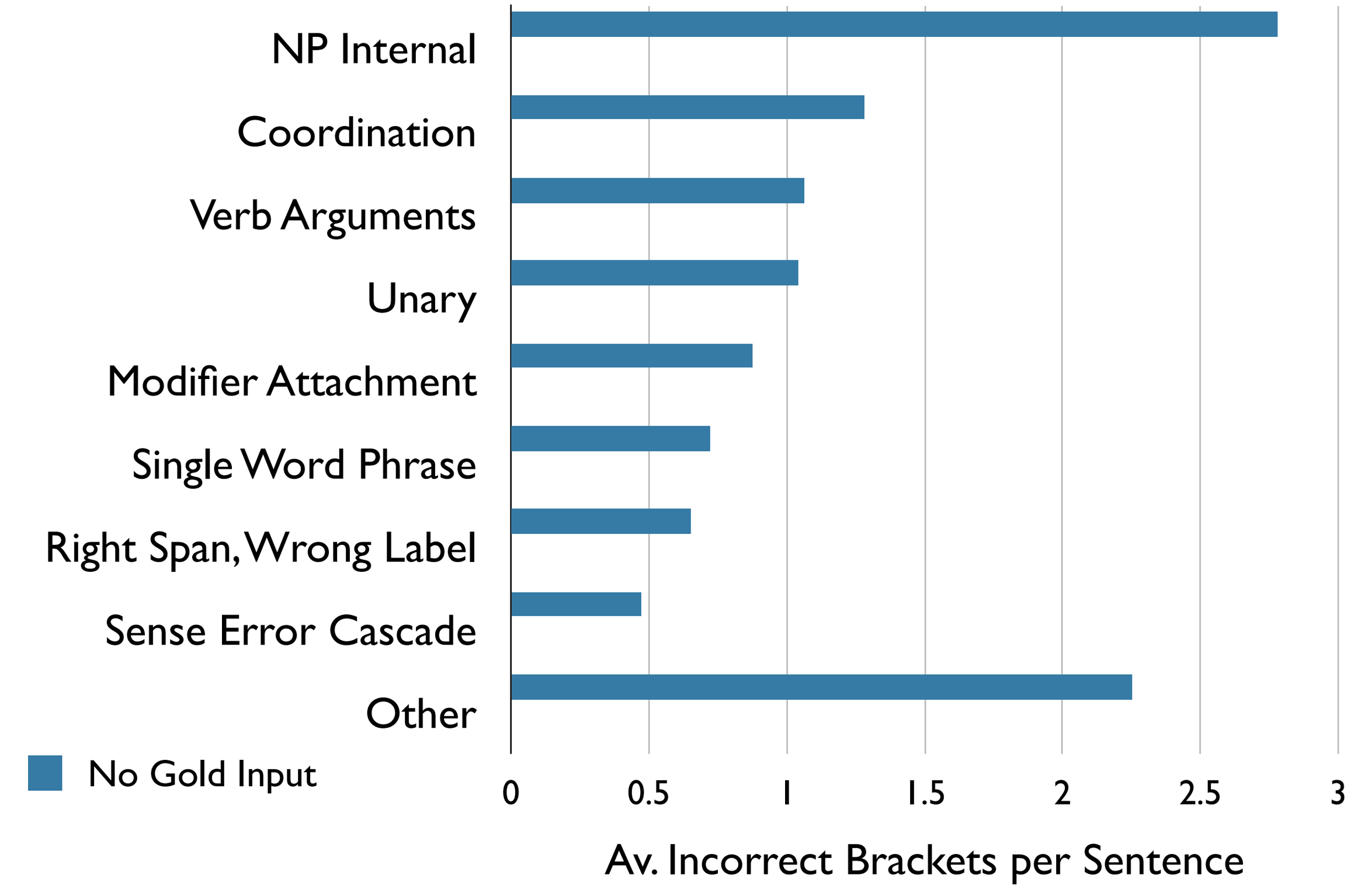


Chinese Parse Errors



Tagging Error Impact

Tagging Error Impact



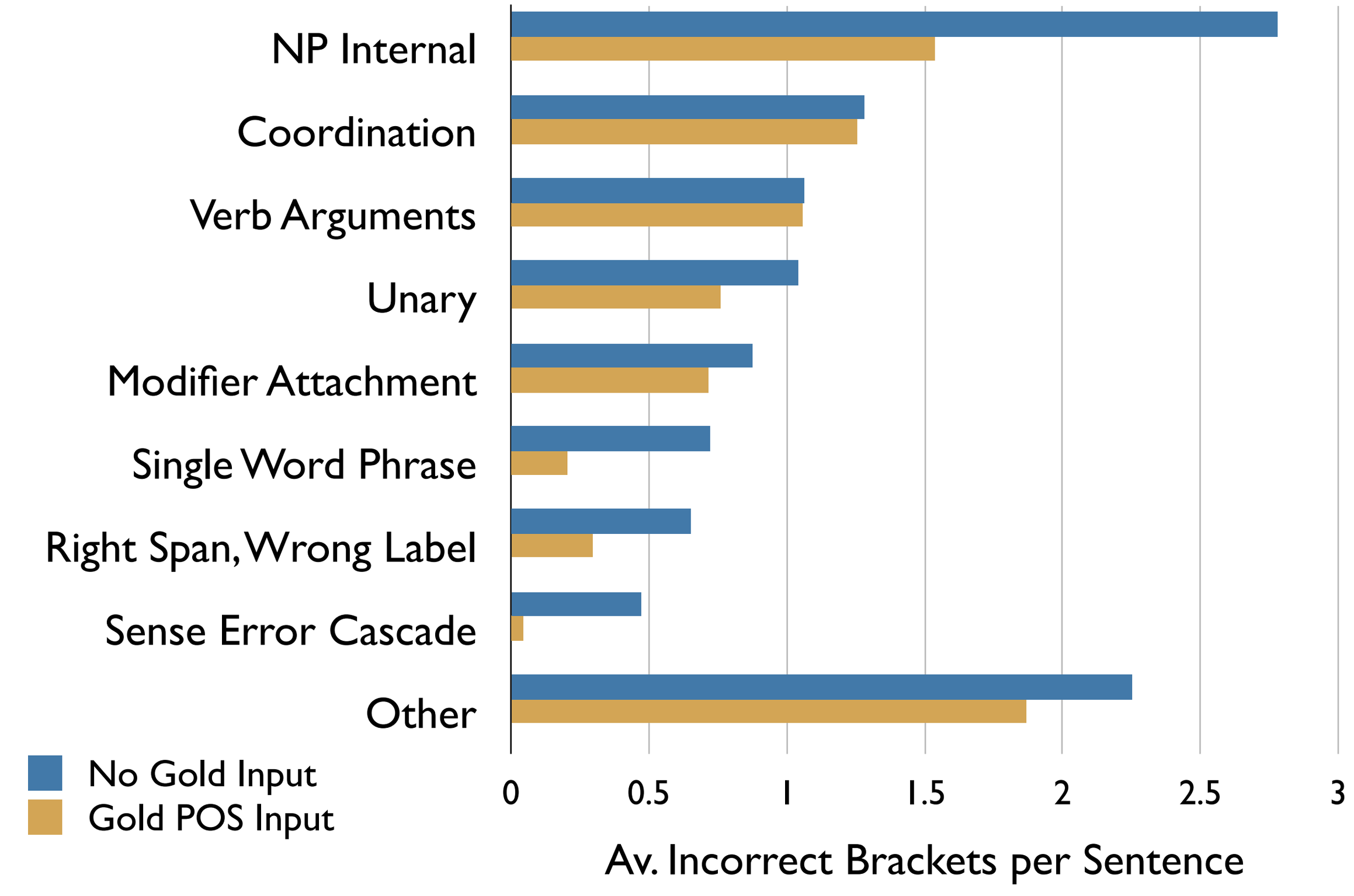
Tagging Error Impact

计划 在 星期六 晚 的 阿拉法特 、 佩雷斯 会晤 …

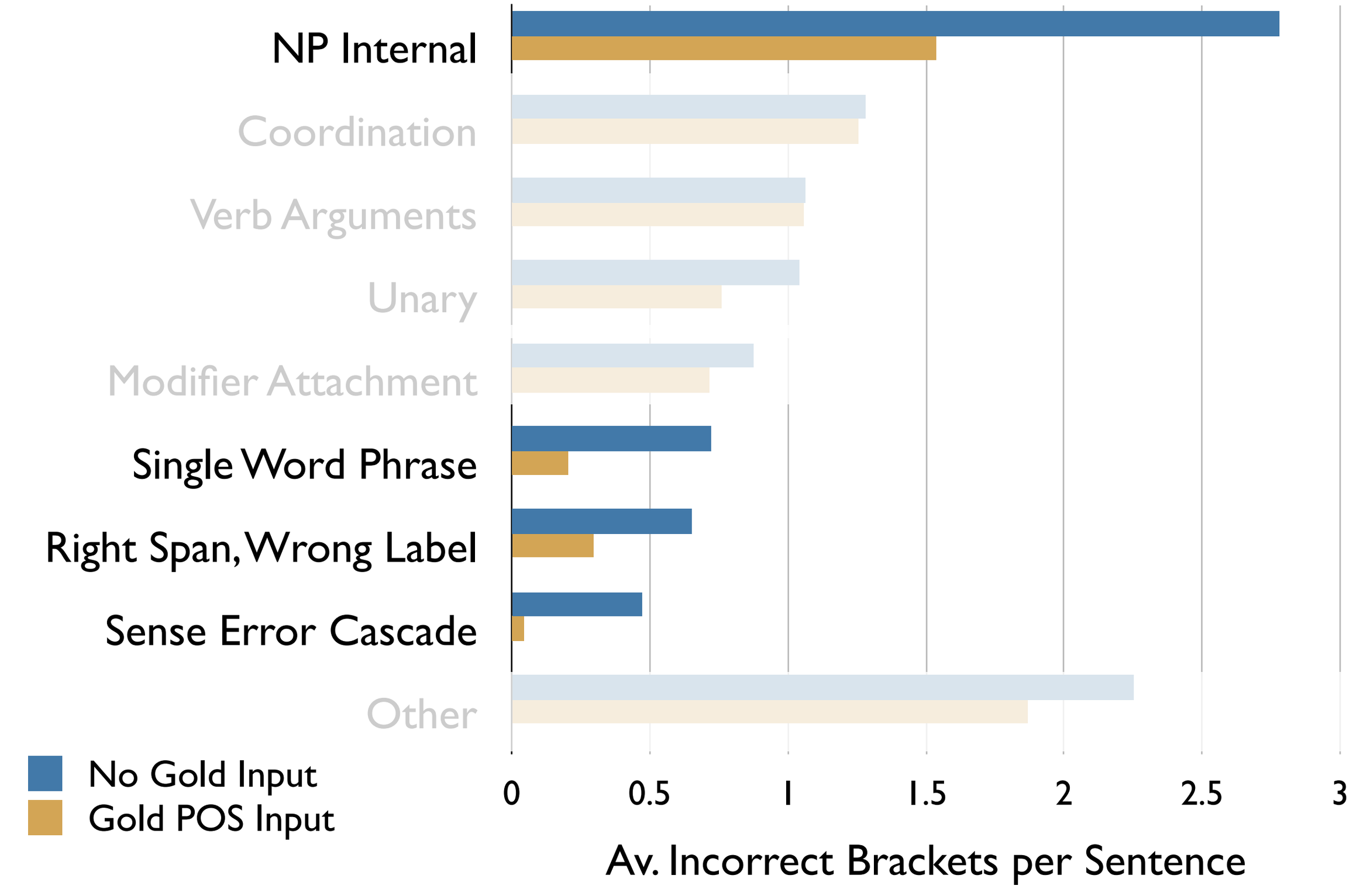
Tagging Error Impact

VV	VV	NT	NT	DEC	NR	PU	NR	NN	
计划	在	星期六	晚	的	阿拉法特	、	佩雷斯	会晤	…

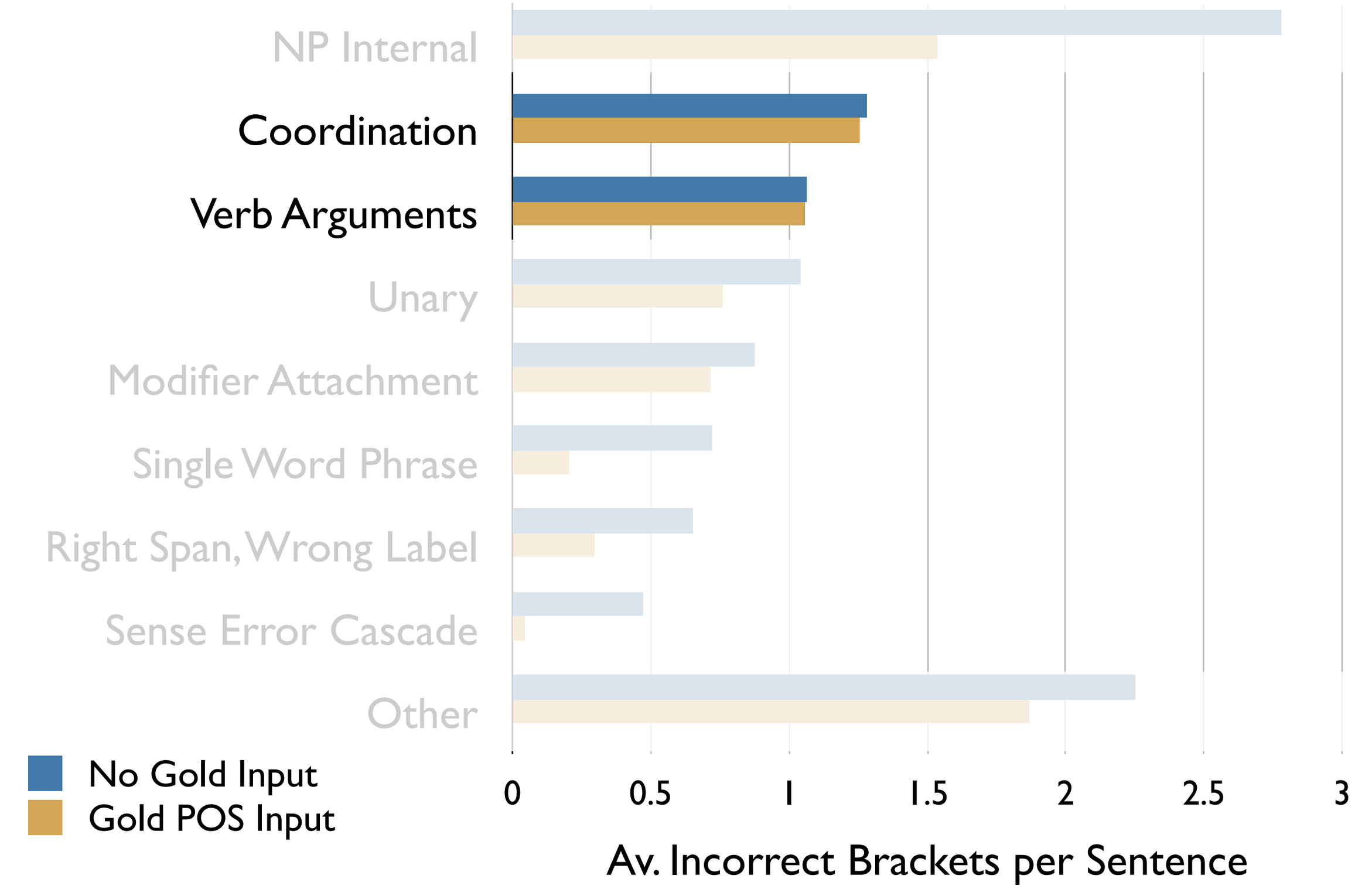
Tagging Error Impact



Tagging Error Impact



Tagging Error Impact



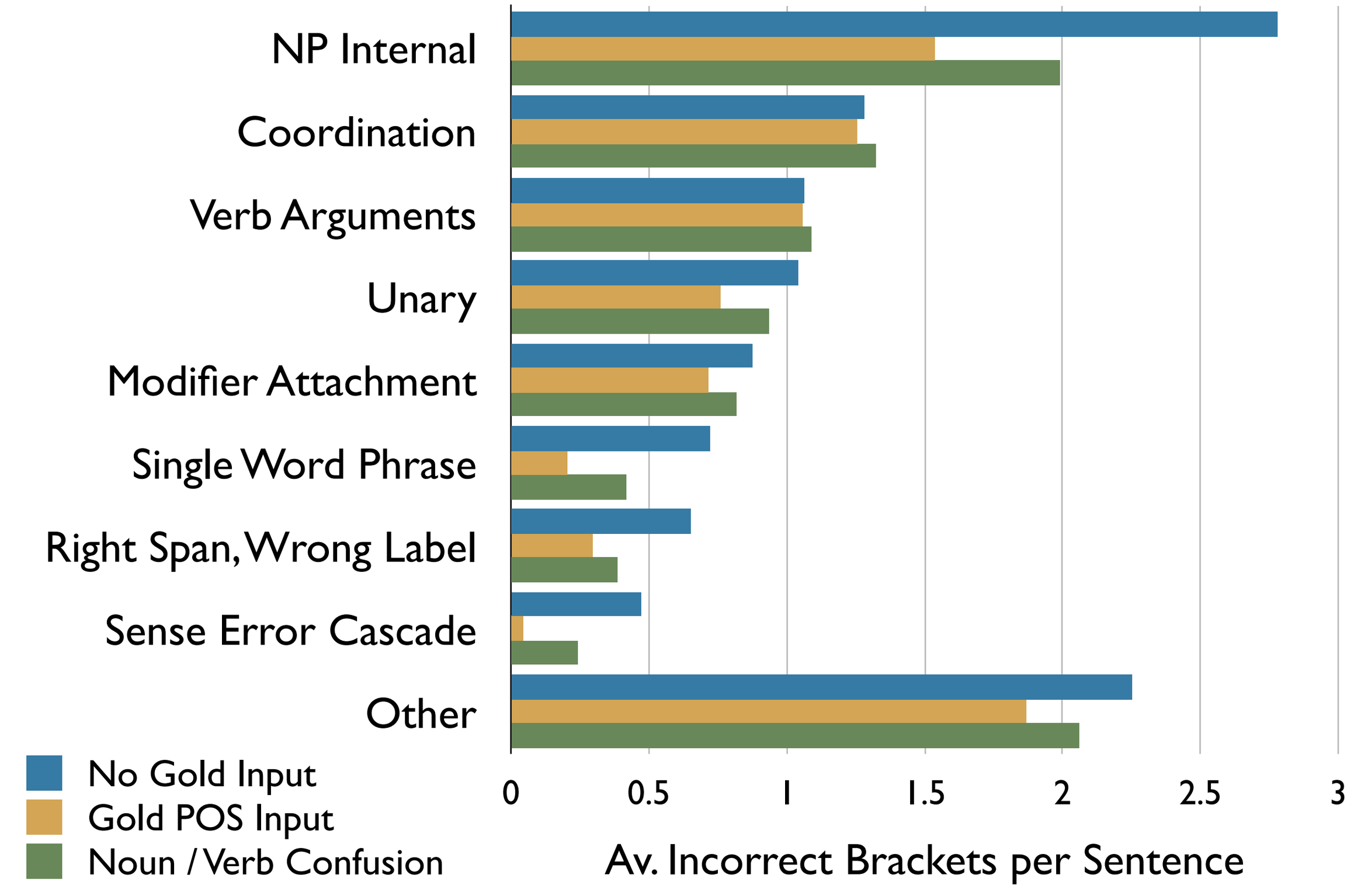
Tagging Error Impact

VV	VV	NT	NT	DEC	NR	PU	NR	NN	
计划	在	星期六	晚	的	阿拉法特	、	佩雷斯	会晤	…

Tagging Error Impact

NN	VV	NT	NT	DEC	NR	PU	NR	VV	…
计划	在	星期六	晚	的	阿拉法特	、	佩雷斯	会晤	…

Tagging Error Impact



Conclusion

Conclusion

- Quantitative validation that Chinese and English parsers deal with different challenges

Conclusion

- Quantitative validation that Chinese and English parsers deal with different challenges
- Perfecting Chinese POS Tagging will only solve certain parsing challenges

Conclusion

- Quantitative validation that Chinese and English parsers deal with different challenges
- Perfecting Chinese POS Tagging will only solve certain parsing challenges
- Download our automatic analysis tool: nlp.cs.berkeley.edu/software

Conclusion

- Quantitative validation that Chinese and English parsers deal with different challenges
- Perfecting Chinese POS Tagging will only solve certain parsing challenges
- Download our automatic analysis tool: nlp.cs.berkeley.edu/software

Thank you!