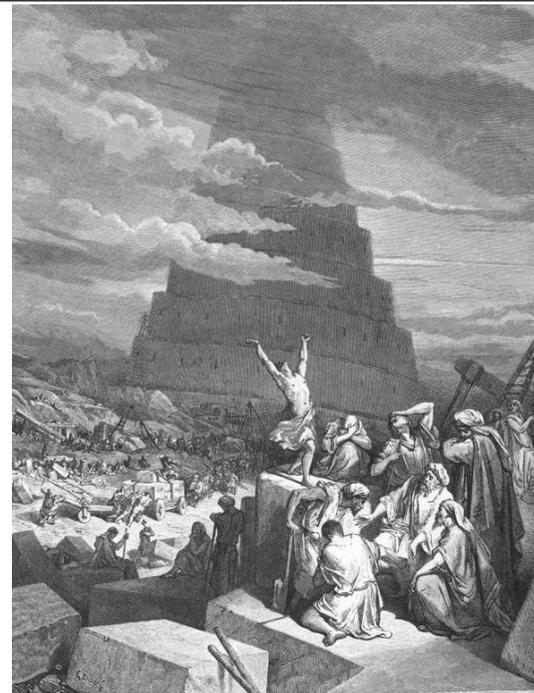


Machine Translation

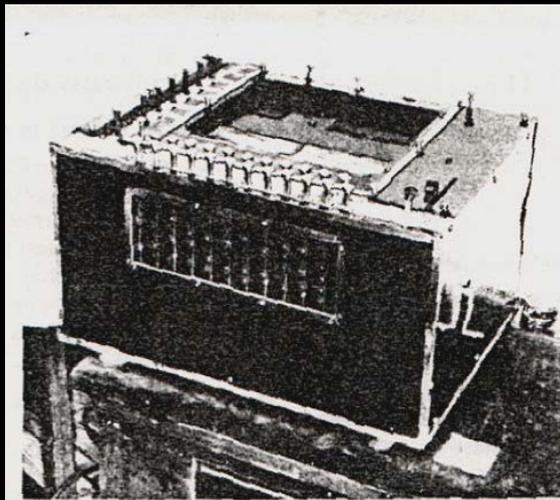
01: Introduction

Rico Sennrich
(slide credit to Adam Lopez)

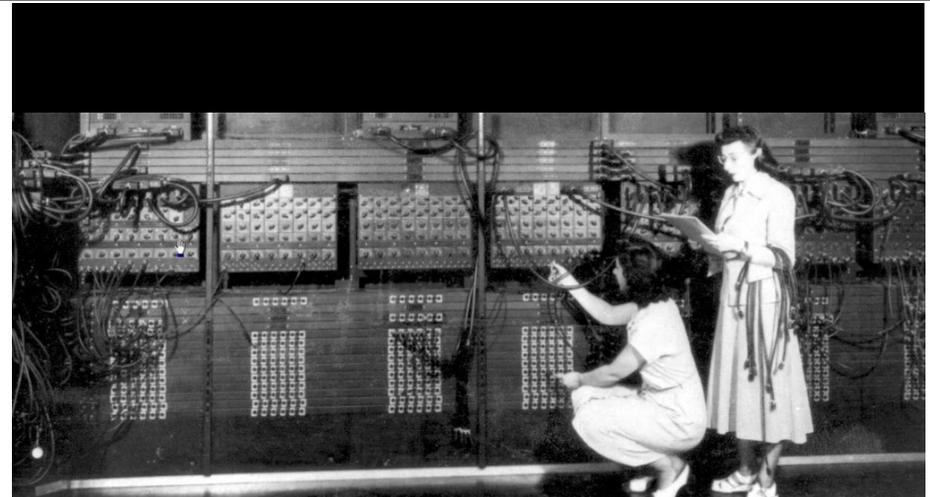
University of Edinburgh



The Confusion of
Tongues
Gustave Doré (1865)



Georges Artsrouni's "mechanical brain",
patented 1933 (France)



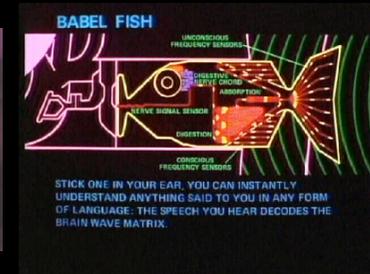
ENIAC (1946), first computer
Gloria Ruth Gordon and Ester Gerson, first programmers



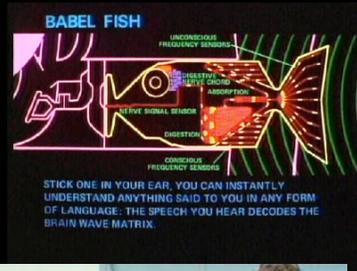
When I look at an article in Russian, I say: "This is really written in English, but it has been coded in some strange symbols. I will now proceed to decode."

Warren Weaver (1949)

Popular view of MT in 2005



Popular view of MT in 2005



Popular view of NLP:

Popular view of MT in 2017



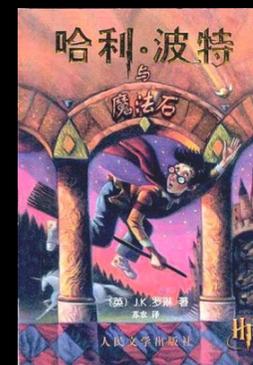
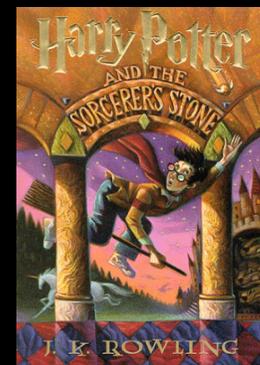
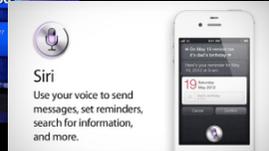
Popular view of MT in 2017



SDL freeTranslation



and NLP:





R. Sennrich

MT – 2018 – 01

11 / 40



R. Sennrich

MT – 2018 – 01

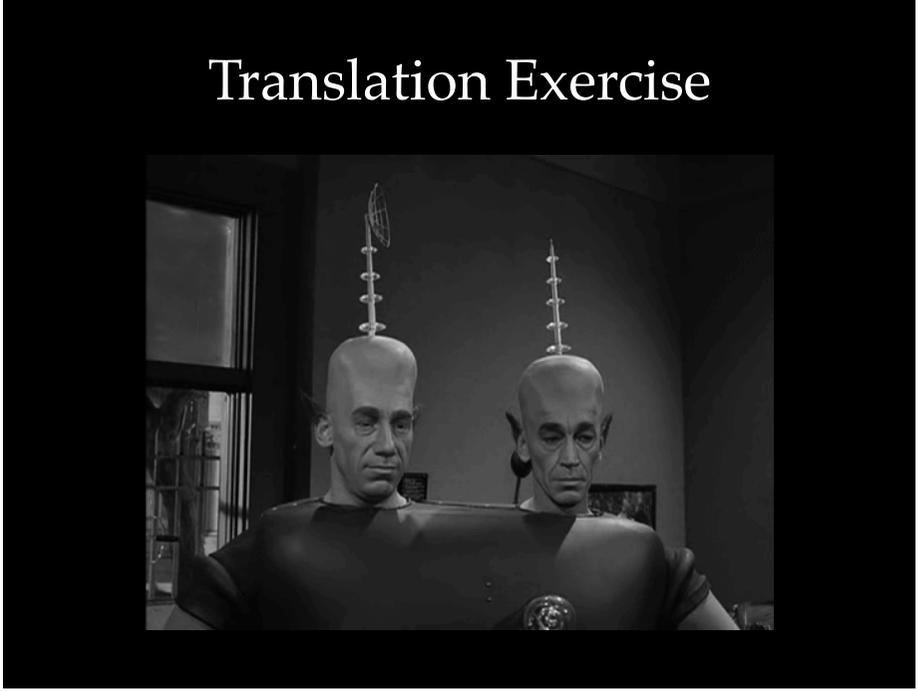
12 / 40



R. Sennrich

MT – 2018 – 01

13 / 40



Translation Exercise

R. Sennrich

MT – 2018 – 01

14 / 40

iat lat pippat eneath hilat oloat at-yurp .



iat lat pippat eneath hilat oloat at-yurp .



iat lat pippat eneath hilat oloat at-yurp .

arrat : hihok	gat : mok	pippat : anak
at-drubel : ok-drubel	hilat : ghirok	quat : izok
at-voon : ok-voon	iat : lalok	rrat : plok
at-yurp : ok-yurp	jjat : farok	sat : drok
bat : klok	krat : jok	totat : erok
bichat : ororok	lat : brok	totat : wiwok
cat : stok	mat : yorok	vat : izok
dat : sprok	nnat : nok	wat : lalok
eneath : enemok	oloath : kantok	zanzanat : zanzanak
forat : rarok		??? : crrrok

iat lat pippat eneath hilat oloat at-yurp .

arrat : hihok	gat : mok	pippat : anak
at-drubel : ok-drubel	hilat : ghirok	quat : izok
at-voon : ok-voon	iat : lalok	rrat : plok
at-yurp : ok-yurp	jjat : farok	sat : drok
bat : klok	krat : jok	totat : erok
bichat : ororok	lat : brok	totat : wiwok
cat : stok	mat : yorok	vat : izok
dat : sprok	nnat : nok	wat : lalok
eneath : enemok	oloath : kantok	zanzanat : zanzanak
forat : rarok		??? : crrrok

iat lat pippat eneat hilat oloat at-yurp .
 lalok brok anak enemok ghirok kantok ok-yurp .

arrat : hihok	gat : mok	pippat : anak
at-drubel : ok-drubel	hilat : ghirok	quat : izok
at-voon : ok-voon	iat : lalok	rrat : plok
at-yurp : ok-yurp	jjat : farok	sat : drok
bat : klok	krat : jok	totat : erok
bichat : ororok	lat : brok	totat : wiwok
cat : stok	mat : yorok	vat : izok
dat : sprok	nmat : nok	wat : lalok
eneat : enemok	oloat : kantok	zanzanat : zanzanok
forat : rarok		??? : crrrok

iat lat pippat eneat hilat oloat at-yurp .
 lalok brok anak enemok ghirok kantok ok-yurp .



iat lat pippat eneat hilat oloat at-yurp .
 lalok anak brok enemok kantok ghirok ok-yurp .



iat lat pippat eneat hilat oloat at-yurp .
 lalok brok anak kantok ghirok enemok ok-yurp .



iat lat pippat eneat hilat oloat at-yurp .
 lalok brok anak kantok ghirok enemok ok-yurp .

5040 possible orderings!

erok sprok izok hihok ghirok .
 totat dat arrat vat hilat .
 ok-voon anak drok brok jok .
 at-voon krat pippat sat lat .

lalok mok nok yorok ghirok klok .
 wat nnat gat mat bat hilat .

lalok sprok izok jok stok .
 wat dat krat quat cat .

lalok rarok nok izok hihok mok .
 wat nnat forat arrat vat gat .

iat lat pippat eneat hilat oloat at-yurp .
 lalok brok anak enemok ghirok kantok ok-yurp .

ok-drubel anak ghirok farok .
 wiwok rarok nok zerok ghirok enemok .
 ok-drubel ziplok stok vok erok enemok kantok ok-yurp zinok jok yorok klok .
 lalok klok izok vok ok-drubel .
 ok-voon ororok sprok .
 ok-drubel ok-voon anak plok sprok .
 erok sprok izok hihok ghirok .
 ok-voon anak drok brok jok .
 wiwok farok izok stok .
 lalok sprok izok jok stok .
 lalok brok anak plok nok .
 lalok farok ororok lalok sprok izok enemok .
 wiwok nok izok kantok ok-yurp .
 lalok mok nok yorok ghirok klok .
 lalok nok crrrok hihok yorok zanzanok .
 lalok rarok nok izok hihok mok .

iat lat pippat eneat hilat oloat at-yurp .
 lalok brok anak enemok ghirok kantok ok-yurp .

ok-drubel anak ghirok farok .
 wiwok rarok nok zerok ghirok enemok .
 ok-drubel ziplok stok vok erok enemok kantok ok-yurp zinok jok yorok klok .
 lalok klok izok vok ok-drubel .
 ok-voon ororok sprok .
 ok-drubel ok-voon anak plok sprok .
 erok sprok izok hihok ghirok .
 ok-voon anak drok brok jok .
 wiwok farok izok stok .
 lalok sprok izok jok stok .
 lalok brok anak plok nok .
 lalok farok ororok lalok sprok izok enemok .
 wiwok nok izok kantok ok-yurp .
 lalok mok nok yorok ghirok klok .
 lalok nok crrrok hihok yorok zanzanok .
 lalok rarok nok izok hihok mok .

iat lat pippat eneat hilat oloat at-yurp .
 lalok brok anak enemok ghirok kantok ok-yurp .

ok-drubel anak ghirok farok .
 wiwok rarok nok zerok ghirok enemok .
 ok-drubel ziplok stok vok erok enemok kantok ok-yurp zinok jok yorok klok .
 lalok klok izok vok ok-drubel .
 ok-voon ororok sprok .
 ok-drubel ok-voon anak plok sprok .
 erok sprok izok hihok ghirok .
 ok-voon anak drok brok jok .
 wiwok farok izok stok .
 lalok sprok izok jok stok .
 lalok brok anak plok nok .
 lalok farok ororok lalok sprok izok enemok .
 wiwok nok izok kantok ok-yurp .
 lalok mok nok yorok ghirok klok .
 lalok nok crrrok hihok yorok zanzanok .
 lalok rarok nok izok hihok mok .

iat lat pippat eneat hilat oloat at-yurp .

lalog brok anok enemok ghirok kantok ok-yurp .

ok-drubel anok ghirok farok .
 wiwok rarok nok zerok ghirok enemok .
 ok-drubel ziplok stok vok erok enemok kantok ok-yurp zinok jok yorok klok .
 lalok klok izok vok ok-drubel .
 ok-voon ororok sprok .
 ok-drubel ok-voon anok plok sprok .
 erok sprok izok hihok ghirok .
 ok-voon anok drok brok jok .
 wiwok farok izok stok .
 lalok sprok izok jok stok .
 lalog brok anok plok nok .
 lalok farok ororok lalok sprok izok enemok .
 wiwok nok izok kantok ok-yurp .
 lalok mok nok yorok ghirok klok .
 lalok nok crrrok hihok yorok zanzanok .
 lalok rarok nok izok hihok mok .

iat lat pippat eneat hilat oloat at-yurp .

lalok brok anok enemok ghirok kantok ok-yurp .

ok-drubel anok ghirok farok .
 wiwok rarok nok zerok ghirok enemok .
 ok-drubel ziplok stok vok erok enemok kantok ok-yurp zinok jok yorok klok .
 lalok klok izok vok ok-drubel .
 ok-voon ororok sprok .
 ok-drubel ok-voon anok plok sprok .
 erok sprok izok hihok ghirok .
 ok-voon anok drok brok jok .
 wiwok farok izok stok .
 lalok sprok izok jok stok .
 lalok brok anok plok nok .
 lalok farok ororok lalok sprok izok enemok .
 wiwok nok izok kantok ok-yurp .
 lalok mok nok yorok ghirok klok .
 lalok nok crrrok hihok yorok zanzanok .
 lalok rarok nok izok hihok mok .

iat lat pippat eneat hilat oloat at-yurp .

lalok brok anok enemok ghirok kantok ok-yurp .

ok-drubel anok ghirok farok .
 wiwok rarok nok zerok ghirok enemok .
 ok-drubel ziplok stok vok erok enemok kantok ok-yurp zinok jok yorok klok .
 lalok klok izok vok ok-drubel .
 ok-voon ororok sprok .
 ok-drubel ok-voon anok plok sprok .
 erok sprok izok hihok ghirok .
 ok-voon anok drok brok jok .
 wiwok farok izok stok .
 lalok sprok izok jok stok .
 lalok brok anok plok nok .
 lalok farok ororok lalok sprok izok enemok .
 wiwok nok izok kantok ok-yurp .
 lalok mok nok yorok ghirok klok .
 lalok nok crrrok hihok yorok zanzanok .
 lalok rarok nok izok hihok mok .

iat lat pippat eneat hilat oloat at-yurp .

lalok brok anok enemok ghirok kantok ok-yurp .

ok-drubel anok ghirok farok .
 wiwok rarok nok zerok ghirok enemok .
 ok-drubel ziplok stok vok erok enemok kantok ok-yurp zinok jok yorok klok .
 lalok klok izok vok ok-drubel .
 ok-voon ororok sprok .
 ok-drubel ok-voon anok plok sprok .
 erok sprok izok hihok ghirok .
 ok-voon anok drok brok jok .
 wiwok farok izok stok .
 lalok sprok izok jok stok .
 lalok brok anok plok nok .
 lalok farok ororok lalok sprok izok enemok .
 wiwok nok izok kantok ok-yurp .
 lalok mok nok yorok ghirok klok .
 lalok nok crrrok hihok yorok zanzanok .
 lalok rarok nok izok hihok mok .

iat lat pippat eneat hilat oloat at-yurp .
lalok brok anak **enemok ghirok** kantok ok-yurp .

ok-drubel anak ghirok farok .
wiwok rarok nok zerok ghirok enemok .
ok-drubel ziplok stok vok erok enemok kantok ok-yurp zinok jok yorok klok .
lalok klok izok vok ok-drubel .
ok-voon ororok sprok .
ok-drubel ok-voon anak plok sprok .
erok sprok izok hihok ghirok .
ok-voon anak drok brok jok .
wiwok farok izok stok .
lalok sprok izok jok stok .
lalok brok anak plok nok .
lalok farok ororok lalok sprok izok enemok .
wiwok nok izok kantok ok-yurp .
lalok mok nok yorok ghirok klok .
lalok nok crrrok hihok yorok zanzanak .
lalok rarok nok izok hihok mok .

iat lat pippat eneat hilat oloat at-yurp .
lalok brok anak **enemok ghirok** kantok ok-yurp .

ok-drubel anak ghirok farok .
wiwok rarok nok zerok **ghirok enemok** .
ok-drubel ziplok stok vok erok enemok kantok ok-yurp zinok jok yorok klok .
lalok klok izok vok ok-drubel .
ok-voon ororok sprok .
ok-drubel ok-voon anak plok sprok .
erok sprok izok hihok ghirok .
ok-voon anak drok brok jok .
wiwok farok izok stok .
lalok sprok izok jok stok .
lalok brok anak plok nok .
lalok farok ororok lalok sprok izok enemok .
wiwok nok izok kantok ok-yurp .
lalok mok nok yorok ghirok klok .
lalok nok crrrok hihok yorok zanzanak .
lalok rarok nok izok hihok mok .

iat lat pippat eneat hilat oloat at-yurp .
lalok brok anak **ghirok enemok** kantok ok-yurp .

ok-drubel anak ghirok farok .
wiwok rarok nok zerok **ghirok enemok** .
ok-drubel ziplok stok vok erok enemok kantok ok-yurp zinok jok yorok klok .
lalok klok izok vok ok-drubel .
ok-voon ororok sprok .
ok-drubel ok-voon anak plok sprok .
erok sprok izok hihok ghirok .
ok-voon anak drok brok jok .
wiwok farok izok stok .
lalok sprok izok jok stok .
lalok brok anak plok nok .
lalok farok ororok lalok sprok izok enemok .
wiwok nok izok kantok ok-yurp .
lalok mok nok yorok ghirok klok .
lalok nok crrrok hihok yorok zanzanak .
lalok rarok nok izok hihok mok .

iat lat pippat eneat hilat oloat at-yurp .
lalok brok anak ghirok **enemok** kantok ok-yurp .

ok-drubel anak ghirok farok .
wiwok rarok nok zerok ghirok enemok .
ok-drubel ziplok stok vok erok enemok kantok ok-yurp zinok jok yorok klok .
lalok klok izok vok ok-drubel .
ok-voon ororok sprok .
ok-drubel ok-voon anak plok sprok .
erok sprok izok hihok ghirok .
ok-voon anak drok brok jok .
wiwok farok izok stok .
lalok sprok izok jok stok .
lalok brok anak plok nok .
lalok farok ororok lalok sprok izok enemok .
wiwok nok izok kantok ok-yurp .
lalok mok nok yorok ghirok klok .
lalok nok crrrok hihok yorok zanzanak .
lalok rarok nok izok hihok mok .

iat lat pippat eneat hilat oloat at-yurp .
 lalok brok **anok ghirok** enemok kantok ok-yurp .

ok-drubel anok ghirok farok .
 wiwok rarok nok zerok ghirok enemok .
 ok-drubel ziplok stok vok erok enemok kantok ok-yurp zinok jok yorok klok .
 lalok klok izok vok ok-drubel .
 ok-voon ororok sprok .
 ok-drubel ok-voon anok plok sprok .
 erok sprok izok hihok ghirok .
 ok-voon anok drok brok jok .
 wiwok farok izok stok .
 lalok sprok izok jok stok .
 lalok brok anok plok nok .
 lalok farok ororok lalok sprok izok enemok .
 wiwok nok izok kantok ok-yurp .
 lalok mok nok yorok ghirok klok .
 lalok nok crrrok hihok yorok zanzanok .
 lalok rarok nok izok hihok mok .

iat lat pippat eneat hilat oloat at-yurp .
 lalok brok **anok ghirok** enemok kantok ok-yurp .

ok-drubel **anok ghirok** farok .
 wiwok rarok nok zerok ghirok enemok .
 ok-drubel ziplok stok vok erok enemok kantok ok-yurp zinok jok yorok klok .
 lalok klok izok vok ok-drubel .
 ok-voon ororok sprok .
 ok-drubel ok-voon anok plok sprok .
 erok sprok izok hihok ghirok .
 ok-voon anok drok brok jok .
 wiwok farok izok stok .
 lalok sprok izok jok stok .
 lalok brok anok plok nok .
 lalok farok ororok lalok sprok izok enemok .
 wiwok nok izok kantok ok-yurp .
 lalok mok nok yorok ghirok klok .
 lalok nok crrrok hihok yorok zanzanok .
 lalok rarok nok izok hihok mok .

iat lat pippat eneat hilat oloat at-yurp .
 lalok brok anok ghirok **enemok kantok** ok-yurp .

ok-drubel anok ghirok farok .
 wiwok rarok nok zerok ghirok enemok .
 ok-drubel ziplok stok vok erok enemok kantok ok-yurp zinok jok yorok klok .
 lalok klok izok vok ok-drubel .
 ok-voon ororok sprok .
 ok-drubel ok-voon anok plok sprok .
 erok sprok izok hihok ghirok .
 ok-voon anok drok brok jok .
 wiwok farok izok stok .
 lalok sprok izok jok stok .
 lalok brok anok plok nok .
 lalok farok ororok lalok sprok izok enemok .
 wiwok nok izok kantok ok-yurp .
 lalok mok nok yorok ghirok klok .
 lalok nok crrrok hihok yorok zanzanok .
 lalok rarok nok izok hihok mok .

iat lat pippat eneat hilat oloat at-yurp .
 lalok brok anok ghirok **enemok kantok** ok-yurp .

ok-drubel anok ghirok farok .
 wiwok rarok nok zerok ghirok enemok .
 ok-drubel ziplok stok vok erok enemok kantok ok-yurp zinok jok yorok klok .
 lalok klok izok vok ok-drubel .
 ok-voon ororok sprok .
 ok-drubel ok-voon anok plok sprok .
 erok sprok izok hihok ghirok .
 ok-voon anok drok brok jok .
 wiwok farok izok stok .
 lalok sprok izok jok stok .
 lalok brok anok plok nok .
 lalok farok ororok lalok sprok izok enemok .
 wiwok nok izok kantok ok-yurp .
 lalok mok nok yorok ghirok klok .
 lalok nok crrrok hihok yorok zanzanok .
 lalok rarok nok izok hihok mok .

iat lat pippat eneat hilat oloat at-yurp .
lalok brok anak ghirok enemok kantok ok-yurp .

ok-drubel anak ghirok farok .
wiwok rarok nok zerok ghirok enemok .
ok-drubel ziplok stok vok erok enemok kantok ok-yurp zinok jok yorok klok .
lalok klok izok vok ok-drubel .
ok-voon ororok sprok .
ok-drubel ok-voon anak plok sprok .
erok sprok izok hihok ghirok .
ok-voon anak drok brok jok .
wiwok farok izok stok .
lalok sprok izok jok stok .
lalok brok anak plok nok .
lalok farok ororok lalok sprok izok enemok .
wiwok nok izok kantok ok-yurp .
lalok mok nok yorok ghirok klok .
lalok nok crrrok hihok yorok zanzanok .
lalok rarok nok izok hihok mok .

iat lat pippat eneat hilat oloat at-yurp .
lalok brok anak ghirok enemok kantok ok-yurp .

ok-drubel anak ghirok farok .
wiwok rarok nok zerok ghirok enemok .
ok-drubel ziplok stok vok erok enemok kantok ok-yurp zinok jok yorok klok .
lalok klok izok vok ok-drubel .
ok-voon ororok sprok .
ok-drubel ok-voon anak plok sprok .
erok sprok izok hihok ghirok .
ok-voon anak drok brok jok .
wiwok farok izok stok .
lalok sprok izok jok stok .
lalok brok anak plok nok .
lalok farok ororok lalok sprok izok enemok .
wiwok nok izok kantok ok-yurp .
lalok mok nok yorok ghirok klok .
lalok nok crrrok hihok yorok zanzanok .
lalok rarok nok izok hihok mok .

totat nnat forat arrat mat bat .

totat nnat forat arrat mat bat .

arrat : hihok	gat : mok	pippat : anak
at-drubel : ok-drubel	hilat : ghirook	quat : izok
at-voon : ok-voon	iat : lalok	rrat : plok
at-yurp : ok-yurp	jjat : farok	sat : drok
bat : klok	krat : jok	totat : erok
bichat : ororok	lat : brok	totat : wiwok
cat : stok	mat : yorok	vat : izok
dat : sprok	nnat : nok	wat : lalok
eneat : enemok	oloat : kantok	zanzanat : zanzanok
forat : rarok		??? : crrrok

totat nnat forat arrat mat bat .

arrat : hihok	gat : mok	pippat : anak
at-drubel : ok-drubel	hilat : ghirook	quat : izok
at-voon : ok-voon	iat : lalok	rrat : plok
at-yurp : ok-yurp	jjat : farok	sat : drok
bat : klok	krat : jok	totat : erok
bichat : ororok	lat : brok	totat : wiwok
cat : stok	mat : yorok	vat : izok
dat : sprok	nnat : nok	wat : lalok
eneat : enemok	oloat : kantok	zanzanat : zanzanok
forat : rarok		??? : crrrok

totat : erok
totat : wiwok

totat nnat forat arrat mat bat .
erok | wiwok nok rarok hihok yorok klok .

arrat : hihok	gat : mok	pippat : anak
at-drubel : ok-drubel	hilat : ghirook	quat : izok
at-voon : ok-voon	iat : lalok	rrat : plok
at-yurp : ok-yurp	jjat : farok	sat : drok
bat : klok	krat : jok	totat : erok
bichat : ororok	lat : brok	totat : wiwok
cat : stok	mat : yorok	vat : izok
dat : sprok	nnat : nok	wat : lalok
eneat : enemok	oloat : kantok	zanzanat : zanzanok
forat : rarok		??? : crrrok

totat nnat forat arrat mat bat .
erok | wiwok nok rarok hihok yorok klok .

ok-drubel anak ghirook farok .
 wiwok rarok nok zerok ghirook enemok .
 ok-drubel ziplok stok vok erok enemok kantok ok-yurp zinok jok yorok klok .
 lalok klok izok vok ok-drubel .
 ok-voon ororok sprok .
 ok-drubel ok-voon anak plok sprok .
 erok sprok izok hihok ghirook .
 ok-voon anak drok brok jok .
 wiwok farok izok stok .
 lalok sprok izok jok stok .
 lalok brok anak plok nok .
 lalok farok ororok lalok sprok izok enemok .
 wiwok nok izok kantok ok-yurp .
 lalok mok nok yorok ghirook klok .
 lalok nok crrrok hihok yorok zanzanok .
 lalok rarok nok izok hihok mok .

total nnat forat arrat mat bat .
erok | **wiwok nok** rarok hihok yorok klok .

ok-drubel anak ghirok farok .
wiwok rarok nok zerok ghirok enemok .
ok-drubel ziplok stok vok erok enemok kantok ok-yurp zinok jok yorok klok .
lalok klok izok vok ok-drubel .
ok-voon ororok sprok .
ok-drubel ok-voon anak plok sprok .
erok sprok izok hihok ghirok .
ok-voon anak drok brok jok .
wiwok farok izok stok .
lalok sprok izok jok stok .
lalok brok anak plok nok .
lalok farok ororok lalok sprok izok enemok .
wiwok nok izok kantok ok-yurp .
lalok mok nok yorok ghirok klok .
lalok nok crrrok hihok yorok zanzanok .
lalok rarok nok izok hihok mok .

total nnat forat arrat mat bat .
wiwok nok rarok hihok yorok klok .

ok-drubel anak ghirok farok .
wiwok rarok nok zerok ghirok enemok .
ok-drubel ziplok stok vok erok enemok kantok ok-yurp zinok jok yorok klok .
lalok klok izok vok ok-drubel .
ok-voon ororok sprok .
ok-drubel ok-voon anak plok sprok .
erok sprok izok hihok ghirok .
ok-voon anak drok brok jok .
wiwok farok izok stok .
lalok sprok izok jok stok .
lalok brok anak plok nok .
lalok farok ororok lalok sprok izok enemok .
wiwok nok izok kantok ok-yurp .
lalok mok nok yorok ghirok klok .
lalok nok crrrok hihok yorok zanzanok .
lalok rarok nok izok hihok mok .

total nnat forat arrat mat bat .
wiwok rarok nok hihok yorok klok .

ok-drubel anak ghirok farok .
wiwok rarok nok zerok ghirok enemok .
ok-drubel ziplok stok vok erok enemok kantok ok-yurp zinok jok yorok klok .
lalok klok izok vok ok-drubel .
ok-voon ororok sprok .
ok-drubel ok-voon anak plok sprok .
erok sprok izok hihok ghirok .
ok-voon anak drok brok jok .
wiwok farok izok stok .
lalok sprok izok jok stok .
lalok brok anak plok nok .
lalok farok ororok lalok sprok izok enemok .
wiwok nok izok kantok ok-yurp .
lalok mok nok yorok ghirok klok .
lalok nok crrrok hihok yorok zanzanok .
lalok rarok nok izok hihok mok .

total nnat forat arrat mat bat .
wiwok rarok nok hihok yorok klok .

ok-drubel anak ghirok farok .
wiwok rarok nok zerok ghirok enemok .
ok-drubel ziplok stok vok erok enemok kantok ok-yurp zinok jok yorok klok .
lalok klok izok vok ok-drubel .
ok-voon ororok sprok .
ok-drubel ok-voon anak plok sprok .
erok sprok izok hihok ghirok .
ok-voon anak drok brok jok .
wiwok farok izok stok .
lalok sprok izok jok stok .
lalok brok anak plok nok .
lalok farok ororok lalok sprok izok enemok .
wiwok nok izok kantok ok-yurp .
lalok mok nok yorok ghirok klok .
lalok nok crrrok hihok yorok zanzanok .
lalok rarok nok izok hihok mok .

totat nnat forat arrat mat bat .
wiwok rarok **nok hihok** yorok klok .

ok-drubel anak ghirok farok .
wiwok rarok nok zerok ghirok enemok .
ok-drubel ziplok stok vok erok enemok kantok ok-yurp zinok jok yorok klok .
lalok klok izok vok ok-drubel .
ok-voon ororok sprok .
ok-drubel ok-voon anak plok sprok .
erok sprok izok hihok ghirok .
ok-voon anak drok brok jok .
wiwok farok izok stok .
lalok sprok izok jok stok .
lalok brok anak plok nok .
lalok farok ororok lalok sprok izok enemok .
wiwok nok izok kantok ok-yurp .
lalok mok nok yorok ghirok klok .
lalok nok crrrok hihok yorok zanzanak .
lalok rarok nok izok hihok mok .

totat nnat forat arrat mat bat .
wiwok rarok **nok hihok** yorok klok .

ok-drubel anak ghirok farok .
wiwok rarok nok zerok ghirok enemok .
ok-drubel ziplok stok vok erok enemok kantok ok-yurp zinok jok yorok klok .
lalok klok izok vok ok-drubel .
ok-voon ororok sprok .
ok-drubel ok-voon anak plok sprok .
erok sprok izok hihok ghirok .
ok-voon anak drok brok jok .
wiwok farok izok stok .
lalok sprok izok jok stok .
lalok brok anak plok nok .
lalok farok ororok lalok sprok izok enemok .
wiwok nok izok kantok ok-yurp .
lalok mok nok yorok ghirok klok .
lalok **nok crrrok hihok yorok** zanzanak .
lalok rarok nok izok hihok mok .

totat nnat forat arrat mat bat .
wiwok rarok **nok crrrok hihok yorok** klok .

ok-drubel anak ghirok farok .
wiwok rarok nok zerok ghirok enemok .
ok-drubel ziplok stok vok erok enemok kantok ok-yurp zinok jok yorok klok .
lalok klok izok vok ok-drubel .
ok-voon ororok sprok .
ok-drubel ok-voon anak plok sprok .
erok sprok izok hihok ghirok .
ok-voon anak drok brok jok .
wiwok farok izok stok .
lalok sprok izok jok stok .
lalok brok anak plok nok .
lalok farok ororok lalok sprok izok enemok .
wiwok nok izok kantok ok-yurp .
lalok mok nok yorok ghirok klok .
lalok **nok crrrok hihok yorok** zanzanak .
lalok rarok nok izok hihok mok .

totat nnat forat arrat mat bat .
wiwok rarok nok crrrok hihok yorok klok .

ok-drubel anak ghirok farok .
wiwok rarok nok zerok ghirok enemok .
ok-drubel ziplok stok vok erok enemok kantok ok-yurp zinok jok yorok klok .
lalok klok izok vok ok-drubel .
ok-voon ororok sprok .
ok-drubel ok-voon anak plok sprok .
erok sprok izok hihok ghirok .
ok-voon anak drok brok jok .
wiwok farok izok stok .
lalok sprok izok jok stok .
lalok brok anak plok nok .
lalok farok ororok lalok sprok izok enemok .
wiwok nok izok kantok ok-yurp .
lalok mok nok yorok ghirok klok .
lalok nok crrrok hihok yorok zanzanak .
lalok rarok nok izok hihok mok .

total nnat forat arrat mat bat .
 wiwok rarok nok crrrok hihok **yorok klok** .

ok-drubel anak ghirok farok .
 wiwok rarok nok zerok ghirok enemok .
 ok-drubel ziplok stok vok erok enemok kantok ok-yurp zinok jok **yorok klok** .
 lalok klok izok vok ok-drubel .
 ok-voon ororok sprok .
 ok-drubel ok-voon anak plok sprok .
 erok sprok izok hihok ghirok .
 ok-voon anak drok brok jok .
 wiwok farok izok stok .
 lalok sprok izok jok stok .
 lalok brok anak plok nok .
 lalok farok ororok lalok sprok izok enemok .
 wiwok nok izok kantok ok-yurp .
 lalok mok nok yorok ghirok klok .
 lalok nok crrrok hihok yorok zanzanok .
 lalok rarok nok izok hihok mok .

total nnat forat arrat mat bat .
 wiwok rarok nok crrrok hihok yorok klok .

ok-drubel anak ghirok farok .
 wiwok rarok nok zerok ghirok enemok .
 ok-drubel ziplok stok vok erok enemok kantok ok-yurp zinok jok yorok klok .
 lalok klok izok vok ok-drubel .
 ok-voon ororok sprok .
 ok-drubel ok-voon anak plok sprok .
 erok sprok izok hihok ghirok .
 ok-voon anak drok brok jok .
 wiwok farok izok stok .
 lalok sprok izok jok stok .
 lalok brok anak plok nok .
 lalok farok ororok lalok sprok izok enemok .
 wiwok nok izok kantok ok-yurp .
 lalok mok nok yorok ghirok klok .
 lalok nok crrrok hihok yorok zanzanok .
 lalok rarok nok izok hihok mok .

Is this exercise realistic?

In real natural language...

Is this exercise realistic?

In real natural language...

...words are ambiguous

English: bank

French: banque



French: rive



Is this exercise realistic?

In real natural language...

...words may have non-standard *orthography*

dats why pluto is pluto it can neva b a star

Is this exercise realistic?

In real natural language...

...word order matters

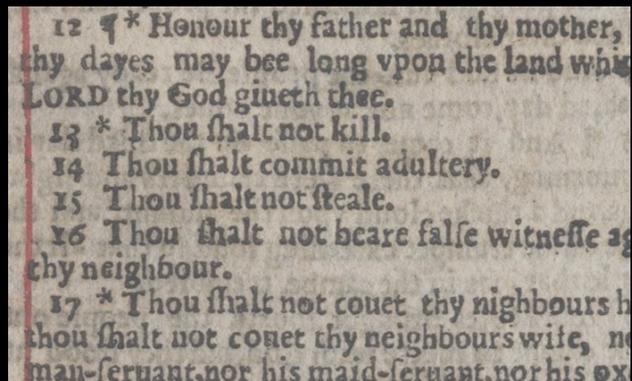
Jaime loves Casey

Casey loves Jaime

Is this exercise realistic?

In real natural language...

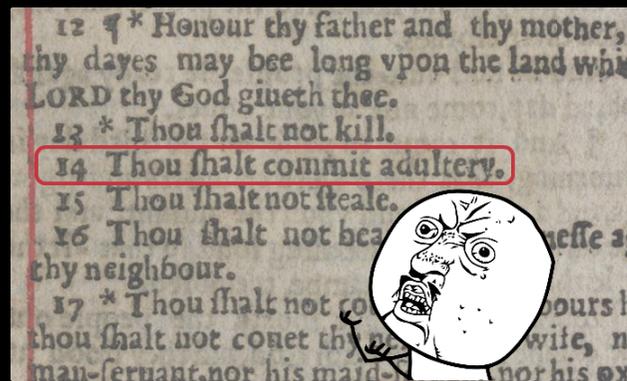
...every word matters



Is this exercise realistic?

In real natural language...

...every word matters



Is this exercise realistic?

In real natural language...

...sentences are longer.

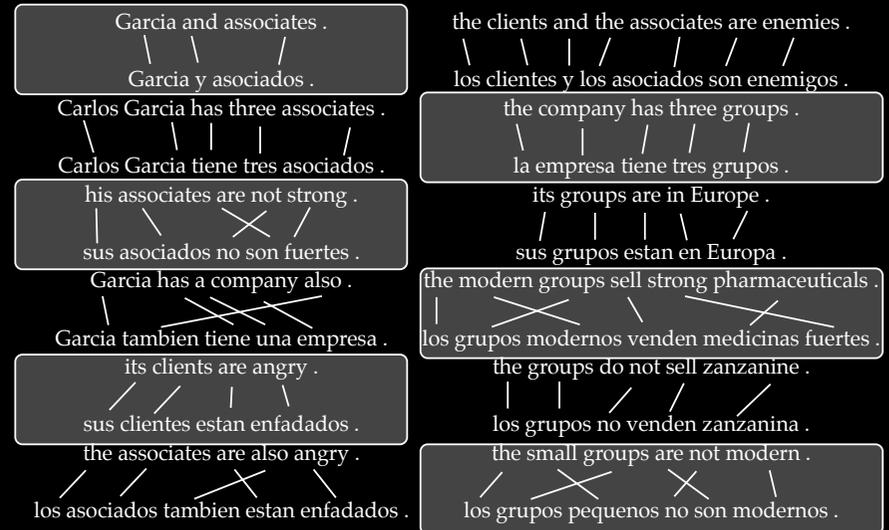
One of Scotland's leading cultural heritage groups will stage the Mary, Queen of Scots Festival, which is hoped to become a major new attraction for the area, which has lost the T in the Park music festival this summer after 18 years.

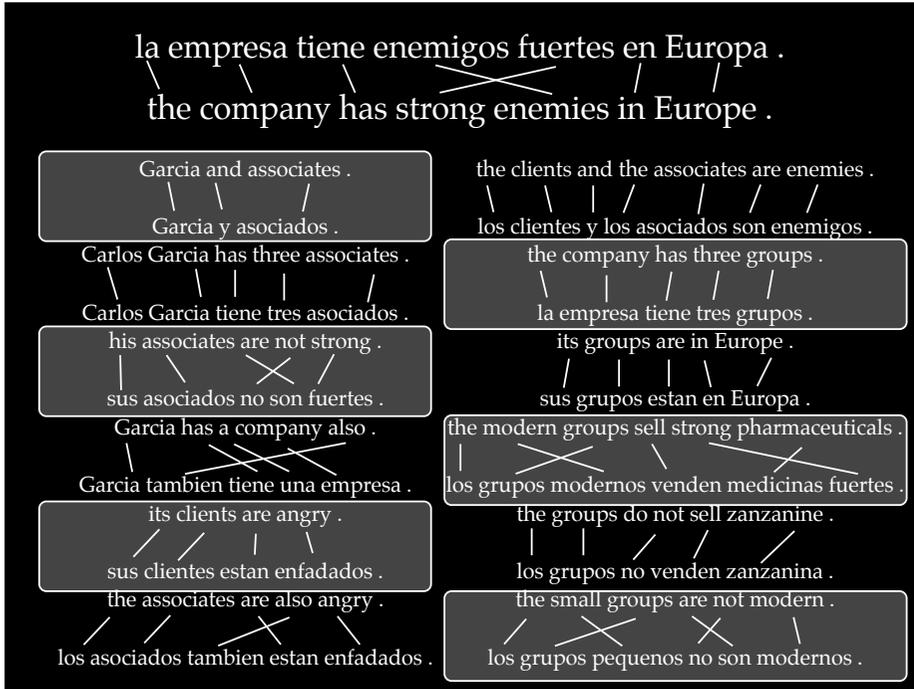
Is this exercise realistic?

iat lat pippat eneath hilat oloat at-yurp .
lalok brok anak ghrok enemok kantok ok-yurp .



la empresa tiene enemigos fuertes en Europa .





Goals of this Course

at the end of this course, you will:

- understand the linguistic and computational challenges of MT
- know how MT is formalized as a machine learning problem
- have experience reading, using and modifying an MT implementation
- understand state-of-the-art approaches to MT and their limitations

our focus:

- sequence-to-sequence models with deep learning
→ applications beyond MT
- probabilistic models of language

“Probabilistic models of language”

This course is about **implementing...**

```
for ej, logprob, fi, fj in alignments[ej]:
    if bitmap(range(fi, fj)) & v == 0:
        new_v = bitmap(range(fi, fj)) | v
        if new_v in chart[ej]:
            chart[ej][new_v] = logadd10(chart[ej][new_v], sums[v]+logprob)
        else:
            chart[ej][new_v] = sums[v]+logprob
```

...**probabilistic models...**

$$p(\theta) = p_{\theta}(\theta) = \prod_{n=1}^N p_{\theta}(\mathbf{f}^{(n)}, \mathbf{a}^{(n)} | \mathbf{e}^{(n)}) = \prod_{n=1}^N p(I^{(n)} | J^{(n)}) \prod_{i=1}^{I^{(n)}} p(a_i^{(n)} | J^{(n)}) p(f_i^{(n)} | e_{a_i}^{(n)})$$

...of **language**

Ngaragana-nguja ngiy-a gujinganjanga-ni jiyawu ngabulu
grog-PROP.IV.ACC 3.SG.NM.A-PST morthor.II.ERG give milk.IV.ACC
‘(His) mother gave (him) milk with grog in it.’ [wmb]

“Probabilistic models of language”

Very few people are experts in computer science, mathematics, and linguistics, and I **guarantee** you will encounter something you don’t know. But if you are unwilling to engage with all of these topics,

DO NOT TAKE THIS COURSE

“Probabilistic models of language”

To future students: There is no way that you should take this course. The reason why is that you need to be competent in linguistics, mathematics and Python programming to do well in the coursework. The lecturer neglected to mention that he expected people to have those skills, but it was clear from the assignments. You will get no help, as there are no tutorials, there is only a forum where you can ask questions, but when a student asked for help because they did not understand the material, the lecturer was not willing to make any changes.

— 2015 course review from someone who didn't listen to the advice on the previous slide.

“Probabilistic models of language”

The teaching team is willing to answer questions on piazza. You will get more out of this if you also answer questions.

If you want answers to questions you must be **specific**. This is not specific (nor is it a question):

Need more help

I would like to write on behalf of a group of people who are finding this assignment impossible.

Requirements

pre-requisites

It is RECOMMENDED that students have passed

- Advanced Natural Language Processing (INFR11059) OR
- Accelerated Natural Language Processing (INFR11125) OR
- Foundations of Natural Language Processing (INFR09028)

if you have no NLP background, you'll have a bad time

Requirements

other requirements

Students are expected to understand the following topics, or be prepared to learn them independently:

- Discrete mathematics: analysis of algorithms, dynamic programming, basic graph algorithms, finite and pushdown automata.
- Other essential maths: basic probability theory; basic calculus and linear algebra; ability to read and manipulate mathematical notation including sums, products, log, and exp.
- Programming: ability to read and modify python programs; ability to design and implement a function based on high-level description such as pseudocode or a precise mathematical statement of what the function computes.
- Linguistics: willingness to learn basic elements of linguistic description; no formal linguistics background required.

warning: you will need these skills to pass the coursework and exam

Still Interested?

course overview

- \approx 15 lectures, Monday and Thursday 4-5pm
- course work, due in week 8 (March 15, 3pm) – 30% of final grade
- written exam – 70% of final grade

tasks until Thursday

- sign up for piazza
- review the course syllabus
- review Koehn, 13.2-3