

# EN.601.647/667

# INTRODUCTION TO HLT

# MACHINE TRANSLATION

---

KENTON MURRAY

9/28/2021

Slides in part thanks to:  
Kevin Duh  
David Chiang  
Kevin Knight

# farok crrrok hihok yorok clok kantok ok-yurp .

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

# farok crrrok hihok yorok clok kantok ok-yurp .

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

# farok crrrok hihok yorok clok kantok ok-yurp .

jjat

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

# farok crrrok hihok yorok clok kantok ok-yurp .

jjat

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

# farok crrrok hihok yorok clok kantok ok-yurp .

jjat

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

# farok crrrok hihok yorok clok kantok ok-yurp .

jjat

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

# farok crrrok **hihok** yorok clok kantok ok-yurp .

jjat

ok-voon ororok sprok .	lalok farok ororok lalok sprok izok enemok .
at-voon bichat dat .	wat jjat bichat wat dat vat eneat .
ok-drubel ok-voon anok plok sprok .	lalok brok anok plok nok .
at-drubel at-voon pippat rrat dat .	iat lat pippat rrat nnat .
erok sprok izok <b>hihok</b> ghirok .	wiwok nok izok kantok ok-yurp .
totat dat arrat vat hilat .	totat nnat quat oloat at-yurp .
ok-voon anok drok brok jok .	lalok mok nok yorok ghirok clok .
at-voon krat pippat sat lat .	wat nnat gat mat bat hilat .
wiwok farok izok stok .   totat jjat quat cat .	lalok nok crrrok <b>hihok</b> yorok zanzanok . wat nnat arrat mat zanzanat .
lalok sprok izok jok stok .	lalok rarok nok izok <b>hihok</b> mok .
wat dat krat quat cat .	wat nnat forat arrat vat gat .

# farok crrrok hihok yorok clok kantok ok-yurp .

jjat

arrat

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

# farok crrrok hihok yorok clok kantok ok-yurp .

jjat

arrat

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

# farok crrrok hihok **yorok** clok kantok ok-yurp .

jjat

arrat

ok-voon ororok sprok .	lalok farok ororok lalok sprok izok enemok .
at-voon bichat dat .	wat jjat bichat wat dat vat eneat .
ok-drubel ok-voon anok plok sprok .	lalok brok anok plok nok .
at-drubel at-voon pippat rrat dat .	iat lat pippat rrat nnat .
erok sprok izok hihok ghirok .	wiwok nok izok kantok ok-yurp .
totat dat arrat vat hilat .	totat nnat quat oloat at-yurp .
ok-voon anok drok brok jok .	lalok mok nok <b>yorok</b> ghirok clok .
at-voon krat pippat sat lat .	wat nnat gat mat bat hilat .
wiwok farok izok stok .	lalok nok crrrok hihok <b>yorok</b> zanzanok .
totat jjat quat cat .	wat nnat arrat mat zanzanat .
lalok sprok izok jok stok .	lalok rarok nok izok hihok mok .
wat dat krat quat cat .	wat nnat forat arrat vat gat .

# farok crrrok hihok **yorok** clok kantok ok-yurp .

jjat

arrat

mat

ok-voon ororok sprok .	lalok farok ororok lalok sprok izok enemok .
at-voon bichat dat .	wat jjat bichat wat dat vat eneat .
ok-drubel ok-voon anok plok sprok .	lalok brok anok plok nok .
at-drubel at-voon pippat rrat dat .	iat lat pippat rrat nnat .
erok sprok izok hihok ghirok .	wiwok nok izok kantok ok-yurp .
totat dat arrat vat hilat .	totat nnat quat oloat at-yurp .
ok-voon anok drok brok jok .	lalok mok nok <b>yorok</b> ghirok clok .
at-voon krat pippat sat lat .	wat nnat gat mat bat hilat .
wiwok farok izok stok .	lalok nok crrrok hihok <b>yorok</b> zanzanok .
totat jjat quat cat .	wat nnat arrat mat zanzanat .
lalok sprok izok jok stok .	lalok rarok nok izok hihok mok .
wat dat krat quat cat .	wat nnat forat arrat vat gat .

# farok crrrok hihok yorok clok kantok ok-yurp .

jjat

arrat

mat

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

# farok crrrok hihok yorok **clok** kantok ok-yurp .

jjat

arrat

mat

ok-voon ororok sprok .	lalok farok ororok lalok sprok izok enemok .
at-voon bichat dat .	wat jjat bichat wat dat vat eneat .
ok-drubel ok-voon anok plok sprok .	lalok brok anok plok nok .
at-drubel at-voon pippat rrat dat .	iat lat pippat rrat nnat .
erok sprok izok hihok ghirok .	wiwok nok izok kantok ok-yurp .
totat dat arrat vat hilat .	totat nnat quat oloat at-yurp .
ok-voon anok drok brok jok .	lalok mok nok yorok ghirok <b>clok</b> .
at-voon krat pippat sat lat .	wat nnat gat mat bat hilat .
wiwok farok izok stok .	lalok nok crrrok hihok yorok zanzanok .
totat jjat quat cat .	wat nnat arrat mat zanzanat .
lalok sprok izok jok stok .	lalok rarok nok izok hihok mok .
wat dat krat quat cat .	wat nnat forat arrat vat gat .

# farok crrrok hihok yorok clok kantok ok-yurp .

jjat

arrat

mat

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

# farok crrrok hihok yorok clok kantok ok-yurp .

jjat

arrat

mat

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

# farok crrrok hihok yorok clok kantok ok-yurp .

jjat

arrat

mat

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

# farok crrrok hihok yorok clok kantok ok-yurp .

jjat

arrat

mat

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

# farok crrrok hihok yorok clok kantok ok-yurp .

jjat

arrat

mat

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

# farok crrrok hihok yorok clok kantok ok-yurp .

jjat

arrat

mat

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

# farok crrrok hihok yorok clok kantok ok-yurp .

jjat

arrat

mat

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

# farok crrrok hihok yorok clok kantok ok-yurp .

jjat

arrat

mat

ok-voon ororok sprok .	lalok farok ororok lalok sprok izok enemok .
at-voon bichat dat .	wat jjat bichat wat dat vat eneat .
ok-drubel ok-voon anok plok sprok .	lalok brok anok plok <b>nok</b> .
at-drubel at-voon pippat rrat dat .	iat lat pippat rrat nnat .
erok sprok izok hihok ghirok .	wiwok <b>nok</b> izok kantok ok-yurp .
totat dat arrat vat hilat .	totat nnat quat oloat at-yurp .
ok-voon anok drok brok jok .	lalok mok <b>nok</b> yorok ghirok clok .
at-voon krat pippat sat lat .	wat nnat gat mat bat hilat .
wiwok farok izok stok .	lalok <b>nok</b> crrrok hihok yorok zanzanok .
totat jjat quat cat .	wat nnat arrat mat zanzanat .
lalok sprok izok jok stok .	lalok rarok <b>nok</b> izok hihok mok .
wat dat krat quat cat .	wat nnat forat arrat vat gat .

# farok crrrok hihok yorok clok kantok ok-yurp .

jjat

arrat

mat

ok-voon ororok sprok .	lalok farok ororok lalok sprok izok enemok .
at-voon bichat dat .	wat jjat bichat wat dat vat eneat .
ok-drubel ok-voon anok plok sprok .	lalok brok anok plok <b>nok</b> .
at-drubel at-voon pippat rrat dat .	iat lat pippat rrat nnat .
erok sprok izok hihok ghirok .	wiwok <b>nok</b> izok kantok ok-yurp .
totat dat arrat vat hilat .	totat nnat quat oloat at-yurp .
ok-voon anok drok brok jok .	lalok mok <b>nok</b> yorok ghirok clok .
at-voon krat pippat sat lat .	wat nnat gat mat bat hilat .
wiwok farok izok stok .	lalok <b>nok</b> crrrok hihok yorok zanzanok .
totat jjat quat cat .	wat nnat arrat mat zanzanat .
lalok sprok izok jok stok .	lalok rarok <b>nok</b> izok hihok mok .
wat dat krat quat cat .	wat nnat forat arrat vat gat .

# farok crrrok hihok yorok clok kantok ok-yurp .

jjat

arrat

mat

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

# farok crrrok hihok yorok clok kantok ok-yurp .

jjat

arrat

mat

ok-voon ororok sprok .	lalok farok ororok lalok sprok izok enemok .
at-voon bichat dat .	wat jjat bichat wat dat vat eneat .
ok-drubel ok-voon anok plok sprok .	lalok brok anok plok nok .
at-drubel at-voon pippat rrat dat .	iat lat pippat rrat nnat .
erok sprok izok hihok <b>ghirok</b> . totat dat arrat vat hilat .	wiwok nok izok kantok ok-yurp . totat nnat quat oloat at-yurp .
ok-voon anok drok brok jok . at-voon krat pippat sat lat .	lalok mok nok yorok <b>ghirok</b> clok . wat nnat gat mat bat hilat .
wiwok farok izok stok . totat jjat quat cat .	lalok nok crrrok hihok yorok zanzanok . wat nnat arrat mat zanzanat .
lalok sprok izok jok stok . wat dat krat quat cat .	lalok rarok nok izok hihok mok . wat nnat forat arrat vat gat .

# farok crrrok hihok yorok clok kantok ok-yurp .

jjat

arrat

mat

ok-voon ororok sprok .	lalok farok ororok lalok sprok izok enemok .
at-voon bichat dat .	wat jjat bichat wat dat vat eneat .
ok-drubel ok-voon anok plok sprok .	lalok brok anok plok nok .
at-drubel at-voon pippat rrat dat .	iat lat pippat rrat nnat .
erok sprok izok hihok <b>ghirok</b> .	wiwok nok izok kantok ok-yurp .
totat dat arrat vat hilat .	totat nnat quat oloat at-yurp .
ok-voon anok drok brok jok .	lalok mok nok yorok <b>ghirok</b> clok .
at-voon krat pippat sat lat .	wat nnat gat mat bat hilat .
wiwok farok izok stok .	lalok nok crrrok hihok yorok zanzanok .
totat jjat quat cat .	wat nnat arrat mat zanzanat .
lalok sprok izok jok stok .	lalok rarok nok izok hihok mok .
wat dat krat quat cat .	wat nnat forat arrat vat gat .

# farok crrrok hihok yorok clok kantok ok-yurp .

jjat

arrat

mat

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

# farok crrrok hihok yorok **clok** kantok ok-yurp .

jjat

arrat

mat

ok-voon ororok sprok .	lalok farok ororok lalok sprok izok enemok . \\ \\ \\\\ wat jjat bichat wat dat vat eneat .
at-voon bichat dat .	lalok brok anok plok nok . /\\ iat lat pippat rrat nnat .
ok-drubel ok-voon anok plok sprok .	wiwok nok izok kantok ok-yurp .   totat nnat quat oloat at-yurp .
at-drubel at-voon pippat rrat dat .	lalok mok nok yorok ghirok <b>clok</b> . \\ \\ \\ \\\\ wat nnat gat mat bat hilat .
erok sprok izok hihok ghirok . totat dat arrat vat hilat .	lalok nok crrrok hihok yorok zanzanok . \\ \\ \\ \\\\ wat nnat arrat mat zanzanat .
ok-voon anok drok brok jok .	lalok rarok nok izok hihok mok . \\ \\ \\ \\\\ wat nnat forat arrat vat gat .
at-voon krat pippat sat lat .	
wiwok farok izok stok . totat jjat quat cat .	
lalok sprok izok jok stok . \\ wat dat krat quat cat .	

# farok crrrok hihok yorok **clok** kantok ok-yurp .

jjat

arrat

mat

bat

ok-voon ororok sprok .	lalok farok ororok lalok sprok izok enemok . \\ \\ \\\\ wat jjat bichat wat dat vat eneat .
at-voon bichat dat .	lalok brok anok plok nok . /\\ iat lat pippat rrat nnat .
ok-drubel ok-voon anok plok sprok .	wiwok nok izok kantok ok-yurp .   totat nnat quat oloat at-yurp .
at-drubel at-voon pippat rrat dat .	lalok mok nok yorok ghirok <b>clok</b> . \\ \\ \\ \\ wat nnat gat mat bat hilat .
erok sprok izok hihok ghirok . totat dat arrat vat hilat .	lalok nok crrrok hihok yorok zanzanok . \\ \\ \\ \\ wat nnat arrat mat zanzanat .
ok-voon anok drok brok jok .	lalok rarok nok izok hihok mok . \\ \\ \\ \\ wat nnat forat arrat vat gat .
at-voon krat pippat sat lat .	
wiwok farok izok stok . totat jjat quat cat .	
lalok sprok izok jok stok . \\ wat dat krat quat cat .	

farok crrrok hihok yorok clok kantok ok-yurp .

jjat

arrat

mat

bat

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

# farok crrrok hihok yorok clok kantok ok-yurp .

jjat

arrat

mat

bat

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

farok crrrok hihok yorok clok kantok ok-yurp .

jjat

arrat

mat

bat

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

# farok crrrok hihok yorok clok kantok ok-yurp .

jjat

arrat

mat

bat

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

# farok **crrrok** hihok yorok clok kantok ok-yurp .

jjat

arrat

mat

bat

ok-voon ororok sprok .	lalok farok ororok lalok sprok izok enemok . \\ \\ \\\\ wat jjat bichat wat dat vat eneat .
at-voon bichat dat .	
ok-drubel ok-voon anok plok sprok .	lalok brok anok plok nok .
at-drubel at-voon pippat rrat dat .	iat lat pippat rrat nnat .
erok sprok izok hihok ghirok . totat dat arrat vat hilat .	wiwok nok izok kantok ok-yurp .   totat nnat quat oloat at-yurp .
ok-voon anok drok brok jok . at-voon krat pippat sat lat .	lalok mok nok yorok ghirok clok . \\ \\ \\\\ wat nnat gat mat bat hilat .
wiwok farok izok stok . totat jjat quat cat .	lalok nok <b>crrrok</b> hihok yorok zanzanok . \\ \\ \\ \\ wat nnat arrat mat zanzanat .
lalok sprok izok jok stok . \\ wat dat krat quat cat .	lalok rarok nok izok hihok mok . \\ \\ \\ \\ wat nnat forat arrat vat gat .

# farok crrrok hihok yorok clok kantok ok-yurp .

jjat

arrat

mat

bat

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

farok crrrok hihok yorok clok kantok ok-yurp .

jjat

arrat

mat

bat

ok-voon ororok sprok .

at-voon bichat dat .

ok-drubel ok-voon anok plok sprok .

at-drubel at-voon pippat rrat dat .

erok sprok izok hihok ghirok .

totat dat arrat vat hilat .

ok-voon anok drok brok jok .

at-voon krat pippat sat lat .

wiwok farok izok stok .

totat jjat quat cat .

lalok sprok izok jok stok .

wat dat krat quat cat .

lalok farok ororok lalok sprok izok enemok .

wat jiat bichat wat dat vat eneat .

lalok brok anok plok nok .

iat lat pippat rrat nnat .

wiwok nok izok kantok ok-yurp .

totat nnat quat oloat at-yurp .

lalok mok nok yorok ghirok clok .

wat nnat gat mat bat hilat .

wat nnat gat mat bat hilat .

lök nok crrrok hihok yorok zanzanok .

wat nnat arrat mat zanzanat .

lalok rarok nok izok hihok mok .

wat nnat forat arrat vat qat .

farok crrrok hihok yorok clok kantok ok-yurp .

jjat

arrat

mat

bat

oloat

at-yurp

ok-voon ororok sprok .

at-voon bichat dat .

ok-drubel ok-voon anok plok sprok .

at-drubel at-voon pippat rrat dat .

erok sprok izok hihok ghirok .

totat dat arrat vat hilat .

ok-voon anok drok brok jok .

at-voon krat pippat sat lat .

wiwok farok izok stok .

totat jjat quat cat .

lalok sprok izok jok stok .

wat dat krat quat cat .

lalok farok ororok lalok sprok izok enemok .

wat jjat bichat wat dat vat eneat .

lalok brok anok plok nok .

iat lat pippat rrat nnat .

wiwok nok izok kantok ok-yurp .

totat nnat quat oloat at-vurp .

lalok mok nok yorok ghirok clok .

wat nnat gat mat bat hilat .

nok crrrok hihok yorok zanz

wat nnat arrat mat zanzanat .

lalok rarok nok izok hihok mok .

wat nnat forat arrat vat qat .

# farok crrrok hihok yorok clok kantok ok-yurp .

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

# clients do not sell pharmaceuticals in Europe .

clientes	no venden	medicinas	en Europa
Garcia and associates .		the clients and the associates are enemies .	
Garcia y asociados .		los clientes y los asociados son enemigos .	
Carlos Garcia has three associates .		the company has three groups .	
Carlos Garcia tiene tres asociados .		la empresa tiene tres grupos .	
his associates are not strong .		its groups are in Europe .	
sus asociados no son fuertes .		sus grupos estan en Europa .	
Garcia has a company also .		the modern groups sell strong pharmaceuticals .	
Garcia tambien tiene una empresa .		los grupos modernos venden medicinas fuertes .	
its clients are angry .		the groups do not sell zenzanine .	
sus clientes estan enfadados .		los grupos no venden zanzanina .	
the associates are also angry .		the small groups are not modern .	
los asociados tambien estan enfadados .		los grupos pequenos no son modernos .	



Warren Weaver

**Warren Weaver,**  
**American scientist (1894-1978)**

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".

# Progress in MT



Warren  
Weaver's  
memo

1947

Founding of SYSTRAN.  
Development of Rule-based MT (RBMT)

1968

Seminal SMT  
paper from IBM

1993

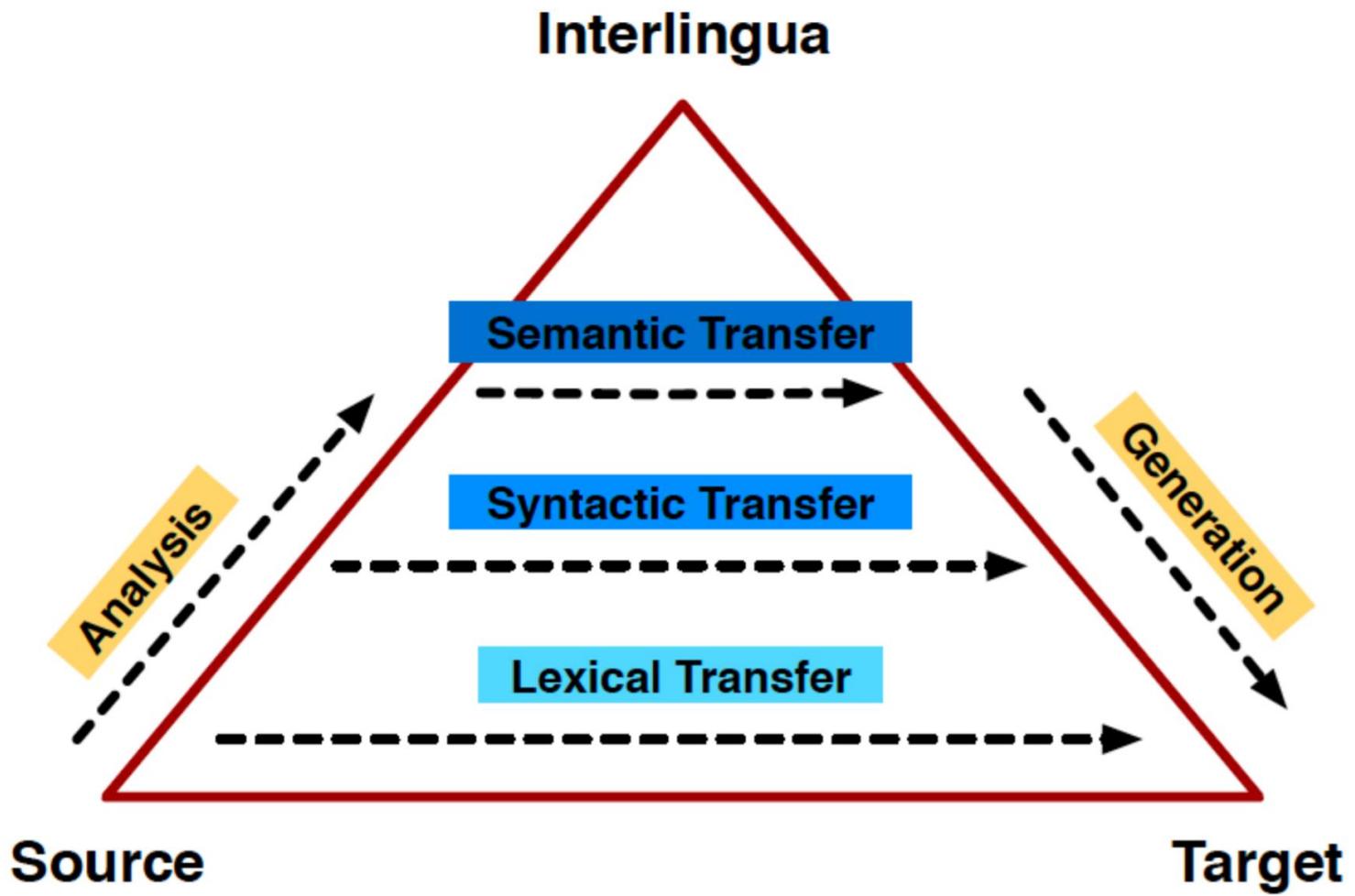
DARPA TIDES, GALE, BOLT programs  
Open-source of Moses toolkit  
Development of Statistical MT (SMT)

Early 2000s

2010s-Present

2011-2012: Early deep learning success in speech/vision  
2015: Seminal NMT paper (RNN+attention)  
2016: Google announces NMT in production  
2017: New NMT architecture: Transformer

# Vauquois Triangle



# RULE-BASED MACHINE TRANSLATION

---

- Build Dictionaries
- Write Transformation Rules

```
"have" :=  
  
if  
    subject/animate)  
    and object(owned-by-subject)  
then  
    translate to "kade... aahe"  
if  
    subject/animate)  
    and object(kinship-with-subject)  
then  
    translate to "laa... aahe"  
if  
    subject(inanimate)  
then  
    translate to "madhye... aahe"
```

# STATISTICAL MACHINE TRANSLATION

---

كيف حالك



How are you?

# STATISTICAL MACHINE TRANSLATION

---



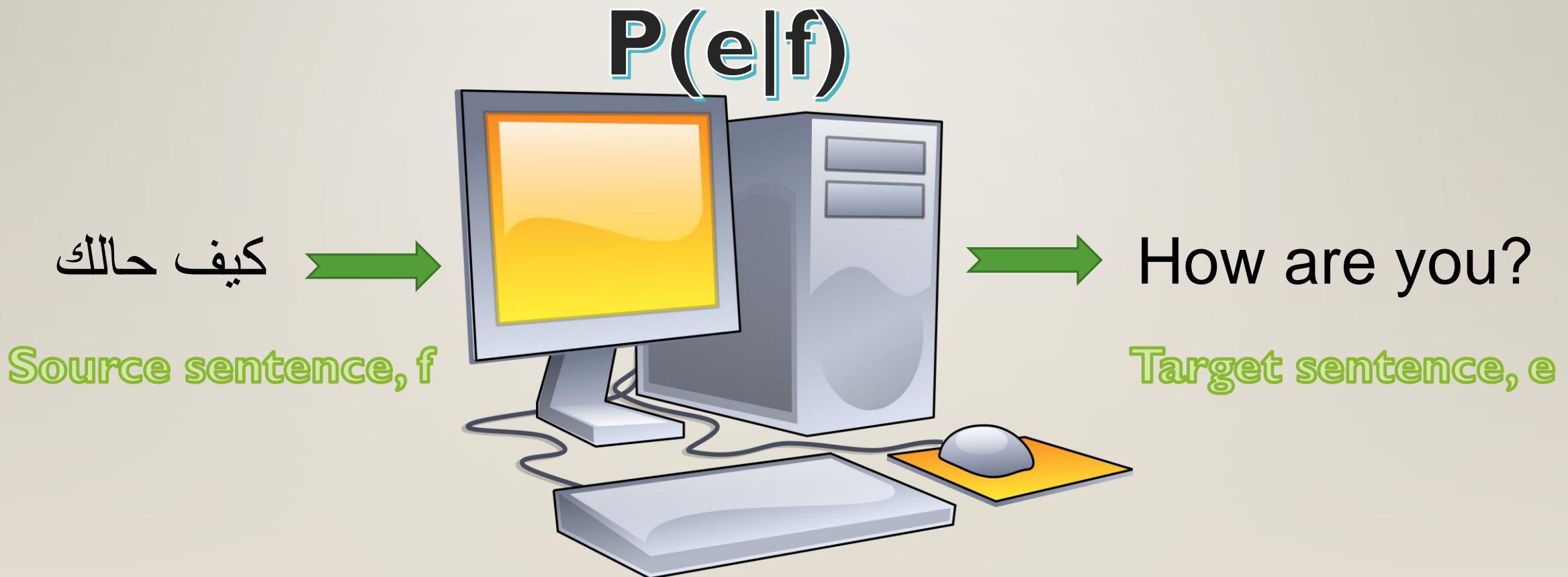
# STATISTICAL MACHINE TRANSLATION

---



# STATISTICAL MACHINE TRANSLATION

---



# STATISTICAL MACHINE TRANSLATION (SMT)

---

Data, Data, Data!

# STATISTICAL MACHINE TRANSLATION (SMT)

---

- Learn Dictionaries from Data

# STATISTICAL MACHINE TRANSLATION (SMT)

---

- Learn Dictionaries from Data “farok” → “jjat”

# STATISTICAL MACHINE TRANSLATION (SMT)

---

- Learn Dictionaries from Data “farok” → “jjat”
- Learn “Rules” from Data
- 1980 - 2015

# BITEXTS

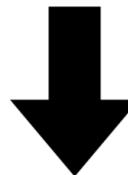
---



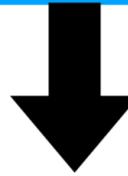
# MACHINE TRANSLATION (ABSTRACTION)

---

**There are 6000 languages in the world**



**Machine Translation (MT)  
System**



**世界には6000の言語があります**

# MACHINE TRANSLATION (SMT) ... SIMPLIFIED

---

**There are 6000 languages in the world**

あります 6000 言語 には 世界

TRANSLATION  
MODEL

LANGUAGE MODEL &  
REORDERING MODEL

世界には 6000 の 言語 が あります

# SMT VERSUS NMT

---

Statistical MT	Neural MT
Input: Source Sentence	

# SMT VERSUS NMT

---

Statistical MT	Neural MT
Input: Source Sentence	Input: Source Sentence

# SMT VERSUS NMT

---

Statistical MT	Neural MT
Input: Source Sentence	Input: Source Sentence
Output: Target Sentence	

# SMT VERSUS NMT

---

Statistical MT	Neural MT
Input: Source Sentence	Input: Source Sentence
Output: Target Sentence	Output: Target Sentence

# SMT VERSUS NMT

---

Statistical MT	Neural MT
Input: Source Sentence	Input: Source Sentence
Output: Target Sentence	Output: Target Sentence
Automatically Learn from Bitext	

# SMT VERSUS NMT

---

Statistical MT	Neural MT
Input: Source Sentence	Input: Source Sentence
Output: Target Sentence	Output: Target Sentence
Automatically Learn from Bitext	Automatically Learn from Bitext

# SMT VERSUS NMT

---

Statistical MT	Neural MT
Input: Source Sentence	Input: Source Sentence
Output: Target Sentence	Output: Target Sentence
Automatically Learn from Bitext	Automatically Learn from Bitext
Probabilistic Translation Model	

# SMT VERSUS NMT

---

Statistical MT	Neural MT
Input: Source Sentence	Input: Source Sentence
Output: Target Sentence	Output: Target Sentence
Automatically Learn from Bitext	Automatically Learn from Bitext
Probabilistic Translation Model	
Probabilistic Reordering Model	

# SMT VERSUS NMT

Statistical MT	Neural MT
Input: Source Sentence	Input: Source Sentence
Output: Target Sentence	Output: Target Sentence
Automatically Learn from Bitext	Automatically Learn from Bitext
Probabilistic Translation Model	
Probabilistic Reordering Model	
Probabilistic Language Model	

# SMT VERSUS NMT

Statistical MT	Neural MT
Input: Source Sentence	Input: Source Sentence
Output: Target Sentence	Output: Target Sentence
Automatically Learn from Bitext	Automatically Learn from Bitext
Probabilistic Translation Model	One Neural Model (Probabilistic)
Probabilistic Reordering Model	
Probabilistic Language Model	

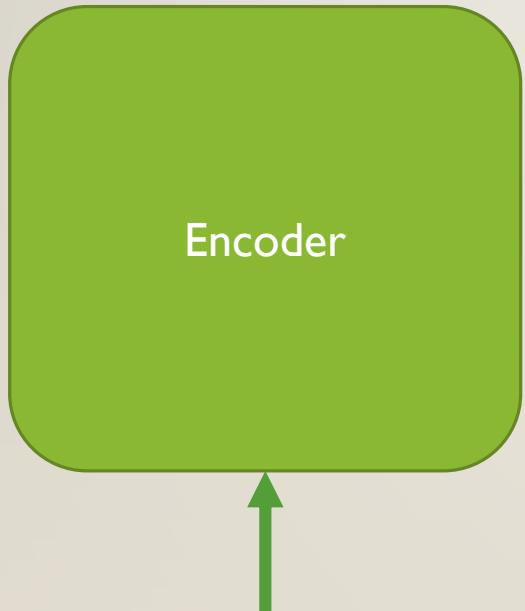
# NEURAL MACHINE TRANSLATION (NMT)

---

- Also a type of Statistical MT
- Represent words in high-dimensional, continuous, space
- $P(e|f)$

# NEURAL MACHINE TRANSLATION (NMT)

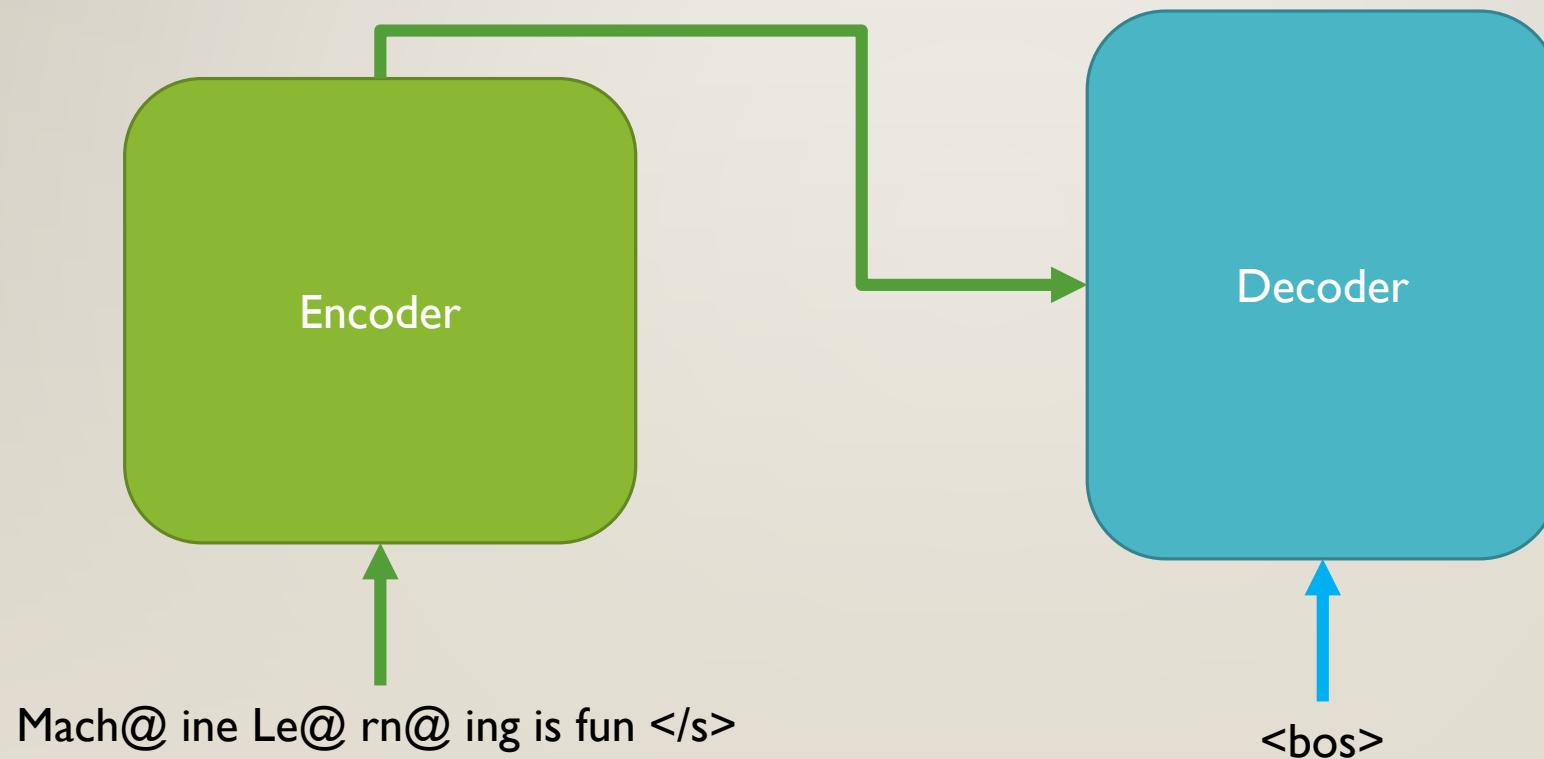
---



Mach@ ine Le@ rn@ ing is fun </s>

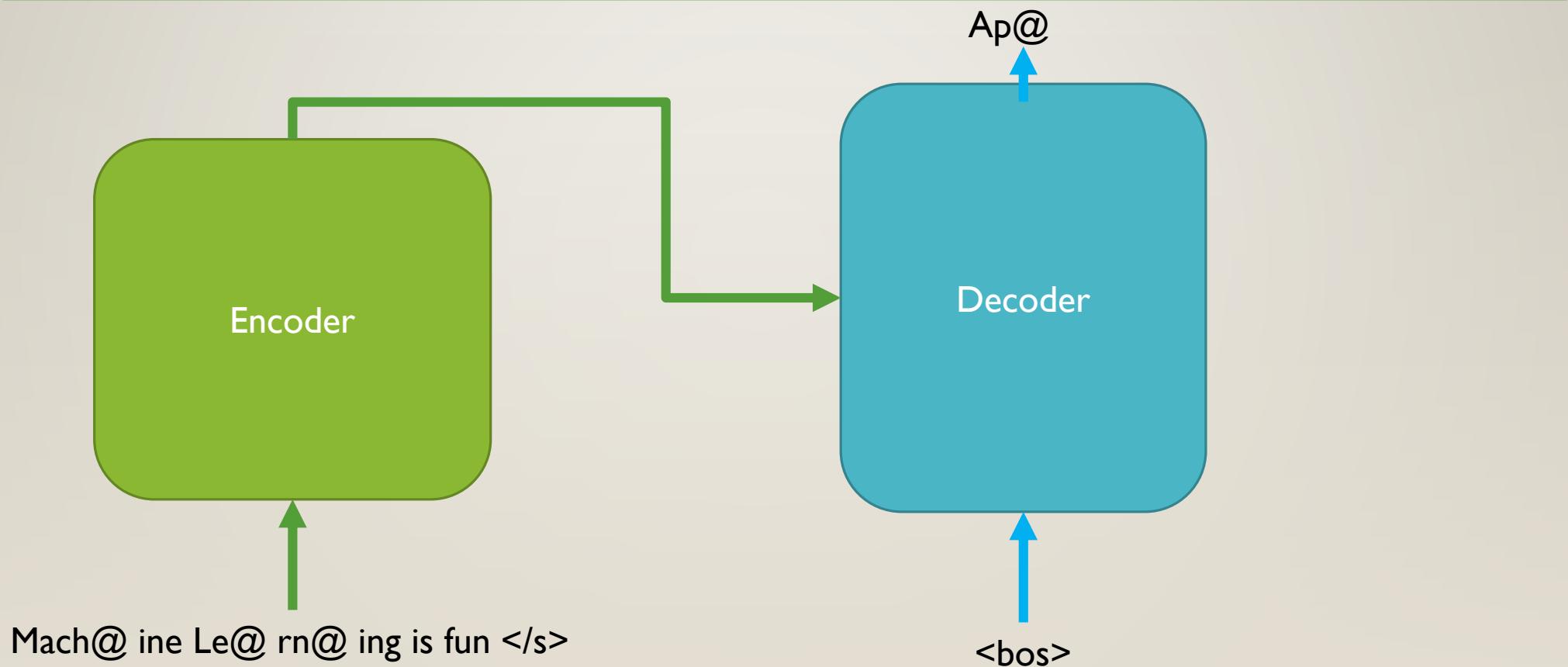
# NEURAL MACHINE TRANSLATION (NMT)

---



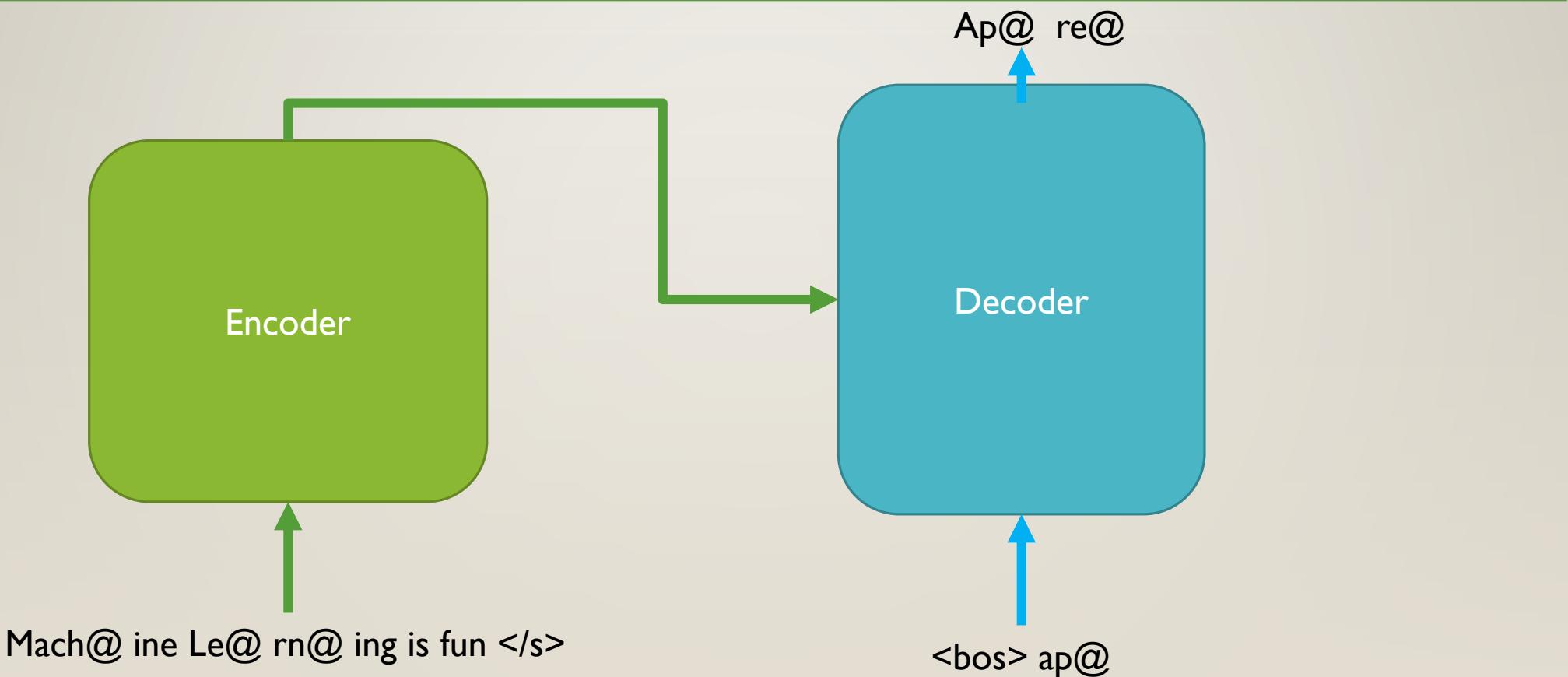
# NEURAL MACHINE TRANSLATION (NMT)

---



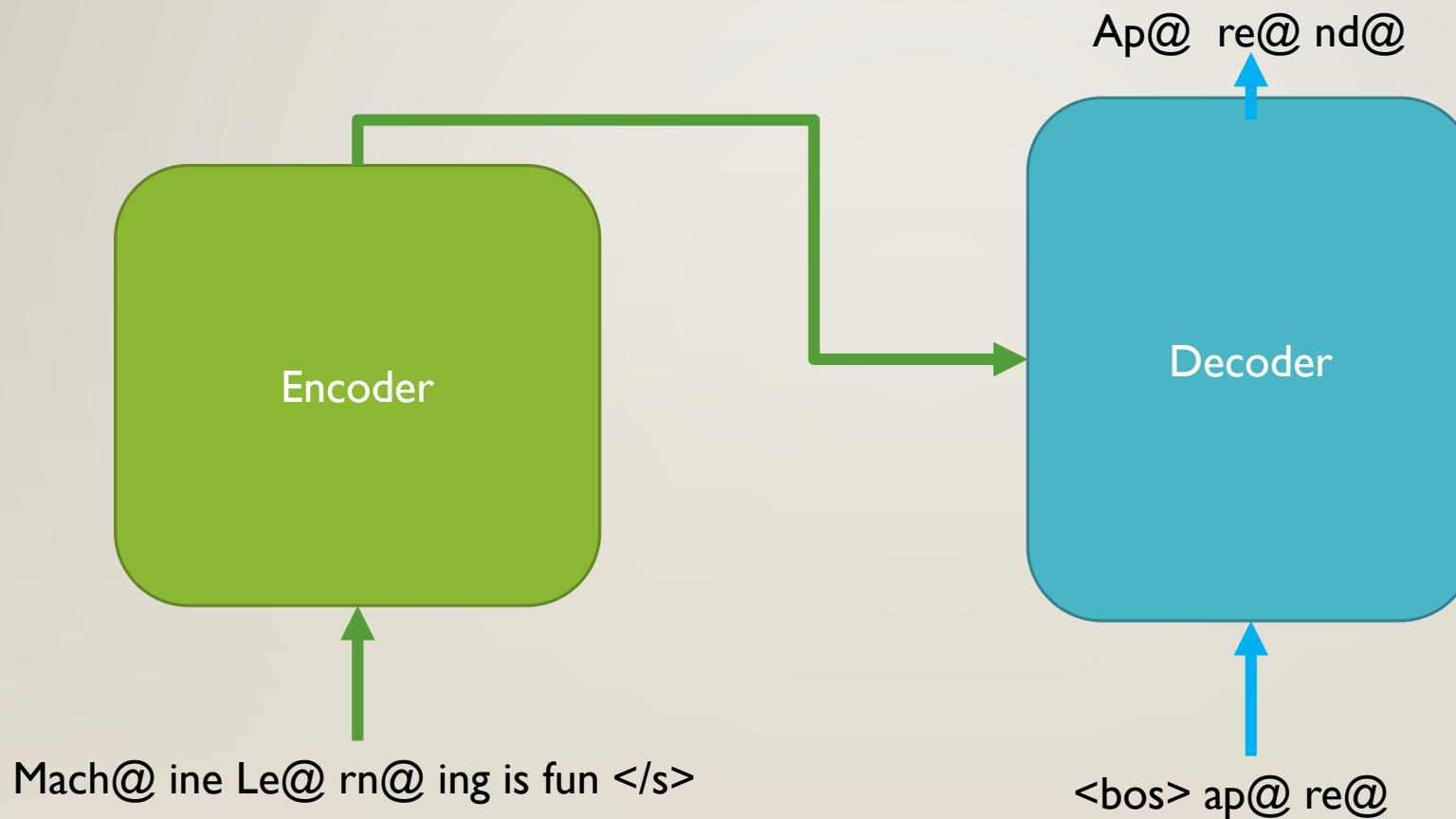
# NEURAL MACHINE TRANSLATION (NMT)

---



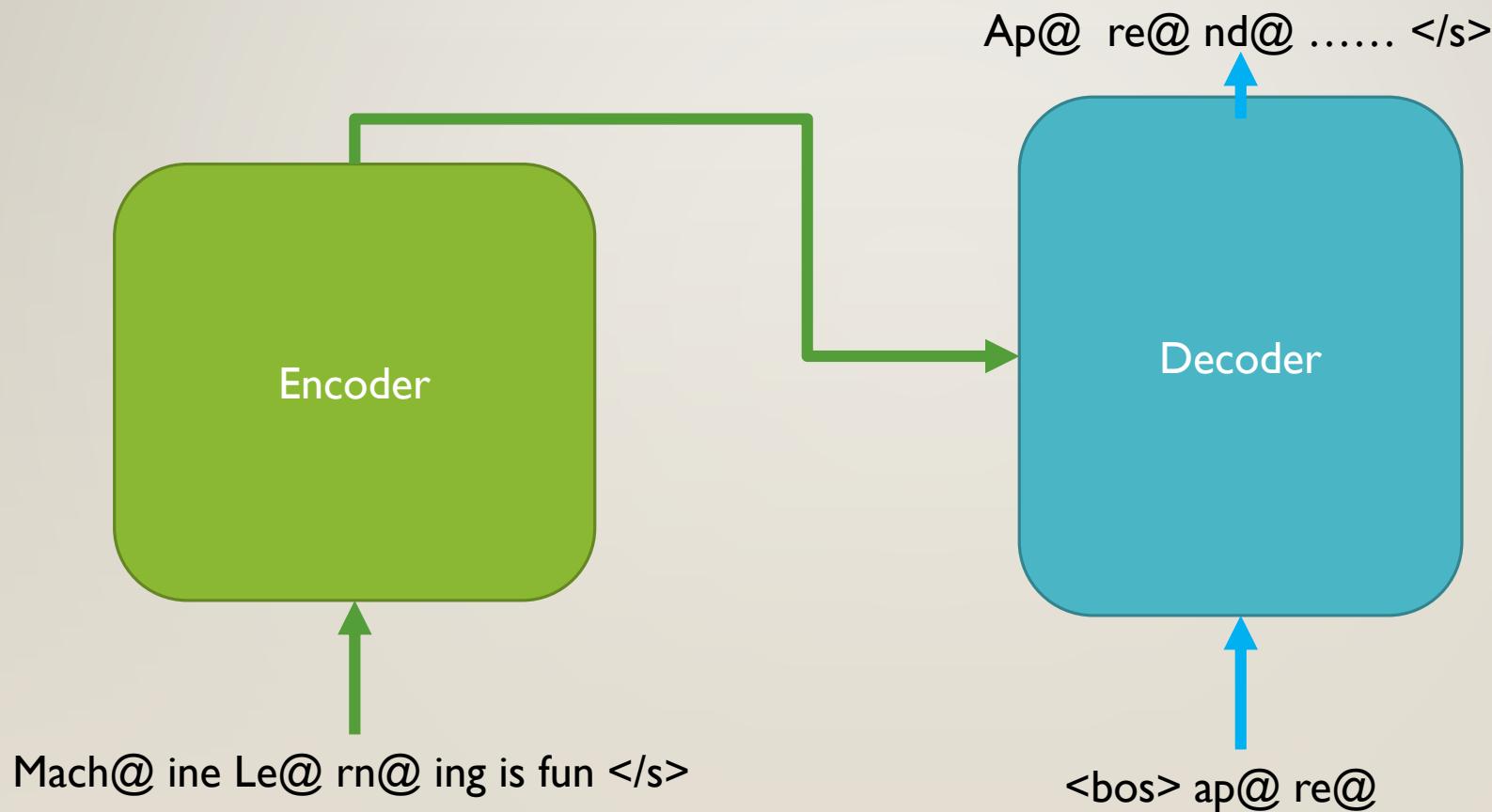
# NEURAL MACHINE TRANSLATION (NMT)

---

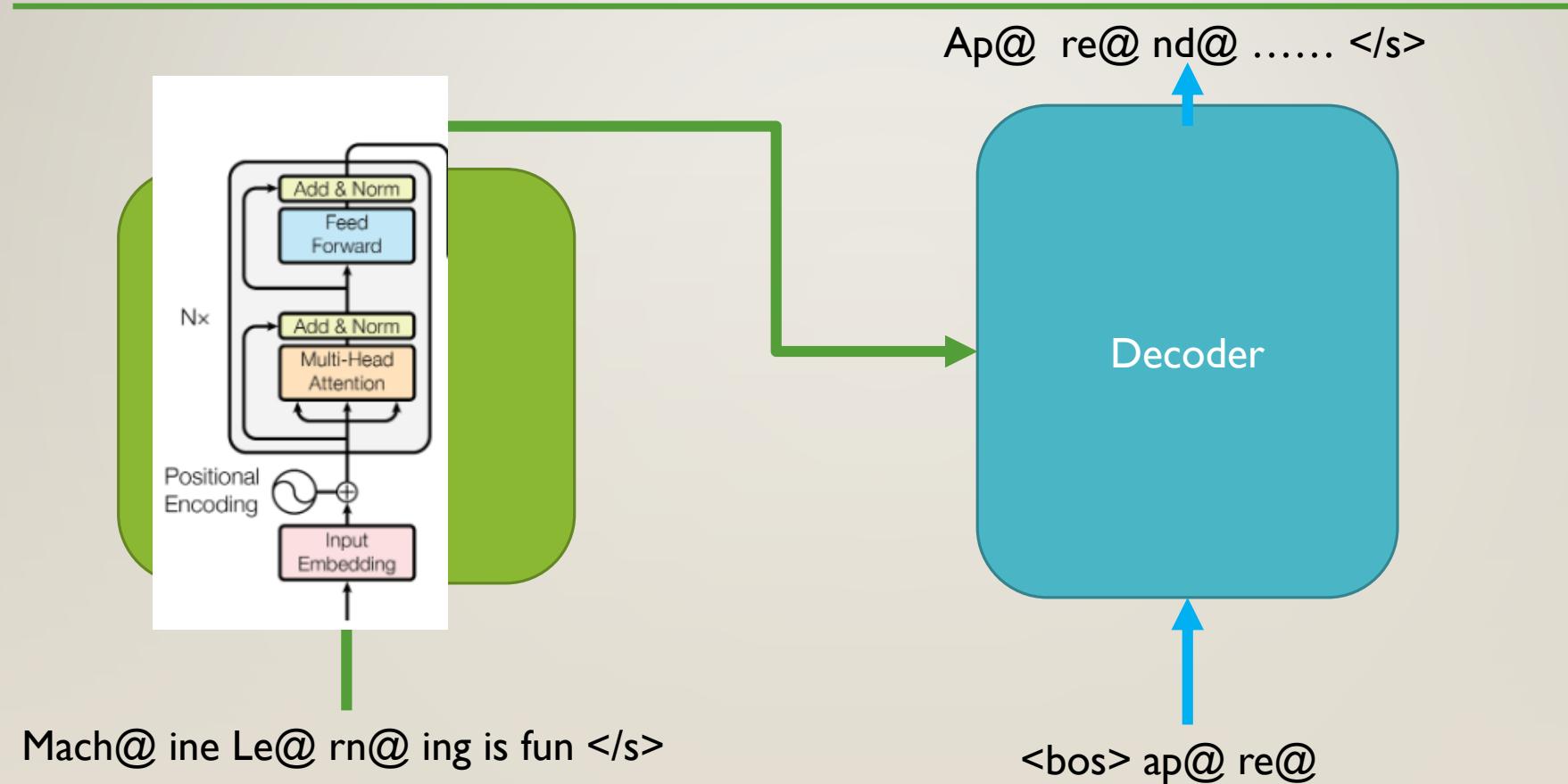


# NEURAL MACHINE TRANSLATION (NMT)

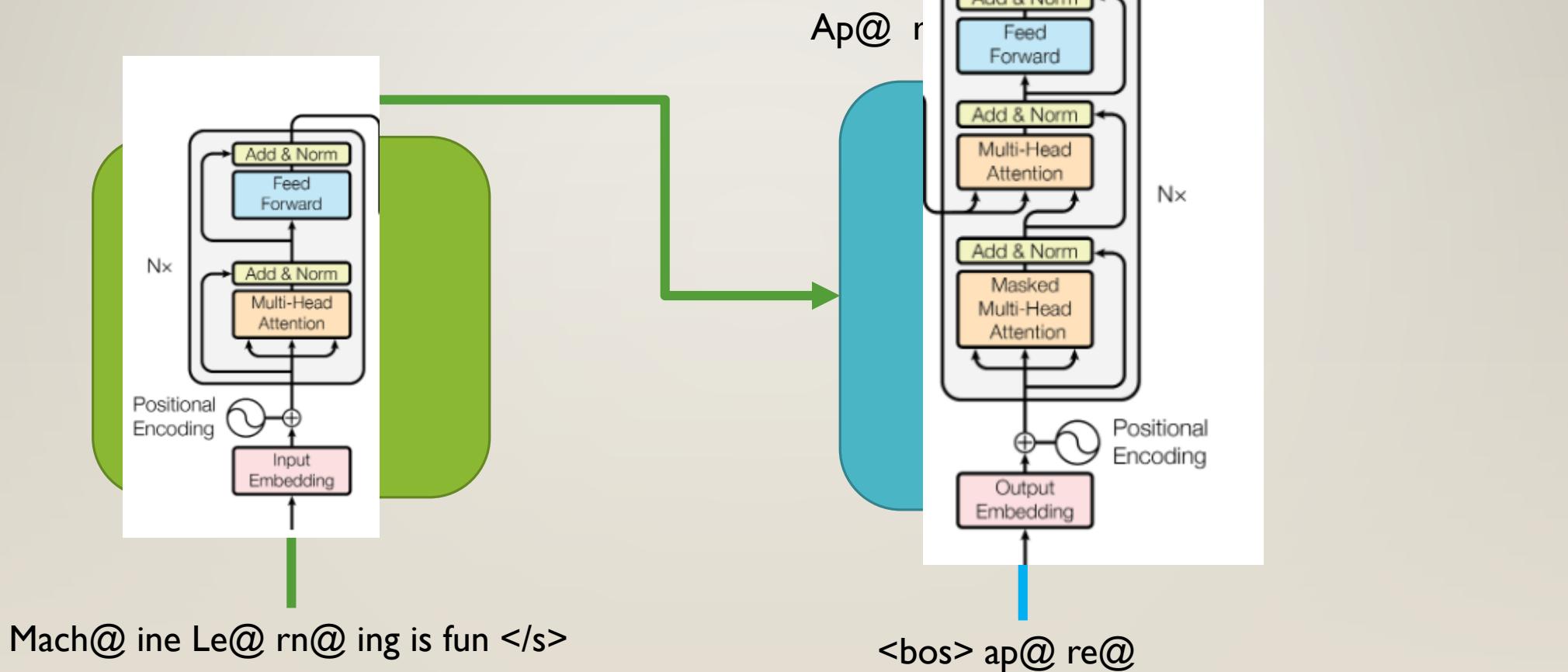
---



# NEURAL MACHINE TRANSLATION (NMT)



# NEURAL MACHINE TRANSLATION



# VOCABULARY

---

# ONE-HOT VECTOR

---

- Words correspond to index in vector
- Fixed size

# ONE-HOT VECTOR Dictionary

---

The	And	I	Dog	Johns	.....	Me	Cat
-----	-----	---	-----	-------	-------	----	-----

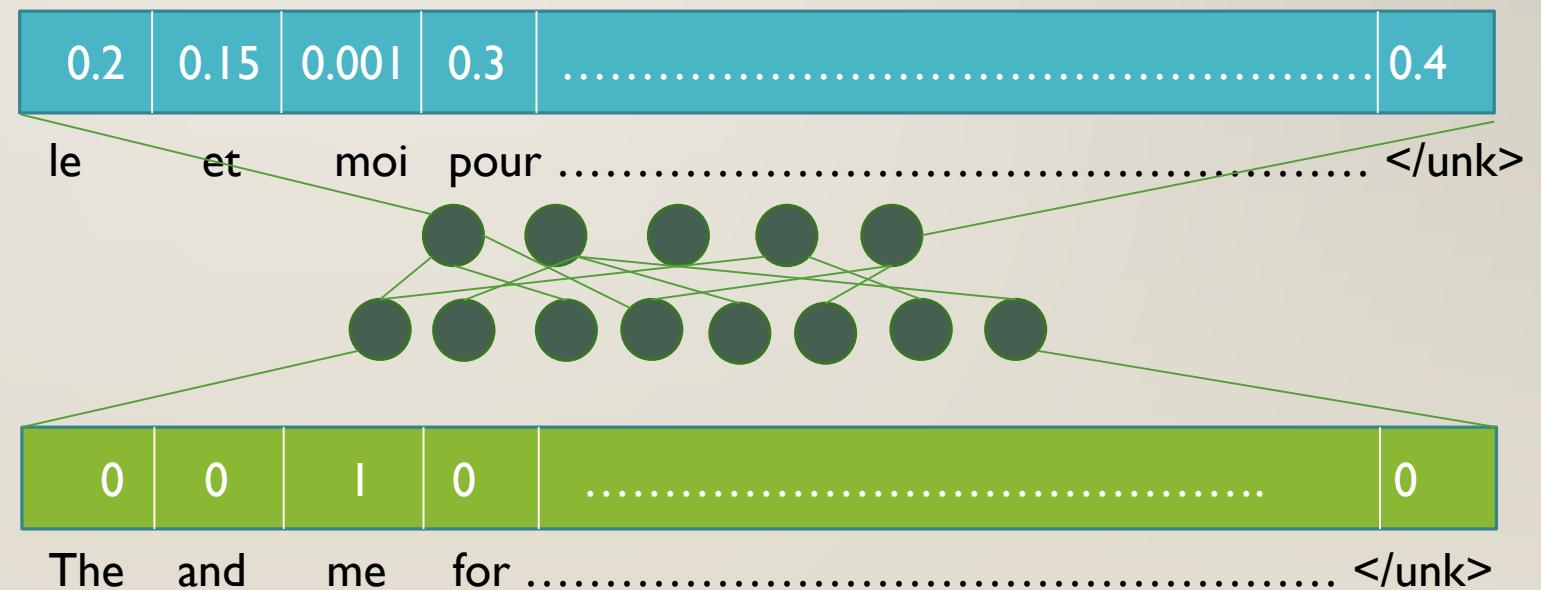
Johns Hopkins was founded in .....

0	0	0	0	I	.....	0	0
---	---	---	---	---	-------	---	---

# VOCABULARY SIZE

---

- Fixed Input & Output Vector Dimensions
- Out-of-vocabulary (OOVs)



# CHARACTER LEVEL

---

- No OOVs
- Very long sequences

# BYTE PAIR ENCODING

---

- Subword Unit
- Based on a compression algorithm
- Start small, repeatedly combine

peter piper picked a peck of pickled peppers

---

p@ e@ t@ e@ r p@ i@ p@ e@ r p@ i@ c@ k@ e@ d a p@ e@ c@ k o@ f p@ i@ c@ k@ l@ e@ d p@ e@ p@ p@ e@ r@ s

---

Vocabulary	Rules
p t i k a f s	e r c d o l

p@ e@ t@ e@ r p@ i@ p@ e@ r p@ i@ c@ k@ e@ d a p@ e@ c@ k o@ f p@ i@ c@ k@ l@ e@ d p@ e@ p@ p@ e@ r@ s

---

Vocabulary	Rules
p t i k a f s	e r c d o l pe@

pe@ t@ e@ r p@ i@ p@ e@ r p@ i@ c@ k@ e@ d a pe@ c@ k o@ f p@ i@ c@ k@ l@ e@ d pe@ p@ pe@ r@ s

---

Vocabulary	Rules
p t i k a f s pi@	e r c d o l pe@

pe@ t@ e@ r pi@ p@ e@ r pi@ c@ k@ e@ d a pe@ c@ k o@ f pi@ c@ k@ l@ e@ d pe@ p@ pe@ r@ s

---

Vocabulary	Rules
p t i k a f s pi@	e r c d o l pe@ pic@

pe@ t@ e@ r pi@ p@ e@ r pic@ k@ **e@ d** a pe@ c@ k o@ f pic@ k@ l@ **e@ d** pe@ p@ pe@ r@s

---

Vocabulary	Rules
p t i k a f s pi@	e r c d o l pe@ pic@

peter piper picked a peck of pickled peppers

---

Vocabulary		Rules
p	e	p@ e@ → pe@
t	r	p@ i@ → pi@
i	c	pi@ c@ → pic@
k	d	e@ d → ed
a	o	.
f	l	.
s	pe@	.
pi@	pic@	.
.....		.
of		o@ f → of

# HOW GOOD ARE OUR TRANSLATIONS?

---

- “Iyunivesithi yasekwa ngonyaka olandelayo ukusweleka kwakhe kwaye yanikezelwa kunyana wabo okuphela kwakhe.”
- The university was founded in the year following his death and was dedicated to their only son.
- The university was following his death and was dedicated to their only son.
- Machine Translation always works perfectly.

# MT EVALUATION

---

- Human Evaluation (Expensive, but best)
- Automatic (Cheap, can be correlated)

# BLEU SCORES

---

- Modified  $n$ -gram precision
- BiLingual Evaluation Understudy

# BLEU SCORES

---

- Gold: The Blue Jays are the mascot of Johns Hopkins University and can be seen around campus.
- Hyp: The Blue Jays are mascot of Johns Hopkins University University and can be seen around.

# BLEU SCORES

---

Gold:

The Blue Jays are the mascot of Johns Hopkins University and can be seen around campus

Hyp:

The Blue Jays are mascot of Johns Hopkins University University and can be seen around

Unigrams	Bigrams	Trigrams	4-grams
14/16			

# BLEU SCORES

---

Gold: The Blue Jays are the mascot of Johns Hopkins University and can be seen around campus

Hyp: The Blue Jays are mascot of Johns Hopkins University University and can be seen around

Unigrams	Bigrams	Trigrams	4-grams
14/16			

# BLEU SCORES

---

Gold:

The Blue Jays are the mascot of Johns Hopkins University and can be seen around campus

Hyp:

The Blue Jays are mascot of Johns Hopkins University University and can be seen around

Unigrams	Bigrams	Trigrams	4-grams
14/16			

# BLEU SCORES

---

Gold:

The Blue Jays are the mascot of Johns Hopkins University and can be seen around campus

Hyp:

The Blue Jays are mascot of Johns Hopkins University University and can be seen around

Unigrams	Bigrams	Trigrams	4-grams
14/16	12/15		

# BLEU SCORES

---

Gold:

The Blue Jays are the mascot of Johns Hopkins University and can be seen around campus

Hyp:

The Blue Jays are mascot of Johns Hopkins University University and can be seen around

Unigrams	Bigrams	Trigrams	4-grams
14/16	12/15	/14	

# BLEU SCORES

---

Gold:

The Blue Jays are the mascot of Johns Hopkins University and can be seen around campus

Hyp:

The Blue Jays are mascot of Johns Hopkins University University and can be seen around

Unigrams	Bigrams	Trigrams	4-grams
14/16	12/15	9/14	

# BLEU SCORES

---

Gold:

The Blue Jays are the **mascot of Johns Hopkins** University and can be seen around campus

Hyp:

The Blue Jays are **mascot of Johns Hopkins** University University and can be seen around

Unigrams	Bigrams	Trigrams	4-grams
14/16	12/15	9/14	/13

# BLEU SCORES

---

Gold:

The Blue Jays are the mascot of Johns Hopkins University and can be seen around campus

Hyp:

The Blue Jays are mascot of Johns Hopkins University University and can be seen around

Unigrams	Bigrams	Trigrams	4-grams
14/16	12/15	9/14	/13

# BLEU SCORES

---

Gold:

The Blue Jays are the mascot of Johns Hopkins University and can be seen around campus

Hyp:

The Blue Jays are mascot of Johns Hopkins University University and can be seen around

Unigrams	Bigrams	Trigrams	4-grams
14/16	12/15	9/14	/13

# BLEU SCORES

---

Gold:

The Blue Jays are the mascot of Johns Hopkins University and can be seen around campus

Hyp:

The Blue Jays are mascot of Johns Hopkins University University and can be seen around

Unigrams	Bigrams	Trigrams	4-grams
14/16	12/15	9/14	/13

# BLEU SCORES

---

Gold:

The Blue Jays are the mascot of Johns Hopkins University and can be seen around campus

Hyp:

The Blue Jays are mascot of Johns Hopkins University University and can be seen around

Unigrams	Bigrams	Trigrams	4-grams
14/16	12/15	9/14	4/13

# BLEU SCORES

---

$$BLEU = \min\left(1, e^{1 - \frac{\text{len}(gold)}{\text{len}(hyp)}}\right) \sqrt[4]{\prod_{i=1}^4 precision(i)}$$

Unigrams	Bigrams	Trigrams	4-grams
14/16	12/15	9/14	4/13

# BLEU SCORES

---

$$BLEU = \min \left( 1, e^{1 - \frac{\text{len}(gold)}{\text{len}(hyp)}} \right) \cdot \text{Brevity Penalty} \cdot \sqrt{\prod_{i=1}^4 precision(i)}$$

Unigrams	Bigrams	Trigrams	4-grams
14/16	12/15	9/14	3/13

# BLEU SCORES

---

$$BLEU = \min \left( 1, e^{1 - \frac{\text{len}(gold)}{\text{len}(hyp)}} \right) \sqrt[4]{\prod_{i=1}^4 precision(i)}$$

$$BLEU = \min \left( 1, e^{1 - \frac{16}{15}} \right) \sqrt[4]{0.138}$$

Unigrams	Bigrams	Trigrams	4-grams
14/16	12/15	9/14	3/13

# BLEU SCORES

---

$$BLEU = \min \left( 1, e^{1 - \frac{\text{len}(gold)}{\text{len}(hyp)}} \right) \sqrt[4]{\prod_{i=1}^4 precision(i)}$$

$$BLEU = 0.94 \sqrt[4]{0.138}$$

Unigrams	Bigrams	Trigrams	4-grams
14/16	12/15	9/14	3/13

# BLEU SCORES

---

$$BLEU = \min \left( 1, e^{1 - \frac{\text{len}(gold)}{\text{len}(hyp)}} \right) \sqrt[4]{\prod_{i=1}^4 precision(i)}$$

$$BLEU = 0.94 * 0.609$$

$$BLEU = 0.57$$

Unigrams	Bigrams	Trigrams	4-grams
14/16	12/15	9/14	3/13

# BLEU SCORES

---

- Calculate over entire test set (not one sentence)
- < 10 .... Pretty useless
- 10 – 20 ... can get some meaning
- 20 – 30 ... looks decent
- > 30 starts getting pretty good

# BLEU SCORES

---

- The large house
- A big mansion

# FAIRSEQ

---

- PyTorch
- <https://github.com/pytorch/fairseq>

