

# När kan man lita på automatisk översättning?

Aarne Ranta

Institutionen för data-och informationsteknik  
Chalmers tekniska högskola och Göteborgs universitet

samt

Digital Grammars AB

Vetenskapsfestivalen, Göteborg, 13 maj 2017



REMU

digital  grammars  
Language technology to rely on.

*Five, perhaps three years hence, interlingual meaning conversion by electronic process in important functional areas of several languages may well be an accomplished fact.*

*IBM press release*

*Five, perhaps three years hence, interlingual meaning conversion by electronic process in important functional areas of several languages may well be an accomplished fact.*

*IBM press release 1954*

[http://www-03.ibm.com/ibm/history/exhibits/701/701\\_translator.html](http://www-03.ibm.com/ibm/history/exhibits/701/701_translator.html)

# The Great A.I. Awakening

How Google used artificial intelligence to transform Google Translate, one of its more popular services — and how machine learning is poised to reinvent computing itself.

BY GIDEON LEWIS-KRAUS DEC. 14, 2016

# Google's Neural Machine Translation System: Bridging the Gap between Human and Machine Translation

Yonghui Wu, Mike Schuster, Zhifeng Chen, Quoc V. Le, Mohammad Norouzi  
`yonghui,schuster,zhifengc,qvl,mnorouzi@google.com`

Wolfgang Macherey, Maxim Krikun, Yuan Cao, Qin Gao, Klaus Macherey, Jeff Klingner, Apurva Shah, Melvin Johnson, Xiaobing Liu, Łukasz Kaiser, Stephan Gouws, Yoshikiyo Kato, Taku Kudo, Hideto Kazawa, Keith Stevens, George Kurian, Nishant Patil, Wei Wang, Cliff Young, Jason Smith, Jason Riesa, Alex Rudnick, Oriol Vinyals, Greg Corrado, Macduff Hughes, Jeffrey Dean

8 October 2016

# Machine translation bug of the week

May 8, 2017 @ 10:02 am · Filed by [Mark Liberman](#) under [Computational linguistics](#)

[« previous post](#) | [next post »](#)



**Emmanuel Macron**

@EmmanuelMacron

Follow

Mes chers compatriotes, vous avez choisi de m'accorder votre confiance et je tiens à vous exprimer ma profonde gratitude.

Translated from French by bing

[Wrong translation?](#)

My fellow Americans, you chose to give me your trust, and I would like to express my deep gratitude.

# Detta föredrag

## Tre metoder

- hur de funkar
- styrkor
- svagheter

## Tips och rekommendationer

# Metod 1: SMT = Statistical Machine Translation

Samla statistik från parallella texter

Översätt genom att klistra ihop textbitar

- Google translate traditionellt



*This big car is yellow.*

*Denna stora bil är gul.*

*This house is clean.*

*Detta hus är rent.*

*This big car is yellow.*

*Denna stora bil är gul.*

*This house is clean.*

*Detta hus är rent.*

*This big house is yellow.*

*This big car is yellow.*

*Denna stora bil är gul.*

*This house is clean.*

*Detta hus är rent.*

*This big house is yellow.*

*Denna stora*

*This big car is yellow.*

*Denna stora bil är gul.*

*This house is clean.*

*Detta hus är rent.*

*This big house is yellow.*

*Denna stora hus är*

*This big car is yellow.*

*Denna stora bil är gul.*

*This house is clean.*

*Detta hus är rent.*

*This big house is yellow.*

*Denna stora hus är gul.*

Swedish English German Detect language ▾



English Swedish German ▾

Translate

This big house is yellow. ×

Denna stora hus är gul.

English Swedish Finnish Detect language ▾

This big house is yellow.



English Chinese (Simplified) Swedish ▾

Detta stora hus är gult.

# Metod 2: NMT = Neural Machine Translation

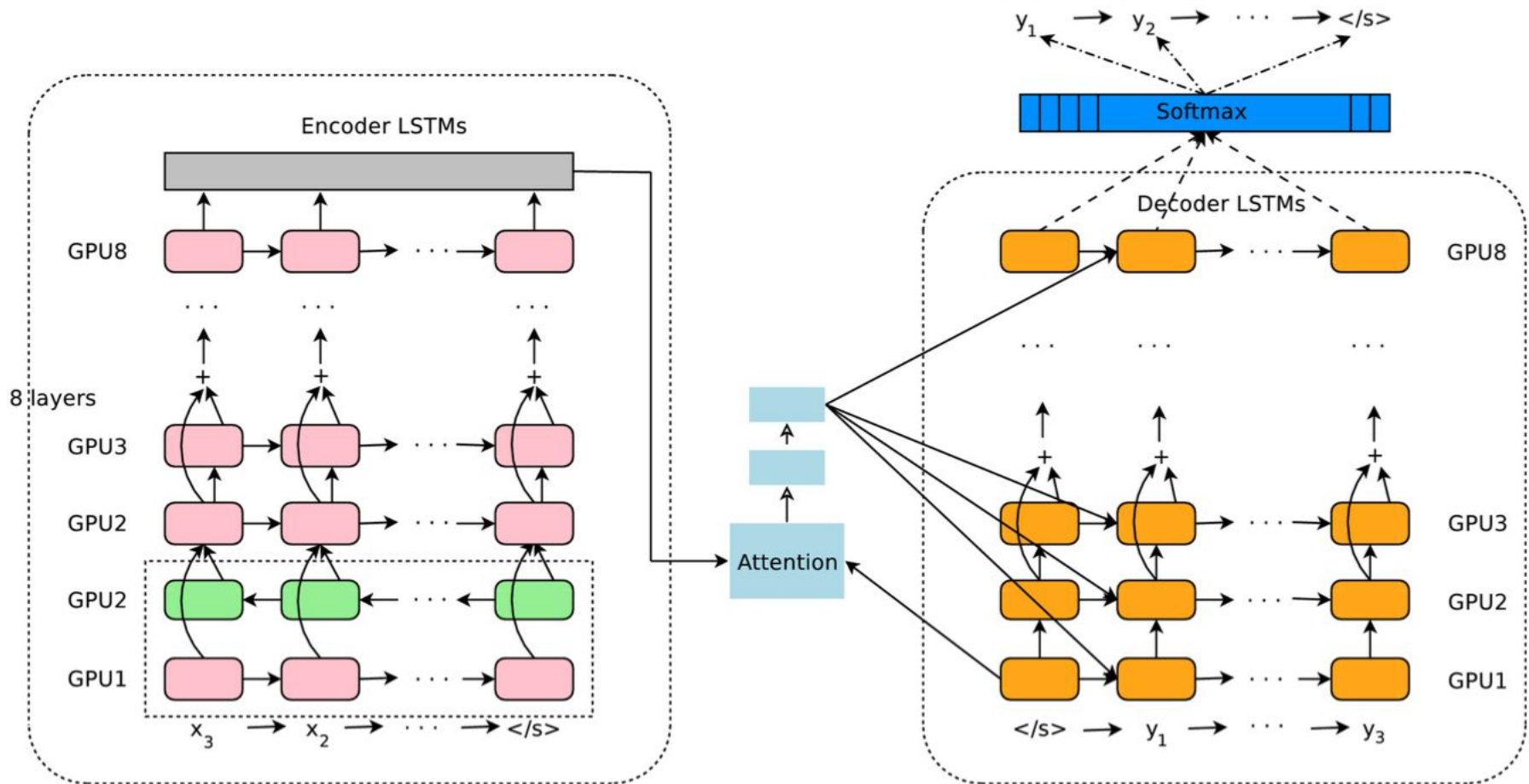
Neurala nätverk: en modell av hjärnan

Lär sig översätta från parallell text

Bättre än SMT på lång distans

- Google translate 2016-





# Gemensamt problem: glesa data

Svenska:

- $\sim 10^6$  ord (med alla böjningar)
- $\sim 10^{12}$  kombinationer av 2 ord
- $\sim 10^{18}$  kombinationer av 2 ord

Finska:

- $\sim 10^8$  ord (med alla böjningar)
- $\sim 10^{16}$  kombinationer av 2 ord
- $\sim 10^{24}$  kombinationer av 2 ord

# Ordformer i SMT

Swedish English French Detect language ▾

↔ English Swedish French ▾ **Translate**

super, sups, söp, söps, sup, sups, supa, supas, supit, supits, supen, supens, supet, supets, supna, supnas, supna, supnas, supna, supnas, supande, supandes, supandet, supandets, supanden, supandens, supandena, supandenas

×

super, sups, drank, SOPs, sup, sups, booze, supas, drunk, supits, the drink, supens, supet, supets, supna, supnas, supna, supnas, supna, supnas, boozing, supandes, binge drinking, drinking of a country, supanden, the booze, booze creatures, supandenas

# Ordformer i NMT

English Swedish French Detect language ▾



Swedish English Italian ▾

Translate

super, sups, söp, söps, sup, sups, supa, supas, supit, supits, supen, supens, supet, supets, supna, supnas, supna, supnas, supna, supnas, supande, supandes, supandet, supandets, supanden, supandens, supandena, supanden  
×

Super sups sup sups sups supa supas supit supits supen supens supet supers supnas supna supnas supna supnas supande supandes supandes supandets supanden supand supas supas supand supands supps supas supit supits supers supens suppet supas supas supas supanden supanden Supandens, supandena, supanden

English Swedish French Detect language ▼



Italian English Swedish ▼

Translate

Det sups mycket på festivalen.



It sups a lot at the festival.

English Swedish French Detect language ▼



Italian English Swedish ▼

Translate

Det sups mycket på festivalen.  
Det sups mycket på Vetenskapsfestivalen.



It sups a lot at the festival.  
It's a lot of fun at the Science Festival.

English Swedish French Detect language ▾



Italian English Swedish ▾

Translate

Det sups mycket på festivalen.  
Det sups mycket på Vetenskapsfestivalen.  
Det sups för mycket på Vetenskapsfestivalen.



It sups a lot at the festival.  
It's a lot of fun at the Science Festival.  
It sups too much at the Science Festival.

English Swedish French Detect language ▾



Italian English Swedish ▾

Translate

Det sups mycket på festivalen.  
Det sups mycket på Vetenskapsfestivalen.  
Det sups för mycket på Vetenskapsfestivalen.  
Det sups inte mycket på Vetenskapsfestivalen.



It sups a lot at the festival.  
It's a lot of fun at the Science Festival.  
It sups too much at the Science Festival.  
It does not suge much at the Science Festival.



English Swedish French Detect language ▾



Italian English Swedish ▾

Translate

Det sups mycket på festivalen.  
Det sups mycket på Vetenskapsfestivalen.  
Det sups för mycket på Vetenskapsfestivalen.  
Det sups inte mycket på Vetenskapsfestivalen.  
Vi super inte.



It sups a lot at the festival.  
It's a lot of fun at the Science Festival.  
It sups too much at the Science Festival.  
It does not suge much at the Science Festival.  
We're not great.

English Swedish French Detect language ▾



Italian English Swedish ▾

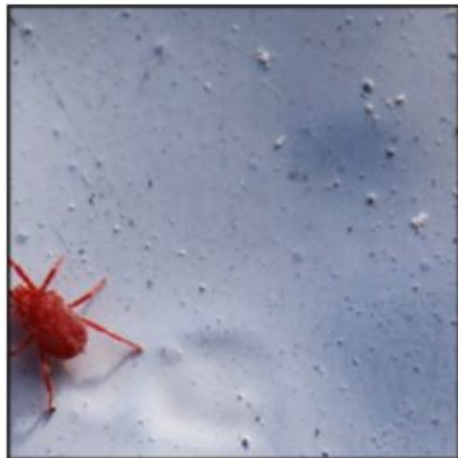
Translate

Det sups mycket på festivalen.  
Det sups mycket på Vetenskapsfestivalen.  
Det sups för mycket på Vetenskapsfestivalen.  
Det sups inte mycket på Vetenskapsfestivalen.  
Vi super inte.  
Vi super.



It sups a lot at the festival.  
It's a lot of fun at the Science Festival.  
It sups too much at the Science Festival.  
It does not suge much at the Science Festival.  
We're not great.  
We are super.

# Neurala nätverk i bildigenkänning



**mite**



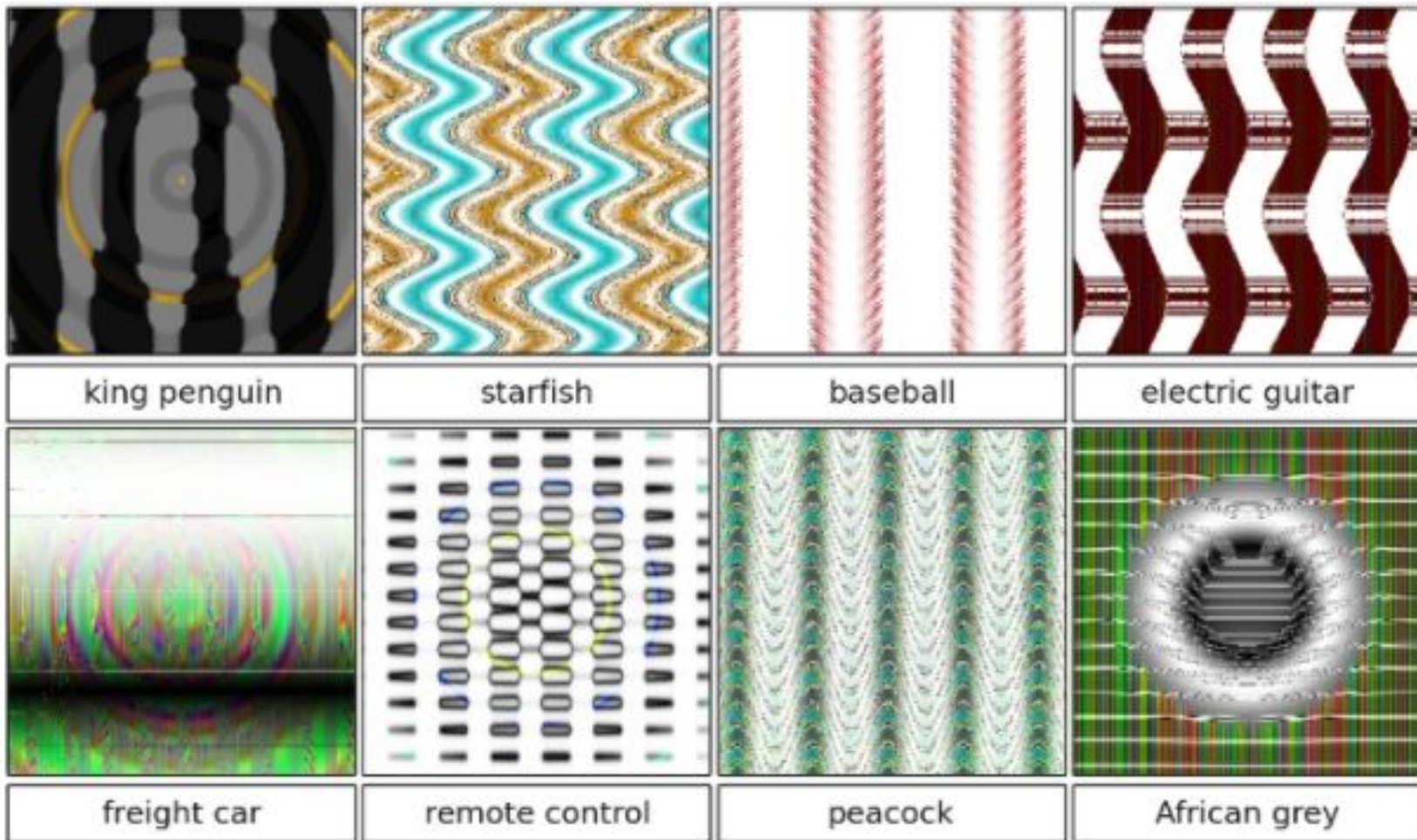
**container ship**



**motor scooter**



**leopard**





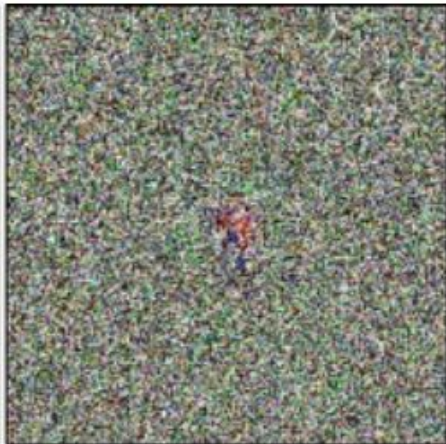
robin



cheetah



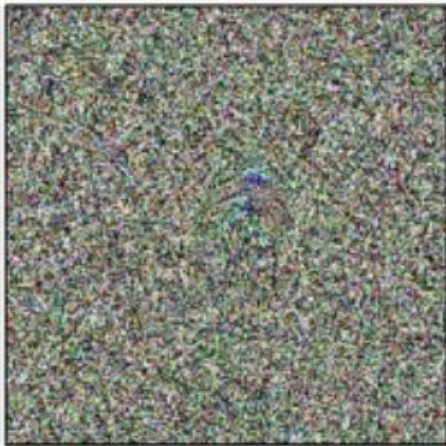
armadillo



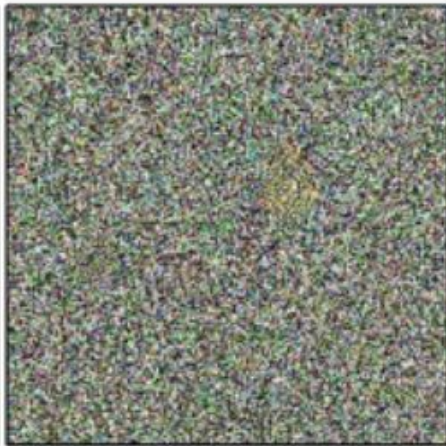
lesser panda



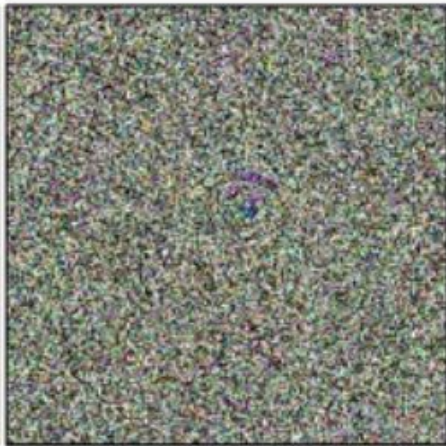
centipede



peacock



jackfruit



bubble

English Swedish Finnish Detect language ▾



English Chinese (Simplified) Swedish ▾

Translate

ä



IA

English Swedish Finnish Detect language ▾



English Chinese (Simplified) Swedish ▾

Translate

ä  
ä ä



IA  
I am

English Swedish Finnish Detect language ▾



English Chinese (Simplified) Swedish ▾

Translate

ia  
ia ia  
ia ia ia



IA  
I am  
No



English Swedish Finnish Detect language ▾



English Chinese (Simplified) Swedish ▾

Translate

ä  
ä ä  
ä ä ä  
ä ä ä ä



IA  
I am  
No  
I do not sleep

English Swedish Finnish Detect language ▾



English Chinese (Simplified) Swedish ▾

Translate

ia  
ia ia  
ia ia ia  
ia ia ia ia  
ia ia ia ia ia



IA  
I am  
No  
I do not sleep  
I do not know

English Swedish Finnish Detect language ▾



English Chinese (Simplified) Swedish ▾

Translate

ia  
ia ia  
ia ia ia  
ia ia ia ia  
ia ia ia ia ia  
ia ia ia ia ia ia



IA  
I am  
No  
I do not sleep  
I do not know  
I do not know

English Swedish Finnish Detect language ▾



English Chinese (Simplified) Swedish ▾

Translate

ia  
ia ia  
ia ia ia  
ia ia ia ia  
ia ia ia ia ia  
ia ia ia ia ia ia  
ia ia ia ia ia ia ia



IA  
I am  
No  
I do not sleep  
I do not know  
I do not know  
Already a son, do not enter the age of your child

**Tips 1:**  
**Lita inte på MT**

English Swedish French Detect language ▾



Translate

Min mor är inte svensk.  
Min mor är svensk.



我的母亲是瑞典的。  
我的母亲是瑞典的。

English Swedish French Detect language ▾



Translate

Min mor är inte svensk.  
Min mor är svensk.



我的母亲是瑞典的。  
我的母亲是瑞典的。

English Swedish French Detect language ▾



German English Norwegian ▾

Min mor är inte svensk.  
Min mor är svensk.



Min mor er svensk.  
Min mor er svensk.

**Tips 2:**

**Prova först med språk som du kan**



English Swedish French Detect language ▾



Swedish English Norwegian ▾

Är du lycklig?  
Är ni lyckliga?  
Är ni lycklig?



Er du lykkelig?  
Er du lykkelig?  
Er du lykkelig?

English Swedish French Detect language ▾



Swedish English Norwegian ▾

Är du lycklig?  
Är ni lyckliga?  
Är ni lycklig?



Er du lykkelig?  
Er du lykkelig?  
Er du lykkelig?

Finnish English Swedish Detect language ▾



English German Norwegian ▾

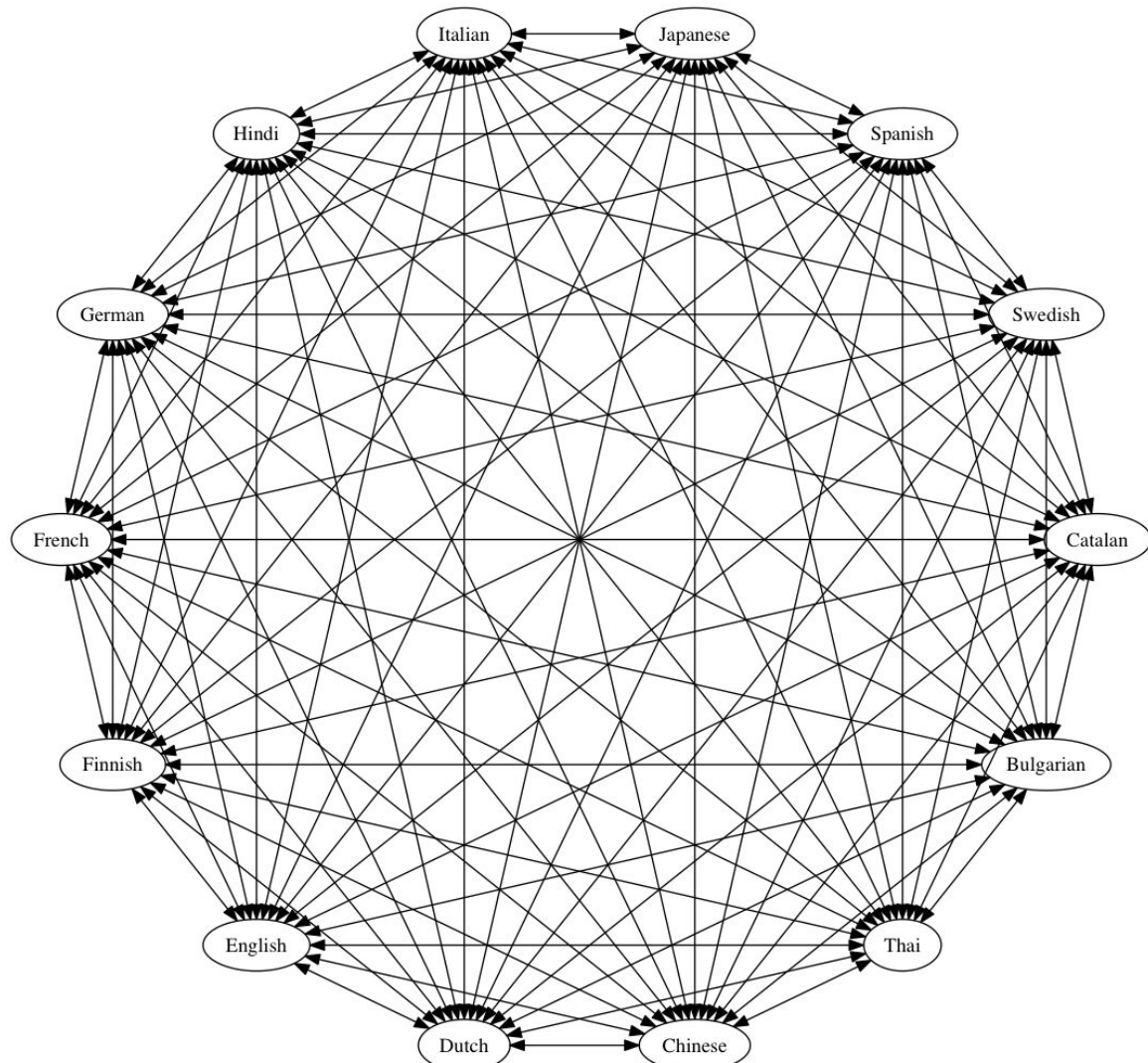
Jag har en fladdermus.

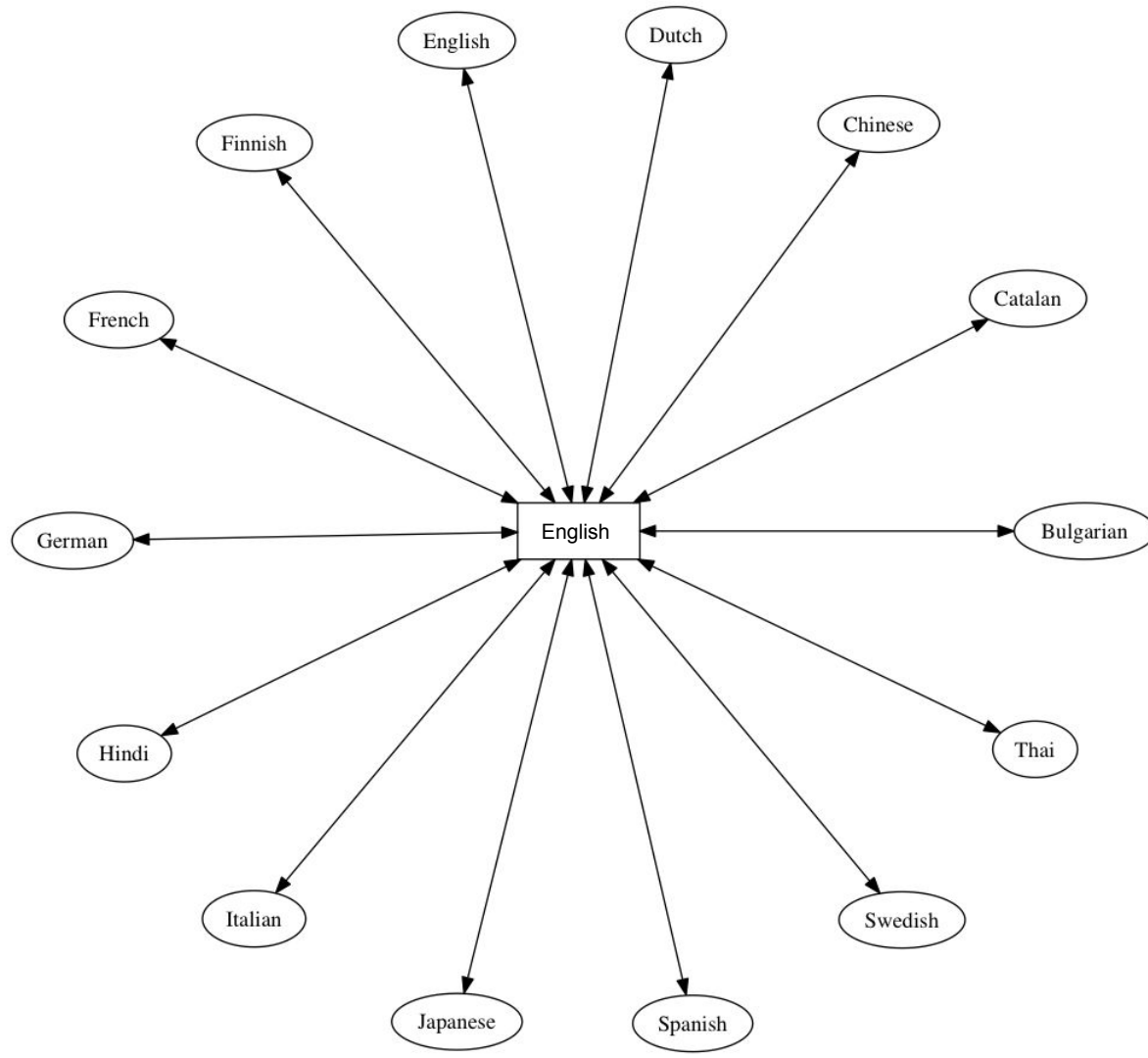


Jeg har et balltre.

**Tips 3:**

**Översätt till eller från engelska**





# Hur bra är MT? BLEU-skalan

Likheten mellan maskinöversättning och facit

original	MT	facit	BLEU
<i>Min mor är svensk.</i>	<i>Min mor er svensk.</i>	<i>Min mor er svensk.</i>	1.0
<i>Min mor är inte svensk.</i>	<i>Min mor er Svensk</i>	<i>Min mor er ikke svensk.</i>	0.41

# SMT vs. NMT

BLEU (max 1.0)

SMT	NMT
0.37	0.41

Fluency (max 6.0)

SMT	NMT	human
3.87	4.44	4.82

# Metoder

SMT



# Metoder

SMT

NMT

# Metoder

SMT

NMT

GSM

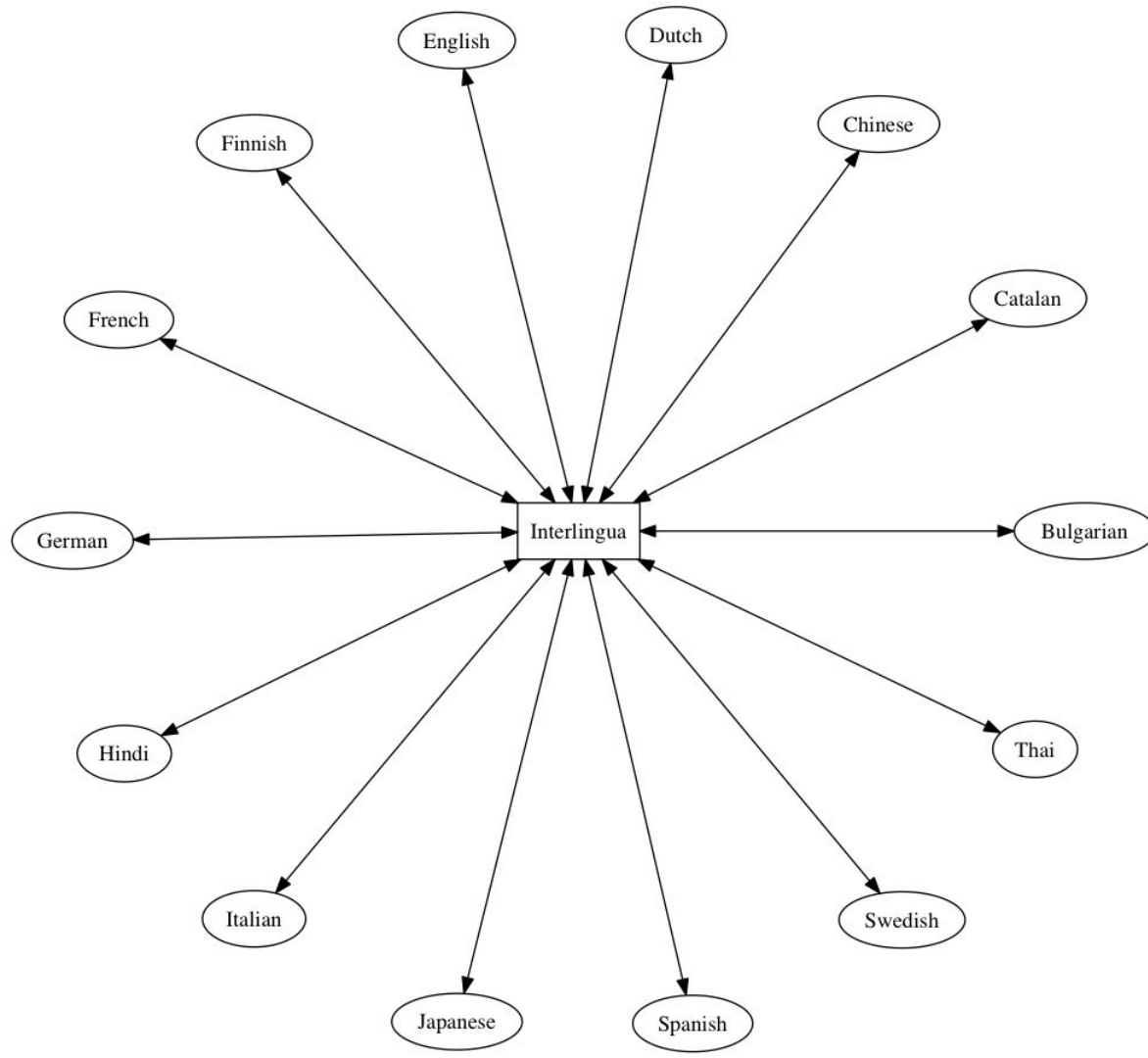
# Metod 3:

SMT

NMT

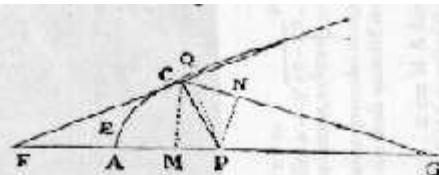
GSM = Grammar-based Semantic Methods

- icke-officiell förkortning





Facen  
graciale  
pour  
trouver  
des lignes  
droites,  
qui coup-  
pent les  
courbes  
données,  
ou leurs  
contin-  
gentes, a  
angles  
droits.



Soit C E  
la ligne courbe,  
& qu'il faille ti-  
rer vne ligne  
droite par le  
point C, qui fa-

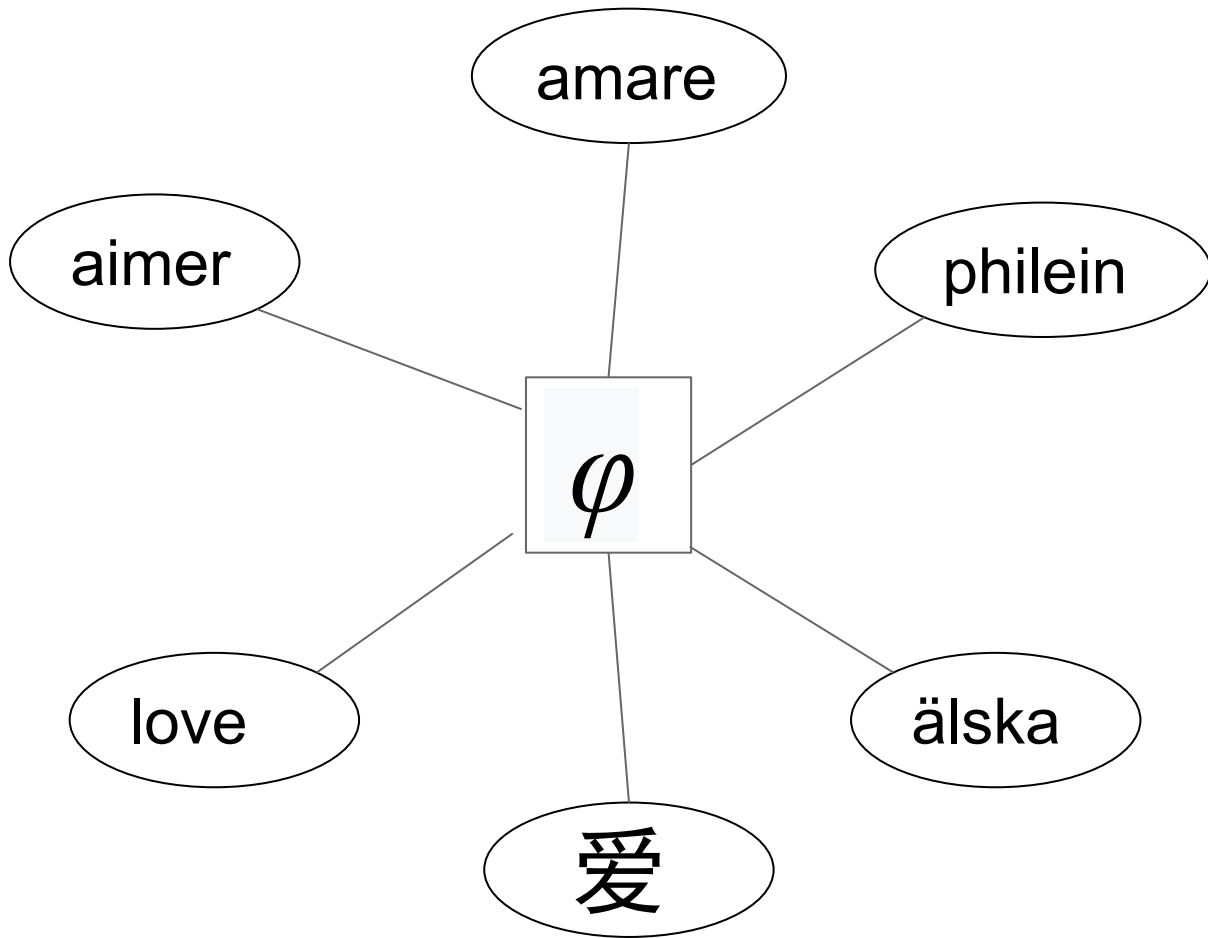
ce avec elle des angles droits. Je suppose la chose desia  
faite, & que la ligne cherchée est C P, laquelle ie pro-  
longe iusques au point P, ou elle rencontre la ligne droi-  
te G A, que ie suppose estre celle aux points de laquelle  
on rapporte tous ceux de la ligne C E: en forte que fai-  
sant M A ou C B  $\propto y$ , & C M, ou B A  $\propto x$ , iay quelque  
equation, qui explique le rapport, qui est entre  $x$  &  $y$ .  
Puis ie fais P C  $\propto s$ , & P A  $\propto v$ , ou P M  $\propto v - y$ , & a  
cause du triangle rectangle P M C iay  $ss$ , qui est le quar-  
ré de la base esgal à  $xx + vv - 2vy + yy$ , qui sont  
les quarrés des deux costés. c'est a dire iay  $x \propto$   
 $\sqrt{ss - vv + 2vy - yy}$ , ou bien  $y \propto v + \sqrt{ss - xx}$ , &  
par le moyen de cete equation, l'oste de l'autre equa-  
tion qui m'explique le rapport qu'ont tous les points de la  
courbe C E a ceux de la droite G A, l'une des deux quan-  
tités indeterminées  $x$  ou  $y$ . ce qui est ayzé a faire en  
mettant partout  $\sqrt{ss - vv + 2vy - yy}$  au lieu d' $x$ , &  
le quarré de cete somme au lieu d' $xx$ , & son cube au lieu  
d' $x^3$ , & ainsi des autres, si c'est  $x$  que ie veuille oster; ou-  
bien

om han hade i sin ordbok en symbol för *aimer, amare, philein* och alla synonymerna, kunde en **bok som är skriven i sådana symboler översättas av alla som hade denna ordbok.**

man kan på en dag lära sig namnen på alla tal och därmed skriva **oändligt många tal i ett okänt språk.**

Detsamma kunde göras med alla andra ord som behövs för att uttrycka alla andra ting

**Descartes, brev till Mersenne, 1629**



# Kompilator: program till maskinkod

1	00000100
+	00000101
2	00000110
*	01101000
3	01100000



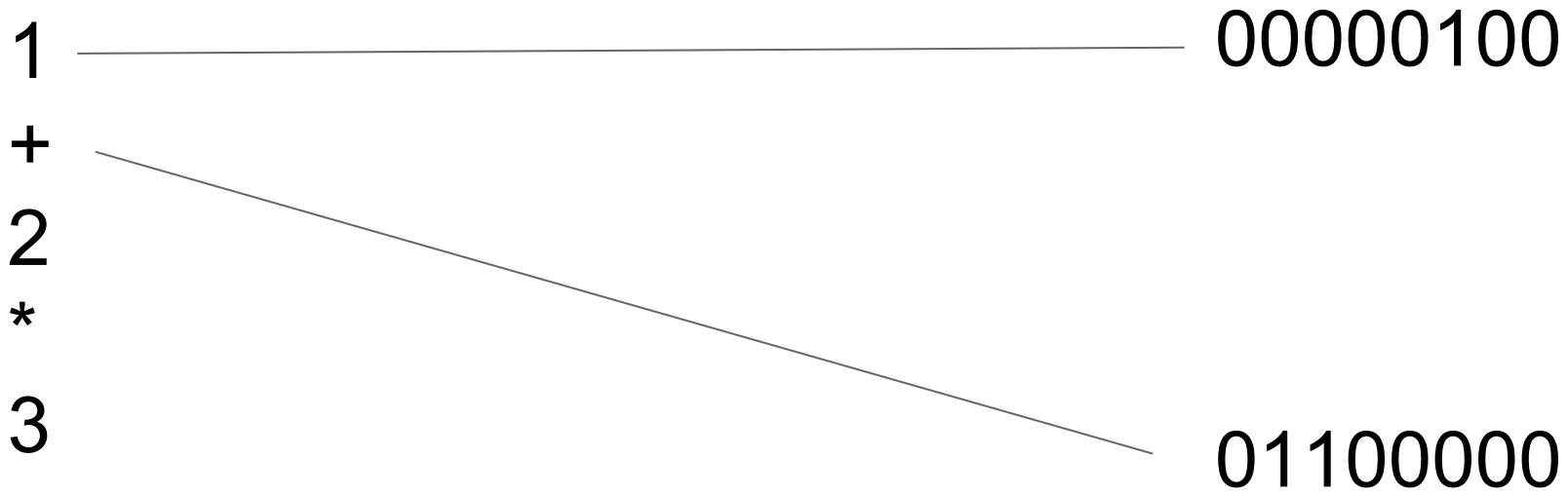
1 \_\_\_\_\_ 00000100

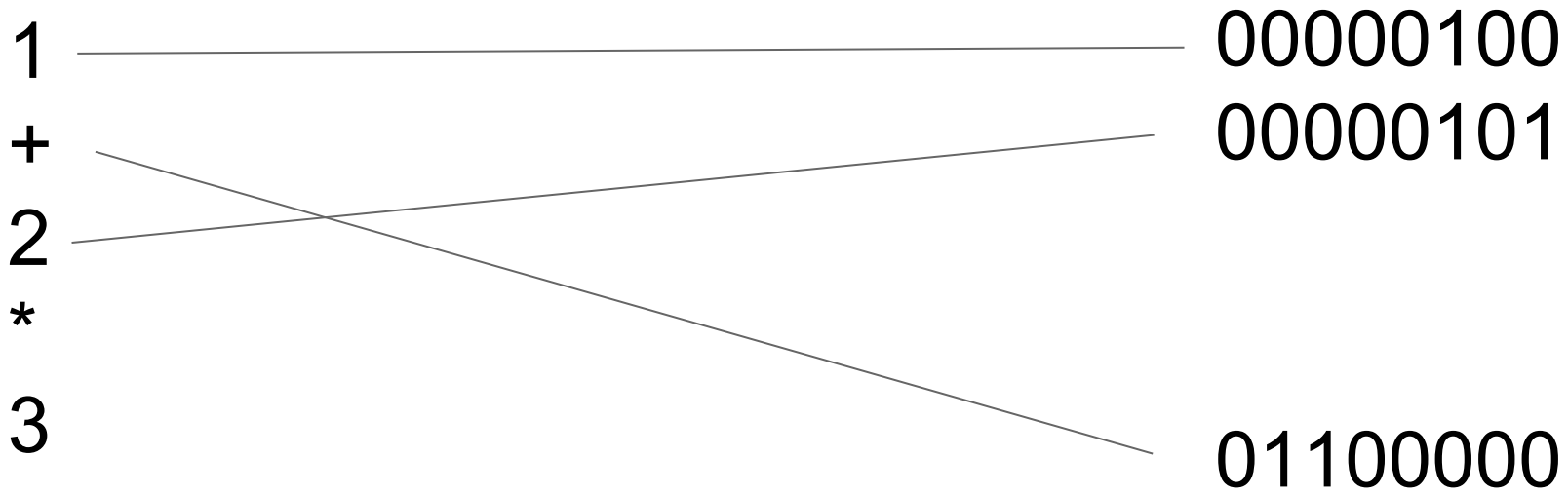
+

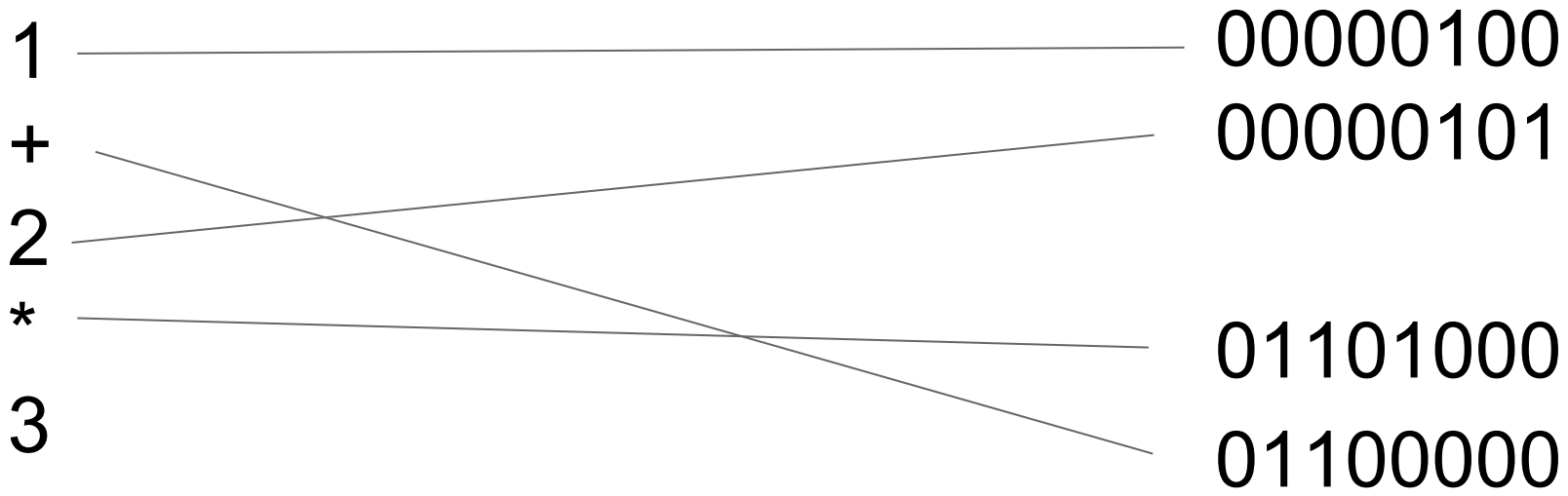
2

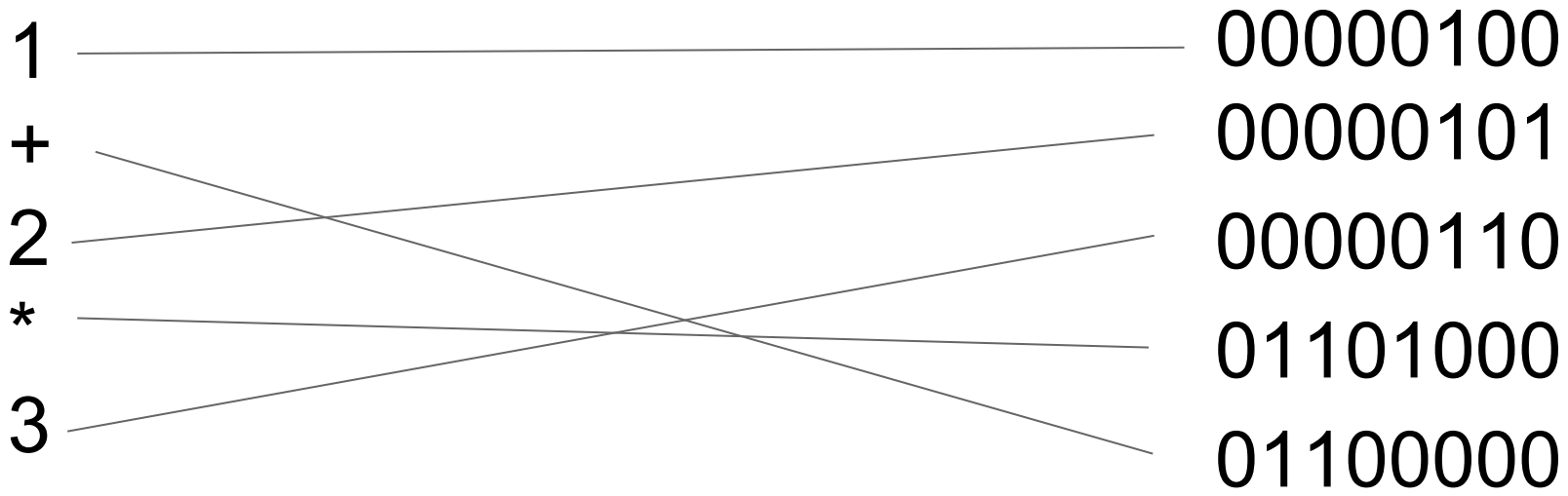
\*

3

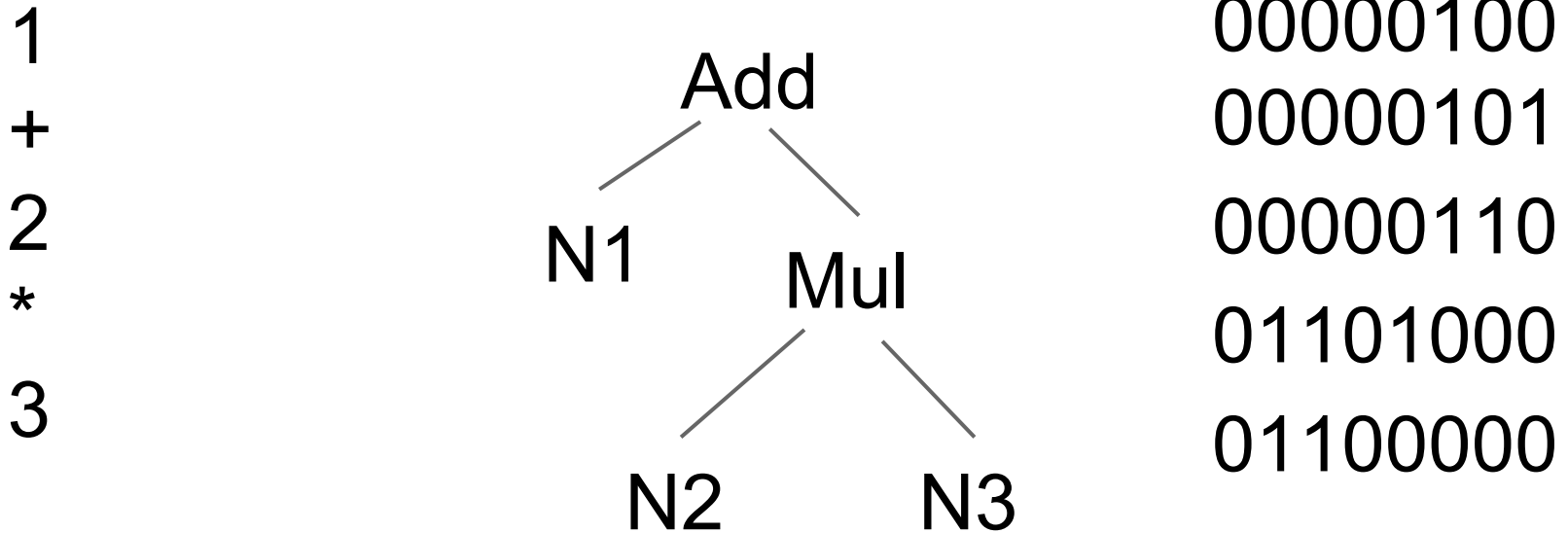








# Interlingua: abstrakt syntax



# Grammatik

## interlingua

Add : Exp -> Exp -> Exp

## Java

Add X Y = X "+" Y

## JVM

Add X Y = X Y "01100000"

# Översättning som kompilering

the  
black  
cats  
see  
me



the  
black  
cats  
see  
me

\_\_\_\_\_ le

the

black

cats

see

me



le

noir

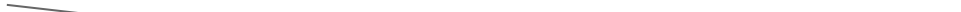
the

black

cats

see

me



le

chat

noir

the

black

cats

see

me



le

chat

noir

voir

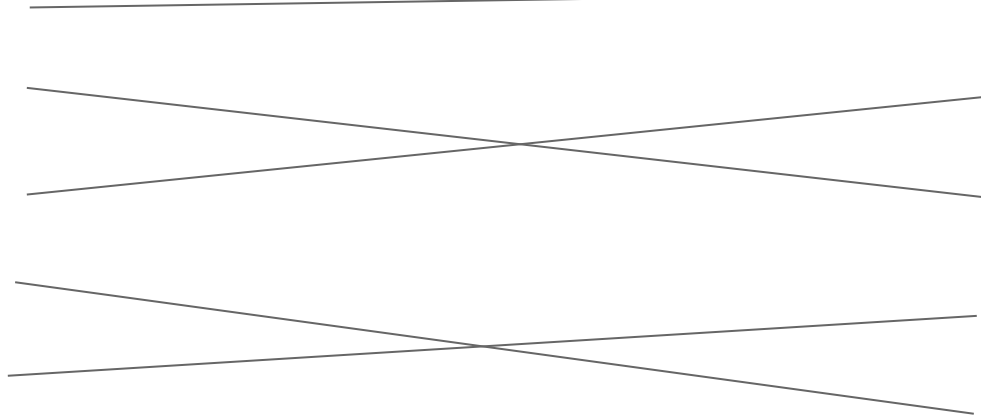
the

black

cats

see

me



le

chat

noir

je

voir

the

black

cats

see

me

\_\_\_\_\_

\_\_\_\_\_

\_\_\_\_\_

\_\_\_\_\_

\_\_\_\_\_

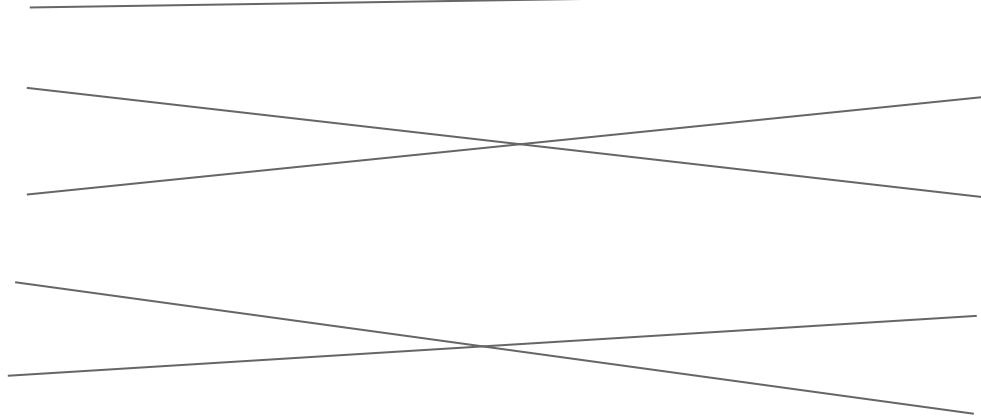
les

chats

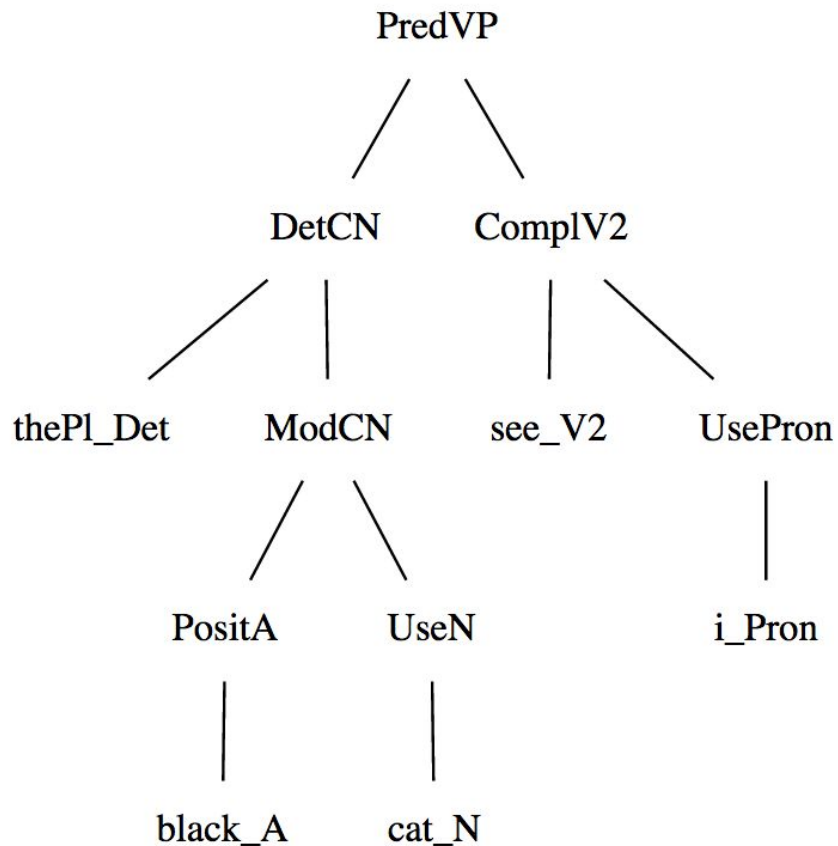
noirs

me

voient



the  
black  
cats  
see  
me



les  
chats  
noirs  
me  
voient

# Grammatik

## abstrakt syntax

ModCN : AP  $\rightarrow$  CN  $\rightarrow$  CN

## engelska

ModCN ap cn =  $\backslash \backslash n \Rightarrow$  ap ++ cn ! n

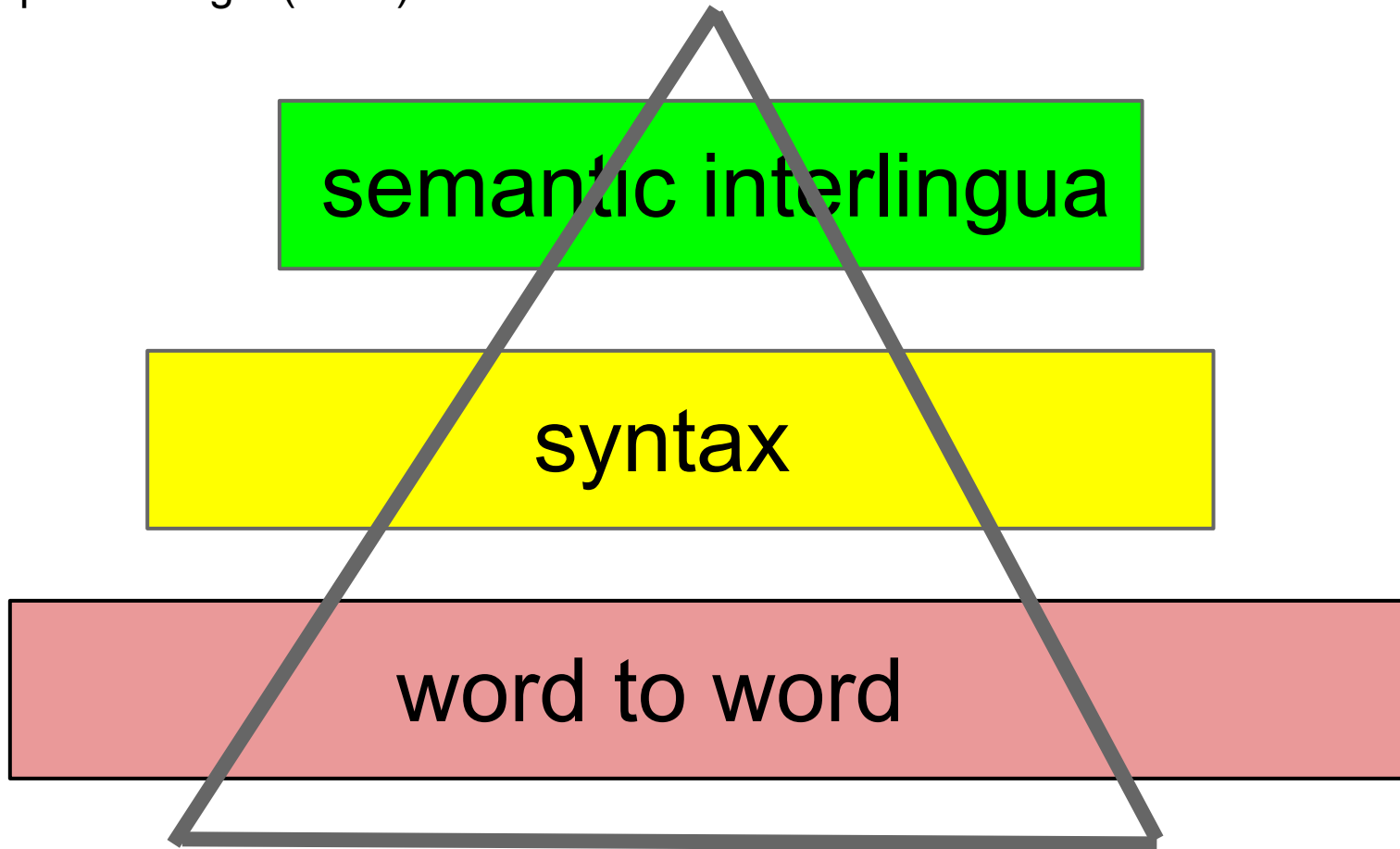
## franska

ModCN ap cn =

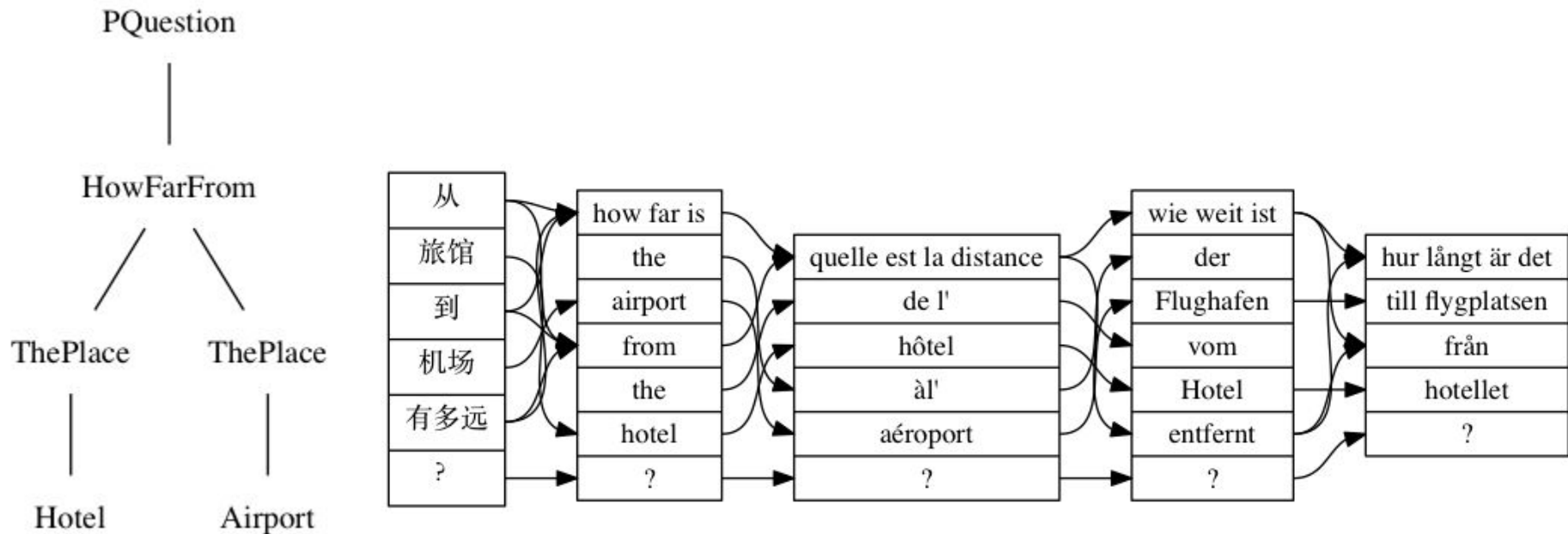
$\backslash \backslash n \Rightarrow$  cn ! n ++ ap ! cn.g ! n



Vauquois triangle (1968)



# Semantisk interlingua





# Wide Coverage Translation Demo

Swedish ▾

Clear



English ▾



Colors

Translate

Grammars...

Vad heter din dotter?  
Min mor är inte svensk.  
Det sups för mycket på Vetenskapsfestivalen.  
Vi kan tyvärr inte översätta vad som helst.

What is your daughter's name??  
My mother is not Swedish.  
Fore of snifter of it very on Vetenskapsfestivalen.  
We entrenchment unfortunately not to translate what  
like [helst].

Enter text to translate above

Try Google Translate

Swedish Finnish English Detect language ▾



Chinese (Simplified) English Norwegian ▾

Translate

Vad heter din dotter?  
Min mor är inte svensk.  
Det sups för mycket på Vetenskapsfestivalen.  
Vi kan tyvärr inte översätta vad som helst.



134/5000

What's your daughter's name?  
My mother is not Swedish.  
It sups too much at the Science Festival.  
Unfortunately, we can not translate anything.



# Intelligens?

SMT: fårskalle

- kan inte generalisera från 2 ord till 3 ord

# Intelligens?

## SMT: fårskalle

- kan inte generalisera från 2 ord till 3 ord

## NMT: galning

- smart men hittar på saker och kan gå hejvilt

# Intelligens?

## SMT: fårskalle

- kan inte generalisera från 2 ord till 3 ord

## NMT: galning

- smart men hittar på saker och kan gå hejvilt

## GSM: fackidiot

- kan bli bra på några få saker men är okunnig om det mesta

# Pålitlighet?

Är systemet förutsägbart?

Kan vi veta varför resultatet blir som det blir?



# Pålitlighet?

Är systemet förutsägbart?

- SMT: ja, i viss mån
- NMT: inte alls
- GSM: ja, i princip fullständigt

Kan vi veta varför resultatet blir som det blir?

# Pålitlighet?

Är systemet förutsägbart?

- SMT: ja, i viss mån
- NMT: inte alls
- GSM: ja, i princip fullständigt

Kan vi veta varför resultatet blir som det blir?

- SMT: ja, vi kan se det bit för bit
- NMT: mycket svårt: bara siffror
- GSM: ja, vi kan se interlinguan

# Evidens i SMT

*This big house is yellow*

*Denna stora hus är gul*



# Evidens i NMT

```
array([-0.05370992, -0.01796519, -0.13489808, -0.00400016, -0.01886696, -0.01855153, -0.0590021,
0.04081715, -0.01833461, -0.01756546, -0.03327245, -0.05934121, -0.13161591, -0.09330324,
-0.0576504, -0.06708767, -0.14609909, -0.06536276, -0.04444694, -0.06847347, -0.0038306,
0.08097503, -0.1450344, -0.0606285, -0.05798667, -0.02206576, -0.02363058, -0.01232632,
0.04450377, -0.0536673, -0.14820194, -0.03370629, -0.00571465, -0.10534635, -0.06061808,
0.05924838, -0.01724624, -0.00195224, -0.08353445, -0.07976257, -0.05860237, -0.02358891,
-0.14326403, -0.02775767, -0.05105672, -0.07834172, -0.01482512, -0.10593458, -0.07428473,
-0.00392154, -0.06843369, -0.0286187, -0.03206379, -0.01065825, -0.0212142, -0.038199,
-0.01821716, -0.16778027, -0.06967456, -0.02450488, -0.03385879, -0.0763156, -0.00977732,
-0.0511325, -0.00714402, -0.07367945, -0.0687027, -0.00737988, -0.00394427, -0.08146569,
-0.03385974, -0.03460994, -0.00039784, -0.0203238, -0.03031046, -0.04941517, -0.09776281,
0.17635746, -0.00446904, -0.05661129, -0.05412859, -0.04316155, -0.07147998, -0.05980725,
-0.06233541, -0.10460561, -0.00153925, -0.04334057, -0.0265348, -0.03904583, -0.06974371,
-0.02253748, -0.00371694, -0.03108814, -0.08722486, -0.08058666, -0.08066339, -0.06889972,
0.05318894, -0.0111025, -0.04847362, -0.04241608, -0.02344587, -0.11333624, -0.01625354,
0.10140302, -0.03682268, -0.09101, -0.01545408, -0.0857216, -0.0635886, -0.01903083,
0.06806806, -0.06100928, -0.08224337, -0.01855342, -0.01142929, -0.0219663, -0.11795305,
-0.05691156, -0.03229586, -0.07092301, -0.10715461, -0.07458216, -0.07924633, -0.08229263,
0.20106314, -0.12279814, -0.03754162, -0.01622134, -0.06508806, -0.06969255, -0.02829286,
-0.02122651, -0.11400309, -0.07765214, -0.03194822, -0.04968892, -0.04011241, -0.02989273,
-0.04679892, -0.13507046, -0.02070364, -0.01047164, -0.03619466, -0.05266512, -0.12246187,
-0.00721033, -0.10127704, -0.00698299, -0.04904006, -0.06502365, -0.01203647, -0.06826507,
0.0650928, -0.01393946, -0.13613988, -0.06799417, -0.03203158, -0.02859124, -0.07497147,
0.018202, -0.07249352, -0.1334185, -0.02979043, -0.02842926, -0.09712628, -0.08914774,
0.15470658, -0.03547382, -0.15360495, -0.01643541, -0.09154803, -0.16215466, -0.14822088,
-0.01966358, -0.04322122, -0.00516653, -0.02396685, -0.0373105, -0.07382059, -0.15486667,
0.01114797, -0.01211035, -0.09367077, -0.02892656, -0.10523268, -0.06287628, -0.05812117,
-0.00592967, -0.01626207, -0.07094574, -0.06422988, -0.01778995, -0.09563628, -0.10500913,
-0.09146846, -0.01761282, -0.02320812, -0.05757652], dtype=float32)
```

Swedish Finnish Chinese Detect language ▾



Chinese (Simplified) English Swedish ▾

极端愚蠢

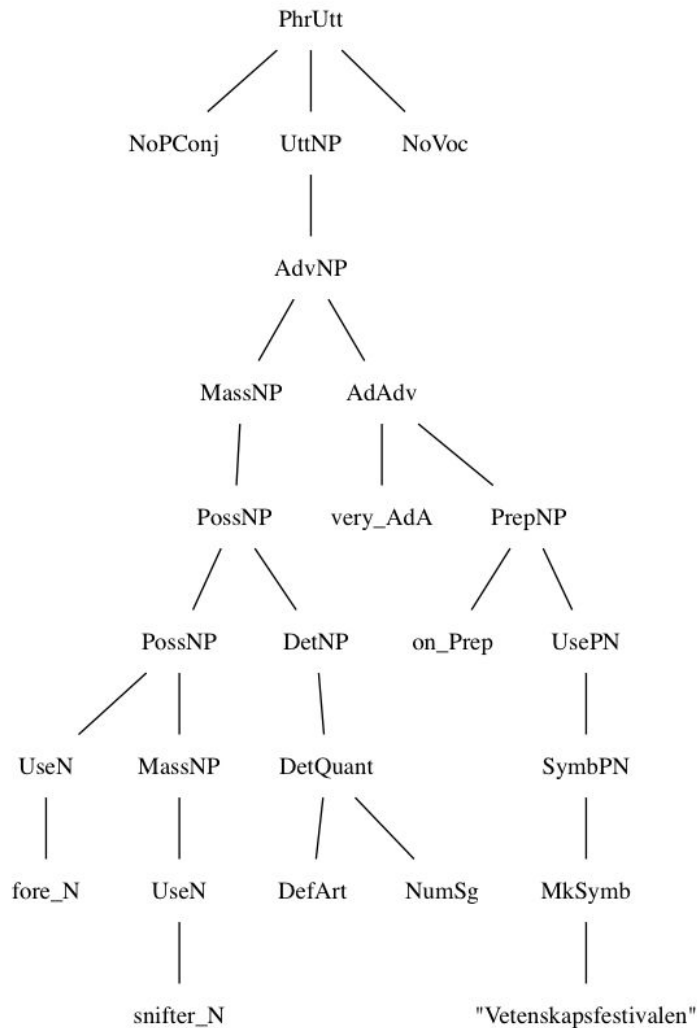


Extreme dårskap

# Evidens i GSM

*Det sups för mycket på  
Vetenskapsfestivalen.*

*Fore of snifter of it very on  
Vetenskapsfestivalen.*



## maskinen fungerar inte på golvet

PhrUtt NoPConj (UttS (UseCl (TTAnt TPres ASimul) PNeg (PredVP (DetCN (DetQuant DefArt NumSg) (UseN machine\_N)) (AdvVP (UseV work\_2\_V) (PrepNP on\_Prep (DetCN (DetQuant DefArt NumSg) (UseN floor\_N)))))) NoVoc

## maskinen arbetar inte på golvet

PhrUtt NoPConj (UttS (UseCl (TTAnt TPres ASimul) PNeg (PredVP (DetCN (DetQuant DefArt NumSg) (UseN machine\_N)) (AdvVP (UseV work\_1\_V) (PrepNP on\_Prep (DetCN (DetQuant DefArt NumSg) (UseN floor\_N)))))) NoVoc

## maskinen ordnar inte om golvet

PhrUtt NoPConj (UttS (PredVPS (DetCN (DetQuant DefArt NumSg) (UseN machine\_N)) (MkVPS (TTAnt TPres ASimul) PNeg (ComplV2 work\_on\_V2 (DetCN (DetQuant DefArt NumSg) (UseN floor\_N)))))) NoVoc

# Resten av föredraget

System som är begränsade men pålitliga

- finns det behov?
- kan de bli lönsamma?
- hur kan de användas?

# Resten av föredraget

System som är begränsade men pålitliga

- finns det behov?
- kan de bli lönsamma?
- hur kan de användas?

Projekt GF = Grammatical Framework

- Göteborg & open source community 1998-
- Digital Grammars AB 2014-



# Konsument vs. producent

## Konsument

- andra språk → mitt språk
- få en idé
- ingen har ansvar

# Konsument vs. producent

## Konsument

- andra språk → mitt språk
- få en idé
- ingen har ansvar

## Producent

- mitt språk → andra språk
- förmedla budskapet
- ansvar för innehållet

# Exempel på producenter

Myndigheter

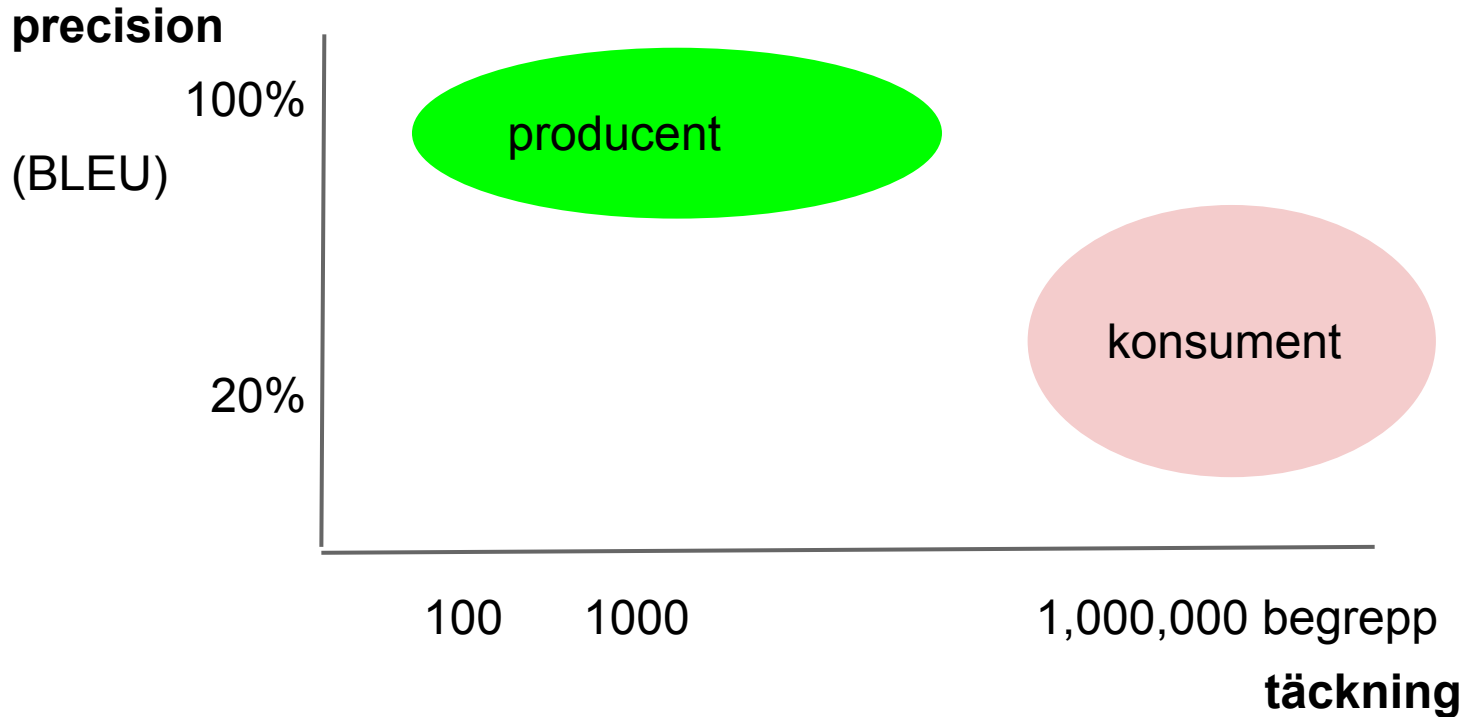
Företag

Sjukvården

Författare av kärleksbrev

Författare av phishing-mail

# Problemets omfattning





- Svängrumsytan utanför dörren lutar 1% i sidled.
- The turning space outside the gate tilts 1% sideways.
- Kääntymätila oven ulkopuolella kallistuu 1% sivusuunnassa.
- Der Schwenkbereich außerhalb der Tür neigt sich um 1% seitlich.
- La zona de giro fuera de la puerta inclina 1 % de lado.

UttSTD (PredSTD (AdvNPTD (DetCNTD (DetQuant DefArt NumSg) (UseNTD svängrumsyta\_NTD)) (PrepNPTD utanför\_Prep (DetCNTD (DetQuant DefArt NumSg) (UseNTD dörr\_NTD)))) (AdvVPTD (luta\_VPTD (procentMeasure 1)) i\_sidled\_AdvTD))

**Talkamatic**  
FREE DIALOGUE



digital **G**rammars  
Language technology to rely on.

`next_membership_level_sys_answer silver (next_membership_points_sys_answer integer0_99_50)`

`test_mockup_travelChi: 您有五十个常旅客点符合会员条件, 您现在是在伦敦.`

`test_mockup_travelDut: je hebt vijftig punten nodig om het zilveren niveau te bereiken`

`test_mockup_travelEng: you need fifty points to reach silver level`

`test_mockup_travelFin: sinä tarvitset viisikymmentä pistettä päästäksesi hopeatasolle`

`test_mockup_travelFre: tu as besoin de cinquante points pour atteindre le niveau argent`

`test_mockup_travelGer: Sie brauchen fünfzig Punkte um das Silberniveau zu erreichen`

`test_mockup_travelIta: avete bisogno di cinquanta punti per raggiungere il livello argento`

`test_mockup_travelSpa: necesitas cincuenta puntos para llegar al nivel plata`

medium-sized enterprises , and to provide natural persons in all Member States with the same level of legally enforceable rights and obligations and responsibilities for controllers and processors , to ensure consistent monitoring of the processing of personal data , and equivalent sanctions in all Member States as well as effective cooperation between the supervisory authorities of different Member States . The proper functioning of the internal market requires that the free movement of personal data within the Union is not restricted or prohibited for reasons connected with the protection of natural persons with regard to the processing of personal data . To take account of the

mittlerer Unternehmen Rechtssicherheit und Transparenz schafft , natürliche Personen in allen Mitgliedstaaten mit demselben Niveau an durchsetzbaren Rechten ausstattet , dieselben Pflichten und Zuständigkeiten für die Verantwortlichen und Auftragsverarbeiter vorsieht und eine gleichmäßige Kontrolle der Verarbeitung personenbezogener Daten und gleichwertige Sanktionen in allen Mitgliedstaaten sowie eine wirksame Zusammenarbeit zwischen den Aufsichtsbehörden der einzelnen Mitgliedstaaten gewährleistet . Das reibungslose Funktionieren des Binnenmarkts erfordert , dass der freie Verkehr personenbezogener Daten in der Union nicht aus Gründen des Schutzes natürlicher Personen bei der Verarbeitung personenbezogener Daten eingeschränkt oder verboten wird . Um der besonderen Situation der Kleinunternehmen sowie der kleinen und mittleren

ostacolare la libera circolazione dei dati personali nel mercato interno , è necessario un regolamento che garantisca certezza del diritto e trasparenza agli operatori economici , comprese le micro , piccole e medie imprese , offra alle persone fisiche in tutti gli Stati membri il medesimo livello di diritti azionabili e di obblighi e responsabilità dei titolari del trattamento e dei responsabili del trattamento e assicuri un monitoraggio coerente del trattamento dei dati personali , sanzioni equivalenti in tutti gli Stati membri e una cooperazione efficace tra le autorità di controllo dei diversi Stati membri . Per il buon funzionamento del mercato interno è necessario che la libera circolazione dei dati personali all' interno dell' Unione non sia limitata né vietata per motivi attinenti alla protezione delle persone fisiche con riguardo al trattamento dei dati personali . Per tener conto della specifica situazione delle micro , piccole e medie imprese , il presente regolamento

## EU-lagtexter

- 3000+ begrepp, 5-24 språk

Grammar: Inlife.pgf i Startcat: S From: Swe To: ... Clear Random

min familj bryr sig inte

Abstract care\_\_S (detNPH (singularDet (PossPron i\_Pron)) family\_0351\_CNH)

Eng ▶ my family doesn't care en-US

Fin ▶ minun perheeni ei välitä fi-FI

Ger ▶ es ist meiner Familie egal de-DE de-DE

Slv ▶ mòja družina ne zanima

Swe ▶ min familj bryr sig inte sv-SE

InLife: kommunikationsstöd för äldre (DART & Digital Grammars AB)

- 1778 begrepp, 5 språk



Ambulanstolken (SU ambulans, MedTech West,  
PreHospiten vid Högskolan i Borås,  
Digital Grammars AB)

- 500-1000 begrepp, 5-6 språk i början
- SBAR:
  - Situation
  - Bakgrund
  - Aktuell information
  - Rekommendation



S	B	A	R	Microphone	Keyboard
Jag heter Stig/... .					
Jag är sjuksköterska/... .					
Vad heter du?					
Vad är ditt personnummer?					
Hur gammal är du?					
Kan någon svenska/... ?					
Vilket språk talar du/... ?					
Kan patienten prata själv?					
Vad har hänt?					
Har det hänt en olycka/... ?					
Var är patienten?					
Är situationen livshotande?					
Vem kallade ambulans?					
Finns det fler patienter?					

```

TitleParagraph DefinitionTitle
DefPredParagraph type_Sort A_Var contractible_Pred (ExistCalledProp a_Var (ExpSort (VarExp A_Var)) (FunInd centre_of_contraction_Fun) (ForAllProp (BaseVar x_Var) (ExpSort (VarExp A_Var)) (ExpProp (equalExp (VarExp a_Var) (VarExp x_Var))))))
FormatParagraph EmptyLineFormat
TitleParagraph DefinitionTitle
DefPredParagraph (mapSort (mapExp (VarExp A_Var) (VarExp B_Var))) f_Var equivalence_Pred (ForAllProp (BaseVar y_Var) (ExpSort (VarExp B_Var)) (PredProp contractible_Pred (AliasInd (AppFunItnd fiber_Fun) (FunInd (ExpFun (ComprehensionExp x_Var (VarExp A_Var) (equalExp (AppExp f_Var (VarExp x_Var)) (VarExp y_Var))))))))))
DefPropParagraph (ExpProp (equivalenceExp (VarExp A_Var) (VarExp B_Var))) (ExistSortProp (equivalenceSort (mapExp (VarExp A_Var) (VarExp B_Var))))
FormatParagraph EmptyLineFormat
TitleParagraph LemmaTitle
TheoremParagraph (ForAllProp (BaseVar A_Var) type_Sort (PredProp equivalence_Pred (AliasInd (FunInd identity_map_Fun) (FunInd (ExpFun (DefExp (identityMapExp (VarExp A_Var)) (TypedExp (BaseExp (lambdaExp x_Var (VarExp A_Var) (VarExp x_Var))) (mapExp (VarExp A_Var) (VarExp A_Var))))))))))
FormatParagraph EmptyLineFormat
TitleParagraph ProofTitle
AssumptionParagraph (ConsAssumption (ForAssumption y_Var (ExpSort (VarExp A_Var)) (LetAssumption (FunInd (ExpFun (DefExp (fiberExp (VarExp y_Var) (VarExp A_Var)) (ComprehensionExp x_Var (VarExp A_Var) (equalExp (VarExp x_Var) (VarExp y_Var)))))) (AppFunItnd (fiberWrt_Fun (FunInd (ExpFun (identityMapExp (VarExp A_Var)))))) (BaseAssumption (LetExpAssumption (barExp (VarExp y_Var)) (TypedExp (BaseExp (pairExp (VarExp y_Var) (reflexivityExp (VarExp A_Var) (VarExp y_Var)))) (fiberExp (VarExp y_Var) (VarExp A_Var))))))
ConclusionParagraph (AsConclusion (ForAllProp (BaseVar y_Var) (ExpSort (VarExp A_Var)) (ExpProp (equalExp (pairExp (VarExp y_Var) (reflexivityExp (VarExp A_Var) (VarExp y_Var))) (VarExp y_Var)))) (ApplyLabelConclusion id_induction_Label (ConsInd (FunInd (ExpFun (VarExp y_Var))) (ConsInd (FunInd (ExpFun (TypedExp (BaseExp (VarExp x_Var)) (VarExp A_Var)))) (ConsInd (FunInd (ExpFun (TypedExp (BaseExp (VarExp z_Var)) (idPropExp (VarExp x_Var) (VarExp y_Var)))) BaseInd))) (DisplayExpProp (equalExp (pairExp (VarExp x_Var) (VarExp z_Var)) (VarExp y_Var))))
ConclusionSoThatParagraph (ForConclusion (BaseVar y_Var) (ExpSort (VarExp A_Var)) (A BaseInd) (ExpProp (equalExp (VarExp u_Var) (VarExp y_Var)))) (PredProp contractible_Pri
ConclusionParagraph (PropConclusion (PredProp equivalence_Pred (FunInd (ExpFun (Type
QEDParagraph

```

**Definition:** A type  $A$  is contractible, if there is  $a : A$ , called the center of contraction, such that for all  $x : A$ ,  $a = x$ .

**Definition:** A map  $f : A \rightarrow B$  is an equivalence, if for all  $y : B$ , its fiber,  $\{x : A \mid fx = y\}$ , is contractible. We write  $A \simeq B$ , if there is an equivalence  $A \rightarrow B$ .

**Lemma:** For each type  $A$ , the identity map,  $1_A := \lambda x:A x : A \rightarrow A$ , is an equivalence.

**Proof:** For each  $y : A$ , let  $\{y\}_A := \{x : A \mid x = y\}$  be its fiber with respect to  $1_A$  and let  $\bar{y} := (y, r_A y) : \{y\}_A$ . As for all  $y : A$ ,  $(y, r_A y) = y$ , we may apply Id-induction on  $y$ ,  $x : A$  and  $z : (x = y)$  to get that

$$(x, z) = y$$

. Hence, for  $y : A$ , we may apply  $\Sigma$ -elimination on  $u : \{y\}_A$  to get that  $u = y$ , so that  $\{y\}_A$  is contractible. Thus,  $1_A : A \rightarrow A$  is an equivalence.  $\square$

**Définition:** Un type  $A$  est contractible, s'il existe un de contraction, tel que pour tous les  $x : A$ ,  $a = x$ .

**Définition:** Une application  $f : A \rightarrow B$  est une équivalence, si pour tous les  $y : B$ , sa fibre,  $\{x : A \mid fx = y\}$ , est contractible. Nous écrivons  $A \simeq B$ , si il existe une équivalence  $A \rightarrow B$ .

**Lemme:** Pour tout type  $A$ , l'identité,  $1_A := \lambda x:A x$ : est une équivalence.

**Démonstration:** Pour tout  $y : A$ , soit  $\{y\}_A := \{x : A \mid x = y\}$  sa fibre par rapport à  $1_A$  et soit  $\bar{y} := (y, r_A y) : \{y\}_A$ . Comme pour tous les  $y : A$ ,  $(y, r_A y) = y$ , nous pouvons appliquer Id-induction sur  $y$ ,  $x : A$  et  $z : (x = y)$  pour obtenir que

$$(x, z) = y$$

. Donc, pour les  $y : A$ , nous pouvons appliquer  $\Sigma$ -élimination sur  $u : \{y\}_A$  pour obtenir que  $u = y$ , de façon que  $\{y\}_A$  soit contractible. Alors,  $1_A : A \rightarrow A$  est une équivalence.  $\square$

# GF Offline Translator



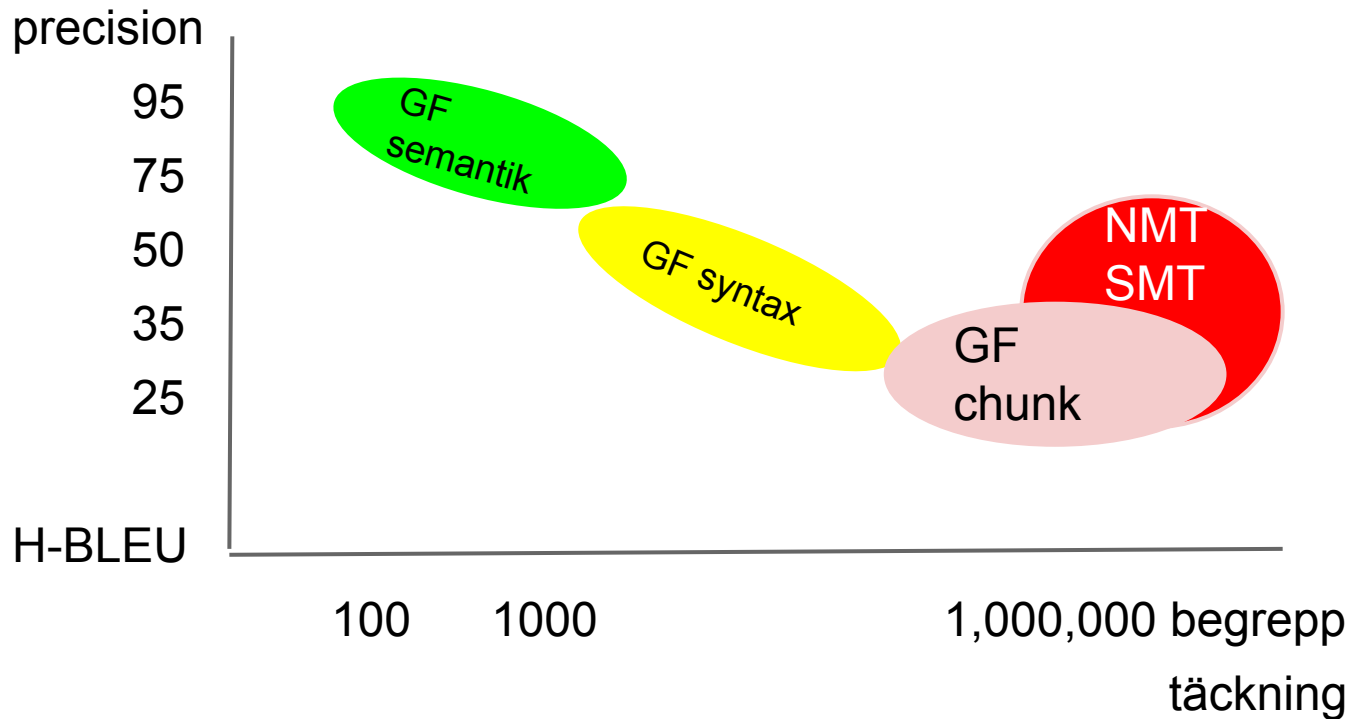
<https://play.google.com/store/apps/details?id=org.grammaticalframework.ui.android>

<https://itunes.apple.com/us/app/gf-offline-translator/id1023328422?mt=8>

K. Angelov, B. Bringert & A. Ranta, Speech-enabled hybrid multilingual translation for mobile devices, EACL 2014.



# Utvärdering



# Den återstående frågan

Hur kan det bli lönsamt i praktiken?

Grammatik: flera års arbete

vs.

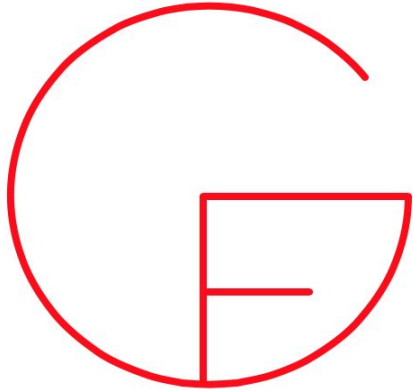
Statistik: några veckors väntande

# GFs svar

Bättre programmeringsmetoder

Grammatikbibliotek

Öppen källkod och Kooperation



*The mission of GF is to formalize the grammars of the world and make them available for computer applications.*

<http://www.grammaticalframework.org>

Since 1998

# Global community of 200+ members



Niemenen, Camilleri 2013-



## **32 languages:**

English Swedish German Dutch French Italian Spanish  
Catalan Bulgarian Russian Finnish Estonian Japanese Thai  
Chinese Hindi Icelandic Norwegian Danish Afrikaans  
Romanian Polish Russian Latvian Mongolian Urdu Punjabi  
Sindhi Greek Maltese Nepali Persian

## **More in progress:**

Latin Slovene Turkish Basque Hebrew Arabic Amharic  
Swahili ...

# Bibliotek för ordböjning

mkV “supa” “söp” “supit”



	active	passive
<b>infitiv</b>	supa	supas
<b>presens</b>	super	sups
<b>preteritum</b>	söp	söps
<b>supinum</b>	supit	supits
<b>imperativ</b>	sup	sups

## Particip Presens

		obest	best
<b>nom</b>	<b>sg</b>	supande	supandet
	<b>pl</b>	supanden	supandena
<b>gen</b>	<b>sg</b>	supandes	supandets
	<b>pl</b>	supandens	supandenas

## Particip Perfekt

		nom	gen
<b>obest</b>	<b>utr</b>	supen	supens
	<b>neut</b>	supet	supets
	<b>pl</b>	supna	supnas
<b>best</b>	<b>sg</b>	supna	supnas
	<b>pl</b>	supna	supnas

# Bibliotek för syntax

Pred

i\_NP

love\_V2

you\_NP

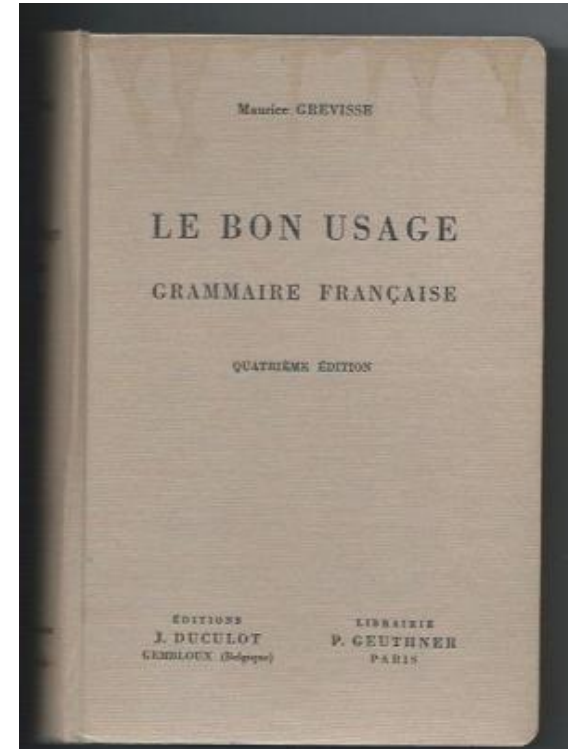
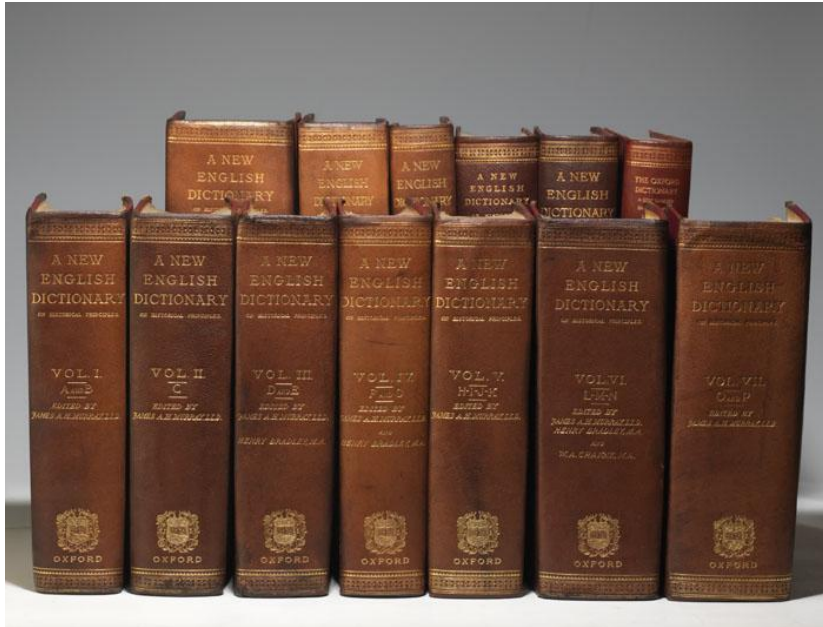


Pres Pos	<i>I love you</i>	<i>je t'aime</i>
Pres Neg	<i>I don't love you</i>	<i>je ne t'aime pas</i>
Pres Quest Pos	<i>do I love you</i>	<i>t'aime-je</i>
Pres Quest Neg	<i>don't I love you</i>	<i>ne t'aime-je pas</i>
Past Pos	<i>I loved you</i>	<i>je t'aimais</i>
Past Anter Pos	<i>I had loved you</i>	<i>je t'avais aimée</i>
Cond Anter Quest Neg	<i>wouldn't I have loved you</i>	<i>ne t'aurais-je pas aimée</i>
...	...	...

# A hundred years perspective

We can build resources that will be useful 100 years from now

Just like dictionaries and grammars written 100 years ago

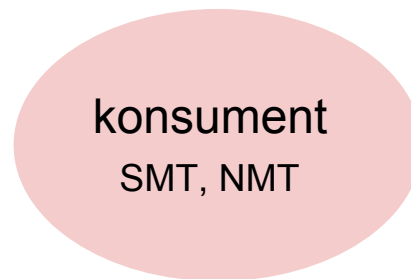


# Att ta hem

precision

100%

20%



100

1000

1,000,000

begrepp

# Att ta hem

precision

100%

producent  
"GSM"

människa  
(+dator)

20%

konsument  
SMT, NMT

100

1000

1,000,000

begrepp

