



digital  grammars
Language technology to rely on.

Grammars as Computer Programs

Aarne Ranta

University of Gothenburg and Digital Grammars AB

Doctoral Conferment Symposium

University of Tampere

16 August 2018

Grammars in science and technology

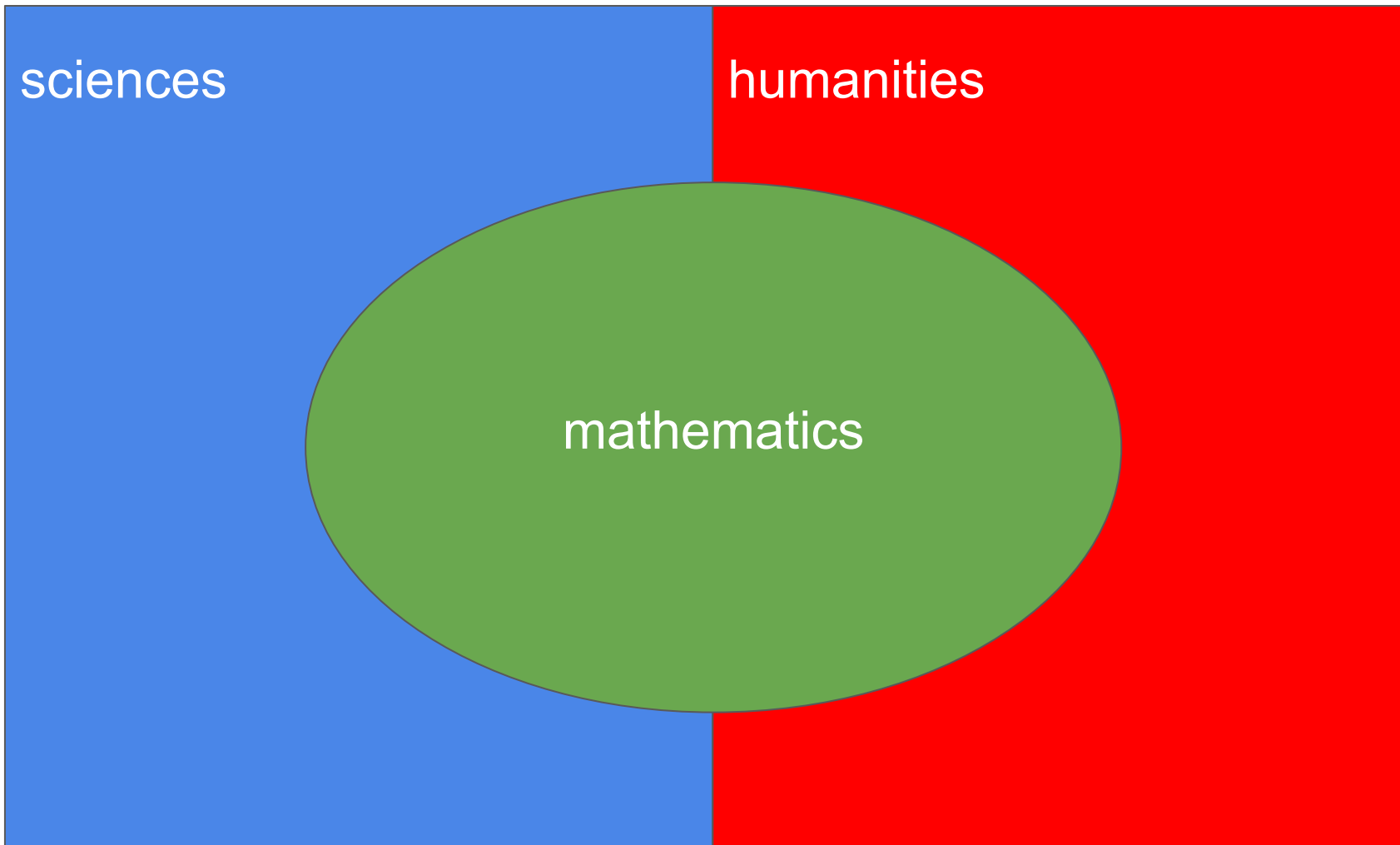
sciences

humanities

sciences

humanities

mathematics



computer science

linguistics

logic

grammars

algorithms

statistics

machine learning

computational linguistics

statistical

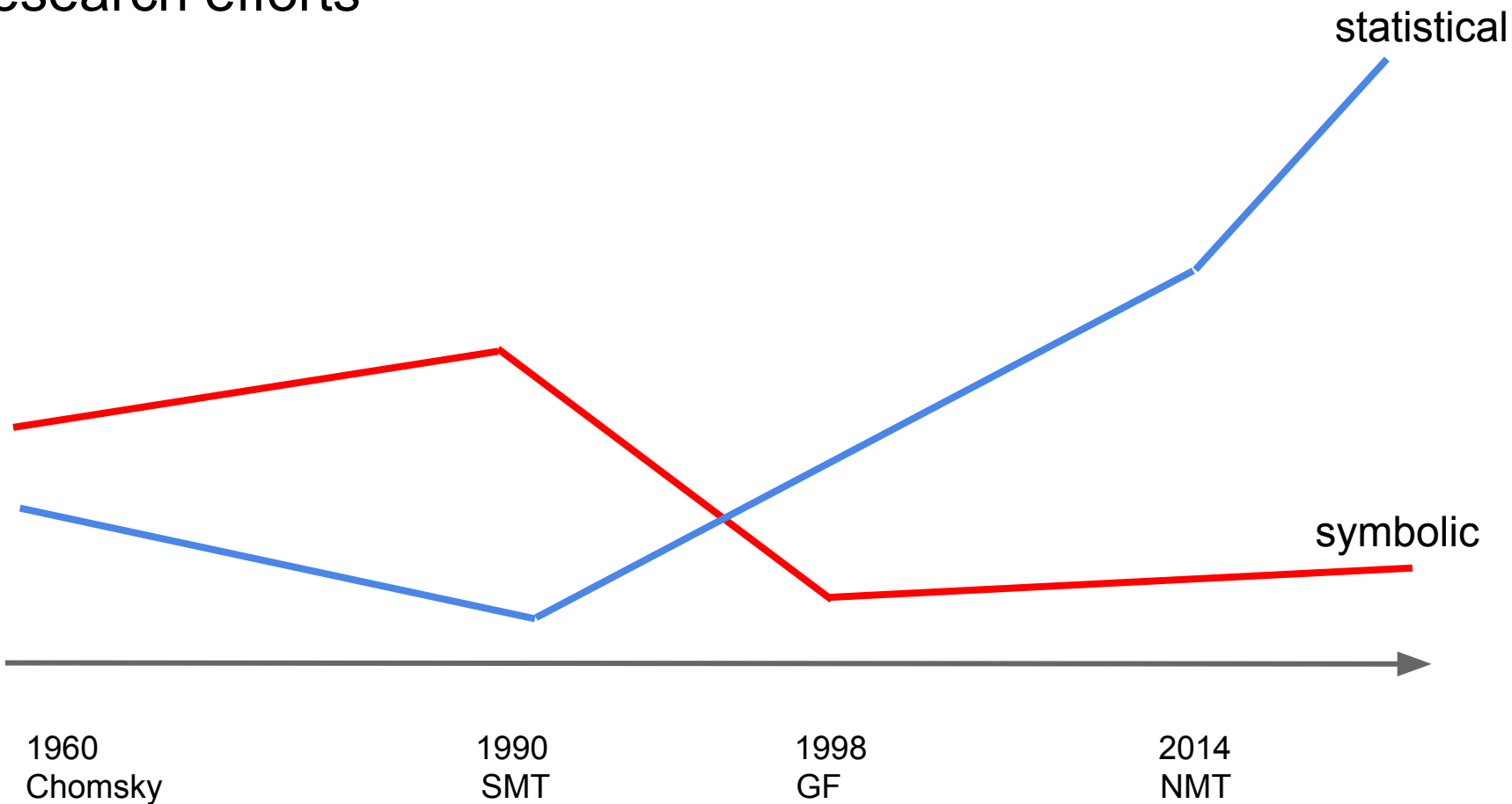
symbolic

statistics
machine learning

algorithms

logic
grammars

Research efforts



pros and cons

statistical

- opaque
- needs a lot of data
- + cheap
- + robust

symbolic

- + explainable
- + can do with little data
- expensive
- brittle

NMT

Neural Machine Translation

极端愚蠢

Extreme dårskap

NMT

Neural Machine Translation

极端愚蠢

```
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, 0.02121242, -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, -0.08516653, 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)
```

Extreme dårskap

English Swedish French Detect language ▾ ↔ Translate

Min mor är inte svensk. × 我的母亲是瑞典的。
Min mor är svensk. 我的母亲是瑞典的。

English Swedish French Detect language ▾ ↔ German English Norwegian ▾

Min mor är inte svensk. × Min mor er svensk.
Min mor är svensk. Min mor er svensk.

English Finnish Swedish Detect language ▾ ↔ Spanish Persian English ▾ Translate

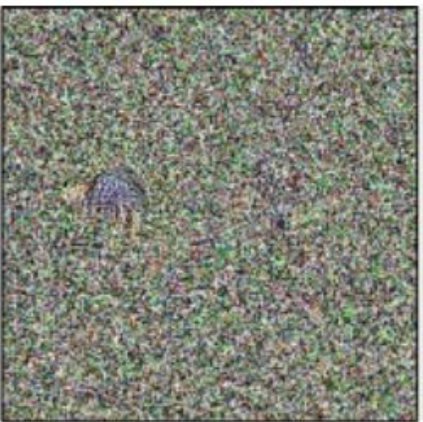
Hon är en professor i akutvård. × She is an acupuncture professor.



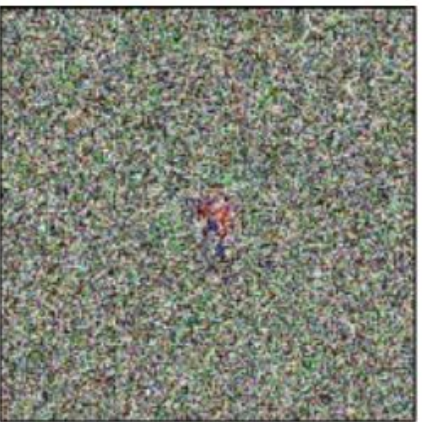
robin



cheetah



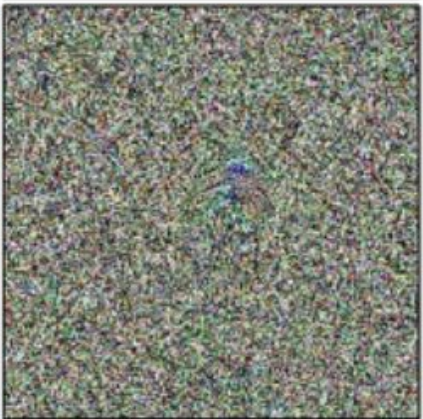
armadillo



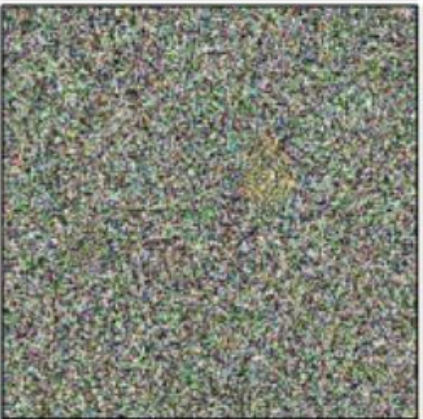
lesser panda



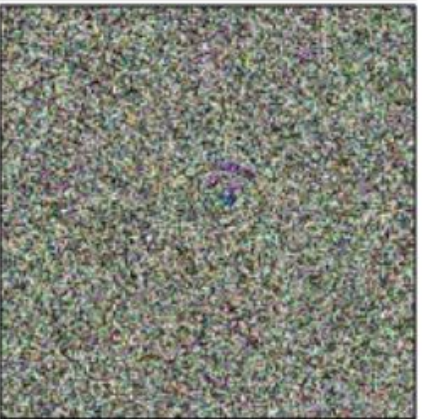
centipede



peacock



jackfruit



bubble

Translate

Chinese English Finnish Detect language ▾



Swedish Finnish English ▾

Translate

ia
ia ia
ia ia ia
ia ia ia ia ia
ia ia ia ia ia iaia ia iaiaiaia



IA
i am
Do not sleep
I do not know
and not a son of a son of a lifetime



64/5000

What comes after NMT?

What comes after NMT?

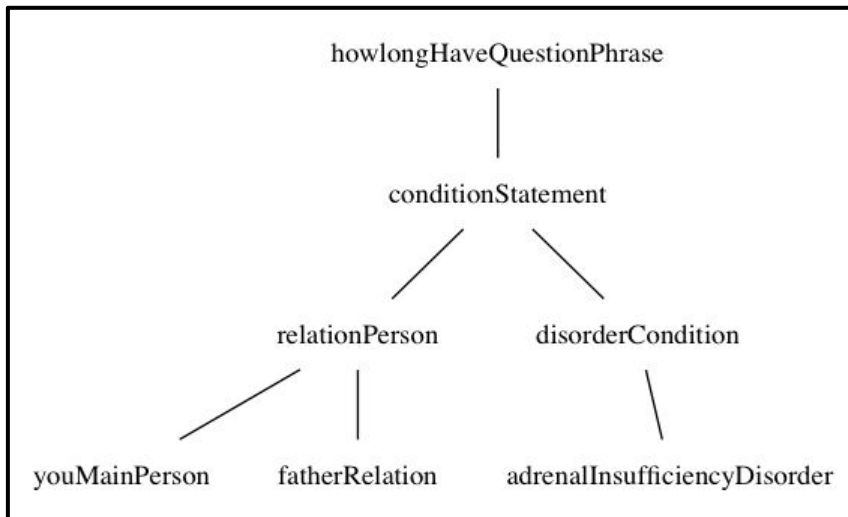
GSM.

What comes after NMT?

GSM

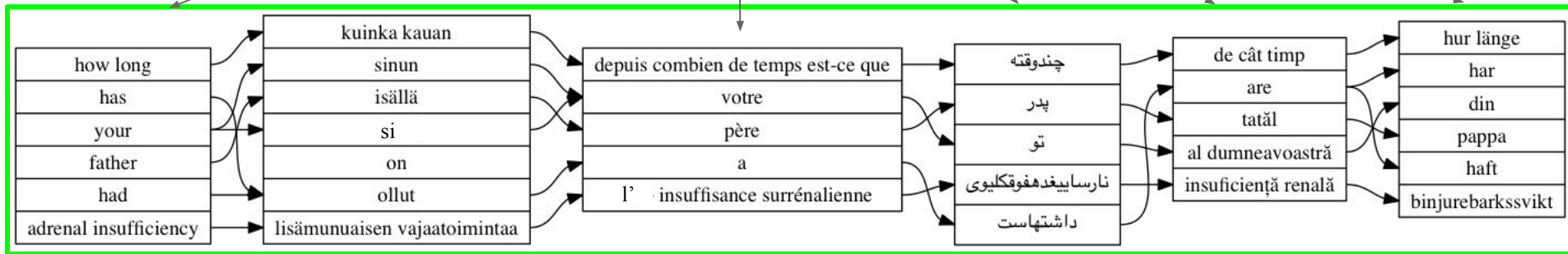
(Grammatical and Semantic Methods)

Grammar-based translation



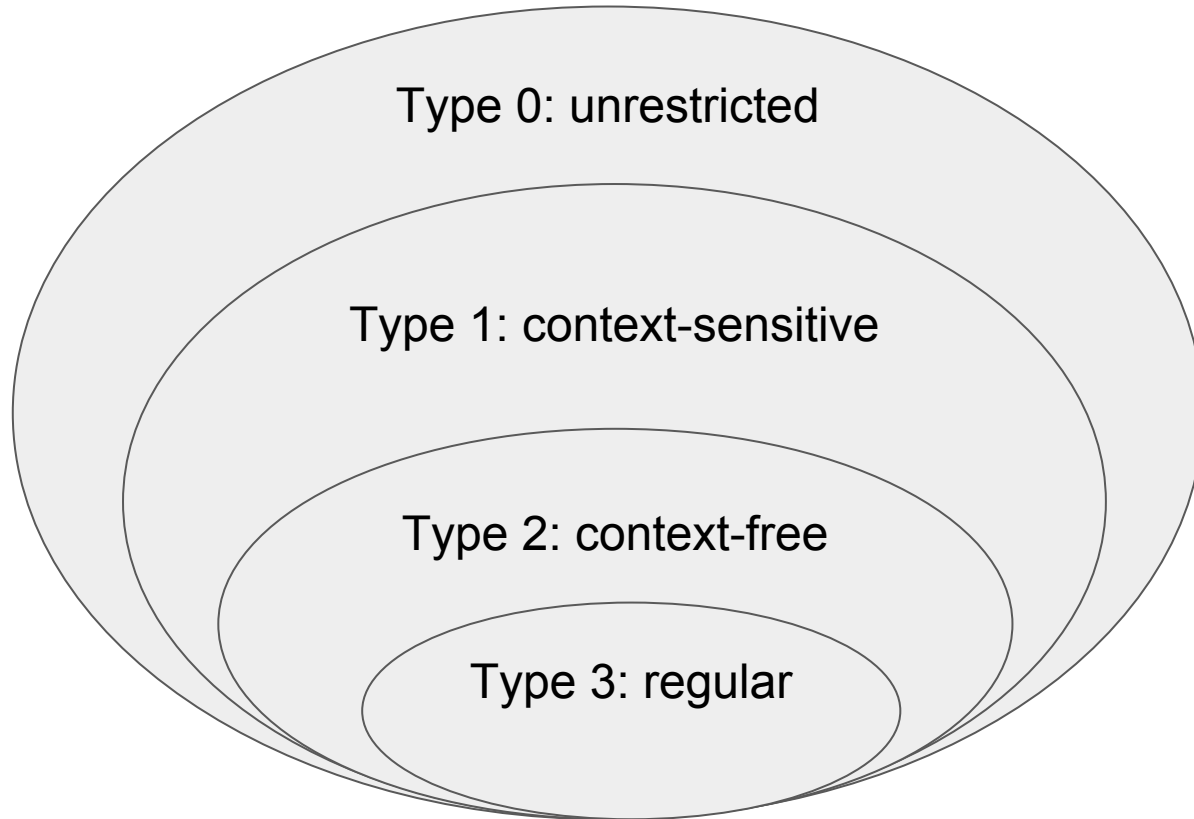
semantic interlingua

grammar

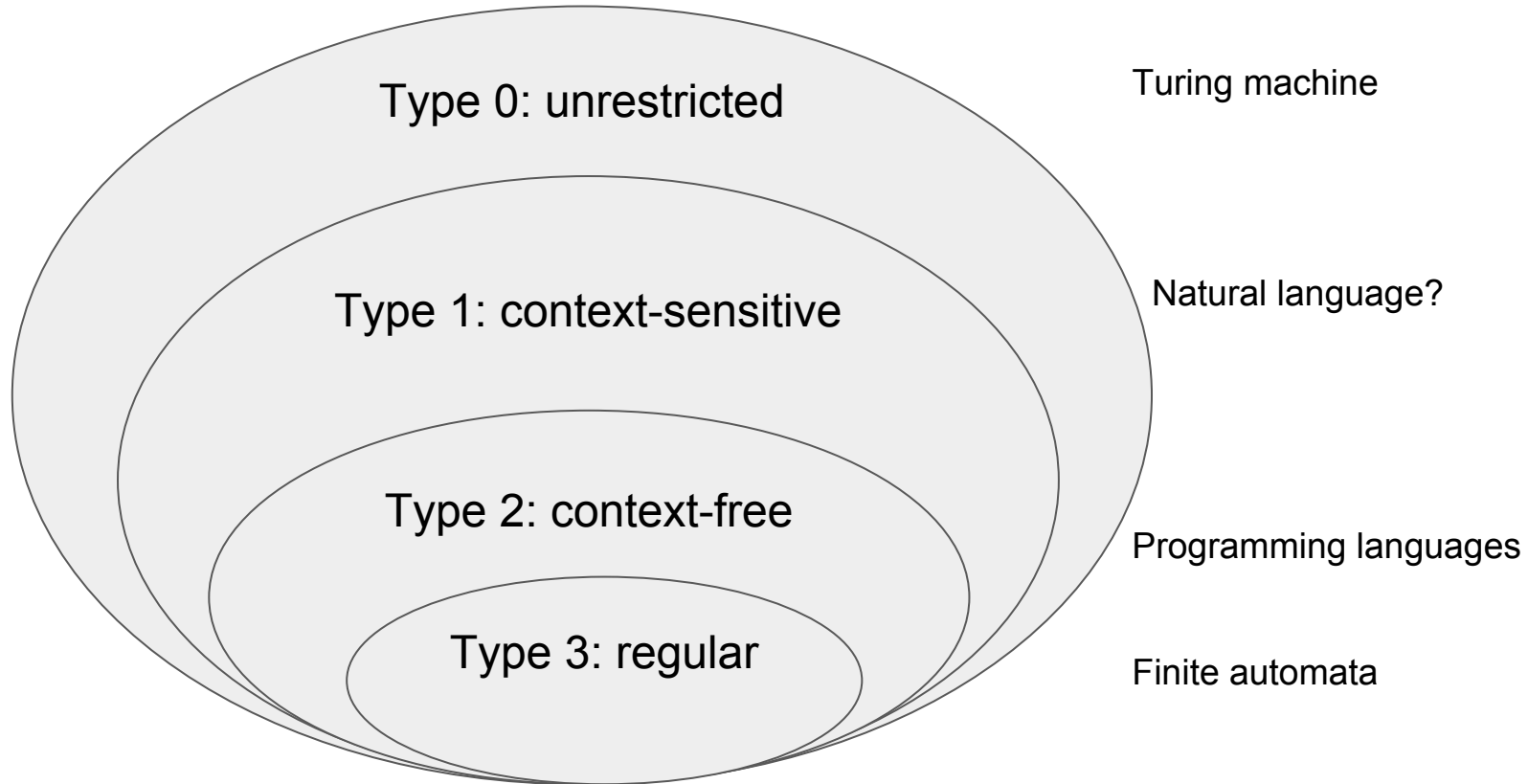


Computational grammars

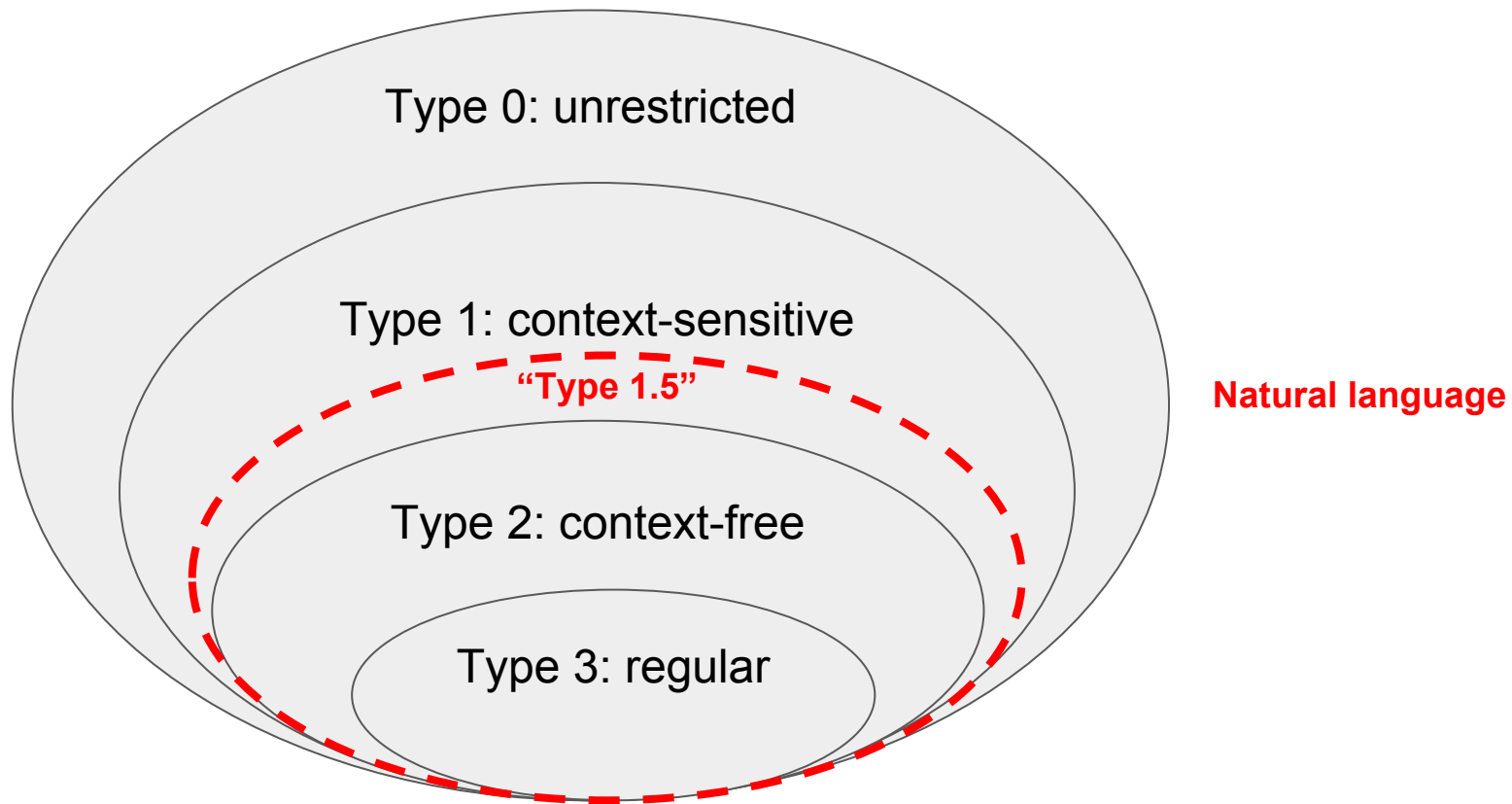
The Chomsky hierarchy (1956)



The Chomsky hierarchy in programming



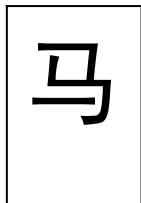
Mildly context-sensitive grammars (Joshi 1985)



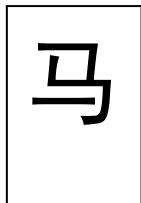
Variation in natural language: morphology



Variation in natural language: morphology



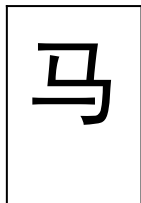
Variation in natural language: morphology



horse

horses

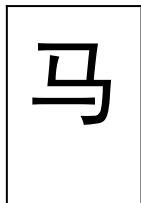
Variation in natural language: morphology



horse	horses
-------	--------

cheval	chevaux
masculine	

Variation in natural language: morphology



horse	horses
-------	--------

cheval	chevaux
masculine	

hevonen	hevoset
hevosen	hevosten
hevosta	hevosia
hevosena	hevosina
hevosiksi	hevosiksi
hevosessa	hevosissa
...	...

Morphology in context-free grammar

$N ::= \text{“马”}$

$N_{Sg} ::= \text{“horse”}$

$N_{Pl} ::= \text{“horses”}$

$N_{MascSg} ::= \text{“cheval”}$

$N_{MascPl} ::= \text{“chevaux”}$

$N_{SgNom} ::= \text{“hevonen”}$

$N_{SgGen} ::= \text{“hevosen”}$

$N_{SgPar} ::= \text{“hevosta”}$

...

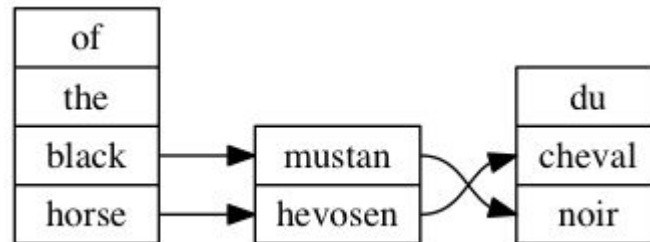
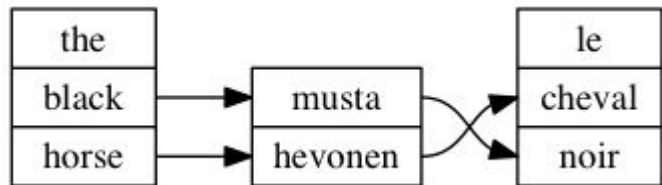
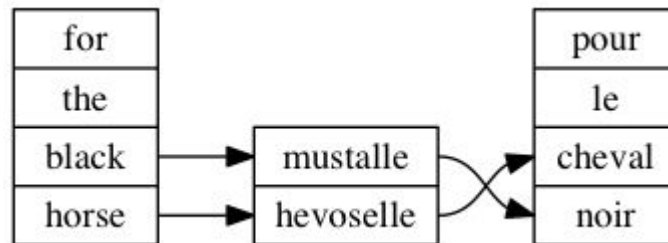
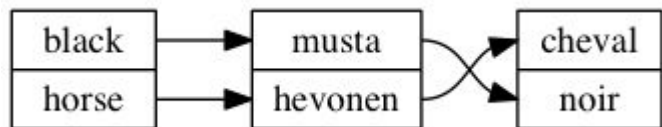
$N_{SgNom} ::= \text{“hevoset”}$

$N_{PlGen} ::= \text{“hevosten”}$

$N_{PlPar} ::= \text{“hevosia”}$

...

Variation in natural language: syntax



Syntax in context-free grammar

$$N ::= A N$$
$$N_{Sg} ::= A N_{Sg}$$
$$N_{Pl} ::= A N_{Pl}$$
$$N_{MascSg} ::= N_{MascSg} A_{MascSg}$$
$$N_{MascPl} ::= N_{MascPl} A_{MascPl}$$
$$N_{FemSg} ::= N_{FemSg} A_{FemSg}$$
$$N_{FemPl} ::= N_{FemPl} A_{FemPl}$$
$$N_{SgNom} ::= A_{SgNom} N_{SgNom}$$
$$N_{SgGen} ::= A_{SgGen} N_{SgGen}$$
$$N_{SgPar} ::= A_{SgPar} N_{SgPar}$$

...

$$N_{PlNom} ::= A_{PlNom} N_{PlNom}$$
$$N_{PlGen} ::= A_{PlGen} N_{PlGen}$$
$$N_{PlPar} ::= A_{PlPar} N_{PlPar}$$

...

Mildly context-sensitive: MCFG (Multiple Context-Free Grammar)

Context-free: strings

$N ::= \text{"horse"}$

Multiple context-free: tuples of strings

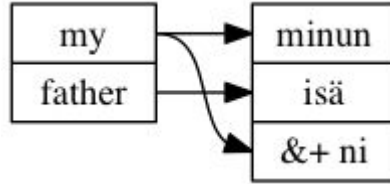
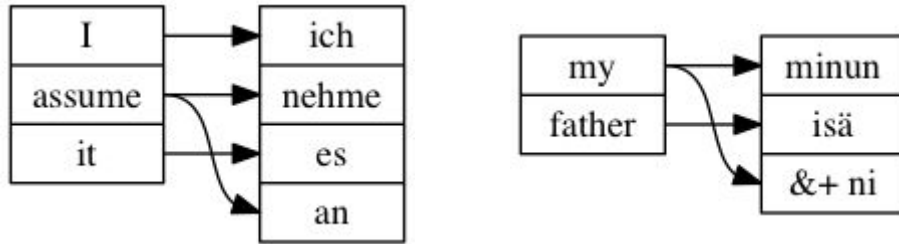
$N ::= \langle \text{"horse"}, \text{"horses"} \rangle$

$N ::= \langle \text{"hevonen"}, \text{"hevosen"}, \text{"hevosta"}, \text{"hevosenä"}, \dots \rangle$

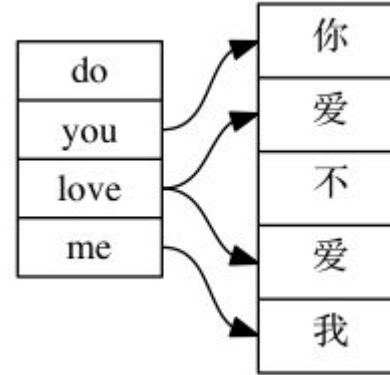
Morphological feature: index in a tuple

1 = singular, 2 = plural

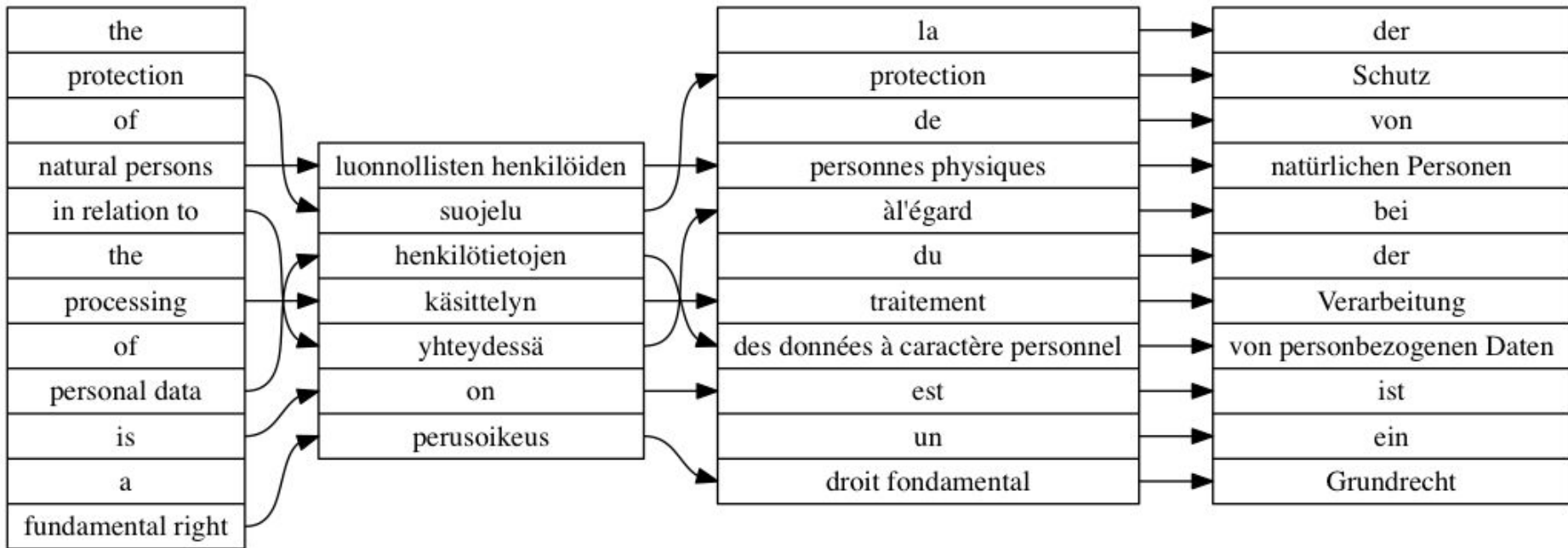
Discontinuous constituents



Reduplication



PMCFG (Parallel MCFG) is able to deal with these variations, unlike CFG.
No more expressive power has been needed to model the world's languages so far.

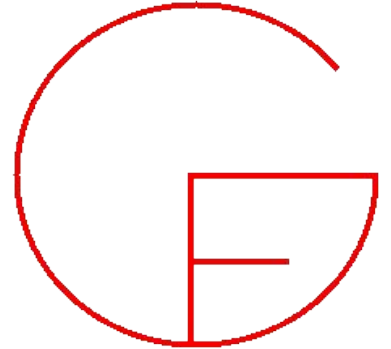


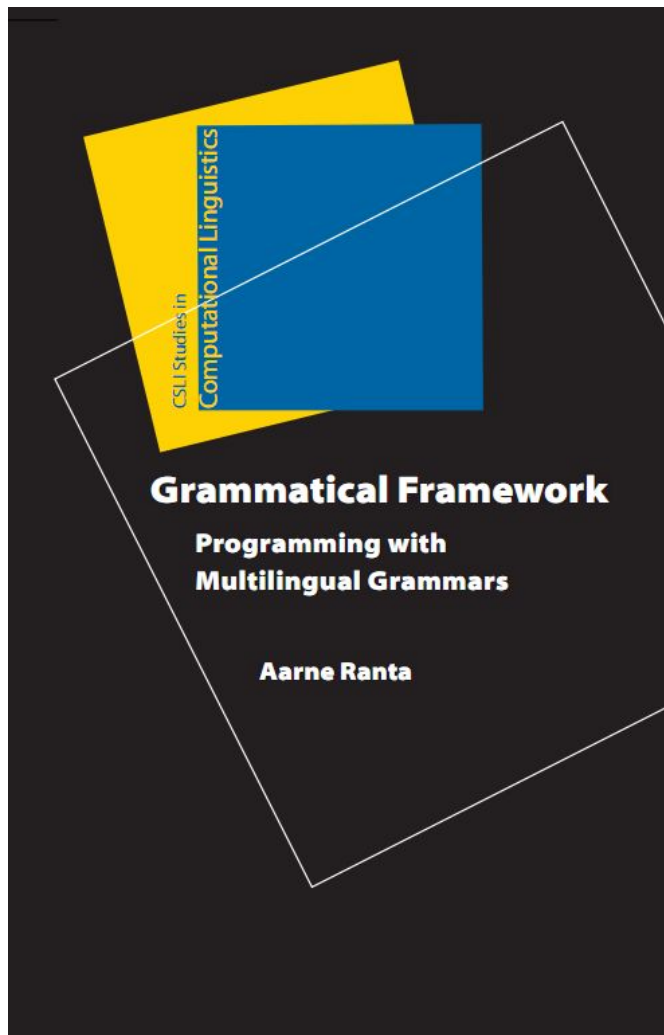
The first sentence of EU's GDPR legislation

<http://eur-lex.europa.eu/legal-content/FI/TXT/HTML/?uri=CELEX:32016R0679&from=EN>

GF, Grammatical Framework

*Mission: to formalize the grammars
of the world and make them
available for computer applications*



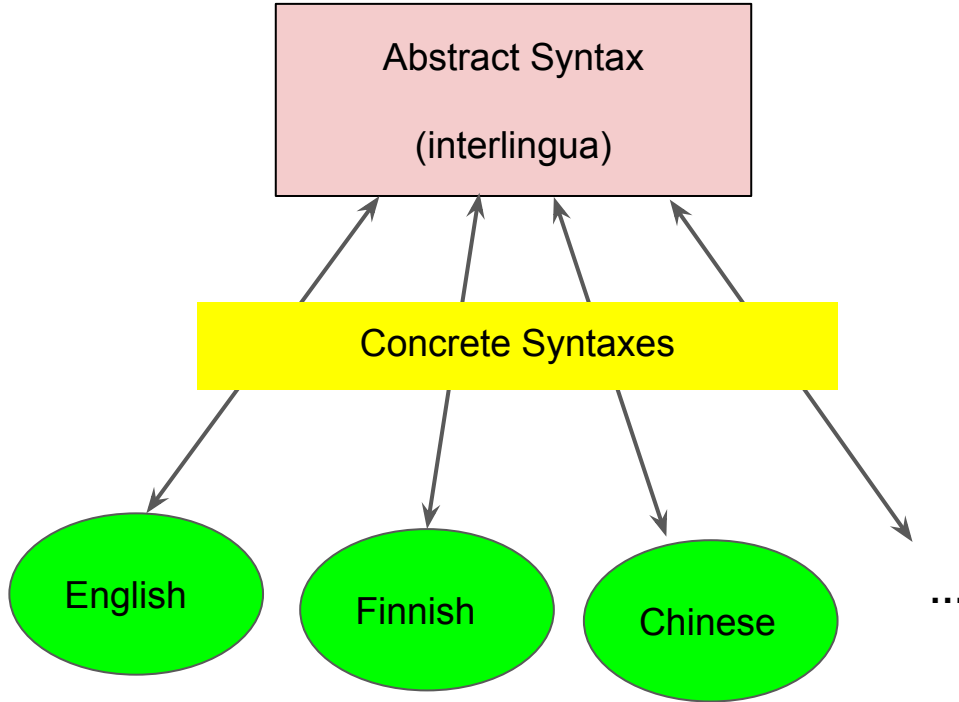


A. Ranta. *Grammatical Framework: Programming with Multilingual Grammars*, CSLI, Stanford, 2011.

Chinese translation by Prof. Yan Tian: 语法框架 为多种自然语言语法编程, Shanghai Jiao Tong University Press, 2014.

<http://www.grammaticalframework.org>

Multilingual grammars in GF



Constructive Type Theory

Parallel MCFG

Morphology in GF

```
cat
  N
fun
  Horse : N
```

```
lincat
  N = Str

lin
  Horse = “马”
```

```
lincat
  N =
    Number => Str

lin
  Horse = table {
    Sg => “horse” ;
    Pl => “horses”
  }
```

```
lincat
  N = {
    s : Number => Str ;
    g : Gender
  }

lin
  Horse = {
    s = table {
      Sg => “cheval” ;
      Pl => “chevaux”
    } ;
    g = Masc
  }
```

```
lincat
  N =
    Number => Case => Str

lin
  Horse = table {
    Sg => table {
      Nom => “hevonen” ;
      Gen => “hevosen” ;
      Part => “hevosta”
      ...
    }
    Pl => table {
      Nom => “hevoset” ;
      Gen => “hevosten” ;
      Part => “hevosia”
      ...
    }
  }
```

Syntax in GF

```
cat
  N ; A
fun
  Mod : A -> N -> N
```

```
lincat
  N = Str
  A = Str
```

```
lin
  Mod adj noun =
    adj ++ noun
```

```
lincat
  N =
    Number => Str
  A = Str
```

```
lin
  Mod adj noun =
    table {
      n => adj ++
        noun ! n
    }
```

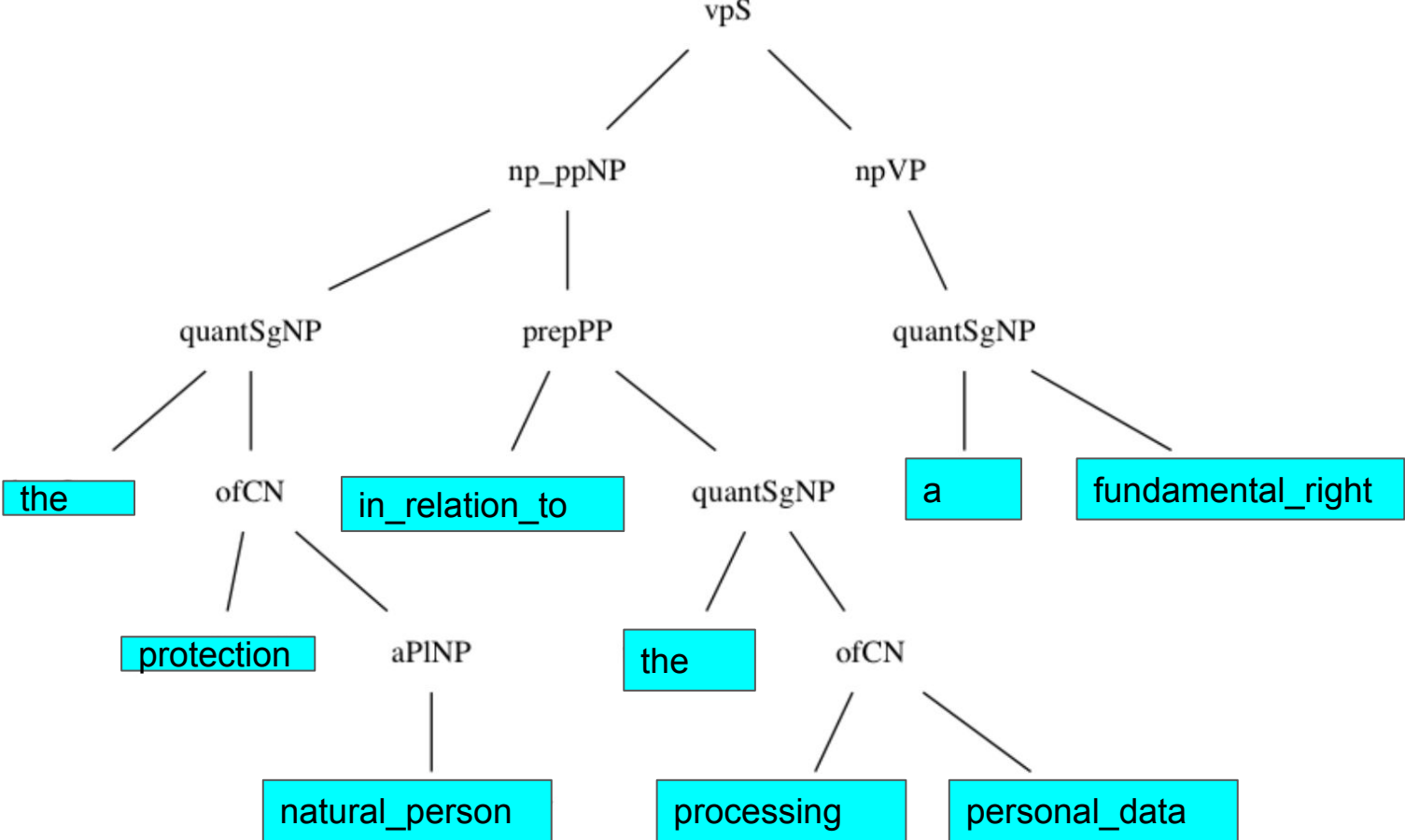
```
lincat
  N = {
    s : Number => Str ;
    g : Gender
  }
  A = Gender => Number
    => Str
```

```
lin
  Mod adj noun = {
    s = table {
      n =>
        noun.a ! n ++
        adj ! noun.g ! n
    } ;
    g = noun.g
  }
```

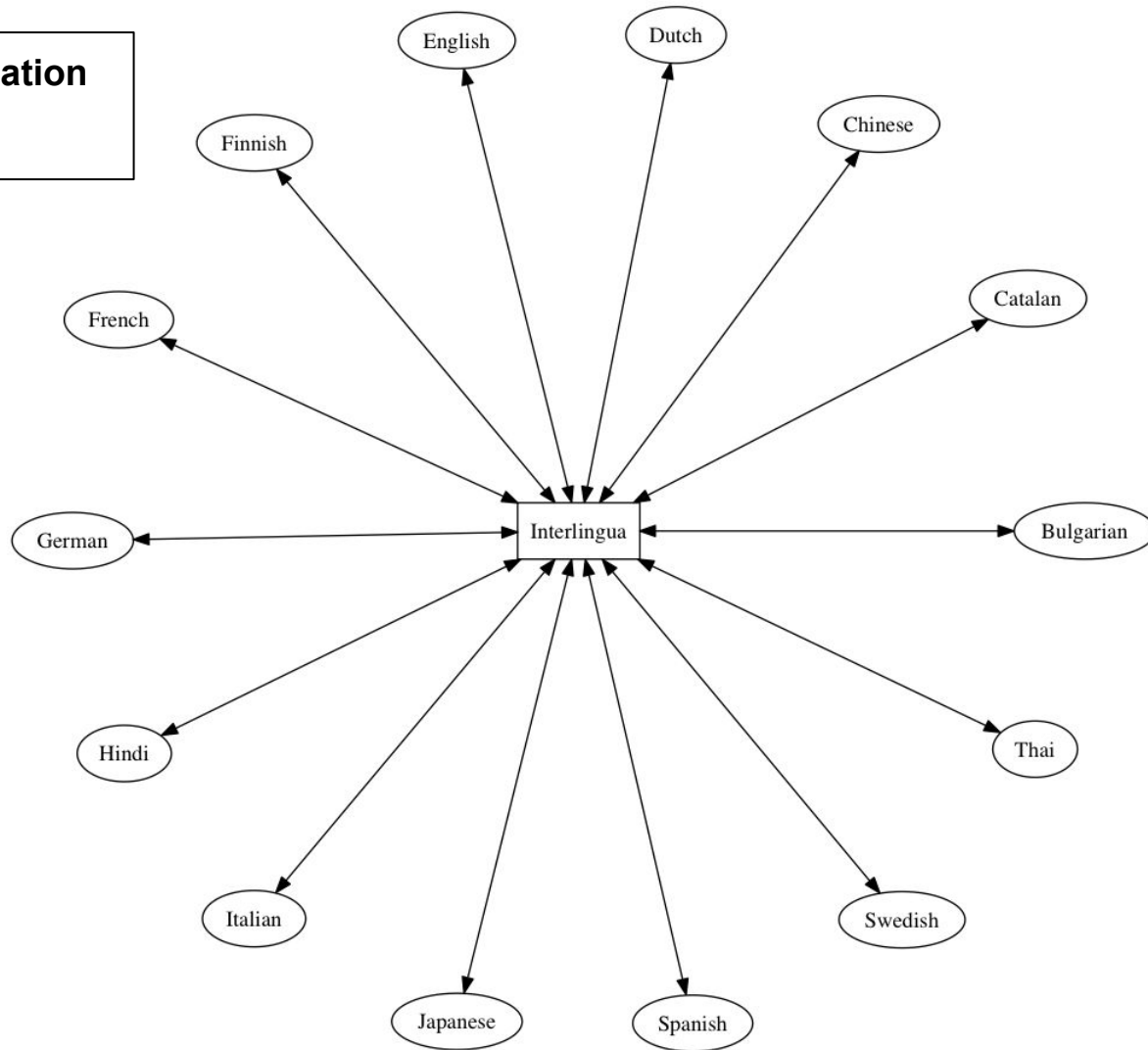
```
lincat
  N =
    Number => Case => Str
  A =
    Number => Case => Str
```

```
lin
  Mod adj noun =
    table {
      n => table {
        c =>
          adj ! n ! c ++
          noun ! n ! c
        }
    }
```

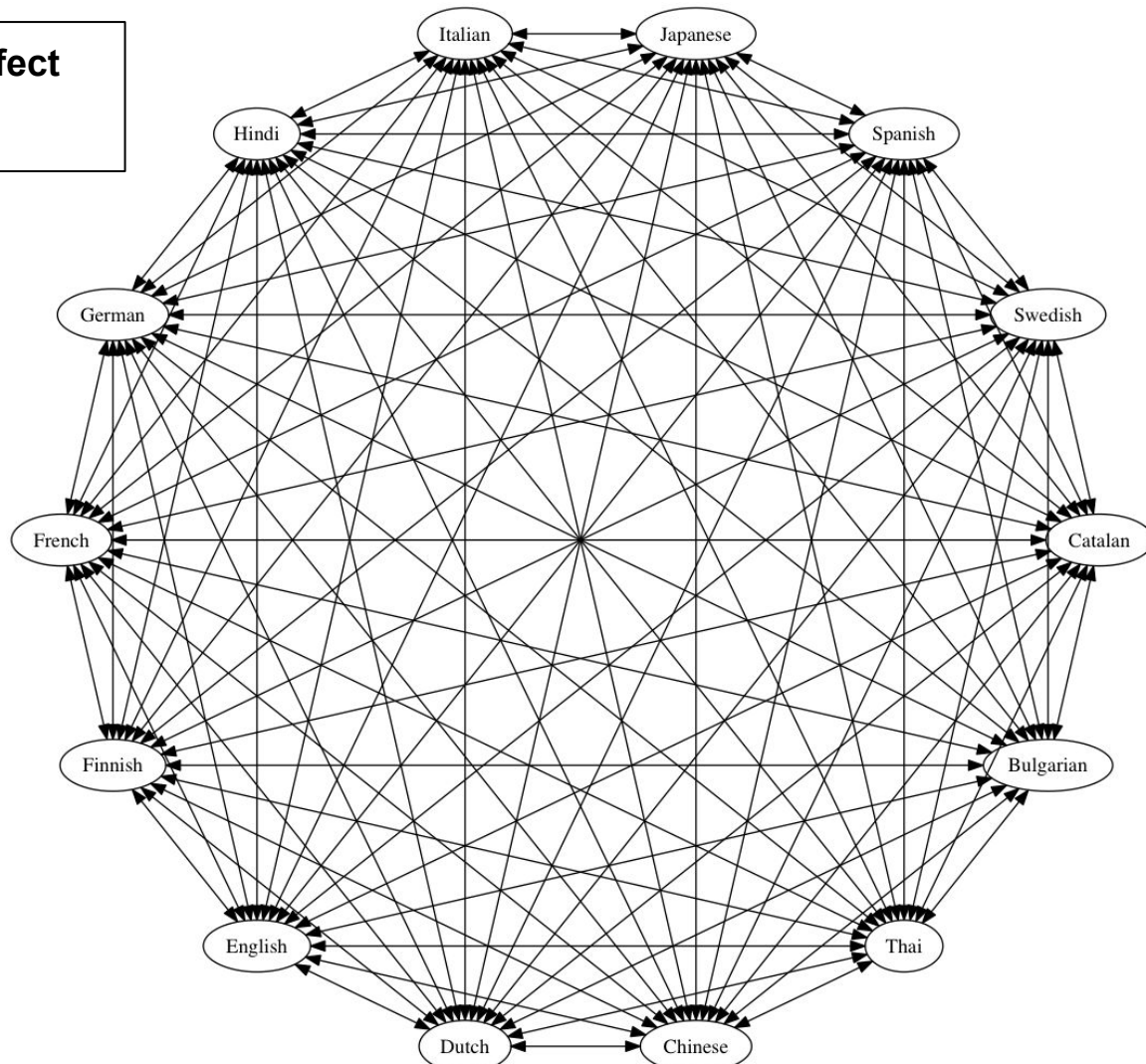
Abstract syntax of the first sentence of GDPR



Interlingual translation



Linear scale-up effect



Multilingual question answering

I want to go from
Pudong Airport to
Hongqiao Station.

I want to go from
Pudong Airport to
Hongqiao Station.

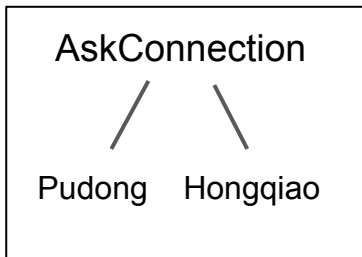
parsing

AskConnection

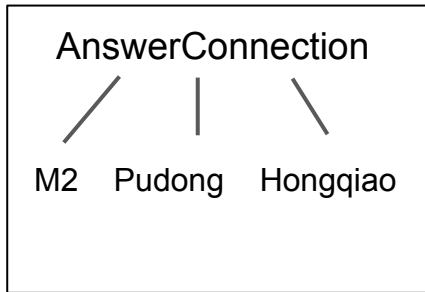
Pudong Hongqiao

I want to go from
Pudong Airport to
Hongqiao Station.

parsing

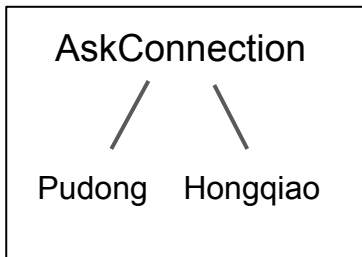


query engine

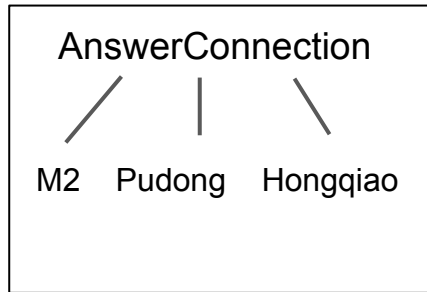


I want to go from
Pudong Airport to
Hongqiao Station.

parsing



query engine



linearization

Take Metro line 2
from Pudong Airport
to Hongqiao Station.

I want to go from
Pudong Airport to
Hongqiao Station.

parsing

AskConnection

Pudong Hongqiao

query engine

AnswerConnection

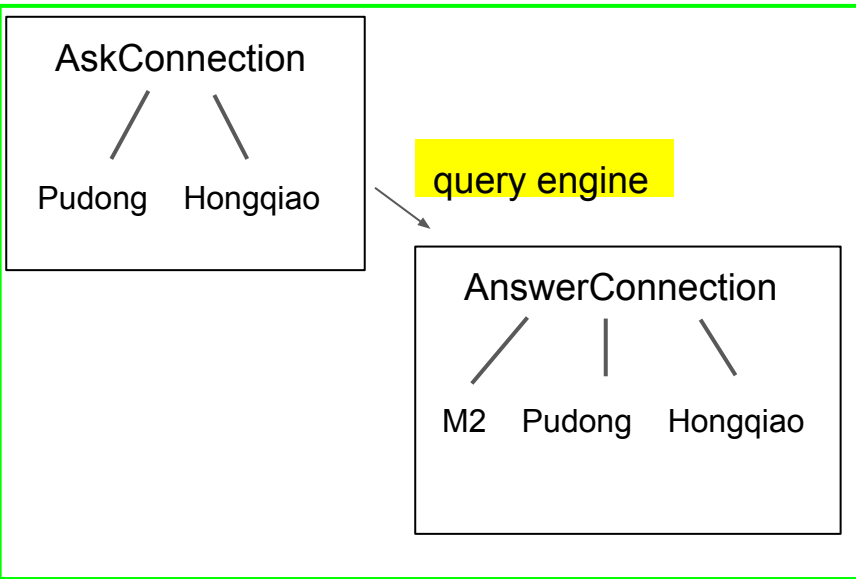
M2 Pudong Hongqiao

linearization

Take Metro line 2
from Pudong Airport
to Hongqiao Station.

从浦东机场到虹桥站怎么走？

parsing



query engine

linearization

在浦东坐2号地铁到虹桥站

Kuinka pääsee
Pudongin lentokentältä
Hongqiao-asemalle?

parsing

AskConnection

Pudong Hongqiao

query engine

AnswerConnection

M2 Pudong Hongqiao

linearization

Mene metrolla 2
Pudongin
lentokentältä
Hongqiao-asemalle.

The cost problem

grammars

PhD in linguistics

5 years of work

limited coverage

brittle

obsolete

statistics

BSc in computer science

5 weeks of waiting

wide coverage

robust

state of the art

grammars

PhD in linguistics

5 years of work

limited coverage

brittle

obsolete

Solution: software technology

Abstraction

- functional programming

Reuse

- grammars as software libraries

Tools

- grammar compiler (from high-level GF to low-level MCFG)
- portable runtime (for MCFG)

Abstraction:

the golden rule of functional programming

*Whenever you find yourself programming by
copy and paste, write a function instead!*

```
table {  
  Sg => table {  
    Nom => "hevonen" ;  
    Gen => "hevosen" ;  
    Part => "hevosta" ;  
    Ess => "hevosena" ;  
    Iness => "hevosessa" ;  
    ...  
  Pl => table {  
    Nom => "hevokset" ;  
    Gen => "hevosten" ;  
    Part => "hevosia" ;  
    Ess => "hevosina" ;  
    Iness => "hevosissa" ;  
    ...  
  }  
}
```

```
table {  
  Sg => table {  
    Nom => "lammas" ;  
    Gen => "lampaan" ;  
    Part => "lammasta" ;  
    Ess => "lampaana" ;  
    Iness => "lampaassa" ;  
    ...  
  Pl => table {  
    Nom => "lampaat" ;  
    Gen => "lampaiden" ;  
    Part => "lampaita" ;  
    Ess => "lampaina" ;  
    Iness => "lampaissa" ;  
    ...  
  }  
}
```

```
mkN "hevonen"
```

```
mkN "lammas"
```

```
mkN : Str -> N      -- a smart paradigm
```

Pres Simul Pos SDecl me juoksemme
Pres Simul Pos SQuest juoksemme me
Pres Simul Neg SDecl me emme juokse
Pres Simul Neg SQuest emmekö me juokse
Pres Anter Pos SDecl me olemme juosseet
Pres Anter Pos SQuest olemmeko me juosseet
Pres Anter Neg SDecl me emme ole juosseet
Pres Anter Neg SQuest emmekö me ole juosseet

Past Simul Pos SDecl me juoksimme
Past Simul Pos SQuest juoksimmeko me
Past Simul Neg SDecl me emme juosseet
Past Simul Neg SQuest emmekö me juosseet
Past Anter Pos SDecl me olimme juosseet
Past Anter Pos SQuest olimmeko me juosseet
Past Anter Neg SDecl me emme olleet juosseet
Past Anter Neg SQuest emmekö me olleet juosseet

Cond Simul Pos SDecl me juoksisimme
Cond Simul Pos SQuest juoksisimmeko me
Cond Simul Neg SDecl me emme juoksisi
Cond Simul Neg SQuest emmekö me juoksisi
Cond Anter Pos SDecl me olisimme juosseet
Cond Anter Pos SQuest olisimmeko me juosseet
Cond Anter Neg SDecl me emme olisi juosseet
Cond Anter Neg SQuest emmekö me olisi juosseet

Pot Simul Pos SDecl me juossemme
Pot Simul Pos SQuest juossemmeko me
Pot Simul Neg SDecl me emme juosse
Pot Simul Neg SQuest emmekö me juosse
Pot Anter Pos SDecl me lienemme juosseet
Pot Anter Pos SQuest lienemmeko me juosseet
Pot Anter Neg SDecl me emme liene juosseet
Pot Anter Neg SQuest emmekö me liene juosseet

```
mkCl we_NP (mkV "juosta")
```

```
mkCl : NP -> V -> Cl      -- a syntactic combination function
```


Reuse: the GF Resource Grammar Library

Morphology: complete paradigms

Basic syntax: 80 categories, 200 functions

Function word lexicon: 100 lemmas

Content word test lexicon: 300 lemmas

34 languages: Afrikaans, Basque, Bulgarian, Catalan, Chinese (simplified), Danish, Dutch, English, Estonian, Finnish, French, German, Greek, Hindi, Icelandic, Japanese, Italian, Latvian, Maltese, Mongolian, Nepali, Norwegian (bokmål), Norwegian (nynorsk), Persian, Polish, Portuguese, Punjabi, Romanian, Russian, Sindhi, Spanish, Swedish, Thai, Urdu

Community of 250+ members



GF Summer Schools



Gothenburg, Sweden, 2009



Barcelona, Catalonia, 2011



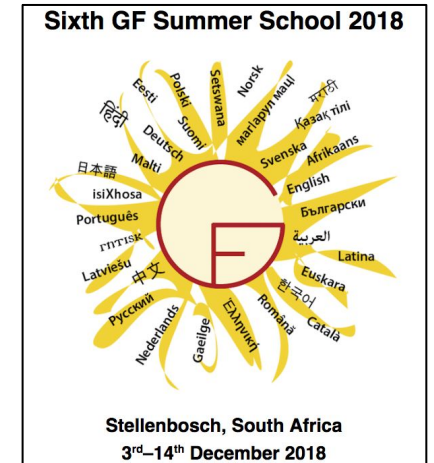
Riga, Latvia, 2017



Frauenchiemsee, Bavaria, 2013



Marsalforn, Gozo, Malta, 2015



Tools: GF as a programming language

methods	constructive type theory functional programming compiler construction static grammar checking
theory	abstract syntax MCFG (Multiple Context-Free Grammars) Resource Grammar Library
algorithms	fast statistical parsing normalization by evaluation almost compositional transformations grammar extraction from data
tools	grammar compiler runtime systems development tools

GF as theory and technology

language technology

- ❑ abstract syntax
- ❑ coding abstractions
- ❑ grammar libraries

linguistic theory

- ❑ universal grammar
- ❑ linguistic generalizations
- ❑ linguistic knowledge

**grammars, as widely
believed**

PhD in linguistics

5 years of work

limited coverage

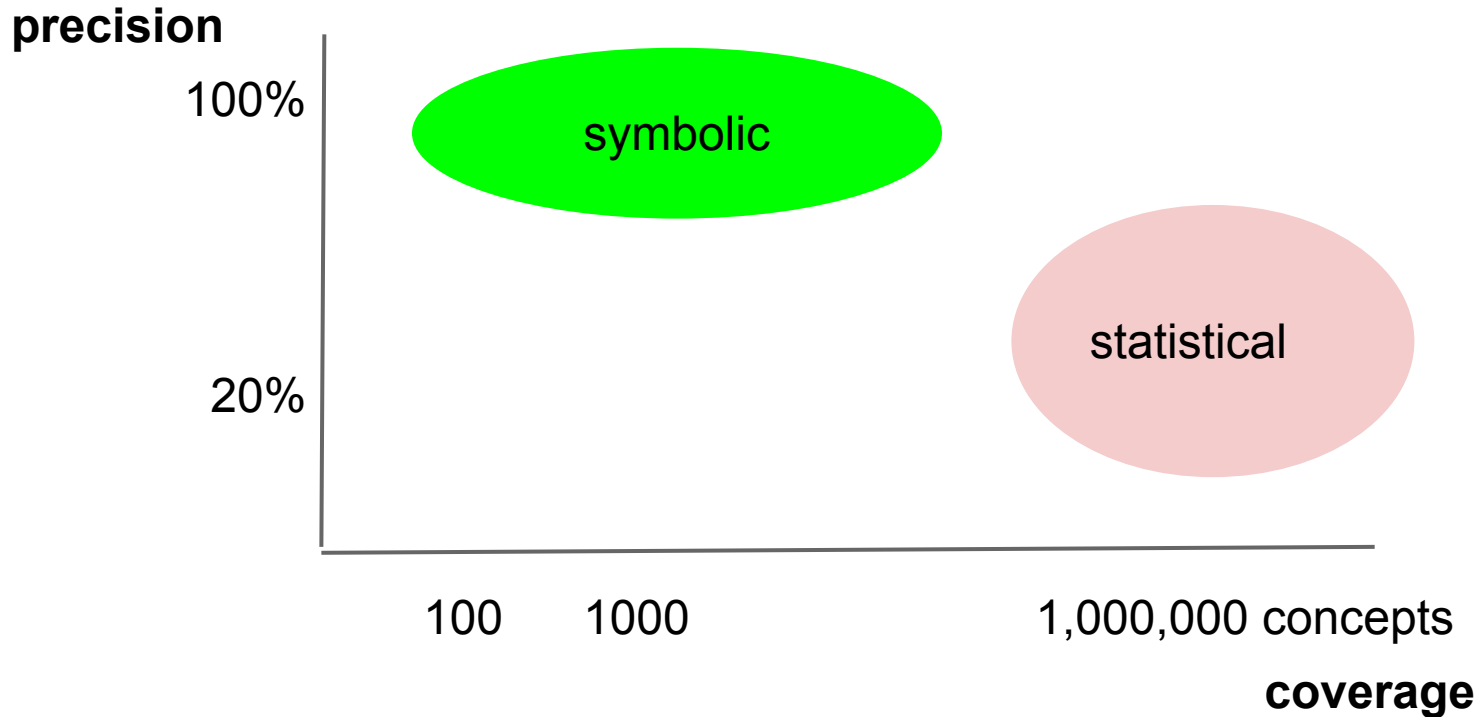
brittle

obsolete

grammars, as widely believed	grammars, in reality
PhD in linguistics	undergraduate to professor in linguistics, computer science, mathematics, philosophy...
5 years of work	work: 5 hours (simple application) to 5 months (new resource grammar)
limited coverage	scalable
brittle	precise and explainable
obsolete	in actual use, both academically and commercially

The coverage problem

The precision-coverage trade-off



Use cases

precision

100%

producer tasks

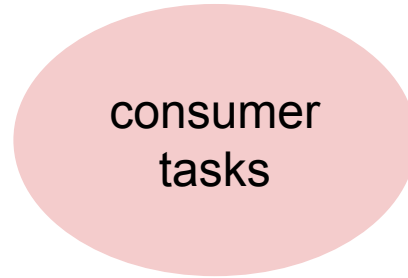
20%

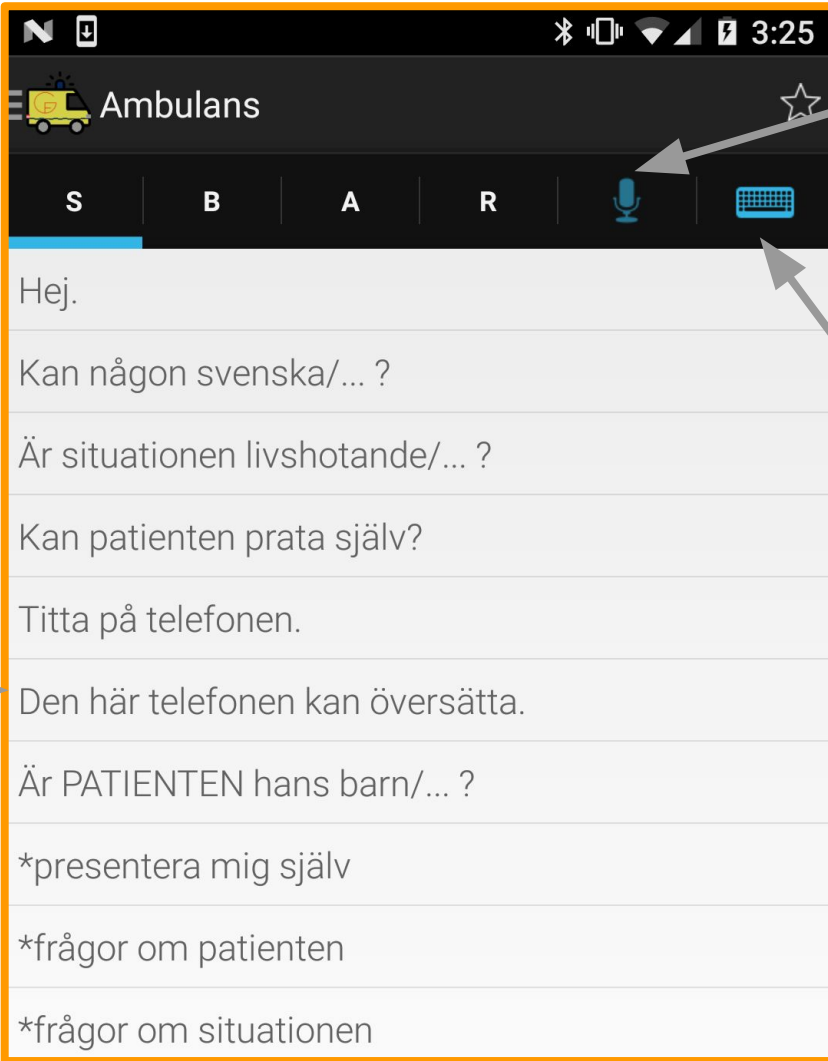
consumer
tasks

100 1000

1,000,000 concepts

coverage





Microphone
speak

Keyboard
write

Menu
build phrases
interactively



Har du haft binjurebarksvikt?

Onko sinulla ollut lisämunuaisen
vajaatoimintaa?



+ fråga: ja/nej (tidigare) ▾

+ du (man) ▾

+ binjurebarksvikt ▾

SU Ambulans

The image displays five overlapping smartphone screens showing the 'SU Ambulans' app interface. Each screen shows a list of medical questions in a different language, with the 'B' (B) tab selected. The languages shown are Swedish, Finnish, Romanian, and Persian. The app interface includes a top bar with the ambulance icon and 'SU Ambulans' text, and a bottom bar with tabs for 'S', 'B', 'A', and 'R'. The questions are listed in a scrollable view.

Swedish	Finnish	Romanian	Persian
Är PATIENTEN gravid/... ?	Onko Potilas raskaana/... ?	Pacientul este însărcinată/... ?	آیا مریض ح
Har vattnet gått/... ?	Onko vesi mennyt/... ?	Apa s- a rupt/... ?	آیا کیسه ی آ
Är PATIENTEN allergisk mot nötter/... ?	Onko Potilas allerginen pähkinöille/... ?	Pacientul are alergie la nuci/... ?	آیا مریض ح
Har PATIENTEN diabetes/... ?	Onko Potilaalla diabetes/... ?	Pacientul are diabet/... ?	آیا مریض ح
Har PATIENTEN feber/... ?	Onko Potilaalla kuumetta/... ?	Pacientul are febră/... ?	آیا مریض ح
Har PATIENTEN haft diabetes/... ?	Onko Potilaalla ollut diabetes/... ?	Pacientul a avut diabet/... ?	آیا مریض ح
Hur länge har PATIENTEN varit gravid/... ?	Kuinka kauan Potilas on ollut raskaana/... ?	De cât timp este Pacientul însărcinată/... ?	آیا مریض ح
Har PATIENTEN ont i hjärtat/... ?	Onko Potilaalla kipua sydämessä/... ?	Pacientul are dureri la inimă/... ?	آیا مریض ح
Har PATIENTEN ont i buken/... ?	Onko Potilaalla kipua vatsassa/... ?	Pacientul are dureri la abdomen/... ?	آیا مریض ح
Blöder PATIENTEN i buken/... ?	Tuleeko Potilaalta vatsasta verta/... ?	Abdomenul său sângerează/... ?	چند وقتہ مر
Vilka läkemedel tar PATIENTEN/... ?	Mitä lääkkeitä Potilas ottaa/... ?	Ce medicamente ia Pacientul/... ?	آیا مریض ح
*när och för hur länge sedan	*milloin ja mistä lähtien	*când și de cât timp	آیا مریض ح
Tar du alkohol/... ?	Nautitko sinä alkoholia/... ?	Dumneavoastră luați alcool/... ?	آیا مریض ح
Har du tagit alkohol/... ?	Oletko sinä nauttinut alkoholia/... ?	Dumneavoastră ați luat alcool/... ?	مریض کدام
Kan PATIENTEN gå/... ?	Voiko Potilas kävellä/... ?	Pacientul poate să meargă/... ?	*از کی و چه
Est-ce que vous prenez de l'alcool/... ?	Do you take alcohol/... ?	آیا مشروبات الکلی مصرف می کنی/... ?	آیا مشروبات الکلی مصرف کرده ای/... ?
Est-ce que vous avez pris de l'alcool/... ?	Have you taken alcohol/... ?	آیا مریض می تواند که راه برود/... ?	

(36) The main establishment of a controller in the Union should be the place of its central administration in the Union, unless the decisions on the purposes and means of the processing of personal data are taken in another establishment of the controller in the Union, in which case that other establishment should be considered to be the main establishment. The main establishment of a controller in the Union should be determined according to objective criteria and should imply the effective and real exercise of management activities determining the main

decisions of person processing. The main the Union

(36) Lo stabilimento principale di un titolare del trattamento nell'Unione dovrebbe essere il luogo in cui ha sede la sua amministrazione centrale nell'Unione, a meno che le decisioni sulle finalità e i mezzi del trattamento di dati personali siano adottate in un altro stabilimento del titolare del trattamento nell'Unione, nel qual caso tale altro stabilimento dovrebbe essere considerato lo stabilimento principale. Lo stabilimento principale di un titolare del trattamento nell'Unione dovrebbe essere determinato in base a criteri obiettivi e implicare l'effettivo e reale svolgimento di attività di gestione finalizzate alle principali decisioni sulle finalità e sui mezzi del trattamento nel quadro di un'organizzazione

(36) L'état que le du re L'état supp dans de m elles-princ dans sont t rester autori tout état d pas être traitem être consi déterminé

(36) El establecimiento principal de un responsable del tratamiento en la Unión debe ser el lugar de su administración central en la Unión, salvo que las decisiones relativas a los fines y medios del tratamiento de los datos personales se tomen en otro establecimiento del responsable en la Unión, en cuyo caso, ese otro establecimiento debe considerarse el establecimiento principal. El establecimiento principal de un responsable en la Unión debe determinarse en función de criterios objetivos y debe implicar el ejercicio efectivo y real de actividades de gestión que determinen las principales decisiones en cuanto a los fines y medios del tratamiento a través de modalidades estables. Dicho criterio no debe depender de si el tratamiento de los datos personales se realiza en dicho lugar. La presencia y utilización de medios técnicos y tecnologías para el tratamiento de datos personales o las actividades de tratamiento no constituyen, en sí mismas, establecimiento principal y no son, por lo tanto, criterios determinantes de un establecimiento principal. El establecimiento principal del encargado del tratamiento debe ser el lugar de su administración central en la Unión o, si careciese de administración central en la Unión, el lugar en el que se llevan a cabo las principales actividades de tratamiento en la Unión. En los casos que impliquen tanto al responsable como al encargado, la autoridad de control principal competente debe seguir siendo la autoridad de control del Estado miembro en el que el responsable tenga su establecimiento principal, pero la autoridad de control del encargado debe considerarse autoridad de control interesada y participar en el procedimiento de cooperación establecido en el presente Reglamento. En cualquier caso, las autoridades de control del Estado miembro o los Estados miembros en los que el encargado tenga uno o varios establecimientos no deben considerarse autoridades de control interesadas cuando el proyecto de decisión afecte únicamente al responsable. Cuando el tratamiento lo realice un grupo empresarial, el establecimiento principal de la empresa que ejerce el control debe considerarse el establecimiento principal del grupo empresarial, excepto cuando los fines y medios del tratamiento los determine otra empresa.

Verantwortlichen bezieht. Wird die Verarbeitung durch eine Unternehmensgruppe vorgenommen, so sollte die Hauptniederlassung des herrschenden Unternehmens als Hauptniederlassung der Unternehmensgruppe gelten, es sei denn, die Zwecke und Mittel der Verarbeitung werden von einem anderen Unternehmen festgelegt.

9 results for "subject"

English	German	French	Italian	Spanish
be subject to X	X unterliegen	être soumise à X	essere sottoposta a X	ser sujeta a X
be subject to X	X unterstehen	être placée sous les ordres de X	sottostare X	ser sujeta a X
be subject to X	X unterworfen sein	faire l' objet de X	essere soggiogata a X	ser sujeta a X
person subject to law	Rechtsunterworfene	justiciable	persona soggetta alla legge	persona sujeta al ley
subject	betroffen	concerné	soggetto	sujeto
subject	Gegenstand	sujet	soggetto	objeto
subject-matter	Gegenstand	objet	oggetto	objeto
subject to X	vorbehaltlich X	sous reserve de X	soggetto a X	siempre que se den X
subject to X	fallend unter X	soumis à X	soggetto a X	interesados a X

Some statistics about the GDPR lexicon

	abstract
functions	3525
atomic	3272
syntactic	139
construction	114

	English	German	French	Italian	Spanish
word tokens	55186	54903	62198	55296	57383
word types	2625	4153	3206	3520	3498
lemma types	2555	3053	2467	2689	2478
multiword functions	590	227	574	559	594
multiwords in corpus	8-13%	3-10%	10-21%	10-20%	11-21%

Is this the largest multilingual corpus ever analysed at this level of precision?


```

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 Σ -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 $A \rightarrow B$.

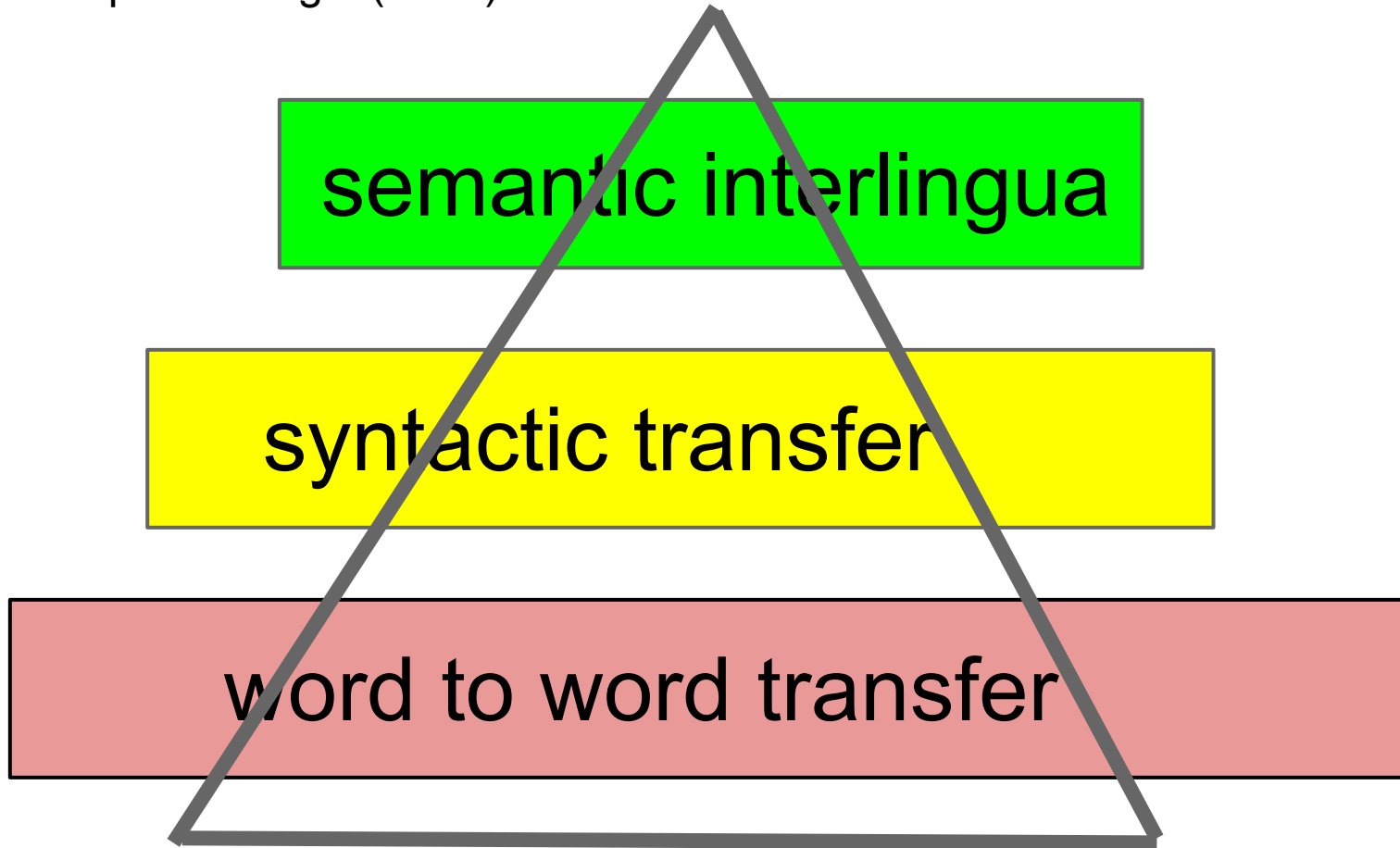
Lemme: Pour tout type A , l'identité, $1_A := \lambda x:A x$: est une équivalence.

Démonstration: Pour tout $y : A$, soit $\{y\}_A := \{x : A \mid x = y\}$ sa fibre par rapport de 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

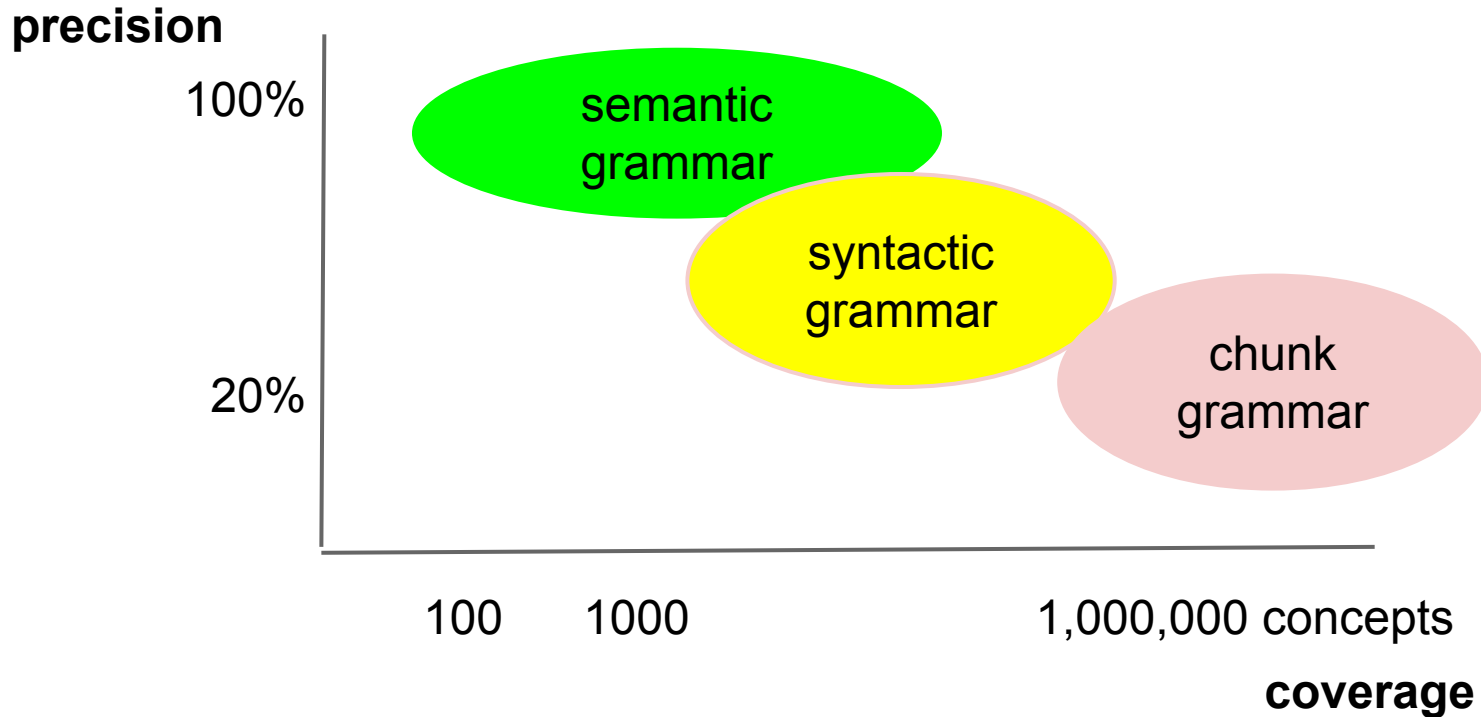
$$(x, z) = y$$

. Donc, pour les $y : A$, nous pouvons appliquer Σ -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

The Vauquois triangle (1968)



Graceful degradation





Wide Coverage Translation Demo

English

Clear



Finnish

Colors

Translate

Grammars...

Is there a train to the airport

The vice dean is in love with the student.

Little boy eat big snake.

Enter text to translate above

Pääseekö lentokentälle junalla?

Pahedekaani on rakkaudessa opiskelijan kanssa.

Pieni poika syökää suuri käärme.

Try Google Translate

<http://cloud.grammaticalframework.org/wc.html>

The GF Offline Translator

<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.



Grammars vs. statistics

Meaning

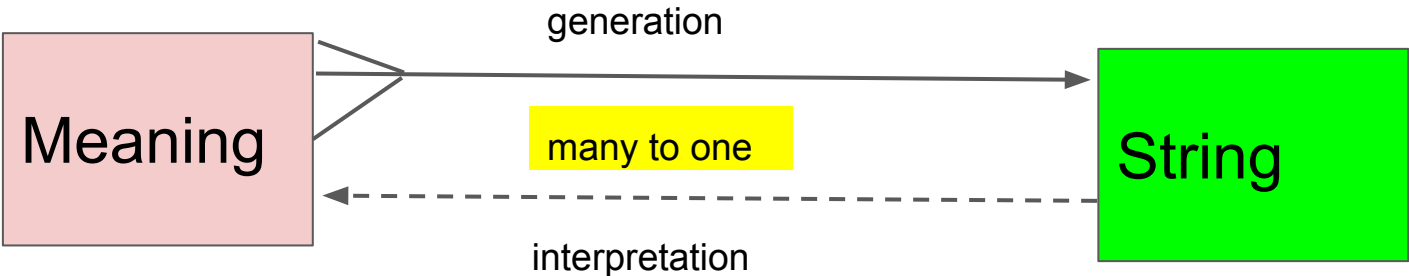
generation

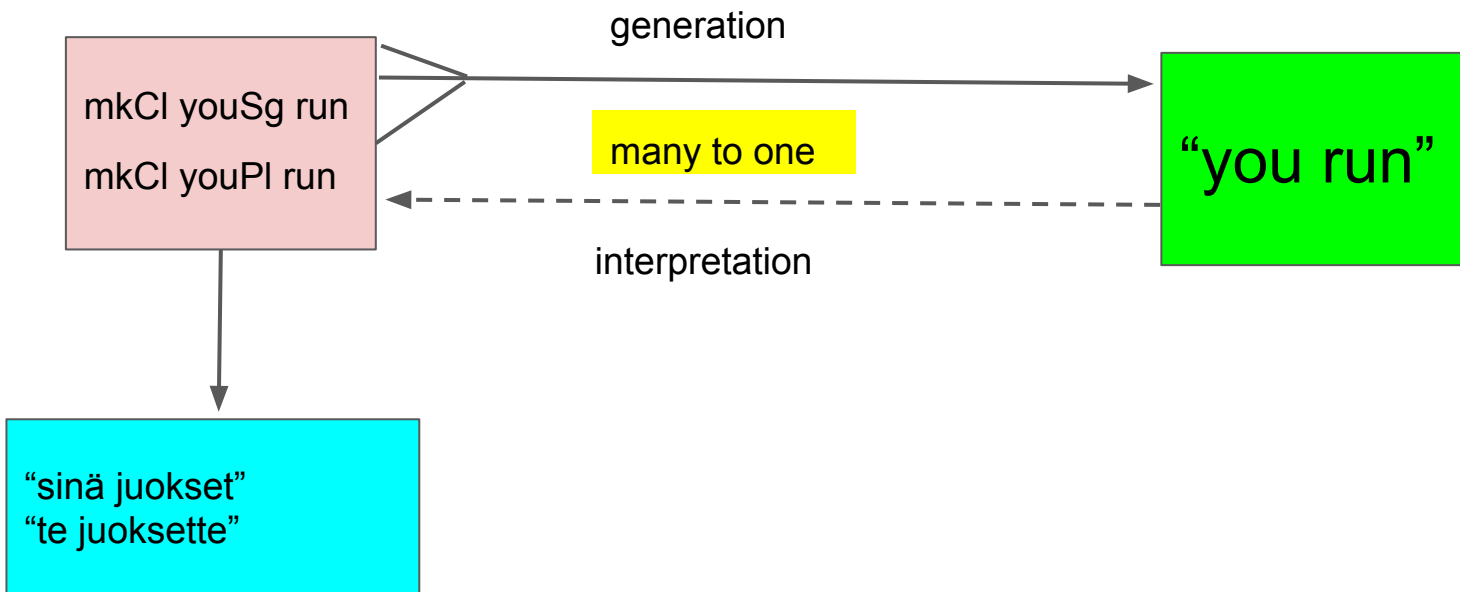


String

interpretation







Meaning

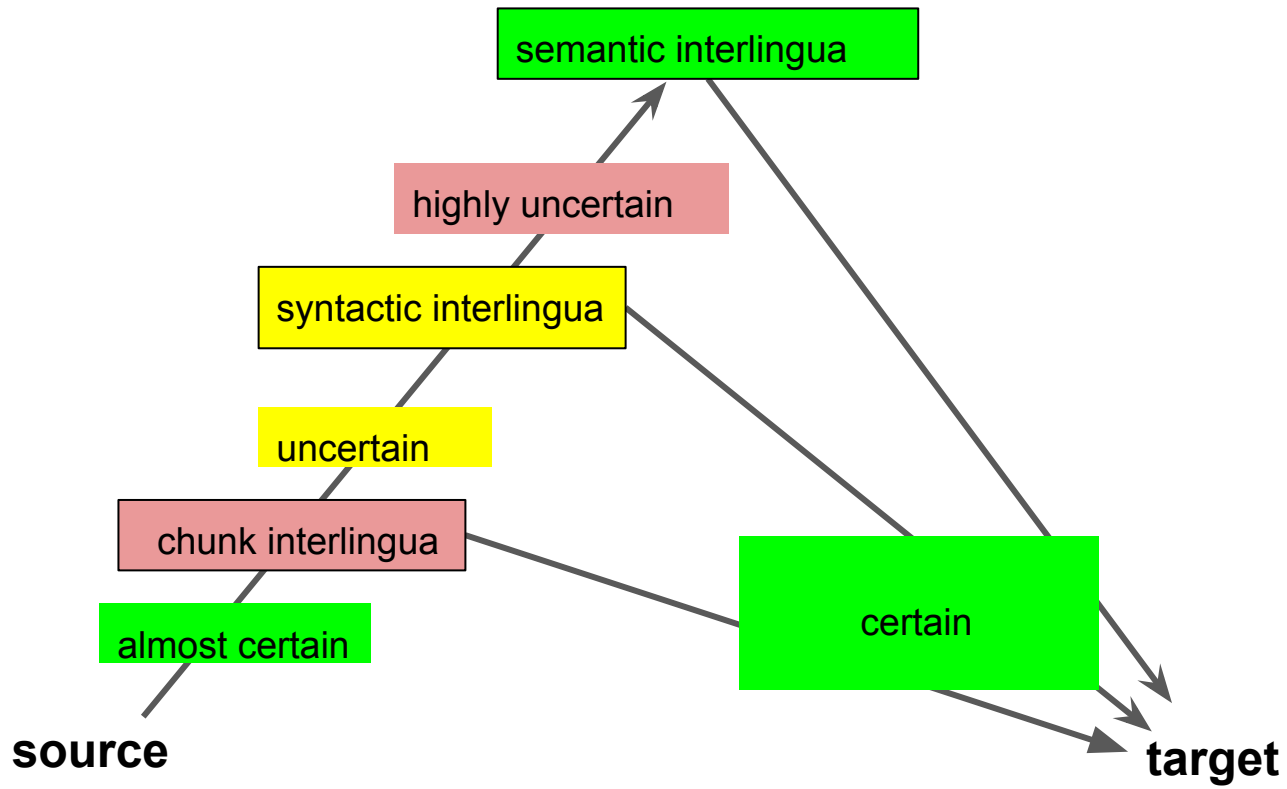
certain, rule-based



Source
string



uncertain: probabilistic



Take home

Grammars have a future

Enable language processing that is

- explainable
- controllable

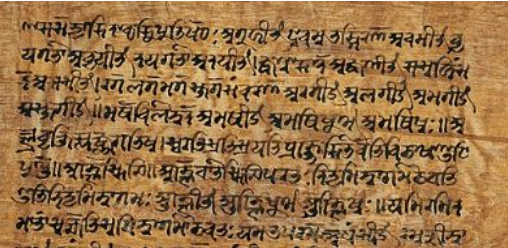
Practically useful for precision tasks

Give insights to the similarities and differences of languages

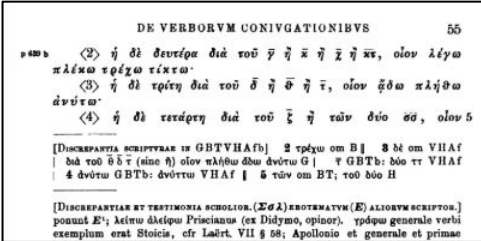
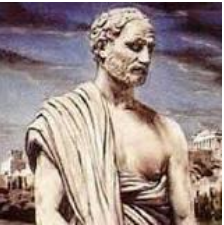
Should be taken seriously as software, with

- powerful programming tools
- open-source libraries
- developer community

Grammars have a past



Panini, 6-4 century BC



Dionysius Thrax, 170-80 BC

Präteritum.					
Indikativ.					
Sg. 1. skaut	fell	safnafa	suaſfa	stýrfa	vakpa
2. skautz	felt	safnaper	suaſper	stýrper,	vakper
§ 524, 2				-ir	
3. skaut	fell	safnape	suaſpe	stýrpe, -i	vakpe
Pl. 1. skutom,	feilom	sofnopom,	suaſpom,	stýrpom,	vakpom,
-um		safnadom	suaſdom	-um	vakpum
2. skutop,-ot,	fellop,	sofnopoj,-ot,	suaſpoj,-ot,	stýrpoj,	vakpoj,-ot,
-ur	-ot,-ur	safnador	suaſdur	-ot,-ur	vakpur
3. skuto,-u	fello	sofnopo,-o	suaſpo,-o	stýrpo,-u	vakpo,-u
		safnado	suaſdo		vakpu



Adolf Noreen, 1854-1925

Formalizing grammars and making them available for computer applications is a way to build systems that can constantly improve and last for centuries.

