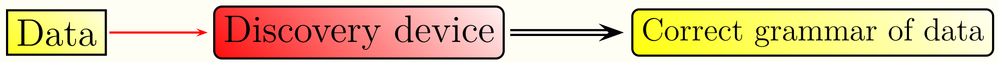
Language and the Mind: Encounters in the Mind Fields

John Goldsmith

April 23, 2014

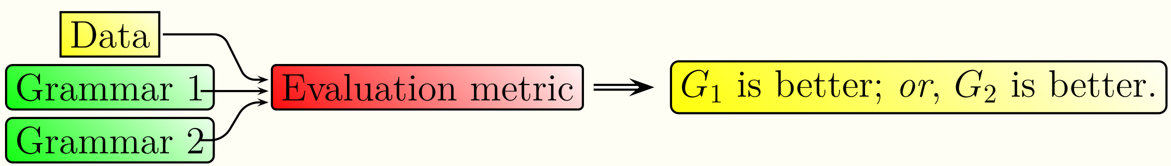

Frank and Ernest
PSYCHIATRY
FIRST WE'LL ANALYZE YOUR INFERIORITY COMPLEX THROUGH SOME WORD ASSOCIATION-- THE FIRST WORD: "MORPHOPHONEMICS".
www.frankandernest.com
2-8 THAVES

1. Strongest, best option:



2. Next best option:

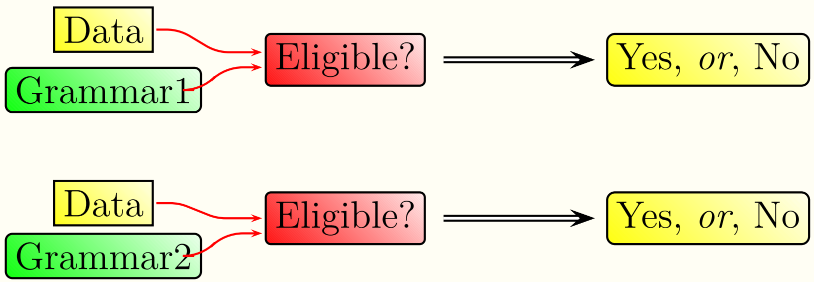
3. Fallback position:



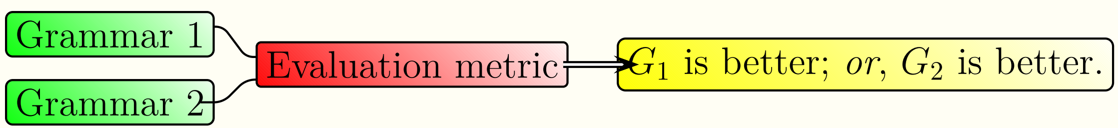
Chomsky's vision of Generative Grammar (1955)

Generative position: a special case of Option 3

First, test grammars' eligibility:



If both grammars are eligible:



Three central questions:

1. **Where do hypotheses come from?** Answer: As far as Linguistic Theory goes, that's none of your business. Ideas come from wherever they come from. As far as individual grammars go, hypotheses may come from anywhere, but mostly they come from looking at what linguists have said about other languages.

2. **How do we determine the extent to which data support a hypothesis?** Generative theory has no answer to this.

3. **How do we determine the goodness of a theory, independent of data?** Formal simplicity, but we have not yet found the right way to calculate this.

Machine learning:
Back to Option 1

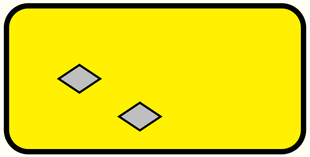
Data $\rightarrow$ Discovery device; $\mathcal{G}$ $\rightarrow$ Best grammar in $\mathcal{G}$ of data

Generative grammar and Machine learning agree:

- Growing the space of grammars when needed is a good thing.
- Shrinking the space of grammars when we jettison unnecessary possibilities is a good thing.

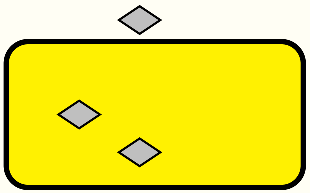
Machine learning:

- A linguistic theory requires a method to *find* the grammar (within the given hypothesis space) that best accounts for the data.



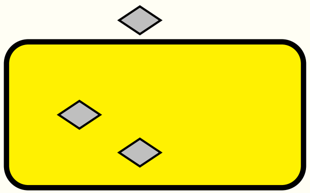
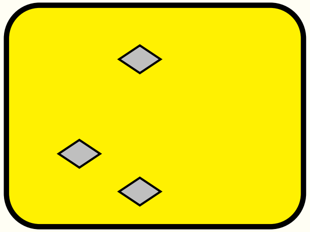
Two languages, two grammars, and a Universal Grammar

The expected evolution of generative theory



A grammar is found that lies outside of Universal Grammar.

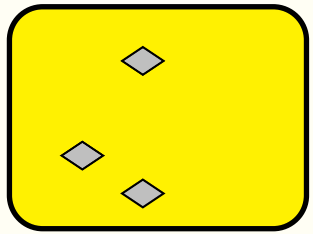
The expected evolution of generative theory

A grammar is found that lies outside of Universal Grammar.

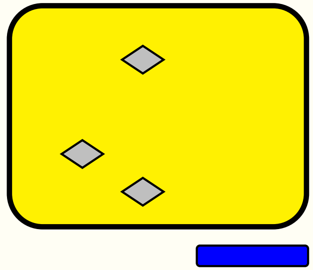
Univeral Grammar is expanded, on empirical grounds.

The expected evolution of generative theory



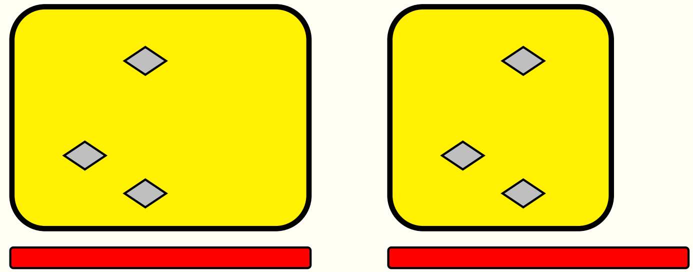
Revised Universal Grammar.

The expected evolution of generative theory



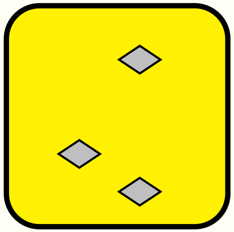
Unused space in Universal Grammar is noticed.

The expected evolution of generative theory



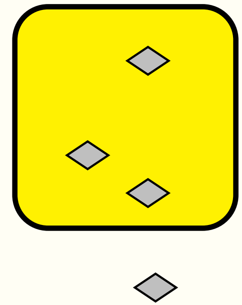
Universal Grammar
is shrunk.

The expected evolution of generative theory



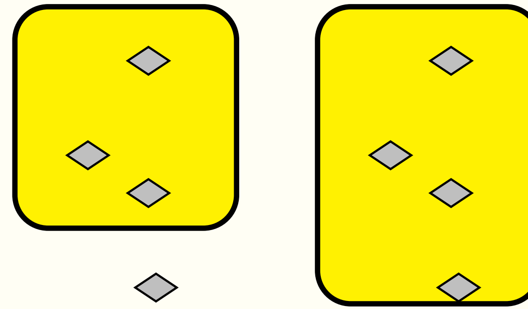
Revised
Universal Grammar.

The expected evolution of generative theory



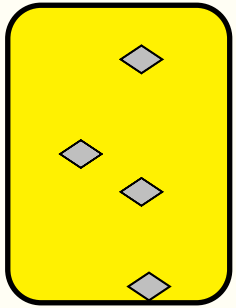
A grammar is found that
lies outside of
Universal Grammar.

The expected evolution of generative theory



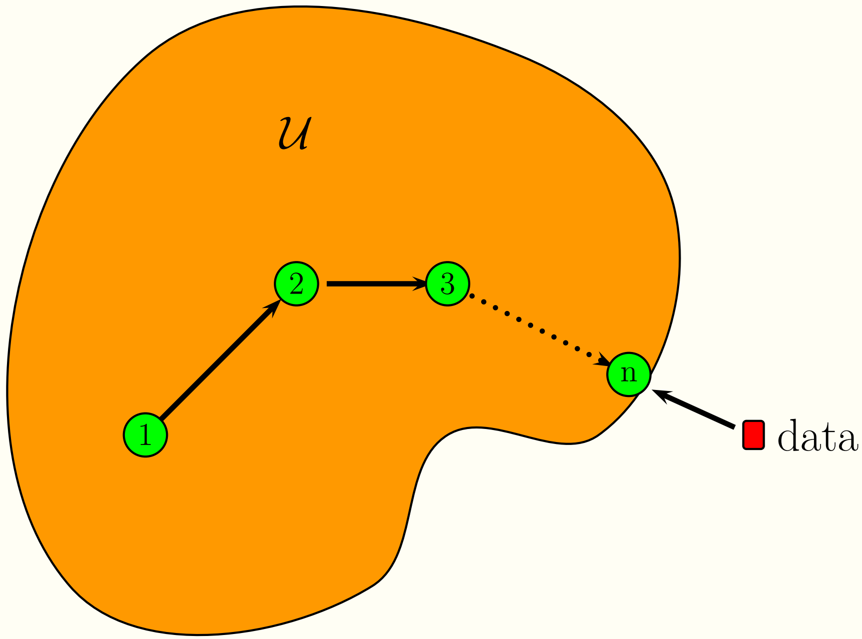
Univeral Grammar is
expanded, on
empirical grounds.

The expected evolution of generative theory



Revised
Universal
Grammar.

The expected evolution of generative theory



Find the grammar within the Universe $\mathcal{U}$ of Universal Grammar which best models the data.

Machine learning world

Example 1: Word learning

Input: A million words without spaces, including:

TheFultonCountyGrandJurysaidFridayaninvestigationo fAtlanta'srecentprimaryelectionproducednoevidenceth...

Desired output:

The Fulton County Grand Jury said Friday an investigation of Atlanta's recent primary election produced no evidence that any irregularities took place.

Actual output:

The F ult on County Gr and Ju ry said Fri day an investig ationof Atlan ta 's recent primary election produc ed no evidence that any ir regular ities took place.

Iteration number 1

piece	count
th	127,717
he	119,592
$\boxed{\textbf{in}}$	86,893
er	81,899
$\boxed{\textbf{an}}$	72,154
re	67,753
$\boxed{\textbf{on}}$	61,275
es	59,943
en	55,763
$\boxed{\textbf{at}}$	54,216
ed	52,893
nt	52,761
st	52,307
nd	50,504
ti	50,253
$\boxed{\textbf{to}}$	48,233
$\boxed{\textbf{or}}$	47,391

Iteration number 1

piece	count
th	127,717
he	119,592
in	86,893
er	81,899
an	72,154
re	67,753
on	
es	59,943
en	55,763
at	54,216
ed	52,893
nt	52,761
st	52,307
nd	50,504
ti	50,253
to	48,233
or	47,391

Iteration number 10

piece	count
In	2,355
vi	2,247
some	2,169
who	2,155
ical	2,130
He	2,119
ure	2,102
ance	2,085
ty	2,061
edthe	2,061
sel	2,053
its	2,053
more	2,034
form	2,023
fac	2,009
act	2,007
cont	1,987

Iteration number 1

piece	count
th	127,717
he	119,592
in	86,893
er	81,899
an	72,154
re	67,753
on	
es	59,943
en	55,763
at	54,216
ed	52,893
nt	52,761
st	52,307
nd	50,504
ti	50,253
to	48,233
or	47,391

Iteration number 10

piece	count
In	2,355
vi	2,247
some	2,169
who	2,155
ical	2,130
He	2,119
ure	2,102
ance	2,085
ty	2,061
edthe	2,061
sel	2,053
its	2,053
more	2,034
form	2,023
fac	2,009
act	2,007
cont	1,987

Iteration number 1

piece	count
th	127,717
he	119,592
in	86,893
er	81,899
an	72,154
re	67,753
on	
es	59,943
en	55,763
at	54,216
ed	52,893
nt	52,761
st	52,307
nd	50,504
ti	50,253
to	48,233
or	47,391

Iteration number 10

piece	count
In	2,355
vi	2,247
some	2,169
who	2,155
ical	2,130
He	2,119
ure	2,102
ance	2,085
ty	2,061
edthe	2,061
sel	2,053
its	2,053
more	2,034
form	2,023
fac	2,009
act	2,007
cont	1,987

Iteration number 399

piece	count
divided	22
minimal	21
ender	21
Baltimore	21
Memor	21
fever	21
WestBerlin	21
thickness	21
contains	21
backin	21
choiceof	21
attentiontothe	21
itthe	21
sophisticated	21
sector	21
jungle	21
Mid	21

Iteration number 1

piece	count
th	127,717
he	119,592
in	86,893
er	81,899
an	72,154
re	67,753
on	
es	59,943
en	55,763
at	54,216
ed	52,893
nt	52,761
st	52,307
nd	50,504
ti	50,253
to	48,233
or	47,391

Iteration number 10

piece	count
In	2,355
vi	2,247
some	2,169
who	2,155
ical	2,130
He	2,119
ure	2,102
ance	2,085
ty	2,061
edthe	2,061
sel	2,053
its	2,053
more	2,034
form	2,023
fac	2,009
act	2,007
cont	1,987

Iteration number 399

piece	count
divided	22
minimal	21
ender	21
Baltimore	21
Memor	21
fever	21
WestBerlin	21
thickness	21
contains	21
$\boxed{\text{backin}}$	21
$\boxed{\text{choiceof}}$	21
$\boxed{\text{attentiontothe}}$	21
$\boxed{\text{itthe}}$	21
sophisticated	21
sector	21
jungle	21
Mid	21

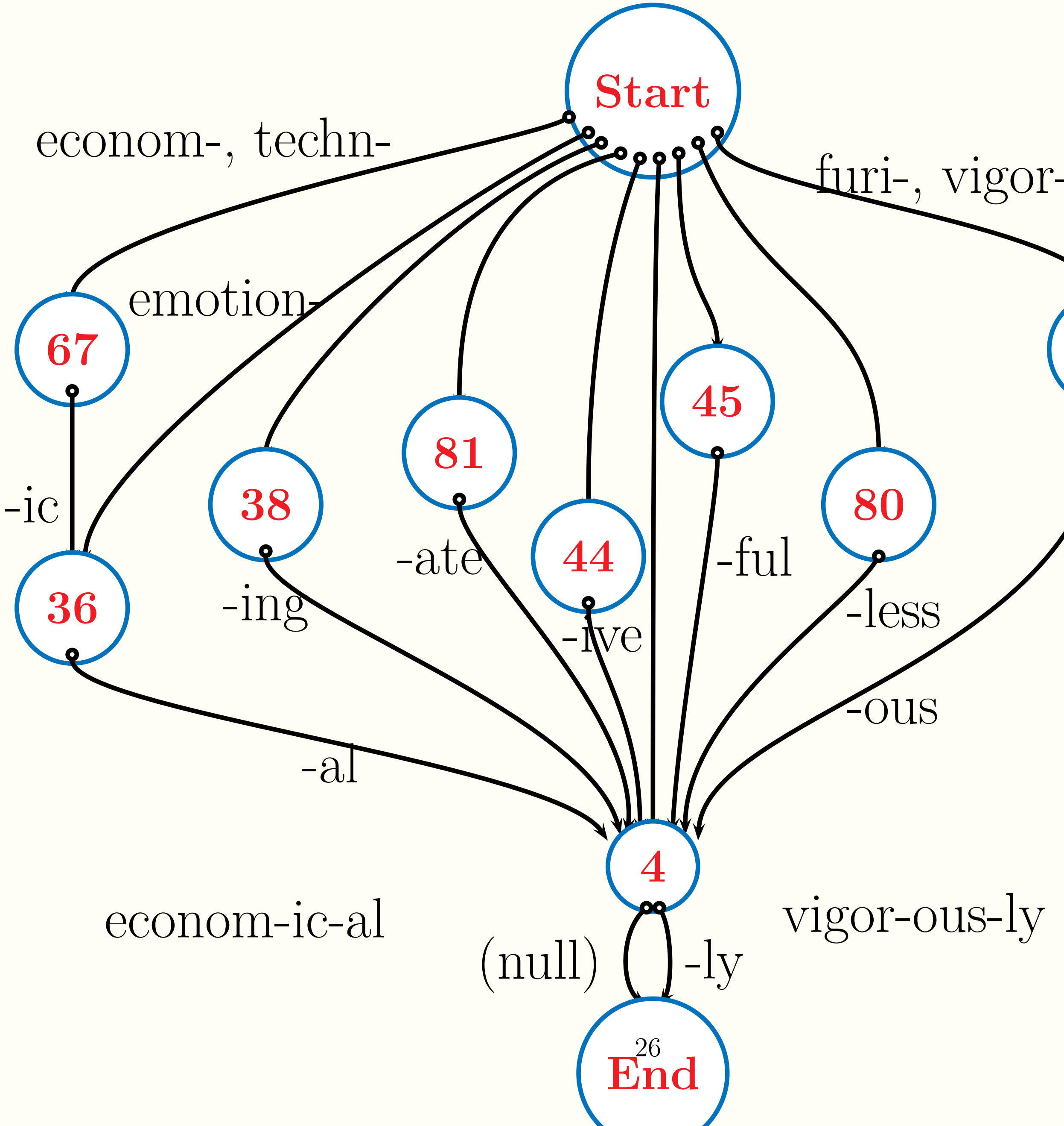
Example 2: Morphology learning

NULL-s	**accomodation**	**accomodations**		
NULL-'s	aunt	aunt's		
NULL-ed-ing-s	account	accounted	accounting	accounts
NULL-s-'s	afternoon	afternoons	afternoon's	
e-ed-ing-es	accuse	accused	accusing	accuses
ies-y	ability	abilities		
NULL-al-s	addition	additional	additions	
NULL-ped-ping-s	drop	dropped	dropping	drops
ied-ies-y-ying	tried	tries	try	trying

guerrilla	camera	suburb	electronic
athletic	poetic	plastic	characteristic
hundred	fluid	field	thousand
ground	method	neighborhood	standard
toward	afterward	hazard	cloud
voice	price	device	service

NULL-s	accomodation	accomodations		
NULL-ly	according	accordingly		
NULL-ed-ing-s	**account**	**accounted**	**accounting**	**accounts**
NULL-s-'s	afternoon	afternoons	afternoon's	
e-ed-ing-es	accuse	accused	accusing	accuses
ies-y	ability	abilities		
NULL-al-s	addition	additional	additions	
NULL-ped-ping-s	drop	dropped	dropping	drops
ied-ies-y-ying	tried	tries	try	trying

proceed	demand	depend	extend
appeal	reveal	level	dream
remain	train	maintain	question
develop	appear	remember	consider
answer	honor	expect	shift
represent	point	print	mount
request	consist	exist	review


Start
econom-, techn-
furi-, vigor-
emotion-
67
31
45
81
38
80
44
36
-ic
-ate
-ful
-ing
-less
-ive
-ous
-al
4
econom-ic-al
vigor-ous-ly
(null)
-ly
End

words			
jump	jumped	jumping	jumps
move	moved	moving	moves
stop	stopped	stopping	stops
try	tried	trying	tries
make	made	making	makes
buy	bought	buying	buys

We need a new device that will show us how words are used. . . a **megascope**.


knate myers


A BELL QUALITY PRODUCT
IT'S OUT OF THIS WORLD!
SPACE COMMANDO MEGASCOPE
©SPACE COMMANDO
INTER-PLANETARY
MEGASCOPE
6-WAY SPACE CONTROL
6 TOYS in one
MADE OF

Tom wrenched himself upward, for one dizzying moment hanging free in space

THE TOM SWIFT INVENTION ADVENTURES

TOM SWIFT
AND HIS MEGASCOPE SPACE PROBER

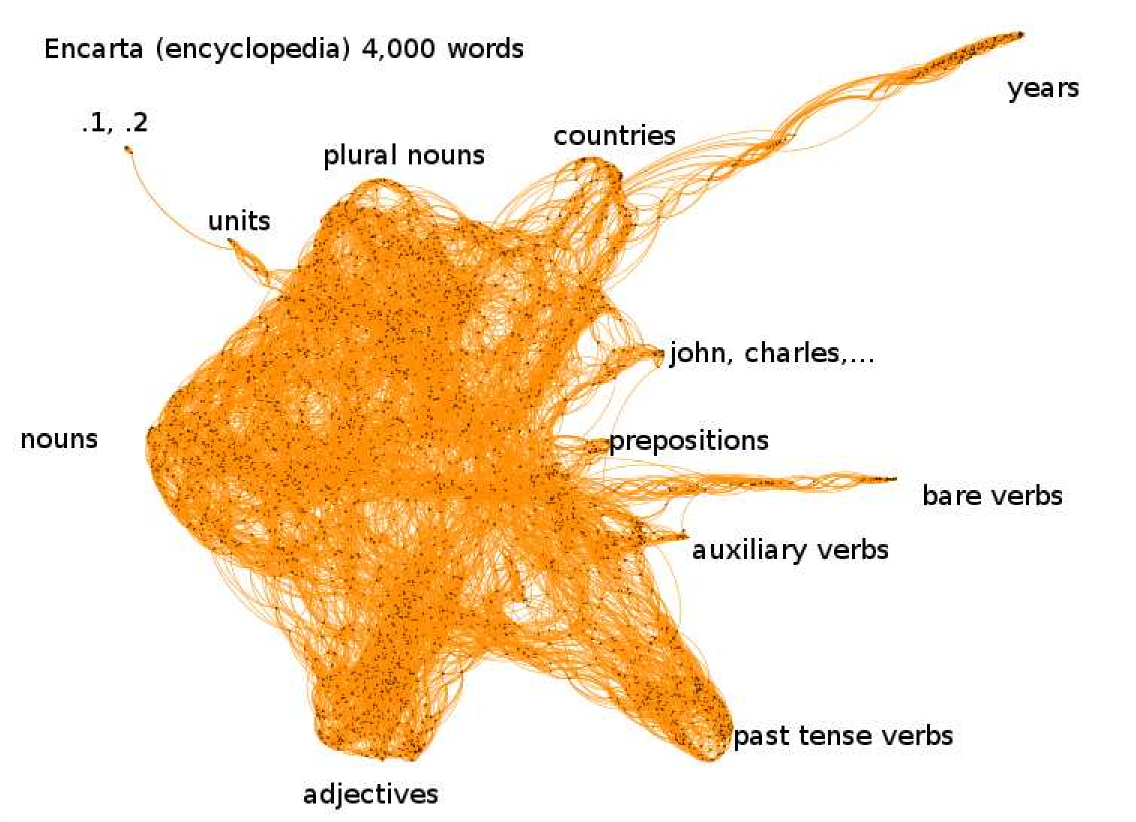
BY VICTOR APPLETON II

ILLUSTRATED BY SCOTT DICKERSON

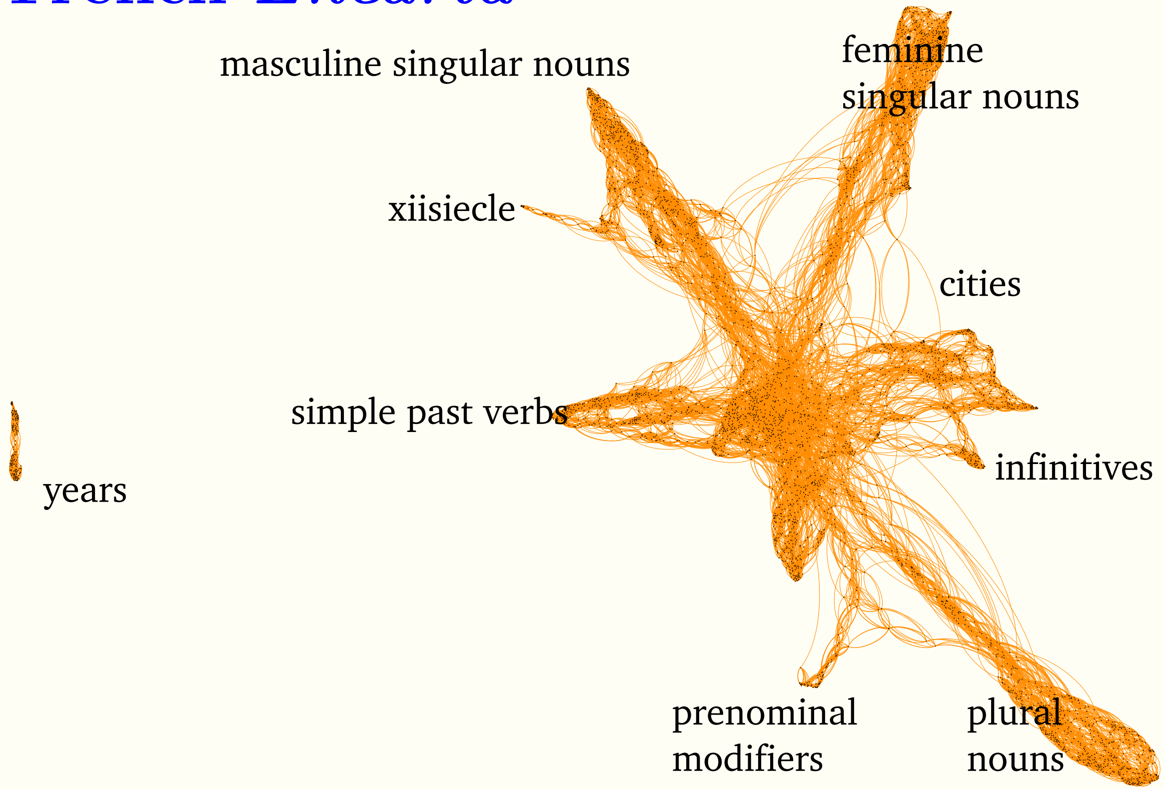
WWW.TOMSWIFTLIVES.COM

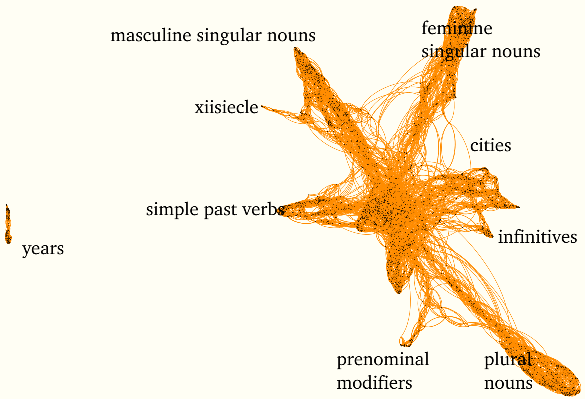
Part 3: The Syntactic Megascope

English *Encarta*

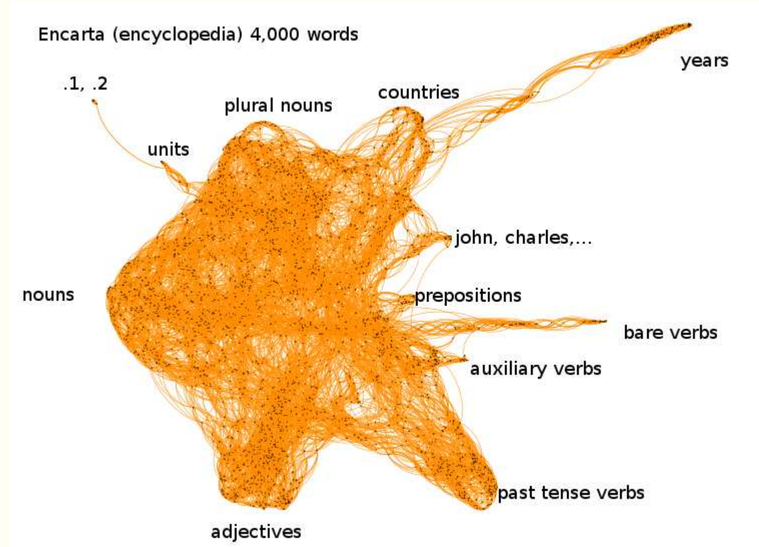


French *Encarta*





French



English

A reminder about English parts of speech

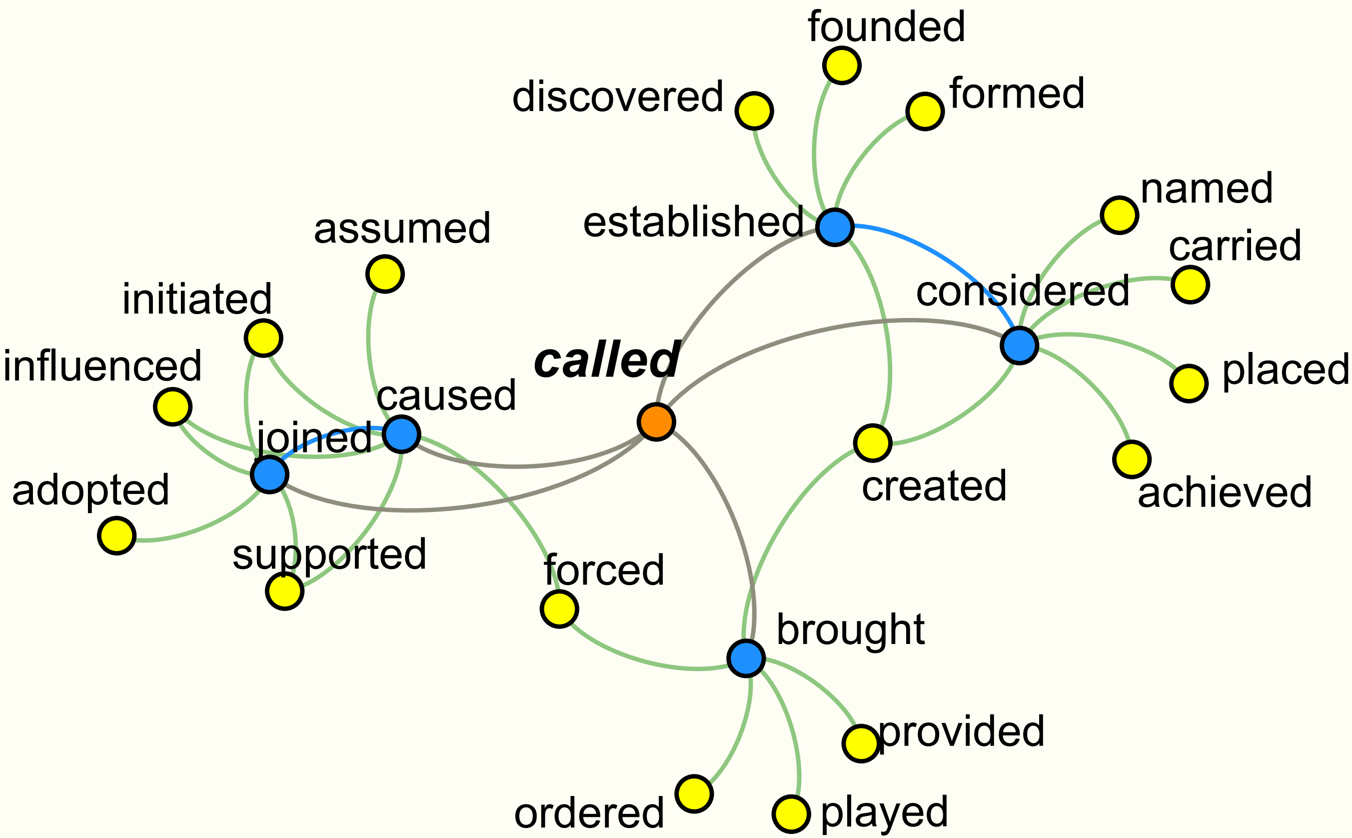
- Prepositions: *to, from, up, down, in, out, of, off*
- Modal auxiliaries: *Can I go outside?* but not *Speak you French?*
  - *I cannot speak Russian* but not *I speak not Russian.*
  - *can, could, must, should, shall, will, would*
  - Forms of *be* also invert, and there is a dummy *do* available as needed.

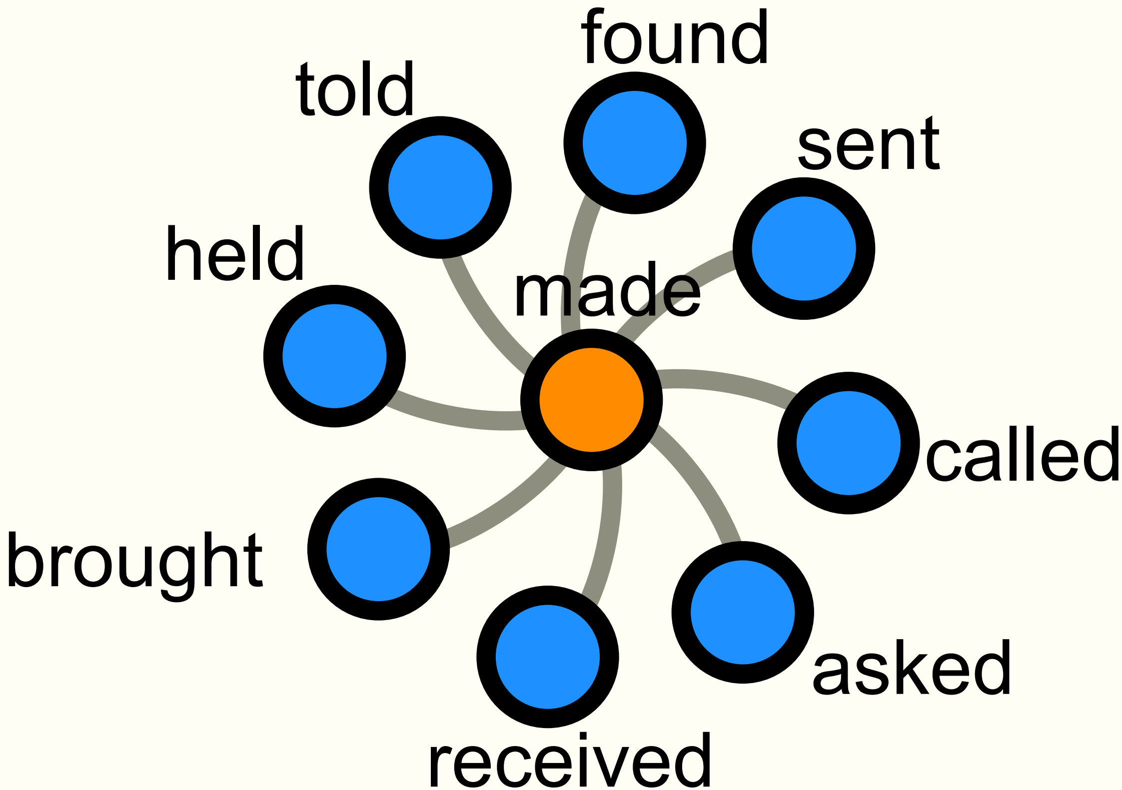
Dynamic view: English color codes

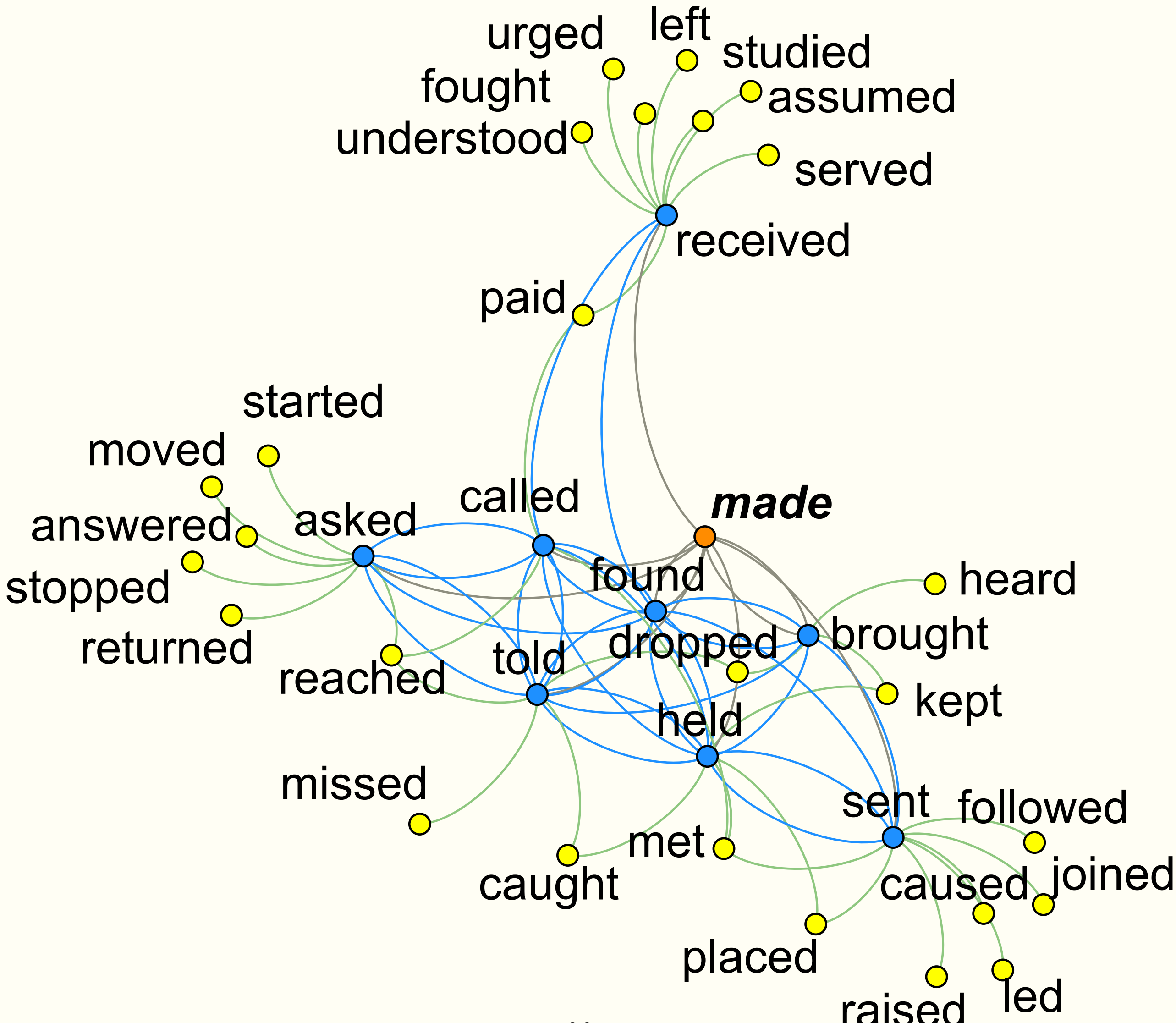
Verbs: ‘bare’ verb (*jump*)	red
Verbs: past tense(*jumped*, *bought*)	blue
Verbs: auxiliary (*should*, *can*)	green
Prepositions (*from*, *to*, *up*, *down*	aqua
Adjectives	purple
Cities	gray
Nouns	pink

Dynamic view: French color codes

Infinitives	red
Prepositions	light blue
Past participles	blue
Adjectives	purple
Cities	gray
Masculine nouns	pink
Feminine nouns	light green
Inflected verbs	light gray


founded
discovered
formed
named
established
carried
assumed
considered
initiated
called
placed
influenced
caused
joined
created
achieved
adopted
supported
forced
brought
provided
ordered
played


found
told
sent
held
made
called
brought
asked
received


urged
left
studied
fought
assumed
understood
served
received
paid
started
moved
called
made
answered
asked
found
stopped
heard
dropped
brought
returned
told
reached
kept
held
missed
sent
followed
met
joined
caused
caught
placed
led
raised

Conclusions

- The importance of asking elementary questions.
- Machine learning: More surprising answers to questions asked of Mother Language.
- Interdisciplinary applications: bioinformatics.
- Data visualization.