

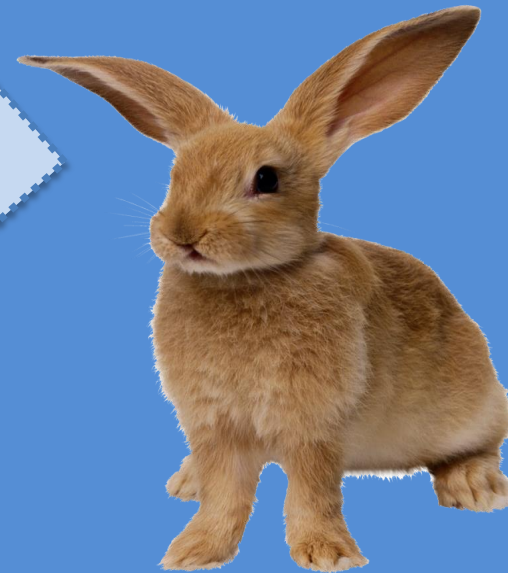
THE BIG PERSPECTIVE

Mr Joseph Gavagai
- the typological bunny

Hedvig Skirgärå

Young Scientists Stockholm

hedvig@lingolympiad.com



Who am I?



- Project leader at Unga Forskare Stockholm (Young Scientists Stockholm)
- Assistant at the Max Planck Institute for Psycholinguistics in Nijmegen, The Netherlands
- Master in Linguistics from Stockholm University (candiatet 3-4 yrs, MA 1-2 yrs, PhD >3-4 yrs)
- Focus on general linguistics
grammatical typology, contact linguistics, complexity



!QUIZ!

172 languages are spoken by more than
3,1 million each.

8 by more than 100 millions each

Guess the top-8!

Answer at the end of presentation

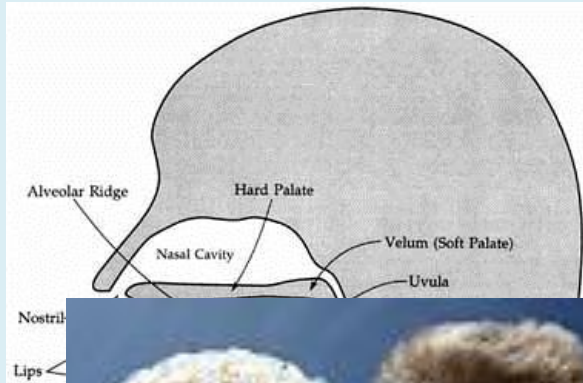


NB only native speakers



What is linguistics?

Scientific study of language as a phenome



, spoken, informally
written, formally



“General”



Psycholinguistics

Field work

What is language typology?



What is linguistic typology?

- classifications of languages into types
- cross-linguistic comparison

in order to find

- Kinship and genealogy and/or
 - in
 - Po
 - cha
- ... behind language
... guage

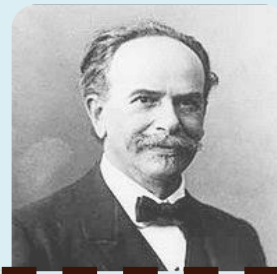
What are the cultural, historical, and geographical constraints on language change?



*[the aim of linguistic theory]
must be to show which
structures are possible, in
general, and why it is just those
structures, and not others, that
are possible*

(Hjelmslev 1970 [1963]: 96)





Important to remember!

All languages can express all concepts

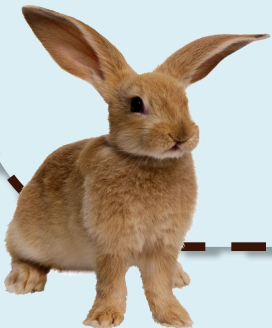
*grammar [...] determines those aspects of each
experience that must be expressed*

(Boas 1938:132)

Other “truths”

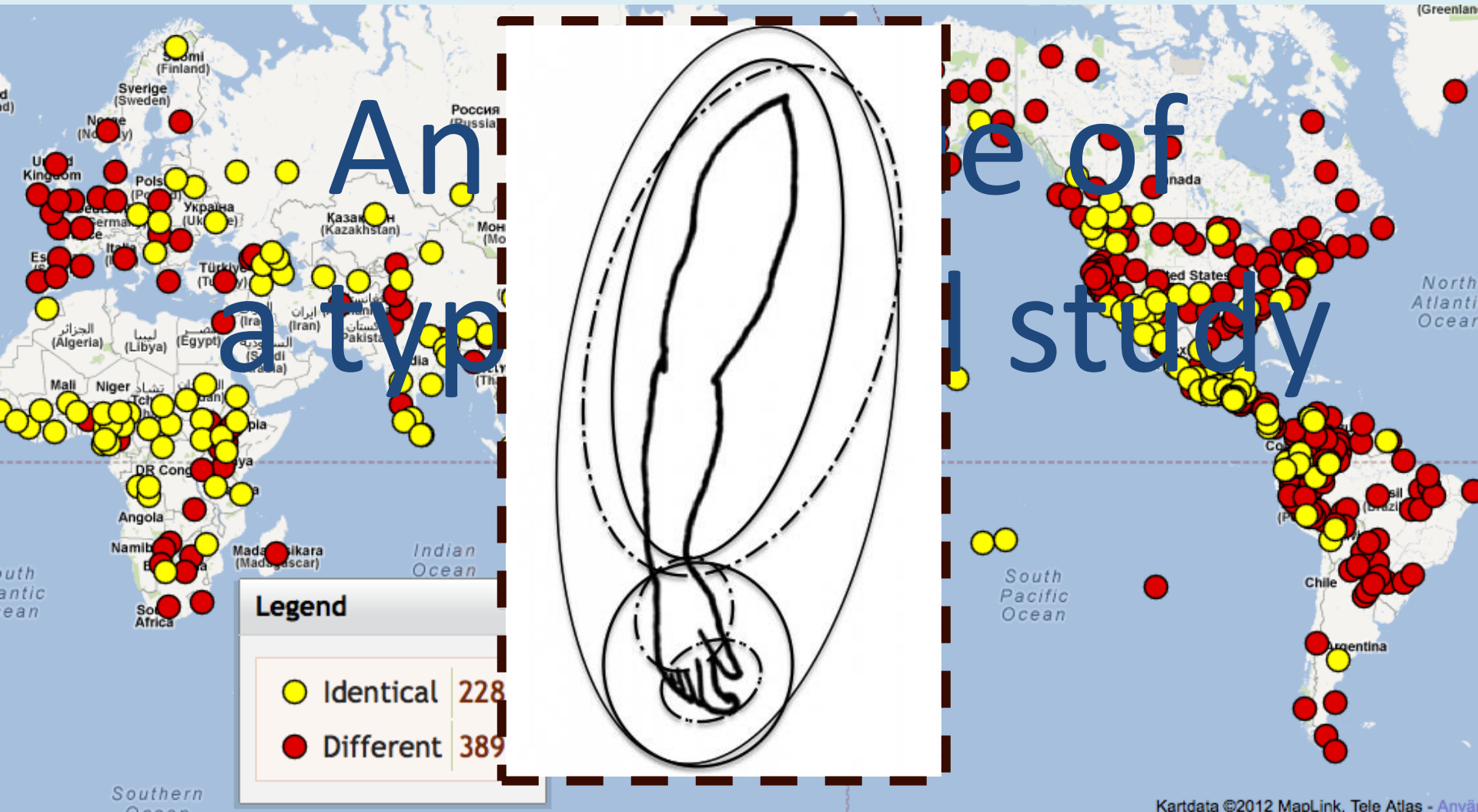
However, for most researchers that in their research assume that the opposite is true these distinctions often don't matter – i.e. they're most often too small to be relevant.

We think...



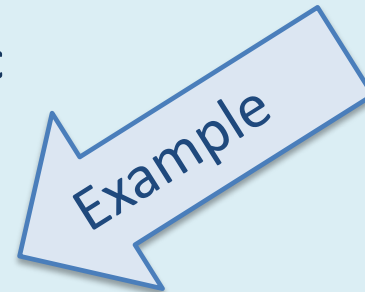
World Atlas of Language Structures (WALS) Chapter 129: Hand & Arm

by Cecil H. Brown 2011



What is a language?

- Human & natural (i.e. not mathematics, codes, Klingon or Python)
- Linguistic criteria: mutual intelligibility (to a certain extent)
- Political issue rather than linguistic



1110 My hovercraft is full of eels 1560 1660 1905 SAMMI

Danish: mit luftpudefartøj er fyldt med ål

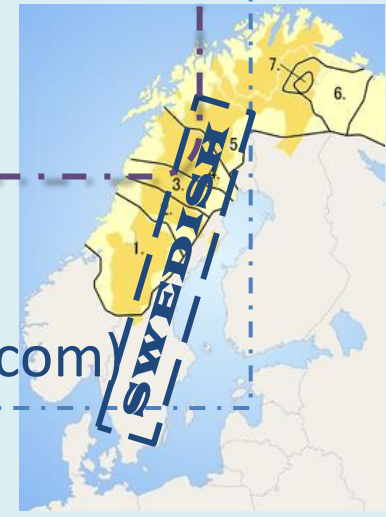
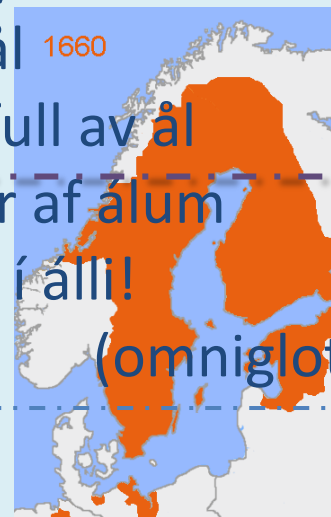
Swedish: min svävare är full av ål 1660

Norwegian: min luftputefartøy er full av ål

Icelandic: svfnökkvinn minn er fullur af álum

Faroese: luftputufar mítt er fult í álli!

(omniglot.com)



IDIOLECT

What is a language? cont.

LANGUAGE A

DIALECT A

DIALECT B

DIALECT C

SOCIOLECT A

SOCIOLECT B

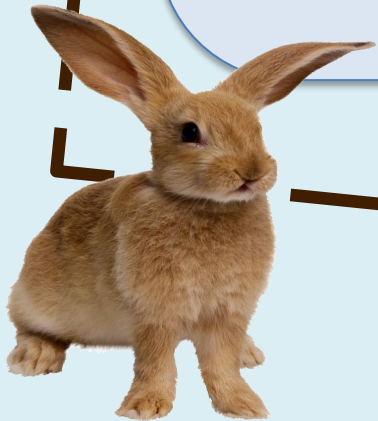
GROUP SLANG A
TELLY WATCHERS

GROUP SLANG B
ACADEMICS



You're welcome to ask questions,
nothing is too dumb! However, please
remember to visit Ethnologue.com,
Glottolog.org, WALS.info and the
other sites on your own ^^!

Glottolog.



Want clues ^^?

- 193 member states of UN
- 203 states recognized by at least one UN member state
- 249 entities of ISO 3166-1
- +7 billion humans on Tellus
- we've had languages for at least 100 000 years
- (we're not sure if language arose in one place or several)
- the distribution of speaker populations are Zipfian



Handy terms to deal with some of these problems

Ethnologue

- Language family
- (Macro language)
- language
- dialect

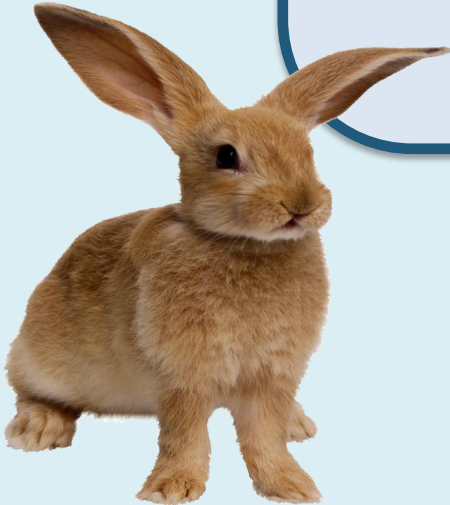
Cysouw & Good (2003)

- Languoid
 - language varieties (sociolect, dialect, etc)
- glossonym
- doculect

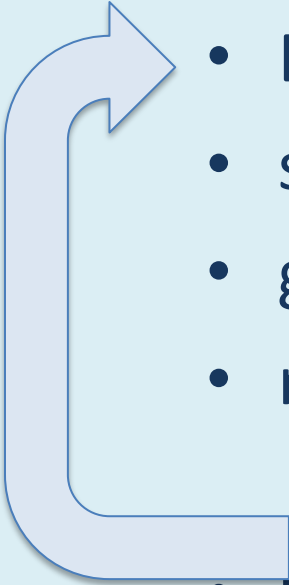
	# lgs	% of all lgs	# speakers	% of all speakers
100,000,000 to 999,999,999	8	0.1%	2,543,460,358	40.4%
10,000,000 to 99,999,999	80	1.1%	2,458,383,987	39.0%
1,000,000 to 9,999,999	305	4.3%	929,591,638	14.8%
100,000 to 999,999	937	13.2%	294,626,823	4.7%
10,000 to 99,999	1,811	25.5%	61,556,414	1.0%
1,000 to 9,999	1,978	27.8%	7,613,358	0.1%
100 to 999	1,062	14.9%	466,128	0.0%
10 to 99	338	4.8%	12,944	0.0%
1 to 9	137	1.9%	541	0.0%
0	204	2.9%	0	0.0%
Unknown	246	3.5%		
Totals	7,106	100%	6,295,712,191	100%

Useful sites








- Ethnologue
- WALS
- MultiTree
- Omniglot
- Glottolog
- Phoible
- The Universals Archive



How do typology?

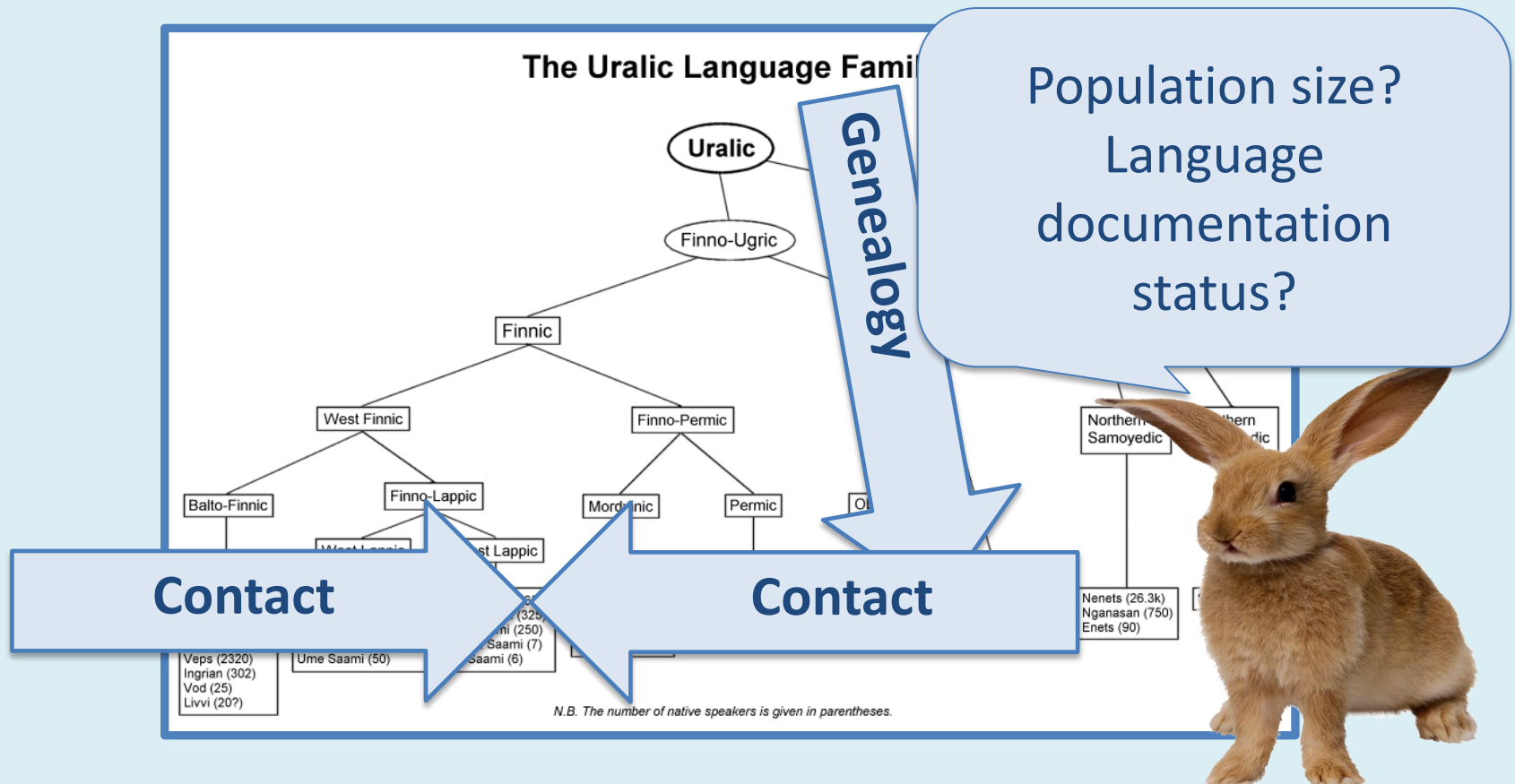
- 
- Define a variable to be investigated
 - sample languages
 - go through the sources
 - realize that the original variable is ill-defined
 - become enough (*lagom*) satisfied and expand sample

Databases of structural features

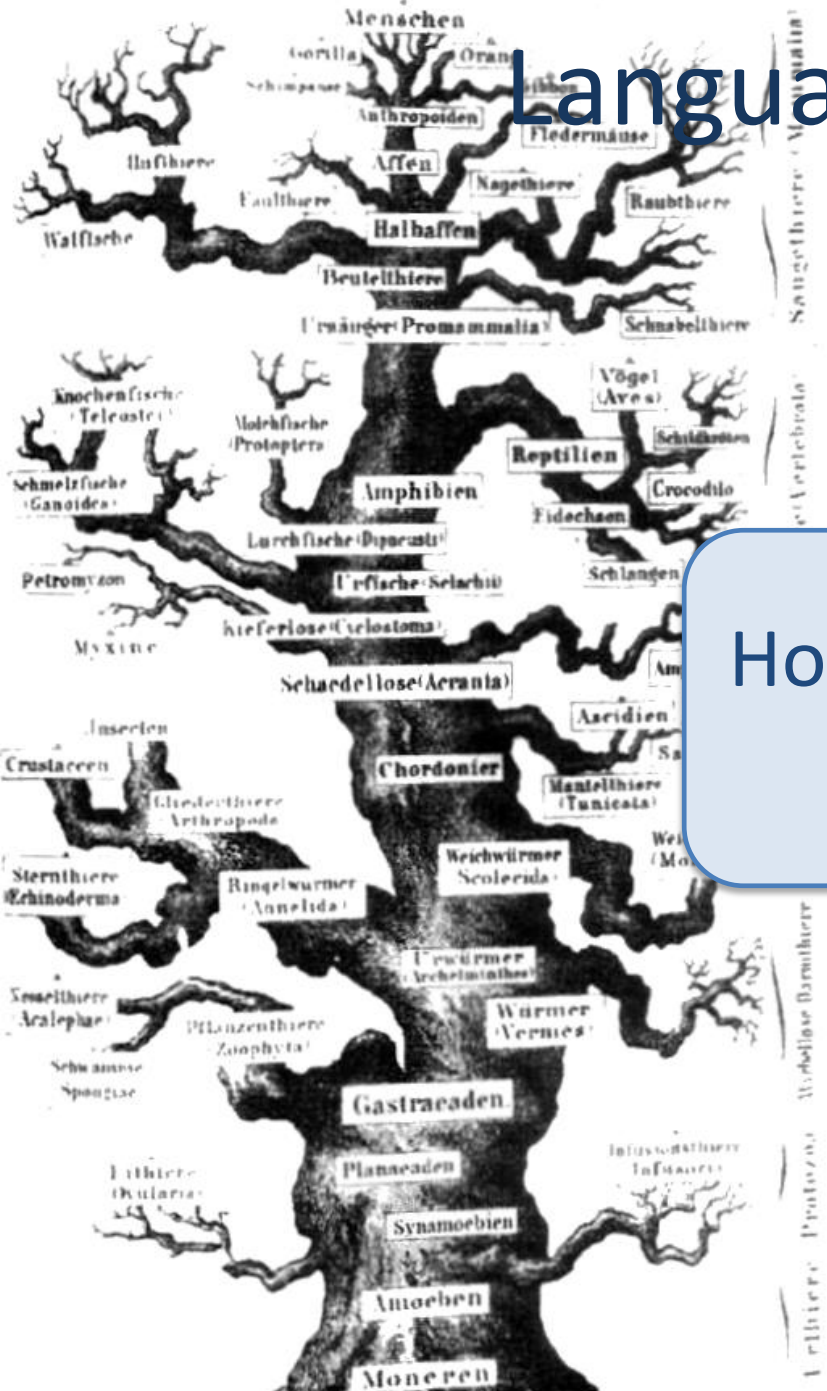
		# languages	# features	datapoints	features with over 200 lgs
Syntactic Structures of the World's Languages (2009)		237	93	14 440	?
World Atlas of Language Structures (2013)		2 679	165 (192)	69 590 (76 465)	129
Nijmegen Typological Survey (2014)		274	279	48 435	145
Atlas of Pidgin and Creole Language Structures (2013)		76	130	TBA	N/A
Valency Patterns Leipzig Online Database (2013)		36	N/A	N/A	N/A
Phonetics Information Base and Lexicon (2012)		1 010	1 680	N/A	N/A
South American Indigenous Language Structures (SAILS)		167	604	31794	N/A

Sampling

- All languages: 5-9 000
- Unbiased with respect to genealogy or contact



Language families



Nichols	(1992)	≈300
Campbell	(2009)	250-300
WALS	(2011)	212

How many language families are there?



Language families cont.

How far back can we go.. ?

Nichols (2003)

about the cognate method: term from WALS 4000-8000 BC
group of languages with a shared history at least 3 500 – 4 000 years ago

Proto-Uraliska:

approx 7000-2000 BC

Proto- Indo-Europeiska:

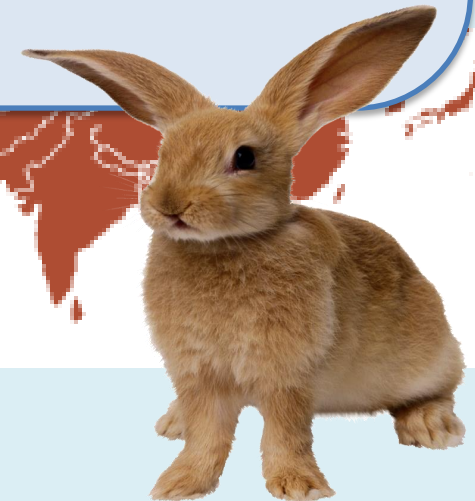
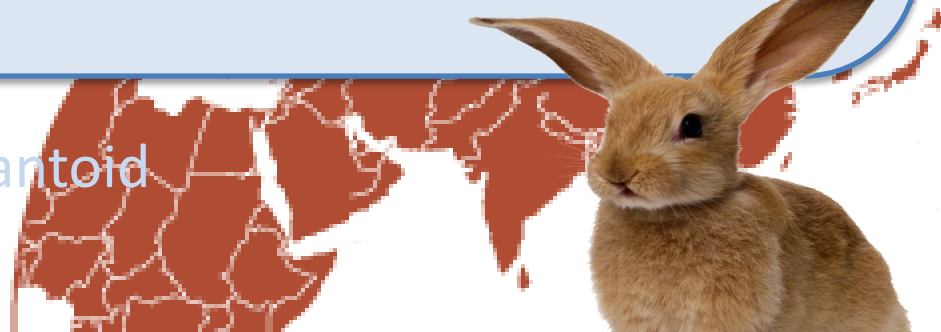
approx 3700 BC

Proto-Dravidiska:

approx 500 BC

THE
OF LANGUAGES STRUCTURES
ONLINE

WALS innehåller 510 gener
Ex: slavisk, germanisk, semitisk, bantoid



More on language families

6 languages families include 63% of all languages and are spoken by 85% of the worlds population

Language families	# languages	% of world population of languages	# speakers	% of world population of speakers
Niger-Congo	1 526	21,47%	435 432 213	6,92%
Austronesian	1 222	17,20%	346 489 508	5,50%
Trans-New-Guinea	476	6,70%	3 540 024	0,06%
Sino-Tibetan	455	6,40%	1 268 218 984	20,14%
Indo-European	437	6,15%	2 925 253 210	46,46%
Afro-Asiatic	367	5,16%	374 573 409	5,95%
Total	4 483	63,09%	5 353 507 348	85,03%

What should we do about “very small families”?

An isolate is a lang

What should we do about “very small families”?



Isolate



Basque

Korean

Burushaski

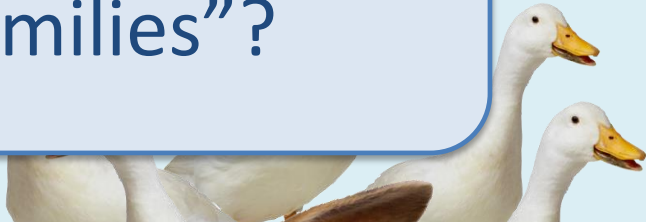
74

189

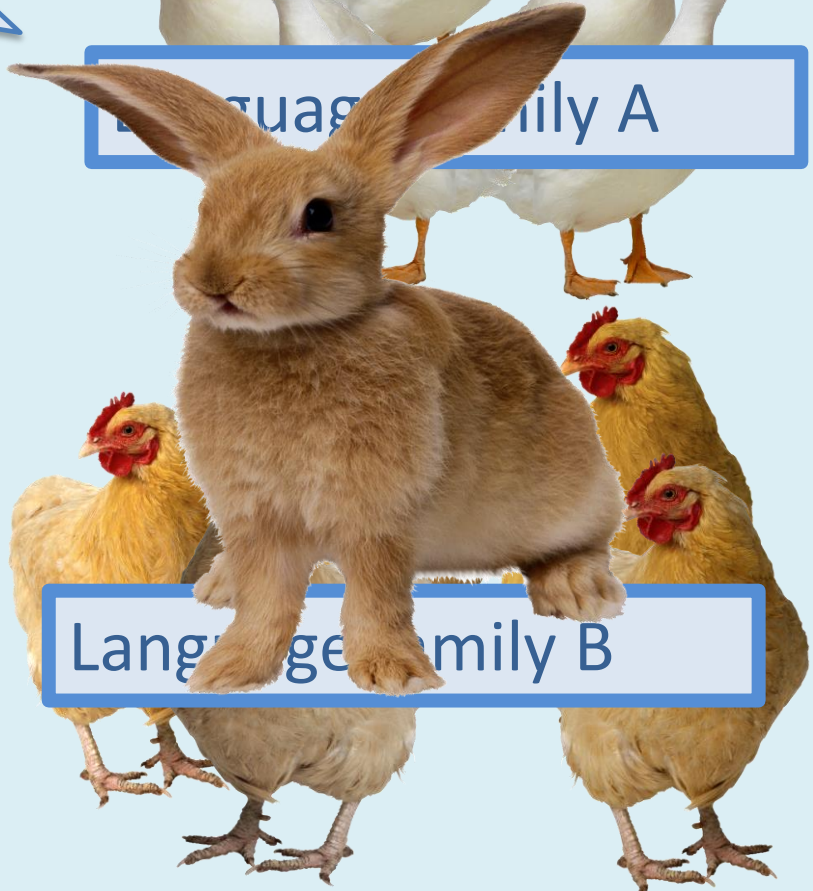
129



Campbell (2009)



Language Family A



Language Family B

Sampling cont: geography

Macroarea	Living languages		Speaker population	
	Number	Percent of totalt	Number	percent of total
Africa	2,146	30.2 %	997,320,660	15.9 %
Americas	1,060	14.9 %	946,060,483	1.5 %
Asia	2,303	32.4 %	4,086,262,000	65.5 %
Europe	285	4.0 %	735,669,330	11.7 %
Pacific	1,312	18.5 %	35,284,389	0.1 %
Total	7,106	100 %	6,236,421,567	100 %

Lewis (2009)



Participating countries in IOL 2014

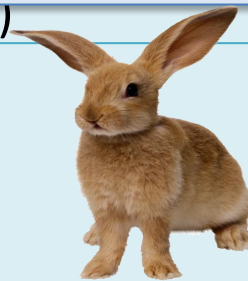
Country	# lgs	GDI
India	449 (442)	0.913
Pakistan	76 (71)	0.795
Singapore	32 (24)	0.756
Kazakhstan	44 (41)	0.699
Latvia	14 (7)	0.628
Canada	172 (88)	0.601
China	301(298)	0.511
Taiwan	26 (21)	0.489
Estonia	10 (4)	0.445
Ukraine	42 (24)	0.420
Turkmenistan	27 (4)	0.385
United States	420 (215)	0.334
Netherlands	38 (15)	0.301
Spain	22 (15)	0.277

Country	# lgs	GDI
Russia	137 (105)	0.251
Bulgaria	17 (11)	0.217
Australia	244 (213)	0.211
Slovenia	10 (8)	0.160
Sweden	23 (13)	0.153
UK	55 (12)	0.144
Georgia	10 (4)	0.133
Czechia	10 (4)	0.122
France	10 (4)	0.111
Italy	10 (4)	0.100
Poland	10 (4)	0.089
Belgium	10 (4)	0.078
Japan	10 (4)	0.067
Isle of Man	2 (2)	0.000

GDI stands for the Greenbergs Diversity Index

High GDI = unlikely that two people in that country speak the same language natively.

Low GDI = likely



Universals

Joseph Greenberg (1915-2001)

- founder of modern typology
- genealogical classification

The Universals Archive

Absolute universals (297/2029)

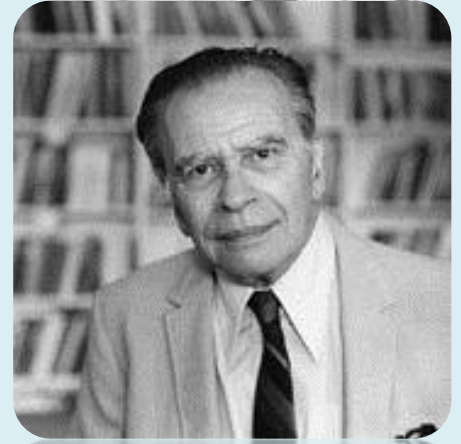
“all languages have X”

Dependent universals (1476/2029)

“if a language has X, if also has Y”

Hierarchies

singular>plural>dual>trial



(Greenberg 1963, universalie 511)

Example of universals

Universals
IF the
terms

Implic
Berlin

White
Black

Dryer
tends

“Languages with normal SOV order
are postpositional”



Man
subject



Dog
object

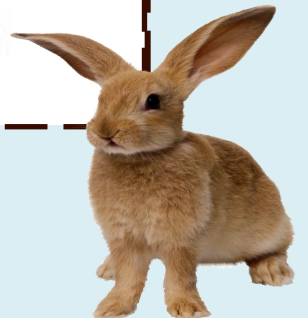


Put
verb



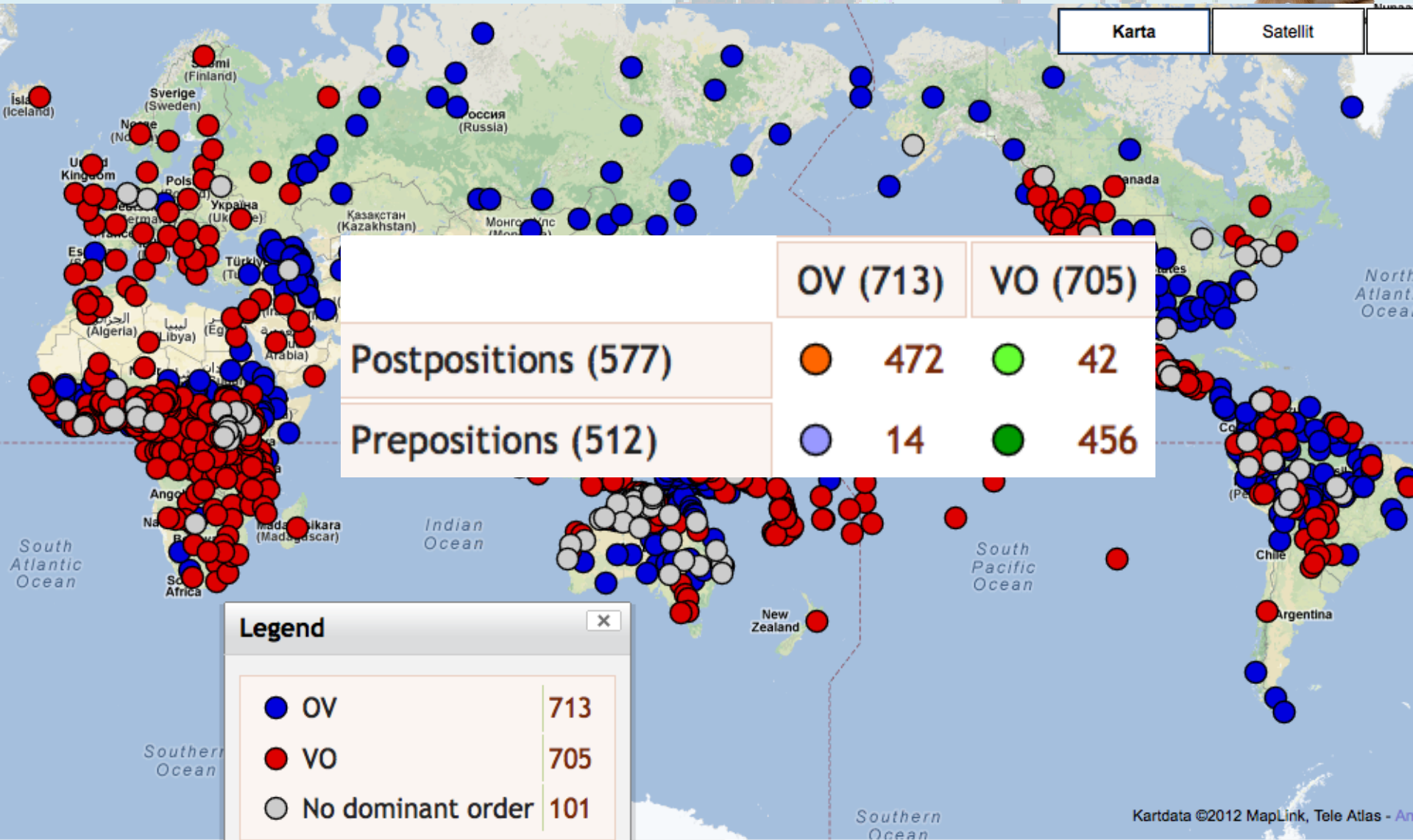
Canoe in
noun

Purple
Pink
Orange
Grey



Chapter 83: Order of Verb and Object

av Matthew Dryer (2011)



How many do we know anything about?

Common size of typological investigations: 200-450 languages



Type	# Languages
< Grammar	2580
< Grammar sketch	1161
< Dictionary	212
< Phonology	106
< Text collection	26
< Wordlist	1073
< Ethnographic	78
< Minimal	21
Unknown	263
All	5520

Lang

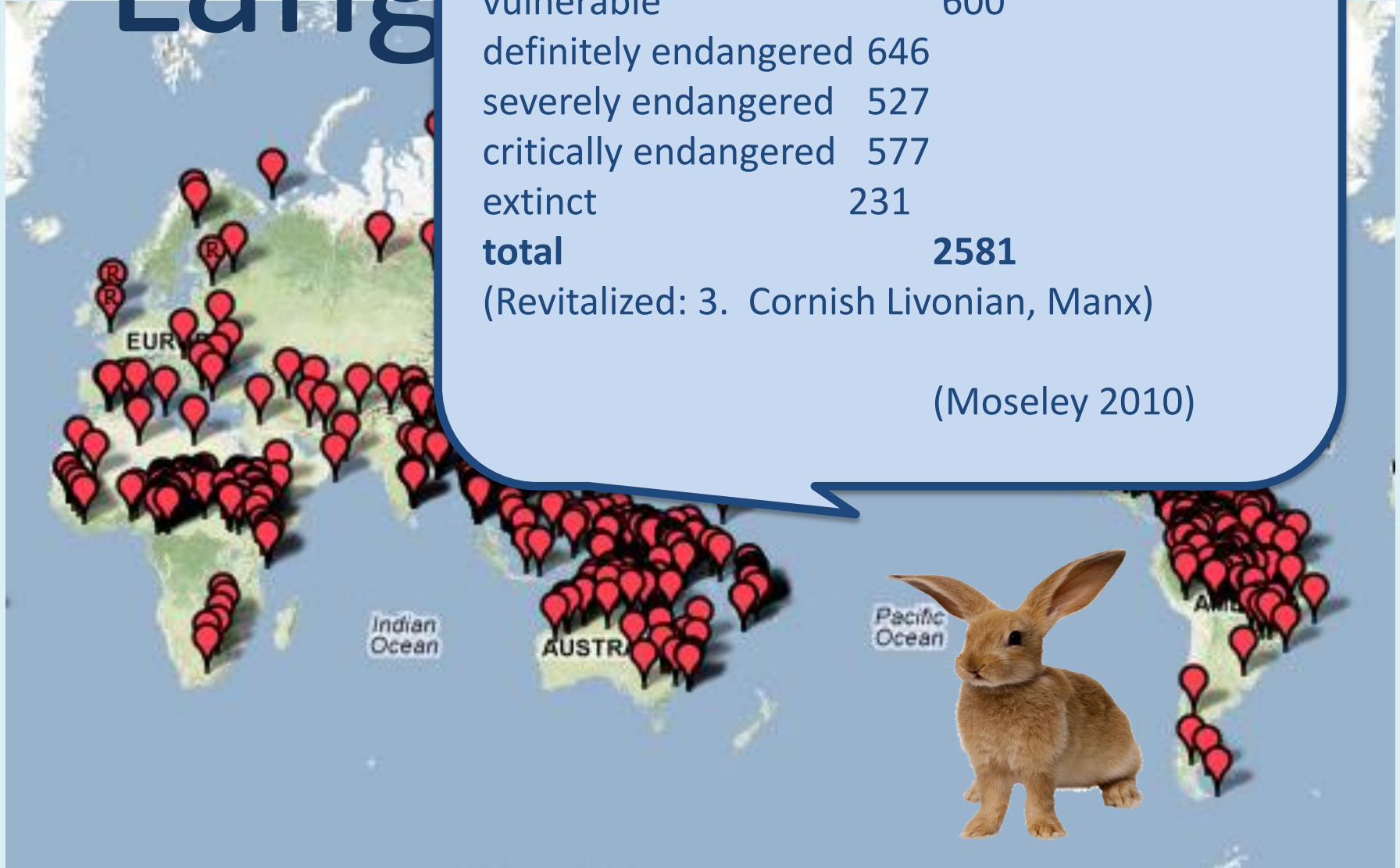
UNESCO

vulnerable	600
definitely endangered	646
severely endangered	527
critically endangered	577
extinct	231

total 2581

(Revitalized: 3. Cornish Livonian, Manx)

(Moseley 2010)



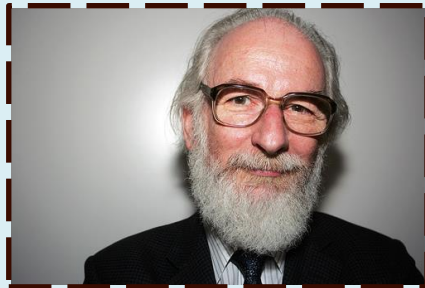
How many are dying?

“Critical limit”: 20 000 speakers

45% of the worlds languages today have less than 20 000 speakers each

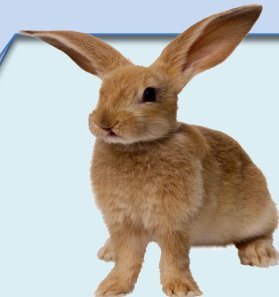
Reasonable estimation

>50% will die in 100 years
or even
90% dead or endangered



(Crystal 2000:1)

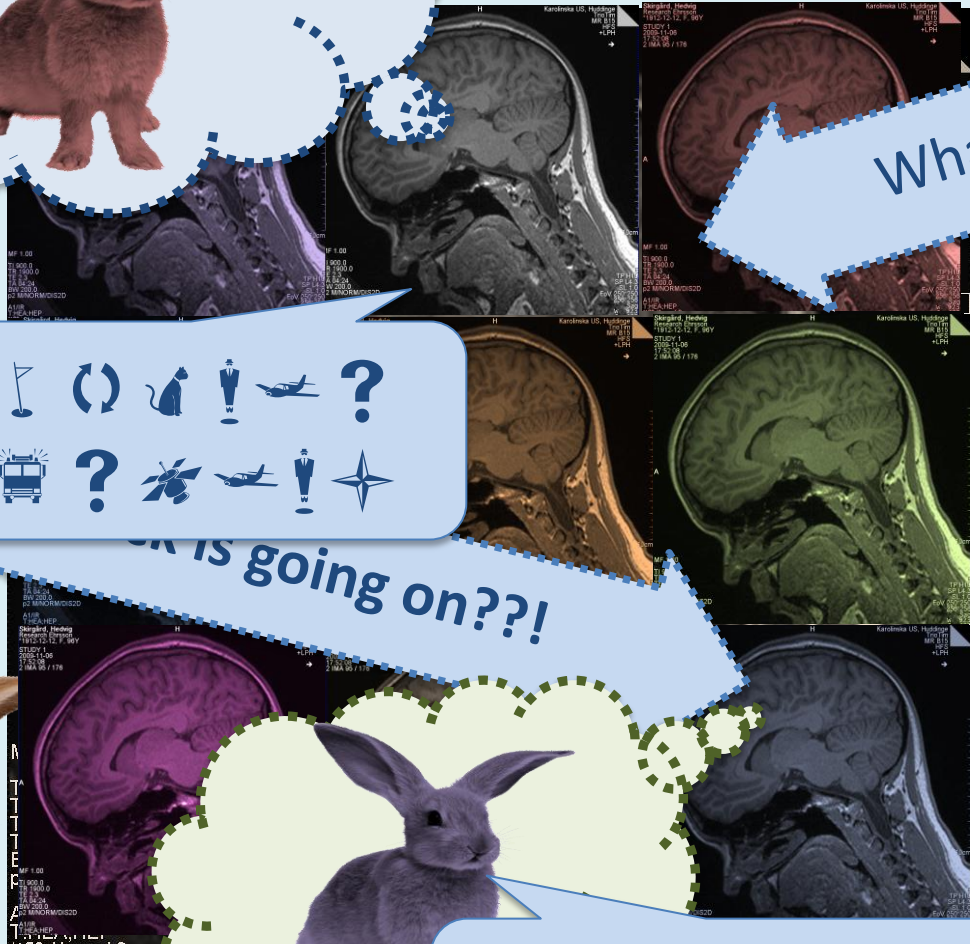
6% of the population of the world today are speaking 90 % of the languages



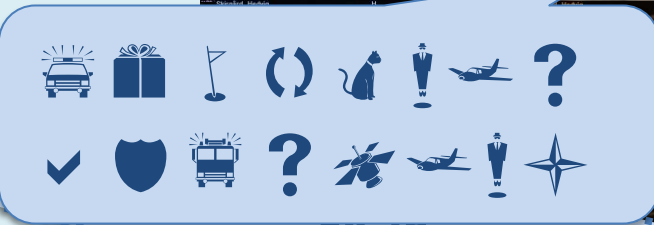
Why is language typology interesting?

- Same species=same brains = same capacity
- Yet, great diversity and disparity!
- Is something impossible?
- What is complexity?
- Is complexity even?
- What influences "complexity"?
- What does this say about our cognitive capacity and categorize our reality?
(cf psychology, anthropology, neurology and biology)





What is going on?



What is going on??!



Answers to the quiz!

(Ethnologue 2014)

1. "Chinese"	1.497 billions
2. Spanish	414 millions
3. English	335 millions
4. Hindi	260 millions
5. "Arabic"	203 millions
6. Portuguese	193 millions
7. Bengali	187 millions
8. Japanese	167 millions

BUT!

In summary

- Linguistics is a very broad field
- It's tricky defining what a language is
- There are many different languages and
- Few are written
- They are all different and vice versa...
- We know very little about many of them
- Some are dying
- Most of them are dying

We've talked about
a lot of things...

:) or :(?



BYE

Hedvig Skirgård
Unga Forskare Stockholm
hedvig@lingolympiad.org



Referenser 1

Berlin, Brent and Paul Kay (1969) *Basic Color Terms: Their Universality and Evolution*. Berkeley and Los Angeles. University of California Press

Boas, Franz (1938) *Language*. In F. Boas (ed), *General Anthropology*. Boston, Nya York: D.S. Heath and Company

Brown, Cecil H. (1983). Where do cardinal direction terms come from? *Anthropological Linguistics* 25: 121-161.

Brown, Cecil H.. (2011) *Hand and Arm*. In: Dryer, Matthew S. & Haspelmath, Martin (eds.) *The World Atlas of Language Structures Online*. Munich: Max Planck Digital Library, feature 129A. Available online at <http://wals.info/feature/129A> Accessed on 2012-06-19.

Campbell, Lyle (1998) *Historical linguistics - an introduction*

Campbell, Lyle (unpublished) *Language Isolates and Their History, or, What's Weird, Anyway?*

Croft, William (1990) *Typology and Universals*. (Cambridge Textbooks in Linguistics.) Cambridge: Cambridge University Press.

Comrie, Bernard & Dryer, Matthew S. & Gil, David & Haspelmath, Martin. (2011). Introduction. In: Dryer, Matthew S. & Haspelmath, Martin (eds.) *The World Atlas of Language Structures Online*. Munich: Max Planck Digital Library, supplement 1. Available online at <http://wals.info/supplement/1> Accessed on 2012-06-17.

Crystal, David (2000). *Language Death*. Cambridge: Cambridge University Press

Referenser 2

Dryer, Matthew S.. (2011). Genealogical Language List. In: Dryer, Matthew S. & Haspelmath, Martin (eds.) The World Atlas of Language Structures Online. Munich: Max Planck Digital Library, supplement 4. Available online at <http://wals.info/supplement/4>. Accessed on 2012-06-17.

Dryer, Matthew S.. (2011) Order of Object and Verb. In: Dryer, Matthew S. & Haspelmath, Martin (eds.) The World Atlas of Language Structures Online. Munich: Max Planck Digital Library, chapter 83. Available online at <http://wals.info/chapter/83> Accessed on 2012-12-06.

Gordon, Jr., R. G., editor (2005). Ethnologue: Languages of the World. Dallas: SIL International, 15 edition.

Hammarström, Harald and Nordhoff, Sebastian (2011) How many languages have so far been described? pp. Paper presented at NWO Endangered Languages Programme Conference, Leiden, April 2011

Hammarström, H. (2008). Automatic annotation of bibliographical references with target language. In Proceedings of MMIES-2: Workshop on Multi-source, Multilingual Information Extraction and Summarization, pages 57–64. ACL.

Hammarström (2009). chapter 5: Counting Languages in Dialect Continua Using the Criterion of Mutual Intelligibility

Referenser 3

Hammarström, H. and Nordhoff, S. (2010). Langdoc: Bibliographic infrastructure for linguistic typology. Oslo Studies in Language 3 (2). Pages: 31–43. (ISSN 1890-9639)

Hammarström, Harald. 2007. The Languages Families of the World: A Critical Synopsis.

www.cs.chalmers.se/~harald2/.

Hjelmslev, Louis (1970) Language. An introduction. Madison: The University of Wisconsin press [translation by Francis. J Whitfield of Hjelmslev (1963) Sproget. En introduktion. Copenhagen: Berlingske Forlag,

Campbell, Lyle. (2008) Language isolat and Their History, or, What's Weird, Anyway?. Paper from workshop "The Historical Investigation of Language isolat, or, What's weird about them?" Spring Workshop in Reconstruction, University of Michigan, April 4-5, 2008.

Lewis, M. Paul, Gary F. Simons, and Charles D. Fennig (eds.), 2013. Ethnologue: Languages of the World, Seventeenth edition. Dallas, Texas: SIL International. Online version:

<http://www.ethnologue.com>.

Maddieson, Ian. (2011a). Tone. In: Dryer, Matthew S. & Haspelmath, Martin (eds.) The World Atlas of Language Structures Online. Munich: Max Planck Digital Library, chapter 13. Available online at <http://wals.info/chapter/13> Accessed on 2012-11-19.

Referenser 4

Maddieson, Ian. (2011b). Absence of Common Consonants. In: Dryer, Matthew S. & Haspelmath, Martin (eds.) The World Atlas of Language Structures Online. Munich: Max Planck Digital Library, chapter 18. Available online at <http://wals.info/chapter/18> Accessed on 2012-11-22.

Moseley, Christopher (ed.). (2010). Atlas of the World's Languages in Danger, 3rd edn. Paris, UNESCO Publishing. Online version:

<http://www.unesco.org/culture/en/endangeredlanguages/atlas>

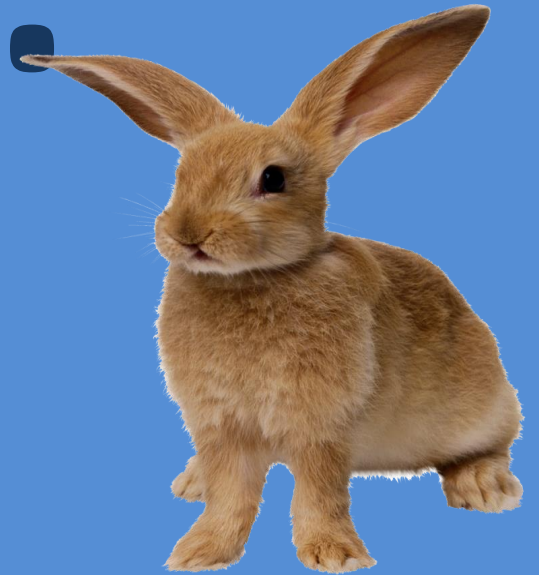
Nichols, J (1992) Linguistic Diversity in Space and Time. Univ. of Chicago Press, Chicago,

Nichols, J (2003) in The Handbook of Historical Linguistics, B. D. Joseph, R. D. Janda, Eds. Blackwell, Oxford, pp. 283–310.

Parkvall, Mikael (2010) Lagom finns bara i Sverige. Stockholm: Telegram bokförlag

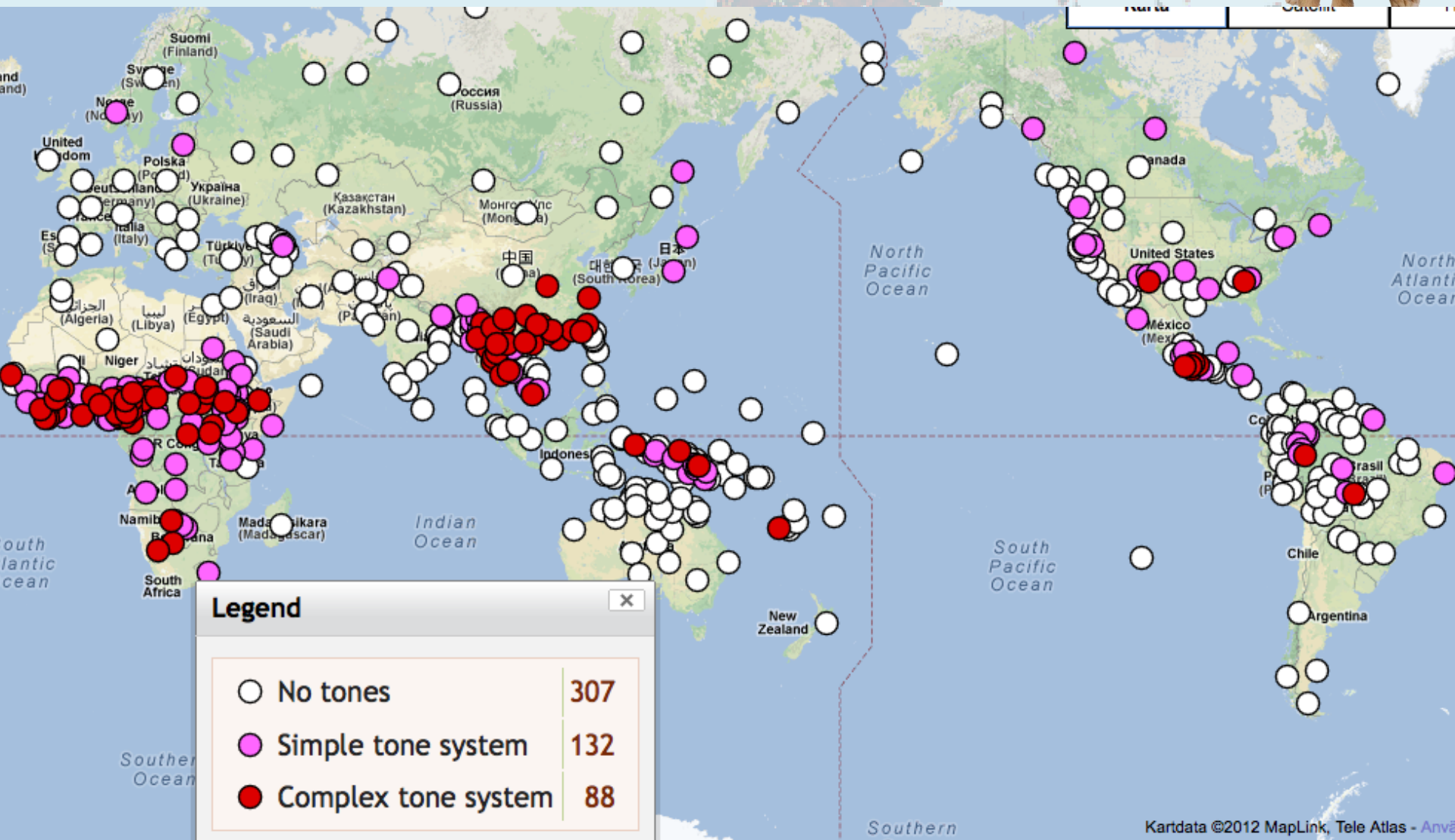


EXTRA-
OST?



Chapter 13: Tone

Ian Maddieson 2011



Kapitel 18: Absence of Common Consonants

av Ian Maddieson 2011 ([LÄNK](#))



Bilabial



○ All present	503
● No bilabials	4
● No fricatives	48
● No nasals	10
● No bilabials or nasals	1
● No fricatives or nasals	1

How many languages per country?

Ethnologue facts

Languages per country

Mean 45,6

Median 18

Max 830

(Papua New Guinea)

Min 1

(St Helena, Vatican, North Korea,
Falklands & British Indian Ocean Territory)

Indigenous languages per country

Mean 38,3

Median 10

40 countries

70% Indigenous

830/830
(Papua New Guinea)





How many has there ever been?

Bickel (p.c. 2012)

If we assume the following

- ❖ At each point in time there is 5 000 languages
- ❖ Each language has a lifetime of 1 000 years
- ❖ We've spoken languages for at least 100 000 years

$$(100\,000 / 1\,000) * 5\,000 = 500\,000$$

At least half
a million!!

This means that
<2 % are alive today



1 language = 1 name?

What's in a name?

Ethnologue 2005 (Hammarström et al.)

39 419 unique names for languages

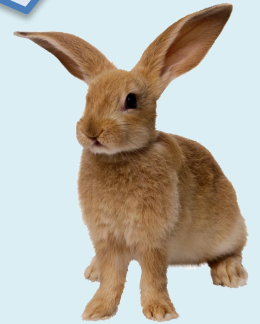
42 768 occurrences of these names

mean 5,86 names per language

max the language Armenian has 77 names

mean 1,08 languages per name

max the name Miao refers to 14 languages



KOMBOWALS

([LÄNK](#))

		130A	
		Identical (72)	Different (521)
129A	Identical (228)	● 15	● 208
	Different (389)	● 57	● 309



129: hand & arm

130: finger & hand

Identical = samma ord

Different= olika ord