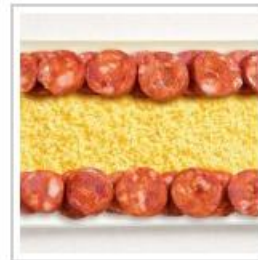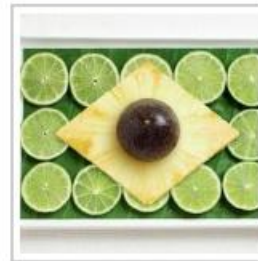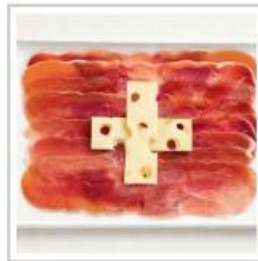# Highlighting cultural distinction from multilingual Wikipedia

Outline

Is an important article in English Wikipedia also important in other Wikipedia?

- Cross-cultural study of Wikipedia
- Ranking algorithms
- Ranking structure of Wikipedia
- Ranking of persons
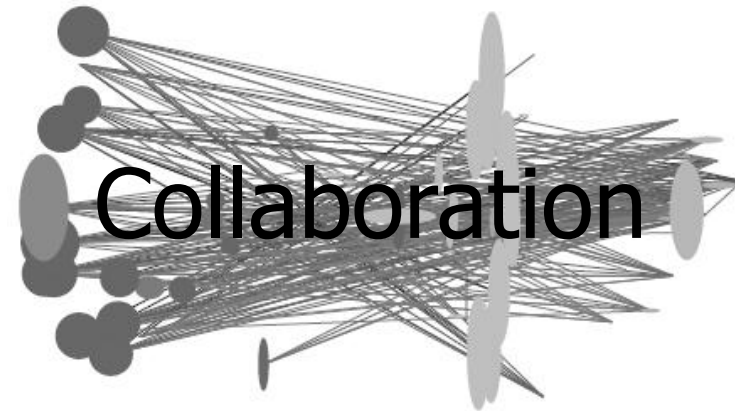- Character-Word interaction in Shakespeare's play

Ranking across Wikipedia

Wikipedia : amazing example of online collaboration

Anybody

Decentralized

Recorded log

Conflict

Collaboration

Language editions of Wikipedia

WIKIPEDIA

English
The Free Encyclopedia
4 240 000+ articles

Español
La enciclopedia libre
1 020 000+ articulos

Русский
Свободная энциклопедия
1 000 000+ статей

日本語
フリー百科事典
860 000+ 記事

Deutsch
Die freie Enzyklopädie
1 590 000+ Artikel

Français
L'encyclopédie libre
1 390 000+ articles

Polski
Wolna encyklopedia
970 000+ haseł

Italiano
L'enciclopedia libera
1 030 000+ voci

Português
A enciclopédia livre
780 000+ artigos

中文
自由的百科全書
690 000+ 條目

275 active editions currently

Language : Key vector of cultural diversity

PART II.
KEY VECTORS OF CULTURAL DIVERSITY

Faced with the challenges of rethinking intercultural dialogue, countering stereotypes as well as tendencies towards withdrawal into closed identities, a deeper examination of the implications of cultural diversity in areas outside of culture in its restricted sense is equally important to public policy-making today. While virtually all human activities are shaped by and in turn help to shape cultural diversity, the prospects for the continued vitality of diversity are crucially bound up with the future of languages, education, the communication of cultural content, and the complex interface between creativity and the marketplace.

From the most immediate manifestation of cultural diversity — languages — to its pervasiveness across the entire social and economic fabric — the marketplace — including its role in education, media and communication policies, Part II explores these often overlapping fields — all central to UNESCO's mandate — with a view to identifying trends and factors that impact on the state of cultural diversity and refining our political agendas for its preservation and promotion in keeping with the complex realities of today's world.

CHAPTER 3 – **Languages**

Chapter 3 addresses the need to safeguard linguistic diversity in the wider context of managing cultural change, while facilitating dialogue and mutual understanding through the promotion of multilingualism and translation capacities.

CHAPTER 4 – **Education**

Chapter 4 argues the case for broadening the compass of educational systems to take account of informal learning environments and learning needs in culturally diverse settings, with a view to enhancing the quality of education and preparing us to live together with our differences through the development of intercultural competencies.

CHAPTER 5 – **Communication and cultural contents**

Chapter 5 analyzes recent trends in the communication of cultural contents, highlighting the challenge of ensuring that the expansion of media diversity and outreach is matched by efforts to redress the imbalances inherent in the digital divide so as to further processes of cultural exchange and mutual understanding.

CHAPTER 6 – **Creativity and the marketplace**

Chapter 6 explores the continuum across artistic creativity, social innovation and economic growth, highlighting the added-value of cultural diversity in key sectors ranging from contemporary art practices, crafts and tourism to corporate activities ranging from management and human resources to marketing and 'cultural intelligence'.
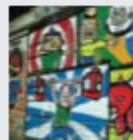
Languages are not just a means of communication but represent the very fabric of cultural expressions

So comparative study across Wikipedia editions can give us some insight for cross-cultural research

Comparison of contents

One-One: contents, length, # of words and photos etc. ➔ Node level

We compare "relation" or "interaction" across Wikipedia

In particular, ranking

How different cultures evaluate a person?

Is an important person in one culture also important in the other culture?

Ranking algorithm based on network structure!

- PageRank
- CheiRank
- 2DRank

Ranking algorithm : PageRank

$$P(i,t) = (1-\alpha)/N + \alpha \sum_j A_{ij}P(j, t-1)/k_{out}(j)$$
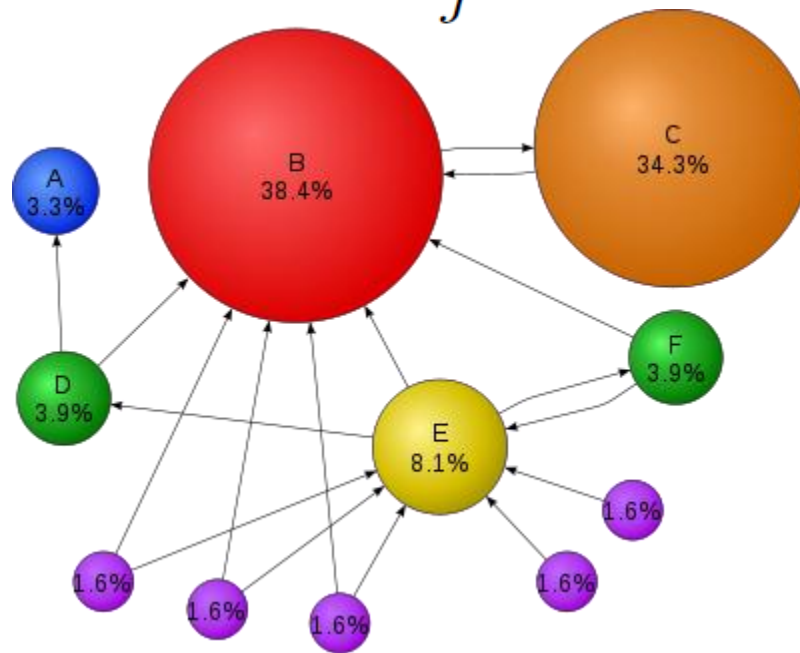
Weighted voting of other Wikipedia articles

Ranking algorithm : CheiRank

PageRank of network given by reversing direction of all links

$$P^*(i) = (1 - \alpha)/N + \alpha \sum_j A_{ji} P^*(j)/k_{in}(j)$$

Ranking algorithm : 2DRank
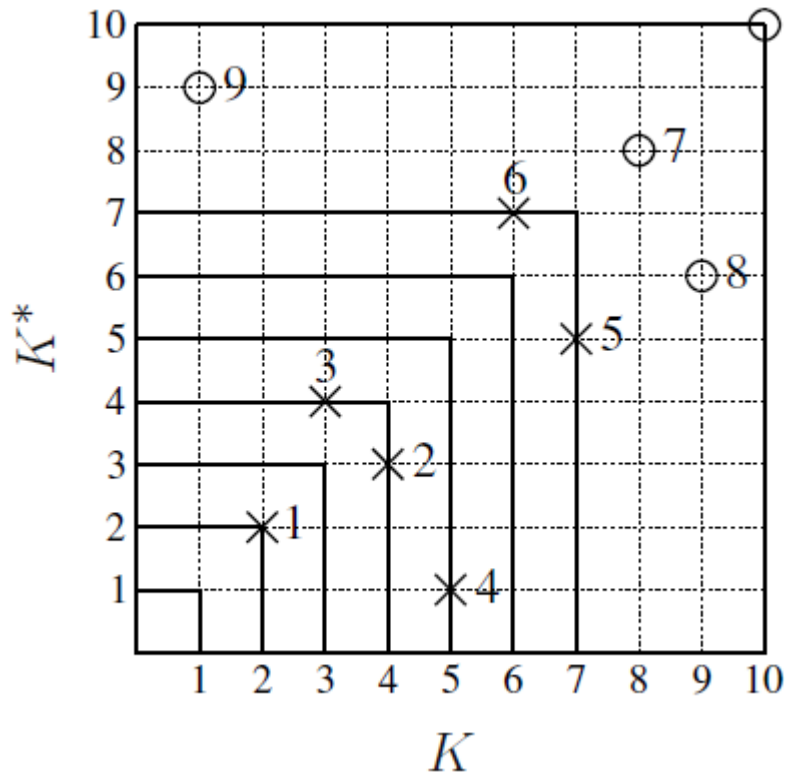
counting nodes
in order of
their appearance
on ribs of squares
in (K, K^*) plane

Wikipedia as a network of articles

Edition	N_A	N_L	κ	Date
EN	3920628	92878869	3.905562	Mar. 2012
FR	1224791	30717338	3.411864	Feb. 2012
DE	1396293	32932343	3.342059	Mar. 2012
IT	917626	22715046	7.953106	Mar. 2012
ES	873149	20410260	3.443931	Feb. 2012
NL	1034912	14642629	7.801457	Feb. 2012
RU	830898	17737815	2.881896	Feb. 2012
HU	217520	5067189	2.638393	Feb. 2012
KO	323461	4209691	1.084982	Feb. 2012

$$\kappa = N \sum_i P(i)P^*(i) - 1$$

General ranking structure
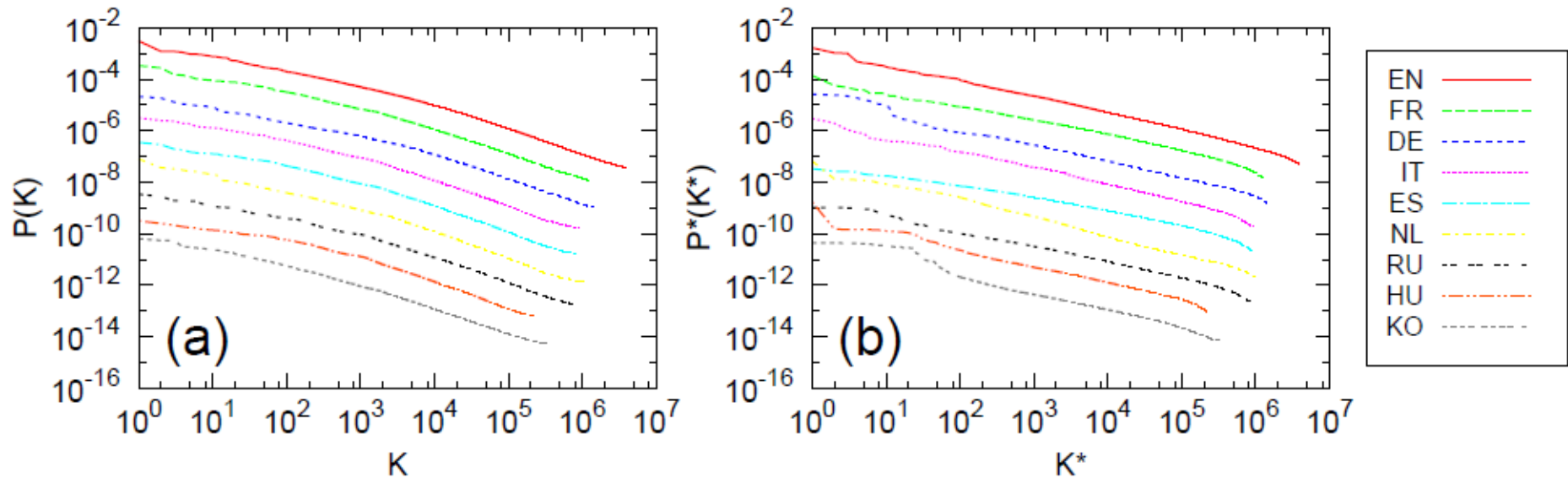
Similar PageRank and CheiRank in spite of various size and density

Wikipedia articles on KK* plane

Similar ranking structure in system level

What about individual article level?

Ranking of persons

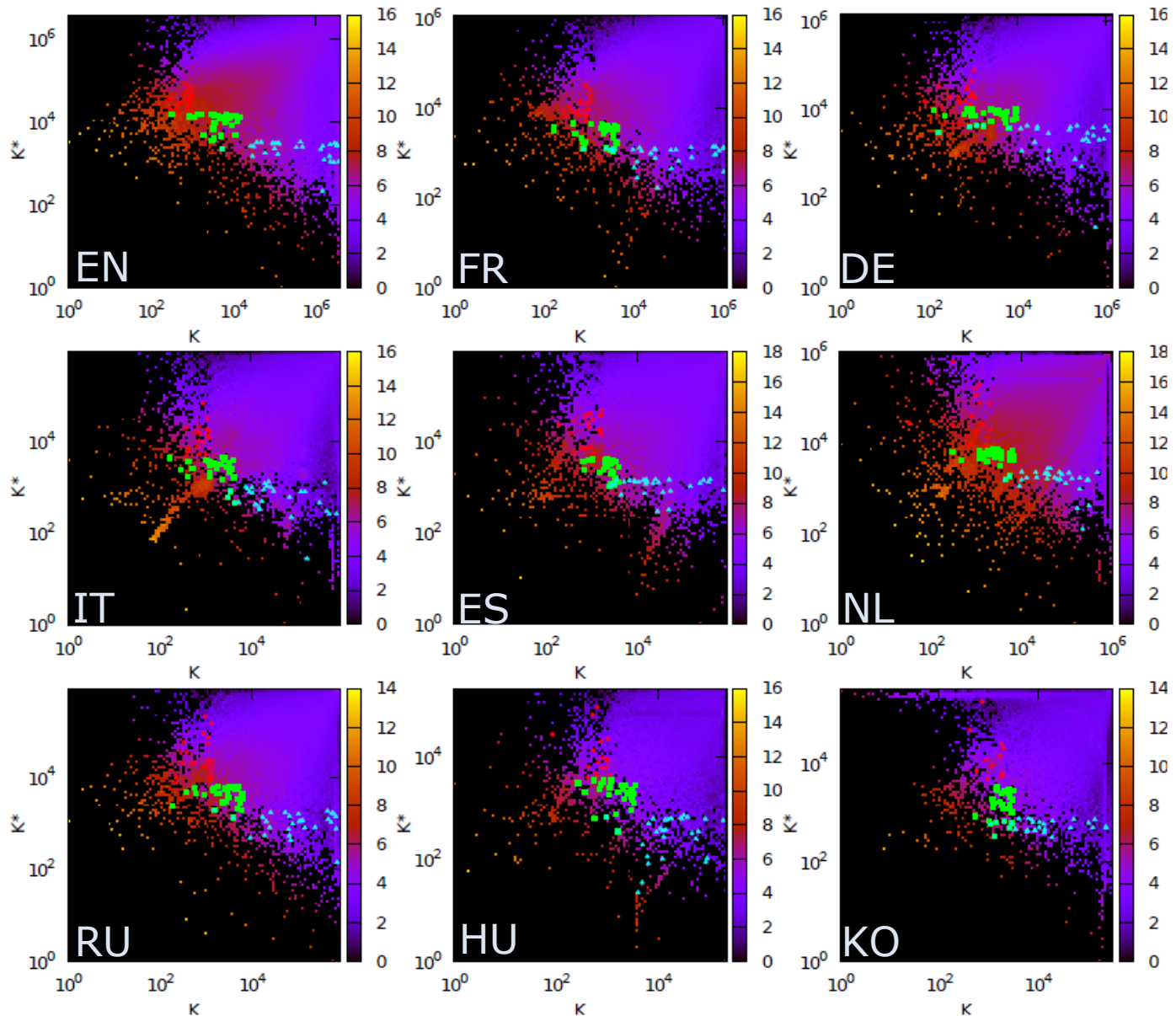|				
1 236 32469 35083 Napoleon	1 218 1449979 1664774 Roger Calmel (composer)	1 106 3050 3658 5401 Frank Sinatra
2 283 38322 2347 Carl Linnaeus	2 1075 2687885 1280698 C. H. Vijayashankar (minister of india)	2 127 3389 4056 1932718 Paul McCartney
3 285 16825 796999 George W. Bush	3 1113 2240193 1473998 Matt Kelley (writer,public speaker)	3 160 1928 4878 2049530 Michael Jackson
4 295 21985 215455 Barack Obama	4 1119 3567841 3850956 Alberto Cavallari (journalist)	4 174 5156 4756 13533 Steven Spielberg
5 310 20267 1747583 Elizabeth II	5 1223 1582025 1615115 Yury Chernavsky (producer, composer)	5 184 5422 2304 11884 Pope Pius XII
6 327 15690 353180 Jesus	6 1511 2134225 3289568 William Shakespeare (inventor) (busines	6 277 3322 6901 16564 Vladimir Putin
7 334 39411 16580 William Shakespeare	7 1516 29345 177409 Kelly Clarkson (singer)	7 453 9189 8310 9673 Mariah Carey
8 353 103484 8 Aristotle	8 1669 120229 242423 Park Ji-Sung (football)	8 492 3535 9644 1034539 John Kerry
9 447 23025 683626 Adolf Hitler	9 1796 59135 368500 Mithun Chakraborty (indian actor)	9 521 6432 9919 7080 Isaac Asimov
10 476 29578 1448 Bill Clinton	10 2226 1311114 3159294 Olga Sedakova	10 625 6576 10950 13541 Stephen King
11 477 17283 5304 Franklin D. Roosevelt	11 2275 100236 877707 Sara García (actress)	11 637 11151 4958 4127 Dolly Parton
12 588 31572 12817 Ronald Reagan	12 2304 5422 11884 Pope Pius XII	12 726 7728 12082 29324 Prince (musician)
13 678 39687 5759 George Washington	13 2438 335090 487907 Andy Kerr (Scottish politician)	13 770 12494 11157 151776 Robert Brown (botanist
14 698 86298 11415 Plato	14 2529 476868 341965 Joe-Max Moore (football)	14 786 4840 12641 16445 Vincent van Gogh
15 737 58221 12831 Richard Nixon	15 2597 30906 3346527 Josef Kemr (Czech actor)	15 838 8277 13204 2308127 Lady Gaga
16 746 54042 7 Abraham Lincoln	16 2635 26509 41129 Darius Milhaud (French composer)	16 842 8799 13261 40452 Beyoncé Knowles
17 753 20580 7606 Joseph Stalin	17 2652 1470479 3229694 Jan Crull, Jr. (advocate, filmmaker)	17 867 1115 13426 11881 Pope John Paul II
18 824 26717 16672 Winston Churchill	18 2725 3161209 3676176 Farshad Fotouhi (Computer Science)	18 893 6099 13684 2290127 Lord Byron
19 829 29222 1034114 John F. Kennedy	19 2745 3299175 3247402 Swaroop Kanchi (Indian film director)	19 914 838 13942 9348 Muhammad
20 830 37948 6900 Henry VIII of England	20 2871 2206140 1921527 Jacques Lancelot (Clarinet player)	20 965 6956 14432 1731878 Johnny Cash
21 838 13942 9348 Muhammad	21 2960 3204324 3840527 František Martin Pecháček (composer)	21 980 14599 13329 83036 Alice Cooper
22 846 77356 15044 Thomas Jefferson	22 3037 129156 182732 George Stephanopoulos (journalist)	22 981 4260 14613 23190 Catherine the Great
23 880 25488 94 Albert Einstein	23 3063 92429 3560521 Chano Urueta (mexican flim director)	23 985 9167 14648 2551927 14th Dalai Lama
24 898 25806 112 Alexander the Great	24 3097 2802095 3840407 Franz Pecháček (composer)	24 994 11717 14729 85014 Christina Aguilera
25 921 54975 345 Augustus	25 3118 107558 160722 Nicolae Iorga (Prime Min. of romania)	25 1031 6813 15017 9604 Marilyn Monroe
26 945 42393 2383 Charlemagne	26 3152 121491 2268820 Arnold Houbraken (Dutch Painter)	26 1048 3392 15228 4165 David Bowie
27 991 35181 8214 Karl Marx	27 3192 90134 202 August Derleth (novelist)	27 1059 3063 15328 22905 John McCain
28 996 18396 1385471 Charles Darwin	28 3296 47914 169446 Javier Solana (politician)	28 1081 1353 15496 974772 Bob Dylan
29 1020 69447 4850 Elizabeth I of England	29 3301 37508 3684 Drew Barrymore (actress)	29 1101 1722 15673 1548348 Johann Sebastian Bach
30 1024 45913 7758 Julius Caesar	30 3331 422398 202328 Kevin Bloody Wilson (singer)	30 1103 327 15690 353180 Jesus

Categories of top persons
by six active fields

(a) PageRank

(b) CheiRank

(c) 2DRank
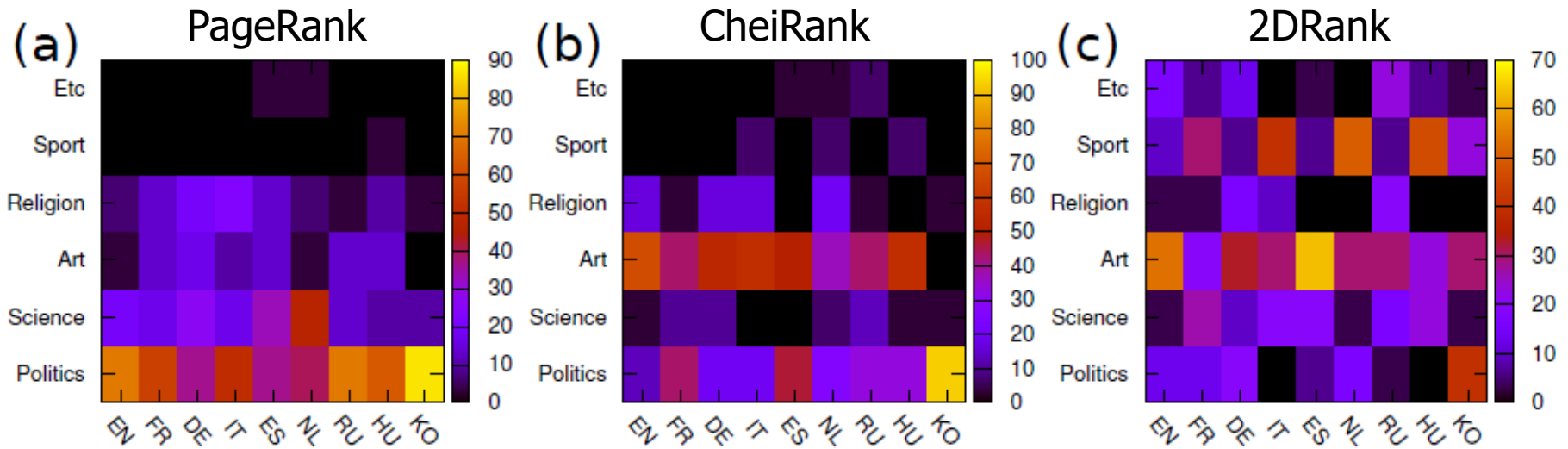
Categories of top persons by language

(a) PageRank

(b) CheiRank

(c) 2DRank
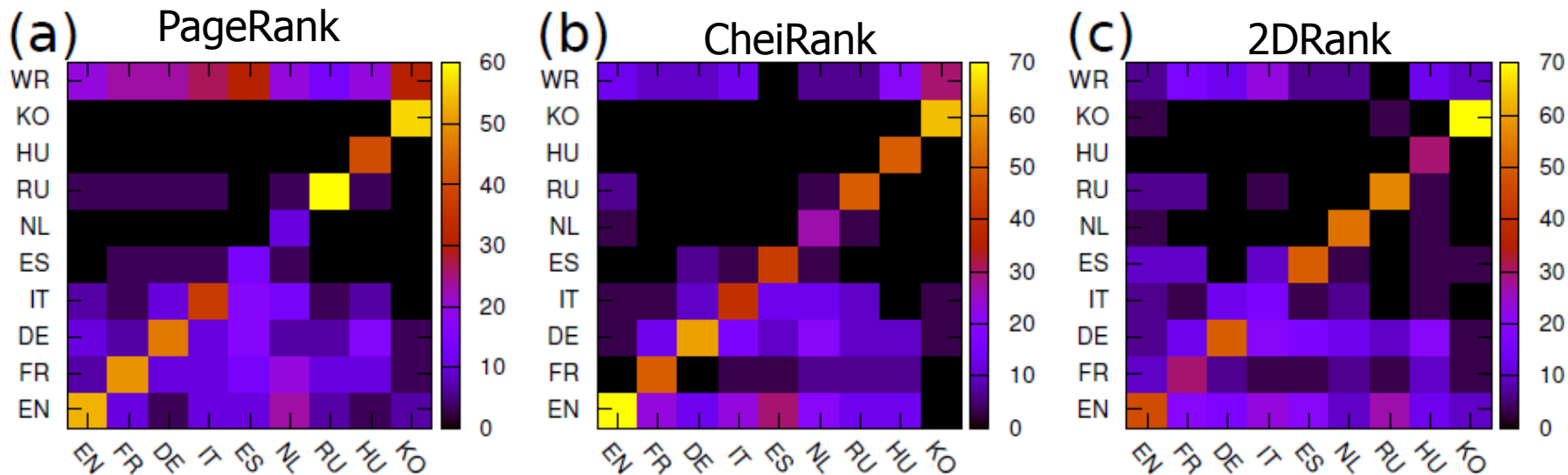
Dominant local persons but there are non-local persons

Top PageRank local persons for each edition

Edition	1st	2nd	3rd
EN	George W. Bush	Barack Obama	Elizabeth II
FR	Napoleon	Louis XIV of France	Charles de Gaulle
DE	Adolf Hitler	Martin Luther	Immanuel Kant
IT	Augustus	Dante Alighieri	Julius Caesar
ES	Charkes V, Holy Roman Emperor	Philip II of Spain	Francisco Franco
NL	Ernst Mayr	William I of the Netherlands	Beatrix of the Netherlands
RU	Peter the Great	Joseph Stalin	Alexander Pushkin
HU	Matthias Corvinus	Szentágothai János	Stephen I of Hungary
KO	Gojong of the Korean Empire	Sejong the Great	Park Chung-hee

$R_{EN,PageRank}$	Person	Field	Culture	Locality
1	Napoleon	Politics	FR	Non-local
2	Carl Linnaeus	Science	WR	Non-local
3	George W. Bush	Politics	EN	Local
4	Barack Obama	Politics	EN	Local
5	Elizabeth II	Politics	EN	Local
6	Jesus	Religion	WR	Non-local
7	William Shakespeare	Art	EN	Local
8	Aristotle	Science	WR	Non-local
9	Adolf Hitler	Politics	DE	Non-local
10	Bill Clinton	Politics	DE	Non-local

Is there any difference between local and non-local top persons?

Categories of top persons Local Vs NonLocal

Edition	N_{Local}	$\langle P(j)/k(j)_{out}\rangle_L$	$\langle P(j)/k(j)_{out}\rangle_{NL}$	$\langle k(i)_{in}\rangle_L$	$\langle k(i^*)_{in}\rangle_{NL}$
EN	16	1.53×10^{-8}	2.19×10^{-8}	5.9×10^3	4.3×10^3
FR	15	4.03×10^{-8}	5.56×10^{-8}	2.9×10^3	2.4×10^3
DE	14	4.28×10^{-8}	4.43×10^{-8}	2.7×10^3	2.9×10^3
IT	11	7.86×10^{-8}	8.44×10^{-8}	1.7×10^3	2.0×10^3
ES	4	5.66×10^{-8}	9.93×10^{-8}	2.0×10^3	1.5×10^3
NL	3	9.33×10^{-7}	4.24×10^{-7}	6.8×10^2	1.0×10^3
RU	18	7.27×10^{-8}	9.60×10^{-8}	1.9×10^3	1.8×10^3
HU	12	3.51×10^{-7}	3.19×10^{-8}	7.8×10^2	8.1×10^2
KO	17	1.71×10^{-7}	6.06×10^{-8}	5.8×10^2	3.3×10^2

Identifying global heroes

$$\Theta_{P,A} = \sum_E (31 - R_{P,E,A})$$

$R_{EN,PageRank}$	Person	Field	Culture	Locality
1	Napoleon	Politics	FR	Non-local
2	Carl Linnaeus	Science	WR	Non-local
3	George W. Bush	Politics	EN	Local
4	Barack Obama	Politics	EN	Local
5	Elizabeth II	Politics	EN	Local
6	Jesus	Religion	WR	Non-local
7	William Shakespeare	Art	EN	Local
8	Aristotle	Science	WR	Non-local
9	Adolf Hitler	Politics	DE	Non-local
10	Bill Clinton	Politics	DE	Non-local

Global heroes

| Rank	PageRank global heroes	Θ_{PR}	N_A	2DRank global heroes	Θ_{2D}	N_A
1st	Napoleon	259	9	Micheal Jackson	119	5
2nd	Jesus	239	9	Adolf Hitler	93	6
3rd	Carl Linnaeus	235	8	Julius Caesar	85	5
4th	Aristotle	228	9	Pope Benedict XVI	80	4
5th	Adolf Hitler	200	9	Wolfgang Amadeus Mozart	75	5
6th	Julius Caesar	161	8	Pope John Paul II	71	4
7th	Plato	119	6	Ludwig van Beethoven	69	4
8th	Charlemagne	111	8	Bob Dylan	66	4
9th	William Shakespeare	110	7	William Shakespeare	57	3
10th	Pope John Paul II	108	6	Alexander the Great	56	3

Conclusion

• Cross-edition analysis of article ranking in Wikipedia

• Dominant local but significant non-local even global

• Alternative way of cross-cultural study with Big data

Thank you!!!
But as I told you,
this is not the end!!!

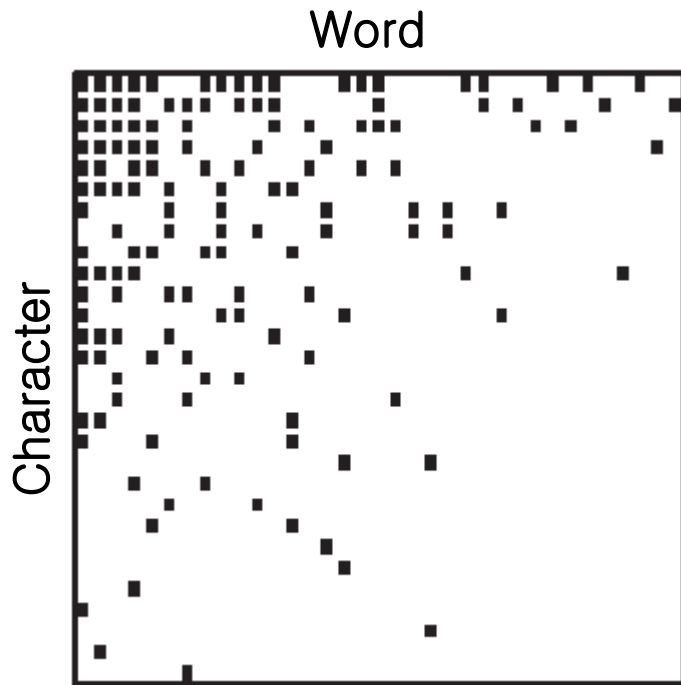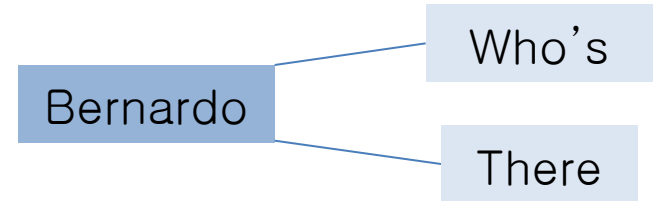# Analysis of character-word bipartite networks in Shakespeare's play
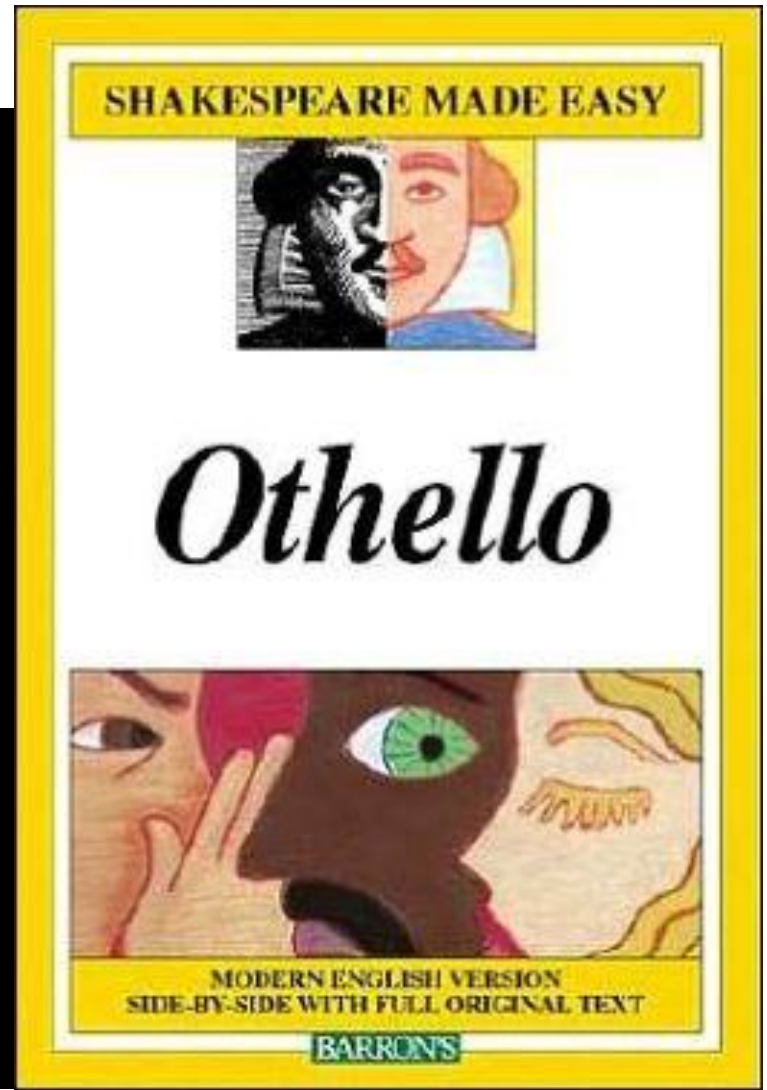
Character-Word interaction matrix

Bernardo. Who's there?

Francisco. Nay, answer me. Stand and unfold yourself.

Bernardo. Long live the King!

Bernardo — Who's
Bernardo — There

Word

Character

How these interactions are organized?

Considered systems : Shakespeare's play

OpenSourceShakespeare

SEARCH TEXTS [   ] GO

Plays ✦ Sonnets ✦ Poems ✦ Concordance ✦ Character Search ✦ Advanced Search ✦ About OSS

The Tragedy of Hamlet, Prince of Denmark (1600)

Scenes
(20 total)

Complete Text

Act I

Scene 1. Elsinore. A platform before the Castle.
Scene 2. Elsinore. A room of state in the Castle.
Scene 3. Elsinore. A room in the house of Polonius.
Scene 4. Elsinore. The platform before the Castle.
Scene 5. Elsinore. The Castle. Another part of the fortifications.

Act II

Scene 1. Elsinore. A room in the house of Polonius.
Scene 2. Elsinore. A room in the Castle.

Act III

Scene 1. Elsinore. A room in the Castle.
Scene 2. Elsinore. hall in the Castle.
Scene 3. A room in the Castle.
Scene 4. The Queen's closet.

Act IV

Scene 1. Elsinore. A room in the Castle.
Scene 2. Elsinore. A passage in the Castle.
Scene 3. Elsinore. A room in the Castle.
Scene 4. Near Elsinore.
Scene 5. Elsinore. A room in the Castle.
Scene 6. Elsinore. Another room in the Castle.
Scene 7. Elsinore. Another room in the Castle.

Act V

Characters
(34 total)

Click on a name to see all of that character's speeches

All
Ambassador
Bernardo, sentinel
Claudius, King of Denmark
Cornelius, courtier
Father's Ghost, Ghost of Hamlet's Father
First Clown
First Player
Fortinbras, Prince of Norway
Francisco, a soldier
Gentleman, courtier
Gentlemen
Gertrude, Queen of Denmark and mother to Hamlet
Guildenstern, courtier
Hamlet, son of the former king and nephew to the present king
Horatio, friend to Hamlet
Laertes, son to Polonius
Lord
Lucianus
Marcellus, Officer
Messenger
Norwegian Captain
Ophelia, daughter to Polonius
Osric, courtier
Player King
Player Queen
Polonius, Lord Chamberlain
Priest

OpenSourceShakespeare

SEARCH TEXTS [   ] GO

Plays ✦ Sonnets ✦ Poems ✦ Concordance ✦ Character Search ✦ Advanced Search ✦ About OSS

The Tragedy of Othello, Moor of Venice (1604)

Scenes
(15 total)

**Complete Text**

Act I

Scene 1. Venice. A street.
Scene 2. Another street.
Scene 3. A council-chamber.

Act II

Scene 1. A Sea-port in Cyprus. An open place near the quay.
Scene 2. A street.
Scene 3. A hall in the castle.

Act III

Scene 1. Before the castle.
Scene 2. A room in the castle.
Scene 3. The garden of the castle.
Scene 4. Before the castle.

Act IV

Scene 1. Cyprus. Before the castle.
Scene 2. A room in the castle.
Scene 3. Another room in the castle.

Act V

Scene 1. Cyprus. A street.
Scene 2. A bedchamber in the castle: DESDEMONA in bed asleep;

Characters
(26 total)

Click on a name to see all of that character's speeches

All
Bianca, Mistress to Cassio
Brabantio, A senator of Venice
Cassio, Othello's lieutenant
Clown, Servant to Othello
Desdemona, Daughter to Brabantio and wife to Othello
Duke of Venice
Emilia, Wife to Iago
First Gentleman
First Musician
First Officer
First Senator, A senator of Venice
Fourth Gentleman
Gentleman
Gratiano, Brother to Brabantio
Herald
Iago, Othello's ancient (?)
Lodovico, Kinsman to Brabantio
Messenger
Montano, Othello's predecessor in the government of Cyprus
Othello, A noble Moor in the service of the Ventian state
Roderigo, A Venetian gentleman
Sailor
Second Gentleman
Second Senator, A senator of Venice
Third Gentleman

Zipf pattern
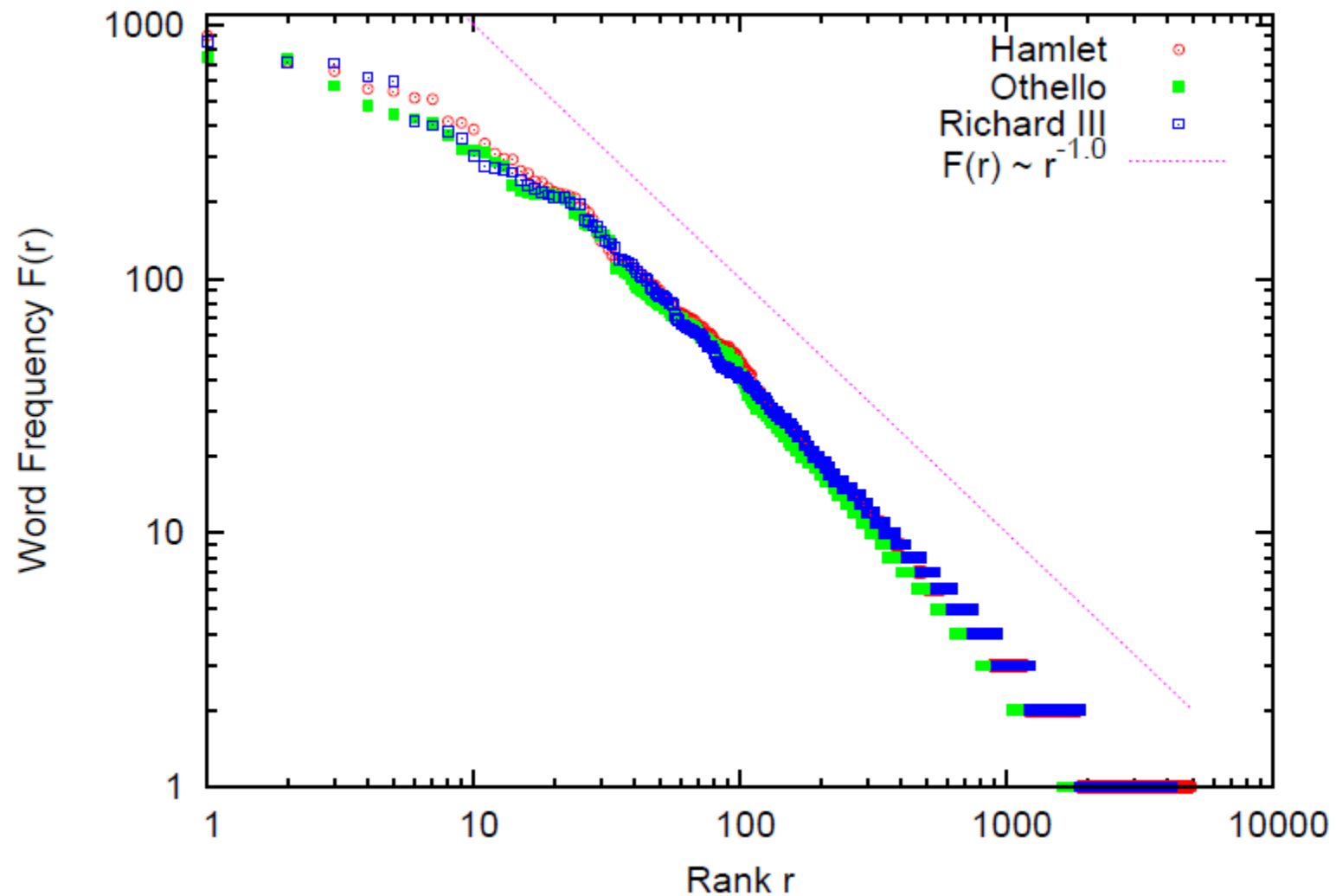
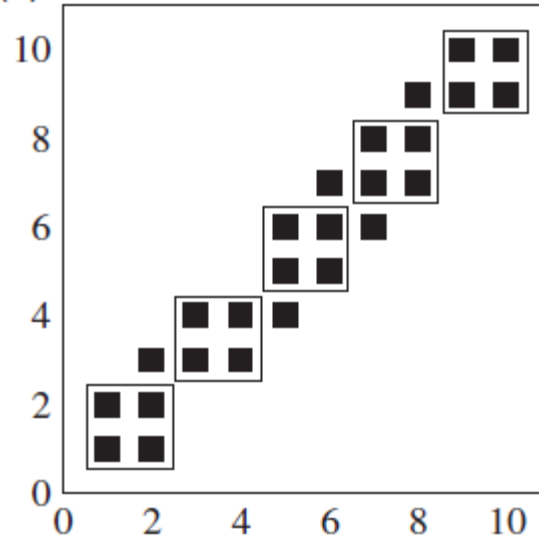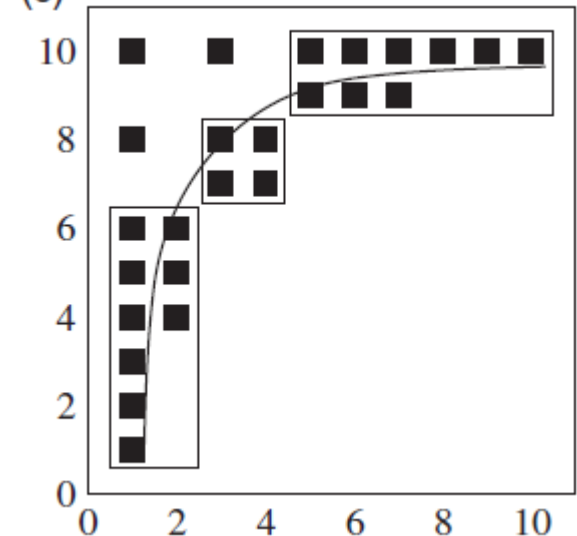# Nestedness in ecology
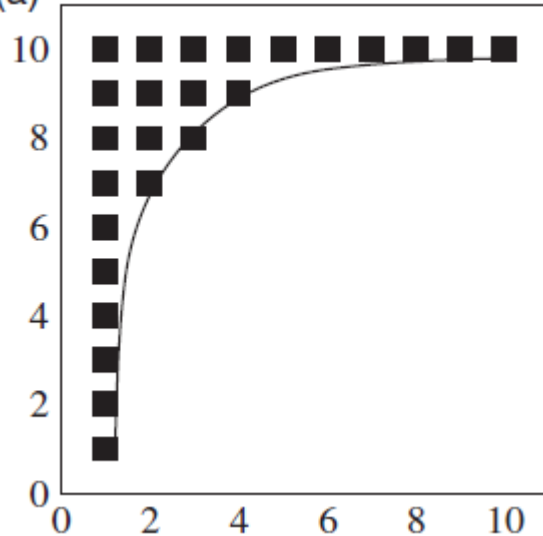
(a) Nested (b) Modular (c) Nested and modular

Nestedness in Character-verb interactions