IFJ PAN – NO4 Seminar – 2025-10-23

Science of Complexity ⇄ Artificial Intelligence

Natural language

An interdisciplinary field of research that asks how, from simple rules and numerous mutual non-linear interdependencies, organized behaviors and structures spontaneously emerge at a level involving many interacting elements, which often bring new qualities

(Nobel 2021)

-> More is different, Philip W. Anderson, Science 1972

Our natural language is eminently like that

Artificial intelligence is an interdisciplinary field - based mainly in computer science – dealing with the creation of systems aimed at learning, reasoning, perception, decision-making, and the use of natural language

It is artificial, but how intelligent is it?

is highly energy-intensive

(Nobel 2024)

Stanisław Drożdż – Zakład Teorii Systemów Złożonych IFJ PAN

Based on collaboration with Jarosław Kwapień and Tomasz Stanisz

Energy Efficiency: Human Brain vs Artificial Intelligence

A comparative look at power consumption and efficiency

The Human Brain: Nature's Supercomputer

- Power consumption: ~20 watts (like a dim light bulb)
- Energy per day: ~0.48 kWh
- ~86 billion neurons, ~10^{15} operations per second
- Efficiency: ~10^{-14} joules per operation
- Remarkably energy-efficient

Modern AI Models

- Large models (e.g., GPT-4) require tens of megawatts to train
- Inference (use phase): kilowatts per data center cluster
- One text generation: hundreds of joules
- Training: millions of kWh (≈ annual energy of hundreds of homes)

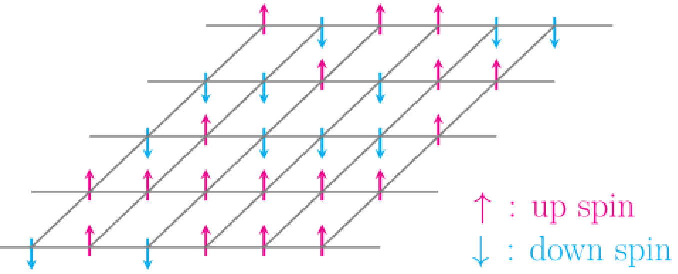
- Level | Human Brain | AI (Large Model) | Ratio
- ------|--------------|------------------|--------
- Per Operation | 10^{-14} J/op | 10^{-9} – 10^{-8} J/op | 10^5 – 10^6× less efficient
- Per Task | few joules | 10^3 – 10^5 joules | 10^3 – 10^5× less efficient

- -> AI systems consume 3–6 orders of magnitude more energy for comparable tasks.

Why the Huge Gap?

- Hardware difference: Brain = analog & adaptive; GPUs = digital & rigid
- Parallelism: Brain has massive parallelism (10^{15} ops/sec)
- Learning style: Brain learns continuously; AI retrains frequently
- Precision: Digital computation uses unnecessarily high precision

From Ising model to contemporary neural networks and Large Langauage Models (LLMs)



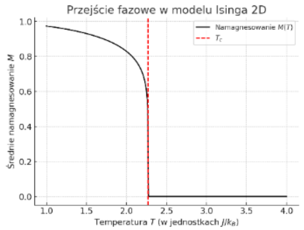
Ising Model
(1920s) ->
System of spins

Hopfield Model
(1982) ->
Associative memory

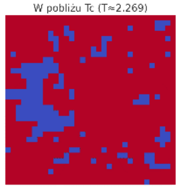
Boltzmann Machines
(1985) ->
Probabilistic networks

Deep leraining
(2000s +) ->
Deep networks
Inspired by
statistical physics

LLMs
(2018+)
Transformers
Language models
Billions of parameters

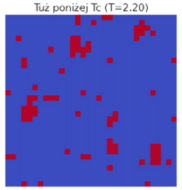


System evolves towards minimum energy states that correspond to stable configurations
(Ising – magnetic domains, Hopfield – memory patterns



Ising Model

(Wilhelm Lenz - 1920)



$$E = -\sum_{i<j} J_{ij} s_i s_j - \sum_i h_i s_i$$

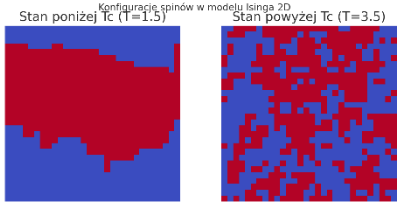
(Spin glasses)

Hopfield (associative) Network

(Nobel 2024 – in physics)

$$E = -\frac{1}{2}\sum_{i \neq j} w_{ij} s_i s_j + \sum_i \theta_i s_i$$

$$w_{ij} = \frac{1}{n}\sum_{\mu=1}^{p} \xi_i^{\mu} \xi_j^{\mu}$$



Spin = ±1

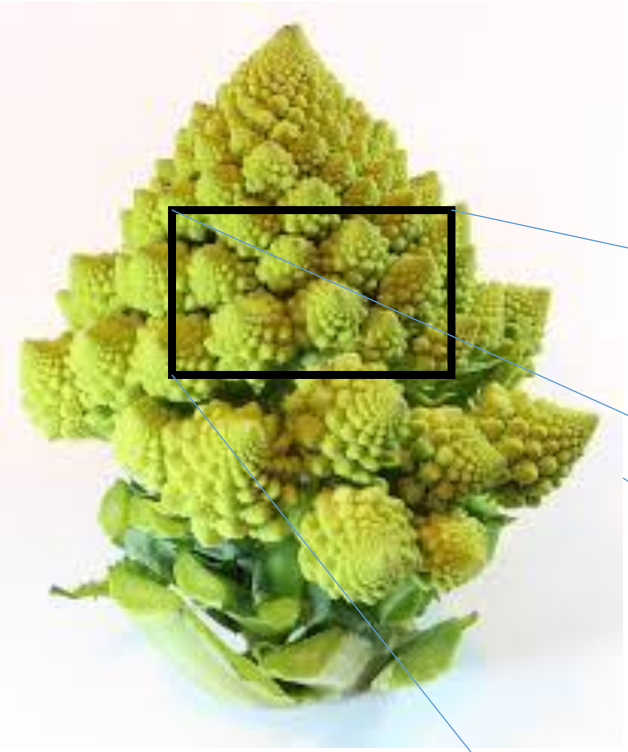
(activation/no activation)

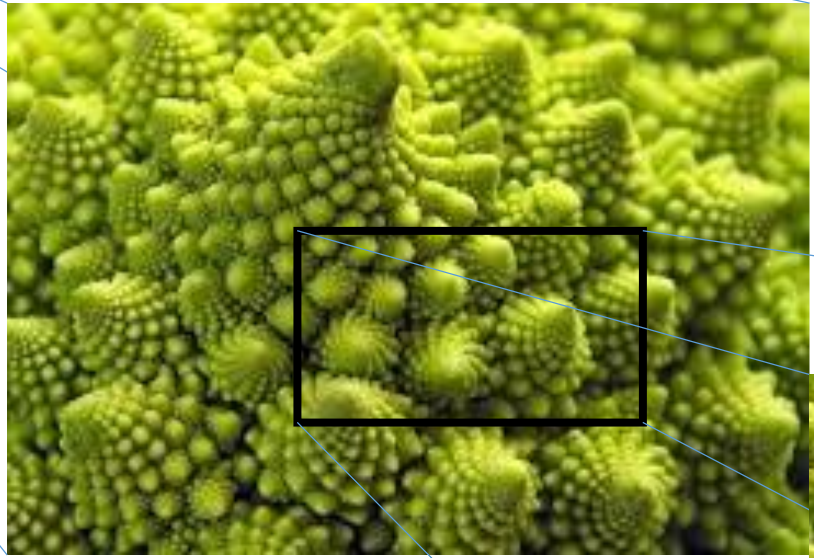
Strength/Wright of connections

(trained according to Hebbs rules
- neurons that fire together, wire together)

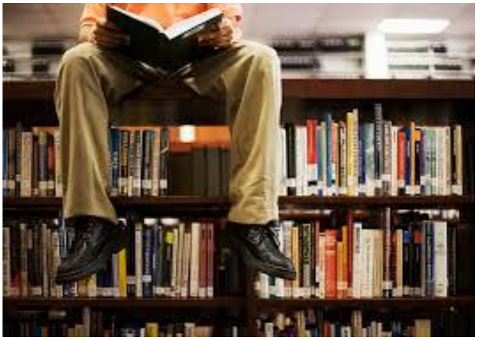
‚external field'

Romanesco broccoli





Fractal



Searching for the Laws of Natural Language

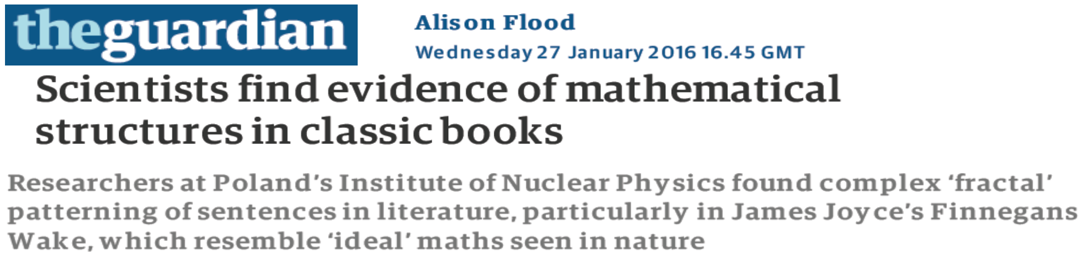
- do they exist ?

Why does language exist? How did it begin?
How do its different variants relate to each other?
Was there a single, common "proto-language"?
And what are its basic characteristics?

Serves interpersonal communication, storing and transferring knowledge, developing abstract thinking, shaping identity and culture, cooperation and organization
-> it is, above all, natural language that makes people human

Is there a mathematical law describing the characteristics of language?

Complexity of Natural Language

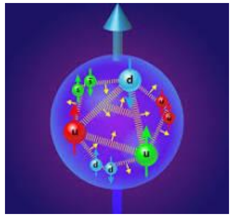

theguardian

Alison Flood

Wednesday 27 January 2016 16.45 GMT

Scientists find evidence of mathematical structures in classic books

Researchers at Poland's Institute of Nuclear Physics found complex 'fractal' patterning of sentences in literature, particularly in James Joyce's Finnegans Wake, which resemble 'ideal' maths seen in nature



Three Quarks for Muster Mark!

Zipf's law

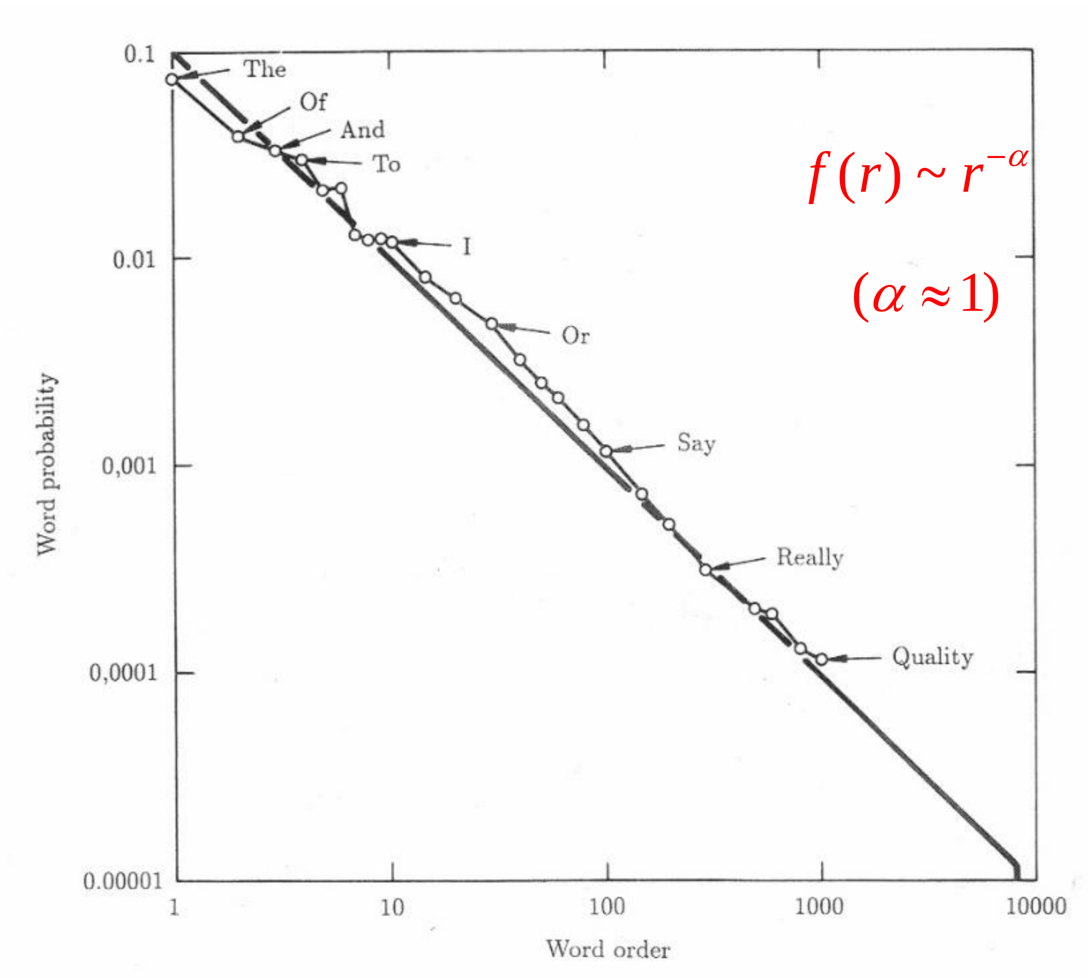
George K. Zipf,
Human behavior and the principle of least effort,
Addison-Wesley, Cambridge, 1949

J.B. Estoup,
Gammes Stenographiques,
Institute Stenographique de France,
Paris, 1916

Zipf-Mandelbrot:

$$f(r) \sim (r+C)^{-\alpha}$$

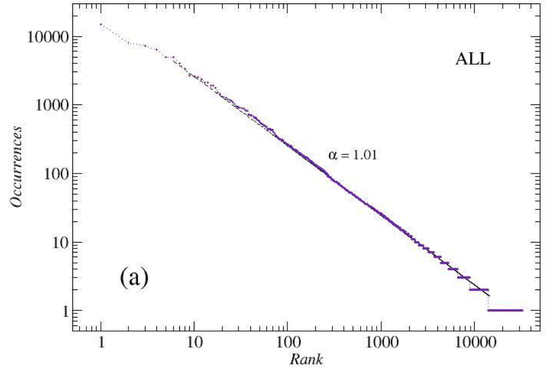
Original Zipf's plot for „Ulysses”



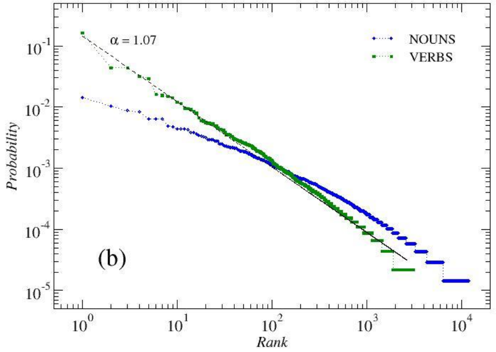
-α

$$f(r) \sim r^{-\alpha}$$

$$(\alpha \approx 1)$$



Different parts of speech



J. Kwapień, S. Drożdż / Physics Reports 515 (2012) 115–226

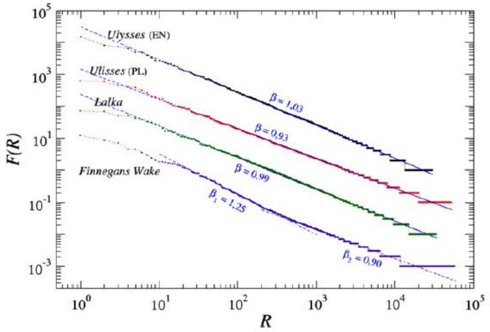
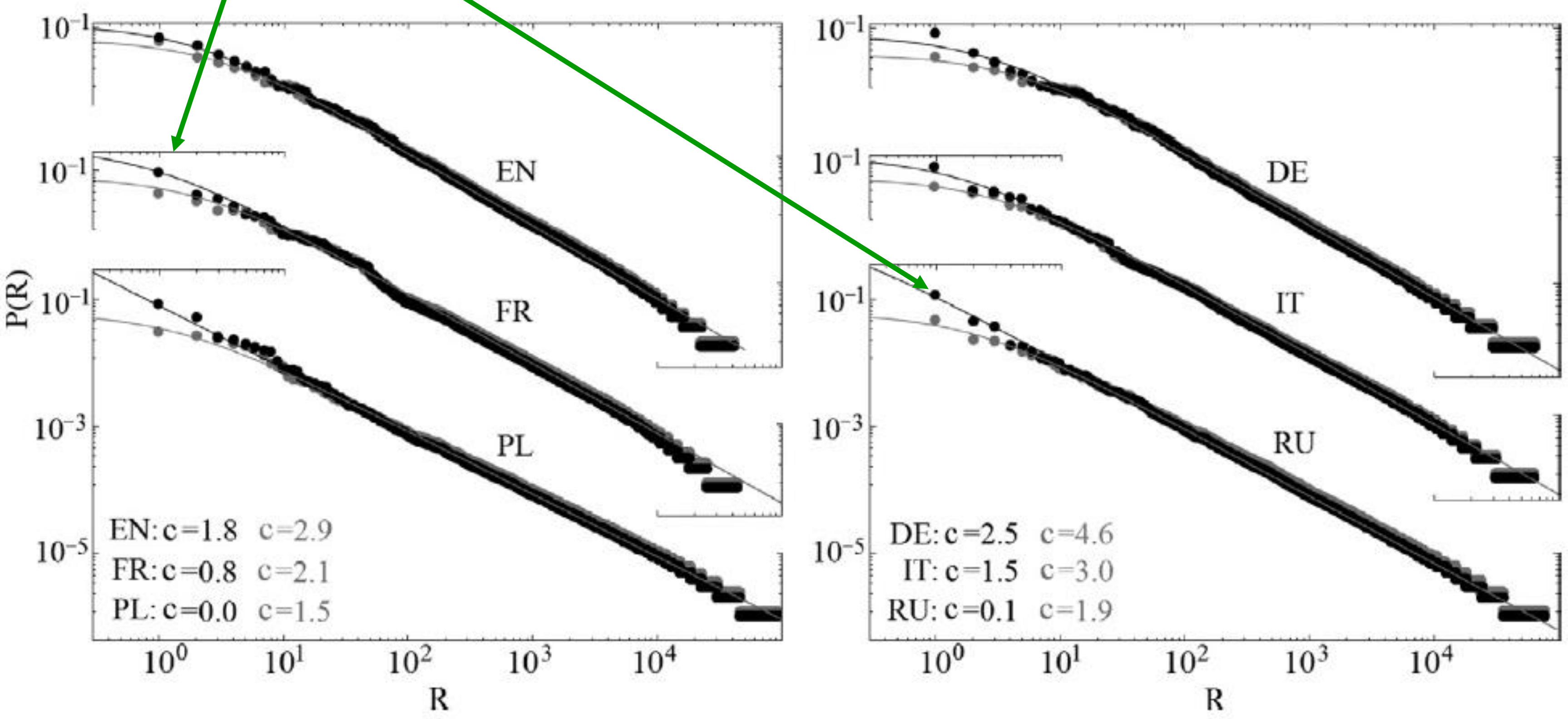


Fig. 63. Frequency-rank plots for the English and Polish versions of "Ulysses" by J. Joyce, for the original Polish text of "Lalka" ("The Doll") by B. Prus, and for the original English text of "Finnegans Wake" also by J. Joyce. Power-law functions defined by the exponent β were fitted to the empirical data. In the case of "Finnegans Wake" fitting a single power-law function is impossible, therefore the corresponding distribution must be characterized by two exponents: β_1 and β_2.

J. Kwapień & S. Drożdż
Physical approach to complex systems
Physivcs Reports 515 (2012)

Adding punctuation to the ‚Zipf' analysis even improves scaling



$$f(r) \sim (r + C)^{-\alpha}$$

Fig. 3. The Zipf–Mandelbrot distribution fitted to the occurrence-probability distributions for corpora representing different European languages. The two following cases are distinguished: the words without the punctuation marks (grey) and the words with the punctuation marks (black). For each language and for both cases, the corresponding value of the best-fitted parameter *c* is given.

A. Kulig, J. Kwapień, T. Stanisz, S. Drożdż, *In narrative texts punctuation marks obey the same statistics as words*
Information Sciences 375 (2017) 98-113

Quantifying patterns of punctuation in modern Chinese prose

Cite as: Chaos 35, 023155 (2025); doi: 10.1063/5.0248520
Submitted: 12 November 2024 · Accepted: 4 February 2025 ·
Published Online: 21 February 2025

Michał Dolina,[1] Jakub Dec,[1] Stanisław Drożdż,[1,2,a)] Jarosław Kwapień,[2] Jin Liu,[3] and Tomasz Stanisz[2]

AFFILIATIONS

[1]Faculty of Computer Science and Telecommunications, Cracow University of Technology, ul. Warszawska 24, 31-155 Kraków, Poland
[2]Complex Systems Theory Department, Institute of Nuclear Physics, Polish Academy of Sciences, ul. Radzikowskiego 152, 31-342 Kraków, Poland
[3]School of Modern Languages, Georgia Institute of Technology, Atlanta, Georgia 30332, USA

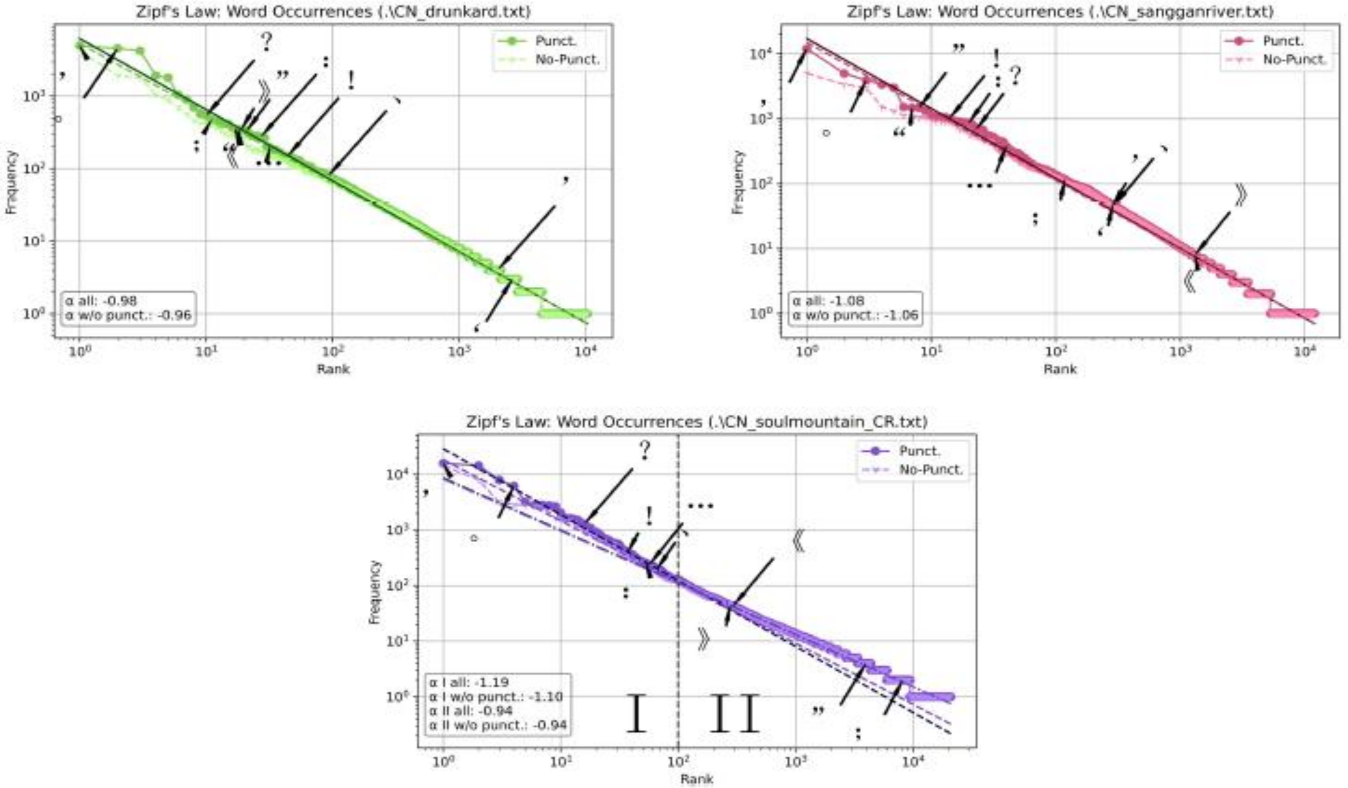


FIG. 1. Rank-frequency distributions of words (crosses) and words + punctuation marks (full circles) for the three Chinese books: *The Drunkard* (top left), *The Sun Shines over the Sanggan River* (top right), and *The Soul Mountain* (bottom). The distributions are fitted with a power-law function whose scaling index γ is given explicitly in each panel.

Quantifying distribution of inter-punctuation intervals

Geometric (Bernoulli)

$$1 - F(k) = (1-p)^k$$

Complementary cumulative distribution

Number of trials

Probability of ‚success' in one trial

generalizing with β

Weibull (discrete)

$$1 - F(k) = (1-p)^{k^\beta}$$

Hazard function:

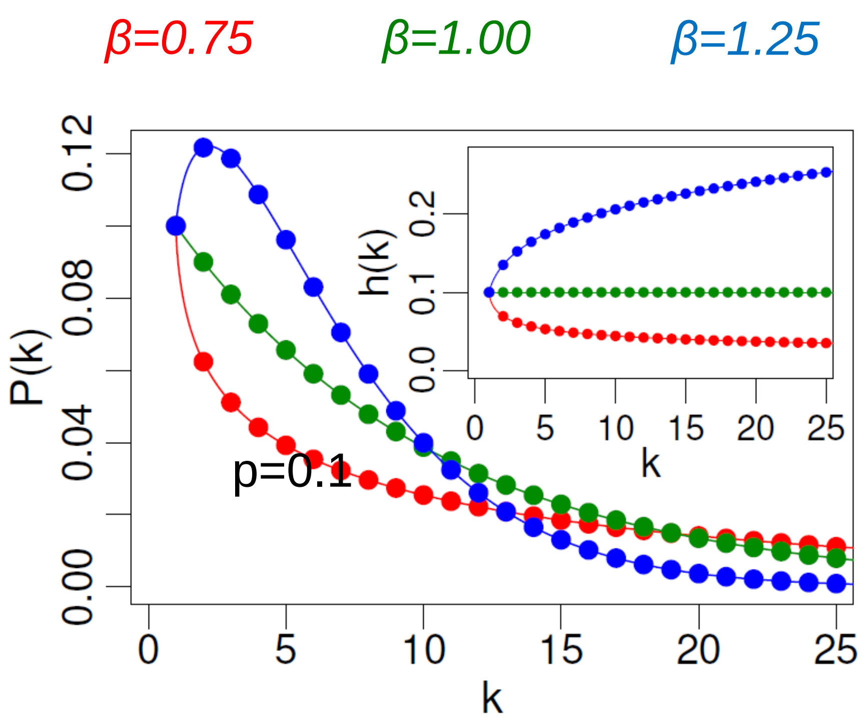
$$h(k) = \frac{P(k)}{1 - F(k-1)} = 1 - (1-p)^{k^\beta - (k-1)^\beta} .$$

$P(k)$ denoting the probability mass function

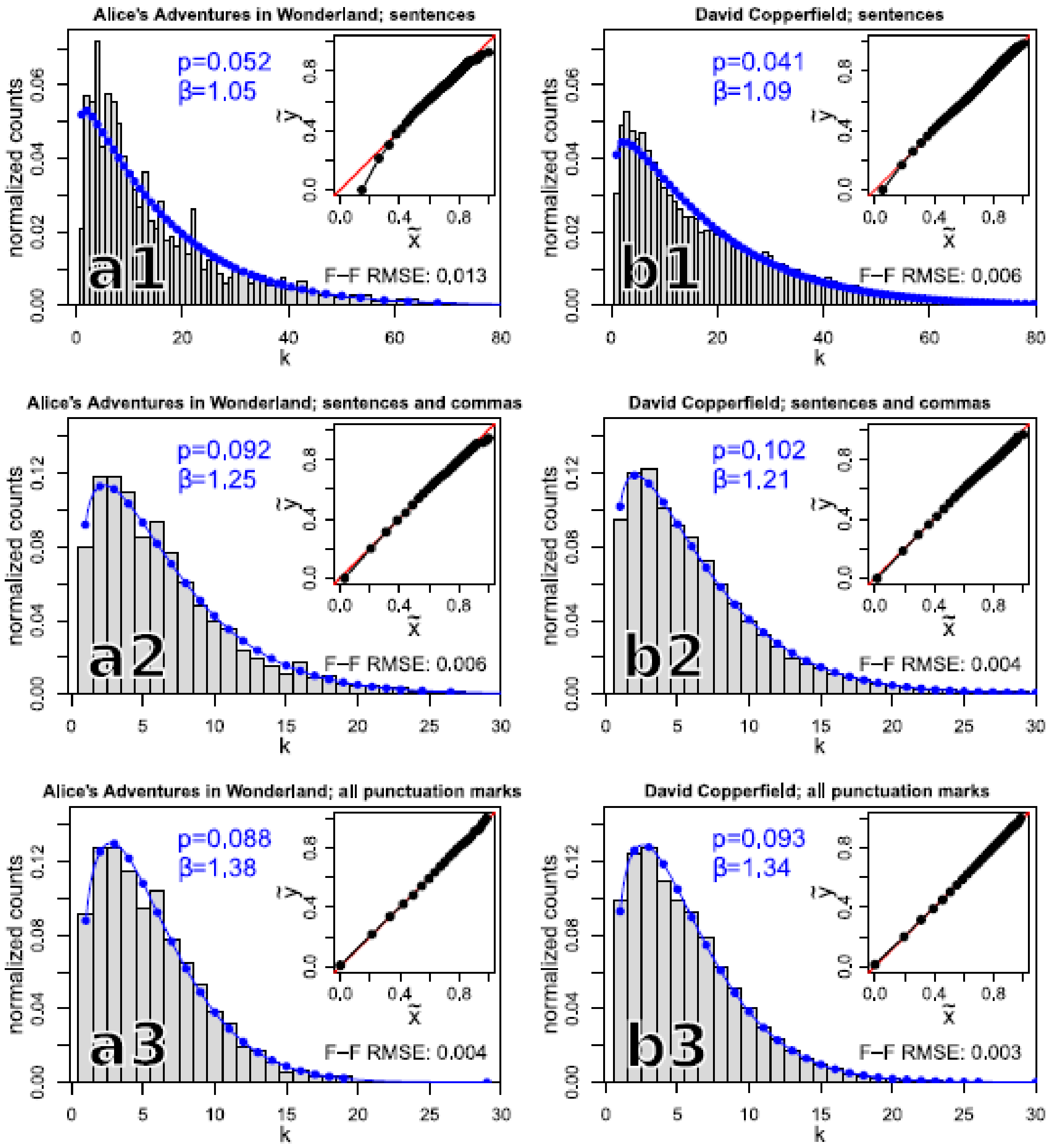
$$f(x;\beta,p) = \begin{cases} \frac{\beta}{p}\left(\frac{x}{p}\right)^{\beta-1} e^{-(x/p)^\beta} & \text{dla } x \geq 0 \\ 0 & \text{dla } x < 0 \end{cases}$$

$$h(x;\beta,p) = \frac{\beta}{p}\left(\frac{x}{p}\right)^{\beta-1}$$

Continuous representation



two books:
distribution of iterpunctuation intervals in three variants
- a1: sentence ending punctuation
- a2: sentence ending + commas
- a3: all punctuation



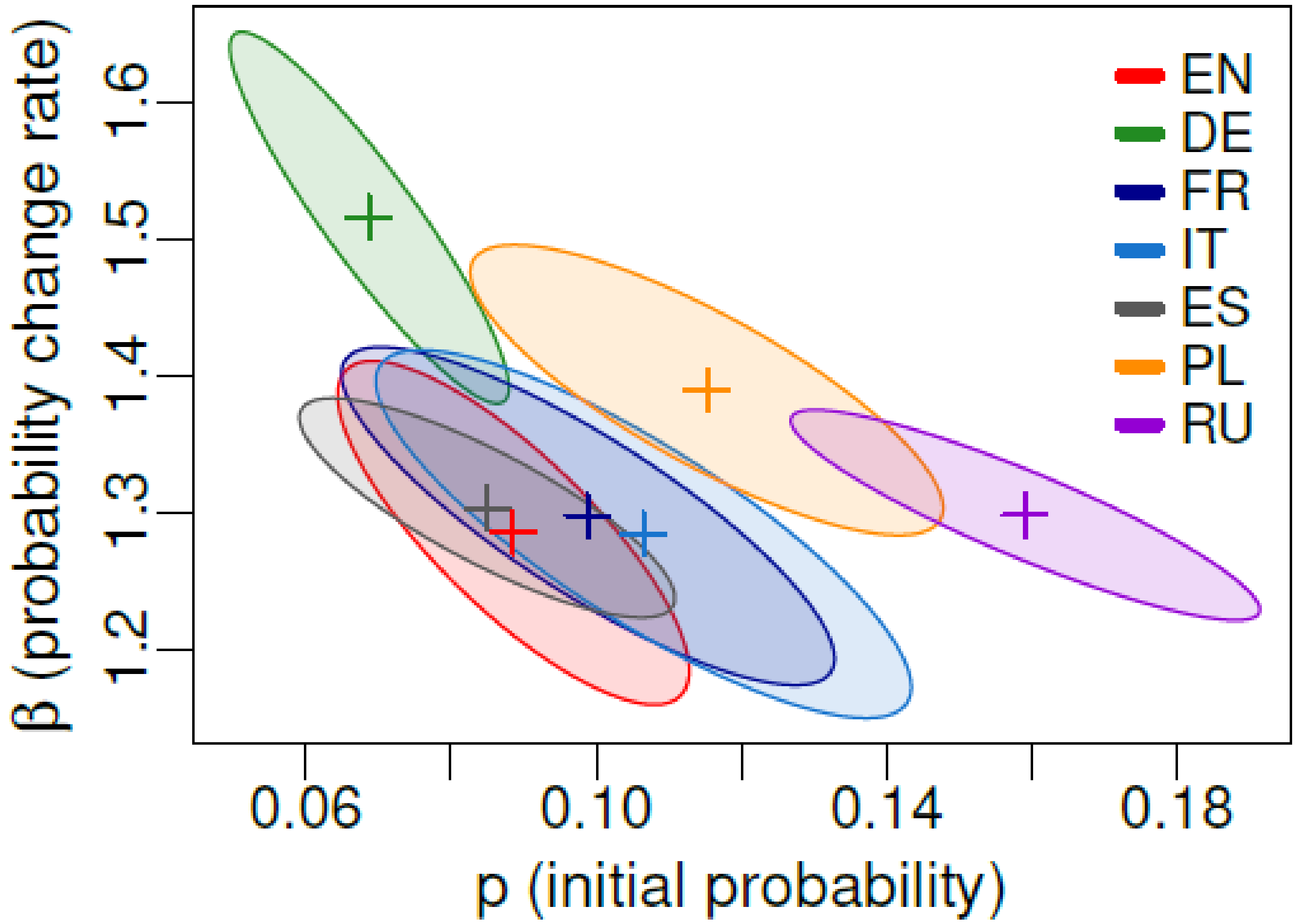
Best (very good) fit in terms of the Weibull distribution is obtained when all punctuation is included

T. Stanisz, S. Drożdż, J. Kwapień
Universal versus system-specific features of punctuation usage patterns in major Western languages
Chaos, Solitons and Fractals 168 (2023) 113168

The Weibull distribution provides a very good representation of the distribution of distances between punctuation marks universally for all 7 major Western languages

240 books in 7 languages

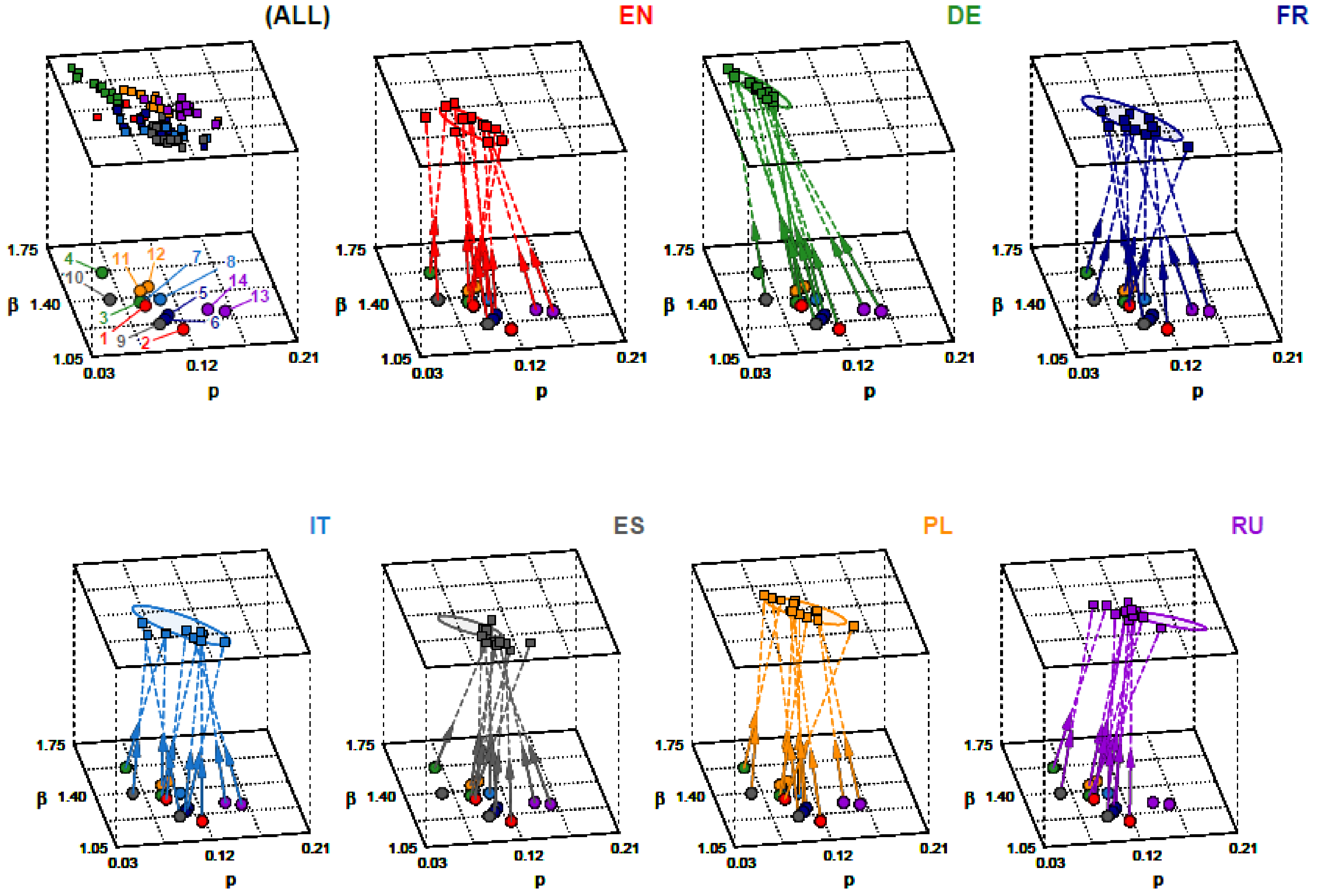
The values of the p – β parameters however, differ significantly and turn out to be specific to languages



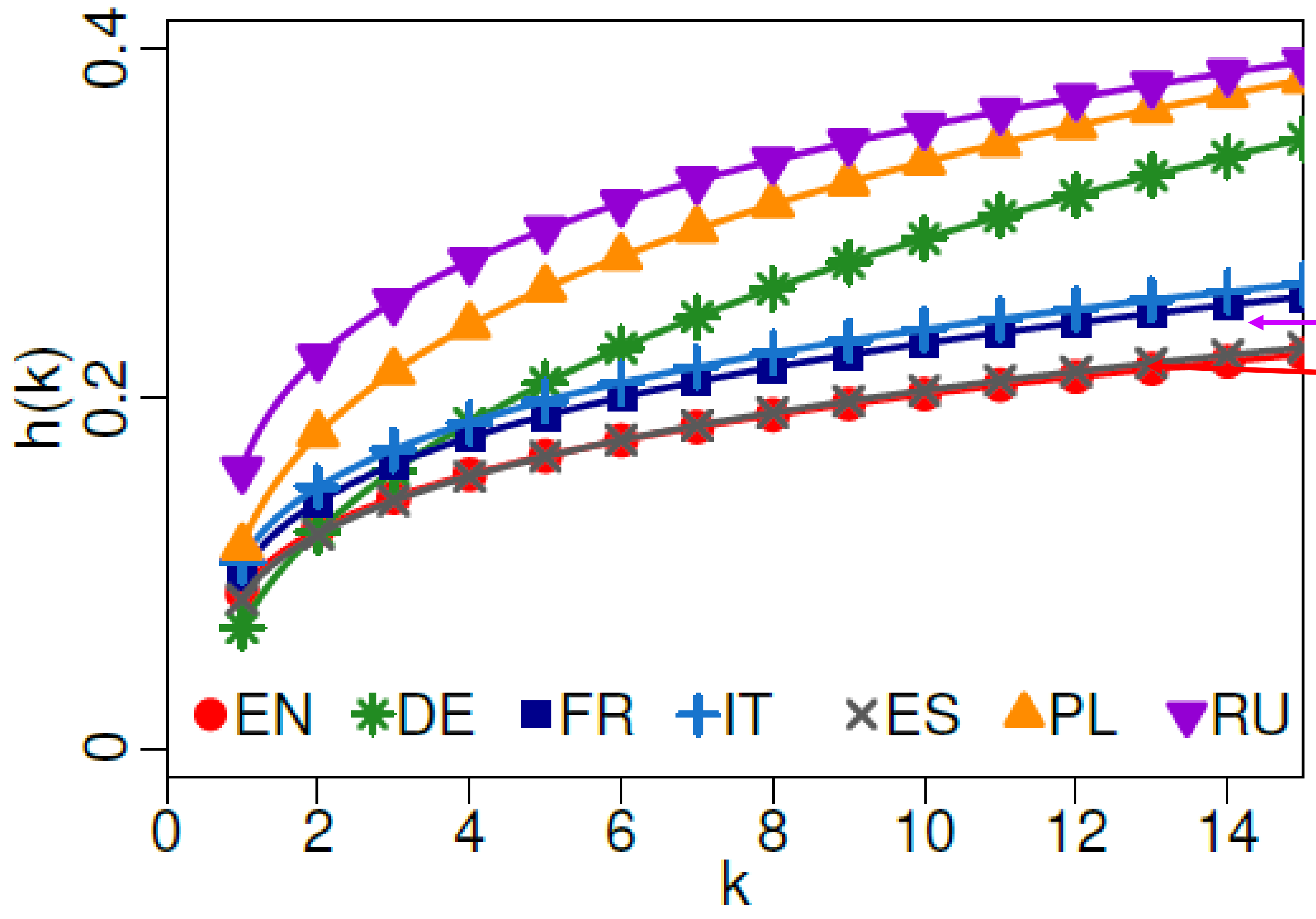
Characteristically:
Translations navigate the p and β towards the ellipsoids of the target languages

The set of texts consists of 14 books and their translations:
2 books in each of the 7 studied languages are translated into 6 remaining languages

The dots on the lower plane mark the values of p and β of the texts in their original languages, and the squares on the upper plane give p and β for the same texts translated into the target language



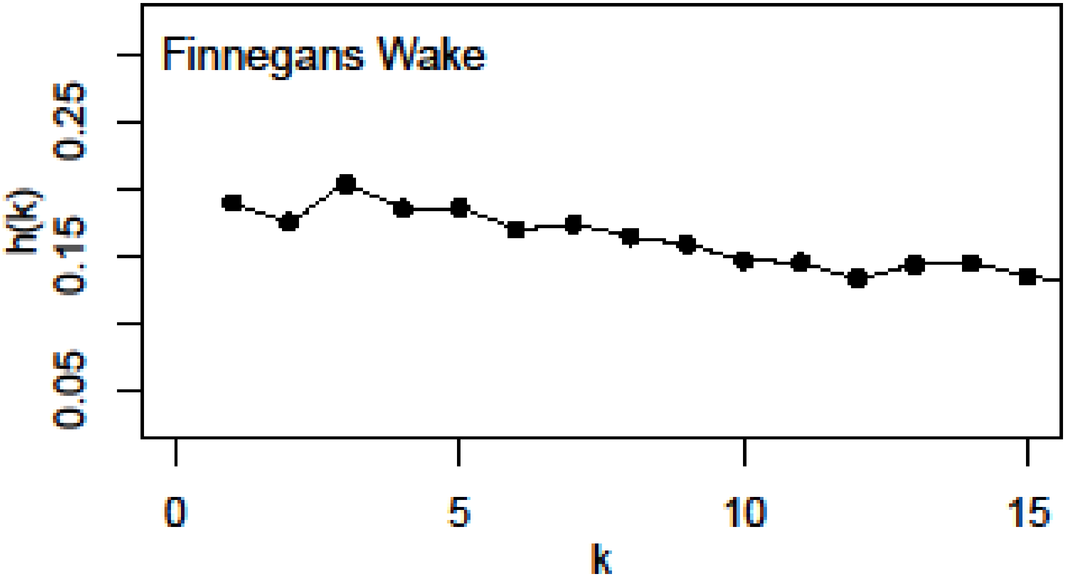
Hazard function: $$h(k) = \frac{P(k)}{1 - F(k-1)} = 1 - (1-p)^{k^{\beta} - (k-1)^{\beta}} .$$

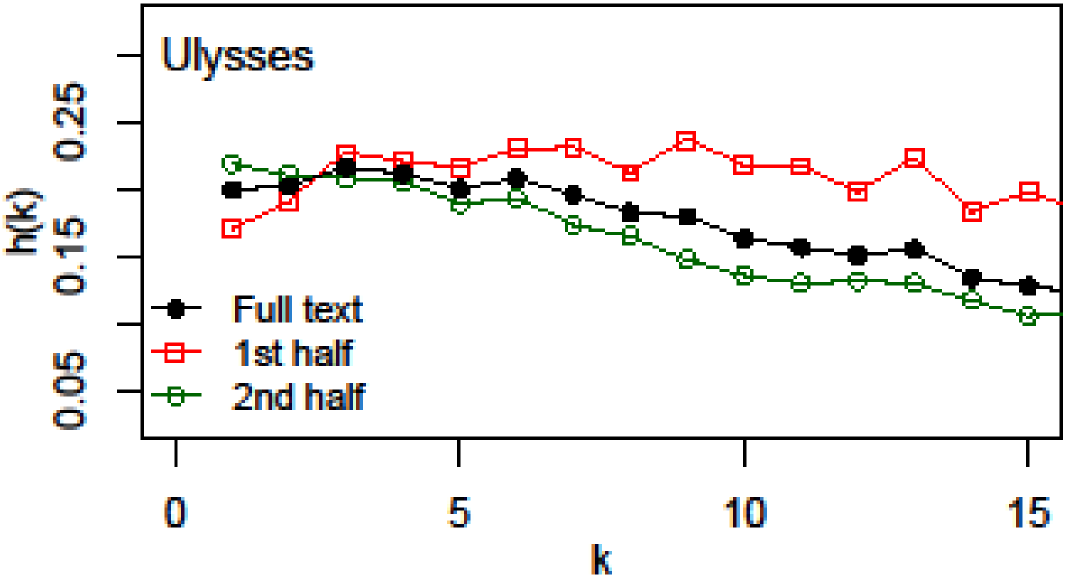


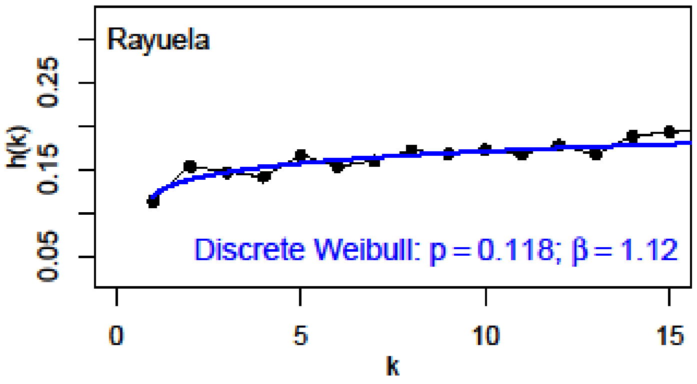
$$h(x;\beta,p) = \frac{\beta}{p}\left(\frac{x}{p}\right)^{\beta-1}$$ - Continuous representation

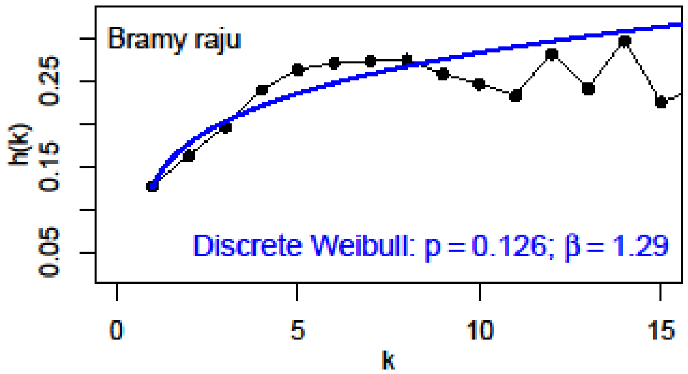
English and Spanish are the ‚least hazardous'

Hazard functions: ‚exotic' books









Statistics of punctuation in experimental literature—The remarkable case of Finnegans Wake by James Joyce

Cite as: Chaos 34, 083124 (2024); doi: 10.1063/5.0203530
Submitted: 14 February 2024 · Accepted: 5 July 2024 ·
Published Online: 20 August 2024

View Online Export Citation CrossMark

Tomasz Stanisz,[1] Stanisław Drożdż,[1,2,a)] and Jarosław Kwapień[1]

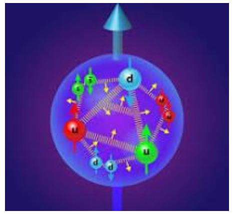
Article

Punctuation Patterns in *Finnegans Wake* by James Joyce Are Largely Translation-Invariant

Krzysztof Bartnicki [1], Stanisław Drożdż [2,3,*], Jarosław Kwapień [2] and Tomasz Stanisz [2]

1 Independent Researcher, 53-201 Wrocław, Poland
2 Complex Systems Theory Department, Institute of Nuclear Physics, Polish Academy of Sciences, 31-342 Kraków, Poland
3 Faculty of Computer Science and Telecommunications, Cracow University of Technology, 31-155 Kraków, Poland

Three Quarks for Muster Mark!



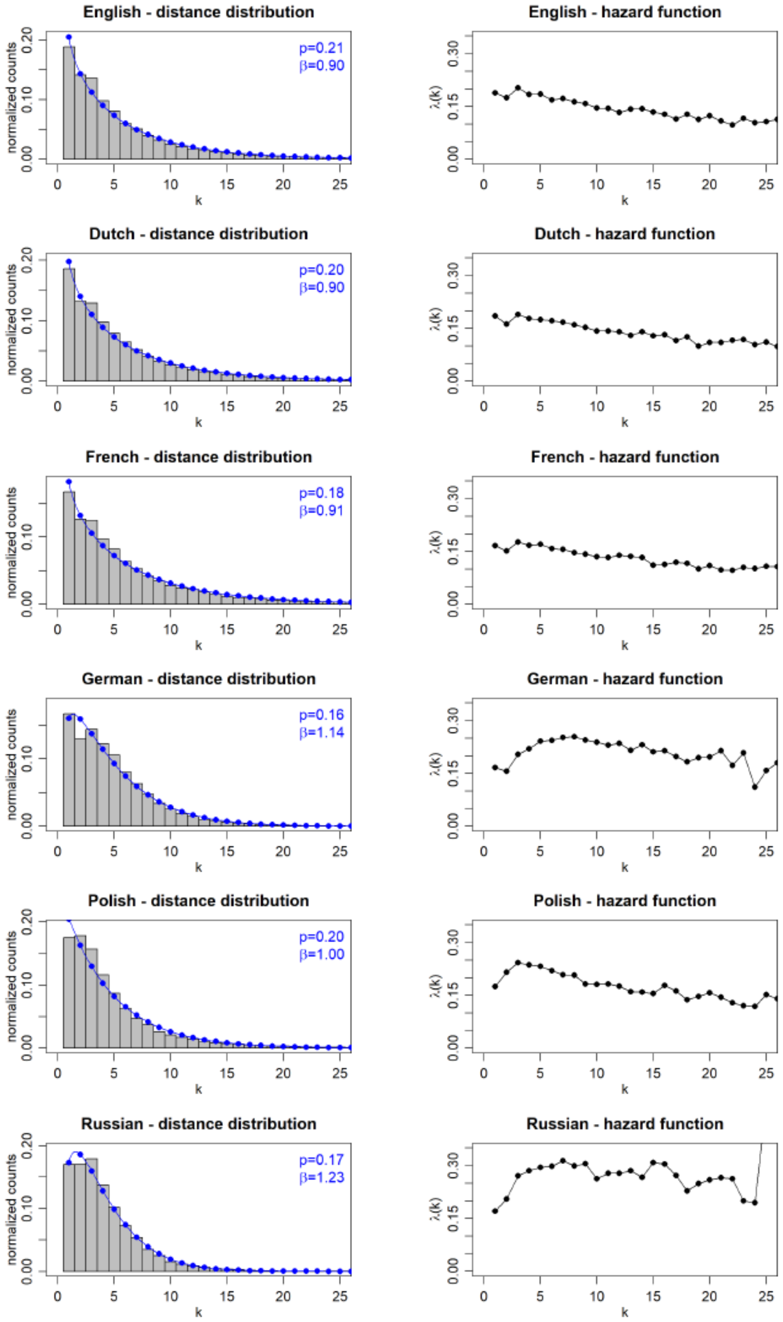
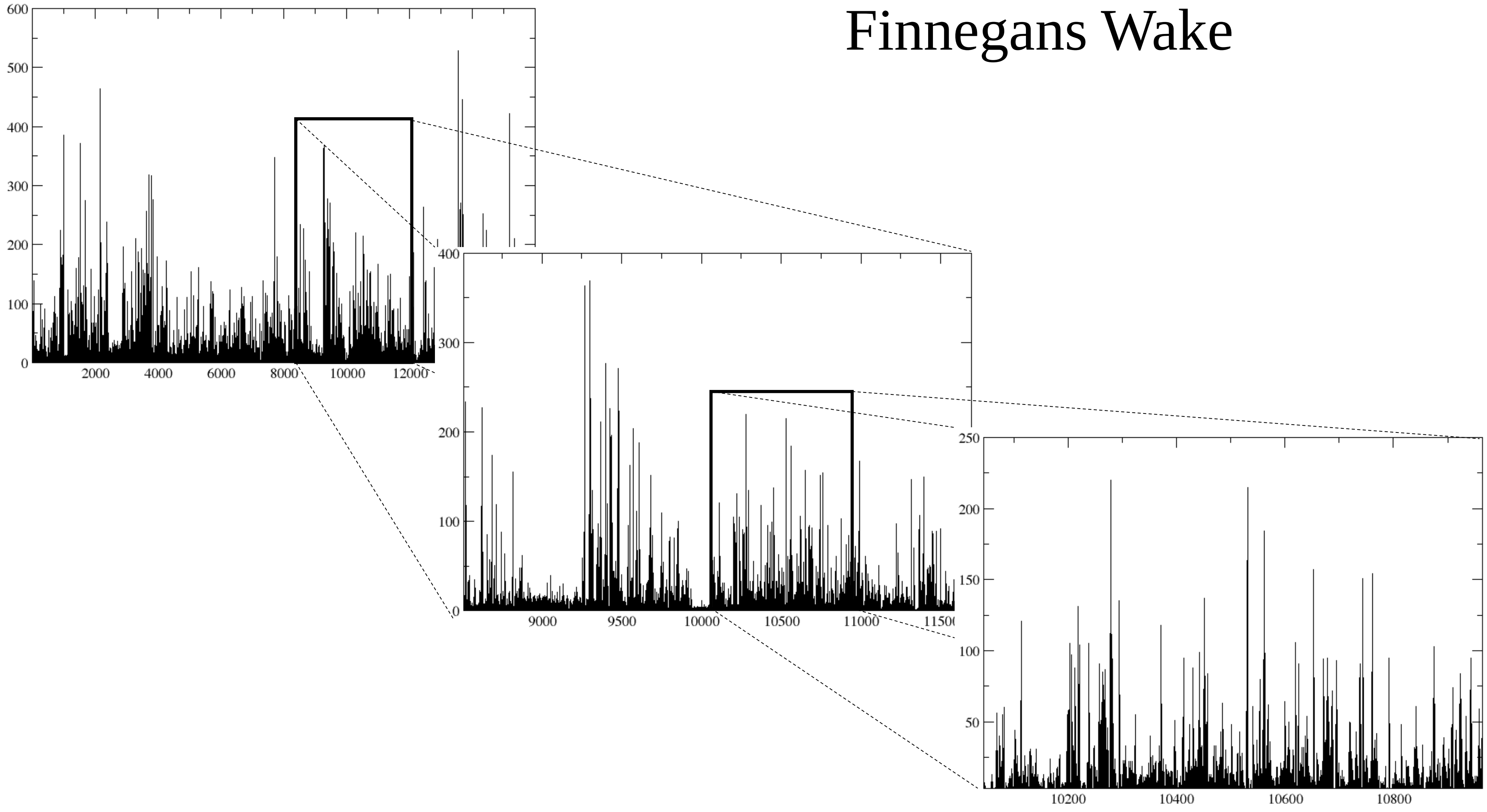


Figure 1. The distributions of the distances between consecutive punctuation marks (**left column**) and the corresponding hazard functions (**right column**) in *Finnegans Wake* and its translations.

Self-similar (fractal) sentence length variability

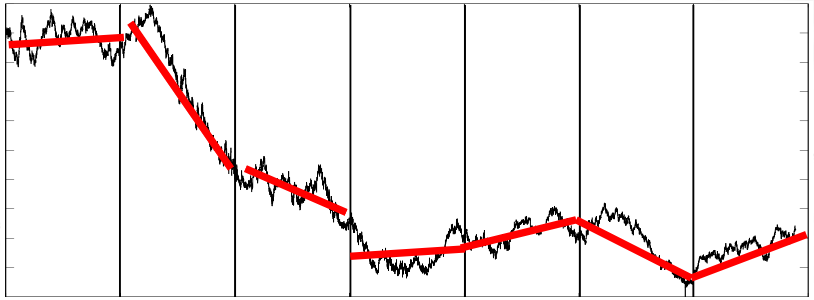
Finnegans Wake



There may exist nonlinear, heterogeneous, self-similarly convoluted structures, undetectable by S(f

→ **Multifractal analysis - methods**

Multifractal Detrended Fluctuations Analysis (MFDFA)



$$F_q(s) \equiv \left\{ \frac{1}{2N_s} \sum_{v=1}^{2N_s} \left[F^2(s,v)\right]^{q/2} \right\}^{1/q} \sim s^{\frac{1+\tau(q)}{q}}$$

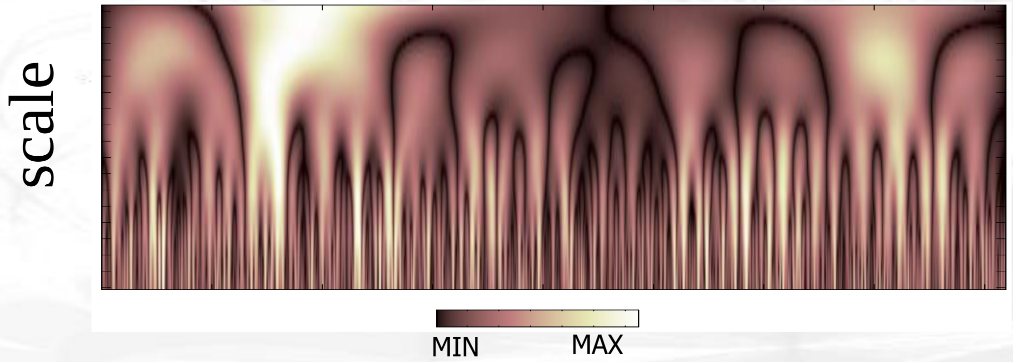
$$F_q \sim s^{h(q)}$$

$$\tau(q) = qh(q) - 1$$

$$\alpha = \tau'(q),$$

$$f(\alpha) = q\alpha - \tau(q)$$

Wavelet Transform Modulus Maxima (WTMM)



$$T_\psi(n,s') = \frac{1}{s'} \sum_{i=1}^{N} x_i \psi[(i-n)/s']$$

$$Z(q,s') = \sum_{l \in L(s')} \left|T_\psi(n_l(s'), s')\right|^q \sim s'^{\tau(q)}$$

PHYSICAL REVIEW E 74, 016103 (2006)

Wavelet versus detrended fluctuation analysis of multifractal structures

Paweł Oświęcimka,[1] Jarosław Kwapień,[1] and Stanisław Drożdż[1,2]
[1]Institute of Nuclear Physics, Polish Academy of Sciences, Kraków, Poland

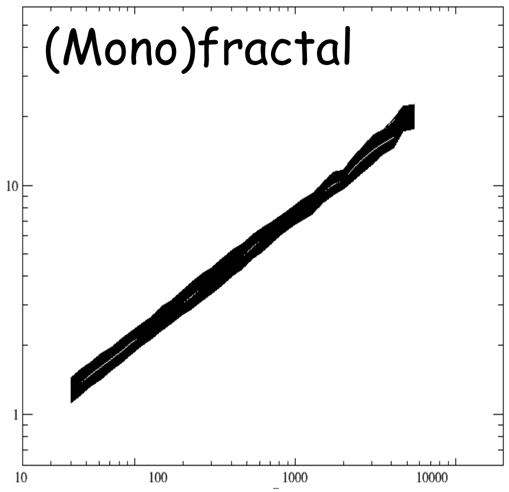
Power-laws, fractals and multifractals

Multifractal spectrum f(α)

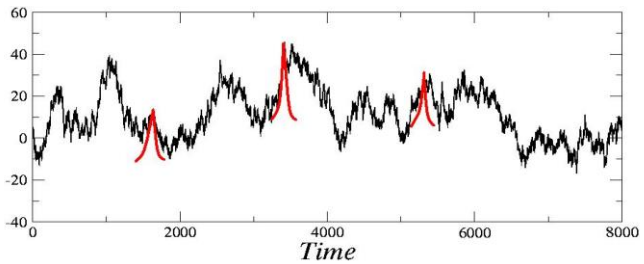
$$F_q(s) \sim s^{h(q)}$$

$$|f(x+l) - f(x)| \sim C\, l^{\alpha(x)}$$

$F_q(s)$



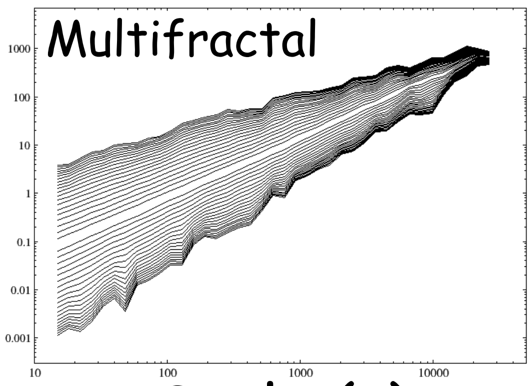
f(X)



X

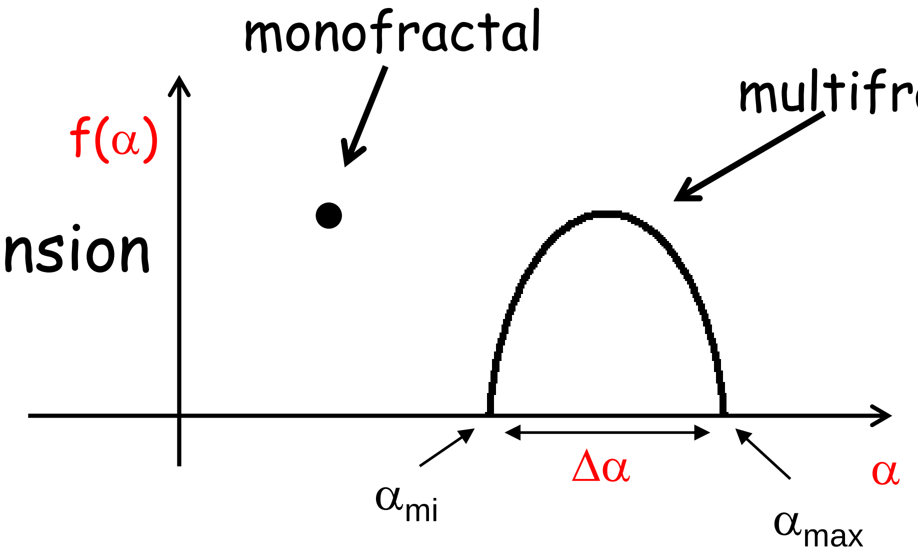
α – Hölder exponent

$F_q(s)$



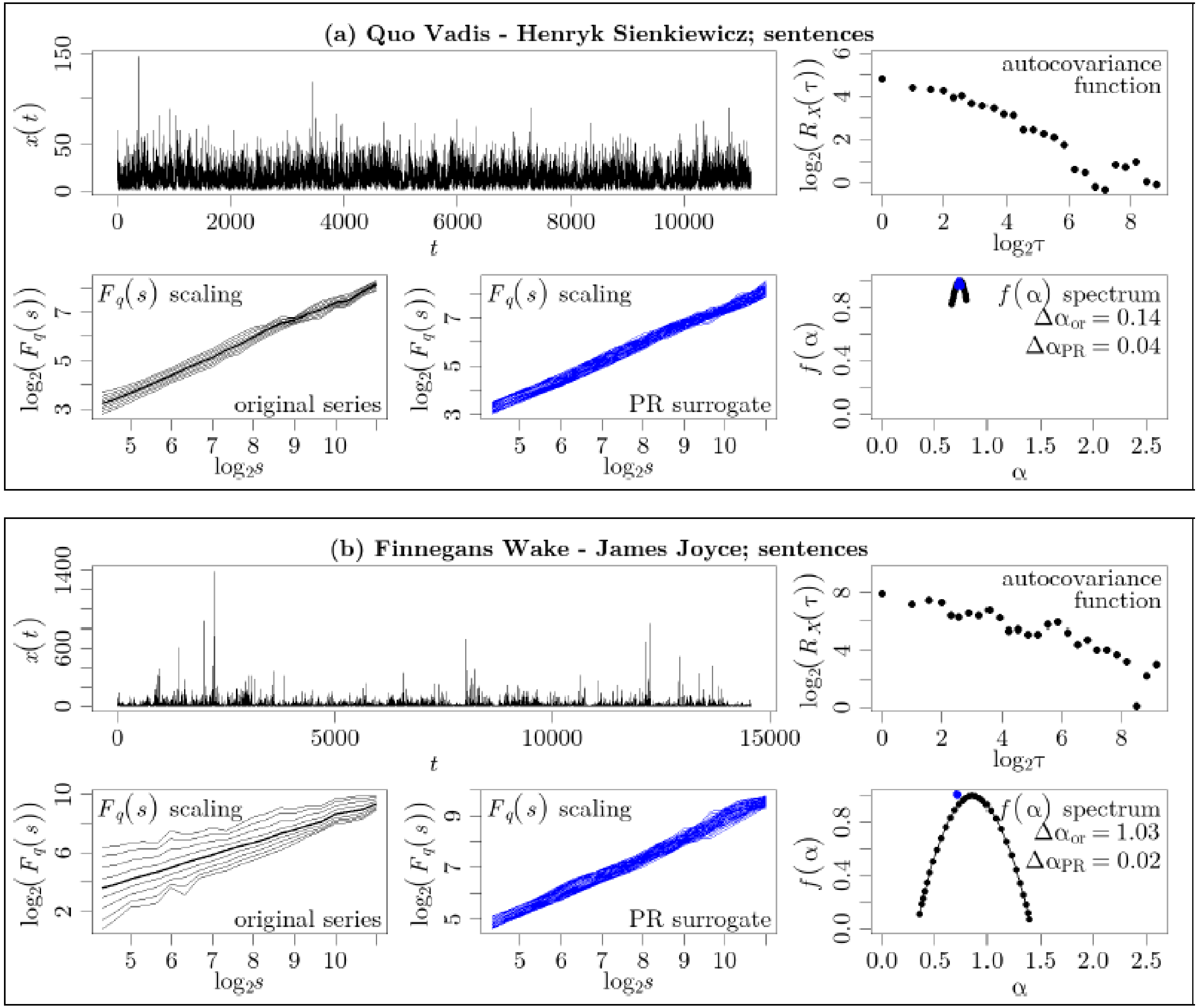
Scale (s)

f(α) – fractal dimension

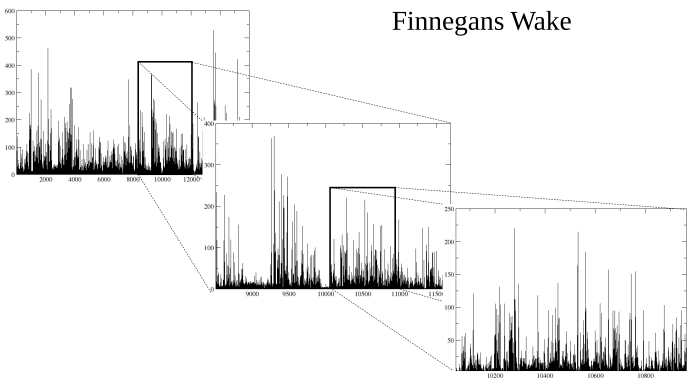


Large-size fluctuations

Small-size fluctuations



Finnegans Wake
is multifractal !
in sentence length
organization



Linguistic texts reveal many
complexity related, quantifiable, correlations

Identifying them may significantly help shaping
Large Language Models
and in stylometric analyses

- Punctuation carries the essence of correlations

„Speech is silver, silence is golden”, indeed

Thank you !