International Quantitative Linguistics Association

Uniwersytet Wrocławski

QUALICO 2018

International Quantitative Linguistics Conference

July 5-8, Wroclaw, Poland

DO LINGUISTIC LAWS EMERGE FROM VOICE?

Antoni Hernández Fernández

antonio.hernandez@upc.edu

UNIVERSITAT POLITÈCNICA DE CATALUNYA BARCELONATECH

Complexity and Quantitative Linguistics Lab,
Laboratory for Relational Algorithmics, Complexity and Learning (LARCA),
Departament de Ciències de la Computació /
Institut de Ciències de l'Educació

UNIVERSITAT POLITÈCNICA
DE CATALUNYA
BARCELONATECH

OUR GROUP

Complexity and Quantitative Linguistics Lab,
Laboratory for Relational Algorithmics, Complexity and Learning (LARCA),
Departament de Ciències de la Computació /Institut de Ciències de l'Educació

https://recerca.upc.edu/larca/en
https://www.cs.upc.edu/~cqllab/

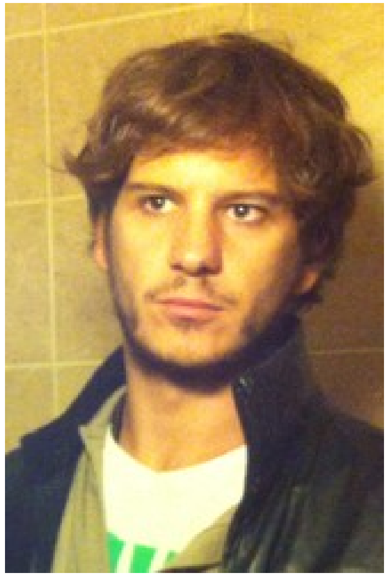
UNIVERSITAT POLITÈCNICA DE CATALUNYA BARCELONATECH

COLLABORATORS IN THIS WORK

Iván González-Torre
(The Hand-Walker Artist)

Bartolo Luque

UNIVERSIDAD POLITÉCNICA MADRID



Lucas Lacasa

Queen Mary University of London

Jordi Luque

Telefónica Investigación y Desarrollo

UNIVERSITAT POLITÈCNICA
DE CATALUNYA
BARCELONATECH

REFERENCES

[1] Luque, J., Luque, B. & Lacasa, L. (2015). **Scaling and universality in the human voice.** *J. R. Soc. Interface* 12, 20141344. doi: 10.1098/rsif.2014.1344

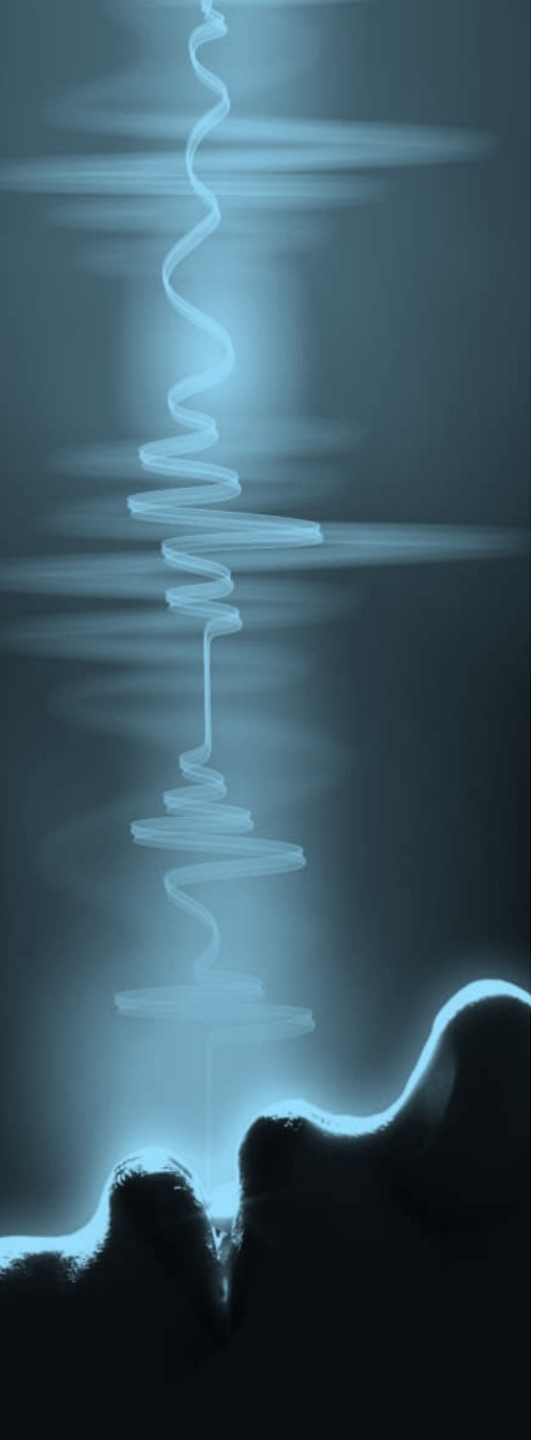
[2] Ferrer-i-Cancho, R., Hernández-Fernández, A., Lusseau, D., Agoramoorthy, G., Hsu, M. J. and Semple, S. (2013). **Compression as a Universal Principle of Animal Behavior**. *Cognitive Science*, 37: 1565–1578. doi:10.1111/cogs.12061

This conference:

[3] González Torre, I., Luque, B., Lacasa, L., Luque, J. & Hernández-Fernández, A. **Emergence of linguistic laws in human voice**. *Scientific Reports* 7, 43862 (2017). doi:10.1038/srep43862

[4] González Torre, I., Luque, B., Lacasa, L., Luque, J. & Hernández-Fernández, A. On the physical magnitudes of phonemes in English and Spanish (2018). ***In preparation***

[5] Hernández-Fernández, A., González Torre, I., Lacasa, L. Luque, J. & Luque, B. Do linguistic laws emerge from voice? (2018). ***In preparation***



CONTENTS

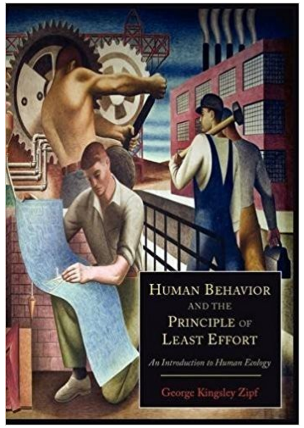
- Theoretical framework
- Scaling laws
- Materials & Methods
- Results
- Discussion and open questions

UNIVERSITAT POLITÈCNICA DE CATALUNYA
BARCELONATECH

THE QUESTION(S)

How to explain Linguistic Laws theoretically?

- **Least Effort** Ferrero (1894) ; Zipf (1949)
- **Compression Principle** and/or other principles from Information Theory (maximization of mutual information, minimization of entropy…)…



Ferrer i Cancho et al (2013)
Dębowski (2015, 2018 in prep)
Ferrer i Cancho (2018)

Evidence from voice or from texts? Are both sources "equal"?

Ferrero, G. (1894). L'inertie mentale et la loi du moindre effort. *Revue Philosophique de la France et de l'Étranger* ,37, 169–182.

Ł. Dębowski, (2015). The Relaxed Hilberg Conjecture: A Review and New Experimental Support. *Journal of Quantitative Linguistics*, vol. 22, pp. 311–337.

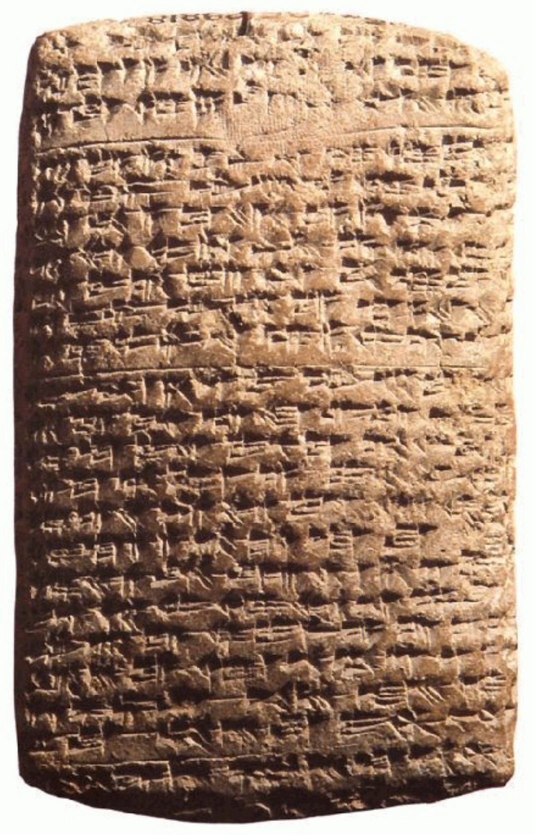
Ramon Ferrer-i-Cancho (2018) Optimization Models of Natural Communication, *Journal of Quantitative Linguistics*, 25:3, 207-237, DOI: 10.1080/09296174.2017.1366095

- Theoretical framework
- Scaling laws
- Materials & Methods
- Results
- Discussion and open questions

TEXT vs VOICE

UNIVERSITAT POLITÈCNICA DE CATALUNYA BARCELONATECH

Empirical evidence of robust linguistic laws holding in written texts across different human languages has been reported many times (Baayen, 2001; Altmann & Gerlach, 2016), and it has been shown that **these laws are not observed in random texts** (Ferrer-i-Cancho & Elvevag, 2010)



Text is interesting but…

…is a product of our **TECHNOLOGY (Scripture)**.

…inferences of statistical patterns of language in acoustics are biased by the arbitrary **segmentation of the signal** (language dependent), and virtually precludes the possibility of making (**not-biased**) comparative studies between human voice and other animal communication systems.

Altmann, E. G. & Gerlach, M. (2016). *Statistical Laws in Linguistics*. In Creativity and Universality in Language, Lecture Notes in Morphogenesis (eds Degli Esposti, M., Altmann, E. & Pachet, F.) 7–26 (Springer, Cham, 2016).

Baayen, H. (2001). Word frequency distributions 18 (Springer Sci. & Business Media, 2001)

Ferrer-i Cancho, R. & Elvevag, B. (2010). Random texts do not exhibit the real Zipf 's law-like rank distribution. PLoS One 5, e9411

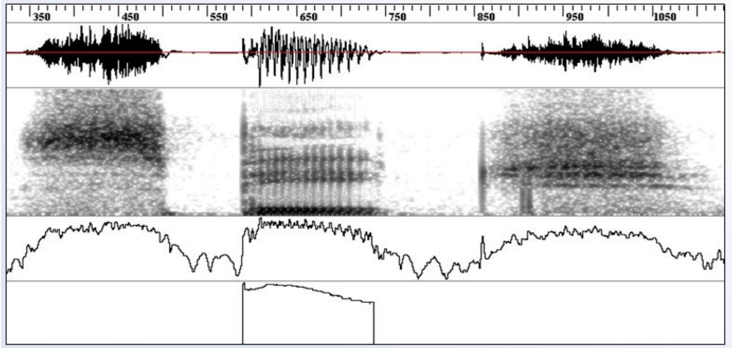
TEXT vs VOICE

UNIVERSITAT POLITÈCNICA DE CATALUNYA BARCELONATECH

Studies with oral corpus are much less abundant, and they imply:

- a **transcription of the acoustical waves** into words (case of **human speech**)
- or **some ill-defined analog of words** (**animal communication**)

... as the main segments to analyze statistically.



Diacritics

n̥ d̥	Voiceless	$t^ɣ$ $d^ɣ$	Velarized	d̚	No audible release	e̙	Retracted tongue root
s̬ d̬	Voiced	$t^ʕ$ $d^ʕ$	Pharyngealized	n̩	Syllabic	ɔ̹	More rounded
t^h d^h	Aspirated	b̤ a̤	Breathy voiced	ẽ	Nasalized	ɔ̜	Less rounded
t̪ d̪	Dental	b̰ a̰	Creaky voiced	ɚ	Rhoticity	u̟	Advanced
t̺ d̺	Apical	t̼ d̼	Linguolabial	e̯	Non-Syllabic	e̠	Retracted
t̻ d̻	Laminal	ɫ	Velarized / pharyngealized	e̝	Raised	ë	Centralized
t^w d^w	Labialized	d^n	Nasal release	e̞	Lowered	e̽	Mid-centralized
t^j d^j	Palatalized	d^l	Lateral release	e̘	Advanced tongue root		

This problem leads researchers to **manually segment acoustic signals** guided by their expertise and prevents to explore signals of unknown origin (Doyle et al, 2011).

Doyle, L., McCowan, B., Johnston, S. & Hanser, S. (2011) Information theory, animal communication, and the search for extraterrestrial intelligence. *Acta Astronaut.*, 68, 406–417.

- **Theoretical framework**
- Scaling laws
- Materials & Methods
- Results
- Discussion and open questions

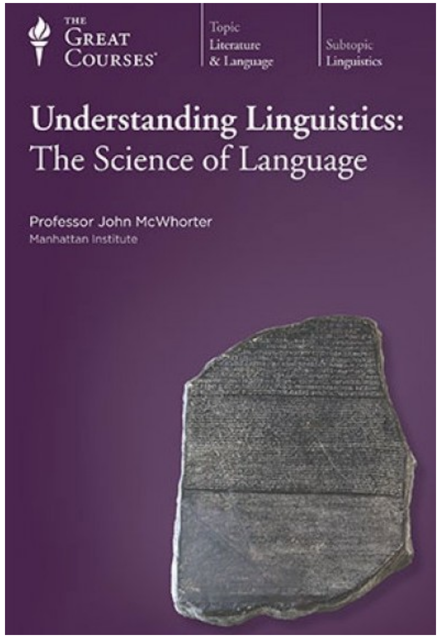
TEXT vs VOICE

UNIVERSITAT POLITÈCNICA DE CATALUNYA
BARCELONATECH

It involves two major problems in communication studies:

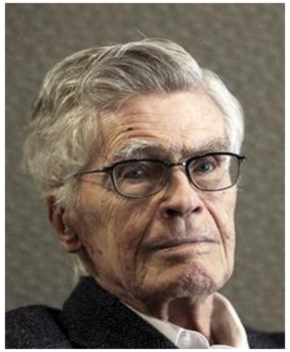
(i) The **impossibility of performing fully objective comparative studies** between human and non-human signals.

(ii) A rather arbitrary definition of the units of study guided by **ortographic conventions** already **produces non-negligible epistemological problems at the core of Linguistics** (Bunge, 1984; Köhler, 2005).



Bunge, M. (1984) What is pseudoscience? *The Skeptical Inquirer* 9, pp. 36–46.

Kohler, R. (2005) Synergetic linguistics. In *Quantitative linguistics* 760774. Berlin: DE Gruyter.



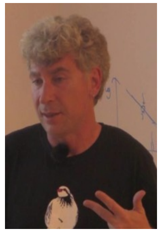
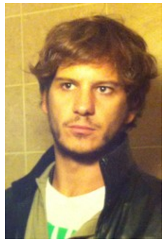
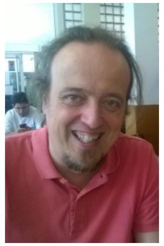
THE QUESTION(S)

UNIVERSITAT POLITÈCNICA DE CATALUNYA
BARCELONATECH


Do linguistic laws emerge from voice?

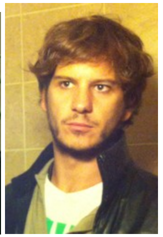
•What is the **origin of the linguistic laws** that we know (Zipf's law, brevity law…)?



•Have they a **physiological** (physical) origin?

Ferrer i Cancho et al (2013); Luque et al (2015)


•Could they be **Scaling Laws/SOC**?

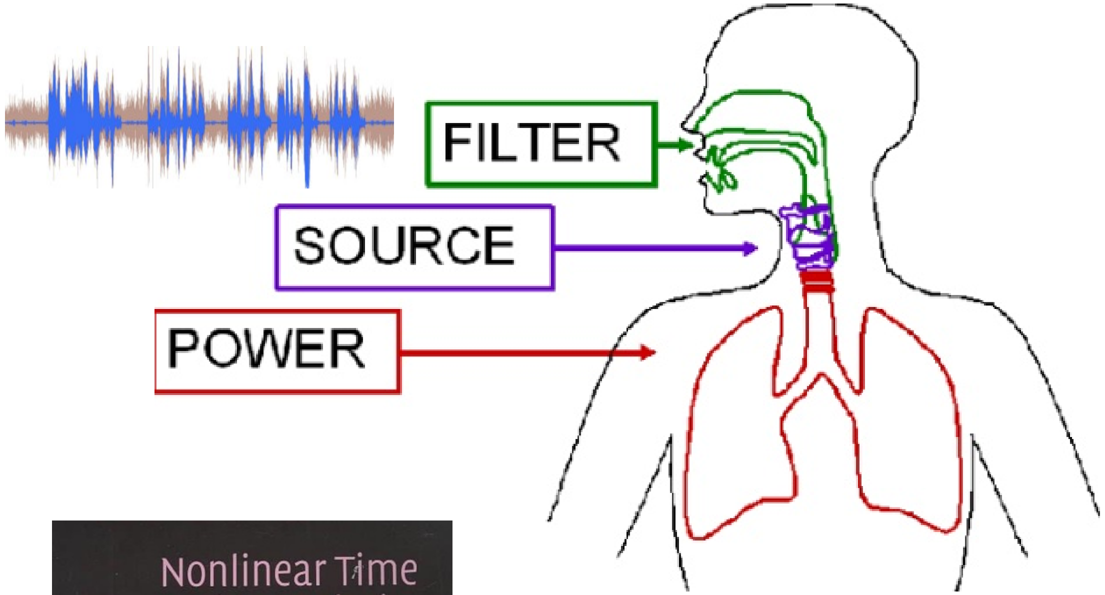
Luque et al (2015)

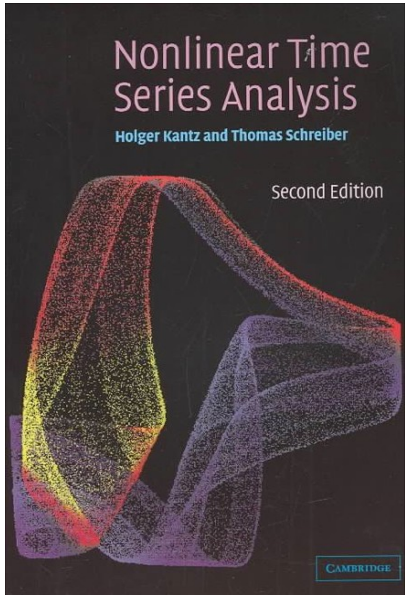
•Can we find computable patterns at levels lower than the phoneme? (TECHNICAL APPLICATIONS, SPEECH TECHNOLOGIES)

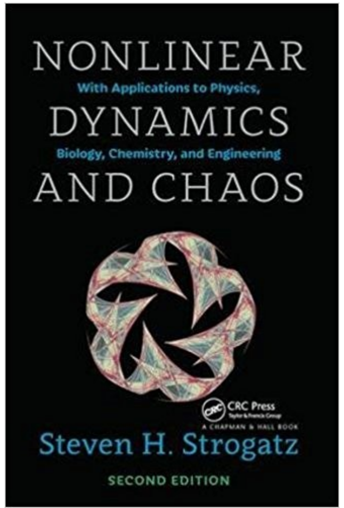
González-Torre et al (2017)

SPEECH SYNTHESIS

UNIVERSITAT POLITÈCNICA DE CATALUNYA
BARCELONATECH







- **Speech synthesis software** fail to be "natural" so engineers introduce "**residuals**" in synthesis algorithms (small pieces of real human voice)

- Sure, because human voice evidences **nonlinearities** at fine grained level (**deviations from source filter theory which is linear and assumes voice is a combination of Gaussians**)

Luque et al (2015)

- Theoretical framework
- Scaling laws
- Materials & Methods
- Results
- Discussion and open questions

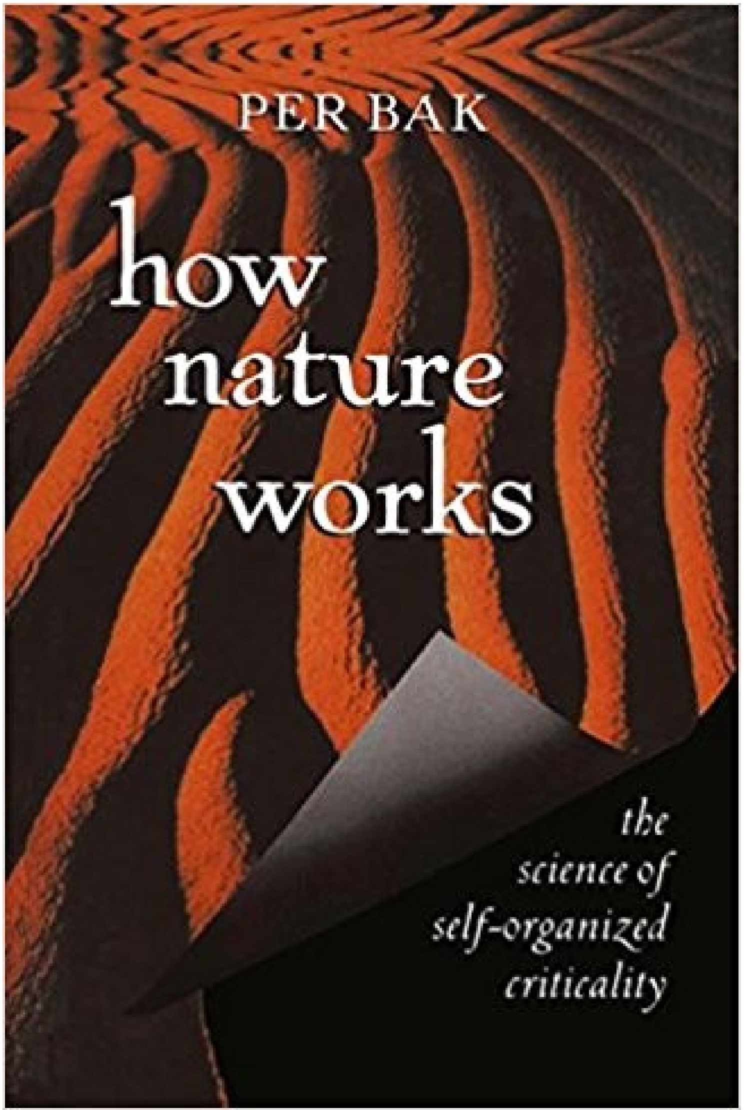
Luque et al (2015)

SCALING LAWS

UNIVERSITAT POLITÈCNICA DE CATALUNYA BARCELONATECH

Scaling law (SL) is a functional relationship between two quantities, independent of the initial size of those quantities: **one quantity varies as a power of another** (POWER LAW).

Self-Organized Criticality (SOC) is a property of dynamical systems that have a critical point as an attractor. (Bak et al, 1987) SOC is a phenomenon observed in complex systems of multiple interacting components, that **produce power-law distributed avalanche sizes**. (Hoffman & Payton, 2018)

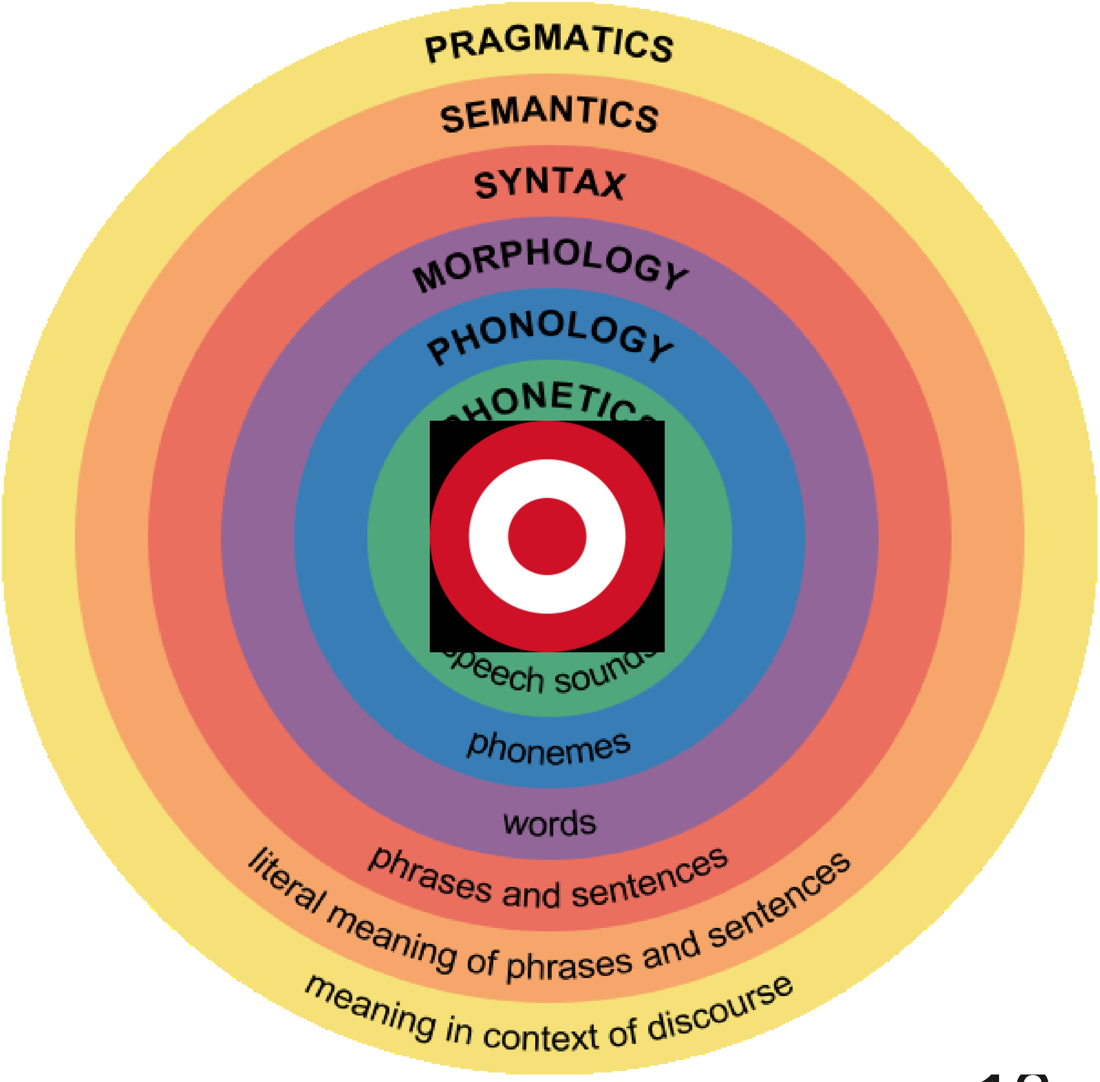


Bak, P., Tang, C. and Wiesenfeld, K. (1987). «Self-organized criticality: an explanation of the 1/f noise». *Physical Review Letters* **59**: 381-384. doi:10.1103/PhysRevLett.59.381

SL in human voice?

UNIVERSITAT POLITÈCNICA DE CATALUNYA
BARCELONATECH

The equivalence of power laws with a particular scaling exponent can have a deeper origin in **the dynamical processes that generate the power-law relation.**



Luque et al (2015)

https://commons.wikimedia.org/wiki/File:Major_levels_of_linguistic_structure.svg

- Theoretical framework
- Scaling laws
- Linguistic laws
- **Materials & Methods**
- Results
- Discussion and open questions

MATERIALS

UNIVERSITAT POLITÈCNICA DE CATALUNYA BARCELONATECH

Telefónica Investigación y Desarrollo

Dataset 1: KALAKA2

- TV broadcast speech dataset
- **4 hours per language**
- **6 languages (Basque, Catalan, Galician, Spanish, Portuguese and English)**
- Different conditions (planned & spontaneous speech, different environments, excluding telephonic channel)
- CD quality (16 bit / 44.1 kHz / stereo) Roland Edirol R-09 ultralight digital audio recorder
- Signals downsampled at 16kHz, left & right channel averaged via SoX and stored in WAV

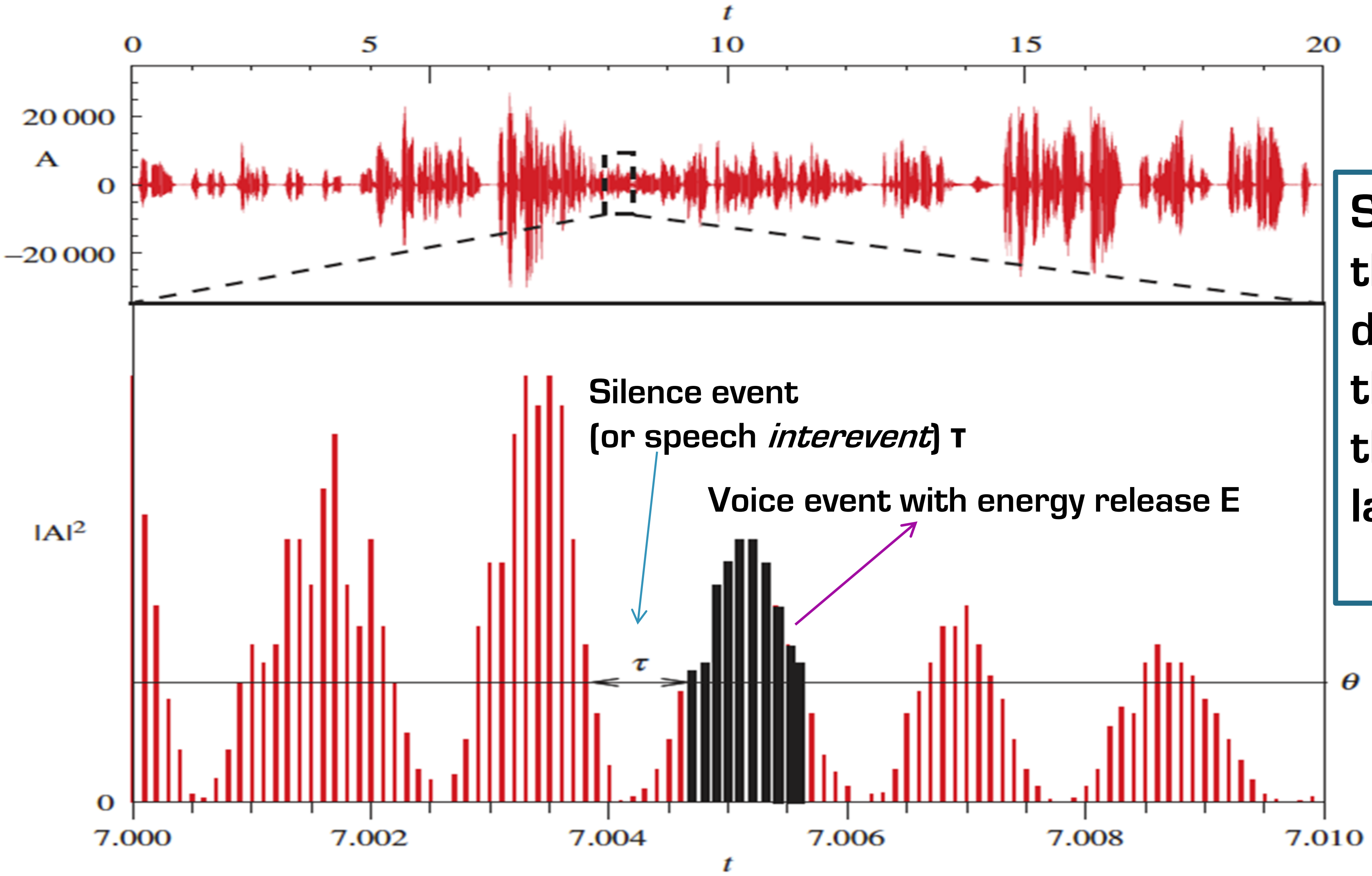
Dataset 2: NIST Language Recognition Evaluation 1996

- conversations drawn mainly from LDC Friendcall corpus
- **2—4 hours per language**
- **11 languages (English, Arabic, French, Mandarin, German, Hindi, Japanese, Spanish, Korean, Tamil, Vietnamese).**
- several speakers from several conversations but speaking the same language
- signals correspond to one side of a 4-wire telephonic conversation
- standard 8 bit 8kHz mu-law digital telephone data
- samples converted into 2byte PCM digital format

- Theoretical framework
- Scaling laws
- **Materials & Methods**
- Results
- Discussion and open questions

METHOD

Defining Energy Releases via thresholding



Set a threshold θ defined as the % of data that are larger than θ

Luque et al (2015)

Luque et al (2015)

METHOD

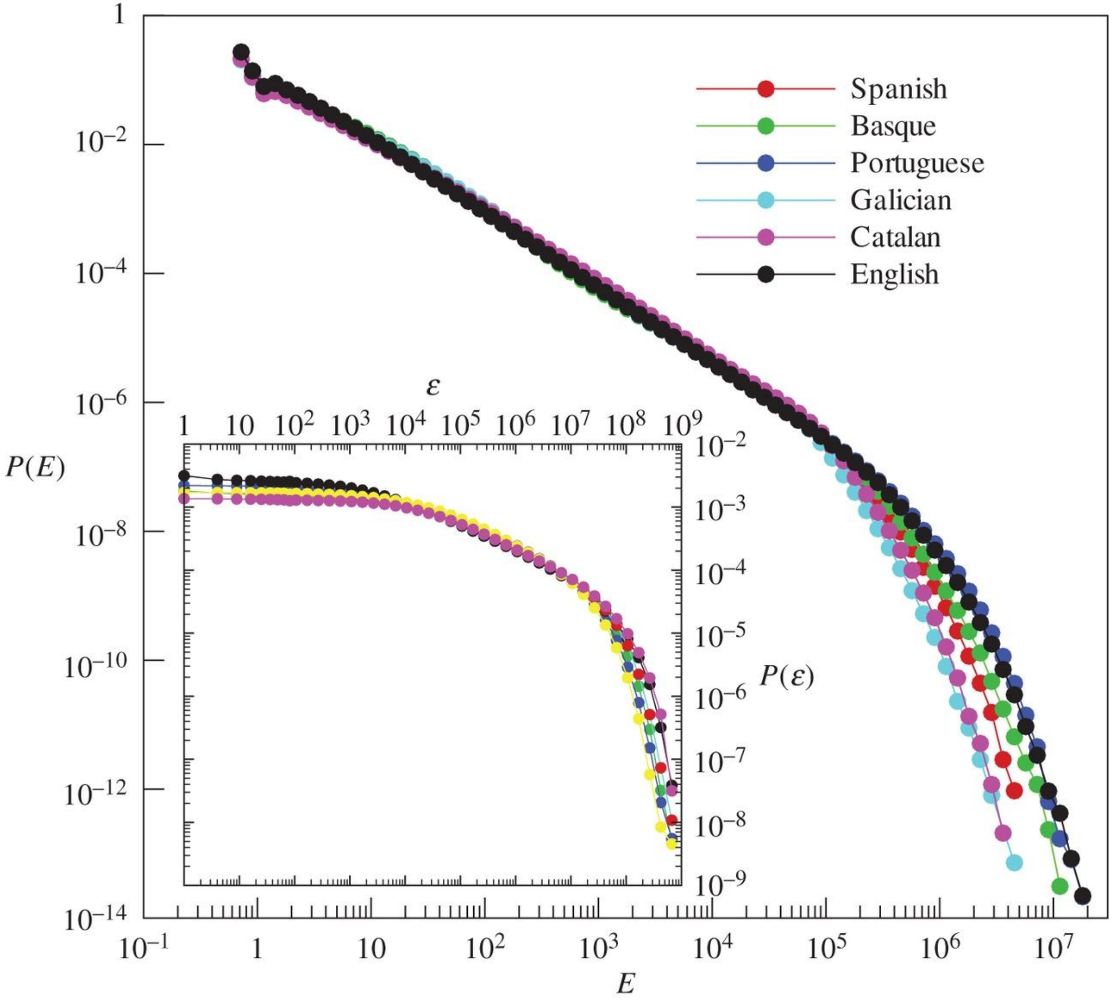
Previous Work

UNIVERSITAT POLITÈCNICA DE CATALUNYA BARCELONATECH

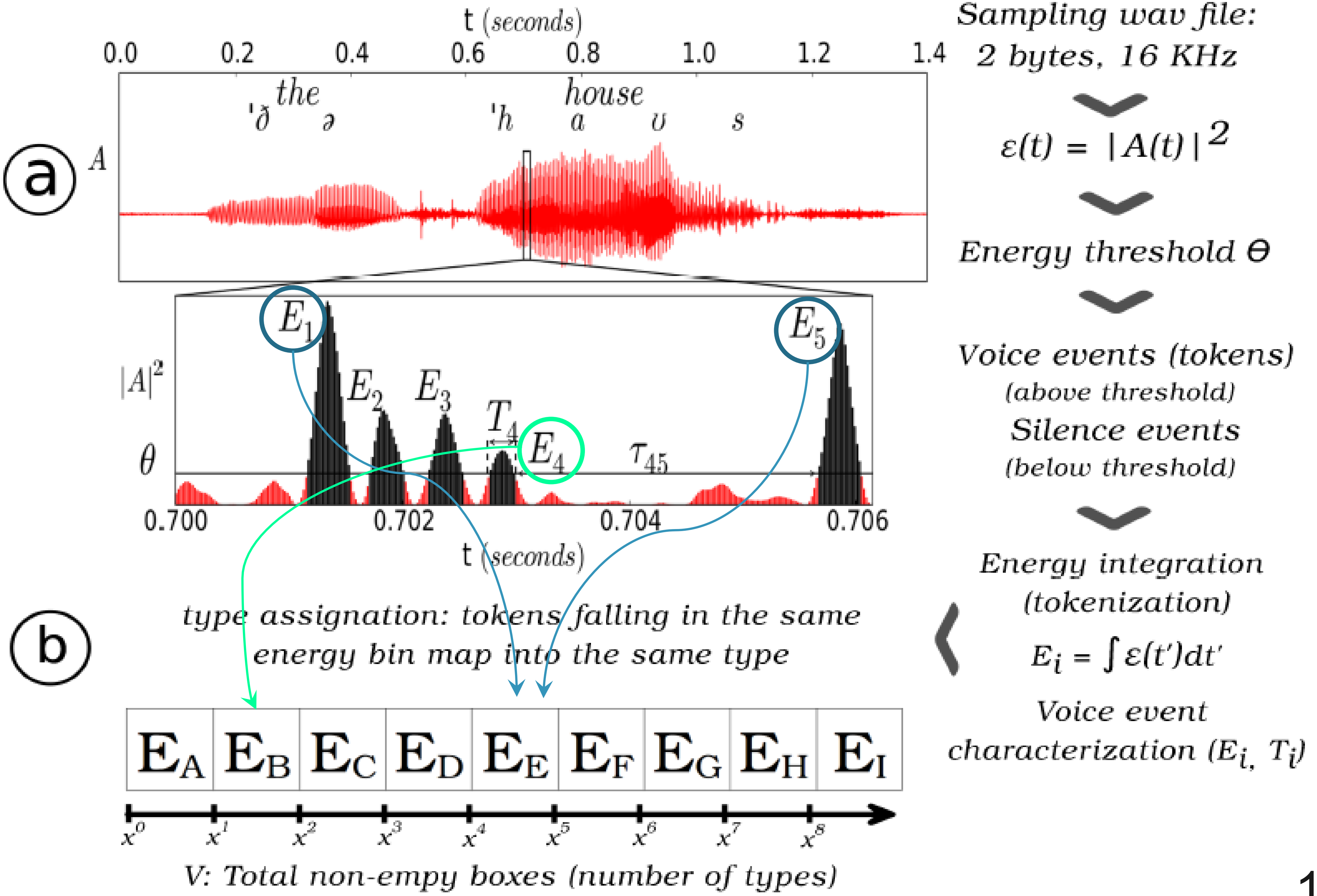
•During speech, **the energy** is unevenly released and power-law distributed. (**Gutenberg–Richter law**)

•**'Earthquakes in speech'** show **temporal correlations** and are power-law distributed.

•The process responsible for this complex **phenomenon is not cognitive, but it resides in the physiological** mechanisms (alveolar) of speech production.



Luque et al (2015)



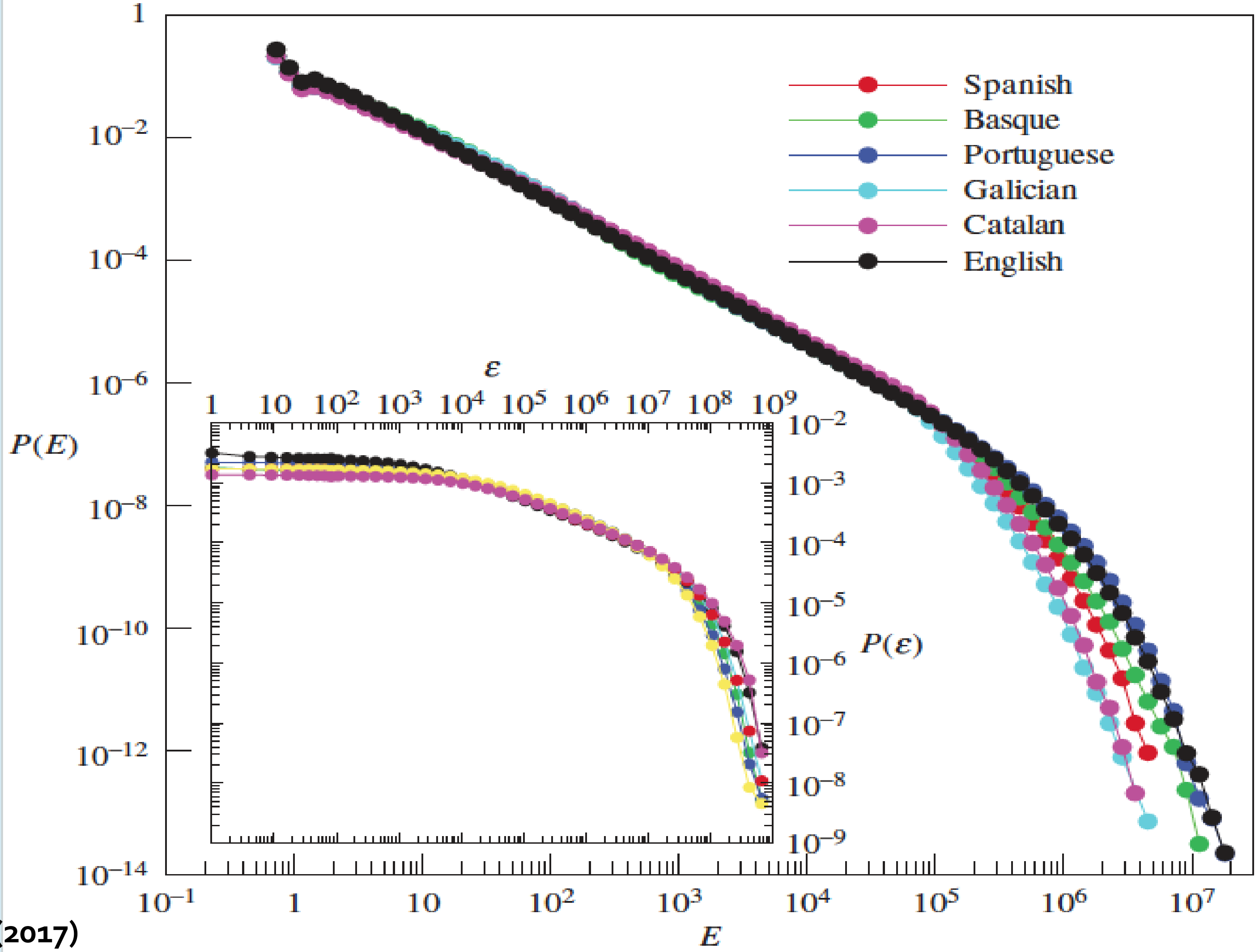
González-Torre et al (2015)

RESULTS

UNIVERSITAT POLITÈCNICA DE CATALUNYA BARCELONATECH

Binned histogram of integrated avalanches P(E) – Gutenberg-Richter-like law

Luque et al (2015); González-Torre et al (2017)



Exponent	ϕ
Basque	1.13 ± 0.04
Catalan	1.17 ± 0.05
English	1.16 ± 0.05
Galician	1.18 ± 0.04
Portuguese	1.16 ± 0.05
Spanish	1.15 ± 0.04

we fix threshold θ=80% constant & small to remove background noise

González-Torre et al (2017)

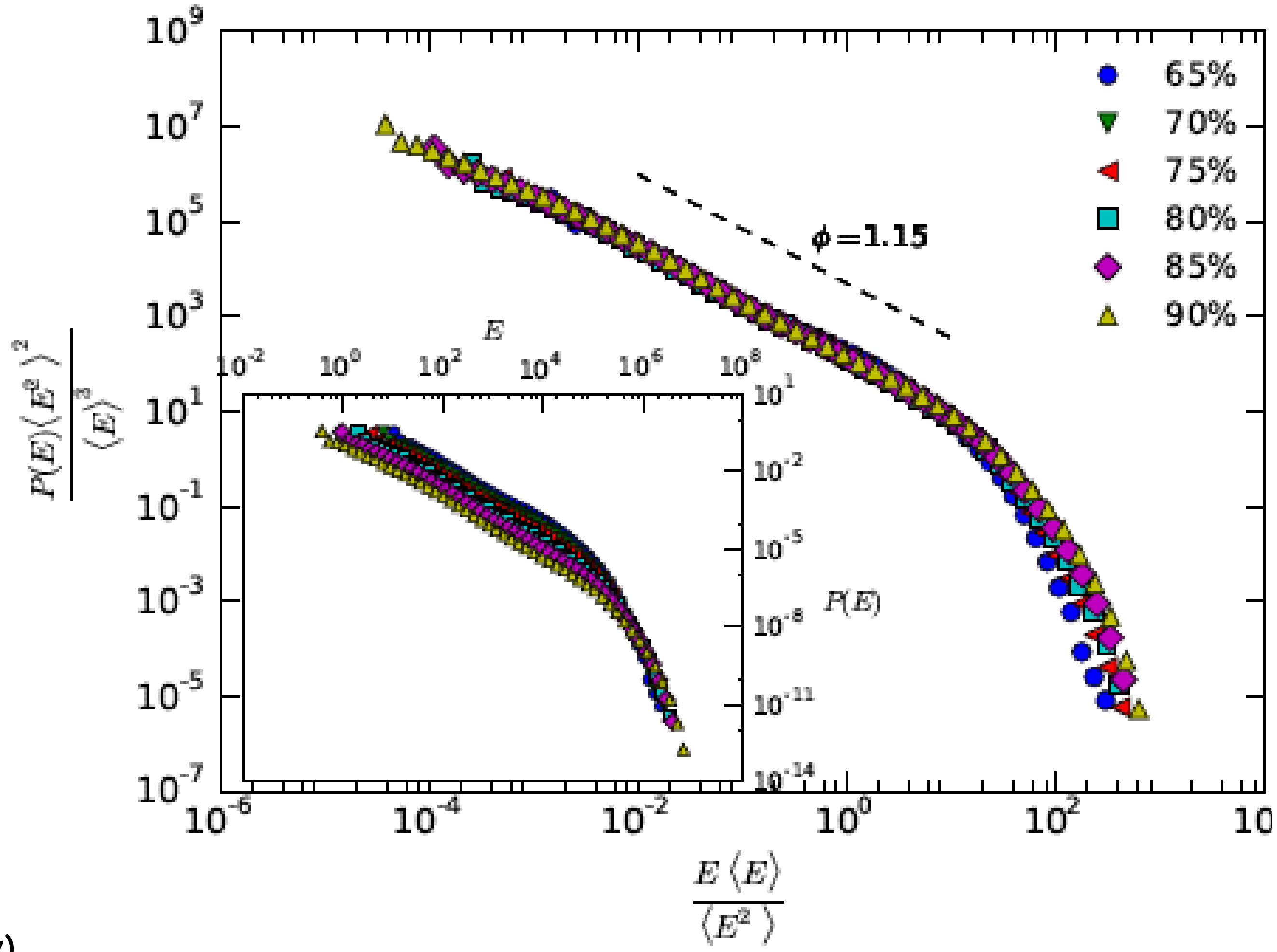
- Theoretical framework
- Scaling laws
- Materials & Methods
- **Results**
- Discussion and open questions

RESULTS

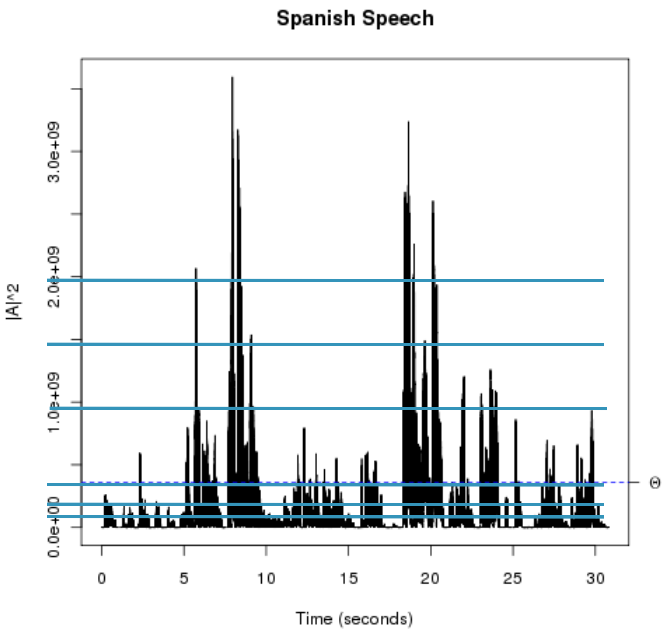
P(E) – Gutenberg-Richter-like law

Results are independent of the threshold (**invariant under rescaling**)

$$E \to E\langle E\rangle/\langle E^2\rangle,\ P_\Theta(E) \to P_\Theta(E)\langle E^2\rangle^2/\langle E\rangle^3.$$



Varying **θ between 90% and 50%** allows to set a threshold that ranges several orders of magnitude in energy.



González-Torre et al (2017)

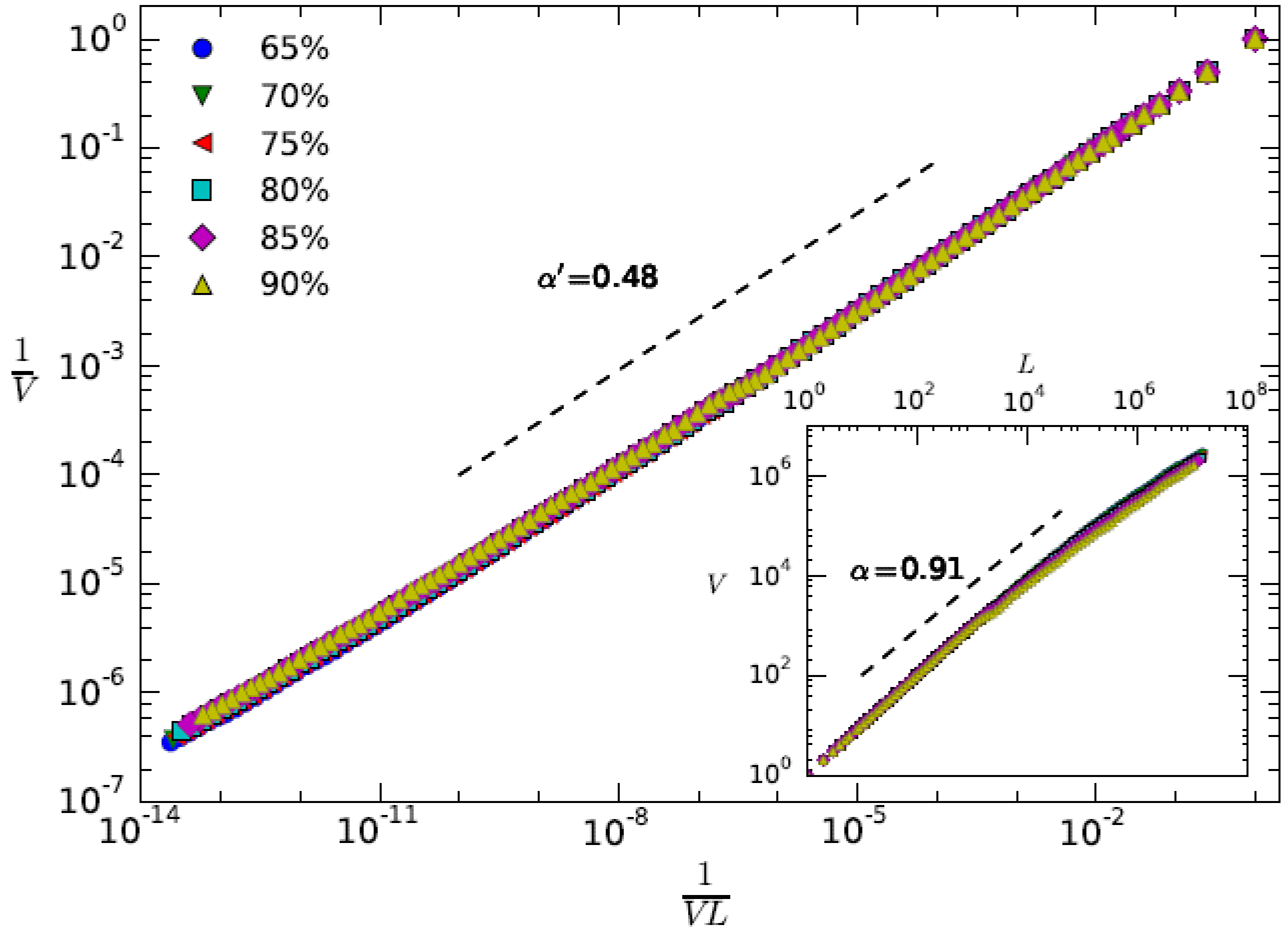
- Theoretical framework
- Scaling laws
- Materials & Methods
- **Results**
- Discussion and open questions

RESULTS HEAPS –HERDAN LAW

UNIVERSITAT POLITÈCNICA DE CATALUNYA BARCELONATECH

Sublinear growth of the number of different **elements V** in a **text** with **text size L**

$$V \sim L^{\alpha},\ \alpha < 1$$



Log-log plot of the Heaps' law for the **Portuguese sample** (KALAKA) and several thresholds. In the inner panel we show how different **tokens (V)** increases sublinearly with **the size of the series (L),** where the slope can be estimated properly for about three decades.

González-Torre et al (2017)

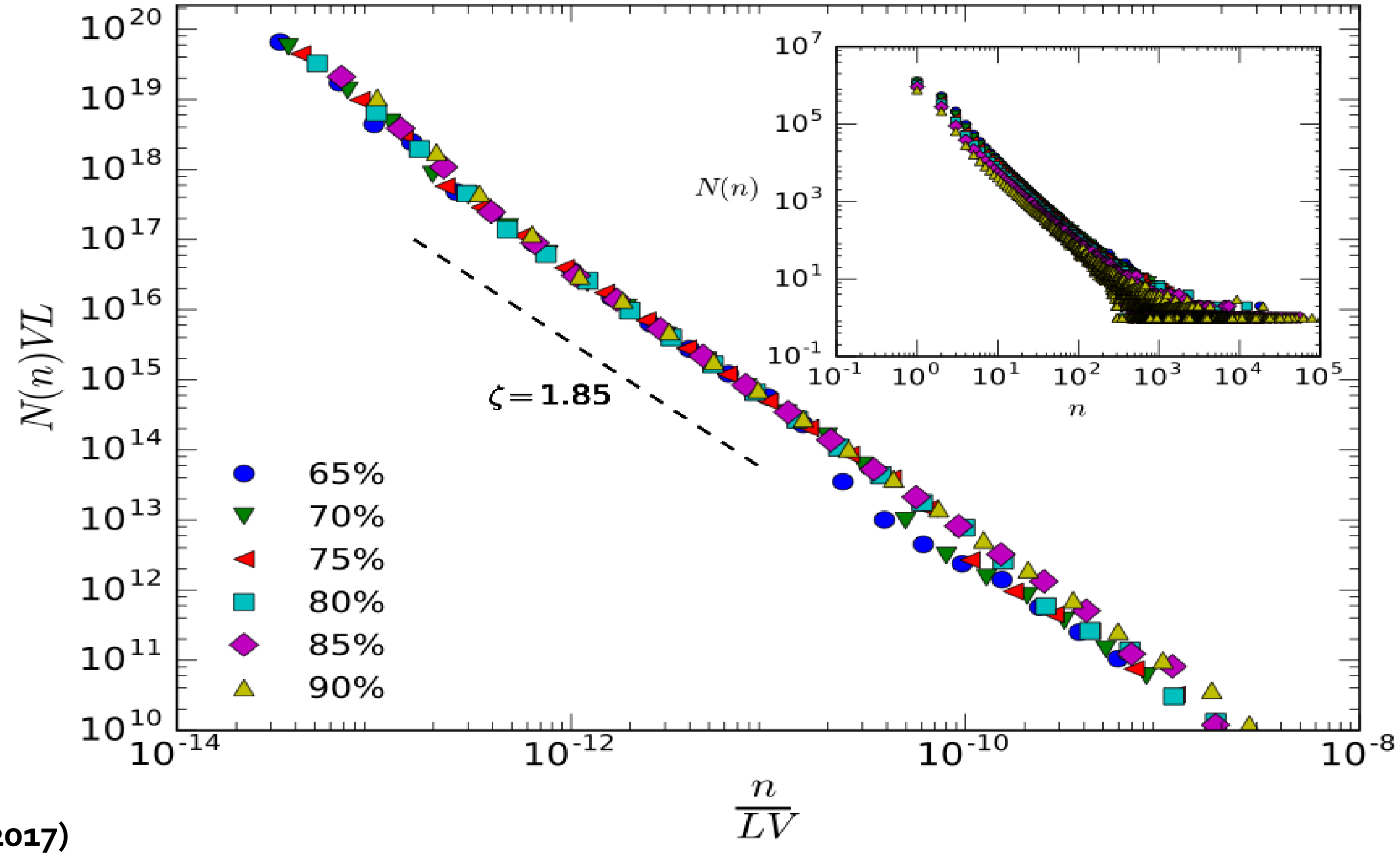
- Theoretical framework
- Scaling laws
- Materials & Methods
- **Results**
- Discussion and open questions

RESULTS

ZIPF's LAW

UNIVERSITAT POLITÈCNICA DE CATALUNYA BARCELONATECH

Number of different ***"words"*** (vocabulary) which occur exactly **n** times decays as $\mathcal{N}(n) \sim n^{-\zeta}$ (or) number of times the word with rank **r** occur decays as $n(r) \sim r^{-z}$ $z = \frac{1}{\zeta - 1}$

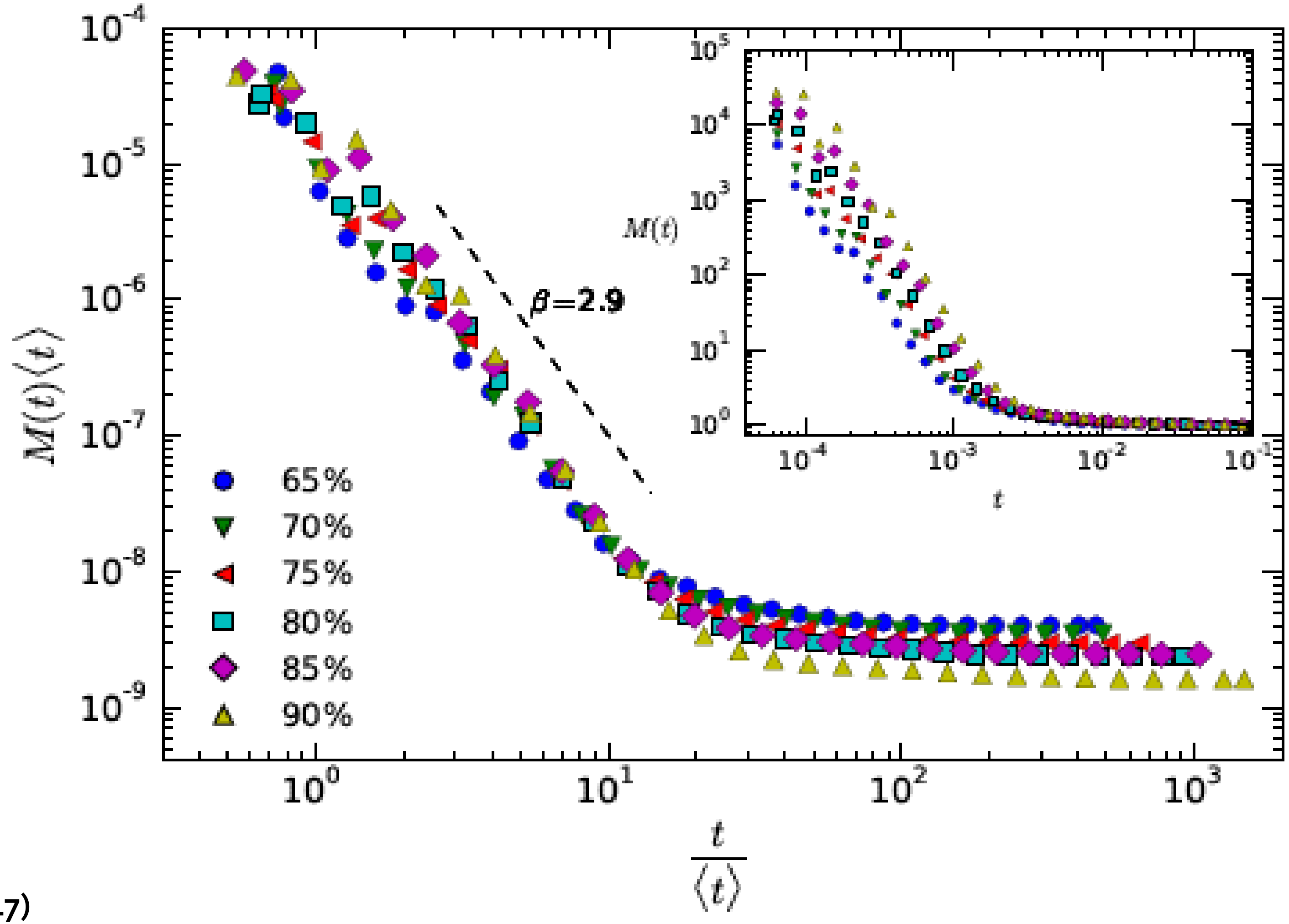


González-Torre et al (2017)

RESULTS

ZIPF'S BREVITY LAW

Tendency of more frequent ***words*** to be shorter or smaller (Zipf 1935).



Log-log plot in the case of **English** (KALAKA), for several thresholds. In the upper panel we plot the histogram M(t) that describes the relative frequency of a type of mean duration t.

González-Torre et al (2017)

RESULTS

SUMMARY

UNIVERSITAT POLITÈCNICA DE CATALUNYA
BARCELONATECH

- Theoretical framework
- Scaling laws
- Materials & Methods
- **Results**
- Discussion and open questions

Exponent	ϕ	ζ	α	β
Basque	1.13 ± 0.04	1.77 ± 0.14	0.90 ± 0.03	3.1 ± 0.3
Catalan	1.17 ± 0.05	1.89 ± 0.14	0.92 ± 0.03	2.8 ± 0.4
English	1.16 ± 0.05	1.85 ± 0.14	0.91 ± 0.01	2.9 ± 0.3
Galician	1.18 ± 0.04	1.80 ± 0.14	0.89 ± 0.03	2.9 ± 0.4
Portuguese	1.16 ± 0.05	1.77 ± 0.14	0.91 ± 0.01	3.0 ± 0.3
Spanish	1.15 ± 0.04	1.79 ± 0.14	0.91 ± 0.03	2.8 ± 0.4

TABLE I: Summary of scaling exponents associated to the energy release distribution (ϕ), Zipf's law (ζ), Heaps' law (α) and Brevity law (β) for the six different languages. Power law fits are performed using maximum likelihood estimation (MLE) following Clauset [71] and goodness-of-fit test and confidence interval are based on Kolmogorov-Smirnov (KS) tests. In all cases, KS are greater than 0.99. Exponents associated to energy release are compatible with those found in rainfall [70]. Results are compatible with the hypothesis of language-independence.

González-Torre et al (2017)

Clauset, A., Shalizi, C. & Newman, M. Power-law distributions in empirical data. SIAM Rev. 51, 661–703 (2009).

DISCUSSION

UNIVERSITAT POLITÈCNICA DE CATALUNYA
BARCELONATECH

- Theoretical framework
- Scaling laws
- Materials & Methods
- Results
- **Discussion and open questions**

• Human voice manifests the **analog of classical linguistic laws** found in written texts (Zipf 's law, Heaps' law and the brevity law) **in this level.**

•These laws **are invariant under changes of the energy threshold Θ.** As Θ is the only free parameter of the method, this invariance determines that the results are not afflicted by ambiguities associated to arbitrarily defining unit boundaries.

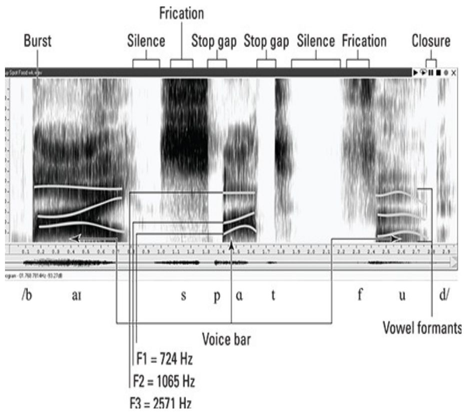
•Results **are robust across a list of 16 different languages** (indoeuropean and non-indoeuropean) and across timescales, energy threshold and conversational modes.

DISCUSSION

UNIVERSITAT POLITÈCNICA DE CATALUNYA BARCELONATECH

- Theoretical framework
- Scaling laws
- Materials & Methods
- Results
- **Discussion and open questions**

•Interpreting linguistic laws as Scaling Laws which emerged in communication systems actually opens the door for speculating on the existence of underlying **scale-invariant (physical) laws** operating underneath.



•The specific and complex alternation of air stops (silences) intertwined with voice production are at the core of the microscopic voice fluctuations **(SOC?).**

• **First observation of scaling behavior with a clear exponent in the case of brevity law in speech**. Our finding of a power law in brevity law differs from the case of **random typing** where a power law doesn't conform.

DISCUSSION

UNIVERSITAT POLITÈCNICA DE CATALUNYA
BARCELONATECH

• We are able **to map an arbitrary acoustic signal into a sequence of types separated by silence events**.

•Standard linguistic laws can then be directly explored in acoustic signals **without needs to have an *a priori* knowledge neither of the signal code nor of the adequate segmentation process** or the particular syntax of the **language** underlying the signal.

•This protocol can be used to make unbiased comparisons across different systems (**comparative studies**): **Universal Segmentation Method.**

Luque et al (2015)

OPEN QUESTIONS

UNIVERSITAT POLITÈCNICA DE CATALUNYA
BARCELONATECH

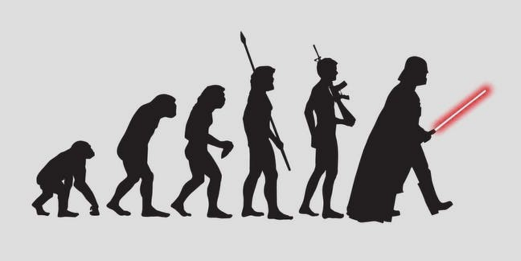
- Theoretical framework
- Scaling laws
- Materials & Methods
- Results
- **Discussion and open questions**

•**What are the values of the exponents indicating** (in this level under the phoneme)?

•How can we connect these findings with **information theory?**

• Emergence of ("linguistic") **Scaling Laws** already at the voice level: another hint of **complexity**? Is the system operating close to a critical point?

•**Is there any evolutive gain?**



• Relation with **traditional linguistic laws (in upper levels)**?

OPEN QUESTIONS

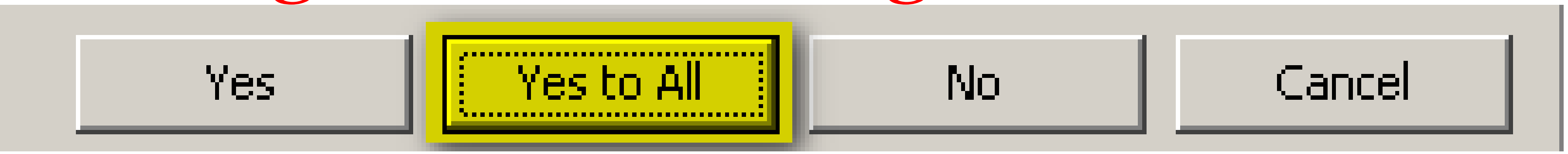
UNIVERSITAT POLITÈCNICA DE CATALUNYA
BARCELONATECH

Is **physiology** the ultimate reason of the onset of complexity and SL (linguistic laws) in communication?

Is it necessary to study other "forgotten" physical magnitudes (**Energy**: Guttenberg-Richter…)?

Instead of introducing pieces of real speech (residuals): Is it better to model speech fluctuations at intraphoneme via simple SOC models?

Do linguistic laws emerge from voice?



González-Torre et al (2017)

International Quantitative Linguistics Association

Uniwersytet Wrocławski

QUALICO 2018

International Quantitative Linguistics Conference

July 5-8, Wroclaw, Poland

July 7th, 2018

THANK YOU FOR YOUR ATTENTION!

Dziękuję bardzo!

Iván González-Torre, Bartolomé Luque, Lucas Lacasa, Jordi Luque & Antoni Hernández-Fernández
Emergence of linguistic laws in human voice. Scientific Reports 7, 43862;
doi: 10.1038/srep43862 (2017).

UNIVERSITAT POLITÈCNICA DE CATALUNYA BARCELONATECH

antonio.hernandez@upc.edu

This research was supported by the grant TIN2017-89244-R from MINECO (*Ministerio de Economía, Industria y Competitividad*; Spanish Government), and the recognition 2017SGR-856 (MACDA) from AGAUR (*Generalitat de Catalunya*).