Iván González Torre[(1)], Bartolo Luque[(1,2)], Lucas Lacasa[(2)], Jordi Luque[(3)], Antoni Hernández-Fernández[(4)]

(1): Department of Applied Mathematics and Statistics, EIAE, Technical University of Madrid, Plaza Cardenal Cisneros, 28040 Madrid (Spain), ivan.gonzalez.torre@upm.es, bartolome.luque@upm.es

(2): School of Mathematical Sciences, Queen Mary University of London, Mile End Road E14NS London (UK), l.lacasa@qmul.ac.uk

(3): Telefónica Research, Edicio Telefónica-Diagonal 00, Barcelona (Spain), jls@tid.es

(4): Complexity and Quantitative Linguistics Lab (LARCA), Institut de Ciències de l'Educacio, Universitat Politècnica de Catalunya, Plaza Eusebi Güell, 08034 Barcelona (Spain), antonio.hernandez@upc.edu

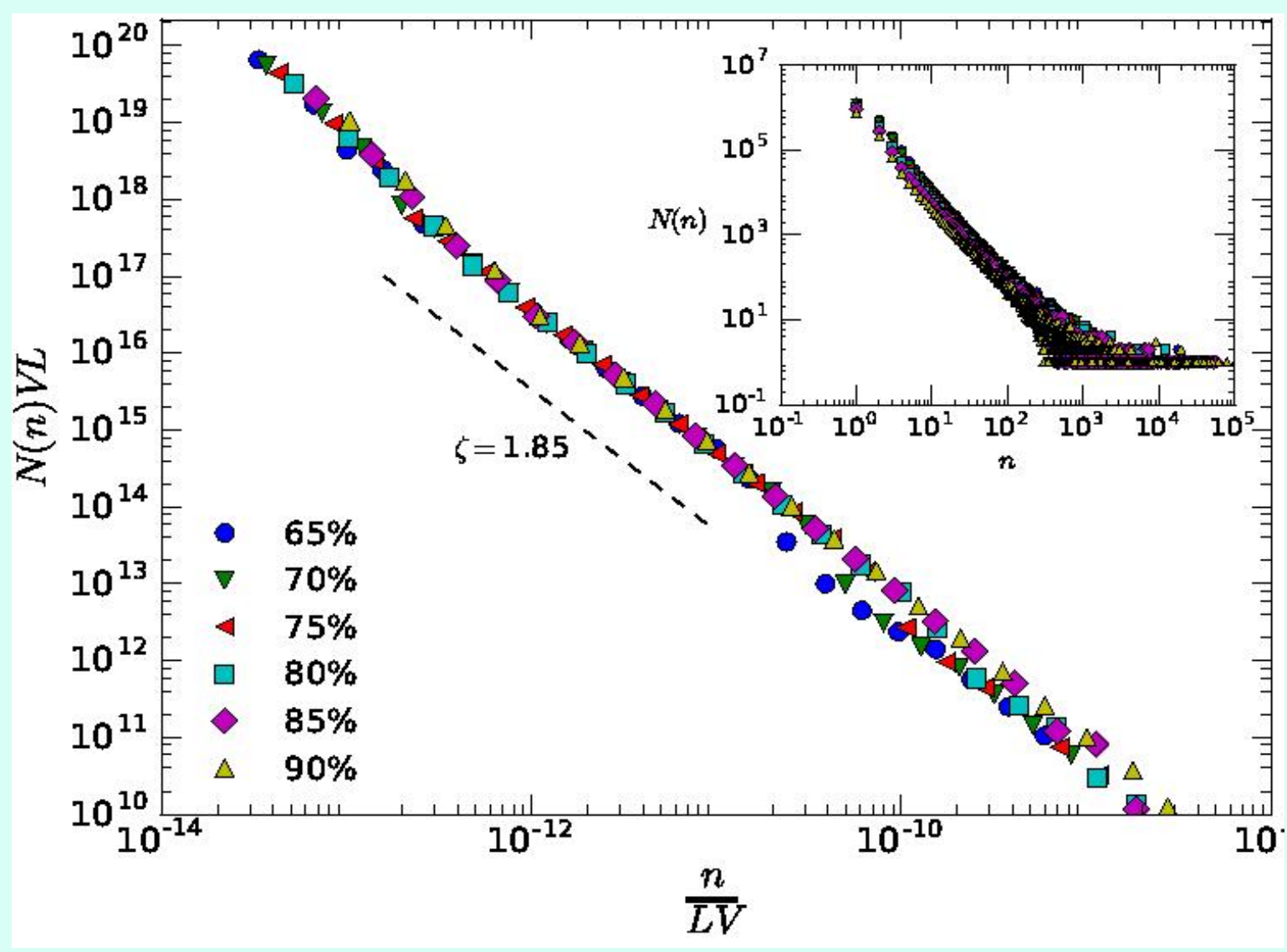
Linguistic laws constitute one of the quantitative cornerstones of modern cognitive sciences and have been routinely investigated in written corpora, or in the equivalent transcription of oral corpora. This means that inferences of statistical patterns of language in acoustics are biased by the arbitrary, language-dependent segmentation of the signal, and virtually precludes the possibility of making comparative studies between human voice and other animal communication systems.

Here we bridge this gap by proposing a method that allows to measure such patterns in acoustic signals of arbitrary origin, without needs to have access to the language corpus underneath. The method has been applied to sixteen different human languages, recovering successfully some well-known laws of human communication at timescales even below the phoneme and finding yet another link between complexity and criticality in a biological system. These methods further pave the way for new comparative studies in animal communication or the analysis of signals of unknown code.

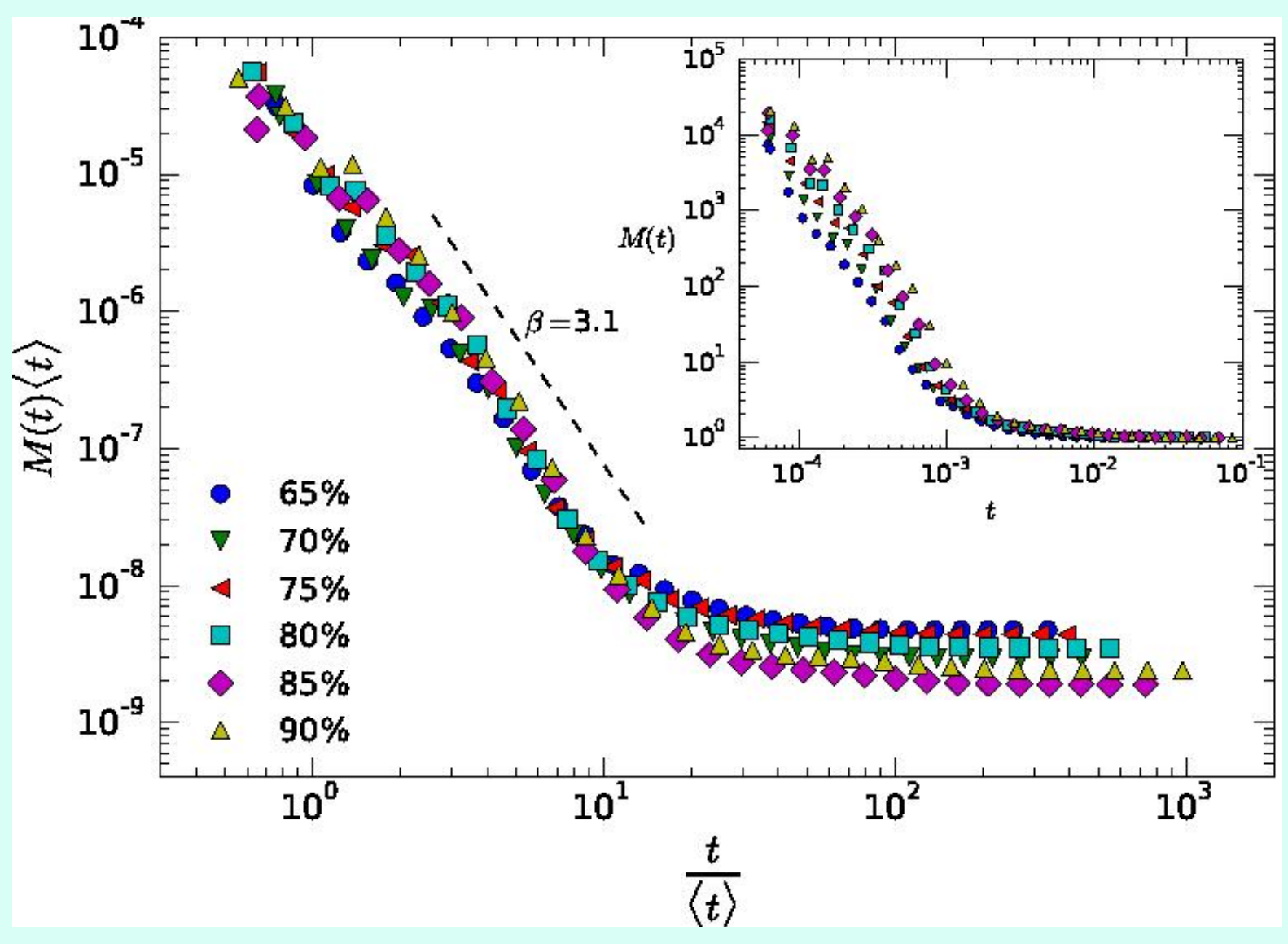
Linguistic Laws in human language

Acoustic communication is fully determined by three physical magnitudes extracted from the signals: frequency, energy and time. Interestingly, it is well known that we use statistical cues to segment the input and probably share with other species some of these mechanisms. When we explore directly human speech we previously found that human voice seems to be operating close to a critical state [2]. Applying the same method we recover linguistic laws directly from the signal without requiring any written transcription.

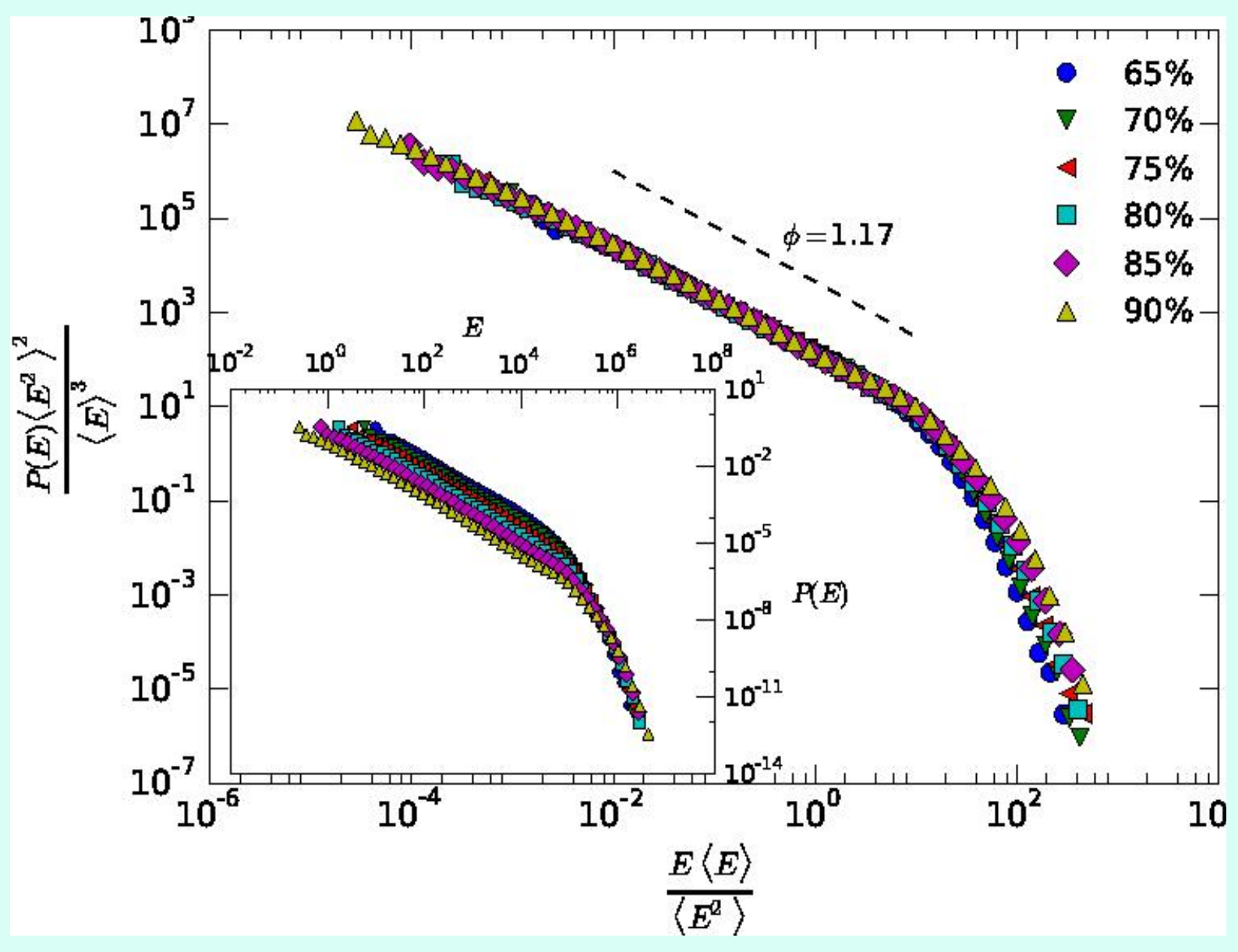
Zipf's law. It establishes that in a sizable linguistic sample the number of different words *N(n)* which occur exactly *n* times decays as $N(n) \sim n^{-\alpha}$, where the exponent α can vary from sample to sample but is usually close to 2..



Brevity Law. The tendency of more frequent words to be shorter can be generalized as the tendency of more frequent elements to be shorter or smaller (both in energy and duration), and its origin has been suggested to be related to optimization and information compression arguments [1].



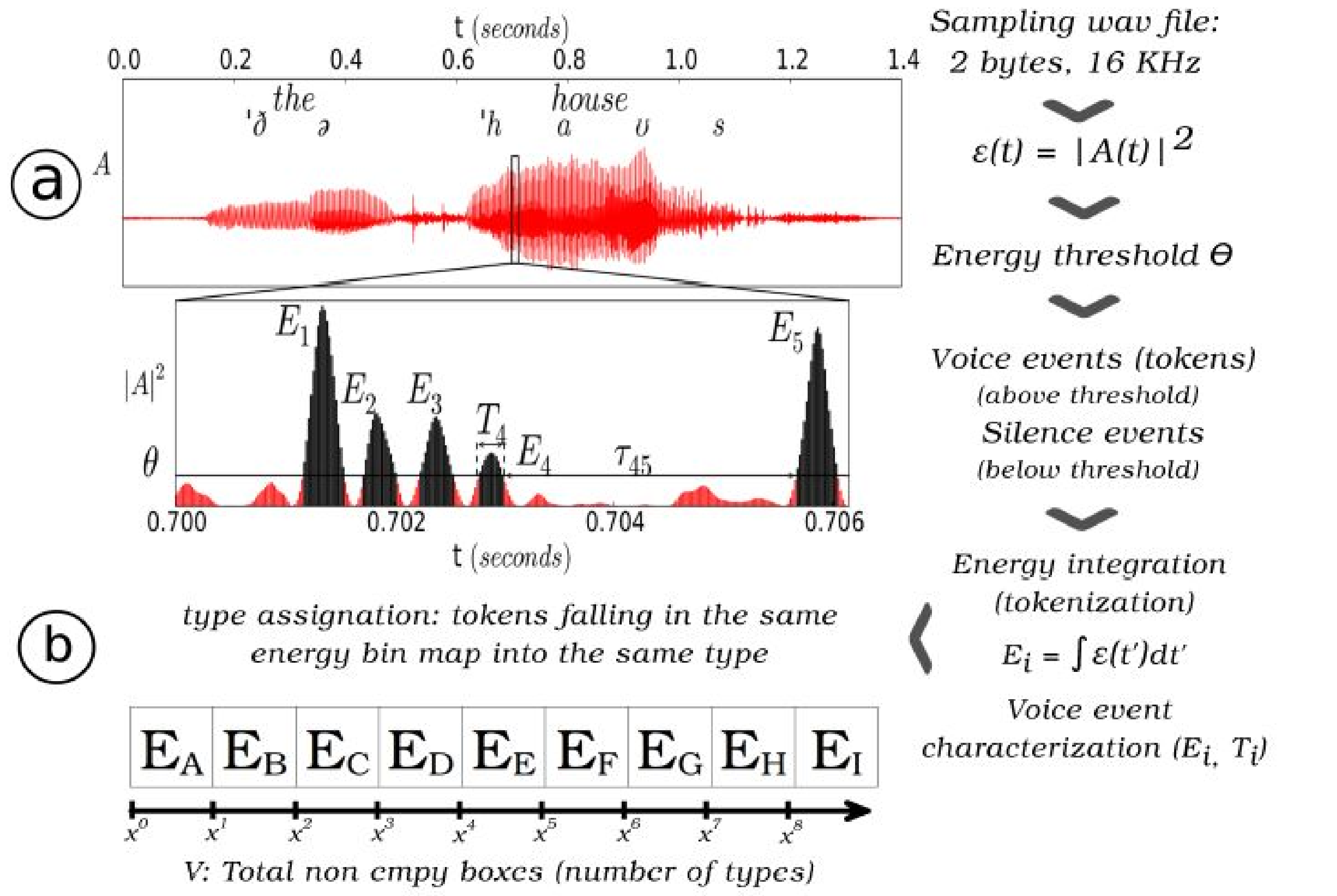
Gutenberg-Richter law. The energy E released during voice events is a direct measure of the vocal fold response function under air pressure perturbations, and its distribution has been shown to be compatible with SOC dynamics [1,2].



Data

KALAKA-2 database: 4 hours of speech recordings in different conditions ranging six different languages (Basque, Catalan, Galician, Spanish, Portuguese and English). LRE database with aditional set of 12 languages, including Japanese, Vietnamese, Mandarin, Korean, Arabic, Hindi or Tamil.

The Method



Sampling wav file: 2 bytes, 16 KHz

$\varepsilon(t) = |A(t)|^2$

Energy threshold Θ

Voice events (tokens) (above threshold)

Silence events (below threshold)

Energy integration (tokenization)

$E_i = \int \varepsilon(t')dt'$

Voice event characterization (E_i, T_i)

type assignation: tokens falling in the same energy bin map into the same type

V: Total non empy boxes (number of types)

Discussion

We have found for the first time that human voice manifests the analog of classical linguistic laws found in written texts (Zipf's law, Heaps' law , Menzerath-Altmann law and the law of abbreviation). These laws are found to be invariant under variation of the energy threshold Θ , and can be collapsed under universal functions accordingly. Our results therefore open the possibility of speculating whether the fact that these laws have been found in upper levels of human communication might be a result of a scaling process and a byproduct of the physics rather than derived from the choice of the typical units of study on the analysis of written corpus. Future work is necessary to extend this protocol to other languages and to other acoustical communication systems (i.e. non-human primates).

ACKNOWLEDGEMENTS

We acknowledge financial support from Universidad Politécnica de Madrid (IGT), Telefónica (JL), the APCOM Project TIN2014-57226-P (AHF) and grant FIS2013-41057P from Ministerio de Economia y Competitividad (BL and IGT). BL acknowledges the hospitality and support of Queen Mary University of London, where part of this research was developed, and a Salvador de Madariaga fellowship.

REFERENCES

[1] 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

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

CROSSROADS IN COMPLEX SYSTEMS

IFISC, Mallorca, June 5-8, 2017