Scaling and Universality in Human Voice

L. Lacasa[1], B. Luque[2], J. Luque[3]

(1) School of Mathematical Sciences, Queen Mary University of London
(2) Departamento de Matematica Aplicada y Estadistica, ETSI Aeronauticos, Universidad Politecnica de Madrid
(3) Telefonica Research

Speech is a distinctive complex feature of human capabilities. In order to understand the physics underlying speech production, in this work [1], we empirically analyse the statistics of large human speech datasets ranging several languages. We first show that during speech, the energy is unevenly released and powerlaw distributed, reporting a universal robust Gutenberg-Richter-like law in speech. We further show that such 'earthquakes in speech' show temporal correlations, as the interevent statistics are again power-law distributed. As this feature takes place in the intraphoneme range, we conjecture that the process responsible for this complex phenomenon is not cognitive, but it resides in the physiological (mechanical) mechanisms of speech production. Moreover, we show that these waiting time distributions are scale invariant under a renormalization group transformation, suggesting that the process of speech generation is indeed operating close to a critical point. These results are put in contrast with current paradigms in speech processing, which point towards low dimensional deterministic chaos as the origin of nonlinear traits in speech fluctuations. As these latter fluctuations are indeed the aspects that humanize synthetic speech, these findings may have an impact in future speech synthesis technologies. Results are robust and independent of the communication language or the number of speakers, pointing towards a universal pattern and yet another hint of complexity in human speech.

[1] J. Luque, B. Luque, L. Lacasa, *J. R. Soc. Interface* **12**, 20141344 (2015).