VU Research Portal

Scale-free amplitude modulation of neuronal oscillations tracks comprehension of accelerated speech

Borges, Ana Filipa Teixeira; Giraud, Anne Lise; Mansvelder, Huibert D.; Linkenkaer-Hansen, Klaus

Link to publication in VU Research Portal

Behavioral/Cognitive

Scale-Free Amplitude Modulation of Neuronal Oscillations Tracks Comprehension of Accelerated Speech

Ana Filipa Teixeira Borges,[1,2] **Anne-Lise Giraud,**[3] **Huibert D. Mansvelder,**[1,2] **and** Klaus Linkenkaer-Hansen[1,2]

[1]Department of Integrative Neurophysiology, Center for Neurogenomics and Cognitive Research, VU University Amsterdam, Amsterdam, Netherlands, [2]Amsterdam Neuroscience, Amsterdam, Netherlands, and [3]Department of Neuroscience, University of Geneva, Biotech Campus, Geneva 1211, Switzerland

Speech comprehension is preserved up to a threefold acceleration, but deteriorates rapidly at higher speeds. Current models posit that perceptual resilience to accelerated speech is limited by the brain's ability to parse speech into syllabic units using δ/θ oscillations. Here, we investigated whether the involvement of neuronal oscillations in processing accelerated speech also relates to their scale-free amplitude modulation as indexed by the strength of long-range temporal correlations (LRTC). We recorded MEG while 24 human subjects (12 females) listened to radio news uttered at different comprehensible rates, at a mostly unintelligible rate and at this same speed interleaved with silence gaps. δ, θ, and low-γ oscillations followed the nonlinear variation of comprehension, with LRTC rising only at the highest speed. In contrast, increasing the rate was associated with a monotonic increase in LRTC in high-γ activity. When intelligibility was restored with the insertion of silence gaps, LRTC in the δ, θ, and low-γ oscillations resumed the low levels observed for intelligible speech. Remarkably, the lower the individual subject scaling exponents of δ/θ oscillations, the greater the comprehension of the fastest speech rate. Moreover, the strength of LRTC of the speech envelope decreased at the maximal rate, suggesting an inverse relationship with the LRTC of brain dynamics when comprehension halts. Our findings show that scale-free amplitude modulation of cortical oscillations and speech signals are tightly coupled to speech uptake capacity.

Key words: accelerated speech; language comprehension; long-range temporal correlations; magnetoencephalography (MEG); principle of complexity management (PCM); scale-free dynamics

Significance Statement

One may read this statement in 20 –30 s, but reading it in less than five leaves us clueless. Our minds limit how much information we grasp in an instant. Understanding the neural constraints on our capacity for sensory uptake is a fundamental question in neuroscience. Here, MEG was used to investigate neuronal activity while subjects listened to radio news played faster and faster until becoming unintelligible. We found that speech comprehension is related to the scale-free dynamics of δ and θ bands, whereas this property in high-γ fluctuations mirrors speech rate. We propose that successful speech processing imposes constraints on the self-organization of synchronous cell assemblies and their scale-free dynamics adjusts to the temporal properties of spoken language.

Introduction

Human perception is remarkably robust to the variations in the way that stimuli occur in the environment. Speech is typically a stimulus from which our brain extracts consistent meaning regardless of whether it is whispered or shouted or pronounced by a male or female or a slow or fast speaker. Natural speech rate varies from two to six syllables/s (Hyafil et al., 2015b) and, depending on the original rate, can easily be decoded when accelerated up to approximately three times (Dupoux and Green, 1997). At higher rates, however, comprehension drops abruptly (Poldrack et al., 2001; Peelle et al., 2004; Ahissar and Ahissar, 2005; Ghitza and Greenberg, 2009; Vagharchakian et al., 2012).

The neural basis of this bottleneck in digesting accelerated speech information is currently unclear. It is unlikely to reflect low-level

auditory processes because speech fluctuations are well represented in auditory cortex even when speech is accelerated to unintelligibility (Nourski et al., 2009; Mukamel et al., 2011). A recent proposal involves a hierarchy of neural oscillatory processes (Ghitza and Greenberg, 2009; Ghitza, 2011; Hyafil et al., 2015a) in which the parsing of speech in the auditory system is limited to a maximal syllable rate determined by the θ rhythm, nine syllables or chunks of information per second, whereas γ oscillations can track speech at rates beyond the bottleneck in comprehension (Hyafil et al., 2015b). Scale-free amplitude modulation is a property of ongoing neuronal oscillations also referred to as long-range temporal correlations (LRTC) (Linkenkaer-Hansen et al., 2001), which may reveal how oscillations underlie the processing of natural and accelerated speech.

Scale-free dynamics are a hallmark of resting-state neuronal activity, when synchronous cell assemblies form in a largely self-organized manner (Pritchard, 1992; Linkenkaer-Hansen et al., 2001; Freeman et al., 2003; Van de Ville et al., 2010; Poil et al., 2012). Scaling analysis provides a summary descriptor of self-similarity in a system, increasingly found to correlate with its functional properties. Sensory and cognitive processing have been observed to mostly decrease (Linkenkaer-Hansen et al., 2004; He et al., 2010; Ciuciu et al., 2012), but also increase (Ciuciu et al., 2008), the strength of LRTC relative to rest, suggesting that scaling analysis can uncover the functional involvement of neuronal oscillations in specific tasks. This is further supported by the findings of LRTC covarying with individual differences in perceptual (Palva and Palva, 2011; Palva et al., 2013) and motor (Smit et al., 2013) performance. As an extension to LRTC, multifractal analysis can also reveal the neuronal involvement in cognitive tasks (Popivanov et al., 2006; Buiatti et al., 2007; Bianco et al., 2007).

Speech dynamics are also characterized by LRTC; for example, in loudness fluctuations across several time scales of radio news (Voss and Clarke, 1975), repetitions of words (Kello et al., 2008), and within-phoneme fluctuations (Luque et al., 2015). Interestingly, analyzing acoustic onsets of two speakers during conversations revealed that the scale-free dynamics of these speech signals approach one another, suggesting that a form of complexity matching underlies speech communication (Abney et al., 2014). The principle of complexity management (PCM) (West and Grigolini, 2011) has further formalized how information transfer between two complex networks derives from the cross-correlation between the output of a given complex network and the output of another complex network perturbed by the former also when their scaling properties do not match. We hypothesize that speech comprehension relates to the interplay between the scaling behavior of the brain and speech.

Here, we investigated the impact of accelerated speech on the scale-free amplitude modulation of neuronal oscillations and comprehensions. Subjects listened to radio news at rates up to an unintelligible speed during MEG acquisition. Silence gaps were also inserted in the fastest condition, creating an additional condition in which comprehension resumed. We examined how the LRTC of neuronal amplitude modulations and the LRTC of speech envelopes varied with speech rate and comprehension.

Materials and Methods

Participants
Twenty-four right-handed, healthy, native French speakers participated in this study (12 females, age 19–45 years). All participants were subject to a medical report and gave their written informed consent according to the Declaration of Helsinki. The local ethics committee approved the study.
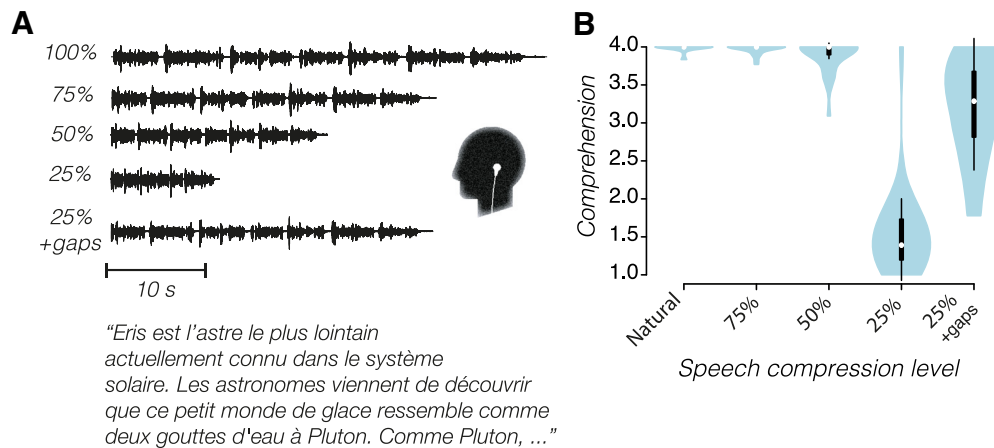
Stimuli
Speech signals consisting of 30–40 s excerpts from French radio news. The audio clips were recorded digitally at a sampling rate of 44.1 kHz in a noise-proof studio and the young female speaker was trained to keep an approximately constant intonation and rate of discourse. The excerpts (n = 7) were time compressed using the PSOLA algorithm (Moulines and Charpentier, 1990) implemented in PRAAT software (Boersma, 2001). The compression alters the duration of the formant patterns and other spectral properties but keeps the fundamental frequency ("pitch") contour of the original uncompressed signals. Four different rates (25%, 50%, 75%, and 100% of natural duration) were created for the speech stimuli. In addition, an extra condition with a 60 ms silent gap inserted at every 40 ms chunk of the most compressed rate was created. The goal was to restore speech comprehension by approaching natural syllabicity (Ghitza and Greenberg, 2009). In natural speech, syllable durations have variable durations of ~50–400 ms with a mean of ~250 ms. The statistical distribution of syllable durations constrains speech to have long-term regularities; in particular, an envelope modulation spectrum <20 Hz with most temporal modulations <6 Hz (Greenberg et al., 2003; Hyafil et al., 2015b). Gap insertion does not mimic syllable statistics but reinstates an artificial slower rhythm that repackages the compressed information into longer time frames. Importantly, this rate of chunking preserves the perception of speech as a continuous stream (Bashford et al., 1988). All stimuli were interleaved in a pseudorandom order, excluding the possibility of presenting the same text consecutively. Each stimulus had a preceding baseline period of 5 s and a period after for comprehension rating by means of a right-hand keypad button press. Subjects had to choose between four possible answers where 1 was nothing; 2 was some words; 3 was some phrases; and 4 was whole text. The scale is subjective, relying on the participant's own comprehension assessment as opposed to an objective rating of comprehended speech; however, the participants did not express difficulty in using the scale and, on the basis of a previous behavioral study using the same type of stimuli (Pefkou et al., 2017), we are confident that this assessment closely reflects the actual comprehension level. Before and after the baseline, subjects had a 2 s period for blinking their eyes if needed, which helped to minimize eye movements during speech listening and to mitigate their effect with the preprocessing. The whole experiment aimed to recreate a naturalistic condition of listening to the news on the radio.

MEG acquisition
Whole-scalp brain magnetic fields of the participants were collected at the Hôpital Pitié Salpetrière (Paris, France) using the whole-head Elekta Neuromag Vector View 306 MEG system (Elekta Neuromag TRIUX) equipped with 102 sensor triplets (two orthogonal planar gradiometers and one magnetometer/position). Before the recordings, four head position indicator coils attached to the scalp determined the head position with respect to the sensor array. The location of the coils was digitized with respect to three anatomical landmarks (nasion and preauricular points) with a 3D digitizer (Polhemus Isotrak system). Then, the head position with respect to the device origin was acquired before each block. Data were sampled at 1 kHz and filtered at 0.1–330 Hz. Stimulus delivery was performed in MATLAB version 2001a (The MathWorks) using the Psychophysics Toolbox Psychtoolbox-3.0.8 extensions (Brainard, 1997; Pelli, 1997). The subjects were sitting comfortably in a magnetically shielded room during the recordings and previously instructed about the task. The sound volume on the earphones was adjusted to a comfortable level subjectively determined for each subject. While listening to the stimuli, the subjects were instructed to look at a fixation cross on the screen; the cross would flicker during the periods allocated to blink/rest the eyes. The experiment involved playing the stimuli twice and was split into 8 blocks of ~5 min with a rest interval after 4 blocks.

MEG preprocessing
Data were preprocessed with Signal Space Separation algorithm implemented in MaxFilter (Tesche et al., 1995; Taulu et al., 2005) to reduce noise from the external environment and compensate for head movements. Manually detected bad channels (noisy, saturated, or with SQUID jumps) were marked "bad" before applying *MaxFilter* and the latter was

Figure 1. Speech stimuli and their intelligibility. ***A***, Example of speech stimuli waveforms and excerpt from one of the radio news used (in French) for each of the five conditions: natural rate, compression to 75%, 50%, or 25% of natural duration, and a condition in which the 25% stimuli had silent gaps of 60 ms inserted every 40 ms. ***B***, Violin plots of the distribution of speech comprehension across participants for each of the five conditions. The scores of comprehension range from 1 (nothing) to 2 (some words) to 3 (some phrases) to 4 (whole text). White dots show the medians; box limits indicate the 25th and 75th percentiles; whiskers extend 1.5 times the interquartile range from the 25th and 75th percentiles; light blue polygons represent density estimates of ratings and extend to extreme values.

also used to identify other potential bad channels. All of these were subsequently interpolated. Head coordinates recorded at the beginning of the first block were used to realign the head position across runs and transform the signals to a standard position framework. Afterward, physiological artifacts in the sensors were identified using principal component analysis and removed with the signal space projection method (Uusitalo and Ilmoniemi, 1997) based on the projections of the ECG and EOG also recorded. The clean files were subsequently processed to extract only the parts of recording corresponding to each of the conditions of stimulation. All speech narratives from the same condition were merged for each participant. The raw files were converted to MATLAB format using the MNE MATLAB toolbox (Gramfort et al., 2013; www.martinos.org/mne/). Finally, all converted files were inspected for transient artifacts probably of muscular origin (e.g., jaw or neck movements) and clipped and discarded from the analysis. At most, data amounting to 3 s were lost by clipping.

Experimental design and statistical analysis
In this within-subject design with $N = 24$ subjects, we wanted to test the effect of speech rate in the spectral power and detrended fluctuation analysis (DFA) quantified from the MEG recordings. The independent variable, speech rate, varied across five conditions, a multiple measurements paradigm. The distributions of the dependent variables, spectral power and DFA exponents of the combined gradiometers in all frequency bands, were tested for normality using the Lilliefors test (Lilliefors, 1967). In at least one of the rate conditions, >20% of sensor locations had either power or DFA biomarkers deviating significantly ($p < 0.05$) from a normal distribution and, for some biomarkers, this was true for >50% of the sensors. Therefore, we opted for nonparametric statistical methods. The Friedman's test (Friedman, 1937) was applied to all conditions and the two biomarkers for each of the neuronal activity bands. *Post hoc* analysis based on Wilcoxon signed-rank tests (Wilcoxon, 1945) was applied to the subset of biomarkers that differed significantly across conditions. Spectral power and DFA was computed as the average of all sensor pairs (global parameters) or for each sensor pair. In the first case, a Bonferroni correction was applied ($p < 0.005$, 10 comparisons). For the sensor-level analysis, to control for type II errors due to multiple comparisons, we used the false discovery rate correction (FDR) as follows:

$$p < \frac{0.05}{N} = 5.0 \times 10^{-4}, N = 102 \text{ combined sensors; Benjamini and}$$

Hochberg, 1995) as implemented in Groppe et al. (2011). No statistical method was used to determine sample size. During acquisition, conditions were randomized to mitigate any carryover effects such as practice or fatigue. The behavioral scores of the different conditions were compared with a Friedman test and *post hoc* analysis was done using the

Wilcoxon signed-rank test with the Bonferroni method for multiple-comparisons correction.

Behavioral analysis
To assess perceived comprehension of the speech narratives within each condition, we computed the mean rating across all 14 stimuli for each participant. This value is referred to as comprehension.

Data analysis
For the MEG data analysis, we used adapted functions from the Neuro-physiological Biomarker Toolbox (NBT Alpha RC3, 2013, www.nbtwiki.net; Hardstone et al., 2012), together with other MATLAB scripts (R2011a; The MathWorks). The analysis of time-averaged spectral power and LRTC in the modulation of amplitude envelopes were performed in the following frequency bands: δ (1–4 Hz), θ (4–8 Hz), α (8–13 Hz), β (13–30 Hz), γ (30–45), and high-γ (55–300 Hz). The band-pass filtering used finite impulse response filters with a Hamming window and filter orders equal to 2000 (δ-band), 500 (θ-band), 250 (α-band), 154 (β-band), 67 (γ-band), and 36 (high-γ-band) ms.

Estimation of spectral power
Spectral power was estimated by applying the Welch's modified periodogram method implemented in MATLAB as *pwelch*() function with non-overlapping Hamming windows of 1 s and the values shown are the square root of the power spectral density obtained. For statistical analysis, we used the vector sum (root-mean-square) of amplitudes in each pair of planar gradiometers (Nicol et al., 2012).

Estimation of LRTC
MEG. The amplitude envelope of the band-passed signals was calculated using the magnitude of the Hilbert transform. Next, we estimated the monofractal scaling exponents using DFA (Peng et al., 1994, 1995), a well established technique for studying the amplitude dynamics of neuronal oscillations (Linkenkaer-Hansen et al., 2001). Details of the method have been described previously (Peng et al., 1994; Kantelhardt et al., 2001; Hardstone et al., 2012). In brief, the DFA measures the power law scaling of the root-mean-square fluctuation of the integrated and linearly detrended signals, $F(t)$, as a function of time window size t (with an overlap of 50% between windows). The DFA exponent (α) is the slope of the fluctuation function $F(t)$ and can be related to the power law scaling exponent of the auto-correlation function decay (γ) and the scaling exponent of the power spectrum density (β) by the following

expression: $\alpha = \dfrac{1+\beta}{2} = \dfrac{2-\gamma}{2}$. DFA exponent values between 0.5 and 1.0 reveal the presence of LRTC, whereas an uncorrelated signal has an exponent value of 0.5. The decay of temporal correlations was quantified over a range of 1–9 s for each of the speech rate conditions using a signal in which the 14 recordings from each condition had been concatenated. The rationale for studying the dynamics within this time-range was to confine the analysis of (auto-)correlations to a period corresponding to the same audio excerpt. The analysis window was thus constrained by the duration of the condition with the fastest rate (compression to 25% of natural duration). Specifically, we calculated the fluctuation function in windows of 1.1, 1.3, 1.7, 2.2, 2.7, 3.5, 4.4, 5.5, 7.0, and 8.9 s. Even though rigorous usage of the term "scale-free dynamics" traditionally relies on assessing the scaling over several orders of magnitude, analyzing the scaling on more limited time frames is technically possible and meaningful (Avnir, 1998) and has proven useful as a quantitative index that captures brain function (Buiatti et al., 2007; Linkenkaer-Hansen et al., 2007; Hardstone et al., 2012). In addition, analysis of task-related activity inherently imposes a time frame restriction because prolonged stimulation can induce confounds such as fatigue/inability to concentrate or even crossovers in scaling behavior. The scaling exponents were obtained separately for each of the gradiometers and an average of each pair calculated.

Speech. Loudness fluctuations in speech are also known to exhibit long-range temporal correlations (Voss and Clarke, 1975). To study the effect of compression on the temporal correlation structure of the speech stimuli, we concatenated the 7 different stimuli and filtered the resulting signal in the audio range 0.1–20 kHz (FIR filter, order 20), in which we determined the amplitude envelope using the Hilbert transform. The strength of LRTC was estimated in this broadband envelope as well as after low-passing the envelope (cutoff frequency 20 Hz; causal FIR filter order 4410) according to the approach described previously (Voss and Clarke, 1975). We estimated the DFA scaling exponents in the same time range as the neuronal oscillations: 1–9 s.

Correlation analysis
To perform correlation analyses between the behavioral scores and biomarkers computed, we applied the Spearman's rank correlation coefficient to quantify associations between power and power law scaling exponents in the condition 25% of natural duration for each sensor pair location and the averaged comprehension of the 14 segments. Interpretation of effect sizes followed Cohen's guidelines (small effect $\rho > 0.1$, medium $\rho > 0.3$, and large $\rho > 0.5$; Cohen, 2013). One subject always reported full comprehension (text comprehension) when the speech was 25% compressed and it was a posteriori excluded from analysis.

The colors used in graphical representation were based on the map colors by Cynthia A. Brewer (https://github.com/axismaps/colorbrewer/). The boxplots and violin plots were produced with the help of the R-based software BoxPlotR (Spitzer et al., 2014).

Estimation of information transfer
The PCM establishes an account of the degree of information transfer between two systems. One measure of information in a network is given by the probability density of its events, which has the following survival probability:

$$\psi(t) = \frac{T^{\mu-1}}{(T+t)^{\mu-1}} \tag{1}$$

Where T and μ characterize the complexity of the system and μ is the power law exponent in the range $1 < \mu < 3$. The PCM predicts that, for systems that are in the so-called ergodic regime ($2 < \mu < 3$), one can quantify the cross-correlation in the asymptotic limit (Φ_∞) between a complex network P and a complex network S being perturbed by P based on the following relationship between the power law exponents of both networks (Aquino et al., 2007, 2010, 2011):

$$\Phi_\infty = \frac{\mu_s - 2}{\mu_s + \mu_p - 4} \tag{2}$$
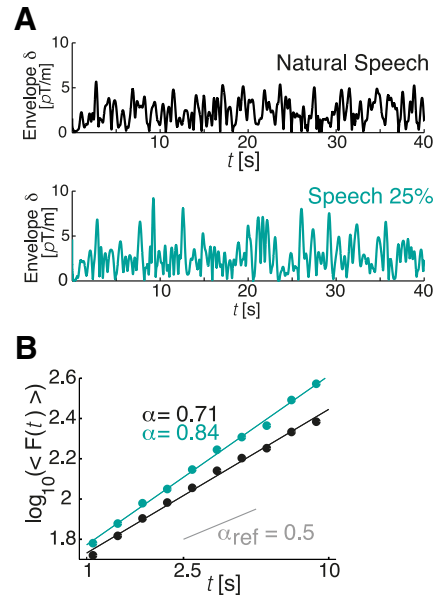
Figure 2. Illustration of the amplitudes and power law scaling of amplitudes of MEG responses to natural and 25% compressed speech. Representative envelopes of MEG activity (40 s) in the δ band recorded while listening to natural rate speech (top) and speech compressed to 25% (bottom). *B*, The traces featured in *A* reveal autocorrelations that follow a power law in the range of 1–9 s and can be characterized by the DFA exponent (α); while listening to speech compressed to 25%, α is higher than while listening to the natural speech.
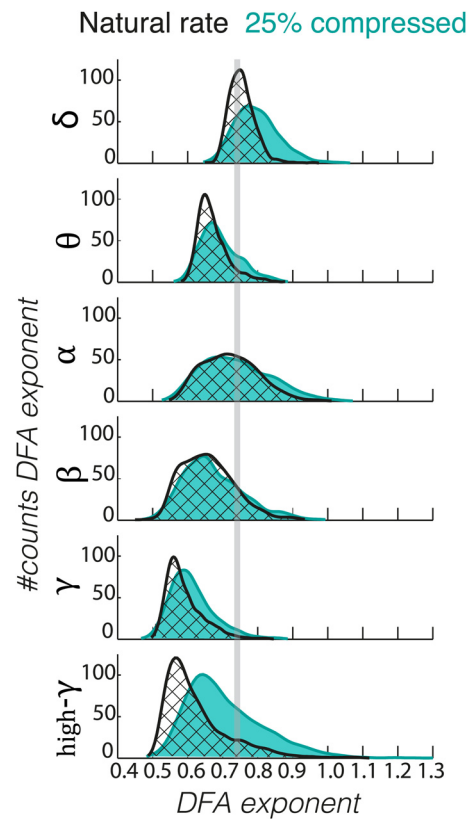
Figure 3. MEG gradiometer signals during natural rate display narrower ranges of scaling exponents compared with unintelligibly compressed speech. Kernel smoothing estimates of the distributions of DFA exponents for all sensor gradiometer pairs and subjects ($n = 24$) for each of the studied frequency ranges (δ, θ, α, β, γ, and high-γ) in natural rate (black line) and 25% compression (green). α and β bands were the least affected by the semantics, whereas δ, θ, and γ oscillations had narrower DFA exponent distributions in the natural rate condition.

Table 1. Global spectral power and DFA (averaged over sensors) for each speech condition

Frequency band (Hz)	Natural rate	75%	50%	25%	25% + gaps	Wilcoxon T^c ($n = 24$)	p	r
Spectral power[a] (pT/m)								
1–4	152 (27.3)	150 (26.8)	156 (27.5)	160 (27.9)	151 (26.7)	0	<0.001*	0.62
4–8	116 (26.0)	117 (25.9)	118 (25.8)	117 (25.6)	116 (24.8)	69	0.021	0.33
8–12	117 (38.2)	118 (38.4)	116 (38.6)	116 (39.1)	116 (38.8)	115	0.317	0.14
12–30	65.8 (14.8)	66.0 (14.7)	65.7 (15.0)	66.1 (15.4)	64.3 (14.7)	130	0.568	0.08
30–45	39.5 (5.9)	39.8 (5.9)	39.4 (6.0)	39.6 (6.1)	39.4 (5.9)	28	<0.001*	0.50
55–300	26.1 (3.5)	26.5 (3.5)	26.1 (3.5)	26.1 (3.5)	26.0 (3.4)	64	0.014	0.35
DFA[b]								
1–4	0.756 (0.010)	0.755 (0.010)	0.766 (0.012)	0.798 (0.017)	0.755 (0.012)	0	<0.001*	0.62
4–8	0.668 (0.010)	0.669 (0.008)	0.666 (0.010)	0.692 (0.011)	0.662 (0.008)	0	<0.001*	0.62
8–12	0.716 (0.030)	0.716 (0.030)	0.716 (0.033)	0.733 (0.032)	0.715 (0.028)	0	<0.001*	0.62
12–30	0.658 (0.031)	0.654 (0.031)	0.658 (0.033)	0.678 (0.032)	0.659 (0.028)	0	<0.001*	0.62
30–45	0.5903 (0.015)	0.595 (0.015)	0.597 (0.017)	0.6141 (0.017)	0.597 (0.018)	0	<0.001*	0.62
55–300	0.613 (0.041)	0.657 (0.050)	0.642 (0.036)	0.702 (0.032)	0.631 (0.018)	0	<0.001*	0.62

The data represent median (SD) across $n = 24$ subjects. The Wilcoxon test indicates the significance and effect size of the difference between the natural rate and the 25% condition.

[a]Friedman test, effect of rate × frequency $\chi^2_{(4)}$ ($n = 24$) = [79, 72, 24, 58, 49, 41], $p < 0.001$.

[b]Friedman test, effect of rate × frequency $\chi^2_{(4)}$ ($n = 24$) = [79, 73, 51, 45, 55, 77], $p < 0.001$.

[c]Comparison natural rate versus 25%.

*Significance $p < 0.005$ (Bonferroni threshold).

Or the equivalent for the DFA scaling exponents following the hyperscaling relation valid for both the ergodic and nonergodic regimes $\left(\alpha = \dfrac{4 - \mu}{2} \right)$ (Kalashyan et al., 2009) as follows:

$$\Phi_\infty = \frac{1 - \alpha_s}{2 - \alpha_s - \alpha_p} \tag{3}$$

with $0.5 < \alpha_s < 1$ and $0.5 < \alpha_p < 1$.

We used this expression to derive an estimate of information transfer during natural speech and speech compressed to 25% of its initial duration. The cross-correlation cube (Aquino et al., 2007, 2010, 2011) demonstrates that, except when a perturbing network P is ergodic and the perturbed system S is nonergodic, all stimuli modify the properties of the responding network. A special case of perfect matching (correlation = 1) occurs when $\alpha \geq 1$. From Equation 3, it follows that the cross-correlation can vary from ∼0.04 to ∼0.96 and that the highest cross-correlation occurs for the highest α_p and the lowest α_s. Furthermore, for perturbations that are nearly white noise, there is less information transfer, but its degree depends on the scaling index of the perturbed system.

Results

Speech comprehension

To study neuronal dynamics and language comprehension while listening to accelerated speech, we compressed a series of radio excerpts to 100%, 75%, 50%, and 25% of their original duration (Fig. 1A; see Materials and Methods). Participants rated comprehension of the news, which was close to perfect for the natural rate, as well as for the compression to 75% or 50% of the original duration (Fig. 1B). At compression to 25%, the news excerpts were nearly incomprehensible (median score of 1.4). Insertion of 60 ms silence gaps every 40 ms increased comprehension markedly. Comprehension differed significantly across conditions (Friedman test; $\chi^2_{(4)} = 86$, $p < 0.001$). In particular, comprehension at the highest speed (25%) abruptly deteriorated compared with 50% compression ($\Delta Mdn = -2.6$, $r = -0.61$, $p = 2.7 \times 10^{-5}$) and was partly restored with the insertion of silence gaps ($\Delta Mdn = 1.9$, $r = 0.60$, $p = 2.7 \times 10^{-5}$). Although comprehension was expectedly high at both 50% and 100% of natural duration, it was more variable in the fastest condition (50%). Finally, there was a larger intersubject variability and the news was harder to understand in the gap condition than during the 50–100% of natural duration rates (e.g., 50% vs 25% + gaps: $\Delta Mdn = 0.7$, $r = 0.61$ $p = 4.0 \times 10^{-5}$).

Spectral power and scaling of neuronal oscillations and speech signals

To investigate the neuronal correlates of speech comprehension, we compared MEG activity during the five speech conditions. The temporal dynamics of MEG activity was quantified using DFA (see Materials and Methods), which estimates the LRTC of the amplitude fluctuations (Fig. 2). Individual DFA exponents (α) across subjects, sensors, and oscillations spanned the range of ∼0.5–1 (Fig. 3), indicating that the amplitude modulation of oscillations exhibited power-law LRTC.

Global changes in spectral power and DFA exponents (averaged across all sensors) varied significantly with speech rate. Specifically, power increased in the δ and γ bands and DFA increased across all frequency bands between the natural and 25% compressed speech (Table 1). At this global level, the DFA of high-γ was the only parameter to show significant changes across consecutive speech rates ($p < 0.002$).

We proceeded with considering regional differences by means of quantifying changes in the MEG sensor pairs. Differences in the strength of LRTC were pronounced for the δ and high-γ activities, with p-values <0.05 (Friedman test) in 88% and 90% of sensors, respectively. DFA exponents of θ, α, β, and γ revealed significant differences ($p < 0.05$) in 44%, 22%, 24%, and 50% of the sensor locations, respectively. For spectral power, however, natural speech only differed from 25% compressed speech in the δ band (Friedman test, 94% of the sensor locations). Together, these data suggest that the spectral power of neuronal oscillations and LRTC probe different aspects of speech processing. To further investigate the direction of the effect of speech acceleration on the amplitude and amplitude dynamics of the neuronal oscillations, we conducted *post hoc* analyses using Wilcoxon signed-rank sum tests.

While listening to the fourfold-compressed speech (25% of natural duration), the spectral power in the δ band increased across all regions of the sensor array compared with the natural rate (Fig. 4A–C,E). The analysis of scale-free dynamics revealed a greater dissociation between speech rates both across frequency ranges and anatomical regions. LRTC of amplitude envelopes exhibited negligible differences across the three fully intelligible conditions (100%, 75%, and 50%) for δ, θ, α, β, and γ oscillations. In contrast, LRTC in the high-γ band increased monotonically with
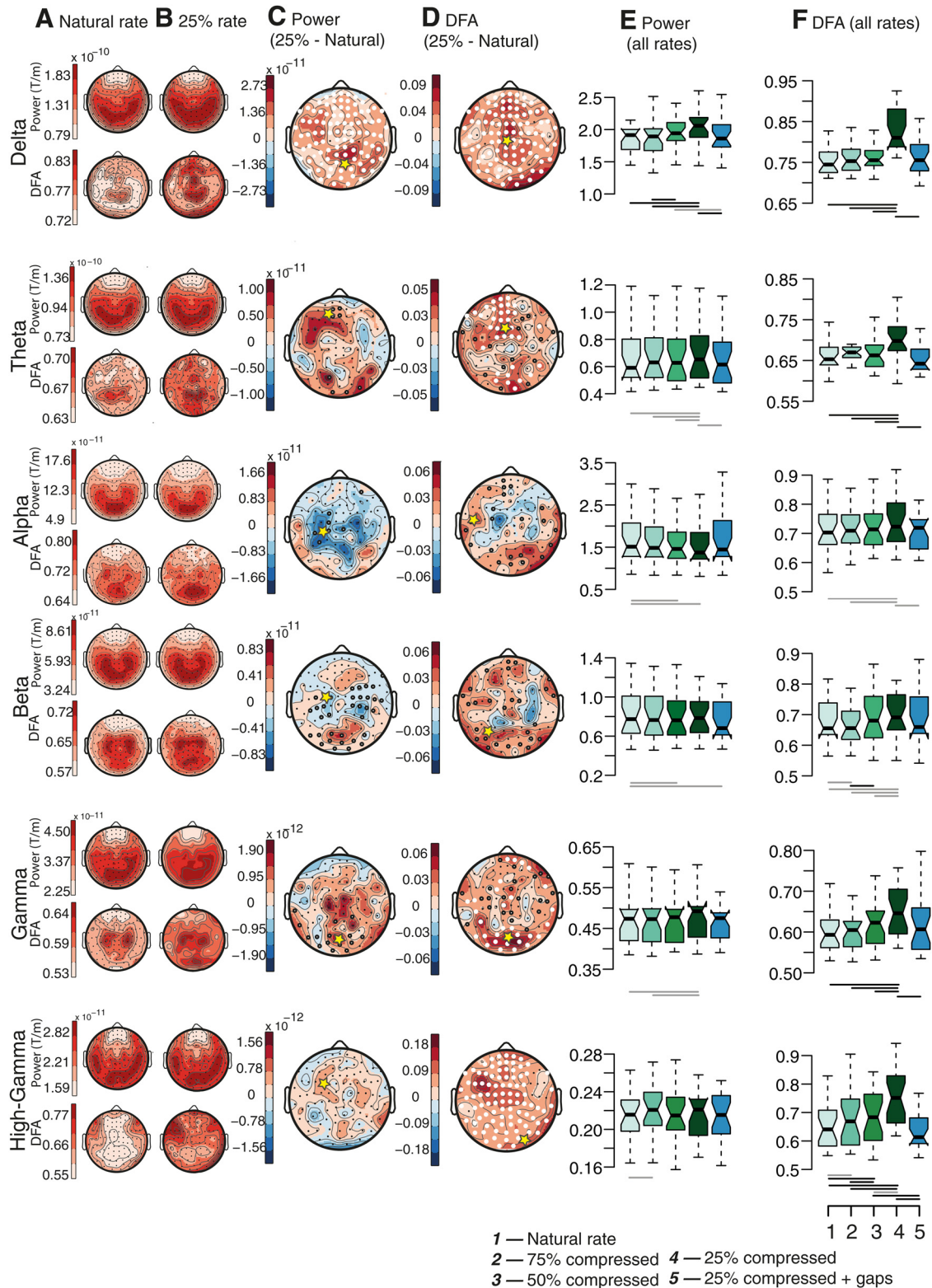
Figure 4. Scale-free dynamics of delta, theta and gamma bands reflects the comprehension of accelerated speech. Topographies of time-averaged power (odd rows) and DFA exponents (even rows) in the δ, θ, α, β, γ, and high-γ range are shown for the natural rate (**A**) and the 25% time compressed (**B**) conditions. The difference topographies of conditions 25% time compressed minus natural rate are shown for amplitude (**C**) and DFA exponents (**D**), with significance levels indicated with black circles (Wilcoxon signed-rank sum tests, significance threshold $p < 0.05$, $n = 24$) or white circles (significant differences after correction for multiple comparisons with FDR $< 5.0 \times 10^{-4}$). Boxplots of spectral power (**E**) and DFA exponents (**F**) for each of the five conditions (natural rate, 75%, 50%, and 25% compression and 25% compression with periodic gaps inserted) for sensors marked with yellow stars in **C** and **D**. A significant difference between any two conditions is shown with a horizontal gray line (uncorrected) or black line (corrected) as in **C** and **D**. Center lines show the medians; box limits indicate the 25th and 75th percentiles, whiskers extend 1.5 times the interquartile range from the 25th and 75th percentiles. Note that DFA exponents of the δ, θ, and γ oscillations are the only biomarkers that specifically reflect comprehension of the speech conditions, whereas high-γ also responds to the acceleration of speech per se.

rate, mainly in the frontal, centroparietal, and cerebellar regions of the sensor array (Fig. 4F). Importantly, responses to 25% compressed speech showed elevated LRTC relative to the natural speed for several oscillations (Fig. 4D,F). For the δ and high-γ bands, the increase occurred nearly ubiquitously across the scalp [e.g., $\Delta Mdn(\delta) = 0.066$, $T = 7$, $r = 0.59$, $p_{corrected}$ (two-tailed) = 0.0005; ΔMdn(high-γ) = 0.111, $T = 3$, $r = 0.6$, $p_{corrected}$ (two-tailed) = 0.0001; Fig. 4D,F]. For θ, α, β, and γ bands, the power law exponents increased in sensors located over occipital, cerebellar, and frontal regions. The magnitude and topography of the DFA exponents increases in the 25% compressed condition relative to the natural speech rate were very similar to those of the contrast between the 25% compressed condition and the intermediate rates (75–50%; Fig. 5). The selective increase that occurred only in the 25% compressed condition and mostly in the δ/θ bands suggests that LRTC of these slow oscillations possibly reflect extraction of meaning from speech. The condition in which speech was compressed four times thus differed both quantitatively and qualitatively from all other conditions. Importantly, the insertion of silence gaps in the most compressed condition, which nearly restored comprehension (Fig. 1B), reduced the scaling exponents (Fig. 4F) to values close to those observed at the natural rate.

To address whether changes in LRTC in the amplitude modulation of oscillations solely reflect a stimulus-driven dynamics such as increasing LRTC with compression, we also analyzed the speech stimuli using the DFA method (see Materials and Methods). The amplitude envelope of the broadband and low-pass-filtered amplitude envelopes exhibited temporally structured fluctuations at all compression levels (Fig. 6A), approximating a power law function (Fig. 6B). In contrast to the scaling of neuronal oscillations, however, DFA exponents decreased with speech compression. Whereas the rates by design decreased linearly, scaling exponents exhibited subtle decreases ($\Delta\alpha \approx 0.01$) at the two moderate compression rates (75% and 50%), followed by an abrupt reduction in the most compressed speech ($\Delta\alpha \approx 0.07$; Fig. 6C). Together, our results show that speech comprehension is closely coupled to the power law scaling of multiple oscillation envelopes in several brain regions in a fashion not trivially linked to the physical features of the stimuli.

## Spectral power and LRTC are associated with comprehension

The behavioral scores showed that comprehension mostly collapsed for the 25% compressed speech, varying from "nothing understood" (1) to "some words understood" (2) (Fig. 1B). To explore whether the individual variation in comprehension was associated with individual variation in the amplitude and power law scaling behavior, we correlated speech comprehension scores in the 25% condition with the amplitudes or DFA exponents across the sensor array. We found that amplitudes in frontal areas correlated strongly and positively with comprehension, especially in the δ, β, γ, and high-γ bands(e.g., $\rho_{\delta(21)} = 0.53$, $p_{(two-tailed)} = 0.01$; Fig. 7A–C). In contrast, DFA exponents exhibited negative correlations in most regions and frequency bands, reaching significance in central and temporoparietal sensor regions of δ and θ bands (peak at $\rho_{\delta(21)} = -0.62$ $p_{(two-tailed)} = 0.002$; Fig. 7D–F). These observations underscore that speech comprehension is associated with a reorganization of the temporal structure of δ and θ oscillations and that the time-averaged spectral power and the LRTC of the amplitude modulation of oscillations have an inverse relationship with speech comprehension.
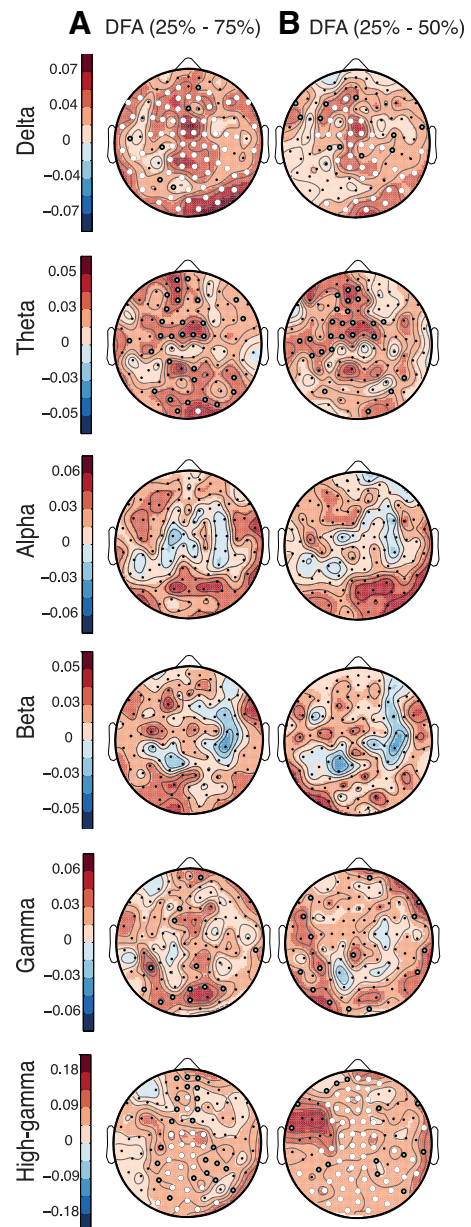
Figure 5. Neural LRTC increase in the most accelerated speech in contrast to moderate speech rates. Topographical differences between the DFA scaling exponents obtained during the 25% minus the 75% (**A**) and 50% (**B**) time-compressed speech conditions for the different frequency bands. Significance levels indicated with black circles (Wilcoxon signed-rank sum tests, significance threshold $p < 0.05$, $n = 24$) or white circles (significant differences after correction for multiple comparisons with FDR $< 5.0 \times 10^{-4}$). Scale-free properties are relatively constant between moderate rates and the fastest rate of speech.

As a group tendency, the scaling exponents of the speech and brain activity (Fig. 8A) diverged with increasing speech rate. Specifically, the α_{brain} approached 0.8 (Fig. 4) and the α_{speech} reached 0.6 (Fig. 6) in the 25% compression condition compared with approximately equal scaling exponents ($\alpha_{brain} \approx \alpha_{speech}$) at the slower rates and when silent gaps were inserted to restore comprehension (Fig. 8B). Using the cross-correlation measure of information transfer (Φ_∞) applied to the interaction of speech and brain (see Eq. 3 in the Materials and Methods; Fig. 8C), we found that information transfer was ~0.6 during the natural rate and merely ~0.2 in the fastest rate of speech (Fig. 8D). Therefore, it may be that comprehension is related to information transfer.
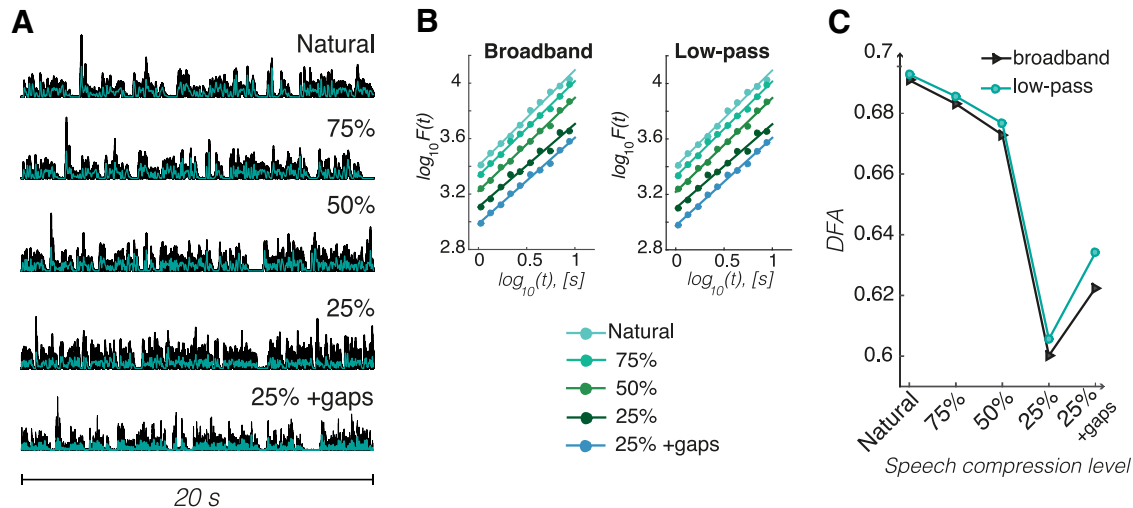
Figure 6. Long-range correlations of speech envelope decrease with time compression. **A**, Envelopes of speech stimuli in the five rate conditions (20 s excerpt) filtered in the range 0.100–20 kHz (black line) and after low-pass filtering at 20 Hz (green line). **B**, Double logarithmic plots of characteristic fluctuation size [$F(t)$] versus box size (t) display power law scaling within the interval [1,9] s for both the broadband envelopes (left) and the 20 Hz low-pass envelopes (right). The linear relationship reveals the presence of self-similarity (scaling) between the fluctuation at small and larger boxes. The slopes are equal to the DFA exponent. The lines were vertically shifted to aid visualization. **C**, DFA exponent of the speech envelopes (broadband and 20 Hz low-pass) decreases with the increase of speech rate and the insertion of gaps in the 25% compressed speech augments the exponent moderately.
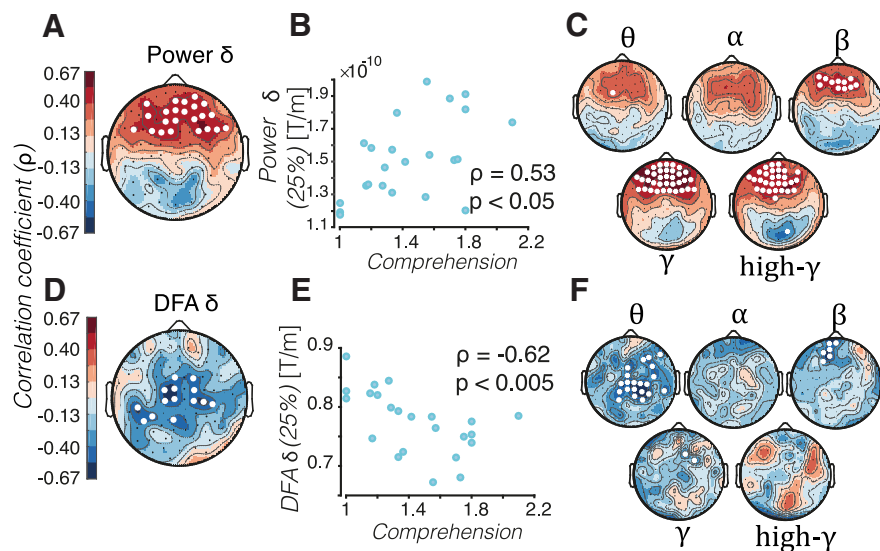
Figure 7. Spectral power and scale-free dynamics of oscillations are linked to comprehension of severely compressed speech. **A**, **C**, Topographical maps denote the Spearman's correlation coefficient (ρ) between the comprehension in the 25% condition and the amplitude (**A**) or DFA (**C**) exponent in the δ band; significant changes are marked with white circles (level of significance 0.05). **B**, **D**, Individual results for the sensor pair marked with a yellow star in **A** and **C**. δ power correlates with comprehension in the frontal region, whereas DFA decreases as comprehension increases in the medial and left temporal regions. **C**, **F**, Same as in **A** and **C**, but, for the remaining five frequency bands labeled with the correspondent Greek letters, the biomarker values are color coded in the same range.

Moreover, following the relationship of Equation 3, with a scaling exponent (α_{speech}) equal to 0.6 during the fastest rate (where comprehension deteriorated), the cross-correlation can amount maximally to ~0.6 and reach zero if the α_{brain} approaches but does not equal one. This low level of information transfer is congruent with the fact that subjects who understood nothing of the speech showed the highest α_{brain} (Fig. 7E).

Discussion

We investigated how speech processing and comprehension are coupled to the spectral power and scale-free amplitude modulation of neuronal oscillations recorded with MEG. We found that

LRTC prominently in the δ, θ, and γ bands mirrored the abrupt change in the speech comprehension at high rates of presentation, whereas high-frequency γ fluctuations (55–330 Hz) displayed a progressive increase in LRTC with speech acceleration when rates were comprehensible. These findings suggest that the scale-free amplitude dynamics of neuronal oscillations can reflect processes associated with speech comprehension. Interestingly, the semantics-related increase in neuronal LRTC occurred when the time-compressed acoustic speech signal had reduced LRTC.

Roles of power and scale-free dynamics in speech processing

We confirmed here that the temporal structure of oscillations can be modulated independently of its amplitude, as several studies characterizing neuronal dynamics in healthy subjects have observed (Nikulin and Brismar, 2005; Linkenkaer-Hansen et al., 2007; Smit et al., 2011). Both when comparing intelligible and unintelligible conditions and when looking at intersubject variability in the fastest condition, we observed that comprehension was associated with weak LRTC. High performance has previously been associated with lower scaling exponents, for example, in an audiovisual coherence detection task (Zilber et al., 2013), an auditory target-detection task (Poupard et al., 2001), or a sustained visual attention task (Irrmischer et al., 2017). Long-range temporal correlations are also suppressed in behavioral time series during tasks with high memory load (Clayton and Frey, 1997) and unpredictability (Kello et al., 2007). One possible interpretation of such findings is that the regime of reduced LRTC facilitates information processing (He, 2011). A reduction of LRTC might

arise from demanding exogenous constraints that prompt rapid reorganization of cortical assemblies and reduce their intrinsic propensity to participate in a large repertoire of spatiotemporal patterns as those observed during self-organized neuronal activity (Plenz and Thiagarajan, 2007). Notwithstanding, reasoning tasks less strictly shaped by external demands are characterized by higher scaling regimes (Buiatti et al., 2007) and behavioral tasks of repetitive nature such as estimating periodic intervals (Gilden et al., 1995) or repeating a word (Kello et al., 2008) approach closely 1/f dynamics, suggesting that the scaling regime depends on the nature of the task. Our finding that neural LRTC either increase or remain unaltered with faster rates challenges the view that LRTC generally reduces with effort (Churchill et al., 2016); however, we cannot exclude that the abrupt increase in LRTC in δ and θ oscillations reflects less responsiveness to the stimuli and the emergence of brain dynamics more reminiscent to rest.

The taxing imposed by the accelerated pace increases memory load, which is known to increase the power of δ, θ, and γ oscillations in the frontal lobes (Gevins et al., 1997; Jensen and Tesche, 2002; Howard et al., 2003; Onton et al., 2005; Zarjam et al., 2011). Under the fastest speed condition, positive correlations were found between the individual subject power in the δ, β, γ, and high-γ activity in frontal regions and speech comprehension. However, at the group level, δ power increased significantly in this condition relative to the other three slower-paced conditions despite hampered comprehension. Therefore, increased power plausibly reflects a domain-general mechanism engaged by the difficulty in language understanding (Fedorenko, 2014), which assists but does not guarantee meaning retrieval. Future studies should include a control condition of equally time-compressed random sound sequences devoid of meaning to dissect thoroughly the relative contribution of comprehension success (semantics) and comprehension effort (commensurate to speech rate) to neural activity.

Contributions of frequency and anatomy

Several studies have implicated δ, θ, and γ oscillations in speech comprehension (Giraud et al., 2007; Luo et al., 2010; Peelle et al., 2013; Henry et al., 2014; Lewis et al., 2015; Lam et al., 2016; Mai et al., 2016; Keitel et al., 2017). Conceivably, sensory selection involves a hierarchical coupling between lower and higher frequency bands (Lakatos et al., 2008; Giraud and Poeppel, 2012; Gross et al., 2013). Temporal modulations (1–7 Hz) of the speech envelope are crucial (Elliott and Theunissen, 2009) and may suffice for comprehension (Shannon et al., 1995). Their representation by neuronal dynamics appears therefore sensible. Con-
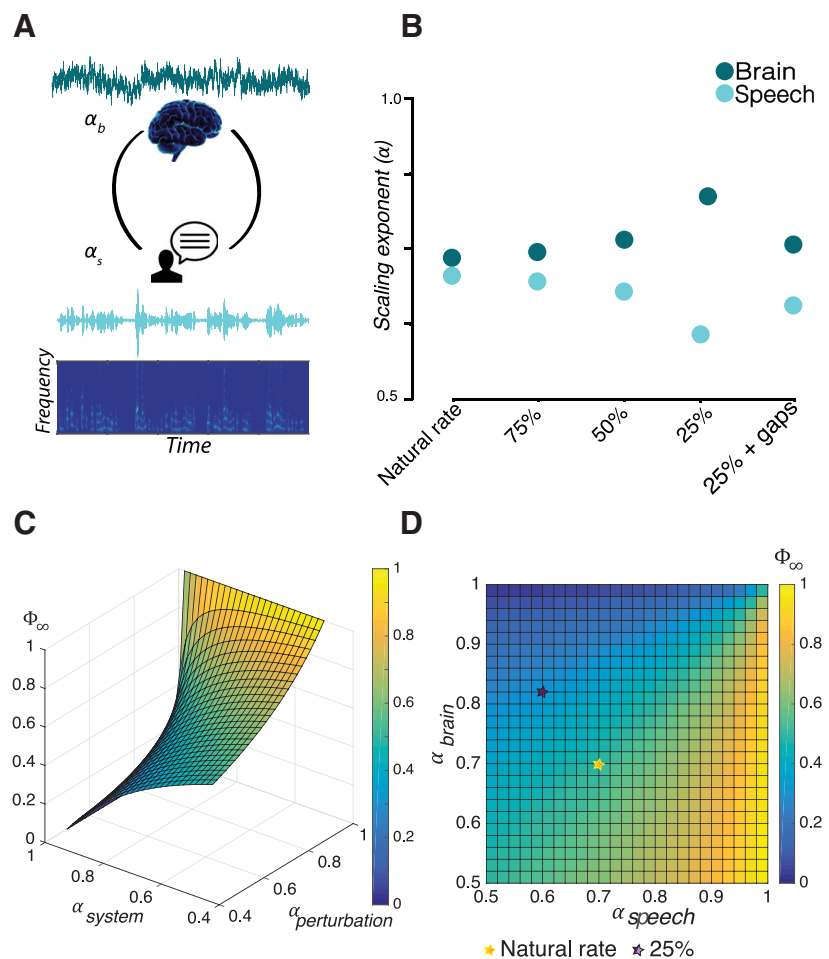
Figure 8. Hypothetical interplay between scale-free dynamics in speech and brain activity and comprehension. *A*, Speech and brain represent two complex systems with amplitude fluctuations characterized by their scaling exponents α_s and α_b, respectively. *B*, With increasing rate, α_s and α_b diverge, being maximally distinct at the fastest rate (25% compression), where comprehension was compromised. *C*, For values of scaling exponents in the range $0.5 < \alpha < 1$, a measure of information transfer between the perturbed system (brain) and a perturbation (speech stimuli) is given by the cross-correlation (Φ_∞), which depends on a relationship between the scaling exponents of speech ($\alpha_{perturbation}$) and brain (α_{system}) as follows: $\Phi_\infty = \dfrac{1 - \alpha_s}{2 - \alpha_s - \alpha_p}$ (Aquino et al., 2007, 2010, 2011). *D*, The cross-correlation is ~0.6 when subjects listen to the speech at the natural rate and minimal (~0.2) when the news pieces were time compressed by 25%.

sidering the overarching role of δ oscillations in large-scale cortical integration (Bruns and Eckhorn, 2004), our finding that δ/θ dynamics change selectively when comprehension deteriorates suggests that it reflects top-down processes (Kayser et al., 2015) and possibly mirrors an internal "synthesis" of the attended speech. In contrast, scaling of high-γ fluctuations varied, not just as a function of comprehension, but also at accelerated intelligible rates, suggesting an involvement of high-γ in the tracking of speech streams (Canolty et al., 2007; Nourski et al., 2009; Honey et al., 2012; Mesgarani and Chang, 2012; Zion Golumbic et al., 2013) and bottom-up processing (Zion Golumbic et al., 2013; Fontolan et al., 2014).

Our findings altogether agree with recent studies indicating that meaning retrieval affects almost the whole brain (Boly et al., 2015; Huth et al., 2016). Although temporal, frontal, and parietal regions encompass language-dedicated areas (Hickok and Poeppel, 2007; Fedorenko et al., 2011; Silbert et al., 2014), occipital and cerebellar regions are selectively involved when comprehension is challenging (Erb et al., 2013; Guediche et al., 2014). In regions of central and parietal cortices, we found that LRTC of

δ and θ correlated with comprehension. This is unsurprising because the inferior parietal region hosts an important hub for multimodal semantic processing (Binder and Desai, 2011) and phenomenal experience (Koch et al., 2016). Finally, whereas the prefrontal cortex controls semantic retrieval with increasing influence over time (Gardner et al., 2012), more transient systems may arise from posterior regions. Accordingly, we found noticeable LRTC increases in occipital/cerebellar regions with effortful comprehension. The change occurred almost linearly with the increase of speech rate mainly in the high-γ, suggesting that neural activity in these regions signals processes dependent both on comprehension and speed of speech. Our findings align with a cerebellar role in semantic integration and sensory tracking (Kotz and Schwartze, 2010; Buckner, 2013; Moberget and Ivry, 2016). Topographical inferences are duly conservative because some level of correlation between magnetic fields on the surface of the brain is warranted. Future studies using source reconstruction may enrich anatomical extrapolation and shed light on how spatially coordinated oscillatory activity subserves speech processing (Gross et al., 2013).

Language recovery in aphasic patients relies on new functional neuroanatomies involving nonlinguistic regions (Cahana-Amitay and Albert, 2014) and no single region is essential for language comprehension (Price and Friston, 2002). Therefore, it may be worthwhile exploring how scale-free dynamics, which our findings showed accounts for a substantial amount of variance in speech comprehension, evolves in, for example, poststroke aphasia.

Dynamical bottleneck in comprehension
We found that a fourfold increase in speech rate compromises comprehension. However, it did not cause irreversible information loss; otherwise, the repackaging of accelerated speech with silence gaps would not permit a nearly full recovery of comprehension. The bottleneck most likely resides in the trade-off between the quality of the speech fragment and the time elapsed to comprehend it (Ghitza, 2014; Ma et al., 2014).

A decrease in the speech LRTC may reflect decreasing information transfer rates (Aquino et al., 2010), which in turn might compromise comprehension. In addition, in the fastest speech condition, neural and acoustic scaling exponents increased and decreased, respectively, which is opposite of what would be expected for optimal information transfer according to the principle of complexity management (Aquino et al., 2007, 2010, 2011). Importantly, the individual relationship between LRTC and comprehension—subjects with lower neural LRTC had a better comprehension than subjects with higher LRTC—further supports this principle of how complex properties of stimuli and the brain interweave with information transfer.

To understand why this divergence of scaling behavior of speech and neuronal oscillations occurred, we consider the dynamical properties of the two systems. Regarding speech, recent evidence indicates that discrimination of sounds relies on the short- to long-range statistics of their envelopes (McDermott et al., 2013). Although it is uncertain why the LRTC decreased disproportionally at the fastest rate, an attenuation of LRTC with time compression appears intuitive because the downsampling mostly preserves the speech envelope while also miniaturizing it (the Matryoshka doll effect). Speech is a redundant signal (Attias and Schreiner, 1997) and the decrease in degeneracy by time compression may compromise its robustness by altering cues for word boundaries needed for speech parsing (Winter, 2014). Like the speech signal itself, functional cortical assemblies are also organized in a temporal hierarchy; fMRI and ECoG stud-

ies of narrative comprehension have shown that the auditory cortices preferentially process briefer stretches of information than higher-order areas (Lerner et al., 2011; Honey et al., 2012). The regions across this hierarchy act as low-pass filters causing the last regions to have slower dynamics because their inputs underwent more filtering stages (Baria et al., 2013; Stephens et al., 2013). We may thus conjecture that the observed bottleneck arises from a fuzzy acoustic-to-abstract reconstruction. When a less degenerate input propagates along the cortical hierarchy, the serial low-pass-filtering process will gradually obliterate its fast-varying temporal structure while retaining mostly slower-varying properties. Cortical dynamics are therefore characterized by larger scaling exponents, as observed here, signaling that slow fluctuations dominate (Peng et al., 1995), are less dependent on the recent past (Keshner, 1982), and yield perception of jibber-jabber sounds but lack the fine details necessary for meaning retrieval.

Overall, we show that speech comprehension relates to the multiscale structure of neuronal oscillations and speech signals, indicating that scale-free dynamics indexes time-based constraints underlying the bottleneck in processing accelerated speech. The results foster studies using multifractal or other complexity-related metrics (Stanley et al., 1999) that may refine the role of the acoustic rate on neural dynamics.

References

Abney DH, Paxton A, Dale R, Kello CT (2014) Complexity matching in dyadic conversation. J Exp Psychol Gen 143:2304–2315. CrossRef Medline

Ahissar E, Ahissar M (2005) Processing of the temporal envelope of speech. In: The auditory cortex: a synthesis of human and animal research (Konig R, Heil P, Budinger, E, Scheich H, Eds.) pp 295–313. Oxfordshire, UK: Taylor and Francis group.

Aquino G, Bologna M, Grigolini P, West BJ (2010) Beyond the death of linear response: 1/f optimal information transport. Phys Rev Lett 105: 040601. CrossRef Medline

Aquino G, Bologna M, West BJ, Grigolini P (2011) Transmission of information between complex systems: 1/f resonance. Phys Rev E Stat Nonlin Soft Matter Phys 83:051130. CrossRef Medline

Aquino, Grigolini, West (2007) Linear response and fluctuation-dissipation theorem for non-Poissonian renewal processes. EPL (Europhysics Letters) 80:10002. CrossRef

Attias H, Schreiner CE (1997) Temporal low-order statistics of natural sounds. In: Advances in neural information processing systems, Ed 9 (Mozer MC, Jordan MI, Petsche T, Eds), pp 27–33. Cambridge, MA: MIT.

Avnir D, Biham O, Lidar D, Malcai O (1997) Is the Geometry of Nature Fractal? Science 279:39–40.

Baria AT, Mansour A, Huang L, Baliki MN, Cecchi GA, Mesulam MM, Apkarian AV (2013) Linking human brain local activity fluctuations to structural and functional network architectures. Neuroimage 73:144–155. CrossRef Medline

Bashford JA Jr, Meyers MD, Brubaker BS, Warren RM (1988) Illusory continuity of interrupted speech: speech rate determines durational limits. J Acoust Soc Am 84:1635–1638. CrossRef Medline

Benjamini Y, Hochberg Y (1995) Controlling the false discovery rate: a practical and powerful approach to multiple testing. J R Statist Soc B 57:289–300.

Bianco S, Ignaccolo M, Rider MS, Ross MJ, Winsor P, Grigolini P (2007) Brain, music, and non-Poisson renewal processes. Phys Rev E Stat Nonlin Soft Matter Phys 75:061911. CrossRef Medline

Binder JR, Desai RH (2011) The neurobiology of semantic memory. Trends Cogn Sci 15:527–536. CrossRef Medline

Boersma P (2001) Praat, a system for doing phonetics by computer. Glot International 5:341–345.

Boly M, Sasai S, Gosseries O, Oizumi M, Casali A, Massimini M, Tononi G (2015) Stimulus set meaningfulness and neurophysiological differentiation: a functional magnetic resonance imaging study. PLoS One 10:e0125337. CrossRef Medline

Brainard DH (1997) The Psychophysics Toolbox. Spat Vis 10:433–436. CrossRef Medline

Bruns A, Eckhorn R (2004) Task-related coupling from high- to low-frequency signals among visual cortical areas in human subdural recordings. Int J Psychophysiol 51:97–116. Medline

Buckner RL (2013) The cerebellum and cognitive function: 25 years of insight from anatomy and neuroimaging. Neuron 80:807–815. CrossRef Medline

Buiatti M, Papo D, Baudonnière PM, van Vreeswijk C (2007) Feedback modulates the temporal scale-free dynamics of brain electrical activity in a hypothesis testing task. Neuroscience 146:1400–1412. CrossRef Medline

Cahana-Amitay D, Albert ML (2014) Brain and language: evidence for neural multifunctionality. Behav Neurol 2014:260381. CrossRef Medline

Canolty RT, Soltani M, Dalal SS, Edwards E, Dronkers NF, Nagarajan SS, Kirsch HE, Barbaro NM, Knight RT (2007) Spatiotemporal dynamics of word processing in the human brain. Front Neurosci 1:185–196. CrossRef Medline

Churchill NW, Spring R, Grady C, Cimprich B, Askren MK, Reuter-Lorenz PA, Jung MS, Peltier S, Strother SC, Berman MG (2016) The suppression of scale-free fMRI brain dynamics across three different sources of effort: aging, task novelty and task difficulty. Sci Rep 6:30895. CrossRef Medline

Ciuciu P, Abry P, Rabrait C, Wendt H (2008) Log Wavelet Leaders Cumulant Based Multifractal Analysis of EVI fMRI Time Series: Evidence of Scaling in Ongoing and Evoked Brain Activity. IEEE J Sel Topics Signal Process 2:929–943. CrossRef

Ciuciu P, Varoquaux G, Abry P, Sadaghiani S, Kleinschmidt A (2012) Scale-free and multifractal time dynamics of fMRI signals during rest and task. Fron Physiol 3:186. CrossRef Medline

Clayton K, Frey B (1997) Studies of mental "noise." Nonlinear Dynamics, Psychology, and Life Sciences 1:173–180.

Cohen J (1988) Statistical power analysis for the behavioral sciences, 2nd ed. Hillsdale, NJ: L. Erlbaum Associates.

Dupoux E, Green K (1997) Perceptual adjustment to highly compressed speech: effects of talker and rate changes. J Exp Psychol Hum Percept Perform 23:914–927. CrossRef Medline

Erb J, Henry MJ, Eisner F, Obleser J (2013) The brain dynamics of rapid perceptual adaptation to adverse listening conditions. J Neurosci 33:10688–10697. CrossRef Medline

Elliott TM, Theunissen FE (2009) The modulation transfer function for speech intelligibility. PLoS Comput Biol 5:e1000302.

Fedorenko E (2014) The role of domain-general cognitive control in language comprehension. Front Psychol 5:335. CrossRef Medline

Fedorenko E, Behr MK, Kanwisher N (2011) Functional specificity for high-level linguistic processing in the human brain. Proc National Acad Sci USA 108:16428–16433. CrossRef

Fontolan L, Morillon B, Liegeois-Chauvel C, Giraud AL (2014) The contribution of frequency-specific activity to hierarchical information processing in the human auditory cortex. Nat Commun 5:4694. CrossRef Medline

Freeman WJ, Holmes MD, Burke BC, Vanhatalo S (2003) Spatial spectra of scalp EEG and EMG from awake humans. Clin Neurophysiol 114:1053–1068. CrossRef Medline

Friedman M (1937) The use of ranks to avoid the assumption of normality implicit in the analysis of variance. Journal of the American Statistical Association 32:675–701. CrossRef

Gardner HE, Lambon Ralph MA, Dodds N, Jones T, Ehsan S, Jefferies E (2012) The differential contributions of pFC and temporo-parietal cortex to multimodal semantic control: exploring refractory effects in semantic aphasia. J Cogn Neurosci 24:778–793. CrossRef Medline

Gevins A, Smith ME, McEvoy L, Yu D (1997) High-resolution EEG mapping of cortical activation related to working memory: effects of task difficulty, type of processing, and practice. Cereb Cortex 7:374–385. CrossRef Medline

Ghitza O (2011) Linking speech perception and neurophysiology: speech decoding guided by cascaded oscillators locked to the input rhythm. Front Psychol 2:130. CrossRef Medline

Ghitza O (2014) Behavioral evidence for the role of cortical θ oscillations in determining auditory channel capacity for speech. Front Psychol 5:652. CrossRef Medline

Ghitza O, Greenberg S (2009) On the possible role of brain rhythms in speech perception: intelligibility of time-compressed speech with periodic and aperiodic insertions of silence. Phonetica 66:113–126. CrossRef Medline

Gilden DL, Thornton T, Mallon MW (1995) 1/f noise in human cognition. Science 267:1837–1839. CrossRef Medline

Giraud AL, Poeppel D (2012) Cortical oscillations and speech processing: emerging computational principles and operations. Nat Neurosci 15:511–517. CrossRef Medline

Giraud AL, Kleinschmidt A, Poeppel D, Lund TE, Frackowiak RS, Laufs H (2007) Endogenous cortical rhythms determine cerebral specialization for speech perception and production. Neuron 56:1127–1134. CrossRef Medline

Gramfort A, Luessi M, Larson E, Engemann DA, Strohmeier D, Brodbeck C, Goj R, Jas M, Brooks T, Parkkonen L, Hämäläinen M (2013) MEG and EEG data analysis with MNE-Python. Front Neurosci 7:267. CrossRef Medline

Greenberg S, Carvey H, Hitchcock L, Chang S (2003) Temporal properties of spontaneous speech–a syllable-centric perspective. J Phon 31:465–485. CrossRef

Groppe DM, Urbach TP, Kutas M (2011) Mass univariate analysis of event-related brain potentials/fields I: a critical tutorial review. Psychophysiology 48:1711–1725. CrossRef Medline

Gross J, Hoogenboom N, Thut G, Schyns P, Panzeri S, Belin P, Garrod S (2013) Speech rhythms and multiplexed oscillatory sensory coding in the human brain. PLoS Biol 11:e1001752. CrossRef Medline

Guediche S, Blumstein SE, Fiez JA, Holt LL (2014) Speech perception under adverse conditions: insights from behavioral, computational, and neuroscience research. Front Syst Neurosci 7:126. CrossRef Medline

Hardstone R, Poil SS, Schiavone G, Jansen R, Nikulin VV, Mansvelder HD, Linkenkaer-Hansen K (2012) Detrended fluctuation analysis: a scale-free view on neuronal oscillations. Front Physiol 3:450. CrossRef Medline

He BJ (2011) Scale-free properties of the functional magnetic resonance imaging signal during rest and task. J Neurosci 31:13786–13795. CrossRef Medline

He BJ, Zempel JM, Snyder AZ, Raichle ME (2010) The temporal structures and functional significance of scale-free brain activity. Neuron 66:353–369. CrossRef Medline

Henry MJ, Herrmann B, Obleser J (2014) Entrained neural oscillations in multiple frequency bands comodulate behavior. Proc Natl Acad Sci U S A 111:14935–14940. CrossRef Medline

Hickok G, Poeppel D (2007) The cortical organization of speech processing. Nat Rev Neurosci 8:393–402. CrossRef

Honey CJ, Thesen T, Donner TH, Silbert LJ, Carlson CE, Devinsky O, Doyle WK, Rubin N, Heeger DJ, Hasson U (2012) Slow cortical dynamics and the accumulation of information over long timescales. Neuron 76:423–434. CrossRef Medline

Howard MW, Rizzuto DS, Caplan JB, Madsen JR, Lisman J, Aschenbrenner-Scheibe R, Schulze-Bonhage A, Kahana MJ (2003) Gamma oscillations correlate with working memory load in humans. Cereb Cortex 13:1369–1374. CrossRef Medline

Huth AG, de Heer WA, Griffiths TL, Theunissen FEE, Gallant JL (2016) Natural speech reveals the semantic maps that tile human cerebral cortex. Nature 532:453–458. CrossRef Medline

Hyafil A, Giraud AL, Fontolan L, Gutkin B (2015a) Neural cross-frequency coupling: connecting architectures, mechanisms, and functions. Trends Neurosci 38:725–740. CrossRef Medline

Hyafil A, Fontolan L, Kabdebon C, Gutkin B, Giraud AL (2015b) Speech encoding by coupled cortical theta and gamma oscillations. eLife 4:e06213. Medline

Irrmischer M, Poil S, Mansvelder HD, Intra F, Linkenkaer-Hansen K (2017) Strong long-range temporal correlations of beta/gamma oscillations are associated with poor sustained visual attention performance. Eur J Neurosci. CrossRef Medline

Jensen O, Tesche CD (2002) Frontal theta activity in humans increases with memory load in a working memory task. Eur J Neurosci 15:1395–1399. CrossRef Medline

Kalashyan A, Buiatti M, Grigolini P (2009) Ergodicity breakdown and scaling from single sequences. Chaos, Solitons and Fractals 39:895–909. CrossRef

Kantelhardt J, Koscielny-Bunde E, Rego H, Havlin S, Bunde A (2001) Detecting long-range correlations with detrended fluctuation analysis. Physica A: Statistical Mechanics and its Applications 295:441454.

Kayser SJ, Ince RA, Gross J, Kayser C (2015) Irregular speech rate dissociates

auditory cortical entrainment, evoked responses, and frontal alpha. J Neurosci 35:14691–14701. CrossRef Medline

Keitel A, Ince RA, Gross J, Kayser C (2017) Auditory cortical delta-entrainment interacts with oscillatory power in multiple fronto-parietal networks. Neuroimage 147:32–42. CrossRef Medline

Kello C, Beltz B, Holden J, Orden G (2007) The emergent coordination of cognitive function. Journal of Experimental Psychology: General 136:551. CrossRef

Kello CT, Anderson GG, Holden JG, Van Orden GC (2008) The pervasiveness of 1/f scaling in speech reflects the metastable basis of cognition. Cogn Sci 32:1217–1231. CrossRef Medline

Keshner MS (1982) 1/f Noise. Proc IEEE 70:212–218. CrossRef

Koch C, Massimini M, Boly M, Tononi G (2016) Neural correlates of consciousness: progress and problems. Nat Rev Neurosci 17:307–321. CrossRef Medline

Kotz SA, Schwartze M (2010) Cortical speech processing unplugged: a timely subcortico-cortical framework. Trends Cogn Sci 14:392–399. CrossRef Medline

Lakatos P, Karmos G, Mehta AD, Ulbert I, Schroeder CE (2008) Entrainment of neuronal oscillations as a mechanism of attentional selection. Science 320:110–113. CrossRef Medline

Lam NHL, Schoffelen JM, Uddén J, Hultén A, Hagoort P (2016) Neural activity during sentence processing as reflected in theta, alpha, beta and gamma oscillations. Neuroimage 142:43–54. CrossRef Medline

Lerner Y, Honey CJ, Silbert LJ, Hasson U (2011) Topographic mapping of a hierarchy of temporal receptive windows using a narrated story. J Neurosci 31:2906–2915. CrossRef Medline

Lewis AG, Wang L, Bastiaansen M (2015) Fast oscillatory dynamics during language comprehension: unification versus maintenance and prediction? Brain Lang 148:51–63. CrossRef Medline

Lilliefors HW (1967) On the Kolmogorov-Smirnov test for normality with mean and variance unknown. J Am Stat Assoc. 62:399–402. CrossRef

Linkenkaer-Hansen K, Nikouline VV, Palva JM, Ilmoniemi RJ (2001) Long-range temporal correlations and scaling behavior in human brain oscillations. J Neurosci 21:1370–1377. Medline

Linkenkaer-Hansen K, Nikulin VV, Palva S, Ilmoniemi RJ, Palva JM (2004) Prestimulus oscillations enhance psychophysical performance in humans. J Neurosci 24:10186–10190. CrossRef Medline

Linkenkaer-Hansen K, Smit DJ, Barkil A, van Beijsterveldt TE, Brussaard AB, Boomsma DI, van Ooyen A, de Geus EJ (2007) Genetic contributions to long-range temporal correlations in ongoing oscillations. J Neurosci 27:13882–13889. CrossRef Medline

Luo H, Liu Z, Poeppel D (2010) Auditory cortex tracks both auditory and visual stimulus dynamics using low-frequency neuronal phase modulation. PLoS Biol 8:e1000445. CrossRef Medline

Luque, Luque, Lacasa (2015) Scaling and universality in the human voice. Journal of the Royal Society Interface 12:20141344–20141344. CrossRef

Ma WJ, Husain M, Bays PM (2014) Changing concepts of working memory. Nat Neurosci 17:347–356. CrossRef Medline

Mai G, Minett JW, Wang WS (2016) Delta, theta, beta, and gamma brain oscillations index levels of auditory sentence processing. Neuroimage 133:516–528. CrossRef Medline

McDermott JH, Schemitsch M, Simoncelli EP (2013) Summary statistics in auditory perception. Nat Neurosci 16:493–498. CrossRef

Mesgarani N, Chang EF (2012) Selective cortical representation of attended speaker in multi-talker speech perception. Nature 485:233–236. CrossRef Medline

Moberget T, Ivry RB (2016) Cerebellar contributions to motor control and language comprehension: searching for common computational principles. Ann N Y Acad Sci 1369:154–171. CrossRef Medline

Moulines E, Charpentier F (1990) Pitch-synchronous waveform processing techniques for text-to-speech synthesis using diphones. Speech Commun 9:453–467. CrossRef

Mukamel R, Nir Y, Harel M, Arieli A, Malach R, Fried I (2011) Invariance of firing rate and field potential dynamics to stimulus modulation rate in human auditory cortex. Hum Brain Mapp 32:1181–1193. CrossRef Medline

Nicol RM, Chapman SC, Vértes PE, Nathan PJ, Smith ML, Shtyrov Y, Bullmore ET (2012) Fast reconfiguration of high-frequency brain networks in response to surprising changes in auditory input. J Neurophysiol 107:1421–1430. CrossRef Medline

Nikulin VV, Brismar T (2005) Long-range temporal correlations in electro-

encephalographic oscillations: relation to topography, frequency band, age and gender. Neuroscience 130:549–558. CrossRef Medline

Nourski KV, Reale RA, Oya H, Kawasaki H, Kovach CK, Chen H, Howard MA 3rd, Brugge JF (2009) Temporal envelope of time-compressed speech represented in the human auditory cortex. J Neurosci 29:15564–15574. CrossRef Medline

Onton J, Delorme A, Makeig S (2005) Frontal midline EEG dynamics during working memory. Neuroimage 27:341–356. CrossRef Medline

Palva JM, Palva S (2011) Roles of multiscale brain activity fluctuations in shaping the variability and dynamics of psychophysical performance. Prog Brain Res 193:335–350. CrossRef Medline

Palva JM, Zhigalov A, Hirvonen J, Korhonen O, Linkenkaer-Hansen K, Palva S (2013) Neuronal long-range temporal correlations and avalanche dynamics are correlated with behavioral scaling laws. Proc Natl Acad Sci U S A 110:3585–3590. CrossRef Medline

Peelle JE, McMillan C, Moore P, Grossman M, Wingfield A (2004) Dissociable patterns of brain activity during comprehension of rapid and syntactically complex speech: evidence from fMRI. Brain Lang 91:315–325. CrossRef Medline

Peelle JE, Gross J, Davis MH (2013) Phase-locked responses to speech in human auditory cortex are enhanced during comprehension. Cereb Cortex 23:1378–1387. CrossRef Medline

Pefkou M, Arnal LH, Fontolan L, Giraud AL (2017) Theta- and beta-band neural activity reflect independent syllable tracking and comprehension of time-compressed speech. J Neurosci 37:7930–7938. CrossRef Medline

Pelli DG (1997) The VideoToolbox software for visual psychophysics: transforming numbers into movies. Spat Vis 10:437–442. CrossRef Medline

Peng K, Havlin S, Stanley E, Goldberger A (1995) Quantification of scaling exponents and crossover phenomena in nonstationary heartbeat time series. Chaos: An Interdisciplinary Journal of Nonlinear Science 5:82. CrossRef

Peng, Buldyrev, Havlin, Simons, Stanley, Goldberger (1994) Mosaic organization of DNA nucleotides. Phys Rev E 49:1685–1689.

Plenz D, Thiagarajan TC (2007) The organizing principles of neuronal avalanches: cell assemblies in the cortex? Trends Neurosci 30:101–110. CrossRef Medline

Poil SS, Hardstone R, Mansvelder HD, Linkenkaer-Hansen K (2012) Critical-state dynamics of avalanches and oscillations jointly emerge from balanced excitation/inhibition in neuronal networks. J Neurosci 32:9817–9823. CrossRef Medline

Poldrack RA, Temple E, Protopapas A, Nagarajan S, Tallal P, Merzenich M, Gabrieli JD (2001) Relations between the neural bases of dynamic auditory processing and phonological processing: evidence from fMRI. J Cogn Neurosci 13:687–697. CrossRef Medline

Popivanov D, Stomonyakov V, Minchev Z, Jivkova S, Dojnov P, Jivkov S, Christova E, Kosev S (2006) Multifractality of decomposed EEG during imaginary and real visual-motor tracking. Biological Cybernetics 94:149–156. CrossRef Medline

Poupard L, Sartène R, Wallet J-C (2001) Scaling behavior in β-wave amplitude modulation and its relationship to alertness. Biol Cybern 85:19–26. CrossRef

Price CJ, Friston KJ (2002) Degeneracy and cognitive anatomy. Trends Cogn Sci 6:416–421. CrossRef Medline

Pritchard WS (1992) The brain in fractal time: 1/f-like power spectrum scaling of the human electroencephalogram. Int J Neurosci 66:119–129. CrossRef Medline

Shannon RV, Zeng FG, Kamath V, Wygonski J, Ekelid M (1995) Speech recognition with primarily temporal cues. Science 270:303–304. CrossRef Medline

Silbert LJ, Honey CJ, Simony E, Poeppel D, Hasson U (2014) Coupled neural systems underlie the production and comprehension of naturalistic narrative speech. Proc National Acad Sci USA 111:E4687–E4696. CrossRef

Smit DJ, de Geus EJ, van de Nieuwenhuijzen ME, van Beijsterveldt CE, van Baal GC, Mansvelder HD, Boomsma DI, Linkenkaer-Hansen K (2011) Scale-free modulation of resting-state neuronal oscillations reflects prolonged brain maturation in humans. J Neurosci 31:13128–13136. CrossRef Medline

Smit DJ, Linkenkaer-Hansen K, de Geus EJ (2013) Long-range temporal correlations in resting-state α oscillations predict human timing-error dynamics. J Neurosci 33:11212–11220. CrossRef Medline

Spitzer M, Wildenhain J, Rappsilber J, Tyers M (2014) BoxPlotR: a web

tool for generation of box plots. Nat Methods 11:121–122. CrossRef Medline

Stanley HE, Amaral LA, Goldberger AL, Havlin S, Ivanov PCh, Peng CK (1999) Statistical physics and physiology: Monofractal and multifractal approaches. Physica A: Statistical Mechanics and its Applications 270: 309–324. CrossRef Medline

Stephens GJ, Honey CJ, Hasson U (2013) A place for time: the spatiotemporal structure of neural dynamics during natural audition. J Neurophysiol 110:2019–2026. CrossRef Medline

Taulu S, Simola J, Kajola M (2005) Applications of the Signal Space Separation Method. IEEE Trans Sign Process 53:3359–3372. CrossRef

Tesche CD, Uusitalo MA, Ilmoniemi RJ, Huotilainen M, Kajola M, Salonen O (1995) Signal-space projections of MEG data characterize both distributed and well-localized neuronal sources. Electroencephalogr Clin Neurophysiol 95:189–200. Medline

Uusitalo MA, Ilmoniemi RJ (1997) Signal-space projection method for separating MEG or EEG into components. Medical and Biological Engineering and Computing 35:135–140. CrossRef Medline

Vagharchakian L, Dehaene-Lambertz G, Pallier C, Dehaene S (2012) A temporal bottleneck in the language comprehension network. J Neurosci 32:9089–9102. CrossRef Medline

Van de Ville D, Britz J, Michel CM (2010) EEG microstate sequences in

healthy humans at rest reveal scale-free dynamics. Proc Natl Acad Sci U S A 107:18179–18184. CrossRef Medline

Voss R, Clarke J (1975) 1/f noise in music and speech. Nature 258:317–318. CrossRef

West BJ, Grigolini P (2011) The principle of complexity management. In: Decision Making: A psychophysics application of network science. pp 181–189. Toh Tuk Link, Singapore: World Scientific.

Wilcoxon F (1945) Individual comparisons by ranking methods. Biometrics Bulletin 1:80–83. CrossRef

Winter B (2014) Spoken language achieves robustness and evolvability by exploiting degeneracy and neutrality. BioEssays 36:960–967. CrossRef Medline

Zarjam P, Epps J, Chen F (2011) Characterizing working memory load using EEG delta activity. Proceedings of the 19th European Signal Processing Conference. pp 1554–1558.

Zilber N,Ciuciu P, Abry P.,van Wassenhove V (2013) Learning-induced modulation of scale-free properties of brain activity measured with MEG. Proc. of the 10th IEEE International Symposium on Biomedical Imaging pp 998–1001.

Zion Golumbic E, Ding N, Bickel S, Lakatos P, Schevon C, McKhann G, Goodman R, Emerson R, Mehta A, Simon J, Poeppel D, Schroeder C (2013) Mechanisms underlying selective neuronal tracking of attended speech at a "cocktail party." Neuron 77:980–991. CrossRef