Complexity patterns underlying speech production activity

Leonardo Lancia[1], Jinyu Li[1], Louis Goldstein[2]
[1] Laboratoire de Phonétique et Phonologie (CNRS/Sorbonne-Nouvelle).
[2] Department of Linguistics, University of Southern California.

Motivation. In a coordination dynamics account, the control of the many interacting processes underlying speech production relies on the balance between two fundamental but complementary tendencies of the sensorimotor system. A tendency toward particulation, underlying for example the emergence of gesture-specific functional modules, and a tendency toward integration, thanks to which coordination spreads across levels and time scales. Integrative tendencies permit for example the emergence of macroscopic rhythmic behaviors that regulate the functioning of the sensorimotor system and that reduce the degrees of freedom actively controlled by the speakers. However, the emerging rhythmic patterns need to be flexible enough in order for speakers to restructure at will their motor plans. The aim of our work is to uncover this flexibility by tracking changes in the entropic content of speech production activity, which is related to the complexity of the underlying dynamics. This work opens a new window on the role of the prosodic hierarchy in speech motor control. Indeed rhythmic patterns underlying speech production are shaped by the linguistic structure and as such vary consistently across languages [1, 2].
Methodology. Recently [3] could show that the average squared change of the MFCC coefficients extracted from the acoustic waveform (henceforth $\overline{\Delta^2 MFCC}$) generates an oscillatory pattern which is strongly correlated with the rate of change of articulator positions. Under the hypothesis that perturbations of articulator coordination result in perturbations of the $\overline{\Delta^2 MFCC}$ signal, the complexity of articulatory dynamics can be estimated by mapping the recurrent activity of the $\overline{\Delta^2 MFCC}$ signal onto that of the underlying processes. By using non-uniform embedding [4] we build a multidimensional representation of the observed data such that patterns of activity that are repeated over time should correspond to repeated patterns of activity of the underlying processes, even when these display multiple time scales. By using the variant of recurrence analysis proposed in [2] we detect repeated patterns of activity of non-stationary processes and represent them through recurrence plots (RPs). An RP is a two dimensional graph sparsely populated by black dots (recurrences) whose coordinates represent the time locations of equivalent states of a time series. From the analysis of RPs we obtain two different estimates of the entropy, one sensitive to spatial, or state entropy (*sENTR*), the other sensitive to temporal entropy (*tENTR*). *sENTR* is related to the richness of the dynamics underlying the $\overline{\Delta^2 MFCC}$ signals, disregarding the durations and the locations in time of the repeated patterns. The computation of *sENTR* is based on [5]. We first sum the exponentials of the negative Euclidean distances between each state of the multidimensional time series and its recurrent states, then we compute the entropy of the obtained sums. *tENTR* is related to the richness of the temporal deformations affecting the repeated patterns of activity. Inspired from [6], we first build a dictionary containing all possible patterns of white and black dots in 3×3 regions of the RP having a black dot in their center location, then *tENTR* is obtained by counting the occurrences of each dictionary entry in the RP.
Experiment In order to illustrate and test our approach we conducted a speech production experiment with delayed auditory feedback (DAF). 6 French speakers were asked to repeat several times three different sentences, each time in a different random order and without interruptions. During each trial, the value of auditory delay was randomly chosen among 0, 60, 90, 150 and 180 ms. The three sentences were mainly composed of voiced segments and differed with respect to the expected likelihood to generate speech errors. Sentence 1, from which we expected the smallest number of errors, was mainly composed of CV syllables with the same onset (/val'mõ vu'lɛ 'vwaʁ lə 'vɔl/). Sentence 2 (/lə nɔʁ'mɑ̃ ɔʁ nɔʁ'mə 'mɔʁ/) was composed by CV and CVC syllables and the last word was close to (mɔʁn) in which the consonants of the penultimate word are inverted. Sentence 3, from which we expected the highest number of errors (/ma'mɑ̃ e ma'mi ma'ni ma'mɑ̃/) contained several repetitions of the sequence mVmV and one repetition of the sequence mVnV. With increasing feedback delay, speakers increase the amount of speech rate slowing aimed at compensating for the delay between the motor commands and their acoustic consequences. We thus predict that that the

temporal complexity of the observed speech patterns increase with feedback delay. On the other hand, speech rate reduction permits more precise target achievement (modulo the occurrence of speech errors). We thus expect a decrease in the spatial complexity of the recurrent states with increasing feedback delay.

Results As expected, speech errors counts increase from sentence 1 to 3 and this difference is strengthened by feedback delay. Mixed effect regression models with maximal random effects structure reveal that, consistently with the literature on DAF, speakers lengthen their utterances to counteract the effect of the delay and they do that more during stressed vowels. Moreover, the same statistical approach shows that, as hypothesized, *sENTR* decreases with increasing feedback delay (less in sentence 2, more in sentence 3; see. Fig. 1, left panels) while *tENTR* increases (although significantly less in sentence 2; see Fig. 1, right panels). Consistently with sentence complexity, without feedback delay, *sENTR* increases from sentence 2 to 3. The tendency to decrease from sentence 1 to 2 displayed by *tENTR* in absence of feedback delay is not significant.

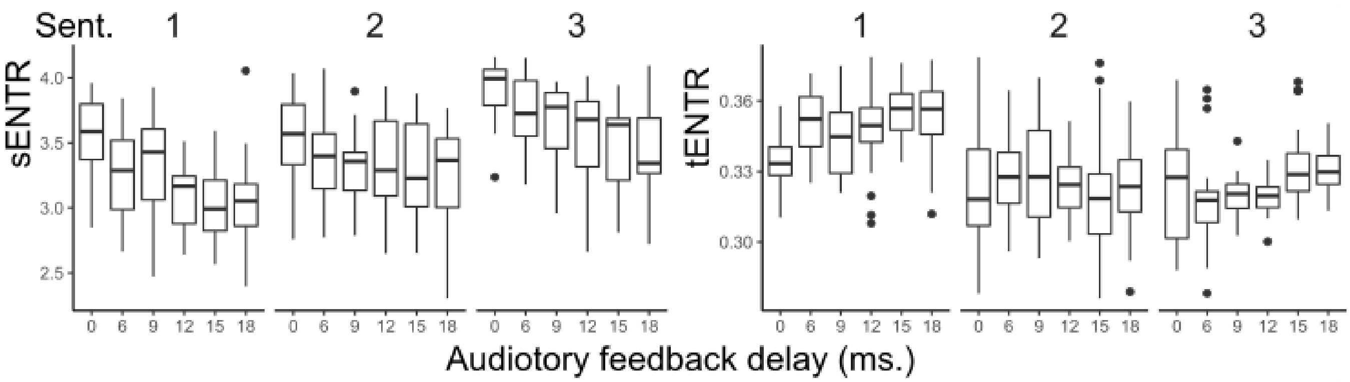

Figure 1: sENTR and tENTR over feedback delay. Data are grouped in panels by measure (leftmost three panels: sENTR , rightmost 3 panels: tENTR), and sentence (from 1 to 3 moving from left to right).

Discussion and conclusions: The observed results demonstrate that our approach is able to capture the entropic content of speech production in a speech task in which the sensory-motor system works in strongly perturbed conditions. At the same time the measures remain sensitive to small differences in the underlying complexity as those produced by small amounts of feedback delays. The approach does not depend on the segmentation of the speech signal and it is based on low level acoustic features, whose robustness guarantees applicability to a wide range of populations (e.g. newborns or pathological speakers) and speech conditions. The data analyzed reveal a complementary relation between the control of temporal features and that of spatial features of speech production. While spatial features seem to be more relevant in the achievement of segmental targets, the control of timing is constrained by prosody (temporal deformation occurs mainly in stressed syllables). Should this finding be generalizable to other tasks and conditions, it would have important implications for our understanding of the role of the prosodic structure in speech motor control.

References

[1] Varnet, L., Ortiz-Barajas, M. C., Erra, R. G., Gervain, J., & Lorenzi, C. (2017). A cross-linguistic study of speech modulation spectra. *JASA*, *142*(4), 1976-1989.

[2] Lancia, L., Krasovitsky, G., & Stuntebeck, F. (2019). Coordinative patterns underlying cross-linguistic rhythmic differences. *JPhon*, *72*, 66-80.

[3] Goldstein, L. (2019). The Role of Temporal Modulation in Sensorimotor Interaction. *Frontiers in Psychology*, *10*, 2608.

[4] Zhang, J. (2018). Low-dimensional approximation searching strategy for transfer entropy from non-uniform embedding. *PloS one*, *13*(3), e0194382.

[5] Eroglu, D., Peron, T. K. D., Marwan, N., Rodrigues, F. A., Costa, L. D. F., Sebek, M., ... & Kurths, J. (2014). Entropy of weighted recurrence plots. *Phys. Rev. E*, *90*(4), 042919.

[6] Corso, G., Prado, T. D. L., Lima, G. Z. D. S., Kurths, J., & Lopes, S. R. (2018). Quantifying entropy using recurrence matrix microstates. *Chaos*, *28*(8), 083108.