

LONG-RANGE SEQUENTIAL DEPENDENCIES PRECEDE SYNTACTICALLY-RICH VOCALIZATIONS IN HUMANS

TIM SAINBURG^{*1}, ANNA MAI¹, and TIMOTHY Q GENTNER¹

^{*}Corresponding Author: tsainbur@ucsd.edu

¹Department of Psychology, University of California San Diego, La Jolla, USA

1. Introduction

As the distances between elements in human language sequences increase, the mutual information (MI) between them decays following a power law. This power-law relationship has been variously attributed to human language syntax (Lin & Tegmark, 2017), semantics (Alvarez-Lacalle, Dorow, Eckmann, & Moses, 2006), and discourse structure (Altmann, Cristadoro, & Degli Esposti, 2012). However, the vocalizations of numerous phylogenetically distant species including humpback whales and songbirds also demonstrate similar long-range statistical dependencies (Sainburg, Theilman, Thielk, & Gentner, 2019). These observations support the hypothesis that the long-range statistical dependencies found in human speech can occur independently of linguistic structure. To test this hypothesis, we computed MI over several child speech corpora (aged 6 months – 12 years) to determine whether long-range relationships are present in human vocalizations prior to the production of syntactically-rich speech.

2. Methods

We computed MI over words and phonemes for 9 corpora of natural speech from English speaking children included in the PhonBank and CHILDES repositories. Although data collection protocols differed across corpora, all data included here were transcripts of spontaneous speech from typically-developing children aged 6 months to 12 years. Transcripts were binned into five 6-month age groups and one age group from 3–12 years. Each transcript was analyzed either as sequences of words or as sequences of phonemes. Sequence distributions were treated independently across speakers to account for individual variation in lexical acquisition.

For each corpus, we calculated sequential MI over the elements of the sequence dataset (i.e., words or phonemes), and fit the MI decay using the methods outlined in Sainburg et al. (2019). The three models fit were (1) a power-law model, (2) an exponential model, and (3) a composite of models (1) and (2).

$$MI = ax^b + c \quad (1)$$

$$MI = ae^{xb} + c \quad (2)$$

$$MI = ae^{xb} + cx^d + f \quad (3)$$

where x represents the inter-element distance between units (e.g. words or phonemes). Context-free languages exhibit power-law decays in MI while Markovian processes exhibit exponential decays (Lin & Tegmark, 2017). The composite model has previously been shown to capture the sequential organization of human speech (Sainburg et al., 2019) where the exponential decay captures organization of speech elements within words and at short distance, and longer-range dependencies are captured by the power-law decay.

3. Results

MI decay for phonemes and in all age groups is best fit by the composite model, which includes a significant power law component. The same is true for words, with the exception of the 6-12 month age group (Figure 1), where decay is best fit by an exponential model. At 12 months, infants produce one-word utterances (Doupe & Kuhl, 1999). After 12 months, we observed a power-law at long distances and an exponential decay at shorter distances, consistent with adult language results. Our observation of the power law decay in infant speech contradicts the notion that it is generated solely by either human language syntax, semantics, or discourse structure.

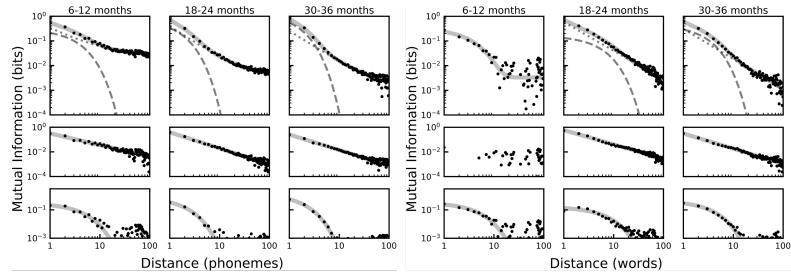


Figure 1.. MI decay over phonemes (left) and words (right) during development. Top row: Composite model fit; Middle row: Power law component fit; Bottom row: Exponential component fit.

4. Conclusion

Developmental emergence of long-range statistical dependencies in human speech precede the production of the linguistic structures hypothesized to generate them. This suggests that the long-range statistical organization of human speech is independent of human language syntax, semantics, or discourse structure. Instead, these dynamics may reflect a general property of biologically-generated sequences (Dawkins, 1976).

References

- Altmann, E. G., Cristadoro, G., & Degli Esposti, M. (2012). On the origin of long-range correlations in texts. *Proceedings of the National Academy of Sciences*, 109(29), 11582–11587.
- Alvarez-Lacalle, E., Dorow, B., Eckmann, J.-P., & Moses, E. (2006). Hierarchical structures induce long-range dynamical correlations in written texts. *Proceedings of the National Academy of Sciences*, 103(21), 7956–7961.
- Dawkins, R. (1976). Hierarchical organisation: A candidate principle for ethology.
- Doupe, A. J., & Kuhl, P. K. (1999). Birdsong and human speech: common themes and mechanisms. *Annual review of neuroscience*, 22(1), 567–631.
- Lin, H., & Tegmark, M. (2017). Critical behavior in physics and probabilistic formal languages. *Entropy*, 19(7), 299.
- Sainburg, T., Theilman, B., Thielk, M., & Gentner, T. Q. (2019). Parallels in the sequential organization of birdsong and human speech. *Nature communications*, 10.