EVOLUTIONARY PERSPECTIVES ON SPEECH AND INFORMATION RATES

CHRISTOPHE COUPE 1[*1,2], FRANÇOIS PELLEGRINO[2], YOONMI OH[3] and DAN DEDIU[2,4]

[*]Corresponding Author: ccoupe@hku.hk
[1]Department of Linguistics, The University of Hong Kong, Hong Kong, China
[2]Laboratoire Dynamique du Langage, CNRS/University of Lyon, Lyon, France
[3]Department of French Language and Literature, Ajou University, Suwon, South Korea
[4]Collegium de Lyon, University of Lyon, Lyon, France

1. Converging information rates in modern languages

Recent cross-linguistic studies have shown that while languages may differ quite significantly when it comes to speech rate, they are actually much closer in terms of information rate, i.e. the quantity of information they convey on average per second (Coupé et al., 2019). The explanation is a trade-off between speech rate and the average amount of information carried by linguistic units. Syllables have in particular been investigated in different languages and assessed with measures such as conditional entropy (Fenk-Oczlon & Fenk, 2010; Pellegrino, Coupé & Marsico, 2011; Coupé et al., 2019). Mandarin thus has a rather low speech rate, but information-dense syllables, while Spanish has a much faster speech rate, but much lighter syllables. Overall, all languages seem to fall into a narrow range of values for their information rate, centered around 39 bits per second.

The previous result is based on averaged unconstrained speech rates, and points toward universally shared capacities to encode, produce and decode speech. In every linguistic community, however, individual speakers display quite a range of variation (Jacewicz et al., 2010), and can additionally easily increase or decrease their usual speech rate in significant proportions. Unless it becomes extreme, a faster or a slower rate does not impact much the interlocutor's comprehension (Dupoux & Green, 1997; Adank & Janse, 2009). Universal functional and cognitive constraints, although they definitely exist, are therefore relatively weak. An underlying mechanism seems to rest on a coupling between the rhythm of cortical activity and the informational bandwidth of our communication system (Bosker & Ghitza, 2018; Hiafyl et al., 2105).

2. Information rates and speech rates in an evolutionary framework

The previous results can be framed in an evolutionary perspective, and offer insights about how language developed through time, in a way complementary to Villasenor et al. (2012). In this contribution, we distinguish the micro, meso and macro time scales (Wang, 1978) and assess our hypotheses with computational stochastic models of sociolinguistic networks (Nettle, 1999; Gong, Minett, & Wang, 2008), simulating the evolution of communities of speakers under constraints of i) linguistic convergence and ii) 'trade-off' between speech rate and information density.

At the micro scale of years or decades, we investigate how some actuated structural changes may be hindered, despite linguistic convergence, if they take idiolects away from the basin of attraction of information rates, e.g. drastic sound changes in some speakers which strongly increase or decrease the average information density.

At the meso scale of centuries or millennia, although the prehistory of modern languages remains mostly out of reach, language change is visible and occurs in speakers all equipped with modern cognitive capacities. It seems safe to assume information rates similar as today for a long period of time, likely at least since the emergence of our species. For this time period, we explore how minor variations in information density drive the evolution of speech rate, keeping a *fixed* optimal information rate, and how changes can spread from a few speakers to the entire community, if not during one's lifetime, across a few generations. Beyond internally-motivated changes, language contact and borrowing between structurally quite different languages can likely also impact speech rates, with the case of creoles requiring particular attention.

Finally, at a macro time scale, the evolution of the language function itself is characterized by a gradual increase in our ancestors' capacity to convey information, with the development of cognitive abilities and of speech physiological structures. An evolutionary scenario can be sketched, where natural selection is made possible by the aforementioned weak constraints and inter-individual variation, and the selective pressure is the sociocultural development pushing speakers to meet increasing communication needs. We investigate this scenario with a *variable* optimal information rate in our population of artificial speakers, and a general tendency for this rate to increase through time.

References

Adank, P., & Janse, E. (2009). Perceptual learning of time-compressed and natural fast speech. *Journal of the Acoustical Society of America, 126*, 2649-2659.

Bosker, H. R., & Ghitza, O. (2018). Entrained theta oscillations guide perception of subsequent speech: Behavioural evidence from rate normalisation. *Language, Cognition and Neuroscience, 33(8)*, 955-967.

Coupé, C., Oh, Y., Dediu, D., & Pellegrino, F. (2019). Different languages, similar encoding efficiency: comparable information rates across the human communicative niche. *Science Advances*, *5(9)*, eaaw2594.

Dupoux, E., & Green, K. (1997). Perceptual Adjustment to Highly Compressed and Rate Changes Speech: Effects of Talker and Rate Changes. *Journal of Experimental Psychology: Human Perception and Performance*, *23(3)*, 914-927.

Fenk-Oczlon, G., & Fenk, A. (2010). Measuring basic tempo across languages and some implications for speech rhythm. In *Proceedings of the 11th Annual Conference of the International Speech Communication Association, Interspeech 2010*, 1537-1540.

Gong, T., Minett, J. W., & Wang, W. S. Y. (2008). Exploring social structure effect on language evolution based on a computational model. *Connection Science*, *20(2–3)*, 135–153.

Hyafil, A., Fontolan, L., Kabdebon, C., Gutkin, B., & Giraud, A. L. (2015). Speech encoding by coupled cortical theta and gamma oscillations. *Elife*, *4*, e06213.

Jacewicz, E. Fox, R. A., O'Neill, C., & Salmons, J. (2009). Articulation rate across dialect, age, and gender. *Language Variation and Change*, *21(2)*, 233-256.

Jacewicz, E. Fox, R. A., & Wei, L. (2010). Between-speaker and within-speaker variation in speech tempo of American English. *J. Acoust. Soc. Am.*, *128(2)*, 839-580.

Nettle, D. (1999). Using Social Impact Theory to simulate language change. *Lingua*, *108(2–3)*, 95–117.

Pellegrino, F., Coupé, C., & Marsico, E. (2011). A cross-language perspective on speech information rate. *Language*, *87(3)*, 539-558.

Villasenor, J., Han, Y., Wen, D., Gonzales, E., Chen, J., & Wen, J. (2012). The information rate of modern speech and its implications for language evolution. In T. Scott-Phillips, M. Tamariz, E. A. Cartmill, and J. R. Hurford (Eds.), *Evolution of Language, The Proceedings Of The 9th International Conference (Evolang9)* (pp. 376-383). World Scientific.

Wang, W. S-Y. (1978). The Three Scales of Diachrony. In B. B. Kachru (ed.), *Linguistics in the Seventies: Directions and Prospects* (pp.63-76). Department of Linguistics, University of Illinois.