Transformers architectures are considered as one of the success stories of modern Natural Language Processing, given that their capacity of representing words in context has helped improving the state-of-the-art performance in several linguistic tasks (Devlin et al., 2019). But are they plausible models of human language processing? According to some of the classical assumptions about how language processing works, they should be not. However, in this talk, I will try to discuss this point in the light of some recent modeling results (Bommasani et al., 2020; Chronis and Erk, 2020; Merkx and Frank, 2021; Chersoni et al., 2021) and with reference to the two classical types of relations in the structuralist tradition in theoretical linguistics, i.e. syntagmatic and paradigmatic relations (De Saussure, 1916).

Abstract

Transformers architectures are considered as one of the success stories of modern Natural Language Processing, given that their capacity of representing words in context has helped improving the state-of-the-art performance in several linguistic tasks (Devlin et al., 2019). But are they plausible models of human language processing? According to some of the classical assumptions about how language processing works, they should be not. However, in this talk, I will try to discuss this point in the light of some recent modeling results (Bommasani et al., 2020; Chronis and Erk, 2020; Merkx and Frank, 2021; Chersoni et al., 2021) and with reference to the two classical types of relations in the structuralist tradition in theoretical linguistics, i.e. syntagmatic and paradigmatic relations (De Saussure, 1916).

Speaker

Emmanuele Chersoni obtained a PhD in Language Sciences from Aix-Marseille University (France) on the topic of a computational model for estimating the semantic complexity of natural language sentences. He is currently working as a Research Assistant Professor at the Department of Chinese and Bilingual Studies of The Hong Kong Polytechnic University. His current research interests include distributional semantic models, thematic fit modeling, automatic discovery of semantic relations and biomedical natural language processing.