Entropy in language comprehension

Tal Linzen
Ecole Normale Supérieure, Paris
tal.linzen@gmail.com

Predictability effects are pervasive in language comprehension. These effects support the view that language users maintain probability distributions over potential representations and use those distributions to allocate more resources to process likely upcoming representations. While it is fairly established that language comprehension is modulated by the probability of the incoming linguistic unit, it is unclear if it is also affected by the way in which probability mass is distributed among potential outcomes.

This talk will investigate whether this question can be addressed using the entropy of the distribution over representations. I will explore implications of existing proposals that have largely been overlooked, and suggest new linking hypotheses that use entropy to predict psycholinguistic measurements. These entropy-based linking hypotheses will be used to analyse empirical data drawn from two domains: reading times in sentence comprehension and electromagnetic recordings during spoken word recognition. I will show that entropy can be profitably used to predict responses in both areas, though not necessarily using the same linking hypotheses; surprisal cannot be subsumed under entropy in either case. In summary, entropy can be a useful component of linking hypotheses in studying language comprehension, but only when its properties and potential limitations are taken into consideration.