

Distinguere le parole? Gioco da bambini

Alla Sissa uno studio analizza (in nove) lingue come il cervello usa metodi statistici

Distinguere le singole parole nel flusso del parlato non è una cosa semplice e secondo i linguisti per farlo il cervello potrebbe usare dei metodi statistici. Un gruppo di scienziati Sissa ha applicato un metodo di segmentazione delle parole basato sulla statistica e ne ha misurato l'efficacia sul linguaggio naturale, in ben nove lingue diverse, scoprendo che il ritmo linguistico ha un ruolo importante. La ricerca è stata pubblicata sul *Journal of Developmental Science*.

Vi è mai capitato di rompervi il cervello cercando cogliere anche una sola parola nel flusso ininterrotto di un discorso in una lingua che conoscete a ma-

lapena? È ingenuo pensare che quando parliamo esista una anche minima pausa fra una parola e l'altra (come lo spazio che mettiamo per convenzione quando scriviamo): in realtà il parlato è quasi sempre un flusso sonoro continuo. Quando però ascoltiamo la nostra lingua madre questa "segmentazione" delle parole non ci costa quasi nessuno sforzo. Quali sono, si chiedono i linguisti, i meccanismi cognitivi automatici alla base di questa capacità? Un contributo certamente lo dà la conoscenza del vocabolario: la memoria del suono delle singole parole ci aiuta a rilevarle, ma, sostengono molti linguisti, esistono mec-

canismi automatici e inconsci di "basso livello" che ci aiutano anche quando non riconosciamo le parole, e anche, come nel caso di bambini molto piccoli, quando abbiamo una conoscenza ancora rudimentale della lingua. Questi meccanismi, credono gli scienziati, si basano sull'analisi statistica della frequenza (stimata in base all'esperienza pregressa) delle sillabe in ogni lingua.

Un indice che potrebbe contribuire ai processi di segmentazione è la "probabilità transizionale" (pt), che dà una stima della probabilità che due sillabe stiano all'interno della stessa parola, basandosi sulla frequenza con cui esse si trovano associa-

te in una data lingua. In pratica, se tutte le volte che sento la sillaba "ta" viene immancabilmente seguita dalla sillaba "da", allora la probabilità transizionale per "da", data "ta" è uno (il massimo). Se invece, ogni volta che sento la sillaba "bu" per metà delle volte capita che la segua la sillaba "di" e per metà delle volte la sillaba "fi", la probabilità transizionale di "di" (e di "fi") data "bu" è 0,5 e così via. Il sistema cognitivo potrebbe computare in maniera implicita questo valore sfruttando la memoria linguistica, dalla quale ricaverebbe le frequenze.

Lo studio condotto da Amanda Saksida, con la collaborazione di Alan Langus, sotto la gui-

da di Marina Nespor, ha usato l'indice pt per segmentare il linguaggio naturale, con due diversi approcci. Saksida ha lavorato con i corpus, ossia collezioni di testi raccolte appositamente per l'analisi linguistiche. In questo caso specifico si tratta di trascrizioni da registrazioni dell'"ambiente sonoro linguistico" a cui sono esposti bambini molto piccoli. «Volevamo avere un esempio del tipo di stimolo nel quale si sviluppa il linguaggio dei bambini», ha spiegato Saksida. «Ci chiedevamo se un meccanismo di basso livello come la probabilità transizionale funzionasse su stimoli linguistici realistici, molto diversi dagli stimoli che si usano normalmente in laboratorio, che sono più schematici e non contengono fonti di 'rumore'». Saksida e colleghi hanno usato corpus di ben 9 lingue diverse, e vi hanno applicato due diversi modelli basati sulla pt.