DOPPIOZERO

La voce: da Neanderthal ad Alexa

Mario Barenghi
8 Settembre 2020

Tradurre un titolo non è mai un'operazione banale. Trevor Cox, studioso di acustica, ha pubblicato nel 2019 *Now You're Talking: Human Conversation from the Neanderthals to Artificial Intelligence*, un agile e accattivante studio sulla storia dell'oralità. Una versione letterale del titolo avrebbe potuto essere qualcosa come «Questo sì che è parlare», «Questo si chiama parlare». Nel presentare il libro ai lettori italiani la casa editrice Dedalo ha optato per una dicitura diversa, e peraltro non impropria: *A ciascuno la sua voce. Come parliamo e ascoltiamo dai Neanderthal all'intelligenza artificiale* (pp. 288, € 17). Come si vede, la rinuncia a rendere l'espressione idiomatica e la scelta di puntare sul tema fondamentale della voce si ripercuote sul sottotitolo, chiamato a mettere in evidenza i due poli della comunicazione verbale, parola e ascolto. Ripresa invece dall'originale – e graficamente migliorata – è l'immagine di un ara dai colori sgargianti, che certo attira più di qualunque automa parlante.

Secondo Trevor Cox, la storia orale dell'umanità si può dividere in tre epoche. La prima è quella che precede l'avvento del linguaggio articolato, e riguarda i nostri progenitori, non sappiamo quanto remoti. La seconda comprende buona parte della preistoria e tutta la storia fino a un secolo e mezzo fa: un lungo arco di tempo in cui il linguaggio ha dispiegato tutta la propria potenza, accompagnando e rendendo possibile l'incredibile espansione e l'eccezionale evoluzione culturale della nostra specie. La terza, che simbolicamente si apre con l'invenzione del primo fonografo di Thomas Edison (1877), è caratterizzata dal massiccio intervento delle innovazioni tecnologiche. Potremmo essere tentati di chiamarla «era della riproducibilità tecnica», non fosse che alla possibilità di registrare voci e discorsi e ritrasmetterli in maniera sempre più accurata si è aggiunta abbastanza presto la possibilità di generare enunciati nuovi, con prospettive sempre più sorprendenti e – come quasi inevitabile – sempre più inquietanti.

I capitoli dedicati all'origine del linguaggio presentano un interessante mutamento di prospettiva rispetto alla maggior parte degli studi. Di solito l'attenzione si appunta sul versante della produzione: lo sviluppo del cervello da un lato, dell'apparato fonatorio dall'altro. Ma non meno indicativa è l'evoluzione dell'apparato uditivo: in particolare, lo sviluppo della sensibilità nell'intervallo di frequenze più importante per il linguaggio. Malgrado l'esiguità della documentazione fossile, è probabile che *Homo heidelbergensis*, comparso 700.000 anni fa (l'antenato che abbiamo in comune con i Neanderthal), avesse capacità uditive simili alle nostre – ferma restando l'avvertenza che «fu il linguaggio […] a trarre vantaggio dalle capacità uditive esistenti, e non il contrario». Per inciso, i dilettanti di paleontologia non potranno non entusiasmarsi all'idea che gli ossicini dell'orecchio siano derivati dalla mandibola di un rettile: un esempio da manuale di cooptazione funzionale, ossia di *exaptation*.

Certo, un nesso particolarmente stretto lega la voce all'identità. Noi siamo la nostra voce: da un lato riconosciamo con facilità le voci delle persone che ci sono più vicine, dall'altro moduliamo toni e modi delle parole a seconda delle circostanze e dell'immagine di noi che intendiamo trasmettere. L'impostazione vocale è un fatto squisitamente culturale: da alcuni studi pionieristici risulta che perfino il vagito dei neonati risente

delle percezioni acustiche prenatali. Ben attestato è invece, negli ultimi decenni, l’abbassamento del tono di voce femminile, un fenomeno da ricondurre alla conquista di una maggiore autonomia sociale (Trevor Cox non riporta confronti tra diversi Paesi, che potrebbero essere molto eloquenti).

Homo heidelbergensis, cranio.

Un registro grave è di norma associato a un’immagine non solo di forza, prestanza fisica, autorità, ma anche di onestà e competenza, come gli esperimenti confermano: dal che i politici maschi possono trarre facili conseguenze. Più complicata è la condizione delle donne, che devono invece trovare un equilibrio tra forza (toni bassi) e fascino (toni alti). A suo tempo, Margaret Thatcher si allenò per abbassare la propria voce.

Uno degli aspetti più interessanti degli studi sulla voce riguarda l'attaccamento agli stereotipi. Ovviamente non sorprende che il cervello vada in cerca di regolarità: gli stereotipi, così come i pregiudizi, garantiscono un risparmio cognitivo, e in un numero cospicuo di casi vengono confermati dall'esperienza. Nel caso degli stereotipi vocali, tuttavia, la frequenza degli errori è molto alta, di poco inferiore a quella che si avrebbe con scelte casuali. Ad esempio, secondo i riscontri sperimentali l'identificazione dell'omosessualità maschile sulla base del modo di parlare risulta corretta solo nel 60% dei casi. Ma, a quanto pare, i *biases* acustici sono tenaci.

La parte più suggestiva del volume è quella dedicata all'ultima fase storica, quella in corso. L'impatto della tecnologia sulla voce è stato enorme, ha cambiato il nostro modo di parlare e di cantare. L'uso dell'amplificazione ha consentito ai cantanti una possibilità di modulazione molto maggiore: ormai siamo largamente assuefatti a stili esecutivi estranei all'esigenza elementare, ineludibile fino a poco tempo fa, di farsi sentire anche dagli spettatori delle file più distanti. Anche nel teatro di prosa si è diffusa la prassi di amplificare le voci, benché in questo campo a volte sorgano dubbi (che personalmente trovo sacrosanti) e discussioni.

Eppure queste innovazioni sono poca cosa rispetto al futuro che si prepara, e che per certi aspetti è già cominciato. I computer stanno imparando ad ascoltare e a parlare. Le ricerche sull'intelligenza artificiale puntano decisamente sull'acquisizione di un uso attivo del linguaggio che potrebbe avere ripercussioni importanti. Sulla capacità di apprendere e di fare confronti si fonda la stessa facoltà di ragionare. Nel 2017 il laboratorio IA di Facebook ha annunciato di avere spento due chatbox (cioè due dispositivi dotati di software per la simulazione delle conversazioni) «perché avevano cominciato a parlare tra di loro abbandonando progressivamente l'inglese». Anche i risvolti emotivi sono sorprendenti. Se Siri, l'assistente vocale di Apple, è una presenza familiare a moltissimi utenti, a quanto pare centinaia di migliaia di persone dichiarano di essere innamorate di Alexa, l'assistente vocale di Amazon Echo («c'è persino chi le ha fatto una proposta di matrimonio»). A questo proposito ci si potrebbe chiedere perché vengano scelte di preferenza voci femminili. Forse perché si ritiene che siano soprattutto gli uomini a lasciarsi abbagliare da queste lusinghe virtuali? Per analogia, mi torna alla mente il comportamento di alcune amiche di Chichita Calvino, che, come lei stessa anni fa mi raccontava, le telefonavano apposta quando sapevano che era fuori casa per riascoltare il messaggio della sua segreteria telefonica, che le era stato registrato dall'amico Jean Sorel.

A parte gli aneddoti e i casi-limite, il problema è ovviamente la progressiva antropomorfizzazione dei dispositivi tecnologici. I fortissimi investimenti sull'Intelligenza Artificiale riguardano anche la simulazione della voce umana, inclusa la capacità di costruire frasi nuove e pertinenti sulla base dei modelli appresi. Turba – anzi, indigna – che sia divenuto ormai possibile utilizzare le registrazioni della voce di qualcuno, ad esempio di un defunto, per fargli dire cose che in vita non aveva mai detto né pensato. Ma inquieta ancora di più che una intelligenza artificiale dotata della capacità di apprendere possa assimilare, nel proprio *stock* di partenza (cioè nei dati sui quali viene addestrata), anche dei pregiudizi sociali. Cox suggerisce un piccolo esperimento con il traduttore di Google. «Provate a usarlo sulle frasi turche "o bir doktor" e "o bir hem?ire": otterrete "lui è un dottore" e "lei è un'infermiera". In turco, però, il pronome di terza persona "o" è neutro». Finché si tratta solo di traduzioni, il problema non è insormontabile. Ma quali competenze saranno attribuite, in futuro, a dispositivi dotati di intelligenza artificiale? Quante decisioni dipenderanno dai dati che sarà loro richiesto di raccogliere? E quante saranno chiamate a prendere direttamente? Quale controllo riusciremo a mantenere su intelligenze artificiali che si rivelino ora molto più rozze e subdolamente prevenute, ora molto più acute, spregiudicate e intraprendenti rispetto alle previsioni dei costruttori? Domande che non sono fatte per migliorare la qualità dei nostri sonni. In confronto, è quasi distensivo chiedersi perché mai non compaia nel frontespizio il nome di chi ha tradotto il libro di Trevor Cox. Ma, per la cronaca, si tratta di Andrea

Migliori, a cui vanno le nostre congratulazioni.

Trevor Cox, *A ciascuno la sua voce. Come parliamo e ascoltiamo dai Neanderthal all'intelligenza artificiale*, trad. di Andrea Migliori, Dedalo, pp. 288, € 17.

Se continuiamo a tenere vivo questo spazio è grazie a te. Anche un solo euro per noi significa molto.
Torna presto a leggerci e SOSTIENI DOPPIOZERO

Trevor Cox

A CIASCUNO LA SUA VOCE

Come parliamo e ascoltiamo dai Neanderthal all'intelligenza artificiale

edizioni Dedalo