

L'hate speech. Definizione e meccanismi di propagazione nei social network

di Luca Corchia*

Il saggio intende esaminare il fenomeno emergente dei discorsi d'odio (*hate speeches*). Da sempre diffusi nelle comunicazioni pubbliche e nelle interazioni quotidiane, da alcuni decenni sono oggetto di disciplinamento al fine di arginare una propagazione che rischia di minare la convivenza sociale. Dapprima, viene posta la questione controversa della loro definizione, superando i limiti del riduzionismo semantico grazie a un approccio pragmatico-linguistico. Quindi, sono specificate le caratteristiche distintive dei discorsi d'odio online e le trasformazioni strutturali della sfera pubblica che li rendono oggi così diffusi e pervasivi. La comunicazione politica, in particolare, sarà un campo di verifica delle tendenze in atto nella società nel suo insieme. Infine, si pone la questione di come prevenire e contrastare gli *hate speeches*, che qui affrontiamo dal punto di vista delle “missioni” delle scienze sociali.

Parole chiave: Hate speech; social network; comunicazione politica; sfera pubblica; sociologia; prevenzione.

Hate speech. Definition and propagation mechanisms in social networks

The paper aims to examine the emerging phenomenon of hate speech. Traditionally popular in public communications and everyday interactions, it has been the subject of regulation for several decades in order to limit a propagation that threatens to erode social coexistence. First, the controversial matter of defining hate speech is posed, overcoming semantic reductionism through a pragmatic-linguistic approach. Then, a specification is given regarding the distinctive characteristics of online hate speech and the structural transformations of the public sphere that makes it so widespread and pervasive today. In particular, political communication will be a field for testing the trends occurring in society as a whole. Finally, the question of how to prevent and counter hate speeches arises, which we address here from the perspective of the “missions” (tasks) of the social sciences.

Keywords: Hate speech; social networks; political communication; public sphere, sociology, prevention.

DOI 10.3280/SISS2023-003003

* Università di Chieti-Pescara. luca.corchia@unich.it.

Sicurezza e scienze sociali XI, 3/2023, ISSN 2283-8740, ISSNe 2283-7523

Copyright © FrancoAngeli.

E' vietata la Riproduzione dell'opera e la sua messa a disposizione di terzi,
sia in forma gratuita sia a pagamento.

Il documento può essere concesso in licenza individuale o istituzionale.

Introduzione

Il registro dell'odio è da sempre diffuso nelle comunicazioni pubbliche, dalla letteratura ai mass media e, soprattutto, nelle interazioni quotidiane. Può sembrare paradossale ma l'odio, così come la rabbia, il risentimento e altri sentimenti di disprezzo, ostilità e ripugnanza, è una delle tonalità emotiva del con-esserci, cioè una qualità della nostra disposizione verso gli altri, della relazione intersoggettiva e della stessa socialità. Al pari degli opposti sentimenti di stima, ammirazione, amicizia, amore, etc., quelli d'odio non solo esprimono un certo rapporto con il prossimo ma possono altresì generare un sostegno eguale da parte di altri ed essere generatori di comunità e subculture. La natura relazionale dell'*hate speech* risponde, infatti, alla funzione di confermare l'appartenenza a un gruppo e distinguerlo da altri, consolidando l'identità personale e sociale del soggetto. Il discorso d'odio è dispositivo retorico che marca sia comunanza che distinzione. Sul piano fenomenologico, si tratta di condotte e discorsi incitanti forme di discriminazione, pregiudizio, disprezzo, odio o violenza per lo più verso minoranze, gruppi o individui a partire da elementi stigmatizzanti che sanciscono un vero e proprio disconoscimento epistemico, morale ed esistenziale (Fricker, 2007). E come altri fenomeni che possiamo qualificare come "patologici" si acuisce sempre nei periodi di crisi economica, sociale e politica, quando più forti sono i sentimenti di insicurezza collettiva e maggiore è il disagio soggettivo.

Nel presente saggio intendiamo, dapprima, affrontare la questione di cosa qualifica precisamente l'*hate speech*. Il concetto, infatti, presenta una vaghezza per intenzionalità di significato e estensionalità di espressioni. La difficoltà a indicare con precisione quali discorsi siano espressioni dell'odio si riscontra nei documenti prodotti da organismi internazionali, comunitari e nazionali che ricorrono solitamente a definizioni onnicomprensive, lasciando ai dispositivi normativi specifici il compito di disciplinare le fattispecie concrete. In Italia, la "Commissione sull'intolleranza, la xenofobia, il razzismo e i fenomeni di odio", attiva dal maggio 2016 al luglio 2017 presso la Camera dei Deputati e presieduta da Laura Boldrini, attraverso audizioni di esperti e consultazione di dati e ricerche, ha approfondito la questione della diffusione di stereotipi, rappresentazioni false o fuorvianti, insulti, linguaggi ostili "normalizzati" o banalizzati e, ai livelli superiori, le discriminazioni e quindi il linguaggio d'odio e i crimini di odio (*hate crimes*). La stessa idea di formulare una "piramide dell'odio" però manifesta il problema più che risolverlo (1).

In secondo luogo, ci concentreremo sui discorsi d'odio online perché hanno specificità che li rendono particolarmente propagabili e pervasivi. Il fenomeno già diffuso sui siti internet è cresciuto esponenzialmente con i social network in cui si assiste alla loro replicabilità ed enfattizzazione. Per comprendere appieno perché le piattaforme online sono un terreno fertile per l'*hate speech* è necessario, tuttavia, considerare alcuni mutamenti in corso nel sistema mediale e come questi stiano ristrutturando l'espressione del sé, la socialità, l'autocomprensione normativa e la cultura del nostro tempo (2).

La comunicazione politica, in particolare, è il *frame* congeniale per i discorsi d'odio, con una "normalizzazione" delle espressioni di violenza verbale che dai circuiti istituzionali si propaga ai social network e viceversa in una spirale di legittimazione di una cultura del disprezzo, discriminatoria ed escludente. I dati del "Barometro dell'odio"¹, ad esempio, confermano la rilevanza degli *hate speeches* nelle campagne elettorali del nostro Paese (3).

Non sarà qui sviluppata la questione di come controllare, prevenire e contrastare i discorsi d'odio. Questa rimanda a strategie multilivello e strumenti giuridici (il diritto europeo e gli ordinamenti penali, civili e amministrativi dei Paesi), negoziali e di *moral suasion* (gli accordi con i mass media e le piattaforme online), deontologici (la *due diligence* giornalistica e le trasformazioni della stampa "di qualità"), formativi (l'aggiornamento dei docenti e i programmi educativi scolastici cui valori dell'inclusione e alla tolleranza, al rispetto reciproco e alla solidarietà tra diversi) e civici (il coinvolgimento delle associazioni della società civile) che richiedono un approfondimento specifico. Chiude il saggio una breve riflessione sulle "missioni" dei sociologi di fronte a un fenomeno che deve essere considerato come "patologico". Seguendo la tipologia formulata da Michael Burawoy sono, quindi, accennati i compiti professionali, di *policy*, di critica e pubblici con cui la sociologia non solo studia ma concorre a contrastare e prevenire gli *hate speeches*.

1. I discorsi d'odio: problemi definitori

¹ Il "barometro dell'odio" è uno strumento di rilevazione dei contenuti offensivi, discriminatori e di incitamento al disprezzo e all'odio presenti nella comunicazione politica. Dal 2018 al 2022, Amnesty International Italia ha realizzato sei report periodici. Per una descrizione delle finalità e metodologie si veda la pagina web: <https://www.amnesty.it/barometro-odio/>

Il problema di definire la specificità dei discorsi d'odio rispetto alle altre forme espressive che rientrano nel genere più ampio del *pólemos* è avvertito anzitutto sul piano giuridico. Raffaella Petrilli ha precisato che, sebbene l'*hate speech* produca «danni anche solo restando in forma di parola, danni individuali psicologici nelle vittime dell'odio e danni sociali [...] è difficile distinguere le molte, “normali” e legittime forme della polemica pubblica, più o meno emotivamente colorite e protette dal principio della libertà di parola, dalla comunicazione intollerante e discriminatoria» (2022: 120-121). Senza entrare nella disamina, da ultimo, il Comitato dei Ministri del Consiglio d'Europa ha definito il termine “discorso d'odio” come «qualsiasi tipo di espressione che incita, promuove, diffonde o giustifica violenza, odio o discriminazione contro un individuo o un gruppo di persone a causa di caratteristiche come la “razza”, il colore della pelle, la lingua, la religione, la nazionalità, l'origine nazionale o etnica, l'età, la disabilità, il sesso, l'identità di genere e l'orientamento sessuale» (*Recommendation CM/Rec(2022)16[1] (97)20*). Questa definizione ricalca quella formulata dall'Office on Genocide Prevention and the Responsibility to Protect delle Nazioni Unite in occasione del lancio della *United Nations Strategy and Plan of Action on Hate Speech*, il 18 giugno 2019. Rispetto al passato, si distinguono meglio i discorsi d'odio per contenuto e per gravità, a seconda che abbiamo una rilevanza penale, siano soggetti al diritto civile e amministrativo o siano espressioni non abbastanza gravi da essere limitate dalla Convenzione Europea dei Diritti Umani ma richiedano, tuttavia, l'adozione di risposte alternative non giuridiche.

Il limite di tali definizioni però è il riduzionismo semantico. Il discorso d'odio viene identificato unicamente per le espressioni verbali di disprezzo rivolte contro categorie di persone le cui “caratteristiche” sono valutate dalla comunità come “sensibili”. Si tratta di persone già vittime di discriminazioni o violenze e storicamente escluse dal discorso pubblico poiché ritenute indegne come interlocutori. Intendiamo, per contro, sostenere qui che l'atto verbale d'odio è riconoscibile per la sua tipica e distintiva “struttura linguistico-enunciativa” e che occorre specificare meglio le condizioni semiotiche di produzione. Spontanei per aggregazione di singole reazioni od organizzati grazie a una regia interessata, gli *hate speeches* rimandano a un'enunciazione “non-descrittiva” con diversi aspetti qualificanti: 1. l'espressione calunniosa intenzionalmente volta a produrre l'esclusione o l'emarginazione dal gruppo; 2. il riferimento esplicito al *parlante-hater*, ossia l'aggressore (io, noi); 3. il riferimento esplicito al *target* da escludere o emarginare dal gruppo (lei/lui, loro); 4. il riferimento esplicito

all'*interlocutore* incitato a condividere l'odio, l'esclusione o l'emarginazione del *target* (tu, voi, noi) (Petrilli, 2022: 125).

Un approccio pragmatico ai modi d'uso del linguaggio – che focalizza l'attenzione sulla funzione metacomunicativa dei verbi performativi – aiuta a identificare il discorso d'odio come atto linguistico che non avanza pretese di validità, anzitutto quella veritativa delle proposizioni constative: «il discorso di incitamento all'odio esprime ma non descrive. Esprime il punto di vista del parlante, evitando però accuratamente di metterlo in dubbio al confronto con i dati della realtà» (ivi: 127-128). Come nelle conversazioni quotidiane, l'*hate speech* presenta una miscela di elementi referenziali, normativi ed espressivi che non consente di tematizzare con precisione le asserzioni sugli stati di fatto nel mondo oggettivo, ma neppure quelle sulla giustezza delle norme di condotta o dei principi morali che regolano il mondo sociale, quelle sull'adeguatezza dei valori etici con cui una particolare comunità sociale in un periodo storico discute e decide ciò che è “bene” e, infine, quelle sulla stessa sincerità delle espressioni relative al mondo interiore – desideri, sentimenti, stati d'animo, etc. – dei soggetti coinvolti nella situazione. Come sottolinea sempre Petrilli, «la forza discriminatoria e intollerante dell'incitamento all'odio introduce nella sfera pubblica fratture e conflitti non argomentabili» (ivi: 120). Prima ancora di essere intenzionale, nel discorso d'odio il rifiuto dell'argomentazione deriva dal modo comunicativo.

Precisata la forma pragmatico-linguistica, vediamo gli elementi costitutivi di un *hate speech*. Il primo è il contenuto e lo stile del messaggio d'odio, volto a produrre l'esclusione o l'emarginazione dell'“obiettivo” dal consesso delle persone rispettabili verso cui avere legami solidali. Sul contenuto si è già detto. Il discorso d'odio è una manifestazione di disprezzo e ostilità verso il *target* stigmatizzato sulla base di caratteristiche giudicate negativamente. Esso incita gli altri componenti a manifestare altrettanta avversione e condotte discriminatorie volte ad escludere il *target*, nel senso ampio di far tacere, isolare, sparire, aggredire, morire, etc. Lo stile comunicativo dipende da una serie di aspetti relativi a tutte le dimensioni del linguaggio verbale e non verbale con cui l'*hater* si rivolge alle vittime e ai suoi sodali. Anche i registri linguistici possono variare molto, dal sarcasmo ai toni aggressivi o a quelli apocalittici tesi a suscitare sensazioni di scandalo, allarmismo, urgenza e bisogno di azione. Il discorso d'odio può anche camuffarsi con codici criptati per sfuggire alla censura. Al di là della forma, ciò che gli è specifico rispetto al discorso polemico, pur volgare e offensivo, è l'intenzione di annichilire il bersaglio in quanto possibile interlocutore. Come ha chiarito Petrilli, nell'*hate speech*, «l'odiatore (*hater*) vuole

eliminare simbolicamente l'altro dalla sfera pubblica, riservandogli la posizione del "muto"» (*ibidem*). Il secondo elemento è il riferimento esplicito al *parlante-hater* – "io" o "noi" e concerne, quindi, l'odiatore, il cui prestigio o potere influenza non poco la "legittimità" e gli effetti del messaggio sui pubblici e sui target degli *speeches*, così come la visibilità e propagazione del discorso d'odio nei circuiti dei social media e il "rimbalzo", tramite processi di re-intermediazione, nel grande spazio dei mass media. Il *target* da escludere o emarginare dal gruppo (lei/lui, loro) è il terzo elemento del discorso d'odio. L'*hate speech* è un discorso istigatorio che "fa odiare" persone connotate da certe caratteristiche vere o presunte. Si riscontra una miscela di variabilità e permanenza sul piano storico-culturale. Le diffamazioni etniche e religiose hanno certamente attirato la maggior parte dell'attenzione in passato, ma sono diffusi anche i discorsi d'odio rivolti all'orientamento sessuale e al genere che generano oggi allarmi crescenti da parte dell'opinione pubblica. I bersagli dei discorsi d'odio però possono essere identificati in contesti del tutto eccezionali. Basti pensare nel periodo di *lockdown* per il Coronavirus ai cinesi, ai lombardi, ai proprietari di seconde case, ai *runners*, disprezzati e accattati come degli "untori". Ogni lista non può considerarsi esaustiva perché "*other forms*" sono sempre possibili. Per tutti questi fenomeni vale il presupposto oggettivo in stigmatizzazioni diverse dell'esistenza di un Altro-da-noi verso il quale l'odio e l'incitamento all'odio sono diretti sino a identificarlo come un "nemico" che rappresenta una minaccia esistenziale per "noi". Il quarto elemento è il destinatario vero e proprio del discorso d'odio, ossia l'*interlocutore* (tu, voi, noi) incitato dall'*hater* a condividere l'esclusione o l'emarginazione del *target*. Non si parla alle vittime, bensì alle persone che la pensano allo stesso modo (Brown, 2018: 302). E la propagazione dipende dall'implicita comprensione tra l'*hater* e gli interlocutori a partire dal comune sfondo culturale pre-riflessivo che rende accettabili i sentimenti e le condotte d'odio.

Se consideriamo da ultimo gli effetti, la capacità di influenza dell'*hater* e la dimensione del pubblico a cui è rivolto il messaggio, così come la sua minore o maggiore ricettività, compattezza, fedeltà, emotività, mobilitazione, etc. emergono variabili che incidono significativamente sulla pericolosità del messaggio. E come precisano Sara Bentivegna e Rossella Rega, vanno presi in esame anche gli elementi contestuali e relazionali: «il potere di influenza di una comunicazione dipende dall'insieme delle dimensioni implicate nel processo comunicativo; non soltanto il messaggio in sé, ma anche le caratteristiche del comunicatore, il contesto in cui rilascia il discorso, il pubblico a cui si rivolge, le sue aspettative e la sua

suscettibilità, valutando contemporaneamente le dinamiche di azione e retroazione tra tutti questi elementi contemporaneamente» (2020: 166). Sarebbe un errore voler contrastare l'*hate speech* senza esaminare il *medium* che lo veicola. Tale avvertenza introduce la questione della comunicazione su Internet e nei social network.

2. La sfera pubblica digitale e l'online hate speech

A grandi linee, l'avvento di Internet e dei social media rappresenta una vera e propria rivoluzione nella storia del genere umano, paragonabile a quella della scrittura e della stampa, e tuttavia con una portata superiore. Considerando solo gli aspetti comunicativi, la digitalizzazione e l'interconnessione, infatti, hanno prodotto un cambiamento radicale con il rapido passaggio dalle nozioni di canale e rete a quella di spazio virtuale inglobante. I nuovi dispositivi "da tutti a tutti", infatti, superano il modello mediale delle trasmissioni astratte, verticali e unilaterali da pochi centri al pubblico indistinto. Una massa di persone può accedere a una quantità massiva di informazioni da loro selezionate e, attivamente, può proporre temi e contenuti, liberamente scelti, potendo assumere spontaneamente, assieme ad altri, i ruoli di lettore, ascoltatore e spettatore, ma anche quelli di autore e critico.

L'abbondanza, tuttavia, provoca una forza centrifuga che tende a strutturare le interazioni in ecosistemi comunicativi reticolari, decentrati e frammentati. La stampa, la radio e la televisione focalizzavano l'attenzione del pubblico su questioni discusse tra gli attori principali (giornalisti, politici, intellettuali, rappresentanti di interessi corporativi, esponenti della società civile, etc.) e, quindi, ridiscusse nella vita quotidiana con letture e prese di posizione diversificate. Con le funzioni di "filtro", "cassa di risonanza", "sistema di allerta", *agenda setting* e *framing* svolte dai media, questo circuito assicurava la formazione di opinioni pubbliche, il mantenimento di legami solidaristici e l'attivazione di processi di apprendimento collettivi (Habermas, 2022, trad. it. 2023). Tale integrazione dall'alto è oggi problematica².

Le preoccupazioni maggiori sono provocate dalla creazione di *echo chambers* regolate da principi di omofilia e discriminazione dei diversi. Si tratta di vere e proprie *filter bubble* in cui gli individui tendono a confinarsi

² Per una rassegna dei mutamenti della sfera pubblica si veda il volume curato da Laura Gherardi (2022), con saggi di Michele Sorice, Carlo Sorrentino, Giovanni Boccia Artieri, Massimo Cerulo e Aurelia Zucaro, Chiara Moroni, Gabriele Giacomini, Olimpia Affuso e Francesco Raniolo, Raffaella Pettrilli, Ilenia Colonna, Fabrizio Tonello, Benedetta Baldi, Luca Corchia, Laura Gherardi e Stefano Cristante.

all'interno di *comfort zone* culturalmente chiuse, entro cui risuonano le stesse parole, gli stessi concetti e le medesime convinzioni (*confirmation bias*), che convalidano e radicalizzano i punti di vista pregiudiziali. Sulle riflessioni e le argomentazioni prevalgono le espressioni di sostegno e attacco in un circolo vizioso che incoraggia la polarizzazione tra fazioni avverse. Spesso, in queste arene “post-discorsive” circolano informazioni alternative rispetto a quelle *mainstream* e che sono sottratte alla falsificazione. L'aspetto che allarma è la distorsione della percezione della sfera pubblica, per cui dal punto di vista di tali gruppi le *fake news* non possono più essere identificate come tali. L'informazione delle istituzioni pubbliche e dei mass media è investita da un senso di sfiducia e, soprattutto in situazioni di crisi ed emergenza, vere e proprie campagne d'odio colpiscono le persone più in vista, come i giornalisti, i ricercatori e i leader di ordini professionali (Scamuzzi *et al.*, 2021: 67).

In quest'ultimo decennio, gli studiosi di molti settori disciplinari hanno esaminato se, quanto e come Internet e i social media costituiscano un terreno favorevole alla propagazione dei discorsi d'odio³. Sono emersi alcuni aspetti ricorrenti: la pervasività, la permanenza e il possibile ritorno, l'anonimato, l'istantaneità e la transnazionalità (Ziccardi, 2016: 78-82). L'*hate speech*, infatti, tende normalmente a essere visualizzato di più e a rimanere virale per molto tempo, potendo ricomparire con nome e/o titolo sulla stessa piattaforma online o altrove, nel caso in cui sia stato rimosso (Gagliardone *et al.*, 2015: 14-15; Tipaldo, 2019: 24). La natura delle relazioni online, certo, facilita la dinamica itinerante della diffusione. La percezione dell'anonimato, la distanza, l'immediatezza e, non ultimo, il senso di irresponsabilità per le conseguenze delle proprie condotte a causa della de-individualizzazione, contribuiscono alla proliferazione di discorsi d'odio. Ricerche *mixed method* hanno mostrato le differenti percezioni tra le azioni online e quelle offline (Hassan *et alii*, 2022). Su Internet, il tempo che intercorre tra l'aver un pensiero o un sentimento e l'esprimerlo ad un particolare individuo che si trova a una certa distanza, o a un gruppo di persone di mentalità affine o ad una platea di ascoltatori, può essere una questione di istanti. L'istantaneità dei contatti in rete contribuisce al proliferare di reazioni “di pancia”, pensieri e giudizi non ponderati, osservazioni spontanee e commenti non filtrati (Brown, 2018: 304). Il “sensazionalismo”, inoltre, è un meccanismo per-verso legato alle logiche di visibilità (Bortone e Cerquozzi, 2017: 24). Le comunicazioni sui social media aumentano la velocità e la portata

³ Sergio A. Castaño-Pulgarín e altri ricercatori (2019) hanno pubblicato un'utile rassegna sistematica di 67 su 2389 articoli relativi agli *hate speeches online*, riscontrando ambiti di interesse ricorrenti e rilevando i limiti metodologici e contestuali delle analisi comparative.

dell'aggregazione attorno ai messaggi orientati da sentimenti di antipatia, disprezzo e odio e sono moltiplicatori di stereotipi e stigmatizzazioni verso i "nemici". Bentivegna e Rega avvertono che, trattando gli elementi dell'*hate speech* – l'odiatore, il messaggio e il ricevente – occorre considerare il contesto dei media digitali e le proprietà e l'*affordance* delle piattaforme veicolari, cioè «il modo in cui essi transitano tra i diversi media e piattaforme, come vengono trasformati e ri-distribuiti da innumerevoli soggetti tra loro diversi (cittadini, attori politici, media, etc.), gli effetti a catena che possono innescare, impossibili da delimitare nel tempo, nello spazio come pure nel numero dei soggetti su cui potrà produrre delle ricadute» (2020: 151).

3. Portata e rilevazioni sul lessico politico

La politica, ovviamente, è un tema che infiamma gli animi più di altri. Va considerato che il discorso politico è il *frame* congeniale per il discorso d'odio proprio perché largamente protetto dal diritto di espressione del libero pensiero. Questa tutela è particolarmente estensiva negli Stati Uniti d'America dove la libertà di parola è sancita dalla Costituzione ed è un valore fondamentale per l'identità nazionale. Benussi sottolinea come in virtù della giurisprudenza del Primo Emendamento la protezione praticata negli USA sia unica al mondo ed oggetto di grande interesse teorico (2019: 40). Il diritto di manifestare liberamente il proprio pensiero con la parola, lo scritto e ogni altro mezzo di diffusione è tuttavia un diritto riconosciuto a tutti i cittadini negli ordinamenti giuridici di matrice liberale. Il confine tra lecito e illecito è quindi sottile. La normalizzazione dei discorsi d'odio nel lessico politico è comunque un fatto ineludibile e riguarda oramai le stesse deliberazioni nelle assemblee di rappresentanza della sovranità popolare. Una ricerca condotta da Marinella Belluati (2018) sugli *hate speeches* nei dibattiti parlamentari è illuminante. Selezionando temi potenzialmente ad alto contenuto di odio, la sociologa torinese ha verificato che i discorsi istituzionali non sono immuni da offese, che i linguaggi offensivi sono parte integrante di una particolare cultura politica e che la scelta dei bersagli dei discorsi d'odio è condizionata da disposizioni discriminatorie verso attori che fungono da capri espiatori. I discorsi di incitamento all'odio sono generati e si riversano per lo più nella sfera pubblica digitale. I social network, infatti, sono un ambiente comunicativo caratterizzato da inasprimenti sistematici del linguaggio politico, ben al di là di normali critiche, polemiche, sarcasmi e ironie verso gli avversari. Un'ampia parte della popolazione, anche in Italia – come confermano le rilevazioni del "Barometro dell'odio" – sembra essersi rassegnata alla

violenza verbale nei social network considerandola un nuovo modo di comunicare ai tempi di internet. Un salto di livello nella legittimazione simbolica di tali atti si verifica quando l'*hate speech* sono più o meno direttamente creati, sostenuti o rilanciati da figure con cariche pubbliche o personaggi influenti nella società. Si potrebbe richiamare il concetto di “spirale del silenzio”, formulato da Elisabeth Noelle-Neumann, per descrivere l'acquiescenza verso una violenza verbale socialmente non sanzionata. Gli utenti, infatti, “sentono nell'aria”, per così dire, che il disprezzo è legittimato. In tali casi, è più probabile l'effetto di “imitazione” da parte delle persone esposte a replicare atteggiamenti ostili verso i *target* e reiterare i messaggi d'incitamento all'odio.

Il “Barometro dell'odio” realizzato da Amnesty International Sezione Italia sulla campagna per le elezioni parlamentari del 2022, ad esempio, conferma la presenza di discorsi d'odio veicolati dai social network (qui solo Facebook e Twitter) da parte dei candidati ai collegi uninominali e dei capolista ai seggi plurinominali. Nelle cinque settimane di monitoraggio (tra agosto e settembre) sono stati esaminati oltre 30.000 messaggi, rilevando che circa 8% ha contenuti offensivi e discriminatori mentre l'1% è definibile come *hate speech*, cioè un messaggio di incitamento all'odio e alla discriminazione basato sulle caratteristiche personali della persona o del gruppo di persone prese di mira. I messaggi segnalati hanno riguardato tutti gli schieramenti elettorali, con una prevalenza del centro-destra (9%) e a scalare Azione-Italia Viva (6%), il Centro-sinistra (4%), e il Movimento 5 stelle (3%). Svettano nella classifica dei discorsi che incitano all'odio i candidati della Lega. Riguardo ai contenuti, più della metà dei messaggi, il 53%, ha avuto per bersaglio i migranti e gli immigrati, ma sono stati rilevati discorsi d'odio contro le minoranze religiose, per lo più di impronta islamofobica e antiggiudaica (36%), attacchi alla comunità rom (35%), alle associazioni solidaristiche (35%), al mondo Lgbtqia+ (31%), e non sono mancate le violenze verbali sessiste contro le donne (26%). Sono confermate le impressioni espresse nel “Barometro dell'odio” 2018 dal direttore di Amnesty International: «Alcune forze politiche si sono servite di stereotipi e incitazioni all'odio per fare propri diffusi sentimenti populistici, identitari e xenofobi, promuovendo la diffusione di un linguaggio incendiario, divisivo, che discrimina anziché promuovere l'eguaglianza, che pensa che minoranze e gruppi vulnerabili siano una minaccia e che i diritti non spettino a tutti» (2018: 3).

In presenza di organizzazioni politiche o sociali che perseguono fini discriminatori con il lessico dell'odio, soprattutto se queste sono emanazione di apparati statali o a loro collegati o ancora da loro favoriti, a queste condizioni si può parlare di vere e proprie “politiche dell'odio” che non si limitano

a episodi isolati o accidentali ma si collocano in strategie propagandistiche. Per i gruppi dominanti al governo di un Paese, il controllo dei media, la strumentalizzazione della cultura e delle politiche educative, i meccanismi di censura associati a forme di indottrinamento sono rinforzati spesso da retoriche che convertono la “paura dell’altro” in “odio dell’altro” (Fotia, 2022: 15).

Conclusioni

È possibile prevenire e contrastare i discorsi d’odio? In quasi tutti i Paesi europei sono in vigore norme giuridiche che qualificano come reati forme diverse di incitamento o istigazione alla violenza e alla discriminazione sulla base di molteplici fattori quali la razza/etnia, la religione, la nazionalità e, più recentemente, l’orientamento sessuale, l’identità di genere, la disabilità, etc. Il quadro normativo però è per lo più incerto quando si tratta dei discorsi d’odio. Nella *Raccomandazione CM/Rec(2022)16[1]*, il Comitato dei Ministri del Consiglio d’Europa ha stabilito alcuni principi, lasciando agli Stati membri il compito di decidere le misure attuative nei loro ordinamenti giuridici⁴. Per prevenire e contrastare i discorsi d’odio sui social network, tuttavia, occorre considerare la posizione e gli interessi dei soggetti intermediari. Perciò sta emergendo la necessità di sottoporre normativamente i “custodi di internet” a obblighi analoghi alla diligenza informativa giornalistica⁵. Per altro verso, si punta a coinvolgere attivamente gli utenti che segnalano i contenuti e a garantire maggior trasparenza sulle notifiche e sulle cancellazioni, grazie alla creazione di una rete di “relatori di fiducia” (*trusted flaggers*). Il problema di fondo però è che il controllo, la censura o le sanzioni intervengono *ex post*, senza incidere sulle loro

⁴ Tra i paesi dell’Ue, la Germania con la l. 536/2017 (*Netzwerkdurchsetzungsgesetz*) e la Francia, con la l. 2020-766 (“legge Avia”), hanno approvato disposizioni per prevenire e contrastare il discorso d’odio sui social media e previsto sanzioni economiche per le piattaforme. Nonostante siano state presentate varie proposte di legge in Parlamento, ad oggi, in Italia, non vi è alcuna disciplina legislativa sul tema dell’*hate speech*. Questo ritardo è stato biasimato dalla “Commissione straordinaria per il contrasto dei fenomeni di intolleranza, razzismo, antisemitismo, istigazione all’odio e alla violenza”, istituita presso il Senato della Repubblica (XVIII Legislatura, 2019-2022) e presieduta dalla senatrice a vita Liliana Segre.

⁵ In tale direzione procede la *Raccomandazione CM/Rec(2022)16[1]* con cui l’UE sollecita gli stessi intermediari della rete a identificare e raccogliere dati e informazioni sui casi di *hate speech* online e a collaborare con le istituzioni statali e la società civile per migliorare le politiche e le campagne per contrastare questo fenomeno. Queste disposizioni renderebbero vincolanti le pratiche previste dal *Code of Conduct on Countering Illegal Hate Speech Online*, sottoscritto, dal 2016 ad oggi, dalla Commissione europea e quasi tutte le *corporation* del web.

sorgenti emotive, sociali, culturali. Questa prospettiva apre l'ampio tema delle risposte di tipo non giuridico: la corretta informazione, la creazione di contro-narrazioni, la sensibilizzazione alla diversità, l'educazione alla tolleranza, la conoscenza reciproca, l'uso critico dei mezzi di comunicazione e dei nuovi media. Si tratta di attività che coinvolgono attivamente gli attori della sfera pubblica, le agenzie di socializzazione e di formazione, le associazioni della società civile, le rappresentanze di interessi, il mondo del terzo settore e le organizzazioni di *advocacy*.

In tale direzione, da ultimo, può essere interessante svolgere una breve riflessione sulle "missioni" dei sociologi. Da Michael Burawoy riprendiamo le quattro figure idealtipiche con cui sono differenziati i loro compiti in base alle due variabili, la forma di sapere strumentale/riflessivo e l'orientamento interno/esterno: la sociologia professionale, la sociologia di *policy*, la sociologia critica e la sociologia pubblica (2005, trad. it. 2007: 13). Si tratta di "missioni" che devono poter essere messe a frutto contro i discorsi d'odio.

Il primo compito dei sociologi in quanto scienziati è quello professionale di descrivere e spiegare il fenomeno in sé e nel quadro più ampio dei processi sociali. Ad esempio, sono state compiute rilevazioni delle forme ricorrenti di *hate speech* (Burnap, Williams, 2015), elaborate delle "mappe dell'odio" attraverso l'analisi delle reti sociali, previste delle stime sulla loro diffusione (Stephens, 2013), identificati i contesti delle campagne d'odio, individuati i profili più influenti e le vittime ricorrenti e altri tipi di monitoraggio e interpretazioni degli eventi nel loro insieme. Queste attività scientifiche diagnostiche, in secondo luogo, vengono realizzate quasi sempre in collaborazione con le istituzioni pubbliche o le associazioni della società civile, nell'ottica di una sociologia di *policy* volta a migliorare gli interventi di controllo, calibrare le politiche educative e prevenire i possibili episodi di violenza. I risultati possono essere utili, infatti, non solo per le ricerche di altri istituti scientifici e accademici, ma anche alla scuola per la formazione di una *media literacy*, ai mass media, agli uffici di comunicazione pubblica, alla polizia e alla magistratura per definire strategie di contrasto (Scamuzzi, 2021: 76).

In terzo luogo, le ricerche sui discorsi d'odio possono generare una maggiore consapevolezza riflessiva e critica sulla *doxa* sociologica e sui presupposti incorporati nell'*habitus* che i sociologi che mettono in opera implicitamente in quanto tali e come attori nella riproduzione dell'ordine sociale. L'oggettivazione delle disposizioni e dei condizionamenti sessuali, etnici, culturali, etc. ma anche delle strutture dei campi e dello spazio

sociale concorre a un decentramento dei punti di vista che riduce il rischio di complicità.

Last but not least, i sociologi dovrebbero impegnarsi – come “terza missione” – nella diffusione dei risultati delle loro ricerche, intervenendo direttamente nella sfera pubblica, al fine di alimentare dibattiti con informazioni elaborate e interpretate attendibilmente, accrescere il capitale culturale della popolazione e concorrere all’elevazione della coscienza civile. Tali interventi sono volti a facilitare sia processi di auto-chiarificazione che il disvelamento delle false rappresentazioni che alimentano l’odio. All’orizzonte si profilano i nessi tra la scienza sociale e la democrazia e, quindi, le condizioni per una migliore circolazione del sapere in una sfera pubblica più discorsiva.

Riferimenti bibliografici

Amnesty International Sezione Italiana (2018). *Conta fino a 10. Barometro dell’odio in campagna elettorale*. Roma.

Amnesty International Sezione Italiana (2022). *Barometro dell’odio. Elezioni politiche 2022*. Roma.

Belluati M. (2018). Hate or Hateful? L’uso del linguaggio d’offesa nelle discussioni politiche. *Comunicazione Politica*, 3: 373-392.

Bentivegna S., Rega R. (2020). I discorsi d’odio online in una prospettiva comunicativa: un’agenda per la ricerca. *Mediascapes journal*, 16: 151-171.

Benussi A. (2019). Hate speech. Una categoria inattendibile. *Biblioteca della libertà*, LIV(224): 39-54.

Bortone R., Cerquozzi F. (2017). L’hate speech al tempo di Internet. *Aggiornamenti Sociali*, vol. 12, n. 68: 818-827.

Brown A. (2018). What is so special about online (as compared to offline) hate speech?. *Ethnicities*, vol. 18, n. 3: 297-326.

Burawoy M. (2005). Per la Sociologia pubblica. *Sociologica*, 1, 2007: 1-45.

Burnap P., Williams M.L. (2015). Cyber hate speech on Twitter: An application of machine classification and statistical modelling for policy and decision making. *Policy & Internet*, vol. 7, n. 2: 223-242.

Castaño-Pulgarín S.A., Suárez-Betancur N., Tilano Vega L.M., Herrera López H.M. (2021). Internet, social media and online hate speech. Systematic review. *Aggression and Violent Behavior*, 58: 1-7.

Fotia L. (2022). Contrastare il discorso e le politiche dell’odio: il contributo del mondo accademico. In: Id., a cura di, *Discorso d’odio e politiche dell’odio tra passato e presente*. Roma: RomaTrE-Press.

Frankfurt H. (2005). *On Bullshit*. Princeton: Princeton University Press.

Fricker M. (2007). *Epistemic Injustice. Power and the Ethics of Power*. Oxford: Oxford University Press.

Gagliardone I., Danit G., Alves T., Martinez G. (2015). *Combattre les discours de haine sur internet*. Paris: Organisation des Nations Unies pour l’éducation, la science et la culture.

Gherardi L., a cura di, (2022). *Lezioni brevi sull'opinione pubblica. Nuove tendenze nelle scienze sociali*. Roma: Meltemi.

Habermas, J. (2022). *Nuovo mutamento della sfera pubblica e politica deliberativa*. Milano: Raffaello Cortina.

Hassan G., Rabah J., Madriaza P., Brouillette-Alarie S., Borokhovski E., Pickup D., Varela W., Girard M., Durocher-Corfa L., Danis E. (2022). PROTOCOL: Hate online and in traditional media: A systematic review of the evidence for associations or impacts on individuals, audiences, and communities. *Campbell Systematic Reviews*, vol. 18, n. 2: 1-16.

Petrilli R. (2022). Discorso d'odio e opinione pubblica. In: L. Gherardi, a cura di, *Lezioni brevi sull'opinione pubblica. Nuove tendenze nelle scienze sociali*. Roma: Meltemi.

Scamuzzi S., Belluati M., Caielli M., Cepernich C., Patti V., Stecca S., Tipaldo G. (2021), Fake news e hate speech. I nodi per un'azione di policy efficace. *Problemi dell'informazione*, 1: 49-81.

Stephens M. (2013). *The geography of hate*, in "Floatingsheep", 13 maggio.

Tipaldo G. (2019). *La società della pseudoscienza. Orientarsi tra buone e cattive spiegazioni*. Bologna: il Mulino.

Ziccardi G. (2016). *L'odio online. Violenza verbale e ossessioni in rete*. Milano: Raffaello Cortina Editore.