

VARIAZIONE LINGUISTICA NEI COMMENTI SU FACEBOOK

*Mirko Tavosanis*¹

1. INTRODUZIONE

La comunicazione elettronica tramite Facebook è una realtà consolidata, che da diversi anni coinvolge circa la metà dei cittadini italiani: nel 2018, sul totale della popolazione, gli iscritti erano il 56% e gli utenti regolari il 48,4% (Censis, 2018: 105; We Are Social, 2018 stimava invece che gli utenti attivi fossero il 60% della popolazione). Non si tratta dei numeri più alti tra i sistemi paragonabili, perché le percentuali relative a WhatsApp sono ancora superiori, con il 64,4% della popolazione che ne fa un uso regolare (Censis, 2018: 105), mentre la televisione arriva addirittura all'89,9% (Censis, 2018: 54). Tuttavia, la diffusione di Facebook è evidente.

Questa situazione non è eterna. Facebook è nato solo nel 2004, è diventato accessibile alla maggior parte degli utenti italiani solo nel 2006 e si è diffuso con numeri significativi solo a partire dal 2008, con la localizzazione del sito in lingua italiana. Se il percorso di vita di altre piattaforme un tempo molto diffuse, da America On Line a MySpace, può servire come termine di confronto, probabilmente Facebook non durerà molto. In diversi paesi, a cominciare dagli Stati Uniti il suo uso presso le fasce d'età più giovani – in particolare in età adolescenziale – è in calo, a favore in particolare di Instagram. In Italia, nella fascia d'età tra i 14 e i 29 anni Facebook è oggi meno diffuso che nella fascia d'età tra i 30 e i 44 (Censis, 2018: 11). Non è detto quindi che il ruolo di questa rete sociale sia destinato a mantenersi a lungo.

Comunque, per il momento, l'importanza della comunicazione elettronica e delle reti sociali nel mondo contemporaneo è molto enfatizzata e a Internet, o specificamente a Facebook, viene spesso assegnato un ruolo chiave nelle trasformazioni sociali e nei fenomeni politici (Censis, 2018: 140-141; 178-179). Le dimostrazioni di questa importanza, però, mancano, e in mancanza di dati è consigliabile molta prudenza nelle conclusioni. Se in buona parte del mondo la globalizzazione ha portato a un'effettiva trasformazione del modo di vita, nelle società sviluppate le analisi sul campo non mostrano vere discontinuità nel comportamento e negli atteggiamenti giovanili dopo l'arrivo dei nuovi mezzi di comunicazione (boyd, 2014). Sembra quindi molto difficile assegnare a questi mezzi di comunicazione un ruolo causale nelle evoluzioni politiche e sociali – mentre è chiaro che sono una componente di rilievo nella rete di strutture e rapporti dell'Italia contemporanea.

In particolare, dal punto di vista linguistico Facebook è uno dei veicoli che, assieme ai precedenti SMS, al più diffuso WhatsApp e ad altri sistemi più di nicchia, negli ultimi vent'anni hanno portato milioni di italiani a un uso quotidiano della lingua scritta. Rispetto a quanto avviene in quasi tutti gli altri sistemi, tuttavia, la comunicazione su Facebook è o è diventata in buona parte pubblica, anche all'insaputa degli utenti o in modi di cui gli utenti stessi non sono in grado di cogliere tutte le implicazioni.

¹ Università di Pisa.

Questi due fattori, il fatto che sia pubblica e il fatto che gli utenti spesso non si rendano conto appieno della cosa, rende la comunicazione su Facebook di particolare interesse per chiunque voglia osservare e controllare in tempo reale l'italiano scritto di una parte molto ampia della popolazione. In particolare, per chi vuole osservare la variazione linguistica.

2. VARIETÀ E VARIAZIONI

Le discussioni sull'esistenza dell'italiano della comunicazione elettronica come varietà di lingua, intesa come «un insieme solidale di varianti di variabili linguistiche» (Berruto, 2011: 1550) sembrano ormai arrivate a un consenso. Nella comunicazione elettronica nel suo assieme non esiste un tale «insieme solidale», ma si spazia per esempio da un massimo a un minimo di formalità, e in generale sono presenti differenze su tutti gli assi di variazione (Tavosanis, 2011: 33-39). Sembra che la rete, più che creare una varietà di italiano della comunicazione elettronica, abbia semplicemente «fatto emergere il modo in cui gli italiani si esprimono» (Pistoiesi, 2014: 364), ma ci si può chiedere se per esempio esista un «italiano digitato» informale o «e-taliano», secondo diverse formulazioni di Giuseppe Antonelli (per esempio, Antonelli, 2014). Anche questo, tuttavia, sembra molto dubbio. In una lingua vista come un «*continuum* con addensamenti» (Berruto, 1987: 29) resta difficile identificare e isolare varietà specifiche. Per esempio, Giuliana Fiorentino (2018: 46) preferisce, «più che mostrare varietà autonome», individuare nella comunicazione elettronica «fattori di variabilità» come i fenomeni di 'allegro', la *lingua selvaggia* e i gerghi giovanili.

Va poi ricordato che nella comunicazione elettronica si possono trovare tratti con un alto grado di convenzionalità, anche se appresi in un contesto informale, secondo i principi della *vernacular literacy* (Barton, Lee, 2012) e delle comunità di apprendimento: per esempio, diverse scelte di punteggiatura e ortografiche (Tavosanis, 2013). Un corollario di questa convenzionalità è il fatto che in alcuni casi i messaggi sono il frutto di una simulazione di parlato e di disinvoltura da parte degli utenti, come è stato notato non solo dai linguisti ma anche da una serie di disegnatori e autori di meme che va da Randall Munroe a Federico Palmaroli.

Altri aspetti della comunicazione dipendono dalla modalità dialogica dello scambio in tempo reale, che produce anche effetti indipendenti dal canale comunicativo usato (Voghera, 2017: 39 e 41). Altri ancora sono viceversa dovuti alle caratteristiche del mezzo comunicativo usato, o alle abitudini trasferite da un genere testuale all'altro, o alle competenze di un determinato scrivente. A quest'ultimo punto è poi connessa la questione dell'esistenza o meno di una categoria tradizionale della sociolinguistica, quella dei «semicolti», e del ruolo che le deviazioni dalla norma hanno oggi (Fresu, 2016: 339-343).

In questo contesto, si potrebbe considerare buona parte della conversazione elettronica la manifestazione scritta di una varietà di italiano parlato colloquiale. Tuttavia, l'incertezza che nell'ultimo decennio ha caratterizzato le classificazioni rende molto difficile una proposta precisa. Basterà ricordare, nella loro varietà, le etichette citate da Berruto (2011: 1551-1552): «italiano regionale» (nelle sue numerose realizzazioni, disposte lungo un *continuum*), «italiano parlato informale», «italiano incolto», «italiano dell'uso medio», «italiano popolare», «italiano colloquiale», «italiano gergale». Non è facile dire quali tra queste etichette siano le più adatte a caratterizzare ciò che viene oggi scritto in rete, e in particolare su Facebook.

3. VARIAZIONE E GENERE TESTUALE

È comunque evidente che i fattori di variabilità si manifestano in modi diversi all'interno dei diversi generi testuali. Descrivere tutti i generi attivi è impossibile, ma approfondire un tipo di comunicazione contribuisce a chiarire i ruoli dei fattori attivi.

La comunicazione su Facebook, in particolare, non si può considerare appartenente a un unico genere: i testi visti dagli utenti attraverso una pagina web o l'app di Facebook, appartengono a diversi generi testuali (su questa situazione, normale nei prodotti di comunicazione elettronica, v. Tavosanis, 2011: 60-61). All'interno del proprio flusso di aggiornamenti, per esempio, e con una presentazione che li mostra sullo stesso piano, un utente può vedere post provenienti da amici, post formali provenienti da un giornale o da un personaggio pubblico e post che sono veri e propri interventi pubblicitari.

Soprattutto, su Facebook è notevole la differenza tra i post e i commenti ai post stessi, sia per i testi prodotti da professionisti sia per i testi prodotti da utenti comuni. I commenti sono parti di un dialogo alla pari; i post sono enunciazioni di un autore e, anche quando sono pensati esplicitamente per aprire una discussione e ottenere commenti, si collocano su un piano diverso rispetto ai commenti stessi (come sarà mostrato, per alcuni indicatori, nel paragrafo 6).

Andando più in dettaglio, ci si può quindi chiedere se il «commento» sia o meno un unico genere testuale, con caratteristiche linguistiche definite. Tra i commenti sono del resto evidenti sia alcuni tratti comuni sia alcune differenze vistose. Le caratteristiche del singolo messaggio dipendono infatti dall'argomento di cui si parla, dall'interazione, dalle competenze delle persone coinvolte e così via.

Elena Pistolesi (2018) ha attirato l'attenzione sulla differenza tra due tipologie di comportamenti in linea e di commenti. Da un lato, gli enunciati piuttosto semplici che commentano su YouTube un video di Favij, e che si manifestano come messaggi

telegrafici, ricchi di emoji, emoticon, lettere maiuscole enfatiche, reduplicazione delle lettere con effetto eco, acronimi (lol); [in questi commenti] abbondano i punti esclamativi e le interiezioni; l'omissione degli accenti e gli errori di digitazione sono frequenti [...] non ci troviamo dinanzi a una comunità, ma a milioni di ragazzi che seguono il loro idolo senza elaborare contenuti, né avere una rete di relazioni che vada oltre la convergenza sul canale di Favij (Pistolesi, 2018: 21).

Dall'altro, i commenti inseriti in un gruppo Facebook attorno al quale si è aggregata una specifica comunità, con le proprie regole linguistiche. All'interno dei testi prodotti da questo gruppo «il ricorso alle emoji e alla punteggiatura modale (punto esclamativo, interrogativo e punti di sospensione) è limitato se confrontato con gli usi grafici degli adolescenti su Youtube o di alcuni utenti di Twitter» (Pistolesi, 2018: 22-23).

Procedimenti di questo tipo sono in continuità con la forte distinzione tra attività «di socializzazione» e «tematiche» che si ritrovano in molti generi testuali, per esempio nella distinzione tra forum «professionali» e forum «non professionali» (Tavosanis, 2011: 176). La possibilità di passare dalle attività di socializzazione informale ad altre più strutturate, spesso in rapporto a un'area di interesse, è del resto stata riscontrata da numerosi osservatori e la relativa progressione condensata nella formula «Hanging out, messing around and geeking out» che dà il titolo a un'importante raccolta di studi relativi alle comunità online americane (Ito *et alii*, 2010).

In sostanza, è ovvio che all'interno dei commenti, inclusi quelli di Facebook, come minimo si nota una differenza tra gli interventi occasionali, magari fatti solo per lasciare un tributo nello spazio appartenente a un personaggio famoso, e quelli più legati a una presenza costante, a una discussione con interlocutori che tornano e si conoscono tra di loro. Tuttavia, è altrettanto facile notare che questa differenza non corrisponde a una rigida distinzione di spazi: anche all'interno di serie di commenti che si limitano in sostanza a esprimere approvazione o condanna per un personaggio si trovano scambi più articolati tra persone che a volte erano già reciprocamente note e formano almeno un embrione di comunità. Andare più in dettaglio richiede un'elaborazione di dati.

4. RACCOLTA DI CORPORA

La costruzione di *corpora* di testi tratti da Facebook è possibile in quanto il sistema stesso consente lo scaricamento ordinato di materiali attraverso le API (*Application Programming Interface*) disponibili sul sito Facebook for developers (<https://developers.facebook.com/>). L'interazione con le API è basata su uno scambio di messaggi secondo il protocollo HTTP. In questo modo si può per esempio usare il metodo HTTP GET applicato a un URI (*Uniform Resource Identifier*) corrispondente a pagine pubbliche, col fine di richiederne i post, e a post pubblici, col fine di richiederne i commenti; i dati restituiti dall'API di Facebook sono forniti in formato JSON (*JavaScript Object Notation*). Negli anni, Facebook ha modificato lo schema di base e ha reso progressivamente più difficile l'accesso ai materiali, anche in rapporto allo scandalo Cambridge Analytica. Tuttavia, con qualche accorgimento, ancora oggi è possibile ottenere risultati soddisfacenti, facendo ricorso a programmi esterni che sfruttano le API stesse.

I dati presentati qui di seguito sono quindi tratti da tre *corpora* di post e commenti assemblati con varianti di questa metodologia negli ultimi anni presso l'Università di Pisa al servizio di elaborati di laurea del Corso di Studi in Informatica umanistica:

1. «post politici» (in Meini, 2014): 221.226 token;
2. «commenti politici» (in Tortorelli, 2017): 263.000 token, include il *sottocorpus* «commenti Martina», elaborato in Nasuto (2018): 18.564 token;
3. «commenti spettacolo» (in Cassarà, 2018): 135.610 token.

Il *corpus* «post politici» è descritto in dettaglio in Tavosanis (2016: 682-683). Il *corpus* «commenti politici» è formato dall'analisi dei commenti ai post di 14 ministri del governo Gentiloni. Il *corpus* «commenti spettacolo» ha preso in esame quattro tra i dieci personaggi con più seguito nelle rispettive pagine su Facebook: i due cantanti con più seguaci, Laura Pausini (7 milioni) ed Eros Ramazzotti (circa 4,5 milioni), e i due personaggi di altro genere con più seguaci, la conduttrice televisiva Belén Rodríguez (6 milioni) e lo youtuber Frank Matano (circa 3,2 milioni). Dalle pagine Facebook sono stati scaricati in totale 18.065 commenti, di cui 13.092 con un contenuto linguistico; tra questi ultimi, 11.215 sono risultati commenti in lingua italiana.

Per testi di questo tipo, occorre ricordare che non sempre si può essere certi dell'autenticità dei commenti. La pratica di gonfiare le statistiche sulle reti sociali, per far apparire un personaggio o un prodotto più seguito di quanto in realtà lo sia, è ben documentata. Su Twitter, in particolare, la creazione di utenti fittizi con profili realistici, capaci anche di rispondere polemicamente in caso di contestazioni, sembra frequente, e

pratiche simili si ritrovano anche sul più controllato Facebook. Tuttavia, esaminando casualmente i profili pubblici di alcuni degli utenti che hanno realizzato i messaggi dei corpora, non è mai venuta l'impressione che i profili stessi fossero falsi. In altri termini, sembra che dietro a questi commenti su Facebook siano presenti soprattutto persone reali.

5. VARIAZIONE DI LINGUA

L'esame dei *corpora* permette innanzitutto la quantificazione di un primo livello di variazione, radicale: l'accettabilità di lingue diverse all'interno di uno stesso spazio comunicativo – e in pratica all'interno di una stessa serie di commenti.

Questa accettabilità va innanzitutto collegata a un contesto in cui la comunicazione su Facebook è molto legata alla lingua nazionale di riferimento, anche a livello di interfaccia (la scelta viene compiuta al momento dell'iscrizione, anche se può essere modificata in seguito). In generale, si ha l'impressione che la maggior parte degli utenti veda contenuti nella propria lingua nazionale di riferimento, e spesso solo in quella.

Per fare un passo oltre le impressioni e le esperienze del singolo individuo, è possibile usare gli strumenti di gestione delle campagne promozionali di Facebook. Con questo sistema si ricava che, a fine 2018, le lingue usate regolarmente dagli iscritti a Facebook residenti in Italia, 34 milioni in totale, erano queste (per la metodologia e i dettagli: Tavosanis, in stampa):

Italiano	Inglese (tutte le varietà)	Rumeno	Spagnolo (tutte le varietà)	Francese (tutte le varietà)
33.000.000	9.900.000	1.000.000	840.000	700.000
Arabo	Albanese	Tedesco	Portoghese (tutte le varietà)	Cinese (tutte le varietà)
510.000	390.000	250.000	190.000	100.000

Interpretando questi dati, l'ipotesi che massimizza il numero degli utenti che usano l'italiano accanto ad altre lingue è che tutti gli iscritti residenti in Italia che usano altre lingue siano anche attivi in italiano. Accettando questa ipotesi, restano 19.120.000 utenti di Facebook in italiano, cioè oltre metà del totale, che apparentemente non hanno contatti regolari con lingue diverse dall'italiano. Il quadro viene comunque complicato dalle difficoltà a determinare la situazione all'estero (Tavosanis, in stampa) o dal ruolo della traduzione automatica, che dal 2016 presenta spesso agli iscritti a Facebook la traduzione italiana di testi originariamente stesi in altre lingue, e viceversa (Tavosanis, 2018).

Tuttavia, esiste evidentemente un certo livello di alternanza nell'uso delle lingue, in rapporto alle diverse situazioni comunicative. Il successo di alcuni personaggi dello spettacolo, per esempio, è legato anche a un pubblico internazionale (e i post sulle loro pagine sono a volte scritti in spagnolo e in inglese, oltre che in italiano) – mentre alcuni tipi di notorietà sono legati al solo pubblico italiano. Tra i casi presi in esame nel *corpus* «commenti spettacolo», i commenti in italiano vanno quindi dal minimo del 79,38% per Eros Ramazzotti al massimo del 97,67% per Frank Matano. Tra i commenti non in italiano, la lingua più rappresentata è lo spagnolo, che arriva al 15,72% nel caso di Laura Pausini, mentre l'inglese, pur essendo la seconda lingua più rappresentata, seguita a breve

distanza dal portoghese, ottiene un livello massimo del 3,91% nei commenti a Eros Ramazzotti. Tutto questo può essere interessante anche in rapporto alle discussioni recenti sull'importanza dell'intercomprensibilità e sul ruolo esclusivo che spesso viene dato all'inglese nelle politiche italiane di internazionalizzazione.

Più in dettaglio, nel caso di Belén Rodriguez, su 4889 commenti, 1194 sono risultati privi di contenuto linguistico: erano cioè composti solo da immagini, sticker, interiezioni, emoji e simili. Inoltre, in 9 casi non è stato possibile assegnare con sicurezza a una lingua il contenuto linguistico presente, e in 18 erano presenti due lingue a titolo più o meno paritario. Tra i commenti residui, 3207 sono risultati in italiano; tra gli altri, la ripartizione è stata:

Tabella 1. *Lingue diverse dall'italiano usate nei commenti ai post di Belen Rodriguez.*

Albanese	1
Bulgaro	1
Croato	1
Dialetti italiani	61
Francese	9
Greco	1
Inglese	122
Maltese	10
Montenegrino	1
Olandese	1
Portoghese	1
Rumeno	4
Slovacco	1
Spagnolo	244
Swahili	1
Tedesco	2
Totale	461

Tra i commenti in italiano, il livello di conoscenza della lingua e di formalità nell'uso si è poi rivelato molto variabile. Alcuni esempi di commento spaziano dalle citazioni, con la ripresa dell'aria *O soave fanciulla* dalla *Bohème* di Puccini (1), all'italiano standard con alcuni errori ortografici, incluso il classico “pò” accentato (2), a testi che sono probabilmente realizzati da stranieri con una ridotta conoscenza dell'italiano (3 e 4):

- (1) O soave fanciulla...o dolce viso di mite circonfuso albalunar...♥
- (2) il social media manager è un pò confuso con l'italiano e spagnolo
A differenza che se lei sbaglia l'italiano è tra virgolette “scusata” visto che non è italiana, ma quando a sbagliare l'italiano sono gli italiani ...Allora si che è una cosa grave 😊😊😊😊😊😊
- (3) Belle .tu sei la miliorellll
- (4) Bella donna conoscere bene donna fetta bela amore mio bela amore mio
bela amore mio bela amore mio

Un alto livello di variazione si ha anche per i commenti in dialetto, principalmente di origine meridionale (ma il dato più interessante è che, rispetto ai commenti in italiano, quelli in dialetto sono meno del 2%). Gli esempi 5 e 6 sono rispettivamente in un dialetto centrale (umbro o marchigiano) e in napoletano:

(5) [nome e cognome] ce sente cerqua???

(6) E vuo' taglià nu poc sti capill a stu criatur??Ten o pagliaron ngap con questo caldoooooo

6. STATISTICHE LINGUISTICHE

Rimanendo sui soli testi scritti in italiano, anche il più superficiale degli esami mostra che i commenti abbondano di frasi fatte e di espressioni riprese da varie aree: linguaggio dei media, linguaggio giornalistico, meme, tormentoni, classici della letteratura e così via. In assenza di una caratterizzazione sociolinguistica delle persone che scrivono commenti su Facebook, si può comunque notare che il linguaggio si allontana molto dalle ipotetiche caratteristiche di una varietà popolare pura o da semicolti.

Per descrivere meglio somiglianze e differenze può quindi essere utile esaminare i testi con strumenti computazionali, e in particolare con il sistema dimostrativo READ-IT elaborato dall'Istituto di Linguistica Computazionale del CNR di Pisa (Dell'Orletta, Montemagni, Venturi, 2011). I dati quantitativi di base relativi ai «commenti spettacolo» sono riportati nella Tabella 2.

Tabella 2. *Statistiche di base per i sottocorpus del corpus «commenti spettacolo».*

	Numero di periodi	Numero di periodi per commento	Numero di token	Lunghezza media periodi	Lunghezza media token
Pausini	5895	1,78	64880	11,0	4,6
Ramazzotti	2837	1,35	21258	7,5	4,6
Rodriguez	3896	1,20	26162	6,7	4,9
Matano	3179	1,24	23310	7,3	4,8

In sostanza, assieme di commenti di questo tipo sono abbastanza simili tra di loro per quanto riguarda le statistiche di base. Al tempo stesso, si differenziano nettamente rispetto ai «post politici», dove per esempio il numero medio di periodi per commento è 3,02 e la lunghezza media dei periodi è 17,93 (Tavosanis, 2016: 683). Evidentemente, le presentazioni ufficiali di temi e opinioni richiedono più spazio e un discorso più articolato rispetto a sintetiche espressioni di ammirazione e partecipazione.

In altri casi, comunque commenti e post si assomigliano. Per esempio, i rapporti tipo/unità e la densità lessicale del *corpus* «commenti spettacolo», presentati in Tabella 3, non si allontanano molto da quelli del *corpus* «commenti politici», che ha una densità lessicale di 0,573 (Tavosanis, 2016: 683).

Tabella 3. *Rapporto tipo / unità e densità lessicale nei sottocorpora del corpus «commenti spettacolo».*

	Rapporto tipo/unità	Densità lessicale
Pausini	0,71	0,550
Ramazzotti	0,64	0,567
Rodriguez	0,87	0,602
Matano	0,81	0,605

Le differenze negli indicatori sono comunque frequenti, dato che la distribuzione dei tratti linguistici difficilmente presenta caratteristiche di isotropia. Nella Tabella 4 sono riportate le percentuali di presenza delle parti del discorso all'interno del *corpus* «commenti spettacolo», affiancate a quelle del *corpus* «post politici» e a due corpora usati come riferimento da READ-IT: un *corpus* di articoli del quotidiano *La Repubblica*, e un *corpus* di articoli del mensile *Due parole*, in italiano semplificato (Dell'Orletta, Montemagni, Venturi, 2011: 78). Le differenze tra i *sottocorpora* sono relativamente contenute. Tra le differenze più significative, si nota che nei commenti ai post di Frank Matano compare un maggior numero di nomi propri, in quelli ai post di Belen Rodriguez un maggior numero di aggettivi.

Tabella 4. *Distribuzione percentuale delle parti del discorso in una serie di corpora.*

	Sostantivi	Nomi propri	Aggettivi	Verbi	Congiunzioni	di cui coordinanti	di cui subordinanti
Pausini	15,4	5,7	6,6	14,2	4,2	68,7	31,3
Ramazzotti	15,4	7,2	8,4	12,0	3,1	69,8	30,2
Rodriguez	17,3	7,3	12,5	12,5	4,2	67,9	32,1
Matano	15,1	12,1	5,6	14,5	3,5	65,7	34,3
<i>Corpus</i> politici	20	8	6	13	4		
<i>Due parole</i>	23,8	5,4	5,9	13,7	3,7	81,9	17,8
<i>La Repubblica</i>	21,5	5,4	6,4	12,9	3,6	74,6	25,1

Decisamente più significative sono le differenze per quanto riguarda la sintassi, presentate nella Tabella 5. Nel *corpus* «commenti spettacolo», la subordinazione è molto più contenuta, in evidente rapporto con la brevità dei commenti: le proposizioni subordinate rappresentano il 32,6% del totale delle proposizioni nel *corpus* di articoli giornalistici, ma nei *sottocorpora* di commenti si collocano tra un massimo del 17,3% e un minimo del 10,8%. Tuttavia, in aree come il numero medio di parole per proposizione o di dipendenti per testa verbale le differenze sono molto limitate.

Tabella 5. *Quantificazione di indicatori sintattici in una serie di corpora.*

	n. medio parole per proposizione	n. medio di dipendenti per testa verbale	% principali	% subordinate	Media delle altezze massime	Profondità media di strutture nominali complesse	Profondità media catene di subordinazione	Lunghezza media	Media delle lunghezze massime
Pausini	8,364	2,125	82,7	17,3	2,940	1,093	1,173	2,931	4,815
Ramazzotti	9,680	1,970	89,2	10,8	2,217	1,080	1,132	2,267	3,194
Rodriguez	9,113	1,940	87,1	12,9	2,096	1,096	1,153	2,177	2,994
Matano	8,113	1,945	86,1	13,9	2,283	1,110	1,174	2,084	3,036
<i>Due parole</i>	9,845	2,221	73,7	26,3	5,292	1,243	1,007	2,161	7,929
<i>La Repubblica</i>	10,153	2,090	67,5	32,5	6,532	1,351	1,171	2,398	10,316

In sostanza, tra questi commenti a post di Facebook e altri tipi di testo esistono alcune differenze linguistiche ben individuabili. In altri casi, però, le differenze individuate dagli indicatori non sembrano particolarmente significative. Corrispondentemente, i commenti hanno caratteristiche simili ma non del tutto unitarie. Per una caratterizzazione più precisa occorre quindi andare più in profondità.

7. NUOVI USI LINGUISTICI E SNOBISMI

Un utile punto di partenza è costituito dai tratti innovativi dell'italiano descritti da Gaetano Berruto (2017: 39-56) per diverse aree sintattiche:

- Marcatezza della struttura delle frasi
- Rinnovamento nell'uso di alcune forme verbali
- Nuovi usi del complementatore *che*
- Nuovo sistema pronominale
- Innovazioni sintattiche
- Nuovi sistemi di formazione di parole
- Mode linguistiche («Linguistic customs»)

L'ultimo elemento della lista in parte corrisponde a quelli che Renzi (2012: 47), sulla scorta di Labov, ha definito «snobismi», spesso di taglio brillante o giornalistico. Gli usi di questo tipo colpiscono spesso l'attenzione. Di fatto, però, sono anche rarissimi in molti tipi di testo, e in particolare nel *corpus* «commenti spettacolo».

Uno degli usi descritti da Berruto (2017: 49-50) è quello di «grazie di» + infinito presente riferito al passato, anziché riferito al presente o al futuro: «grazie di inviarmi i compiti richiesti» anziché «grazie di avermi inviato i compiti richiesti». Nel *corpus* «commenti spettacolo» si trova solo un caso di «grazie di» + infinito presente, ed è un

caso in cui la persona che scrive, con ogni probabilità, intende riferirsi al presente, più che al passato:

(7) Grazie di renderci sempre così' partecipe di te.

È viceversa presente nel *corpus*, ma in forma estremamente ridotta, un altro degli usi menzionati: la grammaticalizzazione della parola *tipo*. Quest'uso è stato descritto da Miriam Voghera (2013) e suddiviso in sei stadi di progressiva grammaticalizzazione:

- come aggettivo
- per la formazione di aggettivo
- come preposizione
- avverbio
- connettivo
- segnale discorsivo

Nel *corpus* l'uso grammaticalizzato di *tipo* è rappresentato da sole due occorrenze, entrambe provenienti da commenti a post di Frank Matano ed entrambe nel senso di «simile a»:

(8) Ma è tipo il disco di platino ?

(9) ... avere episodi tipo come questi ancora oggi.

Nel secondo caso si ha addirittura la duplicazione del significato con «come», a indicare probabilmente una certa incertezza nell'uso della forma innovativa, percepita come troppo debole (Piotti, in questi stessi Atti, cita occorrenze del genere «tipo per esempio»).

Un terzo uso linguistico innovativo, la locuzione «piuttosto che» con valore disgiuntivo, è del tutto assente dal *corpus*. La locuzione compare una sola volta, in un commento a un post di Laura Pausini, e lo fa con il significato tradizionale:

(10) fanno più scalpore gli scoop superficiali piuttosto che le azioni vere come queste

Quarto uso innovativo, la formula *e/o*, che compare una sola volta, anche in questo caso come commento a un post di Laura Pausini:

(11) Spero da tempo in una nuova collaborazione e/o programma insieme

Insomma, da questo punto di vista la scrittura del *corpus* è molto diversa da una scrittura giornalistica o pubblicitaria. La differenziazione si trova però anche in direzione opposta, rispetto al parlato colloquiale. All'interno del *sottocorpus* «commenti Martina», per esempio, esaminando un campione di 1500 commenti che includono la parola *che*, su 687 usi considerabili «relativi» solo 26 sono non standard (5 con funzione di soggetto, 8 di oggetto diretto, 13 di complemento). Quantitativamente, si tratta quindi del 3,8% del totale, una percentuale molto inferiore al 15% indicato da Alfonzetti (2002: 156) anche per il parlato di persone colte. Inoltre, si ha una forte selezione tra le tipologie di *che* polivalente presentate. Il *che* con clitico di ripresa, molto comune nel parlato, è presente con un'unica occorrenza tra gli esempi studiati:

(12) ...la proposta che vuoi proporla agli altri

Per confronto, all'interno di 2573 commenti in cui compare la parola *che*, compaiono anche numerose occorrenze di pronomi standard:

- il quale: 8
- i quali: 3
- la quale: 1
- le quali: 10
- cui: 188
- dove: 172

La situazione dei commenti è quindi un caso di convivenza di forme, esempio delle situazioni in movimento descritte da Lorenzo Renzi (2012), in cui prima o poi la comunità dei parlanti potrebbe prendere una decisione esclusiva in favore dell'una o dell'altra opzione.

8. IL SINGOLO UTENTE

All'interno di questo contesto un promettente percorso di approfondimento, per quanto difficile da seguire, sembra quello del singolo scrivente. In aggiunta allo studio di un genere testuale distinto, le caratteristiche dei post e dei commenti pubblici su Facebook permettono infatti in una certa misura di identificare le competenze e le scelte linguistiche dei singoli.

Tra le difficoltà pratiche di questo lavoro ci sono i limiti imposti dalla privacy e dall'applicazione del recente regolamento europeo GDPR per la protezione dei dati personali. Tuttavia, i controlli fatti occasionalmente e individualmente mostrano l'interesse che potrebbero avere tali informazioni.

L'impressione generale, per quanto ancora non quantificabile, è che gli utenti non modificano molto il proprio modo di scrivere a seconda della situazione. Gestire la variazione diafasica è possibile, ma richiede sforzo: molto spesso le scelte linguistiche dei commenti sembrano costanti nel tempo. Inoltre, la differenziazione tra post e commenti ha in pratica un peso ridotto a livello del singolo scrivente. È vero che i post spesso sono diversi dai commenti a causa di diversi fattori, oltre alla semplice lunghezza; tuttavia, questa diversità in parte si spiega per il fatto che su Facebook i post spesso sono fatti da testo copiato e incollato da altre fonti: altri post, ma anche articoli di giornale o (più spesso di quanto forse si potrebbe pensare) testi collegati al lavoro. Sono piuttosto rari invece i post originali, cioè quelli di persone che scrivono autonomamente sulla propria pagina un testo che genera risposte.

Un esempio tipico può essere costituito da un'utente che pubblica sulla propria pagina post che contengono un testo probabilmente fornito da un'azienda, e forse con leggera rielaborazione individuale. Un post di questo tipo (con la cancellazione dei nomi di azienda) è presentato nell'Immagine 1. La sua formulazione assomiglia infatti molto a quella di altri messaggi usati nelle campagne di reclutamento di aziende di cosmetici (per esempio, «Entra nel mio team ***, Diventa Imprenditrice di te stessa e offri questa tua opportunità ad altre persone che, come te hanno voglia e desiderio di successo nel settore del Wellness&Beauty, un mercato sempre più in crescita»). Nell'Immagine 2 è presentato

un commento della stessa utente. In entrambe le immagini, così come nelle successive 3 e 4, i dati sensibili sono stati cancellati.

Immagine 1. *Post su Facebook.*

Cerco Consulenti di bellezza, manager capo gruppo DONNA O UOMO esperta/o nella vendita diretta a cui affidare la zona per incrementare la rete commerciale.
Questo è il mese giusto... Entra nel mio Team e scopri quanto è meraviglioso lavorare in 🍀🍀🍀🍀🍀❤️❤️❤️
Vi aspettiamo al Meeting per la presentazione dell'offerta Lavorativa, domani, Venerdì 06/10/2017 alle ore 10:00 presso la pizzeria " "...
Confermate la vostra adesione nei commenti 🍷

Immagine 2. *Commento su Facebook.*

Fa solo ridere... C'è una piccola differenza tra il M5S e voi... Nel movimento, le mele marce (che possono essere ovunque) vengono espulsi, voi invece se non sono indagati non li volete... Cerchi di fare pulizia nel suo partito prima di guardare gli altri... Buon Lavoro

In particolare, colpisce che i puntini di sospensione e l'uso di maiuscole in casi in cui lo standard vuole le minuscole siano presenti, sia pure in grado diverso, nei due testi – che peraltro sono evidentemente il frutto di un diverso livello di elaborazione.

Anche dal punto di vista diacronico si nota una certa stabilità. Per esempio, un utente pubblica commenti scritti con incertezze ortografiche simili a sei anni di distanza:

Immagine 3. *Post su Facebook (2011).*

Ma a voi piace questo sistema referendario? Insomma vi sembra giusto che i no abbiano il 66.66...% delle possibilita' mentre il si solo il 33.33...%. Tutti a votare.

Immagine 4. *Commento su Facebook (2017).*

Ancora il coraggio di parlare dopo che i tui rubavano i voti ai vecchietti disabili e compravano i voti a 25 € ciascuno. Ma stai zitto hai perso un'altra occasione per stare mutu

Un'altra scrivente produce due messaggi con caratteristiche molto simili (per punteggiatura, grado di revisione e problemi di codifica) a un anno circa di distanza tra di loro:

(13) ...embè...per fortuna esiste la musica che fra un castello all aria all alto...porta con i piedi a terra o quasi.diciamo sulle punte dei piedi. ..ih... embè... embè...

(14) Io non capisco una cosa ..ma perché le forze dell ordinar nn fanno mai niente. .ma perché anche loro non lottano per portarr rispetto allaa

propria divisa e compito...io nn capisco..che possiamo fare noi? Io vivo a bologna ed è vero è uno schifo...salvini ma davvero potresti aiutare l'Italia almeno a rientrare nella propria dignità? Tu da solo contro tutti? Per la mafia camorra ci marciano non so sono molto inkazzara e demoralizzata...scusate lo sfogo

La variazione linguistica sembra quindi riconducibile più alla presenza di persone diverse, ognuna con i propri usi consolidati, che alle variazioni da parte di una stessa persona.

9. CONCLUSIONI

La varietà dei fenomeni presenti è tale da rendere difficile una classificazione unitaria. Per questo, sembra difficile assegnare al genere testuale dei commenti, con le sue variazioni, tratti linguistici specifici che lo individuino in modo univoco all'interno della comunicazione elettronica. Molto più promettente sembra la possibilità di isolare tipologie specifiche, sia nelle enunciazioni che non prevedono risposta sia in quelle che si inseriscono all'interno di un dialogo.

Ancora più interessante sembra la possibilità di valutare le caratteristiche di singoli utenti, costruendo profili linguistici di riferimento. Confermare le impressioni fornite finora richiede senz'altro un lavoro sistematico di notevole estensione. I risultati di un'analisi del genere potrebbero però essere di grande interesse non solo per l'esame della variazione linguistica su Facebook, ma anche per questioni di portata più ampia.

RIFERIMENTI BIBLIOGRAFICI

- Alfonzetti G. (2002), *La relativa non-standard: italiano popolare o italiano parlato?*, Centro studi filologici e linguistici siciliani, Palermo.
- Antonelli G. (2014), "L'e-taliano: una nuova realtà tra le varietà linguistiche italiane?", in Garavelli E. e Suomela-Härmä E. (a cura di), *Dal manoscritto al web: canali e modalità di trasmissione dell'italiano. Tecniche, materiali e sui nella storia della lingua. Atti del XII Congresso SILFI Società Internazionale di Linguistica e Filologia Italiana* (Helsinki, 18-20 giugno 2012), vol. II, Cesati, Firenze, pp. 537-556.
- Barton D., Lee C. M. K. (2012), "Redefining vernacular literacies in the age of Web 2.0", in *Applied Linguistics*, 33, 3, pp. 282-298.
- Berruto G. (1987), *Sociolinguistica dell'italiano contemporaneo*, Carocci, Roma.
- Berruto G. (2011), "Varietà", in Simone R. (a cura di), *Enciclopedia dell'italiano*, Istituto della Enciclopedia Italiana, Roma, pp. 1550-1553:
[http://www.treccani.it/enciclopedia/variet%C3%A0_\(Enciclopedia-dell%27Italiano\)/](http://www.treccani.it/enciclopedia/variet%C3%A0_(Enciclopedia-dell%27Italiano)/).
- Berruto G. (2017). *What is changing in Italian today? Phenomena of restandardization in syntax and morphology: an overview*, in Cerruti M., Crocco C. e Marzo S. (a cura di), *Towards a new standard. Theoretical and empirical studies on the restandardization of Italian*, De Gruyter - Mouton, Berlino - Boston, pp. 31-60.

- Boyd D. (2014), *It's Complicated: The Social Lives of Networked Teens*, Yale University Press, New Haven e Londra; tr. it. di Bertagna F., *It's complicated: la vita sociale degli adolescenti sul web*, Castelveccchi, Roma, 2014.
- Cassarà G. (2018), *Analisi linguistica di commenti a post Facebook di personaggi del mondo dello spettacolo*, tesi di laurea triennale in Informatica Umanistica, Università di Pisa.
- Censis (2018), *Quindicesimo rapporto sulla comunicazione. I media digitali e la fine dello star system*, FrancoAngeli, Milano.
- Dell'Orletta F., Montemagni S., Venturi, G. (2011), "READ-IT: Assessing Readability of Italian Texts with a View to Text Simplification", in *Proceedings of the 2nd Workshop on Speech and Language processing for Assistive Technologies*, Edinburgh, pp. 73-83.
- Fiorentino G. (2018), "Sociolinguistica: varietà del web nel repertorio in rete", in Masini, F. e Grandi, N. (a cura di), *CLUB Working Papers in Linguistics*, vol. 2. CLUB Circolo Linguistico Università di Bologna, Bologna, pp. 40-60.
- Fresu R. (2016), "L'italiano dei semicolti", in Lubello S. (a cura di), *Manuale di linguistica italiana*, de Gruyter, Berlino, pp. 328-350.
- Ito M et alii (2010), *Hanging out, Messing around, and Geeking out. Kids Living and Learning with New Media*, MIT Press, Cambridge.
- Meini A. (2014), *Aspetti della comunicazione politica su Facebook*, tesi di laurea triennale in Informatica Umanistica, Università di Pisa.
- Nasuto P. (2018), *Applicazione di strumenti di analisi linguistica a un corpus di commenti provenienti da Facebook*, tesi di laurea triennale in Informatica Umanistica, Università di Pisa.
- Pistolessi E. (2014), "Scritture digitali", in Antonelli G., Motolese M., Tomasin L. (a cura di), *Storia dell'italiano scritto. Vol. III. Italiano dell'uso*, Carocci, Roma, pp. 349-375.
- Pistolessi E. (2018), "L'italiano in rete. Usi e generi della comunicazione mediata tecnicamente", in *AggiornaMenti*, 13, pp. 17-26.
- Renzi L. (2012), *Come cambia la lingua. L'italiano in movimento*, il Mulino, Bologna.
- Tavosanis M. (2011), *L'italiano del web*, Carocci, Roma.
- Tavosanis M. (2013), "Non-standard rules: innovation you cannot find on the Italian Web", in Miola E. (a cura di), *Languages Go Web. Standard and non-standard languages on the Internet*, Edizioni dell'Orso, Alessandria.
- Tavosanis M. (2016), "Il linguaggio della comunicazione politica su Facebook", in Rita Librandi R., Piro R. (a cura di), *L'italiano della politica e la politica per l'italiano. Atti del XI Convegno ASLI Associazione per la Storia della Lingua Italiana* (Napoli, 20-22 novembre 2014), Cesati, Firenze, pp. 677-685.
- Tavosanis M. (2018), *Lingue e intelligenza artificiale*, Carocci, Roma.
- Tavosanis M. (in stampa), "L'italiano su Facebook fuori d'Italia: emigrazione recente ed emigrazione di ritorno", in *Lingua italiana d'oggi*, XIV.
- Tortorelli M. C. (2017), *Analisi linguistica di commenti ai post delle pagine Facebook dei Ministri della Repubblica italiana*, tesi di laurea magistrale in Informatica Umanistica, Università di Pisa.
- Voghera M. (2013), "A case study on the relationship between grammatical change and synchronic variation: The emergence of tipo[-N] in Italian", in Giacalone Ramat A., Mauri C. e Molinelli P. (a cura di), *Synchrony and Diachrony. A dynamic interface*, John Benjamins, Amsterdam e Philadelphia, pp. 283-311.
- Voghera M. (2017), *Dal parlato alla grammatica. Costruzione e forma dei testi spontanei*, Carocci, Roma.
- We Are Social (2018), *Digital in 2018. Essential Insights into Internet, Social Media, Mobile and Ecommerce Use around the World*. <https://www.slideshare.net/wearesocial/digital-in-2018-global-overview-86860338>.