

Legge di Heaps e ricchezza lessicale nella storia della Repubblica Italiana

Paolo Brasolin*, Federica Cominetti**, Arianna Bienati***

* Università Cattolica del Sacro Cuore, ** Università dell'Aquila, *** EURAC

Un assunto molto diffuso riguardo il discorso politico italiano è che questo abbia subito una semplificazione nel corso della storia repubblicana, e in particolare nel passaggio dalla Prima alla Seconda Repubblica e poi ancora nell'ultimo decennio. Secondo Antonelli (2007), l'abbassamento del livello stilistico del discorso politico è collegato al passaggio dal "paradigma della superiorità" a quello del "rispecchiamento", in cui i politici tendono a riflettere lo stile comunicativo degli elettori. L'ipotesi della semplificazione è solitamente basata sull'analisi di varietà idiosincratiche o su piccoli corpora non rappresentativi, e non è stata sempre confermata da tentativi comparativi più ampi (Giuliano, 2016; Ondelli, 2020).

Questo lavoro intende valutare la componente lessicale della semplificazione del discorso politico calcolando una misura di *relative vocabulary richness* per 150 personalità politiche della storia della Repubblica, per ognuna delle quali sono considerati 10 discorsi. Tali dati, che ammontano a 2.65M token di parlato trascritto, costituiscono il corpus IMPAQTS (Cominetti *et al.*, 2022). I discorsi sono bilanciati per rappresentare, per ciascun parlante, i principali tipi di monologo politico: discorso d'aula, comizio, intervento in assemblea di partito e dichiarazioni (in presenza, trasmesse in tv/radio e trasmesse sui nuovi media). Il corpus comprende 45 politici attivi nella Prima Repubblica (1946-1994), 51 nella Seconda e 54 nella Terza (dal 15 marzo 2013 a oggi) e consente dunque un'analisi diacronica.

La metodologia ricalca quella applicata da Serra-Peralta *et al.* (2021) allo studio dell'armonia classica, basato sull'interpretazione statistica della legge di Heaps come correlazione a livello inter-testuale tra la lunghezza del testo L e la dimensione del vocabolario V (Corral, 2017); la legge ha forma di *power law* con parametri K ed α , ovvero $V = KL^\alpha$. I valori di L e V per ciascun politico si calcolano sull'insieme della sua produzione presente nel corpus. I valori dei parametri K ed α empiricamente validi sul corpus si stimano mediante una regressione lineare sui logaritmi dei valori di L e V precedentemente ottenuti ($K = 10.31 \pm 1.44$, $\alpha = 0.55 \pm 0.04$, $\rho = 0.77$). Da ciò si ottiene un indice intra-corpus R di *relative vocabulary richness* calcolato come la distanza del logaritmo dei valori di V dalla linea di regressione. Per individuare la tendenza diacronica di R si effettua un'altra regressione lineare utilizzando come variabile indipendente la media delle date dei discorsi di ciascun politico. Notiamo che questo approccio evita i problemi delle misure classiche dovuti alle loro assunzioni implicite ($\alpha = 1$ per il TTR, $K = 1$ per l'indice di Herdan, $\alpha = 0.5$ per l'indice di Guiraud).

I risultati evidenziano una tendenza diacronica alla semplificazione moderata ma statisticamente significativa nel corso della storia repubblicana ($\rho = -0.35$, $p < 0.0001$), con i politici più recenti che tendono a presentare *relative vocabulary richness* inferiore. Per alcuni politici si confermano i

risultati di studi qualitativi precedenti: l'“originalità” dello stile retorico di Bettino Craxi (Desideri 1987) è confermata dall'alta ricchezza lessicale. Anche Silvio Berlusconi, figura rappresentativa dell'innovativo stile linguistico definito “gentese” (Gualdo, 2004), si presenta come *outlier* positivo. L'etichetta giornalistica di “dipietrese”, associata allo stile comunicativo di Antonio Di Pietro, concreto ma anche affezionato ai tecnicismi, trova riscontro in un basso indice di ricchezza lessicale. Dall'analisi emergono anche dati inaspettati. Ad esempio, Aldo Moro – rappresentante prototipico del criptico “politichese” (Desideri, 1998) – presenta ricchezza lessicale significativamente sotto la media. Sandro Pertini, a cui è stata attribuita una “svolta semplificatoria” (Ursini, 2007), si colloca invece sopra la media. L'analisi non conferma neppure la “cura della semplicità” attribuita ai politici del Movimento Cinque Stelle (Petrilli, 2015:114 - Giovinazzo, 2020 - Nasi 2021:74), rivelando che la ricchezza lessicale media degli esponenti del Movimento è pienamente nella media della Terza Repubblica. Si discutono inoltre i casi dei due *outlier* più significativi: Nichi Vendola verso l'alto e Lucia Borgonzoni verso il basso.

Nel contributo si riflette infine sulla differenza tra semplicità linguistica (e in particolare lessicale) e semplificazione contenutistica, sulla scia di Ondelli (2020). A una parziale confusione nell'opinione pubblica dei due livelli può forse essere attribuito il fatto che, mentre in altri contesti della comunicazione pubblica la semplificazione viene normalmente auspicata (Piemontese, 2023), la semplificazione del linguaggio politico viene spesso associata a una presunta “decadenza” della classe politica.

Paolo Brasolin

paolo.brasolin@gmail.com

Federica Cominetti

federica.cominetti@univaq.it

Arianna Bienati

arianna.bienati@eurac.edu

Bibliografia

ANTONELLI, G. (2007), *L'italiano nella società della comunicazione*, Bologna, Il Mulino.

COMINETTI, F., GREGORI, L., LOMBARDI VALLAURI, E., PANUNZI, A. (2022), *IMPAQTS: un corpus di discorsi politici italiani annotato per gli impliciti linguistici*, in Cresti, E. – Moneglia, M. (a cura di), *Corpora e Studi linguistici. Atti del LIV Congresso della Società di Linguistica Italiana*, Milano, Officinaventuno, pp. 151-164.

CORRAL, Á., FONT-CLOS, F. (2017), *Dependence of exponents on text length versus finite-size scaling for word-frequency distributions*, "Phys Rev E" 96(2), 022318.

DESIDERI, P. (1998), *Metalinguaggio e retorica dell'attenuazione nel discorso politico di Aldo Moro*, in Alfieri, G. – Cassola, A. (a cura di), *La «Lingua d'Italia». Usi pubblici e istituzionali*, Atti del XXIX Congresso della Società di Linguistica Italiana, Roma, Bulzoni, pp. 212-252.

DESIDERI, P. (1987), *Il potere della parola. Il linguaggio politico di Bettino Craxi*, Venezia, Marsilio.

GIOVINAZZO, S. (2020), *La lingua debole dei poteri forti: Matteo Salvini e Luigi Di Maio su Twitter*, Quaderni di Linguistica e Studi Orientali 6, pp. 425-442.

GIULIANO, L. C. (2016), *La parola del leader: profili di linguaggio parlamentare a confronto tra la prima e la seconda repubblica*, in R. Librandi, R. Piro (a cura di), *L'italiano della politica e la politica per l'italiano: Atti dell'XI Convegno ASLI, Associazione per la Storia della Lingua Italiana*, Firenze, Cesati, pp. 131-152.

GUALDO, R. (2004), *La comunicazione politica: novità e continuità*, in Gualdo, R. - Dell'Anna, M.V. (a cura di), *La faconda Repubblica, La lingua della politica italiana (1992-2004)*, San Cesario di Lecce, Manni, pp. 10-38.

NASI, F. (2021), *Beppe Grillo: politico o comico?*, in S. Ondelli (a cura di), *Populismi, rottamazioni e social media: sviluppi recenti della comunicazione politica in Italia*, Trieste, EUT Edizioni Università di Trieste, pp. 71-98.

ONDELLI, S. (2020), *Semplicità, semplificazione e semplicismo. Su alcuni fraintendimenti nell'analisi della lingua dei politici*, Quaderns d'Italia 25, pp. 171-188.

PETRILLI, R. (2015), *La lingua della 'nuova politica'. Beppe Grillo e Matteo Renzi a confronto*, in R. Petrilli (a cura di), *La lingua politica. Lessico e strutture argomentative*, Roma, Carocci, pp. 103-123.

PIEMONTESE, M.E. (2023) (a cura di), *Il dovere costituzionale di farsi capire*, Roma, Carocci.

SERRA-PERALTA, M. - SERRÀ, J. - CORRAL, Á. (2021), *Heaps' law and vocabulary richness in the history of classical music harmony*, EPJ Data Science 10, 40.

URSINI, F. (2007), L'italiano dei presidenti e l'italiano degli italiani, in Cortelazzo, M.A. – Tuzzi, A. (a cura di), *Messaggi dal Colle: I discorsi di fine anno dei presidenti della Repubblica*, Venezia, Marsilio, pp. 195-206.