

E SE IL DODICESIMO PRESIDENTE DELLA REPUBBLICA ITALIANA FOSSE UN COMPUTER?

1. INTRODUZIONE	
1.1. Che cos'è l'analisi dei dati testuali	3
1.2. Il discorso di fine d'anno dei Presidenti della Repubblica	9
1.3. Il discorso politico e istituzionale analizzato con metodi statistici	12
1.4. Obiettivi della ricerca	27
2. IL CORPUS	
2.1. Descrizione	29
2.2. Misure lessicometriche e trattabilità del corpus	35
3. ANALISI	
3.1. Analisi con TaLTaC2	
3.1.1. Segmenti ripetuti e Indice IS di Morrone	42
3.1.2. Analisi delle specificità	44
3.1.3. I valori specifici di ogni presidente	56
3.2. Analisi con IRaMuTeQ	
3.2.1. T-Gen sui valori specifici dei presidenti	61
3.2.2. Metodo Reinert e classi semantiche	73
4. GENERATORE DI DISCORSI	
4.1. Rank e discorsi incentrati sulle sfere valoriali	83
4.2. Rank e discorso incentrato su un topic derivato da IRaMuTeQ	110
CONCLUSIONI	123
RIFERIMENTI BIBLIOGRAFICI	128

1. INTRODUZIONE

1.1. Che cos'è l'analisi dei dati testuali

Quando si appresta ad iniziare un percorso di ricerca, la prima operazione che il ricercatore è chiamato a portare a compimento è la consultazione e l'organizzazione di un quantitativo di materiali sufficiente per poter costituire una prima base per le osservazioni future. Poi, col passare delle settimane di lavoro, i materiali di studio vengono continuamente integrati da documenti - non necessariamente solo di natura testuale - che servono a delineare, via via, la profondità che molto spesso definisce, almeno in parte, la portata della ricerca stessa. Nel caso i materiali presi in considerazione siano prodotti audiovisivi, può risultare utile trascriverli in forma testuale per una più agevole consultazione in fase di ricerca. La fase successiva consiste nello studio dei materiali organizzati e nell'elaborazione di un pensiero che, muovendo dalla domanda di ricerca e da quanto assimilato nella fase precedente, fornisca un'interpretazione su ciò che è stato preso in esame. L'insieme di queste operazioni può, dunque, essere schematizzato in un processo di acquisizione, sintesi e restituzione dei dati raccolti (Tuzzi, 2003), laddove la fase intermedia costituisce il momento in cui i semplici dati vengono trasformati in informazione - pertinente con il contesto di ricerca - tramite un ulteriore processo di significazione.

Analogamente, l'Analisi dei Dati Testuali (abbreviato: ADT), procede con fasi strutturate secondo questo schema, automatizzando il processo di archiviazione dei dati tramite calcolatori e lasciando la medesima libertà esplorativa al ricercatore (Bolasco 2010, Giuliano e La Rocca, 2008 e 2010). Tramite l'utilizzo di software appositamente concepiti, è possibile trattare i testi come dati e lavorare con un quantitativo di informazioni molto vasto, riuscendo a dominarlo in modo innegabilmente più semplice rispetto a quanto possibile se i testi presi in esame fossero studiati senza l'ausilio di tali strumenti. I due filoni principali in cui si divide l'ADT, afferenti a due approcci alla ricerca diversi, sono l'analisi qualitativa dei dati tramite l'ausilio di calcolatori, e l'analisi quantitativa (o statistica) dei dati testuali.

I CAQDAS (*Computer Assisted Qualitative Data Analysis Software*) sono programmi per l'analisi del contenuto classica, che consentono di etichettare porzioni di testo di lunghezza variabile, assegnando ad ognuna di esse una parola chiave con la quale identificarla. Si ipotizzi che, per una ricerca sull'accoglienza ricevuta dalle opere di Giacomo Puccini nel suo ultimo periodo di attività, sia necessario estrapolare da numerosi volumi di storia della musica tutte e sole le frazioni di testo in cui si tratta dell'"ultimo Puccini". Tramite i CAQDAS, si possono non soltanto etichettare i frammenti di interesse per il ricercatore, ma anche riorganizzarli in una o più tabelle a partire dalle quali sarà successivamente possibile catalogarli o esaminarli secondo ulteriori criteri di ricerca. Ciò che consentono di fare questi software, dunque, è

di archiviare frammenti di testi che il ricercatore ritiene rilevanti per il prosieguo della sua ricerca, consentendo una gestione del materiale che facilita il confronto fra i contenuti di quelle stesse porzioni considerate. L'utilizzo di questi programmi è prevalentemente diretto a ricercatori che li utilizzino *ex post*, al fine di verificare se intuizioni elaborate *ex ante* trovano riscontro nei materiali esaminati. Tuttavia, nulla vieta di farne uso *ex ante*, esplorando testi di riferimento per gli argomenti trattati nella ricerca, allo scopo di selezionare materiali utili alla formulazione di una o più ipotesi di ricerca. Alcuni esempi di CAQDAS sono ATLAS.ti (1993), NVivo (1999), MAXqda (1989) e QDA Miner (2004).

Completamente diverso è, invece, il caso di software che consentono di analizzare testi con un approccio lessicometrico (Tuzzi 2003, Lebart 2004, Giuliano e La Rocca 2008 e 2010, Bolasco 2013). In questo secondo caso, l'insieme di parole che compone un testo è analizzato da un diverso punto di vista, partendo da altri presupposti e perseguendo fini differenti. Il punto di partenza non è più il significato di un testo, ma il suo significante. Tramite il conteggio delle occorrenze, la suddivisione delle stesse in categorie semantiche, l'individuazione delle specificità di un sub corpus e altre operazioni più complesse, si cerca di risalire a un ordine delle parole, un criterio con il quale sono state selezionate dallo scrittore o, ancora, a sue caratteristiche compositive volontarie o inconsce. Il significante, in quanto "contenitore" di un significato, viene trattato con metodi statistici analogamente a quanto possono fare i fisici nell'analizzare la presenza e il comportamento di

determinati atomi nella materia. Soltanto dopo essere stato opportunamente riordinato e catalogato, se ne “estrae” il contenuto procedendo con la fase successiva dell’indagine.

Una struttura operativa di questo tipo consente di affrontare ricerche a livelli differenti. Un primo livello è quello del contenuto; se dopo aver riunito in un corpus e trattato con metodi quantitativi l’*opera omnia* di Italo Calvino, si evince che la gran parte dei suoi personaggi è fortemente connotata da caratteristiche surreali che sfociano nel mondo del fantastico, e si intende proseguire la ricerca approfondendo tale intuizione, si sta operando ad un livello in cui ciò che viene preso in considerazione è il contenuto delle storie di Calvino. Se, invece, dopo aver eseguito lo stesso trattamento dei testi, si osserva che la terminologia che rimanda al concetto di fantastico è molto più accentuata in particolari anni della vita dello scrittore, si sta svolgendo un’analisi sul linguaggio dell’autore prendendo in considerazione la variabile del tempo. In questo secondo caso, dunque, la ricerca è spostata ad un altro livello, più precisamente, quello stilistico. Questi due esempi mostrano come un approccio statistico all’analisi dei testi, partendo dallo studio dei cosiddetti “profili lessicali” dei *corpora* presi in esame, possa aprire la strada a indagini di diverso tipo a seconda dell’interesse e della specializzazione di chi fa ricerca. Non a caso, un *corpus* di testi che sia stato sottoposto a un banale pre trattamento e riportato in una tabella di frequenza, equivale a una rappresentazione diversa dello stesso testo. Nello specifico, una rappresentazione

funzionale al genere di elaborazioni cui quel testo verrà sottoposto in seguito, utile per diverse applicazioni nell'ambito della statistica testuale.

Tra i software più utilizzati per l'analisi quantitativa dei testi, figurano TaLTaC2 (Bolasco, 2010), IRaMuTeQ (Ratinaud, 2009), e una serie di pacchetti che funzionano tramite l'ambiente di sviluppo noto come "R"¹ (<https://cran.r-project.org/doc/manuals/r-release/R-intro.html>), legati al cosiddetto "Cran Project"². Il punto di contatto tra i CAQDAS e i software per l'analisi statistica dei testi è principalmente costituito dall'enorme vantaggio di cui beneficia chi intende fare ricerca con *corpus* di testi molto ampi. La quantità di informazione processabile e catalogabile con entrambi questi due tipi di software (che, si ricordi, afferiscono a due approcci alla ricerca diversi), rapportata al tempo e all'investimento di energie del ricercatore, è molto più consistente di quella che potrebbe essere presa in considerazione con i soli supporti cartacei. In altre parole, si potrebbe dire che entrambi questi gruppi di software permettono a chi fa ricerca di muoversi molto più velocemente all'interno dei due rispettivi filoni di ricerca. Vi è, tuttavia, un limite che riguarda i CAQDAS e, più in generale, tutti i metodi semiautomatici: oltre un certo numero di occorrenze, non riescono a operare efficacemente, dovendosi, quindi, arrendere. La diversa natura di questi due approcci non esclude, tuttavia, la possibilità da parte del ricercatore di farne un utilizzo combinato. Per poter sfruttare al meglio gli strumenti forniti dai

¹ Più propriamente, R è un linguaggio di programmazione pensato per l'elaborazione statistica di dati nei più disparati campi della ricerca.

² Acronimo di "Comprehensive R Archive Network", si tratta di un progetto volto alla diffusione di pacchetti specifici per i diversi ambiti scientifici.

programmi per l'analisi statistica dei testi è, per questo motivo, consigliabile avere una solida base conoscitiva nell'ambito che si intende esplorare. Infatti, per quanto le informazioni estrapolabili dalle elaborazioni sui testi di Calvino portate come esempio poco sopra, possano essere le medesime indipendentemente dal fatto che il ricercatore abbia letto o meno le opere dell'autore, soltanto nel caso in cui quest'ultimo abbia una profonda conoscenza dell'argomento (o, quanto meno, lo conosca a grandi linee), potrà sfruttare al meglio quelle informazioni. Se è vero che per trasformare un insieme di dati disomogenei in informazioni strutturate è necessario passare per un processo di significazione, le elaborazioni statistiche su un corpus non possono essere considerate, a sé stanti, vere e proprie informazioni, ma soltanto dei dati presentati in forma diversa. Il processo di significazione si ha, infatti, nel momento in cui il ricercatore mette in gioco la propria conoscenza personale sull'argomento e la sfrutta per interpretare i dati elaborati con i software statistici che utilizza. In termini pratici, in un approccio *ex post* in cui si cerchi conferma di un'ipotesi tramite l'utilizzo di programmi per l'analisi statistica dei testi, un'ottima analisi delle corrispondenze potrà fornire risultati limitati, se non adeguatamente supportata da un'altrettanto approfondita analisi a priori del contenuto dei testi esaminati. La conoscenza qualitativa, infatti, consente di aggiungere dimensioni comparative più profonde alla ricerca stabilendo, molto spesso, la differenza tra una ricerca monotona e una più articolata. Di conseguenza, dal momento che in questo ambito si lavora con i testi,

la linea che demarca il confine tra conoscenze qualitative e quantitative non può, e non deve, essere netta; quanto più le due metodologie sapranno compenetrarsi, tanto più ne gioverà la buona riuscita dell'indagine. Facendo riferimento ad approcci *“quantiqualitativi”* nell'ambito dell'ADT, Arjuna Tuzzi (2003) scrive: *“Nell'analisi testuale convivono contesti e significati di parole, di natura puramente qualitativa, con ranghi, frequenze e distribuzioni di probabilità, che sono invece quantitativi, nel rispetto della natura di entrambi”*.

Per quanto riguarda, invece, l'ampia gamma di possibilità fornita dai metodi per l'ADT, nessuno meglio di Luca Giuliano (2008) ne ha sintetizzato il contesto: *“L'analisi dei dati testuali non è più un metodo pionieristico ma è ancora - e forse lo sarà sempre - un metodo di frontiera”*. Cioè a dire, che ogni nuovo utilizzo dei metodi a disposizione, ogni invenzione di nuovi indici statistici o software crea un nuovo spettro di possibilità con cui analizzare testi già presi in considerazione, sfruttando angolazioni e gradi di profondità sempre più raffinati.

1.2. Il discorso di fine d'anno dei Presidenti della Repubblica

Il discorso di fine anno del Presidente della Repubblica è inquadrabile nella categoria del discorso istituzionale e ne costituisce un esempio particolare, con proprie caratteristiche linguistiche e di

contenuto che lo distinguono dagli altri tipi di discorsi politici o istituzionali. La consuetudine del discorso di fine anno, per quanto concerne la Repubblica Italiana, si deve a Luigi Einaudi, che già aveva istituito la consuetudine delle “Considerazioni finali” del Governatore della Banca d’Italia nel 1947 (Cortelazzo e Tuzzi, 2007). Il 31 dicembre 1949, il Capo dello Stato propalò il primo discorso augurale del Presidente della Repubblica, tramite tutte le radio collegate con il Quirinale. Benché il linguaggio utilizzato negli anni Quaranta fosse più aulico rispetto ai discorsi di alcuni presidenti successivi (quali ad esempio Pertini e Ciampi), ciò che si percepisce fin dal primo messaggio di fine anno di Einaudi è il carattere più informale e il tono meno austero dello stesso, rispetto agli altri discorsi della prima carica dello Stato. “Intimi affetti”, “fervore di sentimenti”, “mutua comprensione” e “fraterna solidarietà”, sono alcune delle locuzioni che arricchiscono il brevissimo discorso del presidente Einaudi e avvicinano, come mai prima di allora, la più alta carica dello Stato alla totalità della popolazione in ascolto.

La caratteristica principale e costitutiva di questa occasione è, infatti, l’augurio per un anno migliore del precedente. È il momento in cui si presenta un breve bilancio (in termini generali) dell’anno passato e si auspica che l’anno successivo possa essere, in ogni caso, migliore. Un’occasione di questo genere, presenta una serie di spunti e di passaggi che vengono scanditi dalle frasi del Presidente, al fine di rinsaldare il senso della collettività in ogni angolo del Paese che ascolta. Come scrivono Salvatore La Mendola e Davide Sterchele, il Presidente

della Repubblica “*deve lasciare intendere che una trasformazione è possibile e che questa sarà caratterizzata da un miglioramento, da un progresso, una re-surrezione, un risorgimento*”, indipendentemente dalle valutazioni fatte sull’anno passato (La Mendola e Sterchele, 2007). Tutto ciò ha la dimensione di un rito in cui il Presidente si fa officiante e, convogliando l’attenzione della popolazione sulla propria figura, trasmette sensazioni e immagini di sicurezza, integrità e fiducia; caratteristiche rimaste costanti nel corso degli anni, seppur con diversa intensità.

Il carattere ritualistico del discorso di fine anno del Presidente della Repubblica porta con sé la volontà di avvolgere ogni singolo italiano sotto il metaforico mantello del “*Noi*”, perpetrando la costruzione del senso comune che rende i cittadini *con-cittadini*. Salvatore La Mendola individua tre fasi attraverso cui i discorsi di fine anno dei presidenti si caricano di significati simbolici e rendono questa occasione un evento rituale, oltre che mediale: l’evocazione del *Noi* (che si ha nel momento stesso in cui la prima carica dello Stato si palesa agli italiani), l’affermazione dei valori che lo compongono e, infine, la descrizione delle caratteristiche che nel delineano l’identità (La Mendola e Sterchele, 2007; Zijno, 2007).

Anche da questa dimensione ritualistica, appare evidente che il carattere più informale di questa occasione porti a un netto discostamento linguistico, contenutistico e protocollare dei discorsi trasmessi a reti unificate la sera dell’ultimo giorno dell’anno rispetto a

quelli pronunciati in occasione dell'Assemblea plenaria del Consiglio Superiore della Magistratura o della Festa della Repubblica, così come delle numerose altre cerimonie a cui presenza e interviene il Capo dello Stato.

1.3. Il discorso politico e istituzionale analizzato con metodi statistici

La letteratura a disposizione di chi si avvicina all'analisi statistica dei dati testuali applicata alla categoria del discorso politico cresce di anno in anno, parimenti all'interesse che molti ricercatori di sfere epistemologiche diverse stanno riservando alla metodologia statistica applicata all'eloquio in questo ambito. Se il terreno più fertile (nonché il primo esplorato) per l'analisi statistica dei dati testuali si trova in Francia, è opportuno sottolineare che sono diversi gli accademici italiani che hanno implementato queste metodologie nei propri percorsi di ricerca, in particolar modo in quelli a tema politico: Bernardi e Tuzzi (2007 e 2010), Luxardo e Michieli (2012), Giuliano e Villani (2015), Cortelazzo e Tuzzi (2016). Più di altri, vi sono alcuni casi che ben esemplificano ciò che è possibile estrarre da una mole di testi a carattere politico e le conseguenze che si possono trarre in proposito. Giancarlo Luxardo e Marco Michieli (2012) si sono occupati di ADT e discorsi politici nell'ambito dei partiti europei che manifestano avversità nei confronti dei

molteplici processi di integrazione europea, i cosiddetti “partiti populistici”. Il corpus preso in esame in questa ricerca è stato prelevato dagli archivi dell’attività legislativa del Parlamento Europeo, facendo attenzione ad operare un processo di astrazione concettuale rispetto alle diverse lingue adoperate negli interventi degli europarlamentari, affinché la congruenza statistica delle informazioni ricavate fosse il più possibile intatta. Oltre a semplici dati sulla frequenza delle parole e delle locuzioni più utilizzate nel corpus (e nei relativi subcorpora), i due ricercatori si sono concentrati sull’analisi delle specificità, sugli *hapax legomena* (ovverossia le parole che compaiono nel corpus una sola volta), sulle presenze esclusive (parole o locuzioni che vengono adottate da un solo partito) e sulla ricerca dei neologismi presenti nei discorsi, suddividendoli per partito politico cui afferisce il discorso. Queste ultime due operazioni sono state possibili associando a priori, ad ogni discorso, una variabile che, in seguito, permettesse di tracciare il partito al quale apparteneva l’eurodeputato che ha pronunciato uno specifico discorso.

I risultati di questa ricerca, così costruiti, consentono di osservare alcuni comportamenti particolari. Ad esempio, sono spesso presenti riferimenti ai Paesi confinanti con lo Stato rappresentato dal parlamentare che pronuncia il discorso, con i quali, come è inevitabile, in passato quello stesso Paese ha avuto rapporti diplomatici. Più nel profondo della ricerca, Luxardo e Michieli hanno estratto dal corpus principale un subcorpus incentrato sul tema dell’immigrazione, suddividendolo secondo la variabile relativa al partito di appartenenza. In

secondo luogo, hanno creato una lista di specificità che aiutasse a individuare in quale modo il tema veniva trattato nei discorsi per delineare, infine, cinque approcci diversi al discorso sul tema dell'immigrazione, con i relativi partiti che hanno adottato ognuno di quegli approcci (Luxardo e Michieli, 2012).

Questo lavoro mostra una stratificazione di risultati possibile grazie a software specifici per l'analisi del testo. In particolare, l'estrazione di parole che individuano determinate aree che gravitano attorno a un tema definendolo in modo diverso e la loro associazione a un gruppo di partiti, è un'operazione che coi soli mezzi qualitativi sarebbe soggetta a un rapporto di altissima dipendenza da parte del ricercatore. Se è vero che la discrezionalità è un valore della ricerca, il rischio che si corre quando si lavora manualmente con il conteggio delle parole e la loro catalogazione è che, anche soltanto per disattenzione, alcuni termini possano sfuggire ed essere erroneamente esclusi da una classificazione in cui risultano opportuni, o ancora, inclusi in una lista ed esclusi da un'altra che abbia la medesima congruenza logica. Inoltre, a ciò va sommato il risparmio di tempo che richiederebbe la lettura approfondita dell'enorme mole di discorsi compresi nel corpus e la loro analisi, vocabolo per vocabolo, secondo i criteri di ricerca.

Restando nel campo politico e istituzionale, passando tuttavia a una ricerca condotta in Francia, è rilevante il lavoro di Jacques Savoy (2010) intorno al linguaggio utilizzato da Barack Obama nei suoi discorsi.

La ricerca del professore francese è strutturata come un'analisi lessicale comparata tra i discorsi che Obama ha pronunciato durante gli ultimi mesi di campagna elettorale, nel 2008, e quelli relativi ai primi mesi di mandato come 44° Presidente degli Stati Uniti d'America (Savoy, 2010). Dallo studio è emerso come i lemmi più utilizzati nella prima delle due fasi esaminate siano rimasti per lo più nelle stesse prime posizioni del *rank* definito dal lessico di frequenza, anche nei primi mesi della Presidenza dell'ex Senatore dell'Illinois. Al contrario, la grammatica di Barack Obama è mutata con il ruolo acquisito: mentre da Presidente ha utilizzato numerosi aggettivi e avverbi, da candidato alle elezioni era solito fare un più largo uso di pronomi (tra cui il celeberrimo *we*).

La comparazione tra le due parti del corpus (i discorsi da candidato del 2008 e quelli da presidente del 2009), parte da alcune statistiche di base. Il numero di parole totali (*word tokens*) è superiore nei discorsi del 2009 (303.275, contro le 294.553 dei discorsi della campagna elettorale), e il numero di parole distinte (*word types*) è, in proporzione, molto più corposo nei discorsi dell'ospite della Casa Bianca rispetto a quelli pronunciati l'anno precedente (10.251 a fronte di 7.663, circa il 30% di parole in più). Anche il numero di polirematiche composte da due sole parole è diverso nei due subcorpora, con una considerevole maggior presenza di queste locuzioni nei discorsi da Presidente; l'espressione più usata è "*health care*" in entrambe le partizioni del corpus (Savoy, 2010).

In seconda istanza, lo studio di Savoy si sofferma sul concetto di ricchezza del vocabolario. Il professore registra che esso è più ricco nei discorsi del 2009, come risulta evidente anche dall'operazione di lemmatizzazione delle parole. Si noti che la procedura che consente di estrarre il lemma delle parole presenti in un corpus può essere realizzata in forma automatica o semiautomatica grazie ad determinati software per l'ADT. I lemmi contenuti nel subcorpus del 2008 sono 7.925, mentre quelli rintracciati nel subcorpus dell'anno successivo ammontano a 10.737. L'autore della ricerca ha cura di giustificare i dati osservando che, mentre un Presidente USA deve affrontare una vasta gamma di tematiche, un candidato alle presidenziali ha la necessità incentrare i propri interventi sui temi che ritiene rilevanti per la campagna elettorale, mantenendo un linguaggio il più possibile lineare e comprensibile dall'elettorato. In aiuto di questa ipotesi, giunge un modello probabilistico basato sulla regressione lineare con cui Savoy calcola il numero di lemmi attesi.

Nel panorama italiano, si ritiene appropriato citare una pubblicazione a sei mani in cui Sergio Bolasco, Nora Galli de' Paratesi e Luca Giuliano (2006) hanno analizzato dal punto di vista statistico e linguistico i discorsi di Silvio Berlusconi, ex Presidente del Consiglio dei Ministri in Italia. Partendo, come quasi sempre si è usi in questi casi, da un'analisi dei lemmi più frequenti, i tre studiosi italiani hanno classificato il sovrautilizzo e il sottoutilizzo di svariati termini del vocabolario

berlusconiano, procedendo parimenti con una suddivisione degli stessi in base alle variabili “anno”, “periodo” (inteso come “periodo politico”) e “luogo”. Alcuni di questi risultati mostrano che, se l’ossessione dell’imprenditore milanese per il comunismo ha caratterizzato la sua intera carriera politica, altri termini (e le rispettive aree di riferimento) hanno vissuto fasi altalenanti nel corso degli anni. Mentre a metà degli anni Novanta, durante il suo primo governo, i discorsi di Berlusconi vertevano sulle alleanze e sulla politica interna (“opposizioni”, “governo”, “dialogo”, “senatori”, “deputati”, “manovra”, “Polo” e “Bossi” le parole più significative), due lustri dopo, con il rafforzarsi delle relazioni all’interno della Comunità Europea - accentuate dal passaggio alla moneta unica - , il focus dei suoi discorsi si spostava sulle relazioni internazionali, tendenza testimoniata dal frequente utilizzo di espressioni quali “internazionale”, “europeo”, “Unione Europea”, “Onu”, “politica estera”, “guerra fredda”, “Israele”, “Francia” e “Iraq”.

I tre autori si soffermano, in secondo luogo, sulla parola con la quale più di ogni altra - se si osserva la questione dal mero punto di vista ideologico - Berlusconi ha infarcito i propri discorsi e, più in generale, la propria campagna elettorale e identità politica: il concetto di “libertà”. Dal punto di vista numerico, infatti, “libertà” risulta essere la quinta parola più frequente nell’eloquio del leader del centrodestra, dietro a “governo”, “Paese”, “Stato” e “Italia”; tuttavia, mentre i primi quattro termini riconducono a un linguaggio istituzionale neutro, solo nella parola

“libertà” è possibile rintracciare la personalità e una parte dell’ideologia con la quale l’ex Cavaliere del Lavoro ha costruito il proprio partito.

Nell’ambito di successive analisi su altri temi ricorrenti nella politica italiana, alcuni dei quali caratterizzanti il discorso politico di Berlusconi (lavoro, famiglia, impresa, giustizia, sicurezza e pace), Sergio Bolasco, Nora Galli e Luca Giuliano hanno esaminato la presenza e l’utilizzo del “noi” all’interno del corpus. Il pronome di prima persona plurale, che come ricordano gli autori, in italiano può essere frequentemente omesso, è molto spesso presente negli interventi di Silvio Berlusconi (3,6 occorrenze ogni 1000 parole, contro le 1,2 occorrenze dell’italiano standard, un campione di testi utilizzato dai tre autori come parametro di confronto nella ricerca). L’utilizzo del pronome e tutte le locuzioni e i messaggi che vi ruotano attorno hanno, tuttavia, una sfumatura di significato diversa rispetto a quanto avviene nei discorsi dei Presidenti della Repubblica. Mentre il “noi” più istituzionale utilizzato dai Capi dello Stato succedutisi nel corso degli anni, aveva la funzione di creare una comunità includendo tutti i cittadini italiani nella sfera d’interesse della figura del Presidente, il “noi” utilizzato da Berlusconi è, per lo più, riferito agli elettori di Forza Italia o della coalizione di centrodestra (Bolasco, Galli de’ Paratesi, Giuliano 2006). L’intento, dunque, non è di rivolgersi a tutti i cittadini considerandoli nello stesso modo a prescindere dal colore politico di appartenenza ma, piuttosto, di rimarcare la differenza tra il *noi* che vota o milita nel centrodestra e il *loro* in cui vengono relegati tutti gli altri. Pur essendo

anche questa seconda opzione una modalità di identificazione attorno a una collettività, è evidente come nel caso degli interventi dei Presidenti della Repubblica il discrimine per far parte del gruppo sia costituito dall'essere parte di uno stesso Stato - carattere che comprende l'intero universo di riferimento cui i discorsi sono rivolti -, mentre nel caso di Berlusconi, il punto attorno al quale gravita l'appartenenza o meno al gruppo è l'adesione all'ideologia di cui egli si fa promotore. Se il noi dei Presidenti (la cui presenza oggettiva nel relativo corpus è meno invasiva rispetto a quella rilevata nei discorsi di Berlusconi) ha carattere esortativo e funge da "rassicurante coperta" per i cittadini all'ascolto, quello dell'ex Presidente del Consiglio ha valore descrittivo, spesso accompagnato da verbi quali *siamo*, *vogliamo* e *dobbiamo*. A questo proposito, è tuttavia opportuno ricordare che gran parte delle differenze riscontrate nell'utilizzo del *noi* in questi due corpora è attribuibile alla differente funzione dei discorsi pronunciati; da un lato i ritualistici interventi di un'istituzione, dall'altro le allocuzioni di un leader politico, volte a raccogliere il consenso dei cittadini.

Gli studi sopracitati utilizzano diverse metodologie per l'analisi quantitativa dei testi al fine di estrapolare o verificare osservazioni sulle strategie e gli scenari politici presi come universo di riferimento per le rispettive indagini. Due ulteriori lavori, che aderiscono in modo ancora più opportuno a questa ricerca rispetto agli esempi precedenti per via del soggetto esaminato, sono "Parole lette con misura" di Lorenzo

Bernardi e Arjuna Tuzzi - contenuto nella raccolta di pubblicazioni sui Presidenti della Repubblica, “Messaggi dal Colle”, curata da Michele A. Cortelazzo e Arjuna Tuzzi (2007) - e “L’autografo del Presidente della Repubblica”, dei medesimi autori, pubblicato nell’edizione 2010 del *Journées internationales d’Analyse statistique des Données Textuelles* (Bernardi e Tuzzi, 2010).

L’obiettivo del lavoro pubblicato in “Messaggi dal colle”, è quello di delineare i profili lessicali dei Presidenti della Repubblica, prendendo come riferimento tutti e soli i discorsi di fine d’anno pronunciati da Einaudi, Ciampi e tutti i presidenti succedutisi nel mezzo. Dopo un approccio esplorativo e una prima analisi multivariata dei dati (nel caso specifico, delle parole), Bernardi e Tuzzi (2007) hanno evidenziato punti di continuità e punti di rottura nel lessico utilizzato nei discorsi, tenendo presente anche le osservazioni dei colleghi presenti nello stesso gruppo interdisciplinare, che hanno contribuito ad affrontare il tema dei discorsi di fine d’anno dei Presidenti della Repubblica da prospettive di ricerca diverse. In primo luogo, è stato registrato un progressivo aumento nel corso degli anni della prolissità dei discorsi, seppur con qualche eccezione; dai brevissimi interventi di Einaudi (201 occorrenze in media) si è passati, negli anni Novanta, ai lunghissimi panegirici di Scalfaro (3525 occorrenze in media). Ciampi, nel settennato successivo a Scalfaro, ridimensionerà la media delle parole spese, riportandola a livelli più vicini a Cossiga e Pertini. La tendenza ad allungare sempre più i discorsi nel tempo, è riscontrabile anche all’interno dei subcorpora

relativi ai singoli presidenti, dai quali emerge che i “primi discorsi” di tutti i presidenti hanno dimensioni più contenute rispetto alla maggior parte degli altri discorsi dei rispettivi presidenti.

Tuttavia, è tramite l'utilizzo dell'analisi delle corrispondenze che Bernardi e Tuzzi hanno individuato i profili di frequenza dei presidenti, ponendoli sugli assi cartesiani per compararne le similarità lessicali. Questo processo è stato portato avanti a seguito di tre diverse elaborazioni del corpus (volte a individuare sfumature diverse del linguaggio): l'insieme totale delle parole, le espressioni ricorrenti e, infine, i termini lemmatizzati. Tali analisi evidenziano come tra tutti i Presidenti ce ne siano due in particolare che si discostano in maniera evidente dagli altri, oltrech  tra di loro. Pertini e Scalfaro sostano, infatti, in aree degli assi cartesiani molto distanti dal gruppo che costituisce tutti gli altri Presidenti. Con osservazioni ancora pi  specifiche, i due statistici dell'Universit  di Padova hanno osservato che una cospicua fetta dei 57 discorsi dei presidenti si raggruppano sugli assi cartesiani in modo piuttosto ordinato e coerente rispetto alla variabile “Presidente”. Infatti, tutti e sette i discorsi di Scalfaro si raccolgono attorno a una stessa area, cos  come quelli di Pertini, Cossiga e Ciampi. Ci  che consegue da questo assetto   che *“... sembra esistere pertanto una cifra stilistica personale che travalica il tempo, la fase politica, gli eventi dell'anno;   piuttosto il registro dei sentimenti, della storia individuale, della rete di relazioni e, forse, dell'appartenenza a vicende, itinerari politici, attenzioni sociali, condivisa con particolari realt  del paese, non necessariamente*

comune o prossima a quelle degli altri presidenti, che determina la sfera e la qualità della comunicazione di ciascuno” (Bernardi, Tuzzi 2007).

La successiva fase dello studio, che ne costituisce la porzione più corposa ed elaborata, si sofferma sui profili lessicali dei Presidenti. In particolare, i due docenti notano come gli aggettivi utilizzati da Pertini afferiscano principalmente a due universi espressivi: uno marcatamente drammatico e uno più legato alla condizione sociale o alla dislocazione geografica della popolazione in ascolto. Mentre il linguaggio di Scalfaro unisce l’attenzione per le istituzioni a un *pathos* linguistico che permea tutti i suoi lunghi interventi, quello di Cossiga ha uno sguardo più internazionale e, talvolta, più introspettivo e preoccupato nei confronti di problematiche interne al Paese. In Ciampi, risulta chiara la volontà di cercare il dialogo e mediare tra fazioni che hanno portato il dibattito politico a livelli di esasperazione senza precedenti, forse come strascico del terremoto causato dall’inchiesta “Mani Pulite” del 1992 che ha sconquassato il sistema politico italiano. Tema, quest’ultimo, che riecheggia anche nei discorsi di Scalfaro (in carica come Presidente della Repubblica proprio in quegli anni), seppur con una diversa accezione: non già con la volontà di avvicinare le parti politiche riassestate dopo l’inchiesta, ma con l’auspicio di fare chiarezza e di ripristinare o, più probabilmente, instaurare la trasparenza nel sistema politico del Paese.

La profilazione dei Presidenti della Repubblica costruita da Lorenzo Bernardi e Arjuna Tuzzi, prosegue con l’analisi delle specificità, le espressioni tipiche tramite le quali sarebbe possibile individuare i

singoli autori dei discorsi. Con l'ausilio di modelli probabilistici, è infatti possibile stimare il sovrautilizzo o il sottoutilizzo di parole o espressioni contenute nei subcorpora rispetto al loro utilizzo complessivo all'interno del corpus totale. Così, dalle espressioni immaginifiche ed eleganti di Einaudi, si passa ad una maggior concretezza e attenzione alla comunità nel linguaggio di Gronchi, caratteristica che sublima nell'interesse per le condizioni sociali degli italiani manifestato nelle numerose espressioni utilizzate esclusivamente da Saragat. Del tutto *sui generis* è il linguaggio di Sandro Pertini che, unitamente alle espressioni amichevoli con cui era solito rivolgersi ai cittadini in ascolto, arricchiva i propri interventi con aneddoti relativi al proprio vissuto e alla propria esperienza come partigiano, al fine di trasmettere - soprattutto ai giovani che aveva tanto a cuore - i suoi insegnamenti. Alcune espressioni ricorrono più volte nel corso del settennato pertiniano, rendendo il relativo subcorpus, per certi versi, più simile a sé stesso di quanto tutti gli altri subcorpora siano simili a loro stessi; questa caratteristica è, probabilmente, anche legata al fatto che il presidente socialista amava recitare il proprio augurio di fine d'anno senza l'ausilio del canovaccio scritto. Per tutti questi motivi, si può affermare con certezza che il subcorpus di Pertini è quello che, più di ogni altro, si distingue da tutti i discorsi presidenziali esaminati in questo contesto. Mentre dagli interventi di Scalfaro si evince un senso di gratitudine per la possibilità di esprimersi in qualità di Presidente della Repubblica, in Ciampi emerge, una volta di più, la volontà di esortare la cittadinanza a riporre la propria fiducia nelle istituzioni per sanare le ferite

che hanno martoriato la politica italiana nel decennio precedente. L'intento del presidente livornese, di cui si ha un chiaro segno anche nell'analisi delle specificità, era di utilizzare il concetto di Patria come collante sociale, facendosene convinto promotore al fine di superare fratture politiche tra impianti ideologici che all'alba del nuovo millennio apparivano ormai sbriciolati (Grimaldi, 2012). Congiuntamente alla volontà di sensibilizzare i cittadini italiani rispetto al concetto di Patria, nei discorsi di Carlo Azeglio Ciampi si riscontra la particolare attenzione all'istituzione della famiglia, tanto che nel proprio augurio per l'anno successivo integra anche quello della moglie.

In conclusione al proprio lavoro, con uno sguardo più ad ampio spettro, Lorenzo Bernardi e Arjuna Tuzzi osservano come le specificità di ogni Presidente siano sostanzialmente incasellate in una struttura comune a tutti, composta da saluti iniziali, riflessioni sull'anno passato, auspici per l'anno venturo e saluti finali. Nelle conclusioni parziali si sottolinea come ciò che risalta dall'analisi statistica dei discorsi di fine d'anno dei presidenti, è la mancanza di valori (tradotti, in questo contesto, come parole, espressioni, ma anche prese di posizione) tanto al di fuori dal coro da essere considerati *outliers*, come se il ruolo e l'occasione costituissero un ufficioso e tacito confine entro il quale procedere.

Nel 2010, i due docenti dell'Università di Padova sono tornati ad analizzare il corpus dei discorsi di fine d'anno, integrandolo con i primi

tre discorsi di Giorgio Napolitano, per tracciare quello che hanno ironicamente battezzato “Autografo del Presidente”. Lo scopo di questa seconda ricerca è di costruire un “discorso archetipico” di ogni Presidente, riservando particolare attenzione ai temi che emergono nei discorsi e al modo in cui essi vengono affrontati. Per cercare di ottenere questo risultato, Bernardi e Tuzzi hanno operato un’ulteriore suddivisione dei subcorpus in frammenti che identificassero, in una o più frasi, un pensiero, un intento o una modalità comunicativa. Il metodo con cui tali frammenti sono stati successivamente riorganizzati secondo una logica di significatività all’interno del corpus, sfrutta indici statistici di diversa natura. Dopo aver individuato i poliformi più pregni di significato tramite l’indice *IS* di Morrone, i due statistici hanno sfruttato il modello ipergeometrico (*P-Value*) per verificare il sovrautilizzo di determinate espressioni nei subcorpora dei rispettivi Presidenti. Nel farlo, hanno avuto cura di isolare (ed escludere dalle fasi seguenti dell’analisi) quei frammenti che, oltre a risultare sovrautilizzati nel subcorpus di un Presidente, erano esclusivi di un singolo discorso; la *ratio* che sottende a questa scelta si individua nel fatto che una locuzione che compaia soltanto in un discorso su sette non può considerarsi specificatamente costitutiva e distintiva del relativo subcorpus, quanto piuttosto del singolo intervento.

Costruita una lista di espressioni significative e peculiari che rappresentano lo stile di ogni Presidente, i due docenti hanno proceduto all’estrazione dei frammenti più significativi contenenti quelle espressioni,

sfruttando una versione normalizzata dell'indice statistico TFIDF (*Term Frequency Inverse Document Frequency*), sovente utilizzato nel campo del *text mining* (Bernardi, Tuzzi 2010). Grazie a questo strumento, è stato possibile avere una lista di pensieri caratteristici di ogni Presidente; in più, ad ogni frammento era stato associato in precedenza un numero progressivo che lo collocava all'interno del discorso, dimodoché fosse tracciabile ed eventualmente inseribile in un punto adeguato dell'archetipo in fase di costruzione. A questo punto, dopo aver stabilito alcuni parametri entro cui il discorso si sarebbe dovuto mantenere, i due statistici hanno composto il discorso "tipo" dei Presidenti, con particolare attenzione a Pertini e Ciampi.

Nelle conclusioni a questo lavoro, si sottolinea come persistano problemi riguardanti la coerenza in fase di selezione e riassetto dei frammenti, criticità registrata non tanto in fase di apertura o chiusura del discorso, ma soprattutto all'interno del cuore del cosiddetto "autografo" (Bernardi, Tuzzi 2010). Sembra, infatti, mancare la possibilità di amalgamare il discorso con metodi automatici in modo naturale e coerente come il filo logico che una persona tesse attorno agli argomenti che tocca mentre comunica con qualcuno. Se esiste un modo per superare questo limite, forse, le sue radici possono essere rintracciate nel *machine learning*.

1.4. Obiettivi della ricerca

Lo scopo di questa ricerca, che vorrebbe ispirarsi a quanto pubblicato nel 2010 da Lorenzo Bernardi e Arjuna Tuzzi sul *Journées internationales d'Analyse statistique des Données Textuelles*, seppur con un obiettivo differente, è di comporre tre discorsi di fine d'anno "tipo", ognuno di essi definito, dal punto di vista contenutistico, da una sfera valoriale diversa dagli altri, prendendo spunto dai valori individuati da Selena Grimaldi e Gianni Riccamboni (2007) -, selezionando frasi provenienti dai 69 discorsi dei Presidenti della Repubblica Italiana fino ad oggi pronunciati. Fin da ora si specifica che, mentre il lavoro di Bernardi e Tuzzi mirava a costruire un discorso tipico per ogni Presidente, qui si intende procedere tenendo come discriminante i valori espressi nei messaggi di fine anno. Per poter giungere a questo risultato, si ritiene necessario studiare il tema dei discorsi di fine anno da un punto di vista interdisciplinare (linguistico, sociologico, politico e mediatico) e utilizzando strumenti di analisi quantitativa dei testi uniti a osservazioni di tipo qualitativo. In una prima fase della ricerca si intende utilizzare la prospettiva del *distant reading*, al fine prendere le misure e organizzare il *corpus* testuale preso in esame secondo una struttura utile allo scopo della ricerca.

Dopo aver verificato la trattabilità dell'insieme di testi tramite opportuni indici statistici (*Type Token Ratio* e percentuale di *hapax legomena*) si intende procedere analizzando il *bag-of-words* allo scopo di

individuare i caratteri specifici di ogni Presidente, mediante l'utilizzo di variabili definite a priori al momento della definizione di una dimensione comparativa. Lo strumento utilizzato in questa fase è il *software* TaLTaC2. Dopo aver eseguito la ricognizione preliminare e le prime analisi organizzative del *corpus*, si intende proseguire la fase di ricerca computazionale utilizzando il programma IRaMuTeQ, eseguendo, tra le altre cose, una *cluster analysis*. Infine, a partire dalle elaborazioni ottenute, si intende generare i discorsi tramite l'estrazione di un TFIDF testuale utilizzando nuovamente il *software* TaLTaC2.

2. IL CORPUS

2.1. Descrizione

Il corpus preso in esame per la ricerca è composto dai 69 discorsi pronunciati - in occasione del messaggio di fine anno - dagli 11 Presidenti della Repubblica Italiana che si sono susseguiti a partire dal 1949 al 2017. Prendendo in considerazione la variabile relativa ad ogni singolo Presidente, si contano 7 diversi discorsi per ogni presidente, salvo alcuni casi particolari: Luigi Einaudi, che istituì questa consuetudine nel 1949 e pronunciò 6 discorsi, Antonio Segni che si dimise durante il terzo anno di mandato, con 2 discorsi, Giorgio Napolitano che, a seguito della rielezione nel 2013 rimase in carica per un totale di 9 anni, contando altrettanti discorsi, e Sergio Mattarella, attualmente in carica, con 3 messaggi di fine anno firmati finora.

La prima serie di operazioni sul corpus, come sempre è buona prassi nell'analisi statistica dei testi, è consistita nella pulizia del corpus e nella normalizzazione del testo. Il cosiddetto "*tokenization*" consente di individuare i caratteri che il software dovrà considerare come separatori, ovvero come segni che marcano la separazione tra parole. Il separatore per eccellenza è lo spazio vuoto (digitato tramite la barra spaziatrice), ma appartengono a questa categoria anche tutti i segni di interpunzione. Il

corpus di partenza era già sufficientemente pulito, salvo per il carattere “-“ che, qualora non fosse stato inserito manualmente nella categoria dei separatori mediante la procedura di *tokenization*, sarebbe risultato tra le forme grafiche dell’elenco di frequenza. Dal momento che in questa analisi quantitativa sui testi, l’unità statistica è la parola, è importante assicurarsi di segnalare correttamente questi segni al programma.

In seguito, si è proceduto con una normalizzazione, che associa ad ogni parola corrispondente a una specifica categoria, un’etichetta identificativa, radunando all’interno di quella categoria tutte le parole che vi appartengono. La sfera di competenza di ogni categoria può essere selezionata tra quelle preimpostate nel software da chi svolge la ricerca e, in questo caso, si è limitata a “Nomi propri”, “Celebrità”, “Toponimi”, “Giornali”, “Titoli”, “Sigle” e “Società”. Nel caso in cui si fosse deciso di selezionare anche “Locuzioni grammaticali” e “Gruppi nominali e Polirematiche”, si sarebbe trattato di una Normalizzazione cosiddetta “pesante”, ma in questo caso ci si è limitati alla sua versione detta “leggera”.

In seconda istanza, il corpus è stato suddiviso in subcorpora secondo due variabili: la prima (chiamata “VPres”) individua tutti i discorsi di ogni Presidente, la seconda (“VDisc”) raggruppa tutti i discorsi suddivisi per anno di mandato. Ad esempio, VPres01 fa riferimento al primo discorso di ognuno degli 11 Presidenti, mentre VPres08 e VPres09 contengono i dati relativi all’ottavo e al nono discorso di Giorgio Napolitano, perché unico Capo dello Stato in carica per più di sette anni.

Per quanto riguarda una parte delle ricerche dedicate alle osservazioni generali, è opportuno specificare che si è utilizzata una versione del corpus lemmatizzata in modo semi-automatico. Con l'ausilio della funzione di "tagging grammaticale" integrata nel software TaLTaC2 (Bolasco *et al.*, 2009), si è operata una lemmatizzazione automatica che ha estratto i lemmi delle parole che, per l'algoritmo del programma stesso, non presentavano alcuna ambiguità. Tutte le voci che non sono state processate in modo automatico, sono state successivamente disambiguate manualmente.

L'operazione di lemmatizzazione, non sempre di facile esecuzione, consente di soffermarsi su alcuni aspetti del linguaggio non immediatamente evidenti in un corpus composto di forme flesse, aprendo il margine per alcuni ragionamenti specifici. Se, ad esempio, si è interessati maggiormente a uno studio di tipo semantico, l'analisi dei lemmi è particolarmente efficiente perché permette di riunire in una sola unità statistica tutte le parole che nel corpus originario possono essere ricondotte al lemma cui si è interessati. Se si volesse studiare la ricchezza semantica di un corpus, sarebbe utile accorpate tra loro voci singolari e plurali, o maschili e femminili della stessa parola, perché semanticamente equivalenti. In una ricerca di questo tipo che, ad esempio, prenda come riferimento la parola "amico", infatti, non è rilevante che una parola sia usata per indicare un solo individuo anziché più individui o un individuo di genere maschile anziché uno di genere femminile, perché il *focus* della ricerca riguarda il significato di amico,

non il suo significante. Ma la lemmatizzazione non è efficace solo in ambito semantico; essa è utile anche in campo lessicale quando, ad esempio, si vuole sapere qual è il lessico in comune fra due testi. Si consideri il caso di due *corpora* in cui si desidera verificare la frequenza di utilizzo di un verbo: in una tabella di frequenza, forme flesse diverse dello stesso verbo costituirebbero record diversi, non permettendo un confronto diretto tra la presenza di quel verbo nei due testi. In una versione lemmatizzata del corpus, invece, un unico record rappresenta il verbo oggetto dell'analisi e le occorrenze dello stesso sarebbero riportate in corrispondenza della variabile che individua i due testi. In questi casi, dunque, la lemmatizzazione sarebbe un'operazione pertinente e utile allo scopo della ricerca. Tuttavia, in altri tipi di analisi, il processo di lemmatizzazione può non risultare la scelta più appropriata, perché annulla alcune differenze linguistiche che potrebbero essere rilevanti ai fini della ricerca. Ad esempio, il fatto che un subcorpus sia più ricco di espressioni al plurale anziché al singolare come altri subcorpora, non può essere rilevato con analisi all'interno di un corpus lemmatizzato, dal momento che questa operazione ha lo scopo di riportare al ogni forma flessa al lemma originario, rendendo in tal modo indistinguibile una forma singolare da una plurale.

Un aspetto fondamentale del processo di lemmatizzazione automatica riguarda il fatto che una parola presente nel testo potrebbe ricondurre a più lemmi. Se nei discorsi dei Presidenti pronunciati dalla seconda metà del Novecento ad oggi è molto probabile che la parola

“Parlamento” faccia riferimento al sostantivo maschile singolare che indica il luogo istituzionale che riunisce Camera e Senato della Repubblica e non la prima persona singolare del verbo “parlamentare”, più complicato potrebbe essere, per un programma, individuare il lemma di parole come “fatto”, “scelta”, “ancora” e molte altre. Per questo motivo, poiché - a differenza dell’essere umano - un algoritmo che non implichi un approccio al *machine learning*, non è in grado di comprendere sempre dal contesto quale sia il lemma originario di parole ambigue, è opportuno rifinire il lavoro automatico con un controllo manuale.

Una delle operazioni eseguite sul corpus non lemmatizzato, è quella conosciuta come analisi delle specificità che, grazie alla variabile “Presidente”, individua le espressioni sovrautilizzate (specificità positive) e sottoutilizzate (specificità negative) per ogni Capo dello Stato. Questa funzione sfrutta la distribuzione ipergeometrica, uno specifico modello probabilistico utilizzato in diverse branche della statistica. Sfruttando il *p-value* fornito da questo modello ipergeometrico, è si lavora su gruppi di testi creati *ad hoc* da chi svolge la ricerca, in questo caso sui gruppi di testi individuati dalle variabil “Presidente” e “Discorso”. La soglia del *p-value* oltre la quale una parola non può essere considerata specifica, può essere impostata manualmente ma, in questo caso, si è deciso di lasciare quella proposta dal sistema, vale a dire pari al 2,5%. Per comprendere al meglio il senso di questo valore, si ritiene utile proporre un esempio. Qualora la soglia fosse stata fissata manualmente al 5%, il

programma avrebbe probabilmente considerato specifiche parole che con la soglia pre-impostata vengono catalogate come discretamente comuni a più presidenti. Viceversa, una soglia più bassa, ad esempio dell'1%, avrebbe escluso dalla selezione alcune parole che con le impostazioni di base vengono identificate come specifiche. Allora, il risultato è, ancora una volta, il frutto delle scelte di chi fa ricerca, che delimita i confini entro cui opera un modello statistico e un *software*, a seconda delle proprie supposizioni nel merito della questione.

Il frutto di questa elaborazione porta a un elenco di parole cui sono associate le occorrenze all'interno del relativo subcorpus, le occorrenze totali e un valore che ne determina il sovrautilizzo e il sottoutilizzo nel testo. Nel caso del corpus oggetto di questa ricerca, si consideri l'esempio della parola P2, normalizzata in corso d'opera a "pidue". Essa compare 11 volte all'interno dell'intero corpus ed emerge soltanto nei discorsi di fine d'anno di Sandro Pertini. Più in profondità, si può osservare che è presente 10 volte nel suo quarto discorso e 1 volta nel suo settimo e ultimo; di conseguenza, essa è statisticamente considerata una parola specifica del Presidente Pertini. Tuttavia, questa affermazione è solo parzialmente vera e per dimostrarlo, occorre scendere ulteriormente in profondità, prendendo in considerazione la variabile "Discorso". Dal momento che 10 volte su 11 la parola è inserita nello stesso discorso, risulta ragionevole pensare che ne sia stato fatto un sovrautilizzo nel quarto discorso rispetto a tutti gli altri discorsi del medesimo presidente. Infatti, se si utilizza lo stesso modello

probabilistico considerando il subcorpus relativo a Pertini come il corpus totale da analizzare (soltanto in questo specifico frangente) e i subcorpora individuati dalla variabile “Discorso” come quelli all’interno dei quali si desidera individuare il sovrautilizzo, si ha conferma del fatto che “P2” è un termine strettamente legato al quarto discorso preso in esame (VDisc04). Tornando, dunque, all’affermazione precedente, è più corretto affermare che la parola “P2” *non* è specifica di Pertini - restando comunque *esclusiva* di questo presidente -, ma è specifica del suo quarto discorso.

Da notare, che un ragionamento analogo era già stato proposto da Lorenzo Bernardi e Arjuna Tuzzi (2010), al fine di scartare le forme specifiche dei singoli discorsi dall’ammontare delle espressioni sovrautilizzate da ogni Presidente che sarebbero state utili per la composizione finale dell’“autografo dei Presidenti”.

2.2. Misure lessicometriche e trattabilità del corpus

Affinché un corpus di testi possa essere analizzato con strumenti statistici, è necessario svolgere alcune misurazioni lessicometriche che ne attestino l’effettiva trattabilità dal punto di vista quantitativo. A tal fine, sono stati fissati dei valori entro cui il corpus deve rientrare e nel caso in

cui questi limiti non siano rispettati, il valore delle rilevazioni effettuate non può considerarsi scientificamente significativo. A latere, è sempre opportuno effettuare alcune altre semplici misurazioni che, oltre che preparare il terreno per i successivi calcoli, aiutano il ricercatore a conoscere meglio il corpus dal punto di vista numerico e a prendere confidenza con i confini entro cui lavorare. A questo proposito, si ritiene opportuno sottolineare un aspetto cruciale per quanto riguarda questo tipo di analisi: operare con metodi statistici applicati ai dati testuali, nella maggior parte dei casi, significa utilizzare un approccio diretto al *bag-of-words*, in cui tutti i metodi ideati e sfruttati per esplorare il corpus si basano su liste di forme grafiche composte secondo uno specifico criterio, tabelle di contingenza e catalogazione di parole (o gruppi di parole) realizzate a partire da una matrice di dati avente come unità statistica le parole stesse che compongono il corpus. Per questo motivo, avere a priori una dimensione del “territorio” di ricerca risulta determinante per comprendere quali e quanti metodi sia più opportuno sfruttare in ogni ricerca.

La prima fase di verifica del corpus impone di individuare il numero di *word tokens* e *word types* all'interno dei testi analizzati, rispettivamente il numero di parole complessivo (N) e il numero di parole diverse tra loro, vale a dire il vocabolario del corpus (V). Nel caso del corpus preso in esame, si registra che i *word tokens* (N) sono 116.853, mentre i *word types* (V) sono 11.356. Questi due valori vengono utilizzati nel calcolo del primo indice numerico utile a stabilire se il corpus è

trattabile dal punto di vista statistico, il *Type Token Ratio*. Se questo valore è inferiore al 20% (cioè se il numero di parole diverse non è superiore a 1/5 delle parole totali), il corpus può essere liberamente sottoposto alle analisi seguenti, altrimenti potrebbero pervenire problemi di attendibilità delle misurazioni. Nel caso specifico di questa ricerca, il TTR ha un valore pari a 9,72%, pienamente entro i parametri stabiliti.

Il passo successivo di questa esplorazione preliminare del corpus consiste nell'individuare le alte, medie e basse frequenze del corpus; le alte frequenze sono le forme grafiche - elencate secondo l'ordine di frequenza decrescente - comprese tra la più frequente e quella che ha un numero di occorrenze uguale alla successiva. Le basse frequenze, invece, vengono contate a partire dalla parola meno frequente, fino alla parola il cui numero di occorrenze successivo dista più di 1 valore dalla parola con numero di occorrenze più basso. Le medie frequenze, sono tutte quelle comprese tra le alte e le basse frequenze. Nel caso del corpus preso in esame, si registrano i seguenti valori:

Alte frequenze:	da 1 a 25	=	25
Medie frequenze:	da 26 a 236	=	211
Basse frequenze:	da 237 a 11.356	=	11.120

Le parole incluse nelle alte frequenze sono pressoché tutte grammaticali, ovverosia parole vuote di contenuto il cui scopo è quello

di permettere l'articolazione di un pensiero all'interno del discorso, tranne per il verbo "essere" flesso alla forma "sono". Per quanto riguarda le medie frequenze, spiccano, tra le altre, le parole Italia, giovani, mondo, popoli, vita, lavoro, politica, Europa, libertà, Stato e cittadini.

Infine, tra le basse frequenze - che comprendono anche tutti gli *hapax* - figurano i termini disoccupazione, violenza, coscienza, famiglia, lavoratori, economia, cultura e benessere.

Con l'ausilio di questi dati è possibile calcolare due ulteriori percentuali che danno contezza dell'estensione e della profondità del corpus, cioè a dire i tassi di copertura del corpus e del vocabolario, calcolati sia per le alte, sia per le basse frequenze.

Tasso di copertura del vocabolario per le alte frequenze:

$$\text{HF TCV} = (\text{HF}/V) = 25/11.356 = 0,002\%$$

Tasso di copertura del corpus per le alte frequenze:

$$\text{HF TCC} = [(\text{Somatoria HF})/N] = 36784/116853 = 31,5\%$$

Tasso di copertura del vocabolario per le basse frequenze:

$$\text{LF TCV} = (\text{LF}/V) = 11.120/11.356 = 97,9\%$$

Tasso di copertura del corpus per le basse frequenze:

$$\text{LF TCC} = [(\text{Somatoria LF})/N] = 46.741/116853 = 40,0\%$$

I dati che precedono, se affiancati, rivelano quanto segue: lo 0,002% delle parole diverse (*word types*) contenute nel vocabolario copre il 31,5% delle parole totali utilizzate nel corpus. Analogamente, il 98% del vocabolario, che costituisce l'insieme delle basse frequenze, copre il 40% delle parole totali scritte nei testi analizzati. Di conseguenza, una quantità infinitamente piccola di parole contenute nel vocabolario (lo 0,002%), fa più del 75% dello stesso lavoro che fa il 97% di quello stesso vocabolario.

Il secondo indice che stabilisce la trattabilità o meno di un corpus riguarda la ricchezza lessicale dello stesso e utilizza il concetto di *hapax legomena*, le forme grafiche che compaiono nel corpus una sola volta, cioè la cui occorrenza è uguale a 1. La percentuale di *hapax* è facilmente individuabile con il rapporto tra questi e il numero totale di termini compresi nel vocabolario. Nel caso in cui questa frazione porti a un risultato inferiore al 50%, il corpus è pienamente trattabile; in caso contrario, più essa si allontana dalla soglia del 50%, meno ricca e precisa potrà essere l'analisi. Questo perché una percentuale più alta indica un numero maggiore di parole che compaiono una sola volta e, di conseguenza, meno parole accorpabili tra loro sotto una determinata variabile o meno parole significative per un'eventuale *topic detection*. La percentuale di *hapax legomena* riscontrata nel corpus qui esaminato

equivale al 48% (5.475/11.356), che sancisce la definitiva e completa aderenza della collezione di testi ai limiti stabiliti.

Quando si lavora con la statistica dei testi e si conducono le prime elaborazioni su un corpus, vi è un momento in cui ci si accorge che le regole non scritte e inconsapevoli che governano l'economia delle parole in un qualsiasi gruppo di testi posseggono connotati sorprendentemente vicini a una legge matematica. È questo il caso del *Grafico 1*, che rappresenta la Legge di Zipf.

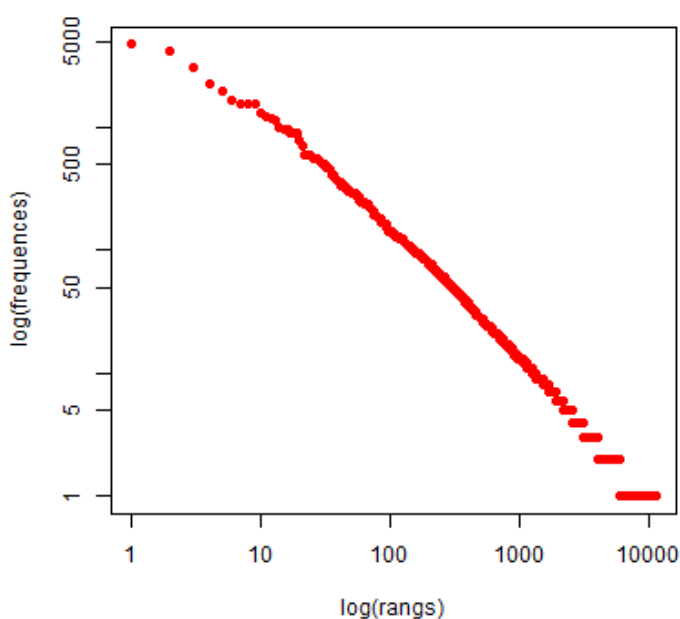


Grafico1. Rappresentazione grafica della Legge di Zipf, ottenuta con il software IRaMuTeQ

Il grafico vede sull'asse delle ascisse il logaritmo del rango delle parole all'interno del corpus ordinato dalla più frequente alla meno

frequente e sull'asse delle ordinate il logaritmo del numero di occorrenze delle rispettive parole. Ciò che è subito evidente dall'andamento descritto non è soltanto che man mano che si procede verso termini di rango inferiore, vi è una progressiva diminuzione del numero di occorrenze (come è naturale supporre, vista la natura stessa del grafico), ma che questo fenomeno procede con regolarità, fino a un certo punto del corpus, a partire dal quale la frequenza delle parole si stabilizza. Il fenomeno cosiddetto della "coda lunga" mostra come, a differenza delle alte frequenze del vocabolario, le parole con 5, ma soprattutto 2 e 1 occorrenza siano molto numerose. Nel caso specifico, le parole con più di 5 occorrenze sono circa 3000, mentre quelle con meno occorrenze superano le 8000 unità. In linea con quanto registrato tramite il tasso di copertura del vocabolario, la somma di queste occorrenze copre l'intero vocabolario del corpus esaminato, cioè a dire 11356 parole. Prendendo come riferimento il grafico descritto dalla Legge di Zipf, si noti che le parole significative sono quelle collocate nella parte centrale della curva, sull'asse delle ascisse. La motivazione alla base di questo fenomeno risiede nel fatto che le parole più vicine all'asse delle ordinate, con un elevato numero di occorrenze, sono troppo comuni per essere considerate significative - e, solitamente, si tratta delle *stop words* o *function words*, ovverosia le parole grammaticali -, mentre quelle che formano la cosiddetta "coda lunga" sono da considerarsi troppo rare e, di conseguenza, non indicative del contenuto di un testo o di un gruppo di testi.

3. ANALISI

3.1. Analisi con TaLTaC2

3.1.1. Segmenti ripetuti e Indice IS di Morrone

Tra le operazioni di più frequente utilizzo nell'ambito dell'analisi quantitativa dei testi figura l'individuazione dei segmenti ripetuti, che consiste nella ricerca dei poliformi ricorrenti nel testo e nella loro catalogazione. Il *software* TaLTaC2 consente di eseguire questa operazione in modo automatico, producendo una lista (esportabile in formato *.txt* al di fuori del programma) di locuzioni utilizzate di frequente nel corpus la quale, tuttavia, presenta anche diverse espressioni non significative. Per esempio, un poliforme come "hanno avuto" non è considerabile come significativo perché espressione molto comune composta dallo stesso verbo, utilizzato in un caso in forma ausiliaria. Per determinare la rilevanza dei poliformi individuati dal programma, giunge in soccorso un indice numerico detto "Indice IS di Morrone" (ideato nel 1996 da) che, dividendo la frequenza delle singole parole all'interno del segmento per la frequenza di quelle stesse parole all'interno del corpus, e sommandole tra loro, la moltiplica per il numero di parole piene che compongono quel poliforme, determinandone la significatività. Il concetto di parola piena è sintetizzabile in una parola che, di per sé,

porta un preciso significato; giocoforza, sono escluse da questa categoria le parole grammaticali, ovverosia quelle che vengono utilizzate per legare tra loro le altre parole e costruire le frasi.

$$IS = \sum_{i=1}^q \frac{f_{segm}}{f_i} * P$$

A titolo d'esempio, un'espressione come "Presidente della Repubblica", rilevata per 20 volte all'interno del corpus, che veda queste tre forme grafiche comparire in totale, rispettivamente, 25, 100 e 20 volte, avrà un indice di IS pari al $\{[(20/25)+(20/100)+(20/20)] \times 2\}$, cioè 4.

Tuttavia, come si è soliti fare in ambito statistico, è opportuno portare il valore assoluto dell'indice IS di Morrone a un valore relativo, in modo da poter confrontare opportunamente tra loro espressioni che abbiano numeri di parole diversi al loro interno.

L'indice IS relativo si ottiene, banalmente, dividendo l'indice IS assoluto per l'indice IS massimo; quest'ultimo, equivale sempre al numero di forme contenute nel poliforme al quadrato. In questo caso, dunque, l'indice IS relativo equivale a $(4/9)$, cioè allo 0,44.

La scelta di considerare l'indice relativo rispetto a quello assoluto, comporta una differenza nel *rank* della significatività dei segmenti considerati. La versione non normalizzata dell'indice (quella assoluta), evidenzierà espressioni più lunghe che, molto spesso, rientrano nella

categoria delle cosiddette “frasi” fatte. Di contro, un ordine stabilito prendendo come punto di riferimento l’indice relativo, tenderà ad affidare alle posizioni più alte del *ranking* espressioni più brevi e pregne di significato, tipiche di un linguaggio più tecnico.

Le rilevazioni effettuate con l’indice IS sul corpus preso in considerazione in questa ricerca, hanno sottolineato un tratto peculiare del Presidente Giuseppe Saragat, riscontrabile nel suo sesto e settimo discorso. Segmenti ripetuti di una lunghezza consistente, quali ad esempio “deriva sostanzialmente da debolezza morale” (IS relativo = 0,39) e “inviolabili diritti della persona umana trovano il loro appagamento”, evidenziano come il suo discorso del 1970 mutui intere frasi (e, di conseguenza, i concetti espressi in esse), direttamente dal suo precedente discorso, quello del 1969. A conferma di questo, un approccio qualitativo consistente nella lettura di entrambi i discorsi, fa emergere la sorprendente similarità di diverse parti di questi due discorsi.

3.1.2. Analisi delle specificità

Quando ci si trova a lavorare su un corpus di grande entità, in particolar modo se su di esso si sono costruite numerose variabili che ne aumentano le dimensioni comparative, può essere utile eseguire l’operazione di “analisi delle specificità”. Sfruttando questa funzione che

mette in luce il sovrautilizzo (o il sottoutilizzo) delle parole e utilizzando come variabile di riferimento i “Presidenti”, si sono registrati i valori riportati di seguito³.

Variabile: VPres - Modalità: **01Einaudi**

.....
Specificità positive:

Forma grafica	Occorrenze totali	Sub-occorrenze	p-value
anno	407	17	0,0000
patria	107	8	0,0000
prove	17	4	0,0000
vie	10	3	0,0001
nella	356	12	0,0004
nuove	74	5	0,0010
domani	45	4	0,0012
voto	21	3	0,0013
state	22	3	0,0015
nuovo	122	6	0,0018
ognuno	32	3	0,0044
percorso	10	2	0,0047
vogliono	10	2	0,0047
consapevoli	10	2	0,0047
pensiero	67	4	0,0052
avvenire	68	4	0,0055
ricostruzione	12	2	0,0068
possa	73	4	0,0071
tuttora	13	2	0,0080

³ Per un criterio di rilevanza, si è deciso di riportare soltanto le specificità positive (cioè a dire i valori sovrautilizzati), con un limite massimo di 25 per ogni modalità definita dalla variabile “VPres”.

Variabile: VPres - Modalità: **02Gronchi**

.....
Specificità positive:

Forma grafica	Occorrenze totali	Sub-occorrenze	p-value
costume	15	9	0,0000
rapporti	37	13	0,0000
aspetto	12	7	0,0000
avvenire	68	15	0,0000
problema	62	14	0,0000
collaborazione	42	11	0,0000
esigenza	24	8	0,0000
soluzione	28	8	0,0000
capacità	85	13	0,0003
progresso	77	12	0,0004
interessi	22	6	0,0006
nazioni	80	12	0,0006
auspicio	10	4	0,0010
sicuro	17	5	0,0012
vicende	17	5	0,0012
possibilità	34	7	0,0013
sociali	54	9	0,0013
errori	11	4	0,0016
considerare	11	4	0,0016

.....
Variabile: VPres - Modalità: **03Segni**

.....
Specificità positive:

Forma grafica	Occorrenze totali	Sub-occorrenze	p-value
commosso	20	4	0,0002
nazioni	80	7	0,0002

collaborazione	42	5	0,0005
profondamente	12	3	0,0007
avvenimenti	12	3	0,0007
sviluppo	99	7	0,0009
partecipazione	31	4	0,0013
beni	16	3	0,0018
popoli	114	7	0,0020
rapporti	37	4	0,0025
essa	38	4	0,0028
tutte	94	6	0,0034
quale	128	7	0,0038
corso	42	4	0,0040
sociale	101	6	0,0049
possa	73	5	0,0055
indipendenza	24	3	0,0059
auguro	25	3	0,0066
progresso	77	5	0,0069
sua	148	7	0,0083
assicurare	29	3	0,0101
mondiali	10	2	0,0102

Variabile: VPres - Modalità: **04Saragat**

.....
Specificità positive:

Forma grafica	Occorrenze totali	Sub-occorrenze	p-value
lavoratori	51	17	0,0000
corso	42	15	0,0000
economico	44	15	0,0000
libere	12	8	0,0000
ritmo	14	8	0,0000
tutti	529	66	0,0000

cara	10	6	0,0000
mezzi	23	8	0,0001
democratica	44	11	0,0002
posti	19	7	0,0002
sviluppo	99	18	0,0003
operai	15	6	0,0004
livello	27	8	0,0005
vengano	11	5	0,0006
tali	17	6	0,0009
numero	23	7	0,0009
sindacati	12	5	0,0010
problemi	141	21	0,0013
umana	46	10	0,0014
anniversario	13	5	0,0016
paesi	106	17	0,0016
conforto	19	6	0,0017

Variabile: VPres - Modalità: **05Leone**

.....
Specificità positive:

Forma grafica	Occorrenze totali	Sub-occorrenze	p-value
ripresa	49	16	0,0000
sicurezza	75	18	0,0000
difficile	41	12	0,0000
situazione	48	13	0,0000
crisi	97	19	0,0000
spesso	25	9	0,0000
possibile	26	9	0,0000
saranno	21	8	0,0000
alcuni	35	10	0,0000
sacrifici	35	10	0,0000

economica	72	14	0,0001
ansia	10	5	0,0002
imprenditori	22	7	0,0003
quella	130	19	0,0005
forme	18	6	0,0006
ambito	13	5	0,0009
industriale	13	5	0,0009
lavoratori	51	10	0,0012
senso	88	14	0,0012
necessario	43	9	0,0013
maggiore	52	10	0,0014
Paese	241	28	0,0014

Variabile: VPres - Modalità: **06Pertini**

.....
Specificità positive:

Forma grafica	Occorrenze totali	Sub-occorrenze	p-value
io	323	189	0,0000
vi	245	107	0,0000
popolo	220	104	0,0000
terrorismo	93	62	0,0000
ebbene	60	46	0,0000
italiano	134	69	0,0000
loro	365	126	0,0000
sono	703	191	0,0000
giovani	252	95	0,0000
quindi	102	52	0,0000
questi	234	86	0,0000
ripeto	40	30	0,0000
amici	31	26	0,0000
mi	361	111	0,0000

mila	19	18	0,0000
qui	105	45	0,0000
fame	28	21	0,0000
guerra	85	39	0,0000
mia	106	44	0,0000
contro	95	41	0,0000
terroristi	19	16	0,0000
domande	24	18	0,0000
ordigni	12	12	0,0000
fianco	12	12	0,0000
vedete	28	19	0,0000

Variabile: VPres - Modalità: **07Cossiga**

.....
Specificità positive:

Forma grafica	Occorrenze totali	Sub-occorrenze	p-value
comunità	106	44	0,0000
società	129	48	0,0000
libertà	190	57	0,0000
civile	99	37	0,0000
Est	11	10	0,0000
cittadini	170	45	0,0000
certo	86	28	0,0000
strutture	27	14	0,0000
culturali	18	10	0,0000
moderna	18	10	0,0000
dovuto	13	8	0,0000
istituzionali	20	10	0,0000
anzi	24	11	0,0000
europeo	24	11	0,0000
convivenza	30	12	0,0001

istituzioni	123	30	0,0001
amministrazione	26	11	0,0001
doveri	35	13	0,0001
verso	91	24	0,0001
sviluppo	99	25	0,0002
culturale	16	8	0,0002

Variabile: VPres - Modalità: **08Scalfaro**

.....
Specificità positive:

Forma grafica	Occorrenze totali	Sub-occorrenze	p-value
grazie	87	62	0,0000
chi	183	99	0,0000
È	1681	517	0,0000
volte	77	53	0,0000
no	42	35	0,0000
dire	114	64	0,0000
auguri	86	53	0,0000
non	1186	362	0,0000
pagina	23	22	0,0000
tema	44	33	0,0000
partiti	61	39	0,0000
mai	118	60	0,0000
capo	52	35	0,0000
vorrei	72	40	0,0000
cosa	64	37	0,0000
politica	191	79	0,0000
gente	35	25	0,0000
amore	35	24	0,0000
Stato	187	75	0,0000
qualche	61	34	0,0000

quanta	15	14	0,0000
magistratura	36	24	0,0000

Variabile: VPres - Modalità: **09Ciampi**

.....
Specificità positive:

Forma grafica	Occorrenze totali	Sub-occorrenze	p-value
identità	15	12	0,0000
secolo	17	12	0,0000
unione	42	19	0,0000
Italia	465	89	0,0000
generazione	23	13	0,0000
nostre	87	27	0,0000
nostri	138	36	0,0000
forte	52	19	0,0000
Euro	13	9	0,0000
europea	82	25	0,0000
stiamo	23	12	0,0000
Europa	191	43	0,0000
provincia	11	8	0,0000
mente	18	10	0,0000
care	12	8	0,0000
patria	107	28	0,0000
europei	31	13	0,0000
nostra	308	59	0,0000
iniziative	31	12	0,0000
fiducia	137	31	0,0000
cari	36	13	0,0001
voi	230	45	0,0001
confronto	32	12	0,0001

Variabile: VPres - Modalità: **10Napolitano**

.....

Specificità positive:

Forma grafica	Occorrenze totali	Sub-occorrenze	p-value
crisi	97	44	0,0000
scelte	34	22	0,0000
guardando	17	14	0,0000
futuro	61	30	0,0000
ricerca	35	21	0,0000
Italia	465	126	0,0000
quel	80	33	0,0000
debito	14	11	0,0000
crescita	37	19	0,0000
opportunità	20	13	0,0000
qualità	23	14	0,0000
globale	13	10	0,0000
coesione	18	12	0,0000
generazioni	24	14	0,0000
nuove	74	29	0,0000
riforme	46	21	0,0000
economia	50	22	0,0000

Variabile: VPres - Modalità: **11Mattarella**

.....

Specificità positive:

Forma grafica	Occorrenze totali	Sub-occorrenze	p-value
vanno	32	12	0,0000
concittadini	15	8	0,0000
riconoscenza	17	8	0,0000
va	80	16	0,0000

persone	64	14	0,0000
nostri	138	20	0,0000
tante	78	14	0,0000
difficoltà	67	12	0,0000
regole	30	8	0,0000
positive	11	5	0,0001
luoghi	11	5	0,0001
elettori	12	5	0,0001
vuol	20	6	0,0002
nostre	87	12	0,0005
esiste	10	4	0,0007
alcuni	35	7	0,0008
donne	36	7	0,0009
Paese	241	22	0,0012
giorno	61	9	0,0015
rispettare	12	4	0,0015
desidero	20	5	0,0016
vita	217	20	0,0018
oltre	30	6	0,0018

Con 3 messaggi di fine anno fino al momento in cui viene realizzata questa ricerca, si ritiene interessante delineare alcune caratteristiche che distinguono il linguaggio dell'attuale Presidente della Repubblica, Sergio Mattarella, rispetto ai suoi predecessori. Tra le specificità del presidente palermitano, spicca la parola "concittadini" (con 8 occorrenze su 15 del corpus totale - le altre appartengono a Cossiga [1], Saragat [4] e Gronchi [2]). Tramite la visualizzazione delle concordanze, che permettono di osservare il

contesto specifico dal quale è stata estrapolata una parola, si può notare che Mattarella sceglie di utilizzare questo termine per introdurre un encomio a cittadini italiani che hanno raggiunto traguardi di rilievo a livello internazionale, o per rivolgersi ai cittadini italiani all'ascolto. Il prefisso "con-" ha una valenza inclusiva che testimonia la volontà di accorciare le distanze tra la più alta carica dello Stato e il normale cittadino medio che figura tra i destinatari dell'augurio di fine d'anno. Se, come evidenziato da Davide Sterchele e La Mendola (2007), spesse volte Sandro Pertini è stato - dal punto di vista della dicotomia tra paleo-televisione e neo-televisione - lo spartiacque tra i suoi predecessori e i suoi successori in alcune innovazioni che hanno contribuito a "costruire il noi" e ad avvicinare l'istituzione del Presidente della Repubblica ai cittadini comuni (dal riferimento "all'intrusione" nelle case degli italiani, alla sedia vuota posta di fianco alla sua), in questo caso, Sergio Mattarella accorcia le distanze tra i cittadini e il Presidente con l'uso di un termine inclusivo che comprende sia il parlante sia i destinatari del messaggio.

Un altro dei termini che risultano specifici dell'attuale Presidente della Repubblica è "riconoscenza", già utilizzato in passato da altri presidenti - in particolare da Pertini, che vede questa parola occorrere 5 volte all'interno dei suoi discorsi - ma sovrautilizzato rispetto a questi e presente 8 volte nel subcorpus individuato dalla modalità associata a Mattarella per la variabile "VPres". Come evidenziato nell'analisi delle concordanze, il contesto di utilizzo di questa parola nei messaggi di fine

d'anno del Presidente attualmente in carica non si discosta da quanto visto degli altri subcorpora. Il sentimento di riconoscenza è rivolto, sia dal punto di vista personale, sia a nome dell'intera collettività in ascolto e rappresentata dal Capo dello Stato stesso, a istituzioni quali il Papa, la magistratura e le forze dell'ordine. In modo analogo, il Pontefice al tempo in carica (Giovanni Paolo II) era stato ringraziato da Oscar Luigi Scalfaro, mentre Sandro Pertini (in particolar modo nel suo settimo e ultimo discorso) aveva manifestato la propria riconoscenza agli Stati Uniti d'America che avevano aiutato l'Italia a liberarsi dall'oppressione del regime nazifascista.

Tuttavia, nel caso di Sergio Mattarella, non sono soltanto le istituzioni a godere della riconoscenza del Presidente; ne beneficiano anche organizzazioni che si sono spese per il bene comune o intere categorie di persone o fasce della popolazione particolarmente meritevoli di gratitudine, come nel caso della totalità delle donne italiane citate nel messaggio del 2015 e ivi rappresentate da Fabiola Gianotti, Samantha Cristoforetti e Nicole Orlando.

3.1.3. I valori specifici di ogni presidente

Al fine di perseguire lo scopo della ricerca - che, si ricorda, consiste nel generare un "discorso tipo" composto da frammenti presi

da più messaggi di fine d'anno, fornendo come parametri un gruppo di valori scelti a priori e utilizzando metodi di selezione statistica - è necessario soffermarsi brevemente sulle categorie valoriali che verranno utilizzate in questa operazione.

Nel capitolo “Valori politici e capitale sociale nei discorsi dei presidenti della Repubblica” contenuto in “Messaggi dal Colle”, Selena Grimaldi e Gianni Riccamboni (2007) - due scienziati politici e docenti dell'Università di Padova - hanno analizzato i messaggi di fine d'anno dei capi dello Stato succedutisi in Italia e ne hanno evidenziato i punti di contatto e di discontinuità rispetto non soltanto al lessico utilizzato, ma anche al significato che portano con sé e dei simboli dei quali sono intrisi. Il punto di vista dal quale muovono gli autori si discosta dalla mera analisi politica di un discorso istituzionale, avvicinandosi più a una individuazione delle leve che i presidenti muovono per trasmettere ai cittadini i valori costitutivi della Repubblica e quelli che ognuno di loro ritiene più importanti con lo scopo di avvicinarli alle istituzioni e tra loro stessi, nell'ottica di un Paese unito a partire dalla sua cittadinanza. Come illustrato da Grimaldi e Riccamboni: “... *l'istituzione presidenziale viene vista nella sua ineliminabile valenza simbolica, capace di suscitare identificazioni, plasmare le visioni del mondo degli individui e dunque di orientare le loro condotte di azione, piuttosto che soffermandosi sulle regole, le norme e le routine con cui essa si caratterizza*”. L'approccio metodologico utilizzato e citato dai due scienziati politici dell'Università di Padova corrisponde al neoistituzionalismo normativista di James G.

March and Johan P. Olsen (Rediscovering Institutions. The Organizational basis of politics, 1989), secondo cui la funzione dell'istituzione non sarebbe soltanto quella di aggregare parti del Paese diverse sotto un unico assetto (politico, normativo, sociale), ma di *integrare* le diverse parti di questo insieme. Partendo da questi presupposti, il contributo inserito in Messaggi dal Colle individua una serie di valori per i quali i presidenti della Repubblica si sono spesi nel corso dei loro mandati e li cataloga seguendo criteri diversi e raggruppandoli in tre insiemi: valori politici, valori morali e valori sociali.

Elenco dei valori individuati da Selena Grimaldi e Gianni Riccamboni:

Valori sociali	Valori morali	Valori politici
lavoro, giovani, famiglia, collaborazione, tradizione, civile, conoscenza, eroismo, comunità, progresso, memoria.	fiducia, ideali, doveri, religione, giustizia, impegno, responsabilità, sacrificio, coraggio, dignità, tolleranza, solidarietà, lealtà, etica, onestà.	Italia, Europa, unità, identità, stabilità, Risorgimento, patria, libertà, nazione, democrazia, Repubblica, pace, diritti, uguaglianza, indipendenza, Costituzione, Resistenza, partecipazione.

Prendendo in considerazione i valori individuati da Grimaldi e Riccamboni in “Messaggi dal Colle” (a cura di Michele A. Cortelazzo e Arjuna Tuzzi, 2007) e incrociandoli con i dati emersi dall’analisi delle

specificità, sono emersi quali valori sono stati trattati con maggiore o minore enfasi da ognuno degli 11 Presidenti. Come già specificato, l'utilizzo del *p-value* consente di individuare sia ciò che è sovrarappresentato in una variabile rispetto alle altre, sia ciò che è sottorappresentato. Tuttavia, è opportuno osservare questi risultati anche da una prospettiva qualitativa per rendersi conto di una differenza di peso tra i due insiemi individuati. Mentre ciò che è sovrarappresentato è indice del fatto che, nel caso specifico dei Presidenti della Repubblica, un Capo dello Stato abbia ritenuto opportuno sottolineare alcuni aspetti legati ai valori rilevati con nelle specificità positive, tutto quello che rientra nell'insieme dei valori sottorappresentati non può dirsi necessariamente trascurato da quello stesso Presidente. In altre parole, data la complessità e la vastità dei temi trattati, la varietà delle situazioni e dei trascorsi politici di un Presidente e il momento storico in cui un discorso viene pronunciato, il caso in cui un tema non venga trattato (o venga trattato marginalmente) all'interno di uno dei subcorpora, è da considerarsi meno rilevante rispetto al caso in cui, invece, su quello stesso tema vi fosse la precisa volontà di soffermarsi ampiamente da parte dell'autore del discorso.

Valori specifici per ogni Presidente selezionati tramite il p-value:

— : specificità positive

— : specificità negative

Einaudi: +: **patria**, -: /

Gronchi: +: **collaborazione**, **progresso** -: **Europa**

Segni: +: **indipendenza**, **partecipazione**, **collaborazione**, **progresso** -: /

Saragat: +: **Repubblica** -: **impegno**, **giovani**

Leone: +: / -: **Italia**, **Europa**, **pace**

Pertini: +: **giovani** -: **Italia**, **Costituzione**, **fiducia**, **impegno**, **responsabilità**, **civile**, **comunità**, **progresso**

Cossiga: +: **libertà**, **doveri**, **impegno**, **tradizione**, **civile**, **comunità** -: **Italia**, **patria**, **fiducia**, **lavoro**, **giovani**

Scalfaro: +: / -: **libertà**, **nazione**, **civile**, **progresso**

Ciampi: +: **Italia**, **Europa**, **unità**, **identità**, **patria**, **fiducia**, **ideali**, **famiglia**

-: **coraggio**

Napolitano: +: **Italia** -: **patria**, **libertà**, **democrazia**, **pace**, **giustizia**

Mattarella: +: **impegno** -: /

Si specifica che per le successive analisi si è deciso di mantenere la suddivisione in valori politici, morali e sociali. Questa categorizzazione tornerà utile, in un secondo momento, per le elaborazioni effettuate con il programma IRaMuTeQ, che porteranno alla generazione di un *rank* in cui

frasi dei discorsi dei vari presidenti saranno catalogate in base al tasso di aderenza rispetto a una delle tre sfere valoriali.

3.2 Analisi con IRaMuTeQ

3.2.1 T-Gen sui valori specifici dei presidenti

Mentre il software TaLTaC2, già utilizzato in precedenza, è particolarmente indicato per esplorare il corpus di una ricerca, creare variabili ed elaborare contenuti a partire dalle dimensioni comparative d'interesse per la ricerca, IRaMuTeQ, il programma utilizzato nelle analisi che seguiranno, è orientato all'utilizzo di diversi metodi (quale, ad esempio, la *topic extraction*) e all'elaborazione di grafici ad essi relativi. Prima di lavorare a elaborazioni più profonde, un corpus di testi importato in IRaMuTeQ può essere trattato in modo sostanzialmente analogo a quanto fatto con TaLTaC2. Tuttavia, dal momento che per questa ricerca il corpus è stato pretrattato e ripulito con il software italiano, si è deciso di utilizzare il quello stesso file di testo per proseguire con le analisi. Dopo l'esportazione da TaLTaC2, l'unico accorgimento necessario perché il corpus fosse riconosciuto da IRaMuTeQ è stato quello di ricodificare le variabili individuate in precedenza e presenti nel file di testo, con un codice comprensibile per il software francese.

Una volta importato il corpus, si è potuto immediatamente procedere con le analisi dal punto in cui si era terminato il lavoro con il software precedente. Considerando i valori specifici di ogni presidente che sono emersi nella ricerca con TaLTaC2, si è utilizzato il modello del Test del Chi quadrato (χ^2) tramite la funzione T-Gen. Lo scopo di questa procedura è di verificare quanto uno o più vocaboli sia stato sovrautilizzato o sottoutilizzato da un determinato Presidente e di presentare il risultato con un istogramma. Ogni colonna dell'istogramma rappresenta l'utilizzo della parola per il quale il T-Gen è stato calcolato, all'interno del subcorpus del relativo Presidente. Questa misura statistica utilizza la frequenza reale e la frequenza attesa di uno o più caratteri individuati da una variabile e calcola la deviazione nella distribuzione della prima dalla seconda. Per sua natura, questo procedimento è strettamente influenzato dal modo in cui si calcola la frequenza attesa che, in questo caso, è stata generata moltiplicando il totale di riga della parola oggetto dell'analisi per il totale delle parole utilizzate da un presidente e dividendo questo risultato per il numero totale delle parole del corpus.

In prima istanza, la funzione T-Gen è stata applicata a tutti i valori specifici positivi per ogni modalità individuata dalla variabile "Presidente".

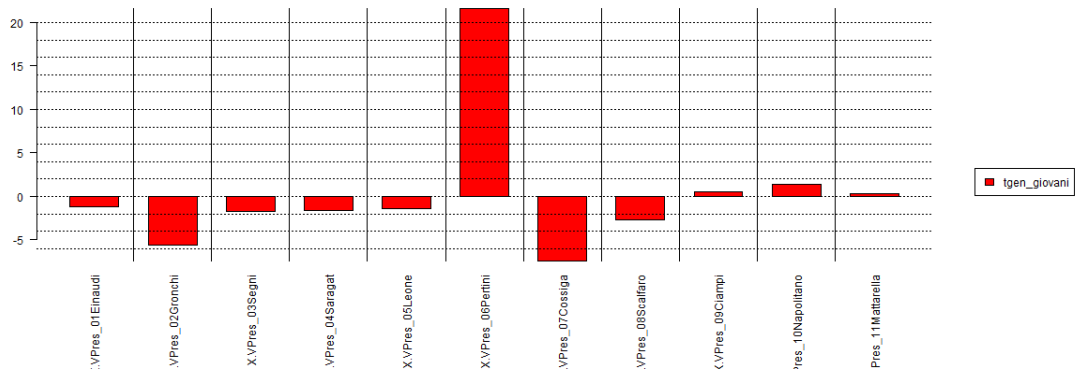


Fig. 1. T-Gen per la parola “giovani”.

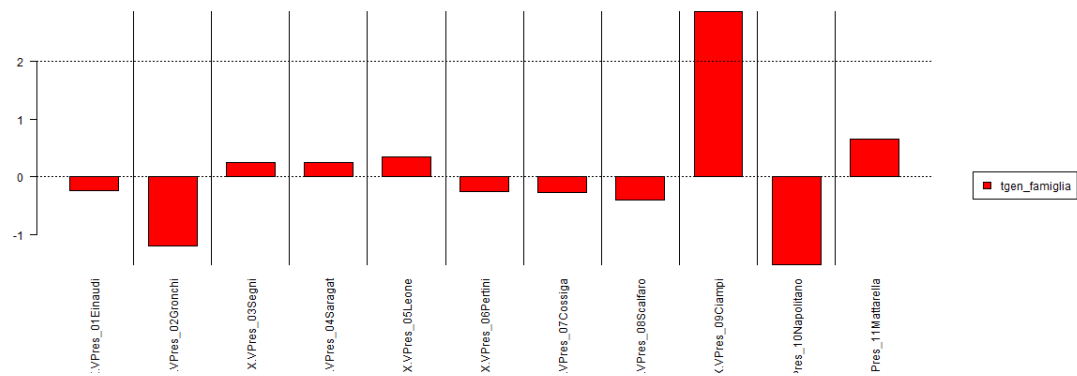


Fig. 2. T-Gen per la parola “famiglia”.

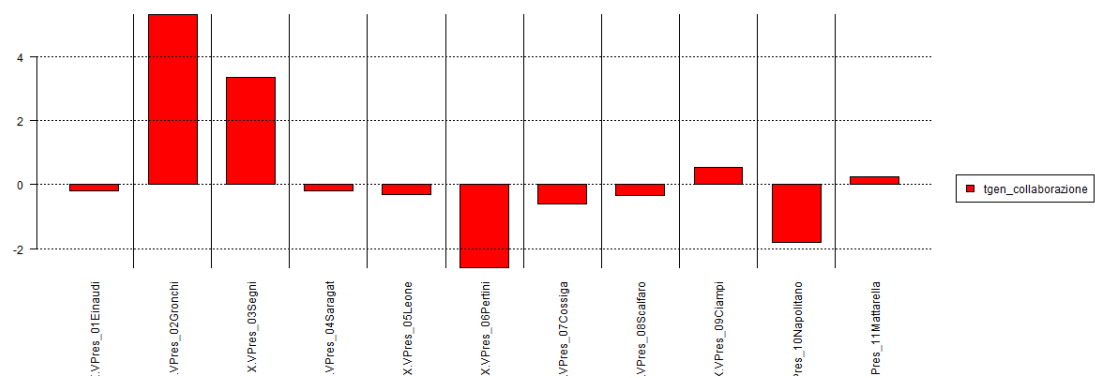


Fig. 3. T-Gen per la parola “collaborazione”.

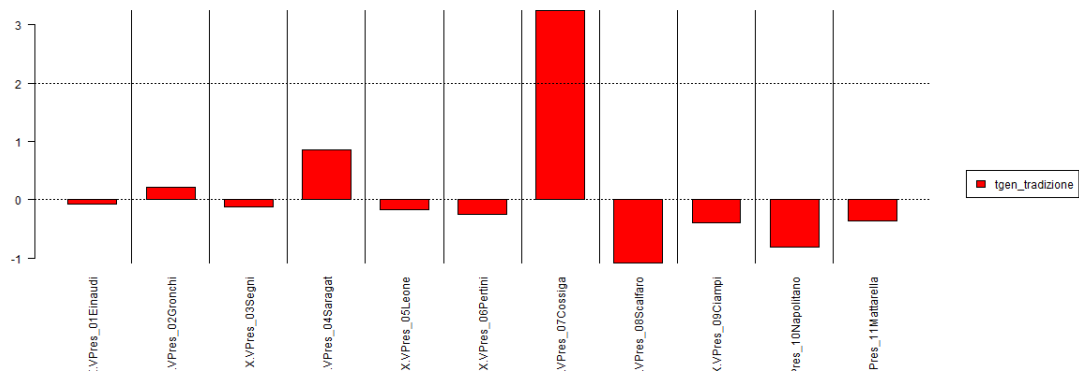


Fig. 4. T-Gen per la parola “tradizione”.

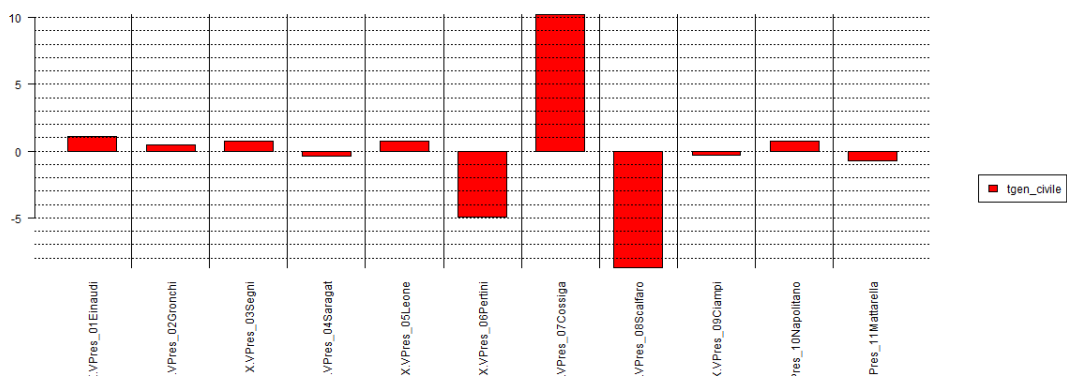


Fig. 5. T-Gen per la parola “civile”.

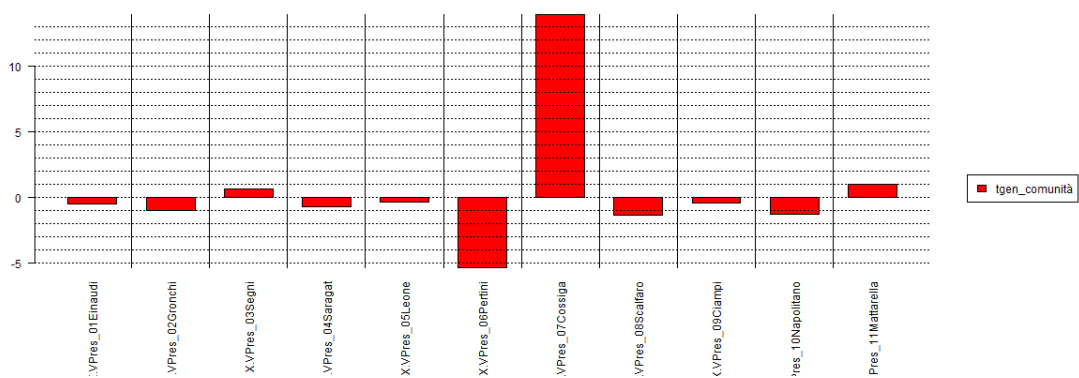


Fig. 6. T-Gen per la parola “comunità”.

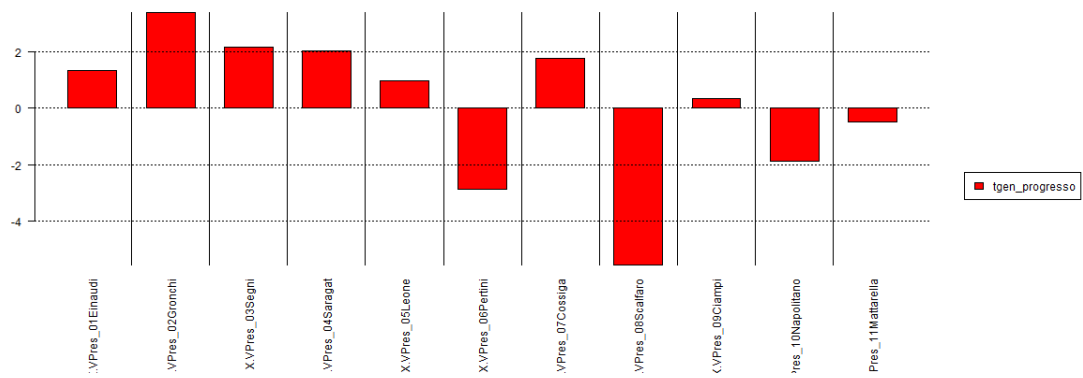


Fig. 7. T-Gen per la parola “progresso”.

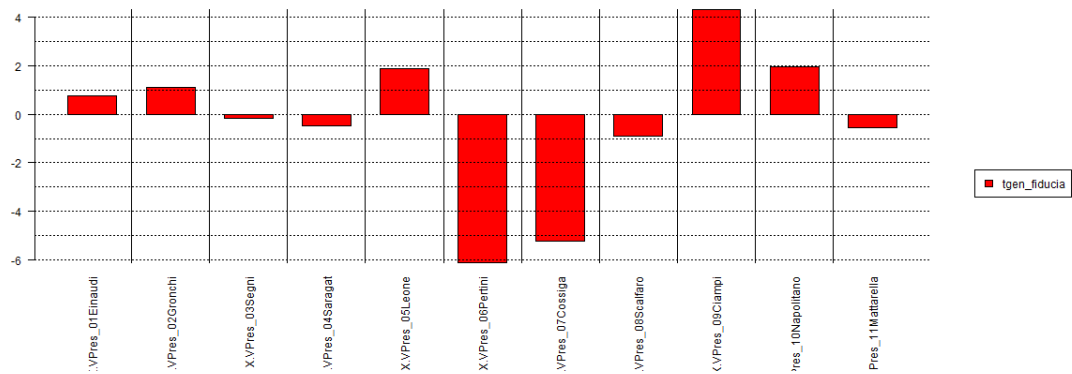


Fig. 8. T-Gen per la parola “fiducia”.

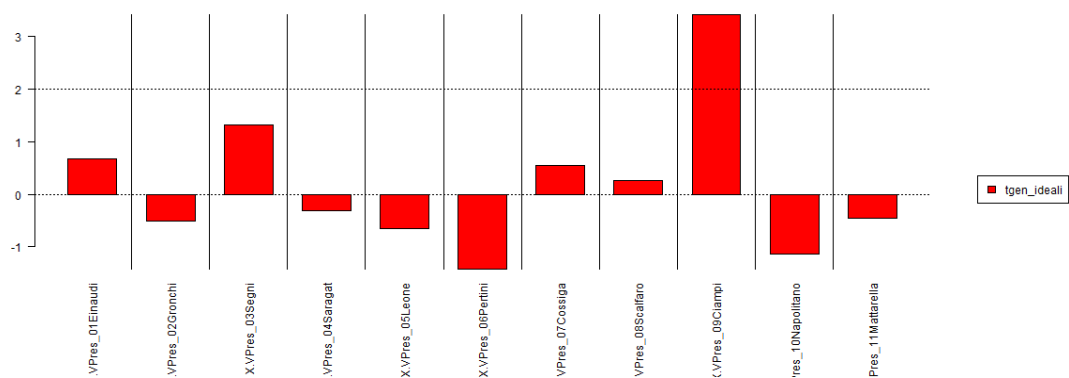


Fig. 9. T-Gen per la parola “ideali”.

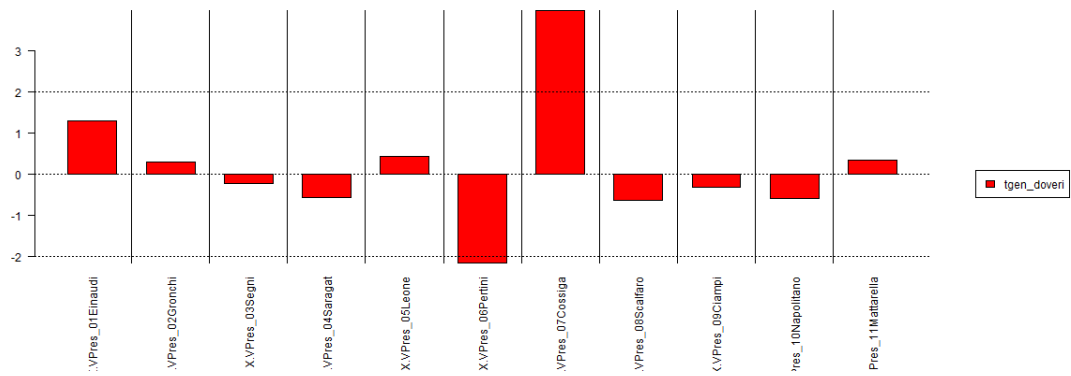


Fig. 10. T-Gen per la parola “doveri”.

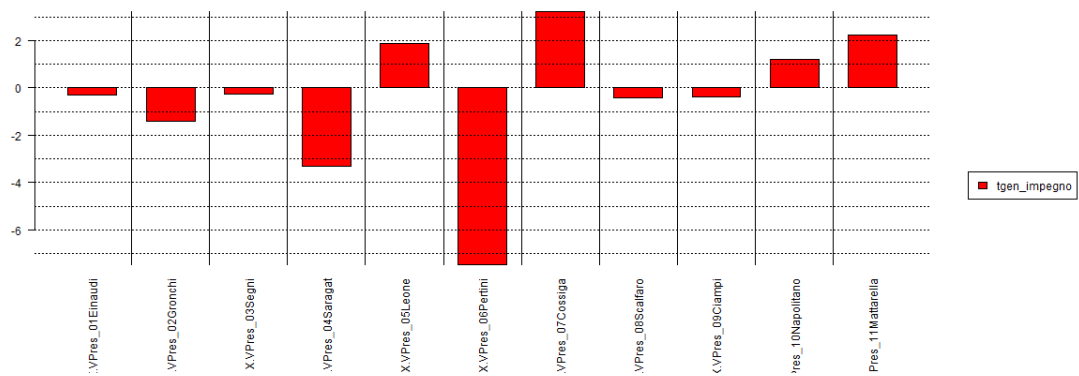


Fig. 11. T-Gen per la parola “impegno”.

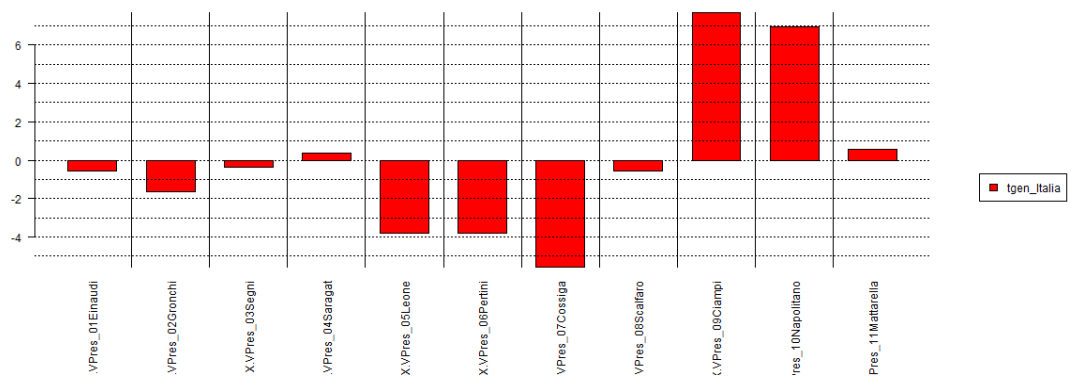


Fig. 12. T-Gen per la parola “Italia”.

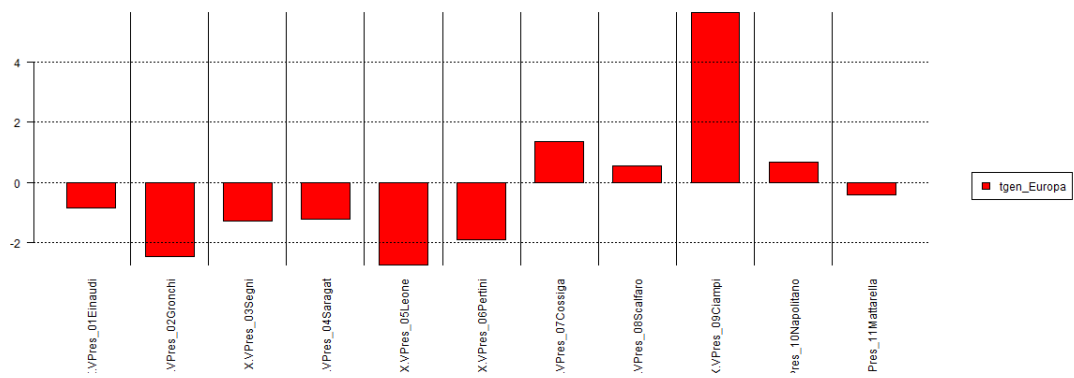


Fig. 13. T-Gen per la parola “Europa”.

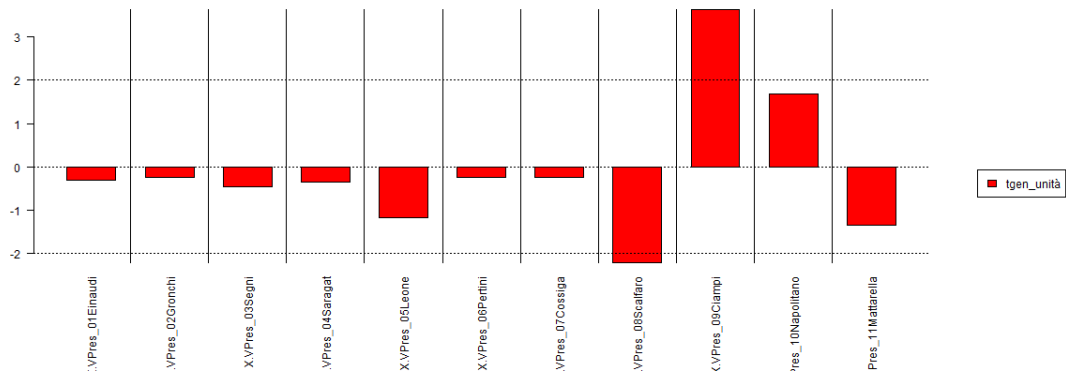


Fig. 14. T-Gen per la parola “unità”.

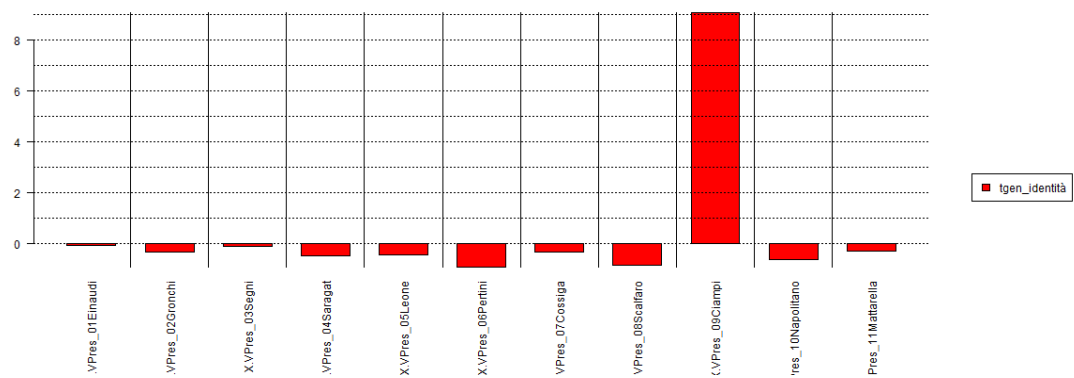


Fig. 15. T-Gen per la parola “identità”.

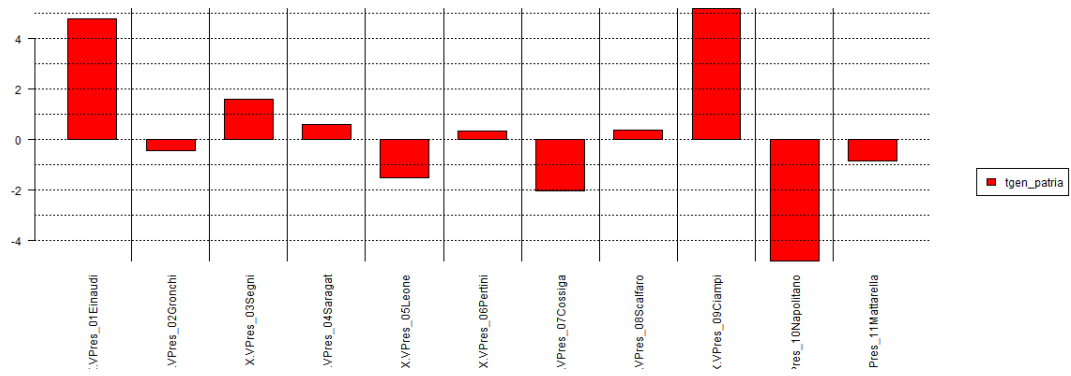


Fig. 16. T-Gen per la parola “patria”.

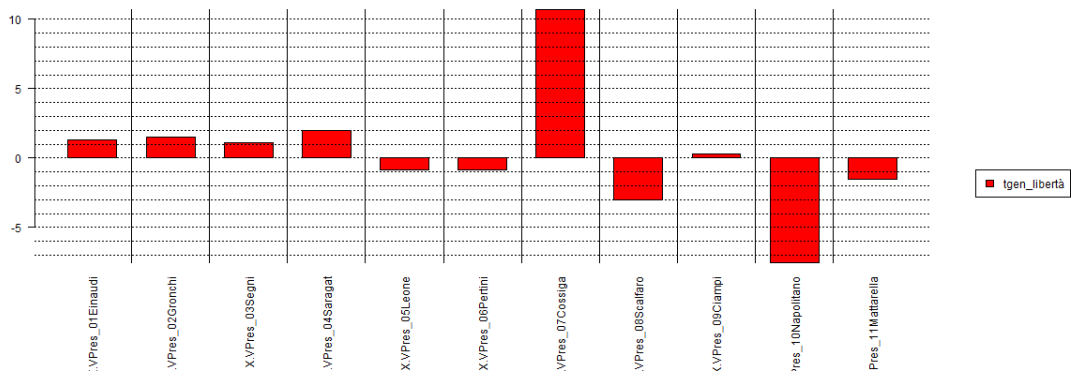


Fig. 17. T-Gen per la parola “libertà”.

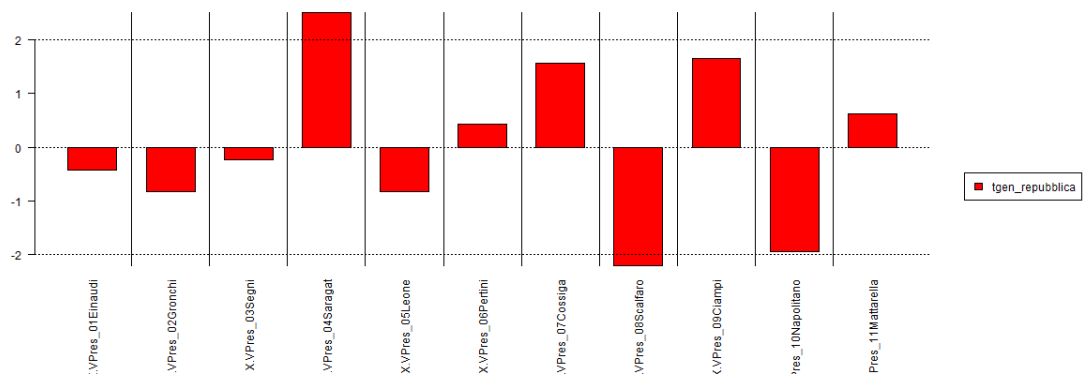


Fig. 18. T-Gen per la parola “Repubblica”.

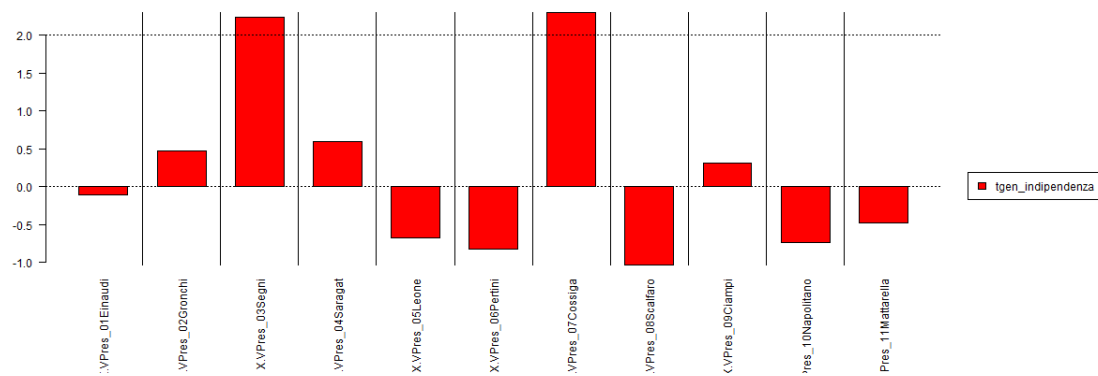


Fig. 19. T-Gen per la parola “indipendenza”.

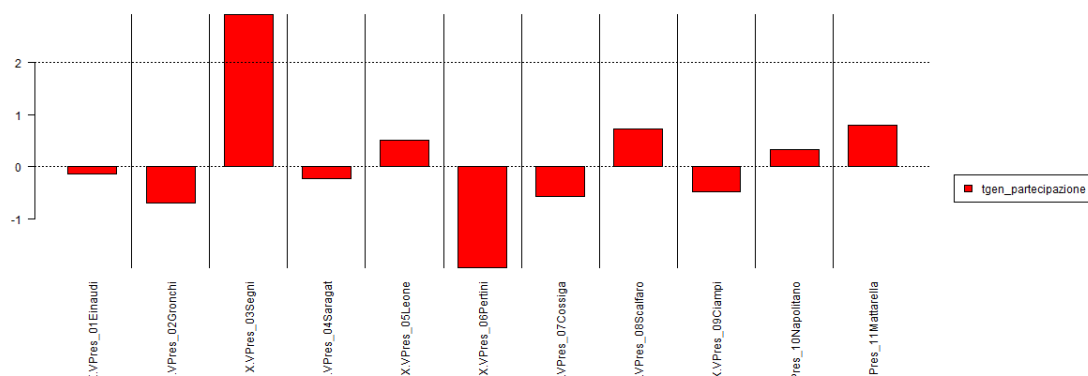


Fig. 20. T-Gen per la parola “partecipazione”.

Una volta analizzati i singoli valori, si è sfruttata la stessa funzione T-Gen per elaborare i tre grafici relativi ai raggruppamenti delle tre sfere valoriali individuate in precedenza. Lo scopo di questa ulteriore elaborazione è di avere una visione della distribuzione dei valori nei singoli Presidenti nel loro complesso e con un ingrandimento minore, in modo da ottenere già da questi grafici un’indicazione su quali saranno i Presidenti dai quali, più probabilmente, verrà estratta la maggior parte delle frasi che comporrà i discorsi tipo per ognuna delle tre classi valoriali.

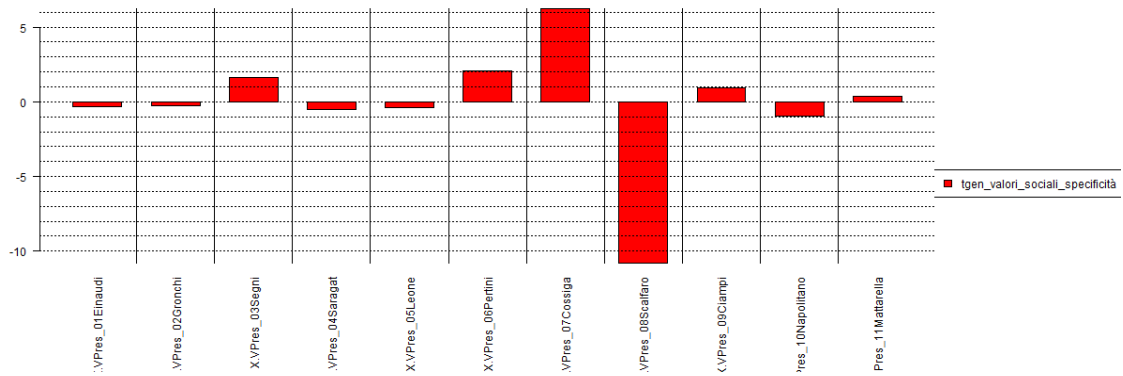


Fig. 21. T-Gen per la classe dei valori sociali.

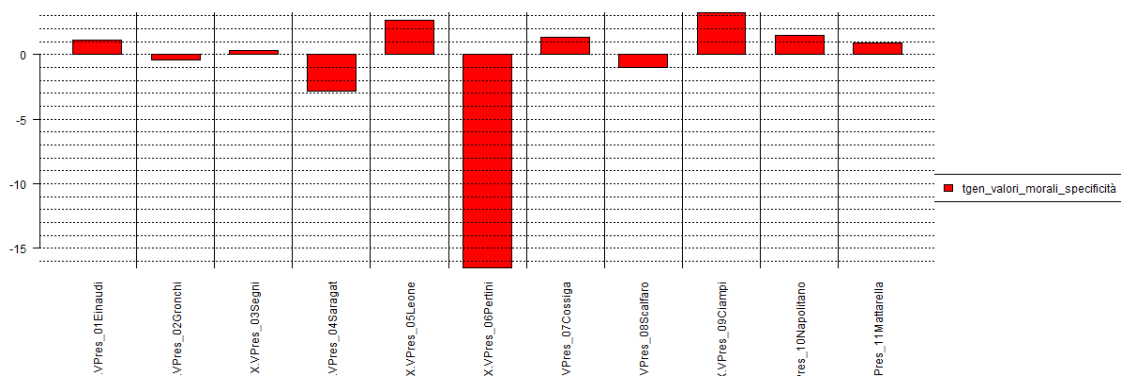


Fig. 22. T-Gen per la classe dei valori morali.

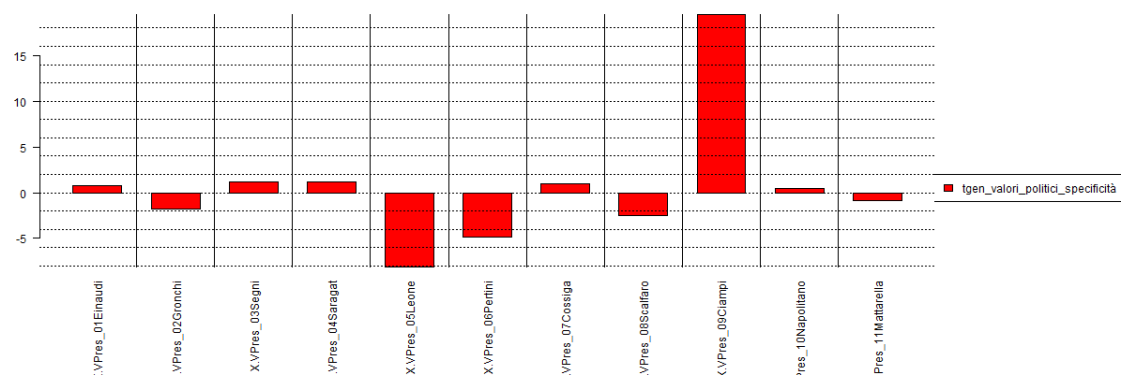


Fig. 23. T-Gen per la classe dei valori politici.

Come emerge dai grafici precedenti, Ciampi è il Presidente in cui, più di ogni altro, emergono i valori politici; nello specifico e in particolare quelli individuati nelle parole Italia, Europa, unità, identità e patria. Ciò non stupisce e conferma le osservazioni qualitative proposte da Selena Grimaldi nel suo “I Presidenti nelle forme di governo” (2012). Infatti, nel periodo del suo settennato, Carlo Azeglio Ciampi fu autore e promotore di un processo di valorizzazione culturale dell’Italia che riportò in auge il concetto di “Patria”, esortando tutte le parti della società a riscoprire - in chiave repubblicana e non più risorgimentale - i simboli costitutivi dell’identità nazionale che si erano persi nel travagliato susseguirsi delle legislature e dei conflitti politici interni alla cosiddetta Prima Repubblica.

Per quanto concerne i valori sociali, in modo inaspettato rispetto a quanto si ipotizzava, spicca la figura di Francesco Cossiga che ha il grado di specificità più elevato tra gli altri presidenti, in questa categoria. Se i giovani e la famiglia non costituiscono il perno dei suoi discorsi, i concetti di tradizione e comunità appaiono centrali, così come l’aggettivo “civile”. Al contrario, non stupisce che la famiglia sia un vocabolo fondamentale in Ciampi, il quale spesso evocava il concetto di famiglia come istituzione che rafforza l’identità patriottica cui si è fatto riferimento in precedenza. Inoltre, unico tra tutti i Capi di Stato italiani, il presidente livornese era solito congedarsi con un saluto finale agli italiani che comprendeva idealmente anche la moglie Franca. Analogamente alla predilezione di Ciampi per la famiglia, vi è quella di Sandro Pertini per i giovani. Sovente, il Presidente che si presentava ai cittadini italiani con la

pipa in bocca si rivolgeva ai giovani dimostrando fiducia nelle future generazioni, ma anche preoccupazione per la disoccupazione diffusa in quella fascia della popolazione.

Ed è proprio su Pertini che si focalizza una riflessione di carattere metodologico che si intende ribadire e che coinvolge la categoria dei valori morali. Come evidenziato dal T-Gen relativo a questo raggruppamento, nessun Presidente presenta specificità positive particolarmente pronunciate in questa categoria. Piuttosto, quello che viene evidenziato è la sottorappresentazione di questi valori all'interno del subcorpus di Pertini. Ciò che si intende ribadire e che riconduce alla scelta operata in precedenza di scartare le specificità negative dalle successive analisi, è che il sottoutilizzo di determinati termini individuato dalle specificità non è rilevante quanto il sovrautilizzo, perché il fatto che Pertini abbia valori così bassi in questa categoria può non significare che ne abbia parlato poco, ma che gli altri presidenti ne abbiano parlato di più. Infatti, nella scala di misura del grafico, il valore corrispondente al picco negativo presente in Pertini è -15 mentre il picco positivo (prerogativa di Ciampi) supera il valore +3. Queste misurazioni descrivono un grafico sbilanciato verso le specificità negative ma, dal punto di vista di quelle positive, non molto lontano da quello relativo ai valori sociali che vede il suo massimo nel valore +6 di Cossiga.

3.2.2 Metodo Reinert e classi semantiche

Tra i metodi di ricerca messi a disposizione da questo software vi è una procedura di *topic extraction* che, dato un corpus di testi in input, individua dei gruppi di parole il più possibile coerenti al loro interno, basandosi sui segmenti di testo che estrae dal corpus totale. Questo significa che, attraverso l'analisi delle parole e dei segmenti di un gruppo di testi, è possibile delineare una serie di tematiche (*topics*) di cui quel corpus si occupa a partire dalle parole che connotano quegli stessi temi. Per quanto riguarda la trasposizione grafica dei risultati, è importante sottolineare che essa ha diverse forme, alcune delle quali più idonee a rappresentarne aspetti numerici, altre a mostrare il linguaggio specifico che compone ogni *topic*. Prima di effettuare questa procedura, si sono scorporate le parole grammaticali (*function words*) dalle parole piene, dimodoché tutto ciò che non ha a che vedere con le aree semantiche dei discorsi, ma con la struttura del linguaggio, non intaccasse il risultato della *cluster analysis*.

L'analisi effettuata sul corpus dei 69 messaggi di fine d'anno dei Presidenti della repubblica ha portato all'individuazione di 8 *cluster* di parole, ognuno dei quali può essere ricondotto a un argomento o a un gruppo di più tematiche. Come regolarmente accade nella *cluster analysis*, alcuni gruppi hanno un'identità maggiormente definita, sia in termini numerici, sia nel momento in cui si definisce qualitativamente l'area di riferimento del *cluster*, mentre altri sono più difficilmente

classificabili. Non solo, i gruppi che al loro interno contengono maggiore eterogeneità in termini linguistici e contenutistici, solitamente, sono i *cluster* residuali, vale a dire quelli in cui il software raccoglie le porzioni di testo che non è riuscito a classificare altrimenti.

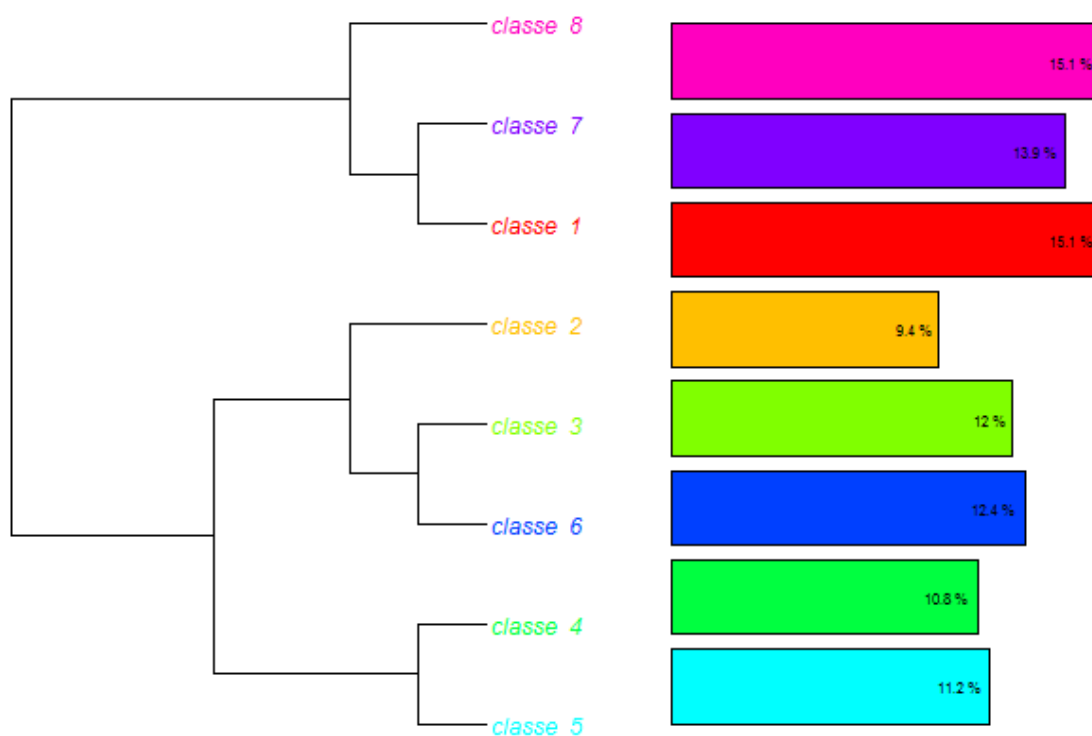


Fig. 24. Dendrogramma di classificazione degli 8 *topics* estratti dal *corpus* con le percentuali di unità testuali associate.

Il dendrogramma della Figura 24 illustra la suddivisione degli ammassi individuati con IRaMuTeQ, mostrando la rispettiva distanza tra ognuno di essi e la percentuale di unità testuali contenuta al loro interno. I topics 1 e 7 sono i più simili tra loro, così come il numero 3 e il numero

6. In seconda istanza, vi è una somiglianza - seppur meno marcata rispetto alle precedenti - tra il *topic* numero 8 e il gruppo composto dai *topics* 1 e 7, tra il *cluster* 2 e la coppia 3 e 6, e tra gli ammassi 4 e 5. Numericamente, i due gruppi che contengono il maggior numero di porzioni di testo, sono il numero 1 e il numero 8, con il 15,1% di unità testuali al loro interno, mentre il *topic* di dimensioni più ridotte è il numero 2, con il 9,4% delle porzioni.

Il meccanismo attraverso il quale si delineano queste somiglianze parte dal *corpus* totale e, ricorsivamente, lo suddivide in parti sempre più piccole e sempre più simili tra loro, fermandosi al punto in cui un'ulteriore suddivisione porterebbe a un'identificazione di ammassi più ambigui rispetto a quelli del livello precedente. Da un punto di vista metodologico è doveroso precisare che, impostando dei parametri specifici all'interno del *software*, è possibile incrementare questo limite. Benché possa sembrare un controsenso rispetto a quanto spiegato poche righe sopra, quest'ultima operazione risulta particolarmente efficace - se non addirittura salvifica - in circostanze in cui il corpus analizzato sia estremamente eterogeneo e sfaccettato da costringere il programma a individuare soltanto due *cluster*, che difficilmente potrebbero produrre un risultato sufficientemente profondo e soddisfacente. In tal caso, il limite dei *cluster* individuabili potrebbe essere ignorato, in modo da produrre tanti ammassi di dimensioni ridotte che al loro interno contengono parole che fanno riferimento a ognuna delle complesse sfaccettature che compongono il *corpus*.

Come emerge dalla rappresentazione grafica della Figura 25, i *cluster* 4 e 5 sono identificativi di quella parte della politica che nel mondo anglosassone viene chiamata *politics* e che riguarda i meccanismi istituzionali e il lavoro degli organi legislativi e di governo. In particolare, il *cluster* 4 sembra più orientato ai valori promossi dalla Costituzione - della quale il Presidente della Repubblica è il garante - e dei termini ad essa associati (Costituzione, cittadino, istituzione, responsabilità, diritto, Repubblica, garanzia, democratico, autonomia), mentre il *cluster* 5 sembra racchudere al suo interno parole che fanno riferimento al funzionamento della Repubblica Italiana, alla sfera legata alle elezioni e, in maniera blanda e in chiave non complottistica, alla cosiddetta politica “di palazzo” (Governo, Parlamento, elettorale, maggioranza, partito, Camera, opposizione, programma, consultazione, Senato, approvare).

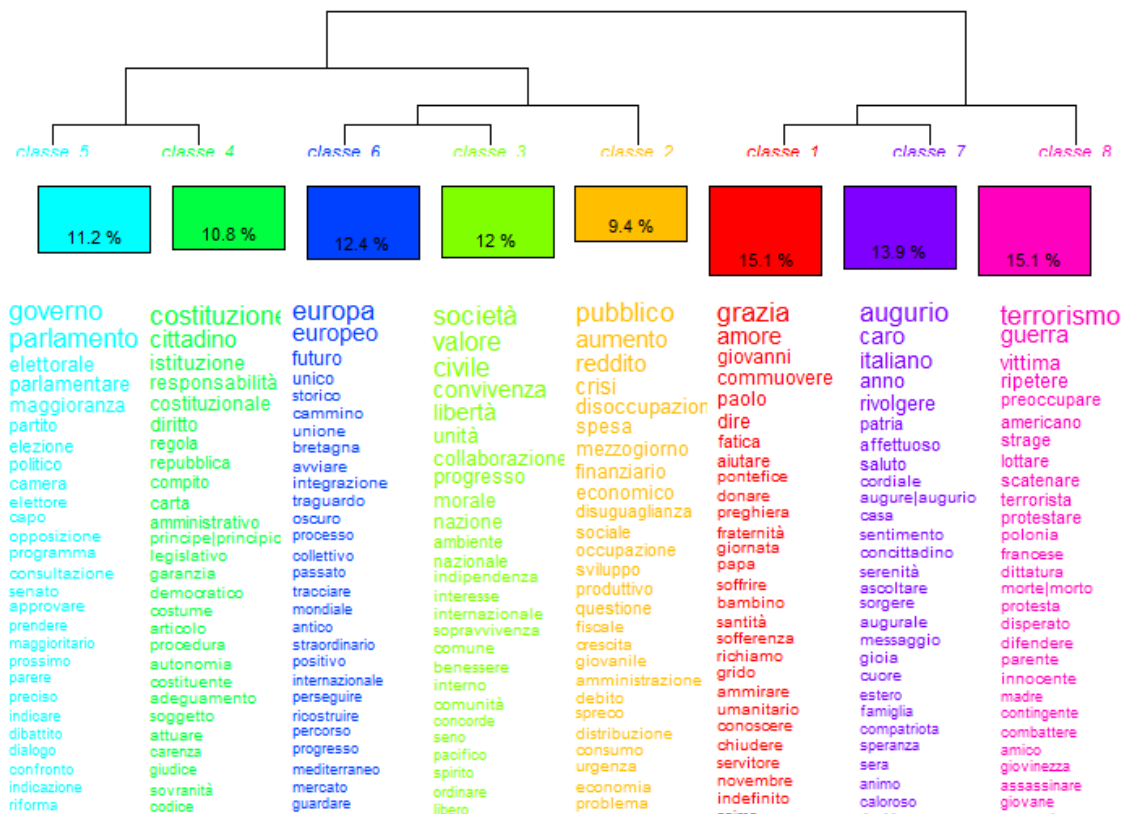


Fig. 25. Dendrogramma della *topic extraction* con le parole più ricorrenti e identificative per ogni *topic*.

Di tematiche radicalmente opposte si occupa il *cluster* numero 2 che, al suo interno, contiene parole legate a tematiche economico-sociali e, in qualche caso, a vere e proprie problematiche sociali (aumento, reddito, crisi, disoccupazione, spesa, Mezzogiorno, economico, disuguaglianza, sociale, occupazione, sviluppo, produttivo, fiscale, giovanile, debito, spreco). Il *cluster* 3 - che, si ricorda, insieme al numero 6 compone la coppia più vicina al *cluster* 2 - tratta anch'esso le tematiche sociali - oltre che morali -, ma osservandole da un punto di vista più legato ai valori e in chiave sostanzialmente più ottimistica e

ideale rispetto alla concretezza dell'ammasso precedente; caratteristica in particolar modo evidenziata dalla presenza di termini quali società, valore, civile, convivenza, libertà, unità, collaborazione, progresso, morale, ambiente, benessere. Inoltre, è anche presente un ristretto numero di parole che si ricollega alla radice di "nazione" (nazione, nazionale, internazionale). Analogamente, nel *cluster* 6, sono presenti chiarissimi riferimenti al concetto di Europa (Europa, europeo, Unione, Bretagna, integrazione, mondiale, internazionale), oltre a una serie di parole legate alla collaborazione futura, alle relazioni internazionali e alla volontà di costruire un percorso comune di concerto con altri Paesi (futuro, cammino, avviare, integrazione, processo, collettivo, tracciare, positivo, ricostruire, percorso, progresso).

L'ultima parte della cluster analysis individua 3 gruppi più slegati da quelli analizzati in precedenza, ma altrettanto distintamente categorizzabili secondo un certo criterio. L'ammasso numero 8 è legato alla tematica del terrorismo e più in generale della violenza, caratteristica che non stupisce, dal momento che molti dei discorsi analizzati riguardano anni in cui lo stragismo nero, rosso e mafioso costituiva una delle piaghe più angoscianti per il Paese. A termini quali terrorismo, guerra, vittima, strage, lottare, terrorista, morte e assassinare, si affiancano parole che vedono la questione dal punto di vista delle vittime e delle famiglie (vittima, preoccupare, disperato, parente, innocente, madre) e, ancora, parole con riferimenti all'estero e a contingenti di guerra transnazionali (americano, Polonia, francese, dittatura,

contingente, combattere). Di tutt'altra composizione sono, invece, i gruppi 1 e 7. Il primo raccoglie un insieme di sentimenti di varia natura, dall'amore alla sofferenza, ed è strettamente legato alla sfera spirituale vicina al cattolicesimo, come messo in rilievo dalla presenza di termini come Giovanni, Paolo, aiutare, Pontefice, donare, preghiera, fraternità, Papa e Santità. Il secondo - il *cluster* numero 7 - è una collezione di termini che sembrano definire in modo molto chiaro gli obiettivi, lo stile comune e la funzione dei messaggi di fine d'anno stessi. Augurio, caro, italiano, anno, rivolgere, patria, affettuoso, saluto, cordiale, casa, sentimento, concittadino, serenità, ascoltare, messaggio, gioia, cuore, famiglia, compatriota e speranza sono tutte parole che definiscono ciò che il discorso di fine anno dei Presidenti della Repubblica è ai loro occhi e, di riflesso, come viene percepito dai cittadini italiani.

Con l'ausilio del piano cartesiano mostrato in Figura 26, è possibile osservare la suddivisione in *cluster* da una prospettiva più lontana e rendersi conto di una caratteristica peculiare del *corpus* esaminato. La figura è il risultato di un'analisi delle corrispondenze che confronta le parole in base alla frequenza nelle diverse classi semantiche (topics) e permette di evidenziare somiglianze e differenze tra classi.

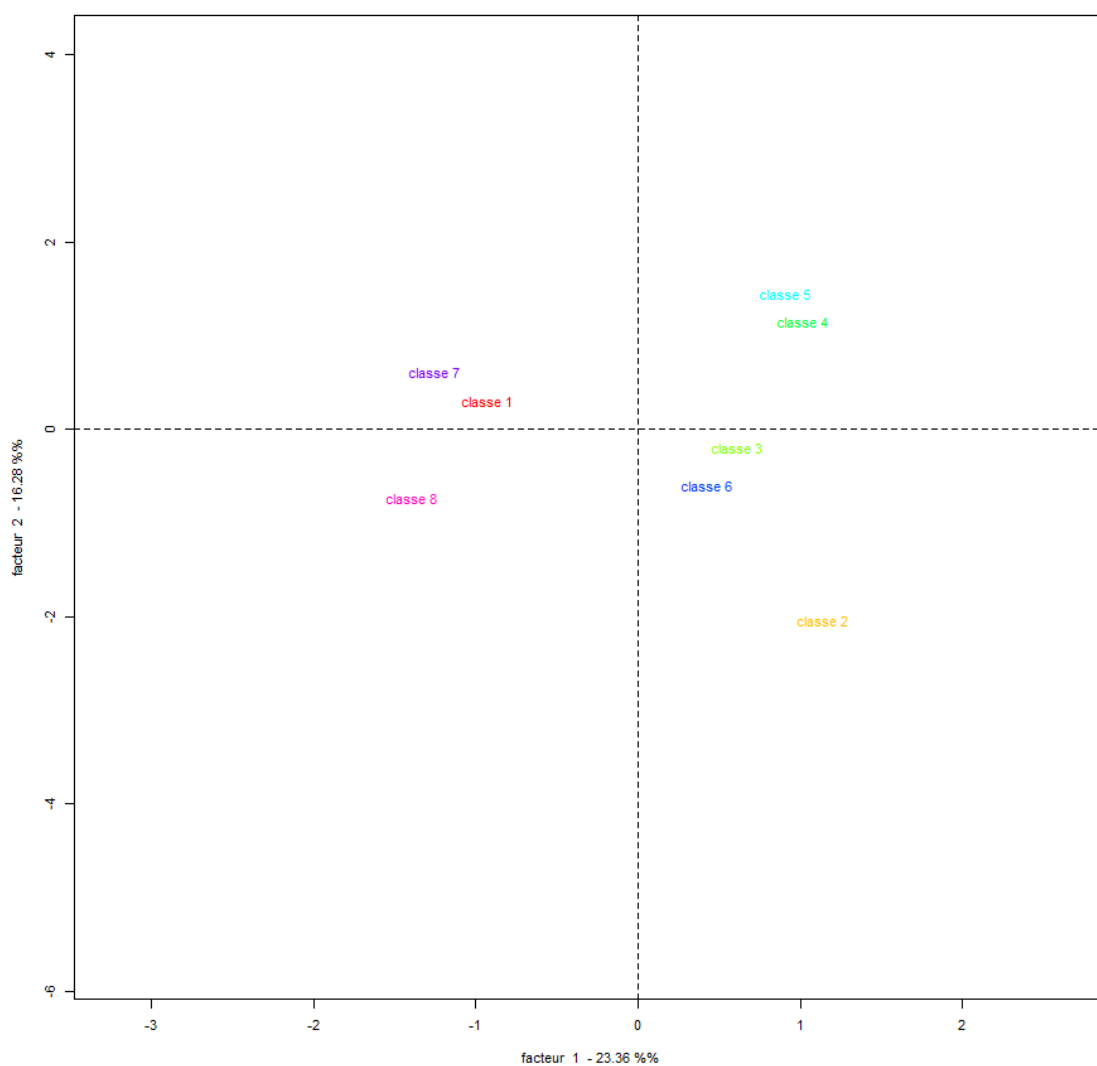


Fig. 26. Disposizione sugli assi cartesiani degli 8 *topics* individuati.

Tre classi (1,7 e 8), formano un gruppo separato dalle restanti cinque (2, 3, 4, 5 e 6); stando a quanto riportato nell'analisi dei termini costitutivi delle classi fatta in precedenza, è ragionevole supporre che la discriminante che separa queste due parti può essere rappresentata dalla sfera emotiva. Mentre nelle prime 5 classi analizzate si trattano aspetti politici, istituzionali, amministrativi, sociali e morali, nelle altre 3

parole individuate si compenetrino tra loro. Mentre gli ammassi 2, 5, 7 e 8 (rispettivamente, quello relativo ai problemi sociali, al funzionamento della Repubblica, ai messaggi di fine d'anno e al terrorismo), gli altri tre gruppi vedono le parole al loro interno diluite in altri *topics*. Un esempio di questo fenomeno riguarda la parola "mercato", contenuta nel gruppo di termini vicini all'Europa, ma pienamente immersa nell'area relativa ai problemi sociali e vicina a molte altre parole afferenti alla sfera economica quali "crescita", "prezzo", "finanziario", "spesa" e "reddito". Dalla parte opposta del piano cartesiano c'è la parola "bambino" - contenuta nella classe vicina all'aspetto religioso - che si colloca vicino a "genitore", "moglie", "fraterno", tutti termini che riportano all'istituzione della famiglia. Ma la parte del grafico che registra la maggior fusione tra due classi è, senza ombra di dubbio, quella che coinvolge il *topic* numero 3 e il *topic* numero 6. Concetti come quelli di identità, passato, ricostruzione, collettività, sostanza e soluzione (tutti appartenenti alla classe 6), sembrano essere legati a doppio filo a termini come valore, unità, società, sopravvivenza, convivenza e progresso (afferenti alla classe 3), quasi ad indicare che il percorso attraverso il quale l'Italia ha costruito la propria identità a partire dalle macerie del dopoguerra ed è progredita concretamente nel corso degli ultimi 70 anni, non è potuto prescindere dai valori sociali costitutivi della Repubblica stessa.

4. GENERATORE DI DISCORSI

4.1. *Rank e discorsi incentrati sulle sfere valoriali*

L'ultima parte della ricerca consiste nella generazione dei tre discorsi basati sulle altrettante sfere valoriali esplorate fino a questo momento. Per procedere con questa operazione è necessario trasferire nuovamente le analisi sul software TaLTaC2, ma prima di farlo è fondamentale far sì che ogni paragrafo del corpus totale che entrerà nel gruppo di frasi che potrebbero essere selezionate per produrre i tre discorsi finali, venga riconosciuta come singolo frammento di testo. Poiché la frammentazione eseguita tramite IRaMuTeQ produce frammenti troppo brevi, spezzettando non soltanto i paragrafi, ma anche le singole frasi e producendo porzioni di frasi lasciate a metà - e, quindi, in cui il senso logico del discorso veniva meno -, si è deciso di procedere con una frammentazione manuale. Nei pochi casi in cui è stato possibile, si è mantenuta la paragrafazione originaria; negli altri casi, si sono suddivisi i testi in modo che ogni frammento costituisse una porzione di testo il più possibile autoconclusiva al suo interno dal punto di vista del contenuto. A partire dai 69 discorsi iniziali, questa operazione ha generato 474 frammenti.

A questo punto, reimportando questo nuovo corpus modellato in maniera da essere funzionale allo scopo ultimo di questa ricerca, si è eseguito un tipo di TFIDF cosiddetto “testuale” che differisce da quello cosiddetto “lessicale” perché dà la possibilità di creare una lista di parole contenute nel corpus, sulle quali calcolare il TFIDF stesso. Questa operazione è nota anche come categorizzazione di frammenti.

Il TFIDF (*Term Frequency Inverse Frequency*) è un indice statistico che consente di verificare la pregnanza di uno o più termini all'interno del testo, tenendo in considerazione non soltanto il numero di occorrenze dei termini in un testo, ma anche la distribuzione di questi all'interno di tutti i subcorpora presi in considerazione: nel caso specifico, in ognuno dei 474 frammenti individuati con la procedura precedente.

Procedendo con la ricerca, si sono create tre liste, ognuna con le parole facenti parte del gruppo di valori oggetto delle analisi precedenti. Per ognuna delle tre liste il risultato è stato un rank di frasi calcolato col TFIDF in cui la prima rappresenta quella con la densità maggiore di valori contenuti in quella lista.

Frammenti con var categoriali e/o ex-post (testuali/quantitative)

ID Frammento	Etichetta del Frammento	Informazioni aggiuntive	TFIDF Valori Sociali
473	69_MATTARELLA_2017_III		1,97668
336	53_CIAMPI_2001_III		1,97583
423	63_NAPOLITANO_2011_VI		1,96956
304	49_SCALFARO_1997_VI		1,95643
▶ 195	33_PERTINI_1981_IV		1,93923
362	57_CIAMPI_2005_VII		1,93363
351	55_CIAMPI_2003_V		1,88911
348	55_CIAMPI_2003_V		1,87937
228	38_COSSIGA_1986_II		1,86673
242	40_COSSIGA_1988_IV		1,84084
257	42_COSSIGA_1990_VI		1,80144
344	54_CIAMPI_2002_IV		1,79297
249	41_COSSIGA_1989_V		1,77934
222	37_COSSIGA_1985_I		1,73941
65	15_SEGNI_1963_II		1,73083
447	66_NAPOLITANO_2014_IX		1,73083
140	23_LEONE_1971_I		1,72414
224	37_COSSIGA_1985_I		1,72079
229	38_COSSIGA_1986_II		1,72078
58	14_SEGNI_1962_I		1,71523
47	12_GRONCHI_1960_VI		1,71304
412	62_NAPOLITANO_2010_V		1,70975
366	57_CIAMPI_2005_VII		1,70643
325	51_CIAMPI_1999_I		1,70606

Fig. 28. Rank di frammenti generato prendendo come riferimento i valori sociali

Frammenti con var categoriali e/o ex-post (testuali/quantitative)				
	ID Frammento	Etichetta del Frammento	Informazioni aggiuntive	TFIDF Valori morali
▶	458	67_MATTARELLA_2015_I		1,71145
	443	66_NAPOLITANO_2014_IX		1,71145
	366	57_CIAMPI_2005_VII		1,66930
	411	62_NAPOLITANO_2010_V		1,66930
	299	48_SCALFARO_1996_V		1,65234
	65	15_SEGNI_1963_II		1,65113
	324	51_CIAMPI_1999_I		1,62644
	228	38_COSSIGA_1986_II		1,62639
	163	29_LEONE_1977_VII		1,61778
	344	54_CIAMPI_2002_IV		1,58217
	278	46_SCALFARO_1994_III		1,52747
	336	53_CIAMPI_2001_III		1,52073
	264	44_SCALFARO_1992_I		1,43865
	257	42_COSSIGA_1990_VI		1,41165
	421	63_NAPOLITANO_2011_VI		1,41165
	149	25_LEONE_1973_III		1,41165
	431	64_NAPOLITANO_2012_VII		1,41165
	400	61_NAPOLITANO_2009_IV		1,41165
	393	60_NAPOLITANO_2008_III		1,41165
	423	63_NAPOLITANO_2011_VI		1,41165
	317	50_SCALFARO_1998_VII		1,41165
	472	69_MATTARELLA_2017_III		1,41165
	377	58_NAPOLITANO_2006_I		1,41165
	323	51_CIAMPI_1999_I		1,41165

Fig. 29. Rank di frammenti generato prendendo come riferimento i valori morali

Frammenti con var categoriali e/o ex-post (testuali/quantitative)			
ID Frammento	Etichetta del Frammento	Informazioni aggiuntive	TFIDF Valori politici
225	37_COSSIGA_1985_I		1,96685
360	57_CIAMPI_2005_VII		1,87841
336	53_CIAMPI_2001_III		1,81819
123	20_SARAGAT_1968_V		1,72492
234	39_COSSIGA_1987_III		1,72292
342	54_CIAMPI_2002_IV		1,70152
322	51_CIAMPI_1999_I		1,70085
329	52_CIAMPI_2000_II		1,70001
136	22_SARAGAT_1970_VII		1,69675
272	45_SCALFARO_1993_II		1,69675
278	46_SCALFARO_1994_III		1,68951
24	08_GRONCHI_1956_II		1,65630
394	60_NAPOLITANO_2008_III		1,65630
328	52_CIAMPI_2000_II		1,64310
257	42_COSSIGA_1990_VI		1,63545
65	15_SEGNI_1963_II		1,63545
116	20_SARAGAT_1968_V		1,63545
340	54_CIAMPI_2002_IV		1,62484
335	53_CIAMPI_2001_III		1,62484
128	21_SARAGAT_1969_VI		1,60170
344	54_CIAMPI_2002_IV		1,54595
49	12_GRONCHI_1960_VI		1,52250
362	57_CIAMPI_2005_VII		1,41416
252	42_COSSIGA_1990_VI		1,41263

Fig. 30. Rank di frammenti generato prendendo come riferimento i valori politici

Una peculiarità che emerge da questi *rank* riguarda la presenza, in tutti e tre, del frammento numero 336 relativo a Carlo Azeglio Ciampi. In quel frammento, il Presidente livornese racconta di aver girato l'Italia in veste di Capo dello Stato e, nel farlo, rammenta i valori costitutivi della Repubblica Italiana. Non è strano, dunque, che in un lavoro di selezione semi-automatica di testi in cui sia presente un alto tasso di contenuti

valoriali, questo paragrafo specifico compaia nelle fasce più alte di tutti i rank.

Come criterio per la selezione del numero di paragrafi con cui comporre un discorso si è deciso di sceglierne, seguendo l'ordine del *rank*, tanti quanti permettevano di arrivare a comporre un discorso di circa 2500 parole; eliminando i discorsi dei primi Presidenti, pronunciati quando la consuetudine si era appena costituita e quindi era ancora in fase di elaborazione, questa è la lunghezza media dei messaggi di fine anno degli ultimi 50 anni. Di conseguenza, nel discorso che contiene il tasso più elevato di valori morali, il frammento 336 relativo a Ciampi non è stato inserito, perché in una posizione del rank più bassa rispetto al numero di frammenti necessari per la composizione del discorso stesso.

Un'altra precisazione riguardo ai discorsi generati riguarda le frasi che costituiscono il saluto iniziale e il commiato finale di ogni discorso: si precisa che la selezione di queste, ininfluenti ai fini dell'analisi perché prive di alcun valore, è stata effettuata manualmente.

Seguono, dunque, i tre discorsi generati secondo la procedura sopra descritta.

Discorso incentrato sui valori sociali:

“Care concittadine e cari concittadini, un saluto cordiale e un grande augurio. A tutti coloro che sono in Italia e agli italiani che si trovano all' estero.

La velocità delle innovazioni è incalzante; e ci conduce in una nuova era, che già cominciamo a vivere. Un'era che pone anche interrogativi sul rapporto tra l'uomo, lo sviluppo e la natura. Basti pensare alle conseguenze dei mutamenti

climatici, come la siccità, la limitata disponibilità di acqua, gli incendi devastanti. Si manifesta, a questo riguardo, una sensibilità crescente, che ha ricevuto impulso anche dal magistero di Papa Francesco, al quale rivolgo gli auguri più fervidi. Cambiano gli stili di vita, i consumi, i linguaggi. Mutano i mestieri, e la organizzazione della produzione. Scompaiono alcune professioni; altre ne appaiono. In questo tempo, la parola "futuro" può anche evocare incertezza e preoccupazione. Non è stato sempre così. Le scoperte scientifiche, la evoluzione della tecnica, nella storia, hanno accompagnato un'idea positiva di progresso. I cambiamenti, tuttavia, vanno governati per evitare che possano produrre ingiustizie e creare nuove marginalità. L'autentica missione della politica consiste, proprio, nella capacità di misurarsi con queste novità, guidando i processi di mutamento. Per rendere più giusta e sostenibile la nuova stagione che si apre. La cassetta degli attrezzi, per riuscire in questo lavoro, è la nostra costituzione: ci indica la responsabilità nei confronti della repubblica e ci sollecita a riconoscerci comunità di vita. L'orizzonte del futuro costituisce, quindi, il vero oggetto dell'imminente confronto elettorale. Il dovere di proposte adeguate - proposte realistiche e concrete - è fortemente richiesto dalla dimensione dei problemi del nostro Paese. Non è mio compito formulare indicazioni. Mi limito a sottolineare, ancora una volta, che il lavoro resta la prima, e la più grave, questione sociale. Anzitutto per i giovani, ma non soltanto per loro. È necessario che ve ne sia in ogni famiglia. Al tempo stesso va garantita la tutela dei diritti e la sicurezza, per tutti coloro che lavorano.

Nei due anni e mezzo trascorsi dalla mia elezione ho compiuto un primo viaggio in Italia. Ne ho visitato oramai quasi tutte le regioni. Continuerò, di provincia in provincia. È un viaggio bellissimo: ne traggio vigore, fiducia, orgoglio sempre più forte di essere italiano. Ovunque avverto, nella ricchezza delle diversità delle nostre contrade, quel sapore d'Italia che viaggiatori del presente e del passato hanno sempre avvertito, che è natura, arte, lingua, cultura, modo di vita. Le radici dell'italianità sono antiche. È antica la nostra nazione. Ma le origini del nostro Stato sono assai più vicine. Risalgono all'inizio dell'ottocento, allorché uno stuolo di uomini di pensiero, poeti, letterati, filosofi, economisti, mossi da un grande amore per l'Italia, animati da un profondo senso etico, da grandi ideali e principi, diventarono anche uomini d'azione, e uomini di Stato. Quel movimento si pose chiari obiettivi: libertà; unità; indipendenza della patria,

dell'Italia. Si diffuse e fu vissuto con intensa passione civile. Si nutrì della consapevolezza delle radici profonde della nostra storia, della nostra civiltà. Non a caso fu chiamato Risorgimento. L'inno di Mameli divenne l'inno della nazione italiana, l'inno del risveglio di un popolo. I grandi del Risorgimento non fecero sogni di conquista. Sognarono l'unità e la libertà d'Italia, e l'indipendenza di tutti i popoli. Vi è continuità fra gli ideali del Risorgimento e la Costituzione repubblicana, che l'Italia si è data dopo avere riconquistato la libertà con la Resistenza. Così come vi è continuità con la costruzione di un'Unione Europea che sia una federazione di Stati-nazione. Sono imprese grandi. Con esse noi, eredi dei padri fondatori dell'Italia e dell'Europa, dobbiamo confrontarci. L'Italia è sempre stata ed intende rimanere all'avanguardia nell'integrazione europea. Non possiamo sfuggire alle sfide che la storia del ventesimo secolo ci propone. Per preparare le nuove generazioni ad affrontarle bene, accanto alla famiglia, che è l'istituzione base della nostra società, deve operare una scuola capace di svolgere, con rinnovato impegno, il suo ruolo insostituibile di servizio pubblico: una scuola volta a formare i giovani, a prepararli ad assolvere responsabilmente i loro compiti di cittadini, e a favorire il loro inserimento, operoso e creativo, in una società che cambia ed avanza con tempi sempre più rapidi. Ci guidano alcuni principi, che uniscono gli italiani, al di là delle diversità d'idee politiche. Siamo una democrazia parlamentare. Chi ha avuto la maggioranza, abbia modo, governando, di dimostrare quanto vale, quanto sa fare per il progresso del nostro popolo. Chi è minoranza eserciti con impegno e responsabilità il compito indispensabile dell'opposizione: di controllo, di critica, di proposta.

Il dialogo fra le due parti, per essere costruttivo, presuppone che nella maggioranza la disponibilità all'ascolto, attento e aperto, della voce dell'opposizione, prevalga sulla tentazione di affidarsi sbrigativamente al rapporto di forza parlamentare; e che nell'opposizione la consapevolezza del diritto del Governo di portare avanti il proprio programma prevalga sulla tentazione del ricorso sistematico all'ostruzionismo. Una democrazia funziona bene se ciascuna istituzione esercita il proprio compito rispettando i limiti delle proprie competenze. La separazione dei poteri, il giudizio della Corte Costituzionale sulla costituzionalità delle leggi, la soggezione dei giudici esclusivamente alla legge, la neutralità e l'imparzialità delle pubbliche

amministrazioni, garantiscono la libertà di tutti i cittadini. Il passaggio di funzioni dal Governo centrale alle autorità di Governo regionali e locali avvicina le istituzioni ai cittadini, valorizzando le autonomie. Questo passaggio deve avvenire razionalmente, al fine di rafforzare, non indebolire, l'unità nazionale. La Repubblica è una e indivisibile. Nel nostro ordinamento, il Presidente della Repubblica non ha, fra i suoi compiti, quello di governare. Egli rappresenta l'unità nazionale; vigila ed opera perché siano rispettati i principi costituzionali; ha il diritto-dovere di consigliare. Avverto tutta la responsabilità di rappresentarvi. Come guida, ho la Costituzione; le nostre tradizioni democratiche; il giuramento prestato dinanzi ai rappresentanti eletti della nazione; la mia coscienza.

Ci si pongono dunque acute necessità di scelte immediate e di visioni lungimiranti. Occorre una nuova forza motivante perché si sprigioni e operi la volontà collettiva indispensabile; occorrono coraggio civile e sguardo rivolto con speranza fondata verso il futuro. Questo ci hanno detto nei giorni natalizi alte voci spirituali. Esse si sono in effetti rivolte al più vasto mondo in cui si collocano i travagli della nostra Italia e della nostra Europa. Un mondo nel quale sono emerse di recente nuove correnti e forze portatrici di aspirazioni alla libertà e alla giustizia, ma anche difficoltà e tensioni, e ancora feroci repressioni. Mentre restano aperti antichi focolai di contrapposizione e di conflitto, e si manifestano ciechi furori religiosi, fino a dar luogo a orribili stragi di comunità cristiane. L' Italia non può restare, e non resta, estranea a ogni possibile iniziativa di pace e umanitaria: come dice la nostra partecipazione - anche con dolorosi sacrifici di giovani vite - a quelle missioni militari e civili internazionali che vedono migliaia di nostri connazionali farsi onore. Nel salutarli e ascoltarli in occasione del Natale, ho colto accenti confortanti di alto senso di responsabilità e di forte vocazione al servizio del bene comune. Sono accenti che colgo, qui in Italia, in tante occasioni di incontro con le molteplici espressioni dell'universo della solidarietà, del volontariato, dell' impegno civile. Sono accenti che trovo in lettere toccanti che mi vengono indirizzate da persone anziane, da giovani e ragazzi, da uomini e donne che raccontano i loro propositi operosi e le loro esperienze. Lasciatemi dunque ripetere: la fiducia in noi stessi è il solido fondamento su cui possiamo costruire, con spirito di coesione, con senso dello stare insieme di fronte alle difficoltà, dello stare

insieme nella comunità nazionale come nella famiglia. E allora apriamoci così al nuovo anno: facciamone una grande occasione, un grande banco di prova, per il cambiamento e il nuovo balzo in avanti di cui ha bisogno l'Italia.

Ecco, nello sguardo che noi diamo all'indietro per qualche momento, mi consentite di fermarmi su un fatto, che io incastonò a questo punto della mia chiacchierata, perché è un fatto che, nel mondo, in tanti Paesi del mondo, io l'ho trovato ovunque, dà veramente prestigio all'Italia: il volontariato, eroico, con dei caduti per la pace. con persone giovani e non giovani che riescono a scoprire delle capacità, delle possibilità di servire la gente, di aiutarla. Incredibili. È vero che il volontariato nasce di dentro, è una vocazione di dentro. Grazie. Gran splendida pagina. In Italia, non c'è dubbio, ma nel mondo. Quante volte Capi di Stato mi hanno detto: qui vostre persone, nei vari settori, svolgono dei compiti di loro iniziativa, non chiedono nulla. Ecco, su questo, vorrei dire, lo Stato ha qualche dovere, il dovere di tutelarlo il volontariato. Attenzione, desidero essere chiaro. Ci sono delle iniziative che sono anche ottime, che servono, che ottengono dei risultati. Ma non possono chiamarsi volontariato, perché c'è una retribuzione o perché si è sistemata tutta la famiglia o perché si sono sistemate una serie di persone, amici o conoscenti, anche capaci. Questo non è volontariato. Il crisma, il marchio del volontariato, si chiama gratuità. E lo Stato interessandosi del volontariato difenda questo che è un pregio altissimo. Se dico difenda è perché so che ogni tanto si presentano proposte diverse sotto il manto di voci varie, ma nel contenuto rompono questa gratuità. Beh, non è volontariato. Adesso, diamo uno sguardo avanti. Il primo sguardo, se mi consentite, è un antico mio pensiero, ancora da Ministro dell' interno: diamo uno sguardo al Mediterraneo. Questo mare, che è ricchezza di civiltà millenarie, civiltà - plurale - millenarie! Incrocio, scambi di pensieri, di arte, di cultura. Scambi sul piano commerciale. In questo mare, oggi, quanto sangue, quanta tragedia, quanta! Lo so che quando si tocca questo tema, si teme, da parte di vari Stati, o di qualche Stato in particolare, che si voglia andare a mettere il naso nelle faccende loro. No, no! Però devo dire, con molta chiarezza, che l'Europa, la comunità europea non può stare a guardare, non può, mentre si massacrano a centinaia, a migliaia le persone. Non è civile, non è umano, non è pensabile. Si tratta di trovare una strada che non sia quella di infilare il naso dove non si ha diritto di metterlo, ma anche di impedire di stare a guardare.

Questo mare, ancora l'altro giorno, è servito perché arrivassero i Curdi; e quante volte persone di colore di varie parti. E questi aguzzini, è un termine assolutamente benevolo, questi manigoldi, che speculano su questa gente, poi fanno incagliare la nave e scappano con i soldi, questa criminalità incredibile. Perché i Paesi non trovano un punto comune per trovare la strada per punire questa criminalità, questa speculazione sulla povera gente? L' Italia allarga le braccia. Lo so che i temi sono molti e delicati. Mi pare che il Ministero dell'interno abbia trovato un punto di equilibrio. È sempre difficile, non vi è dubbio alcuno, ma se le persone vengono perché vivono una persecuzione: beh, allora le porte devono essere spalancate. Anche i nostri che lottavano contro la dittatura hanno trovato le porte aperte. Guai a chiudere le porte a chi è perseguitato, a chi fugge per poter sopravvivere. Non è umano! Ma, attenzione, attenzione a un certo tipo di politica che per punire degli Stati che hanno l'imputazione di non rispettare i diritti umani propongono in fondo l'isolamento. Attenzione! Non credo a questa politica. Da nessuna parte avvenga! Non credo. Non credo che faccia bene all'Unione Europea se a un certo momento con questa motivazione... non spezziamo mai il dialogo. Non spezziamolo mai! Ed eccoci allora a guardare i problemi in prospettiva. Su questo del Mediterraneo aggiungerò solo una cosa: che l'Europa senza il Mediterraneo è monca. Bisogna che i Paesi dell'Europa pensino a questo: è monca! Non è solo un problema dei Paesi che vi si affacciano. No. Io sento di dire una parola di gratitudine al Presidente Herzog, Presidente della Germania. Mi disse, quando ho fatto la visita di Stato: Scalfaro, ogni qualvolta lei parla di Mediterraneo, io sono con lei. E sento e ho sentito questo appoggio.

Vi è poi, si è aggiunta a tutte queste preoccupazioni, italiane e italiani, la questione della P2. Mi si intenda bene perché non voglio che ancora una volta il mio pensiero sia travisato. Quando io parlo della P2 non intendo coinvolgere nella P2 la massoneria propriamente detta, con la sua tradizione storica. Per me almeno, una cosa è la massoneria, che non è in discussione, un'altra cosa è la P2, questa P2 che ha turbato, inquinato la nostra vita. I giuristi stanno discutendo se la P2 cada o non cada sotto il codice penale, se è un'associazione a delinquere. Sono cose che a me non interessano per il momento. Io guardo ad un altro codice, che è il codice morale, il codice morale che ogni uomo, specialmente ogni uomo politico, dovrebbe portare scritto nella

sua coscienza. Ebbene, la P2 cade sotto questo codice morale. Vi è un proverbio che si usa dire: che la moglie di Cesare non deve essere sospettata. Ma prima di tutto è Cesare che non deve essere sospettato. Ed allora ogni sospetto devono allontanare dalla loro persona gli uomini politici, non possono rimanere, non può rimanere al suo posto chi è stato indiziato in questa trappola della P2. La P2 si prefiggeva di compiere atti contro la Costituzione, contro la democrazia e contro la Repubblica. E quindi coloro che facevano parte della P2 dovranno risponderne prima di tutto dinanzi alla loro coscienza, dinanzi ai loro partiti e, soprattutto, dinanzi al Parlamento. Non vi può essere in questo caso alcuna comprensione ed alcuna solidarietà. E ripeto quello che ho detto altre volte: qui le solidarietà personali, le solidarietà di partito, diventano complicità. Eppure nonostante queste preoccupazioni le speranze non si sono spente nel mio animo, italiane e italiani. Io credo nella rinascita del nostro popolo. Credo che il nostro popolo conoscerà giorni più sereni di quelli di oggi, potrà avere lavoro ogni cittadino, ed ogni famiglia potrà finalmente avere una casa. Mi dicono leggendo le statistiche, con mia sorpresa ho constatato che 300 mila coppie di sposi sono senza casa. E questo è veramente una cosa grave che non sia stato risolto questo problema, ad esempio, della casa, che non sia stato risolto durante il cosiddetto miracolo economico, è cosa che mi stupisce, e forse anch'io, siccome ho sempre fatto parte della classe politica, della classe dirigente, assumo anch'io la mia parte di responsabilità. Ma ripeto io guardo ancora al domani, pur essendo il mio animo angosciato, io guardo ancora all'avvenire del popolo italiano con speranza e con fiducia, riuscirà questo popolo nostro a rialzarsi, riusciremo a rialzarci, italiane e italiani. Perché io credo nel popolo italiano. È un popolo buono, generoso, laborioso, non chiede che lavoro, una casa e di poter curare la salute dei suoi cari. Non chiede quindi il paradiso in terra. Chiede quello che dovrebbe avere ogni popolo, ogni popolo civile, ogni popolo che ha raggiunto un certo progresso, come lo ha raggiunto il popolo italiano. Io credo nel popolo italiano. Vedete, italiane e italiani, io ho fatto parecchi viaggi. Ebbene, in tutte queste nazioni dove sono stato io ho incontrato degli italiani, laggiù, che purtroppo hanno dovuto lasciare l'Italia per andare a trovare all'estero il lavoro che non trovavano qui nel nostro Paese. Ho parlato con i dirigenti di quei Paesi e questi dirigenti mi hanno fatto sempre le lodi dei nostri connazionali. Hanno detto: sono dei lavoratori che amano

lavorare, che amano la casa, che appena vengono qui cercano un posto di lavoro e cercano la casa, poi, il modo di poter educare e dare una cultura ai loro figlioli. Abbiamo la lode degli stranieri, e talvolta noi ci autoflagelliamo invece. No, il nostro popolo non è superiore agli altri popoli, ma non è neppure inferiore agli altri popoli.

Vorrei rivolgere, in chiusura, un saluto a quanti, questa sera, non stanno festeggiando perché impegnati ad assolvere compiti e servizi essenziali per tutti noi: sulle strade, negli ospedali, nelle città, per garantire sicurezza, soccorso, informazione, sollievo dalla sofferenza. A loro, ringraziandoli, esprimo un augurio particolare. Auguri a tutti; e buon anno”.

Discorso incentrato sui valori morali:

“Care italiane cari italiani,

nell'anno che sta per aprirsi si svolgerà il maggior percorso del Giubileo della Misericordia, voluto da Francesco, al quale rivolgo i miei auguri ed esprimo riconoscenza per l'alto valore del suo magistero.

È un messaggio forte che invita alla convivenza pacifica e alla difesa della dignità di ogni persona.

Con una espressione laica potremmo tradurre quel messaggio in comprensione reciproca, un atteggiamento che spero si diffonda molto nel nostro vivere insieme.

Sappiamo tutti che quando si parla di noi italiani le prime parole che vengono in mente sono genio, bellezza, buon gusto, inventiva, creatività. Sappiamo anche che spesso vengono seguite da altre, non altrettanto positive: scarso senso civico, particolarismo, individualismo eccessivo accentuato.

Ricevo ogni giorno molte lettere e, in questo mio primo anno di presidenza, in giro per l'Italia e al Quirinale, ho incontrato tante persone e conosciuto le loro storie. parlano di coraggio, di impegno, di spirito d'impresa, di dedizione agli altri, di senso del dovere e del bene comune, di capacità professionali, di eccellenza nella ricerca. E non si tratta di eccezioni. Nei miei colloqui con i

rappresentanti di altri Paesi, in Italia e all'estero, ho sempre colto una considerazione e una fiducia nei confronti dell'Italia e degli italiani maggiori di quanto, a volte, noi stessi siamo disposti a riconoscere.

Il messaggio augurale di fine d'anno che ormai dal 2006 rivolgo agli italiani, presenterà questa volta qualche tratto speciale e un po' diverso rispetto al passato. Innanzitutto perché le mie riflessioni avranno per destinatario anche chi presto mi succederà nelle funzioni di Presidente della Repubblica. Funzioni che sto per lasciare, rassegnando le dimissioni: ipotesi che la Costituzione prevede espressamente. E desidero dirvi subito che a ciò mi spinge l'aver negli ultimi tempi toccato con mano come l'età da me raggiunta porti con sé crescenti limitazioni e difficoltà nell'esercizio dei compiti istituzionali, complessi e altamente impegnativi, nonché del ruolo di rappresentanza internazionale, affidati dai padri costituenti al Capo dello Stato. A quanti auspicano - anche per fiducia e affetto nei miei confronti - che continui nel mio impegno, come largamente richiestomi nell'aprile 2013, dico semplicemente che ho il dovere di non sottovalutare i segni dell'affaticamento e le incognite che essi racchiudono, e dunque di non esitare a trarne le conseguenze. Ritengo di non poter oltre ricoprire la carica cui fui chiamato, per la prima volta nel maggio del 2006, dal Parlamento in seduta comune. Secondo l'opinione largamente prevalente tra gli studiosi, si tratta di una valutazione e di una decisione per loro natura personali, costituzionalmente rimesse al solo Presidente, e tali da non condizionare in alcun modo Governo e Parlamento nelle scelte che hanno dinanzi né subendone alcun condizionamento. Penso che questi semplici chiarimenti possano costituire una buona premessa perché Parlamento e forze politiche si preparino serenamente alla prova dell'elezione del nuovo Capo dello Stato. Sarà quella una prova di maturità e responsabilità nell'interesse del Paese, anche in quanto è destinata a chiudere la parentesi di una eccezionalità costituzionale. Personalmente resto convinto che la disponibilità richiestami e offerta nell'aprile 2013, in un momento di grave sbandamento e difficoltà post-elettorale, sia risultata un passaggio determinante per dare un Governo all'Italia, rendere possibile l'avvio della nuova legislatura e favorire un confronto più costruttivo fra opposti schieramenti politici. Ma è positivo che ora si torni, per un aspetto così rilevante, alla normalità costituzionale, ovvero alla regolarità dei tempi di vita delle istituzioni, compresa la Presidenza della Repubblica. L'aver

tenuto in piedi la legislatura apertasi con le elezioni di quasi due anni fa, è stato di per sé un risultato importante: si sono superati poi momenti di acuta tensione, imprevisti, alti e bassi nelle vicende di maggioranza e di Governo; si è in sostanza evitato di confermare quell'immagine di un'Italia instabile che tanto ci penalizza, e si è messo in moto, nonostante la rottura del febbraio scorso, l'annunciato, indispensabile processo di cambiamento.

Un anno fa, nel messaggio del 31 dicembre, avevo detto: "Spero di poter vedere nel 2014 almeno iniziata un'incisiva riforma delle istituzioni repubblicane". Ebbene, è innegabile che quell'auspicio si sia realizzato. E il percorso va, senza battute d'arresto, portato a piena conclusione. Non occorre che io ripeta ora - l'ho fatto ancora di recente in altra pubblica occasione - le ragioni dell'importanza della riforma del Parlamento, e innanzitutto del superamento del bicameralismo paritario, nonché della revisione del rapporto tra Stato e Regioni. Ma sul necessario più vasto programma di riforme - istituzionali e socio-economiche - messo in cantiere dal Governo, sulle difficoltà politiche che ne insidiano l'attuazione, sulle possibilità di dialogo e chiarimento con forze esterne alla maggioranza di Governo - anche, s'intende, e in via prioritaria, per il varo di una nuova legge elettorale - non torno ora avendovi già dedicato largamente il mio intervento, due settimane fa, all'incontro di fine anno con i rappresentanti delle istituzioni, delle forze politiche e della società civile. Vorrei piuttosto ragionare con voi su come stiamo vivendo questo momento in quanto generalità dei cittadini, uniti dall'essere italiani.

Ma sento ancora una volta il dovere, il bisogno di rivolgermi ai giovani. Siete il nostro domani. La nostra speranza. La mia generazione si è impegnata nel salvaguardare e trasmettervi lo spirito che ci animò all'indomani di una guerra orrenda. Lo spirito che ci diede la forza di ricostruire le nostre città, di dar vita alle istituzioni di libertà che contraddistinguono la Repubblica Italiana, e l'Unione Europea, che abbiamo creato insieme con altri popoli. Dai tanti incontri che ho avuto con voi ho tratto motivi di fiducia nell'avvenire della nostra Italia. So quanto siate impegnati nel prepararvi ad affrontare le sfide del futuro, insieme con i giovani di altri popoli, che condividono le vostre aspirazioni di progresso, di giustizia, di pace. La pace: mi sono rimaste impresse le parole rivoltemi da una bambina nella piazza di Corleone: la pace ti nasce dal cuore e si diffonde nell'aria. Preservate i valori della nostra civiltà, che non soggiacciono

al mutare delle mode. Primo fra essi l'amore per la famiglia, nucleo fondamentale della società, punto sicuro di riferimento per ciascuno di noi. Siete nati, vivete in un'Europa di pace, di libertà. Tenete alti, e diffondete nel mondo, i suoi ideali. Toccherà a voi completarne e rafforzarne le istituzioni. Per tutti gli europei non c'è un domani se non in un'Unione Europea più coesa. Questi sono i sentimenti, le riflessioni che, nell'approssimarsi del congedo, affollano il mio animo. Li affido a voi che mi ascoltate.

Gli ultimi dati ci dicono che le persone in cerca di occupazione sono tornate a superare i due milioni, di cui quasi uno nel Mezzogiorno; e che il tasso di disoccupazione nella fascia di età tra i 15 anni e i 24 - ecco di nuovo il discorso sui giovani, nel suo aspetto più drammatico - ha raggiunto il 24,7 per cento nel Paese, il 35,2 nel Mezzogiorno e ancor più tra le giovani donne. Sono dati che debbono diventare l'assillo comune della nazione. Se non apriamo a questi ragazzi nuove possibilità di occupazione e di vita dignitose, nuove opportunità di affermazione sociale, la partita del futuro è persa non solo per loro, ma per tutti, per l'Italia: ed è in scacco la democrazia. Proprio perché non solo speriamo, ma crediamo nell'Italia, e vogliamo che ci credano le nuove generazioni, non possiamo consentirci il lusso di discorsi rassicuranti, di rappresentazioni convenzionali del nostro lieto vivere collettivo. C'è troppa difficoltà di vita quotidiana in diverse sfere sociali, troppo malessere tra i giovani. Abbiamo bisogno di non nasconderci nessuno dei problemi e delle dure prove da affrontare: proprio per poter suscitare un vasto moto di energie e di volontà, capace di mettere a frutto tradizioni, risorse e potenzialità di cui siamo ricchi. Quelle che abbiamo accumulato nella nostra storia di centocinquanta anni di Italia unita. Celebrare quell'anniversario, come abbiamo cominciato a fare e ancor più faremo nel 2011, non è perciò un rito retorico. Non possiamo come nazione pensare il futuro senza memoria e coscienza del passato. Ci serve, ci aiuta, ripercorrere nelle sue asprezze e contraddizioni il cammino che ci portò nel 1861 a diventare Stato nazionale unitario, ed egualmente il cammino che abbiamo successivamente battuto, anche fra tragedie sanguinose ed eventi altamente drammatici. Vogliamo e possiamo recuperare innanzitutto la generosità e la grandezza del moto unitario: e penso in particolare a una sua componente decisiva, quella dei volontari. Quanti furono i giovani e giovanissimi combattenti ed eroi che risposero, anche

sacrificando la vita, a quegli appelli per la libertà e l'unità dell'Italia! Dovremmo forse tacerne, e rinunciare a trarne ispirazione? Ma quello resta un patrimonio vivo, cui ben si può attingere per ricavarne fiducia nelle virtù degli italiani, nel loro senso del dovere comune e dell'unità, e nella forza degli ideali.

Prima degli auguri, ancora un'osservazione, non secondaria. Parlando al corpo diplomatico, per i tradizionali auguri, ho detto tra l'altro: la politica è anzitutto pensiero, illuminato da ideali che indicano una meta; è pensiero che genera e sostiene un progetto di azioni, dà vita a un programma e a un'opera di convincimento per ottenere adesioni a quel progetto e a ciò che lo anima. E ancora: la politica fa sintesi tra volontà, azione, risposta a desideri, a speranze e indica vie strategiche e ne accompagna i passi esecutivi. Per questo, la politica ha bisogno di un'anima: il mondo attuale soffre per una grave mortificazione dei valori dello spirito. Allora la politica non può esprimersi in frasi volgari, in denigrazioni personali, insinuazioni, calunnie, veleni. Tutto questo non diventerà mai politica, mai! Questo produce confusione, crea sconcerto, rende faticosi i rapporti umani. Parliamo, discutiamo, intrecciamo idee, intrecciamo idee e proposte se vogliamo fare politica. Liberiamoci dall'intreccio tra politica e affari: è il peggiore intreccio che è stato in anni recenti e rimane sempre apportatore di danni gravi. Gli affari leciti hanno diritto a tutto lo spazio necessario, ma non assurgono, di per sé, a dignità politica, perché non possono confondersi con l'interesse generale. Mescolare le cose vuol dire far uscire la politica dal suo alveo, imbastardirla, degradarla. Anche questo aumenta nei cittadini il senso dell'insicurezza, quasi un timore generico, una preoccupazione di un pericolo indefinito, ma capace di portare pena e turbamento. Penso che questo timore del domani, che fa sentire incerto il presente e aumenta le ombre sul futuro, che non aiuta a spingere lo sguardo fiducioso in avanti, questo timore del domani, penso, possa anche essere vinto con un intelligente sforzo di collaborazione per una realtà pacata, per creare una convivenza non conflittuale, non esasperata, non avvelenata. Qui vi è anche l'essenziale compito dei mezzi di comunicazione: occorre libertà, occorre dovere di informare, occorre rispetto della verità e rispetto della dignità della persona. Aumentiamo il nostro senso di responsabilità nelle diversità del pensiero, delle valutazioni, delle proposte e dei progetti; diamo prova di saper dominare le situazioni con saggezza e con forza e non di esserne dominati. Anche questo è

impegno ed è augurio. Soprattutto diventa augurio! Buon anno! a ogni famiglia, a chi è solo, a chi soffre. Auguri perché si realizzi la speranza di chi attende lavoro, la speranza di chi attende giustizia, serena ed oggettiva, la speranza di chi attende clemenza. Buon anno a ogni cittadino, a ogni straniero che vive e opera tra noi. Buon anno a tutti gli italiani nel mondo. Un buon anno particolare ai bambini, a tutti i bambini; che la loro innocenza non sia mai più ferita da tante violenze, che la loro innocenza non debba mai più soffrire per mancanza di amore. Nei diversi compiti, nelle distinte responsabilità, lavoriamo insieme perché l'Italia superi e vinca difficoltà e ostacoli. Insieme, perché l'Italia è di tutti e la sua libertà, la sua unità, la sua giustizia, il suo benessere, il suo domani devono stare a cuore a tutti!

Un altro recente avvenimento ha avuto larghe e profonde ripercussioni: intendo riferirmi al Trattato di Mosca per la limitazione degli esperimenti nucleari, al quale il nostro Paese ha aderito e che mi auguro possa essere apportatore di ulteriori, benefici frutti. L' Italia, da parte sua, ha continuato, intanto, a svolgere - con fiducia - opera di buona volontà, di pace, di civile progresso in seno al concerto delle nazioni. Sempre conscia dei suoi diritti e ferma nel suo impegno di difesa della propria indipendenza e della propria libertà da ogni minaccia, essa, insieme con le altre nazioni alle quali la uniscono le necessità della comune sicurezza e gli ideali comuni, ha continuato ad adoperarsi con tutte le sue forze - e secondo la sua naturale vocazione - per una sempre maggiore comprensione fra i popoli d'ogni razza e d'ogni ideologia, indipendentemente da ogni divergenza di interessi.

Ed è confortante constatare come il nostro contributo in seno a tutti i consessi internazionali, dalle Nazioni Unite alla Nato, dalla conferenza del disarmo alle organizzazioni europee, abbia conseguito non pochi riconoscimenti ed il nostro Paese abbia saputo acquistarsi ovunque rispetto, fiducia e simpatia. I numerosi incontri che ho avuto a Roma con Capi di Stato, Capi di Governo ed autorevoli personalità di tutto il mondo costituiscono gradita conferma dell'importanza che essi attribuiscono all'Italia nel concerto delle nazioni. Non meno significative sono state le visite da me compiute in Marocco, prima, e in Germania, poi: visite che, mentre hanno contribuito a rinsaldare i nostri vincoli di cordiale amicizia già esistenti con quei popoli, mi hanno consentito, altresì, di

raccogliere - con animo commosso - il vibrante saluto di quelle comunità italiane alla patria lontana.

Della disoccupazione, delle disuguaglianze di sviluppo soffrite innanzitutto voi giovani. Ne soffre soprattutto il Mezzogiorno. La mia generazione, la generazione dei vostri padri, avverte il disagio, sente la responsabilità di questi perduranti squilibri. L' articolo primo della Costituzione recita: l'Italia è una Repubblica democratica, fondata sul lavoro. Da troppi anni il numero dei senza lavoro supera l'11% di coloro che vogliono lavorare. Ancor più doloroso, il 30% dei giovani che cercano lavoro non lo trovano. Ci danno speranza i segni recenti di aumento dell'occupazione. Sappiamo che il lavoro durevole si crea innovando, stando al passo di un mondo che muta ed avanza con tempi sempre più rapidi. Sta agli imprenditori, alla loro capacità progettuale di esprimere e attuare nuove iniziative. Ma lo Stato deve assicurare le condizioni, in primo luogo quelle della sicurezza, affinché gli imprenditori possano operare con fiducia. Lo sviluppo economico è sempre più fondato sulla conoscenza. Il capitale principale è quello umano. Voi giovani ne siete l'essenza. È nell'interesse di tutti accrescere la vostra professionalità. Ma quel capitale è capitale vero, quanto più sia ricco di energie ideali, che si esprimano non solo in iniziative economiche ma anche in progetti generosi e alti: l'impegno per il rispetto dei diritti umani, la lotta per un ambiente migliore, la lotta contro il razzismo e l'esclusione sociale, l'opera di volontariato dentro e fuori i nostri confini. Il secolo che ci lasciamo alle spalle ha visto crescere l'Italia in benessere, in solidarietà, in spirito civile. Cento anni fa quattro militari di leva su dieci non conoscevano l'alfabeto; oggi milioni di giovani dialogano attraverso gli strumenti informatici e il loro numero cresce continuamente. Cinquant'anni fa le donne erano escluse dal diritto di voto; oggi partecipano in misura crescente alla vita politica, sociale, economica, arricchendone la qualità. Dobbiamo essere consapevoli non soltanto delle insufficienze, dei ritardi della nostra società, ma anche dei successi. Affronteremo così con maggiore sicurezza le sfide del nuovo secolo. Ci attendono i grandi problemi del mondo. La scienza è giunta alle frontiere della vita, animale e vegetale. Oltre, vi sono la seduzione e i pericoli di manipolazioni aberranti. La mondializzazione dei mercati economici e finanziari produce nuovo benessere, ma rischia di mortificare vocazioni produttive, mestieri, culture di popoli, di acuire

disuguaglianze. La globalizzazione dei mezzi di comunicazione di massa crea spazi di informazione per tutti i popoli, ma rischia di provocare dipendenza da scelte altrui, pericolosa soprattutto per la formazione delle nuove generazioni.

Il mio auspicio, è che nell'anno che sta per cominciare si faccia tutti insieme, con l'aiuto di Dio, un altro passo perché la Repubblica sia sempre di più, con una maggior ricchezza di libertà, di giustizia e di eguaglianza, la casa comune di ogni italiano. Buon anno tutti! Grazie”.

Discorso incentrato sui valori politici:

“Italiani e italiane, miei compatrioti, come è consuetudine, eccomi che ancora una volta devo entrare attraverso la televisione nelle nostre case per gli auguri della fine d'anno.

Dobbiamo avere la coscienza che il rientro dall'inflazione è un problema tutto nostro, quasi tutto nostro, essendo in gran parte venute meno le sue cause esterne. Uguale impegno richiedono il risanamento della finanza pubblica, l'avvicinamento del Mezzogiorno alle regioni di maggiore sviluppo del Nord, la lotta contro ogni forma di spreco delle risorse nazionali che sono risorse di tutti i cittadini. Sono convinto che nessun obiettivo è arduo se noi cittadini, noi istituzioni della Repubblica, troveremo tensione morale, operosità e passione civile, unità nell'essenziale. Possiamo farcela, anzi dobbiamo farcela. Da parte mia, nel doveroso rispetto dei limiti posti dalla Costituzione alle mie funzioni, ma nel pieno e responsabile esercizio dei poteri e delle prerogative che mi sono conferite dalla Costituzione, mai dimenticherò che la mia autorità deriva dall'essere stato eletto dai senatori, dai deputati, dai delegati dei Consigli Regionali riuniti in libero Parlamento, in nome e per conto del popolo sovrano, cioè in nome, e per conto vostro, uomini e donne di ogni età e di ogni condizione. Quest'autorità, in democrazia non può essere un mio privilegio personale: è una autorità di servizio, di servizio alla comunità, alla patria italiana che amo, da cittadino eguale ad ogni altro cittadino, e che ho giurato solennemente di servire con lealtà. Formulo a voi tutti gli auguri più affettuosi e mi faccio tramite degli auguri che idealmente, questa notte, ognuno di voi invia

ad ogni altro cittadino di questo Paese. Di questo Paese che, certo, non è una grande potenza politica, né militare, né economica, ma che è un Paese ricco quant'altri mai di umanità, di generosità, di intelligenza, di fantasia, di laboriosità. Noi, l'Italia, siamo una grande e libera nazione; siamo un nobile, antico e coraggioso popolo europeo. Non dimentichiamolo nell'asprezza del vivere quotidiano; non dimentichiamolo per l'amarezza di un ideale ancora non realizzato; ricordiamolo per operare con serenità e lottare con fermezza per il bene comune. Un augurio particolare e profonda gratitudine vanno ai cittadini che, nelle forze armate e in quelle di polizia, operano per la difesa della patria e per la tutela dell'ordine civile e delle libertà dei cittadini. Anche nel tragico episodio di Fiumicino le forze dell'ordine hanno dimostrato, in condizioni difficilissime, prontezza ed efficienza. Ad esse va, a nome della nazione, il mio più vivo elogio. Tra pochi mesi celebriamo il quarantesimo anniversario di fondazione della Repubblica, di quella Repubblica che fu la prima forma ideale della lotta italiana per le libertà civili e politiche e che è stata coronamento della causa nazionale, dell'indipendenza e dell'unità, coronamento del rinnovamento democratico del nostro Paese.

Più volte mi sono riletto il testo dell'impegno preso in Parlamento il 18 maggio del 1999, il giorno del mio giuramento. Quell'impegno si ispirava alle iscrizioni scolpite sui frontoni del Vittoriano, l'Altare della Patria: per la libertà dei cittadini, per l'unità della patria. Non è retorica, è l'essenza stessa del nostro convivere civile. L'essere chiamato a rappresentare l'Italia, a essere garante della sua Costituzione, l'ho vissuto non solo come un altissimo mandato, ma soprattutto come un dovere, una missione. Per questo ho voluto abitare, con mia moglie, sin dal primo giorno, nel Quirinale: da sette anni è la mia casa, la casa del Presidente della Repubblica, la casa degli italiani. Per questo ho insistito nel richiamare i simboli più significativi della nostra identità di nazione, dal tricolore all'inno di Mameli, l'inno del risveglio del popolo italiano; e nel rievocare il nesso ideale che lega il Risorgimento alla Resistenza, alla Repubblica, ai valori sanciti nella sua Carta Costituzionale. Per questo ho visitato ogni provincia d'Italia e ho voluto che agli incontri nelle città capoluogo partecipassero tutti i sindaci dei Comuni della provincia.

Nei due anni e mezzo trascorsi dalla mia elezione ho compiuto un primo viaggio in Italia. Ne ho visitato oramai quasi tutte le regioni. Continuerò, di provincia in

provincia. È un viaggio bellissimo: ne traggio vigore, fiducia, orgoglio sempre più forte di essere italiano. Ovunque avverto, nella ricchezza delle diversità delle nostre contrade, quel sapore d'Italia che viaggiatori del presente e del passato hanno sempre avvertito, che è natura, arte, lingua, cultura, modo di vita. Le radici dell'italianità sono antiche. È antica la nostra nazione. Ma le origini del nostro Stato sono assai più vicine. Risalgono all'inizio dell'ottocento, allorché uno stuolo di uomini di pensiero, poeti, letterati, filosofi, economisti, mossi da un grande amore per l'Italia, animati da un profondo senso etico, da grandi ideali e principi, diventarono anche uomini d'azione, e uomini di Stato. Quel movimento si pose chiari obiettivi: libertà; unità; indipendenza della patria, dell'Italia. Si diffuse e fu vissuto con intensa passione civile. Si nutrì della consapevolezza delle radici profonde della nostra storia, della nostra civiltà. Non a caso fu chiamato Risorgimento. L' inno di Mameli divenne l'inno della nazione italiana, l'inno del risveglio di un popolo. I grandi del Risorgimento non fecero sogni di conquista. Sognarono l'unità e la libertà d'Italia, e l'indipendenza di tutti i popoli. Vi è continuità fra gli ideali del Risorgimento e la Costituzione repubblicana, che l'Italia si è data dopo avere riconquistato la libertà con la Resistenza. Così come vi è continuità con la costruzione di un'Unione Europea che sia una federazione di Stati-nazione. Sono imprese grandi. Con esse noi, eredi dei padri fondatori dell'Italia e dell'Europa, dobbiamo confrontarci. L'Italia è sempre stata ed intende rimanere all'avanguardia nell'integrazione europea. Non possiamo sfuggire alle sfide che la storia del ventesimo secolo ci propone. Per preparare le nuove generazioni ad affrontarle bene, accanto alla famiglia, che è l'istituzione base della nostra società, deve operare una scuola capace di svolgere, con rinnovato impegno, il suo ruolo insostituibile di servizio pubblico: una scuola volta a formare i giovani, a prepararli ad assolvere responsabilmente i loro compiti di cittadini, e a favorire il loro inserimento, operoso e creativo, in una società che cambia ed avanza con tempi sempre più rapidi. Ci guidano alcuni principi, che uniscono gli italiani, al di là delle diversità d'idee politiche. Siamo una democrazia parlamentare. Chi ha avuto la maggioranza, abbia modo, governando, di dimostrare quanto vale, quanto sa fare per il progresso del nostro popolo. Chi è minoranza eserciti con impegno e responsabilità il compito indispensabile dell'opposizione: di controllo, di critica, di proposta.

Il dialogo fra le due parti, per essere costruttivo, presuppone che nella maggioranza la disponibilità all'ascolto, attento e aperto, della voce dell'opposizione, prevalga sulla tentazione di affidarsi sbrigativamente al rapporto di forza parlamentare; e che nell'opposizione la consapevolezza del diritto del Governo di portare avanti il proprio programma prevalga sulla tentazione del ricorso sistematico all'ostruzionismo. Una democrazia funziona bene se ciascuna istituzione esercita il proprio compito rispettando i limiti delle proprie competenze. La separazione dei poteri, il giudizio della Corte Costituzionale sulla costituzionalità delle leggi, la soggezione dei giudici esclusivamente alla legge, la neutralità e l'imparzialità delle pubbliche amministrazioni, garantiscono la libertà di tutti i cittadini. Il passaggio di funzioni dal Governo centrale alle autorità di Governo regionali e locali avvicina le istituzioni ai cittadini, valorizzando le autonomie. Questo passaggio deve avvenire razionalmente, al fine di rafforzare, non indebolire, l'unità nazionale. La Repubblica è una e indivisibile. Nel nostro ordinamento, il Presidente della Repubblica non ha, fra i suoi compiti, quello di governare. Egli rappresenta l'unità nazionale; vigila ed opera perché siano rispettati i principi costituzionali; ha il diritto-dovere di consigliare. Avverto tutta la responsabilità di rappresentarvi. Come guida, ho la Costituzione; le nostre tradizioni democratiche; il giuramento prestato dinanzi ai rappresentanti eletti della nazione; la mia coscienza.

Hanno responsabilità politica, hanno cioè responsabilità verso la cosa pubblica, tutti i cittadini, ma in particolare quei cittadini che a qualsiasi titolo fanno parte delle categorie dirigenti. Hanno responsabilità politica tanto i dirigenti delle aziende quanto i direttori degli strumenti di informazione di massa che influiscono sull'opinione pubblica, tanto i Magistrati che giudicano nelle aule giudiziarie, quanto i docenti che, nelle scuole di ogni grado, impartiscono insegnamento ai giovani; tanto i liberi professionisti quanto i pubblici funzionari; tanto gli intellettuali quanto i direttori di coscienze; e l'esemplificazione potrebbe continuare. Questa crisi nasce dal fatto che la classe dirigente non è stata sufficientemente integrata dall'apporto di elementi provenienti dai ceti popolari. Si ritorna così al problema della scuola che - per il ritardo di una risolutiva riforma in attuazione del dettato costituzionale, secondo il quale i capaci e meritevoli, anche se privi di mezzi, hanno diritto di raggiungere i gradi

più alti degli studi - non ha potuto integrare in modo sufficiente le categorie dirigenti con nuove energie provenienti da quei ceti che pur costituiscono la maggioranza della nazione. Le riforme in corso e quelle più ampie e risolutive che si delibereranno, risolveranno questo fondamentale problema; ma nell'attesa deve supplire il senso di responsabilità di tutti, e in particolare il senso di responsabilità degli attuali dirigenti. Si tratta ovviamente di una crisi di sviluppo che ha in sé i suoi rimedi, giacché i cittadini sono più che mai solidali con la Repubblica democratica, erede del Risorgimento e nata dalle lotte di liberazione. Democrazia alla quale dobbiamo il bene supremo della libertà nell'indipendenza della patria, lo sviluppo senza precedenti dell'economia del Paese, l'impostazione e l'avviamento a soluzione dei grandi problemi propri delle società moderne più progredite. Italiani, abbiamo dunque fede in questa democrazia, capace di sopperire alle sue stesse lacune, capace di assicurare effettivo progresso, nella libertà e nella giustizia, alla nostra cara patria.

L'Italia del 1988, proseguendo nella necessaria internazionalizzazione del suo sistema economico, dovrà lavorare con sincera convinzione per raggiungere l'importante traguardo del mercato unico europeo: il 1992 dovrà vederci ben pronti ad affrontare quest'ulteriore tappa verso l'unità dell'Europa, una strada lunga e difficile che intendiamo continuare a percorrere con fermezza perché è quella giusta. Nel Paese, al migliore profilo internazionale, corrisponde uno spirito diverso, cioè un più vivo desiderio e una più decisa volontà di crescita sociale, economica e culturale, di partecipazione e di equità. Tale spirito costruttivo costituisce una conferma della validità dello sforzo comune dal quale quarant'anni orsono nacque la nostra Repubblica, impegno collettivo che ci ha permesso di compiere in libertà e democrazia un grande cammino di pace, di sviluppo civile ed economico, nell'immediato dopoguerra difficilmente immaginabile. Quasi a fare da contrappeso al successo conseguito, esiste però un problema rilevante e tuttora irrisolto: l'adeguamento delle nostre istituzioni alle esigenze di oggi e soprattutto di domani. Non solo in Parlamento e nell'università, ma anche nei luoghi di lavoro e di incontro, nelle case di tutti noi, si ripete che le istituzioni non funzionano come dovrebbero, che non rispondono alle attese di una società avanzata, inoltre, le risorse impiegate dal settore pubblico non danno sempre utilità, e rendimenti che sarebbe giusto pretendere. E il fatto che molti problemi della vita quotidiana non trovino

concrete soluzioni attraverso le specifiche strutture istituzionali e amministrative, può creare, e spesso crea, un senso di delusione e di disaffezione per le vicende della politica, un inconcepibile e inaccettabile distacco tra Paese reale e Stato che è allarmante poiché allontana i cittadini dall'idea - guida della democrazia: quella che indica nella politica l'esercizio consapevole di scelte che riguardano il destino di tutti, che devono essere adottate col concorso di tutti. Dinanzi alle carenze e alle distorsioni, dinanzi a quello che può essere definito il malessere delle istituzioni, gli italiani, e soprattutto le nuove generazioni, sentono il bisogno di una democrazia ancora più moderna, ancora più matura, ancora più consapevole. Gli italiani, da cittadini liberi e uguali che funzionino in modo più dinamico e più efficiente. Si rafforza e si precisa, in altre parole, la richiesta che le istituzioni possano lavorare meglio e in modo più incisivo. Di fronte a questo bisogno di Stato, si tratta ormai di porre mano, con intelligente determinazione, ma anche con gradualità e la prudenza connaturata alle nostre tradizioni e al nostro sistema politico, ad un'opera complessa, articolata e realistica di modernizzazione degli strumenti legislativi e amministrativi previsti dalla nostra Costituzione e dalle nostre leggi. A questa impresa, siamo tutti chiamati a partecipare, perché la democrazia è impegno e responsabilità di tutti. La posta in gioco è quella, fondamentale, del superamento del senso di estraneità fra gente comune e Stato, e richiede uno sforzo convergente sia da parte delle istituzioni sia da parte dei cittadini. Non ritengo che, su questi argomenti, al Presidente della Repubblica competa formulare specifiche proposte, ma credo sia suo dovere auspicare che, alla denuncia giustificata e motivata delle disfunzioni istituzionali, se si vuole essere credibili e creduti, sia ormai indispensabile far seguire un coerente impegno per procedere ad una migliore razionalizzazione del nostro sistema di Governo parlamentare, per rendere l'opera del Parlamento e dell'esecutivo più efficace ed adeguata all'esigenza di una democrazia ad alto sviluppo industriale quale la nostra, alle richieste espresse in termini di nuovi diritti e di nuove libertà che provengono dalla comunità civile e dalle forze produttive. Non si tratta, a mio avviso, di fondare un nuovo ordinamento costituzionale.

Stiamo sviluppando uno Stato democratico ispirato ai principi del federalismo solidale. Esso ha radici nella nostra storia comunale, una storia che non ha

l'eguale, e nella pluralità degli Stati da cui è nata l'Italia unita. Creare regole di buongoverno e rapporti di leale collaborazione, anche in campo fiscale, fra lo Stato centrale e i livelli comunale, provinciale, regionale, come fra l'uno e l'altro degli organi di Governo locale, è compito difficile. Va affrontato con animo disposto al confronto e libero da pregiudizi. Si può essere aperti a ogni innovazione, purché resti fermo il principio di solidarietà, e non si metta a rischio, in nessun modo, l'unità nazionale. Vi sono principi intangibili, che non ammettono compromessi. Ancor prima che nella Costituzione, essi sono scolpiti nel marmo del Vittoriano, e in modo ancor più indelebile nel cuore di tutti gli italiani: la libertà dei cittadini, la libertà della patria. Soltanto mobilitando gli sforzi di tutti, grazie a una giusta collaborazione tra le varie sedi di Governo, locale e centrale, come fra pubblico e privato, potremo rafforzare quei territori e quei settori della nostra Italia che sono più deboli. È pur vero che il tradizionale divario fra Nord e Sud va attenuandosi, e che oggi vi è ovunque vitalità, e spirito d'iniziativa. È vero anche che da alcuni anni il tasso medio di disoccupazione in Italia va diminuendo. Ma le disuguaglianze fra regioni rimangono inaccettabilmente alte. Non possiamo rassegnarci all'idea di avere una parte d'Italia a corto di manodopera, tanto da dover ricorrere, nel nostro stesso interesse, a una crescente immigrazione - ben vengano gli immigrati, nel rispetto delle leggi - e un'altra parte che soffre di livelli ancora troppo elevati di disoccupazione. Inoltre, nella presente congiuntura non dobbiamo nasconderci la necessità, per tutto il Paese, di un potenziamento delle infrastrutture, di una accelerazione degli investimenti produttivi e di una maggiore prontezza nel reagire alle iniziative di una concorrenza internazionale agguerrita. A tal fine dobbiamo, in primo luogo, accrescere l'impegno nei settori della ricerca e della formazione, a tutti i livelli: dalla scuola primaria fino all' università. Soltanto così ci manterremo competitivi in un mercato globale che offre molte opportunità, ma che propone anche ardue sfide. La provincia italiana ha dato vita a originali modelli di sviluppo, mettendo a buon uso le grandi risorse naturali e culturali e rivelando una diffusa capacità imprenditoriale.

Ecco, con queste considerazioni italiane e italiani, io vi auguro che l'anno nuovo che sta per sorgere sia un anno di serenità per tutti voi”.

Se letti nella loro interezza, risulta evidente che i discorsi presentano discontinuità temporali piuttosto marcate, tratto inevitabile data la natura stessa dell'analisi. Tuttavia, il fatto che in un discorso vengano citati eventi storici distanti tra loro o che addirittura vengano richiamati periodi temporali incongruenti, non intacca il risultato specifico della ricerca. Nella lettura di ognuno di essi sembrano emergere i valori distintivi di ognuna delle tre sfere, seppur in modo disomogeneo rispetto a quanto si potrebbe riscontrare in un discorso preparato da un singolo oratore. Risulta singolare, ad esempio, che un messaggio di fine anno si concluda con dei riferimenti alla sfera imprenditoriale, che mal si addicono a chiosa di un messaggio il cui proposito e il cui *auditorium* sono molto più ampi rispetto a ciò che gravita attorno alla sola sfera imprenditoriale italiana.

Dal punto di vista della distinzione valoriale, invece, la differenza fra i tre prodotti è da considerarsi netta. Nel discorso incentrato sui valori sociali è molto chiara la tendenza a trattare temi volti al miglioramento della società nel suo insieme e alle innovazioni che l'Italia dovrebbe realizzare per progredire con la crescita di tutte le fasce sociali. Inoltre, è ben evidente come la parte del discorso derivata dal frammento 195 estratto dal quarto discorso di Pertini si focalizzi sulle ripercussioni che si sono avute sulla società tutta dall'emergere della loggia massonica P2 nel sottobosco del Paese. Per quanto riguarda il discorso basato sui valori morali, in esso sono rappresentati valori quali solidarietà, etica,

onestà, responsabilità e impegno, anche se il discorso si conclude in modo più tecnico e meno coerente con questi valori, rispetto al resto di quell'enunciato. Infatti, escluso il commiato finale, il frammento con cui si conclude questo messaggio è estrapolato dal discorso del 1999 di Ciampi, nel quale il Presidente coglie l'occasione del passaggio al nuovo millennio per parlare delle sfide attendono il Paese, soprattutto a livello scientifico e culturale. Il discorso sui valori politici, infine, è quello che, come ci si attendeva, tratta in maniera più approfondita ciò che riguarda la Repubblica, la Costituzione, le province, l'Europa e le sue istituzioni.

4.2. Rank e discorso incentrato su un topic derivato da IRaMuTeQ

Avendo a disposizione le elaborazioni precedenti che sfruttano il metodo Reinert di IRaMuTeQ, si è poi deciso di approfondire la ricerca sfruttando uno dei *topic* individuati per operare un confronto. Questo secondo approccio metodologico prevede di generare un discorso basato su alcuni valori specifici contenuti nel *topic* numero 3 utilizzando lo stesso procedimento con cui si sono prodotti i precedenti, ovvero immettendo una lista di parole in TaLTaC ed estraendo un indice TFIDF sulla base delle stesse. Ciò che varia rispetto a quanto eseguito con i discorsi generati in precedenza è il modo in cui si sono selezionate queste parole: mentre nei casi precedenti si è deciso di utilizzare liste di valori scelti qualitativamente dai due scienziati politici Grimaldi e

Riccamboni, in questo caso il TFIDF viene calcolato sulla lista di parole che la *topic extraction* di IRaMuTeQ ha riconosciuto come coerenti tra loro a tal punto da congregarle in uno stesso *topic*.

Si consideri, dunque, il *topic 3*, che al suo interno contiene i valori politici “unità” e “libertà”. Presa la lista delle prime venti parole del cluster che, insieme a queste due, lo compongono, si sono estratti dai 69 discorsi - tramite il TFIDF di TaLTaC - i frammenti più pertinenti rispetto a questa lista. Infine, si è composto un discorso seguendo lo stesso criterio osservato in precedenza, ovvero la selezione di tanti frammenti quanti sono necessari per la costruzione di un discorso di circa 2500 parole.

Frammenti con var categoriali e/o ex-post (testuali/quantitative)			
ID Frammento	Etichetta del Frammento	Informazioni aggiuntive	TFIDF cluster 3 IR
335	53_CIAMPI_2001_III		3,11075
225	37_COSSIGA_1985_I		2,96882
222	37_COSSIGA_1985_I		2,96672
247	41_COSSIGA_1989_V		2,94914
242	40_COSSIGA_1988_IV		2,90477
447	66_NAPOLITANO_2014_IX		2,89153
236	39_COSSIGA_1987_III		2,88841
249	41_COSSIGA_1989_V		2,88750
259	43_COSSIGA_1991_VII		2,83798
116	20_SARAGAT_1968_V		2,81158
257	42_COSSIGA_1990_VI		2,71122
248	41_COSSIGA_1989_V		2,69720
255	42_COSSIGA_1990_VI		2,69254
234	39_COSSIGA_1987_III		2,54765
336	53_CIAMPI_2001_III		2,54654
446	66_NAPOLITANO_2014_IX		2,53523
221	37_COSSIGA_1985_I		2,53099
148	25_LEONE_1973_III		2,50194
52	13_GRONCHI_1961_VII		2,50143
224	37_COSSIGA_1985_I		2,49936
56	13_GRONCHI_1961_VII		2,48325
27	09_GRONCHI_1957_III		2,45372
299	48_SCALFARO_1996_V		2,44314
47	12_GRONCHI_1960_VI		2,42111

Fig. 31. Rank di frammenti generato prendendo come riferimento il cluster 3 estratto con IRaMuTeQ

Discorso incentrato sulla lista di parole che compongono il topic 3:

“Buona sera e buon anno a tutti in quest'incontro che è diventato tradizionale tra il Capo dello Stato, che ha l'onore, l'onore di presentare gli auguri ai cittadini. Auguri a tutti, auguri a ciascuno: auguri a chi è cittadino d'Italia in patria e a chi vive all'estero; a chi si trova in questa terra per lavoro come immigrato.

L' Europa unita è già oggi, ma deve diventare ancor più in avvenire, una grande forza di pace, per sé e per tutti i popoli. Per esserlo, l'Unione deve trasformarsi in un soggetto politico unitario. Deve poter parlare con una sola voce sui grandi problemi. Deve operare per la crescita di un sistema di istituzioni di governo mondiale. In questi ultimi anni europei, americani, russi abbiamo lavorato insieme nei Balcani, dove odi insensati avevano fatto esplodere conflitti, creando una minaccia gravissima per tutti. C'erano stati massacri, deportazioni di interi popoli. Per porre fine a quelle tragedie, per proteggere i perseguitati, per permettere loro di ritornare alle loro case, non esitammo a impiegare le nostre forze armate. Non c'era altra scelta. Possiamo essere orgogliosi di ciò che hanno fatto e fanno, in Albania, in Bosnia, nel Kosovo, in Macedonia, in Eritrea, i nostri ragazzi in uniforme, e i nostri volontari civili, impegnati in opere a favore dei profughi, dei più deboli. Siamo orgogliosi dello spirito con cui gli uni e gli altri hanno svolto e svolgono il loro compito, riuscendo a farsi stimare perché sono portatori di pace. Lo sono anche le unità ora destinate all'Afghanistan, impegnate nel quadro di una missione internazionale in un compito difficile, ma necessario: aiutare a ricostruire uno Stato nella legalità. A tutti i nostri militari e volontari nel mondo va il mio fervido augurio. Oggi, dopo L'11 settembre, non dobbiamo esitare a combattere un nemico particolarmente insidioso, una rete terroristica internazionale, ispirata da un fanatismo irrazionale. Questa lotta non giungerà al pieno successo, se affidata soltanto alle armi. È necessario il sostegno concorde dei popoli. Essi chiedono una maggiore giustizia, per ridurre le enormi disuguaglianze che caratterizzano la società moderna. Il progresso, la cosiddetta globalizzazione, hanno avvicinato l'umanità, nel tempo e nello spazio. Il confronto fra le condizioni di vita dei popoli ricchi, e di quelli privi dei beni essenziali per la sopravvivenza, si è fatto intollerabile. Se guardiamo la terra dallo spazio, con i nostri astronauti, ci sentiamo padroni del mondo. Ma la televisione ci porta ogni giorno immagini che ci sconvolgono, di guerra, di fame, di malattie. È necessario mobilitare tutte le nostre risorse per eliminare la miseria, fonte di disperazione, terreno di coltura della violenza; così come per salvaguardare l'ambiente, nell'interesse dell'intero genere umano. Oggi abbiamo i mezzi per farlo, dobbiamo e possiamo farlo. L'Europa propone al mondo il principio del dialogo: a

cominciare da quello con il mondo islamico, che ci è così vicino, sull'altra sponda del Mediterraneo. È necessario per il bene comune che si parlino, in spirito di amicizia e di tolleranza reciproca, tutte le nazioni della terra. Nel Medio Oriente, vi è una disperata necessità di dialogo fra Israele e l'autorità nazionale palestinese: senza dialogo, come si può sperare di porre fine alla cieca spirale di sangue che lascia i popoli senza un futuro? Toccano i nostri cuori le parole che il papa rivolge all'umanità. Egli ha invitato tutti a mobilitare le migliori energie, perché l'amore prevalga sull'odio, la pace sulla guerra, la verità sulla menzogna, il perdono sulla vendetta. Al Santo Padre invio il mio saluto grato e augurale per il nuovo anno. Insieme con lui respingiamo ogni idea di una guerra di religione. Una siffatta guerra non c'è e non ci sarà; ripugna alle nostre coscienze, contraddice il fondamentale principio che è il rispetto dei diritti di ogni essere umano. Vengo all'Italia. L'amore della libertà, la volontà di dialogo, sono i principi ai quali si ispira l'idea di Stato che i padri della nostra Repubblica hanno disegnato quando hanno scritto insieme, pur divisi com'erano da dissensi politici, il testo della nostra Costituzione. La loro ispirazione veniva da lontano, dalla nostra identità di popolo, come l'avevano costruita secoli e millenni di una grande storia, che ha visto sempre l'Italia all'avanguardia della civiltà.

Adesso dobbiamo aggredire con determinazione questo male che ritarda il cammino della nostra economia e quindi del nostro progresso. Dobbiamo avere la coscienza che il rientro dall'inflazione è un problema tutto nostro, quasi tutto nostro, essendo in gran parte venute meno le sue cause esterne. Uguale impegno richiedono il risanamento della finanza pubblica, l'avvicinamento del Mezzogiorno alle regioni di maggiore sviluppo del Nord, la lotta contro ogni forma di spreco delle risorse nazionali che sono risorse di tutti i cittadini. Sono convinto che nessun obiettivo è arduo se noi cittadini, noi istituzioni della Repubblica, troveremo tensione morale, operosità e passione civile, unità nell'essenziale. Possiamo farcela, anzi dobbiamo farcela. Da parte mia, nel doveroso rispetto dei limiti posti dalla Costituzione alle mie funzioni, ma nel pieno e responsabile esercizio dei poteri e delle prerogative che mi sono conferite dalla Costituzione, mai dimenticherò che la mia autorità deriva dall'essere stato eletto dai senatori, dai deputati, dai delegati dei Consigli Regionali riuniti in libero Parlamento, in nome e per conto del popolo sovrano,

cioè in nome, e per conto vostro, uomini e donne di ogni età e di ogni condizione. Quest'autorità, in democrazia non può essere un mio privilegio personale: è una autorità di servizio, di servizio alla comunità, alla patria italiana che amo, da cittadino eguale ad ogni altro cittadino, e che ho giurato solennemente di servire con lealtà. Formulo a voi tutti gli auguri più affettuosi e mi faccio tramite degli auguri che idealmente, questa notte, ognuno di voi invia ad ogni altro cittadino di questo Paese. Di questo Paese che, certo, non è una grande potenza politica, né militare, né economica, ma che è un Paese ricco quant'altri mai di umanità, di generosità, di intelligenza, di fantasia, di laboriosità. Noi, l'Italia, siamo una grande e libera nazione; siamo un nobile, antico e coraggioso popolo europeo. Non dimentichiamolo nell'asprezza del vivere quotidiano; non dimentichiamolo per l'amarezza di un ideale ancora non realizzato; ricordiamolo per operare con serenità e lottare con fermezza per il bene comune. Un augurio particolare e profonda gratitudine vanno ai cittadini che, nelle forze armate e in quelle di polizia, operano per la difesa della patria e per la tutela dell'ordine civile e delle libertà dei cittadini. Anche nel tragico episodio di Fiumicino le forze dell'ordine hanno dimostrato, in condizioni difficilissime, prontezza ed efficienza. Ad esse va, a nome della nazione, il mio più vivo elogio. Tra pochi mesi celebriamo il quarantesimo anniversario di fondazione della Repubblica, di quella Repubblica che fu la prima forma ideale della lotta italiana per le libertà civili e politiche e che è stata coronamento della causa nazionale, dell'indipendenza e dell'unità, coronamento del rinnovamento democratico del nostro Paese.

In questo nuovo fronte è di incoraggiamento la vittoria politica contro il terrorismo interno ottenuta, con un grande sacrificio di vite e di dolore; con l'impegno unitario delle istituzioni costituzionali, dei magistrati del pubblico ministero e della giurisdizione, di tutte le istituzioni preposte all'ordine e alla sicurezza, dei lavoratori nei sindacati e nelle fabbriche, delle forze politiche e sociali, delle comunità religiose, del mondo della produzione, della cultura e dell'informazione, di tutto il popolo italiano desideroso di pace e di libertà. Questi sanguinosi eventi non ci fanno, ne ci debbono far dimenticare che lo scenario internazionale appare rasserenato dalle prospettive di distensione seguite al recente vertice di Ginevra tra gli Stati Uniti d'America e l'Unione Sovietica: vertice che è comunque un segno di speranza. Se non saranno

superati i precari equilibri, se non ci sarà vera pace, se la pace sarà soltanto una non-guerra basata sul timore reciproco dell'annientamento totale, tutto sarà costruito sulla sabbia. Il mondo, però, è purtroppo dolorosamente segnato da conflitti armati o latenti, mentre alcuni popoli, alcuni Stati, vasti settori di uomini e donne vivono privi dei fondamentali diritti di indipendenza, libertà, di giustizia e di eguaglianza; e altri popoli ancora sono drammaticamente alle soglie della sopravvivenza, quasi distrutti dalla fame e dalle malattie. Pace, libertà, progresso, diritto alla vita sono i valori unitari propri di ogni popolo e di ogni uomo perché essi affondano le loro radici nella dignità di ogni essere umano. L'Italia è stata ed è sempre più presente con un impegno attento e generoso per la pacifica soluzione dei conflitti, per l'affermazione dei diritti civili, nella lotta contro ogni forma di indigenza, di emarginazione e di discriminazione. Senza presunzione, con la chiara consapevolezza di ciò che tutto questo significa in termini di civiltà e di umanità per la nostra nazione, quest'impegno dovrà essere rinnovato e rinvigorito nelle iniziative internazionali del nostro Governo, nell'attenzione del Parlamento nazionale, nel lavoro volontario di tante associazioni; ma esso dovrà essere soprattutto un impegno civile e morale di ciascuno di noi. Associati alle Nazioni Unite, parte importante della comunità europea, componenti dell'Alleanza Atlantica, in coerenza e in fedeltà alle scelte che liberamente abbiamo fatto insieme, viviamo questa nostra presenza nel mondo e operiamo nella comunità internazionale, quale Paese europeo e mediterraneo, con realismo, intelligenza e prudenza, per la nostra sicurezza e per la nostra libertà, per la pace, per la libertà e per il progresso dei popoli, beni non divisibili dalla famiglia umana. Uno sforzo particolare dovrà essere fatto perché la costruzione della comunità europea ritrovi lo spirito e l'impulso dei suoi inizi di speranza e di coraggio, dopo le tragiche esperienze dell'ultimo grande conflitto mondiale.

Questa nuova stagione della libertà e della pacifica convivenza costituisce certo una grande sfida per i popoli dell'Est europeo impegnati duramente nella edificazione di Stati da riformare e da rifondare e di società da rinnovare: ma costituisce anche, è bene ricordarlo, una grande sfida per noi popoli dell'Occidente. Certo, in quanto accade, e giustamente, la storia riconoscerà parte importante alla serena fermezza ed alla convinta determinazione con la quale noi, nazioni dell'Occidente, nei nostri ordinamenti interni, e con le nostre

libere alleanze, abbiamo serbato, radicato e protetto, non solo per noi ma per tutti, insieme alla nostra sicurezza, i valori della libertà e del pacifico progresso. Ma ora queste nazioni, le nazioni dell'Occidente, devono dimostrare che gli ideali di libertà e di democrazia non possono essere assunti a difesa di una concezione nazionale ed internazionale statica, a protezione di un benessere morale e materiale, magari anche vasto, ma comunque circoscritto e conchiuso. Certo, caduti i muri, resteranno da colmare i fossati. E fra questi, il più immediato ed il più preoccupante è senz'altro quello economico. Nell'avviarci verso la fine di questo millennio, si prospetta per tutti un compito immane, in particolare per i Paesi delle comunità europee che si trovano impegnati in prima linea. Il disegno dei padri fondatori dell'Europa comunitaria, di De Gasperi, di Schumann, di Spaak e di Adenauer, va indubbiamente sollecitato, temprato, ampliato, forse rivisto. Anche sul piano della pace e del disarmo, le grandi trasformazioni in atto hanno improvvisamente aperto prospettive che, solo qualche mese fa, potevano sembrare utopistiche. Per la prima volta in quarant'anni, l'umanità può sperare in un mondo in cui la sicurezza di ciascuno non sia più basata sull'equilibrio del terrore e sull'incubo dell'annientamento nucleare. L'Italia, tutto il popolo, ciascuno di noi ha vissuto e partecipato con il cuore e con l'intelletto agli avvenimenti sconvolgenti dell'Est europeo: dall'avvio pacifico che le riforme avevano avuto in molti Paesi, a cominciare dalla Polonia e dall'Unione Sovietica, sino alla tragedia immane del popolo romeno, che ha dovuto iniziare il suo nuovo cammino verso la democrazia con una rivoluzione che è costata la perdita di tante vite e un così alto sacrificio di umano dolore. Noi, noi Italia, siamo una nazione di grande ed antica civiltà, che ha saputo darsi in questi anni strutture di libertà e vivere una vita democratica ricca ed aperta, che ha sviluppato una società moderna con una economia avanzata. Per questo, noi siamo chiamati a svolgere un ruolo importante nella storia dell'Europa: per questo, ma anche a motivo della nostra posizione all'incontro fra le culture latina, slava e germanica e la civiltà europea, araba ed africana. Ed in Europa, ciò potremo e dovremo fare, nell'anno che inizia, nell'esercizio del nostro turno di presidenza delle comunità europee. Si è parlato molto, sia in Italia che all'estero, della eccezionalità del nostro sviluppo economico, delle luci e delle ombre che hanno caratterizzato questo esempio, forse unico in Europa, di crescita e di progresso. Questi traguardi

sono stati certo raggiunti grazie alla straordinaria attitudine a padroneggiare le difficoltà che ci viene riconosciuta nel mondo intero e che abbiamo sempre dimostrato, e che tante volte ha fatto gridare al miracolo. Ma l'affermazione e il consolidamento della democrazia e della libertà, non dimentichiamolo, hanno influito in maniera determinante nel consentire a questo sviluppo di realizzarsi. E proprio in chiave di maggiore democrazia e di maggiore libertà, la società italiana sembra ora richiedere più chiarezza di orientamenti, più rigore di comportamenti, individuali e collettivi. Per progredire ancora verso quella che è stata chiamata la nuova primavera della nostra società, dobbiamo superare la logica di una crescita economica che sia basata soltanto sulla produzione di nuove ricchezze e sulla circolazione di nuovi prodotti: il perseguimento di più alti livelli di servizi collettivi e di convivenza civile è anche legato alla necessità di operare nel segno di una nuova misura di solidarietà, concreta e attiva, senza la quale noi non saremo mai una società pienamente civile. Penso in particolare al bisogno di più ricche e calde relazioni umane, che viene espresso da tante persone che vivono l'inascoltata malinconia del dolore e della solitudine, penso allo sconcerto di tanti di noi di fronte ad una evoluzione rapida e radicale di modelli di convivenze e di stili di vita. Penso agli anziani, ai minori trascurati, spesso sottoposti a violenza, ai malati e a tutti gli altri che vivono in condizioni di esclusione. E sono questi bisogni che spesso non hanno voce, che raramente giungono alle prime pagine dei giornali, salvo che non vi sia qualche caso drammatico e che quindi spesso non fanno opinione. Ma una società che voglia essere adulta, quale noi vogliamo che essa sia, deve avere occhi ed orecchie anche per queste esigenze. Su di esse si giocano indubbiamente la qualità stessa del progresso e la legittima aspettativa di ciascuno alla propria parte di serenità e di benessere. Ai giovani, sempre giustamente ansiosi di idealità, io credo che possano e debbano essere additati questi obiettivi. Se essi non hanno dovuto lottare in gran parte in prima persona per conquistare il diffuso benessere degli anni 80, possono però prefiggersi di lottare per il superamento dei costi culturali, sociali e civili, che certamente il così rapido sviluppo del Paese ha inevitabilmente comportato. Per farlo, ad esempio, un modo nobile e certamente prezioso per tutti sarà quello di dedicarsi a conservare l'instimabile patrimonio ambientale e culturale dell'Italia, con amore per i monumenti del passato, con amore per la nostra natura. Per citare

un esempio di grande attualità, realizzare un efficace sistema di parchi nazionali e di riserve naturali, significa rispondere ad una domanda di natura sempre più intensamente avvertita e significa contribuire a costruire una nuova primavera anche nel nostro ambiente, che è la casa di tutti. Ricordiamoci che questi non sono discorsi astratti. La natura non è soltanto un fatto fisico. Il rispetto della natura, la riconquista della natura, è una misura della moralità della vita di ciascuno di noi e di moralità della vita della comunità. Su una via però così impegnativa, non si procede da soli. Su una via che deve assicurare il proseguimento del progresso, occorre che noi continuiamo a confrontarci con il mondo esterno, con gli altri Paesi, con il mondo più vasto. Il confronto, su basi di pari dignità con tutti i nostri partners mondiali, richiede un costante aggiornamento della nostra cultura nazionale, che non si può richiudere in se stessa, in uno sforzo che è certo di realismo, ma anche di curiosità, di immaginazione e di fantasia, doti che certo al nostro popolo non mancano. Questo impegno assume, nell'anno a venire, un respiro particolare, perché si innesta in una stagione di grandi speranze e di innovazioni per la comunità internazionale.

Ho detto l'anno scorso: l'Italia risorgerà, ed era augurio fatto con il cuore. Oggi mi sento di poter dire l'Italia sta risorgendo! E allora, Italia auguri!"

Il discorso riportato sopra e ottenuto con lo stesso metodo dei tre precedenti - ma a partire da una lista di parole diversa -, condivide un solo frammento con il discorso generato prendendo come punto di riferimento i valori politici. In questo frammento, il numero 225, occorrono più di una volta le parole libertà e unità che, secondo la categorizzazione eseguita da Grimaldi e Riccamboni (2007), corrispondono a valori politici. Confrontando in modo qualitativo questi due discorsi, è possibile giungere ad una conclusione: mentre il discorso

che si basa sui valori politici appare coerente con questo tipo di tematiche soprattutto dal punto di vista delle parole contenute (Repubblica, Italia, Europa, Costituzione, ecc.), quello generato con il *topic 3*, non meno ricco di concetti legati a questa sfera valoriale, aderisce a quelle tematiche da un punto di vista più strettamente legato al contenuto anziché alla presenza di parole che riconducono a quell'ambito. Ciò porta a pensare che se, ad esempio, la parola "identità" veicola di per sé un concetto legato a una determinata sfera valoriale, al contempo, un ragionamento che tratti la tematica dell'identità può prescindere da quella stessa parola.

Partendo da queste constatazioni e tenendo come punto di riferimento il *topic 3* utilizzato per l'analisi precedente, si è, infine, deciso di estrarre i frammenti di testo che il software IRaMuTeQ ha utilizzato per comporre il *topic* stesso. In particolare, i frammenti relativi alle parole "libertà" e "unità"; questo procedimento non porta alla generazione di un discorso ma fornisce una sorta di "taccuino degli appunti". È opportuno specificare che questa lista di frammenti non può essere direttamente confrontata con quelle da cui si sono generati i discorsi precedenti. Questo perché le liste a partire dalle quali si sono generati i discorsi sono state prodotte dopo avere suddiviso i testi in paragrafi il più possibile coerenti al loro interno, mentre i frammenti estratti direttamente dal cluster 3, con IRaMuTeQ, sono porzioni di testo molto più piccole ricavate in modo automatico da analisi fatte prima della suddivisione in paragrafi. Ciononostante, tramite la variabile VPres, questi

“microframmenti” mantengono traccia del discorso dal quale sono stati estrapolati e consentono di verificare se essi appartengano o meno agli stessi “macroframmenti” utilizzati per comporre i discorsi precedenti. Ad esempio, si prendano come punto di riferimento i tre microframmenti che seguono, i quali hanno consentito a IRaMuTeQ di inserire la parola “libertà” all’interno del cluster 3:

**** *id_37_COSSIGA_1985_I *VPres_07Cossiga *VDisc_1 *id1_frammento_37
*id4_zero *id2_1985Cossiga *id3_zero *idfra= 0222

dei lavoratori nei sindacati e nelle fabbriche delle forze politiche e sociali delle comunità religiose del mondo della produzione della cultura e dell'informazione di tutto il popolo italiano desideroso di pace e di libertà

**** *id_37_COSSIGA_1985_I *VPres_07Cossiga *VDisc_1 *id1_frammento_37
*id4_zero *id2_1985Cossiga *id3_zero *idfra= 0222

mentre alcuni popoli alcuni stati vasti settori di uomini e donne vivono privi dei fondamentali diritti di indipendenza libertà di giustizia e di eguaglianza e altri popoli ancora sono drammaticamente alle soglie della sopravvivenza quasi distrutti dalla fame e dalle malattie

**** *id_37_COSSIGA_1985_I *VPres_07Cossiga *VDisc_1 *id1_frammento_37
*id4_zero *id2_1985Cossiga *id3_zero *idfra= 0222

pace libertà progresso diritto alla vita sono i valori unitari propri di ogni popolo e di ogni uomo perché essi affondano le loro radici nella dignità di ogni essere umano.

È interessante notare come tutti e tre questi microframmenti facciano parte del macroframmento 222, il quale compone il discorso generato a partire dal cluster 3, ma non quello generato a partire dai valori politici, nonostante tra questi valori figurì la parola libertà. Al contrario, si registra che tra i microframmenti estratti con IRaMuTeQ direttamente dal *topic* numero 3, non vi è traccia di alcuna porzione di testo che possa essere riconducibile al macroframmento 225, ovvero sia quello condiviso dal discorso basato su valori politici e da quello costruito con la lista di parole del *topic* 3.

CONCLUSIONI

Lo scopo della ricerca, prevalentemente metodologico e orientato al confronto tra i risultati ottenuti con procedure diverse, consisteva nella generazione semi-automatica di discorsi fondati su sfere valoriali differenti, eseguita tramite metodi statistici per l'analisi dei testi. L'obiettivo era ottenere un prodotto centrato interamente su uno specifico argomento e capace di inglobare le parti di numerosi discorsi diversi, maggiormente pertinenti ad esso. Gli approcci metodologici utilizzati per arrivare a questo risultato sono stati due:

1. il primo è consistito nel selezionare tre liste di valori individuati da Selena Grimaldi e Gianni Riccamboni (in M. A. Cortelazzo e A. Tuzzi, 2007) e di generare un discorso basato su quelle parole;
2. il secondo si è basato su una lista di termini contenute in uno dei *topic* individuati tramite il software IRaMuTeQ che al suo interno comprendeva alcuni di quei valori.

Come *vademecum* e fonte d'ispirazione per gli approcci analitici seguiti, si è tenuto in considerazione il lavoro di Lorenzo Bernardi e Arjuna Tuzzi (2010), che ha portato alla generazione semiautomatica del cosiddetto "Autografo del Presidente". Tuttavia, è importante sottolineare ancora una volta come i presupposti di partenza dei due statistici dell'Università di Padova fossero diversi da quelli all'origine di questa ricerca. Mentre lo

scopo di Bernardi e Tuzzi era quello di generare discorsi che fossero caratteristici di un singolo presidente (cioè generare un discorso che rappresentasse in sintesi tutte le peculiarità di una specifica figura, anche a partire da contributi di altri presidenti), in questo caso il *focus* è stato basato sui valori per trovare nell'insieme dei discorsi dei presidenti un bacino da cui estrarre un discorso centrato interamente su un argomento.

Nel caso dei discorsi composti sulla base di tre sfere valoriali differenti, si è scelto di procedere partendo da un punto di vista qualitativo, ovvero la suddivisione in sfera sociale, morale e politica di un insieme di parole, operata da Grimaldi e Riccamboni (in M. A. Cortelazzo e A. Tuzzi, 2007). La categorizzazione dei due scienziati politici è parsa efficace perché riassume i valori condivisi dalla società, quelli più legati ai singoli individui e quelli la cui diffusione è demandata - o dovrebbe esserlo - soprattutto alla classe politica di un Paese. Partendo da questo presupposto qualitativo e sfruttando i metodi descritti in fase di analisi, si sono prodotti tre discorsi che, nonostante i problemi descritti, rappresentano i valori da cui si è partiti.

Il secondo percorso di analisi, al contrario del precedente, non è partito da una suddivisione in valori fatta con criteri qualitativi, ma da una *topic extraction* effettuata in maniera automatica con uno degli strumenti quantitativi a disposizione. Più precisamente, dopo aver importato il corpus di testi totale formato dai 69 messaggi di fine anno, il *software* IRaMuTeQ ha individuato 8 *topics* coerenti al proprio interno, fornendo la

lista di parole che ha utilizzato per raggruppare i frammenti di testo all'interno di ogni *cluster*. E qui vi è la differenza nell'approccio rispetto ai discorsi precedenti: anziché costruire un messaggio sulla base di una lista di parole formata con criteri qualitativi, si è generato un discorso che prendesse come input per il *TFIDF* testuale di TaLTaC2 - lo stesso strumento utilizzato per la creazione degli altri discorsi - la lista di parole che IRaMuTeQ aveva considerato coerenti tra loro, fra cui alcuni valori presenti anche in una delle liste valoriali create qualitativamente; nello specifico, le parole "libertà" e "unità" presenti nel raggruppamento dei valori politici. Il risultato di questa seconda elaborazione è stato un messaggio diverso da quello incentrato sui valori politici, ad eccezione di un frammento che compare in entrambi, il numero 225, riconducibile al primo messaggio di fine d'anno di Francesco Cossiga.

Dal punto di vista linguistico e storico, i discorsi generati non sono esenti da alcune incoerenze che, tuttavia, riguardano aspetti su cui i metodi di ricerca utilizzati non possono intervenire. L'apertura e la chiusura dei frammenti di testo che compongono i discorsi presentano problemi di coerenza, limite con il quale si erano già misurati Bernardi e Tuzzi (2010). I periodi storici e gli accadimenti ad essi legati che si susseguono nei prodotti finali sono spesso distanti tra loro, motivo per cui ogni discorso appare più come un *collage* di eventi diversi anziché un testo coeso. Tuttavia, queste caratteristiche non intaccano i risultati ottenuti dal punto di vista dei metodi utilizzati. Infatti, le analisi effettuate mostrano che, utilizzando metodi quantitativi per analizzare un *corpus* di

testi, è possibile frammentare un complesso insieme di testi secondo più variabili e, da questa frammentazione, costruire altri discorsi che seguano precisi criteri. Nonostante le criticità citate in precedenza, il risultato è un insieme di testi coerenti con gli argomenti canonici dei messaggi di fine anno. In particolar modo, è interessante sottolineare come a partire da una classe semantica individuata con metodi semiautomatici, in seguito all'estrazione dei frammenti che più la rappresentano, si sia costruito un discorso capace di trasmettere alcuni dei valori che emergono nei discorsi di fine anno originari.

Dal punto di vista metodologico, è opportuno precisare quanto segue. I momenti della ricerca in cui si sono operate delle scelte che hanno successivamente influito sui risultati, sono numerosi: dalla fase di creazione e utilizzo delle variabili, alla scelta della soglia del *p-value* che discrimina tra parole specifiche e parole condivise, passando per la suddivisione manuale in paragrafi, per il numero di *topic* creati con il metodo Reinert, e molti altri. Tuttavia, ciò che è importante sottolineare è la riproducibilità dei risultati: se un'altra ricerca sullo stesso *corpus* di testi dovesse seguire le stesse scelte metodologiche e operative, poiché gli strumenti utilizzati si basano su dati concreti elaborati tramite la statistica dei testi, ci si dovrebbe aspettare che i risultati di quella ricerca siano coerenti con quelli presentati in questo lavoro.

In conclusione, benché la domanda posta a titolo di questa tesi nasconda una vena di ironia, si potrebbe affermare che, riuscendo a ottimizzare sempre più gli strumenti di ricerca e a raffinare precisi aspetti

intrinseci dei discorsi (riguardanti, soprattutto, il collegamento e la coesione tra le frasi), l'idea di produrre un discorso utilizzando metodi statistici per l'analisi dei testi applicati a un *corpus* di partenza, non è così peregrina come ci si potrebbe aspettare. Se gli algoritmi sulla base dei quali un computer elabora le proprie risposte difficilmente potrebbero gestire situazioni istituzionali delicate con la stessa lungimiranza, saggezza ed esperienza di un essere umano, alcuni di essi - tra cui quelli utilizzati per questa ricerca - potrebbero, nondimeno, essere impiegati per produrre quanto meno la traccia di un discorso che sappia far leva su precise caratteristiche, che tocchi determinati argomenti e che sia in grado di rispecchiare una forma appropriata al contesto comunicativo scelto. Sviluppi di questo genere potrebbero, negli anni a venire, aprire la strada alla formazione di nuove figure professionali le cui abilità nello sviluppo e nell'utilizzo di metodi di ricerca sempre più focalizzati su questi obiettivi, potrebbero innovare l'approccio alla costruzione di discorsi a qualunque livello, non soltanto istituzionale. Tuttavia, nonostante ancora molto possa essere fatto dalla ricerca per progredire in questo senso, sembra ragionevole ritenere che l'arricchimento di un discorso che si manifesta con l'accesso alla sfera intima dello scrivente, il suo vissuto e i suoi sentimenti, resteranno una prerogativa umana per molti anni ancora.

RIFERIMENTI BIBLIOGRAFICI

Bernardi L., Tuzzi A. (2007), Parole lette con misura (statistica), in: M.A. Cortelazzo, A. Tuzzi (eds), *Messaggi dal Colle. I discorsi di fine anno dei presidenti della Repubblica*, Venezia: Marsilio Editori, pp. 109-134.

Bernardi L., Tuzzi A. (2010), *L'autografo del Presidente della Repubblica: un archetipo del discorso di fine anno prodotto mediante ADT*, in: S. Bolasco, I. Chiari, L. Giuliano (eds), *JADT 2010 – Statistical Analysis of Textual Data. Proceedings of 10th International Conference*, vol. 2, Milano: LED, pp. 665-676.

Bolasco S., Giuliano L., Galli de' Paratesi N. (2009), *Parole in libertà. Un'analisi statistica e linguistica dei discorsi di Berlusconi*, Roma: Manifestolibri.

Bolasco S. (2010), *Taltac2. 10. Sviluppi, esperienze ed elementi essenziali di analisi automatica dei testi*. Milano: LED.

Bolasco S. (2013), *L'analisi automatica dei testi. Fare ricerca con il text mining*. Roma: Carocci, pp. 410.

Cortelazzo M.A., Tuzzi A. (eds, 2007), *Messaggi dal Colle. I discorsi di fine anno dei presidenti della Repubblica*, Venezia: Marsilio Editori.

Cortelazzo M.A., Tuzzi A. (2007), *Considerazioni finali*, in: M.A. Cortelazzo, A. Tuzzi (eds), *Messaggi dal Colle. I discorsi di fine anno dei presidenti della Repubblica*, Venezia: Marsilio Editori, pp. 231-237.

Cortelazzo M. A., Nadalutti P., Tuzzi A. (2013), *Improving Labbé's Intertextual Distance: Testing a Revised version on a Large Corpus of Italian Literature*, *Journal of Quantitative Linguistics*, 20(2), 125-152.

Giuliano L., La Rocca G. (2008), *L'analisi automatica e semi-automatica dei dati testuali I. Software e istruzioni per l'uso*, Milano: LED.

Giuliano L., La Rocca G. (2008), *L'analisi automatica e semi-automatica dei dati testuali II. Strategie di ricerca e applicazioni*, Milano: LED.

Giuliano L., Villani P. (2015), *Il linguaggio della leadership politica tra la Prima e la Seconda Repubblica, problemi di metodo e linee di ricerca*. Roma: Camera dei Deputati.

Grimaldi S. (2012), *I Presidenti nelle forme di governo*, Roma: Carocci.

Grimaldi S., Riccamboni G. (2007), Valori politici e capitale sociale nei discorsi dei presidenti della Repubblica, in M. A. Cortelazzo, A. Tuzzi (eds), *Messaggi dal Colle*. I discorsi di fine anno dei presidenti della Repubblica, Venezia: Marsilio Editori, pp. 135-178.

Kuckartz U. (1989), MAXqda, VERBIGmbH. (<https://www.maxqda.com/about>)

La Mendola S., Sterchele D. (2007), Costruire Noi. I presidenti della Repubblica Italiana come cerimonieri di un rituale mediale, in: M.A. Cortelazzo, A. Tuzzi (eds), *Messaggi dal Colle*. I discorsi di fine anno dei presidenti della Repubblica, Venezia: Marsilio Editori, pp. 55-86.

Lebart L. (2004), Validation technique in Text Mining. In : Spiros Sirmakessis (ed.), *Text Mining and its Application*, Springer Verlag, Berlin – Heidelberg, pp. 169-178.

March, J. G., Olsen, J. P. (1989), *Rediscovering Institutions, The Organizational basis of politics*, New York: The Free Press.

Muhr, T. (1993), ATLAS.ti, Scientific Software Development GmbH. (<https://atlasti.com/>)

Provalis Research (2004), QDA Miner. (<https://provalisresearch.com/products/qualitative-data-analysis-software/>)

Ratinaud P. (2009), IRaMuTeQ : Interface de R pour les Analyses Multidimensionnelles de Textes et de Questionnaires. (<http://www.iramuteq.org>).

Richards T. (1999), NVivo, Georgia State University Library. (<http://www.qsrinternational.com/about-us/our-history>)

Tuzzi A. (2003), *Analisi del contenuto*. Roma: Carocci.

Tuzzi A. (2007), Introduzione, in: M. A. Cortelazzo, A. Tuzzi (eds), *Messaggi dal Colle*. I discorsi di fine anno dei presidenti della Repubblica, Venezia: Marsilio Editori, pp. 9-21.