

Statistica Presidenziale

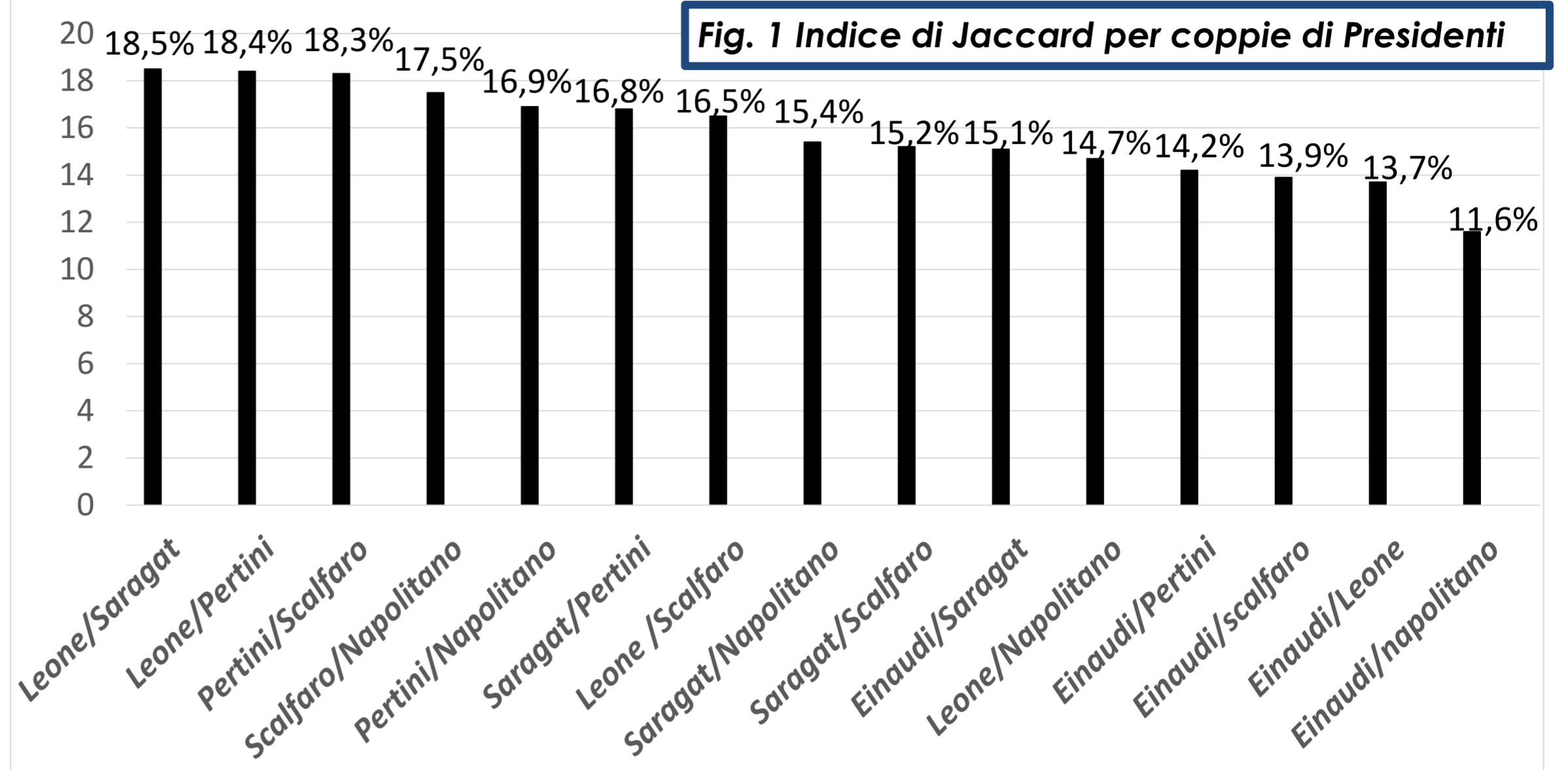
Un'analisi statistica sul linguaggio dei Presidenti della Repubblica Italiana nel tempo

Nell'analisi abbiamo preso in considerazione 6 (su 12) Presidenti della Repubblica Italiana, uno per ogni decennio a partire dal 1948. Dall'analisi è stato escluso l'attuale decennio e dunque l'ultimo Presidente della Repubblica (Mattarella). Sono stati selezionati per ciascuno, in formato digitale dal sito del Quirinale, un discorso per l'inizio dell'anno scolastico, uno di fine anno solare, uno in occasione della festa della Repubblica e un discorso per altre feste nazionali (o 4/11 o 25/04). Per i primi due Presidenti non è stato possibile trovare tali discorsi e abbiamo usato altri discorsi pubblici. Abbiamo scansionato i discorsi digitali mediante i tool offerti gratuitamente dal sito web Lexicoool.com e abbiamo successivamente elaborato i dati estrapolati con programmi scritti da noi in Python.

INDICE DI SIMILARITA'

Quanto sono diversi i discorsi dei Presidenti? Un indice di similarità descrive di quanto i discorsi presi in analisi si discostino per l'uso delle parole. Ai testi, privati di parole vuote e di "stop words" (articoli, congiunzioni...) non portatrici di significato autonomo, abbiamo applicato l'indice di similarità di Jaccard, calcolato da un nostro programma in linguaggio Python. I confronti sono riportati in Fig.1 per coppie di Presidenti, in ordine decrescente di similarità. L'indice di Jaccard è un indice statistico utilizzato per confrontare la similarità/diversità di insiemi campionari. Si trova attraverso il rapporto dell'intersezione dei due insiemi con la differenza tra l'unione e l'intersezione delle parole:

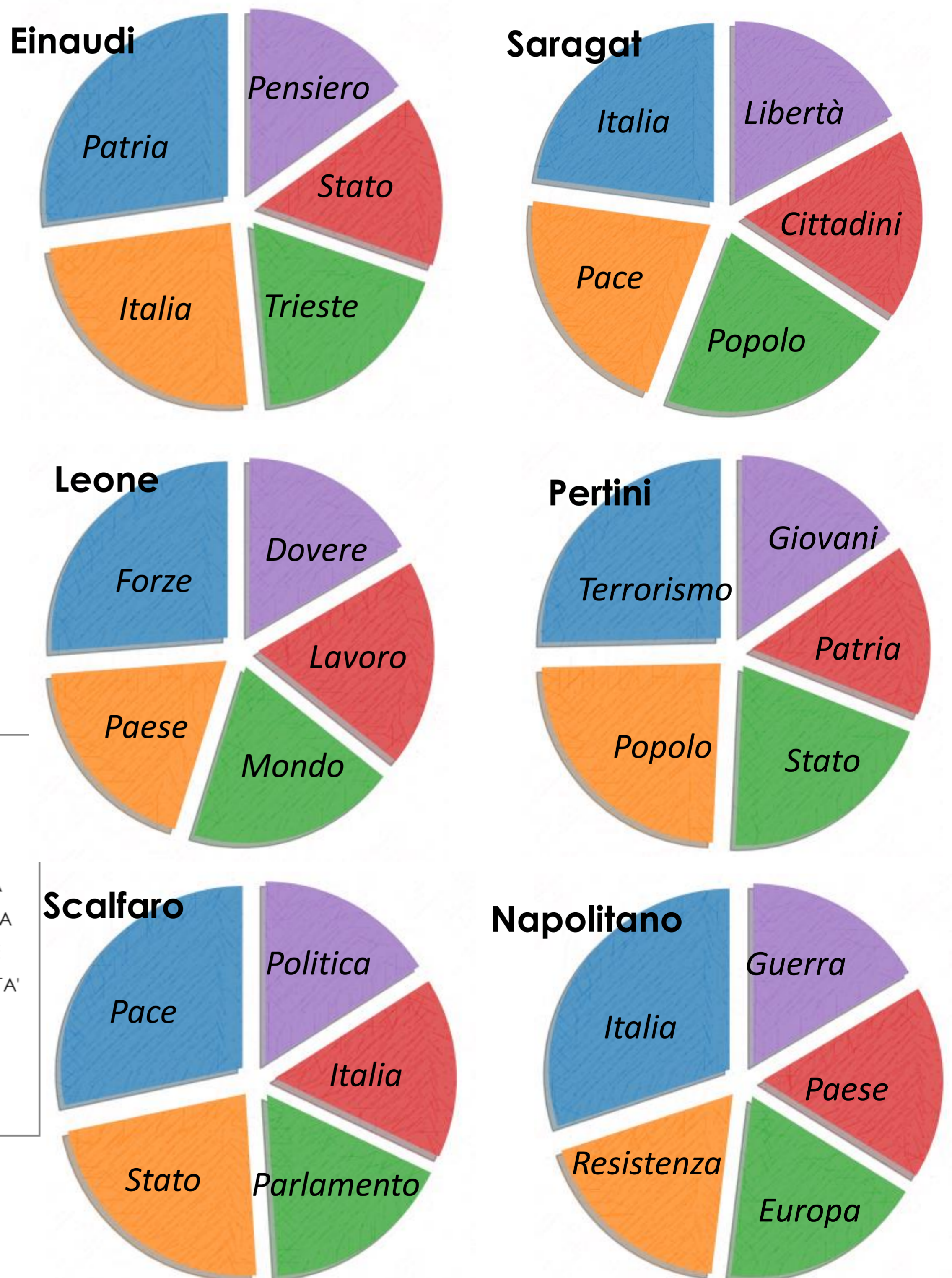
$$J(A,B) = \frac{|A \cap B|}{|A \cup B| - |A \cap B|}$$



LE PAROLE PIU' AMATE DAI PRESIDENTI

Quali sono le parole più amate dai 6 Presidenti scelti? La Fig. 2 mostra i diagrammi a torta per le 5 parole maggiormente preferite nei discorsi pubblici da ciascun Presidente della Repubblica Italiana. Anche i grafici sono stati costruiti in linguaggio Python

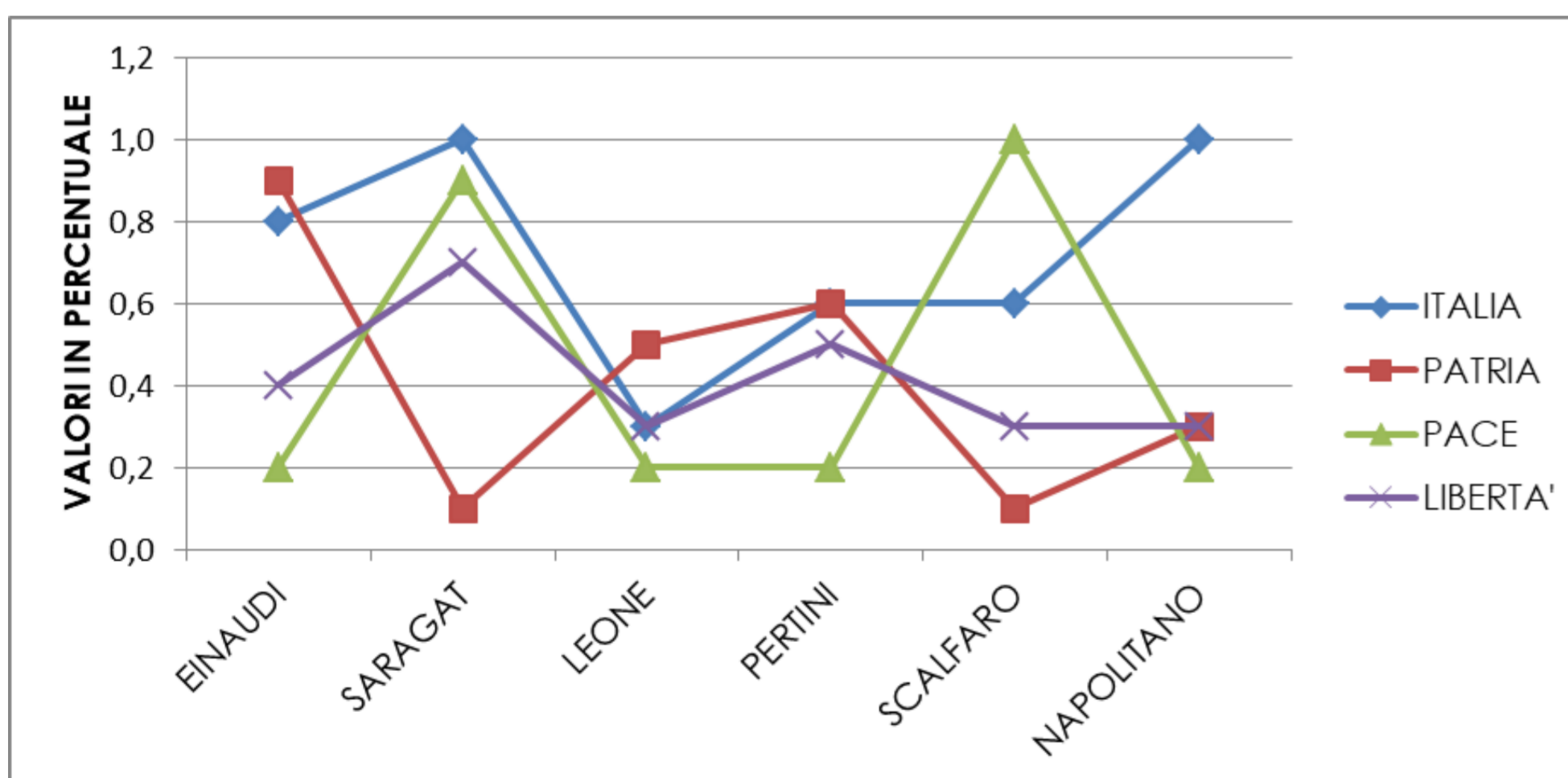
Fig. 2 Le 5 parole preferite dai 6 Presidenti



LE PAROLE NEL TEMPO

Le parole chiave usate dai vari Presidenti a partire da Einaudi fino a Napolitano mostrano una differenza al passare del tempo? Il buonsenso dice di sì, la statistica lo conferma e lo quantifica mediante l'analisi dell'andamento dell'uso di quattro parole chiave (Italia, Patria, Pace, Libertà). I dati di Fig.3, in percentuale, sono estratti dai discorsi privati di «stop words» mentre il grafico è stato creato con Excel.

Fig. 3 Le 4 parole chiave nel tempo



CONCLUSIONI La Statistica Presidenziale mostra che il linguaggio dei Presidenti della Repubblica cambia inevitabilmente nel tempo. In particolare troviamo che il primo presidente preso in analisi, Einaudi, usava un linguaggio che si discosta da quello di tutti gli altri (per esempio con la parola «Trieste»). Spesso Presidenti che si sono succeduti utilizzano un linguaggio più simile rispetto a loro predecessori più lontani nel tempo. Le parole più usate da ciascuno esprimono le problematiche storiche (dopoguerra, terrorismo, nuovo millennio, Europa) ma anche l'estrazione politica e gli ideali incarnati da ciascun Presidente. L'analisi potrebbe essere approfondita utilizzando un numero maggiore di discorsi pubblici accedendo agli archivi del Quirinale o estendendo l'analisi anche ai Presidenti non considerati in questo lavoro. Ulteriore approfondimento potrebbe essere svolto analizzando i discorsi dei Presidenti del Consiglio o dei principali leader politici della storia della Repubblica.

Bibliografia e sitografia

[1] S. Bolasco, "Estrazione automatica d'informazione dai testi", Rivista Mondo digitale n. 1., Marzo 2004, disponibile in rete a questo link:

http://archivio-mondodigitale.aicanet.net/Rivista/04_numero_due/Bolasco_p.27-43.pdf

[2] Implementing the five most popular similarity measures in Python *disponibile in rete a questo link:* <https://dataconomy.com/2015/04/implementing-the-five-most-popular-similarity-measures-in-python/>

[3] An Introduction to Bag-of-Words in NLP <https://medium.com/greyatom/an-introduction-to-bag-of-words-in-nlp-ac967d43b428>

[4] Per i discorsi dei Presidenti: <http://presidenti.quirinale.it>