

# PREPRINT

## L'apporto dell'analisi quantitativa nello studio di alcune forme linguistiche dell'argomentazione<sup>1</sup>

CLAUDIA RICCI  
CORINNE ROSSARI

*Università di Neuchâtel*

### 1. Introduzione

La ricerca qui presentata ha come fondamento la prospettiva secondo la quale diversi generi testuali pongono la/il parlante in primo o in secondo piano, ovvero rivelano o meno la sua presenza all'interno dell'enunciato prodotto. Tale questione è il fulcro dell'approccio alla descrizione del senso nell'ambito di diverse teorie linguistiche; limitandoci a quelle direttamente in linea con il nostro *background* teorico, possiamo qui citare la *Théorie de l'argumentation dans la langue* (ANSCOMBRE/DUCROT 1977), nell'ambito della quale si è sviluppata la polifonia linguistica, e in particolare la nozione di *point de vue*, la quale prevede l'emergere di una molteplicità di 'voci' all'interno dell'enunciato (DUCROT 1984); sulla stessa linea, ma con una concezione più referenzialista, la teoria della ScaPoLine (NØLKE *et al.* 2005), elaborata dai linguisti scandinavi per una descrizione della presa di posizione della/del parlante rispetto al contenuto comunicato; i lavori di Rabatel (tra gli altri, RABATEL 2005, 2006), con le

---

<sup>1</sup> La ricerca presentata in questo contributo è stata svolta in seno al progetto di ricerca *L'ancrage argumentatif des formes modales. Etude sur corpus avec un éclairage comparatif entre français et italien* (n. 100001F\_192247) finanziato dal Fondo Nazionale Svizzero per la Ricerca Scientifica.

nozioni di *sur-énonciation* e di *effacement énonciatif*, rendono la questione dell'adesione di chi parla al suo stesso 'dire', anche in questo caso, centrale nell'analisi. In linea con queste diverse correnti della linguistica enunciativa, questo lavoro intende affrontare la questione del legame tra la presenza/assenza della/del parlante nell'enunciazione e la dimensione argomentativa del genere testuale attraverso l'analisi di una famiglia di forme linguistiche che forniscono un contributo centrale in questa dimensione, ovvero i connettivi argomentativi, qui scelti in funzione della varietà delle relazioni argomentative che indicano: relazione tra argomento e conclusione (*donc, par conséquent, c'est pourquoi, de ce fait; quindi, dunque, perciò, di conseguenza*); tra conclusione e argomento (*car, parce que, puisque; poiché, dato che*); tra contro-argomento e argomento o conclusione (*mais, pourtant, ma, però, eppure*); a queste aggiungiamo la relazione di esemplificazione/specificazione (*par exemple, en particulier; per/ad esempio, in particolare*), che può essere considerata come un sottotipo di relazione argomentativa (cfr. la finalità argomentativa dell'esempio evidenziata in particolare da LANDOLSI 2018)<sup>2</sup>.

Il nostro obiettivo sarà quello di evidenziare il modo in cui i connettivi qui considerati contribuiscono alla dimensione argomentativa di un genere discorsivo, manifestando la presenza della/del parlante in modo più o meno visibile. A questo scopo, prenderemo in considerazione due corpora: un corpus caratterizzabile come neutro dal punto di vista argomentativo (corpus enciclopedico) e un corpus caratterizzabile invece come fortemente

---

<sup>2</sup> Il fatto di aver qui preso in considerazione una selezione limitata di connettivi è da ascrivere al metodo quantitativo utilizzato, che non consente di evitare sistematicamente l'ambiguità dovuta a polisemia/omonimia (ne è un esempio il connettivo *perché*); ciò risulta nell'impossibilità, per alcune forme, di essere recuperate in modo automatico.

argomentativo (corpus di discorsi politici)<sup>3</sup>.

Il nostro lavoro si articolerà in tre fasi. In una prima fase, partiremo dai nostri studi sui connettivi (ROSSARI/RICCI in stampa) effettuati nell'ambito della *textométrie*<sup>4</sup> (LEBART/SALEM 1994, PINCEMIN 2011, RASTIER 2008) per classificarli su una scala di soggettività, a seconda che il connettivo sia più o meno fortemente caratterizzato da un tratto indicante che la relazione è da intendersi come concepita dalla locutrice oppure come una relazione che attiene al mondo esterno e che la locutrice si limita a descrivere. Osserveremo se vi siano o meno differenze quanto alla posizione dei diversi connettivi su tale scala in funzione del genere testuale e di ognuna delle lingue considerate. Nella seconda sezione, misureremo il grado di rappresentazione in diversi generi testuali di alcuni tra i connettivi presenti sulla scala di

---

<sup>3</sup> I corpora analizzati sono: per il discorso enciclopedico, Wikipedia francese (18'715'455 tokens, 10'571 articoli, dal 2019, senza sezioni bibliografiche) e Wikipedia italiana (18'394'987 token, 15'905 articoli dal 2019, senza sezioni bibliografiche); per il discorso politico, un corpus di allocuzioni presidenziali italiane (2'639'007 tokens, che coprono il periodo 1955-2021) e uno di allocuzioni presidenziali francesi (2'843'577 tokens, che coprono il periodo 1958-2020); consideriamo questi discorsi come appartenenti alla categoria dello scritto – anche se sono scritti a fini di oralizzazione, il che comporta, nei termini di FERRARI 2022, un'oralità simulata – essendo caratterizzati (in particolare in italiano) dalla forte presenza di strutture tipiche dello scritto e da un alto grado di formalità (un esempio per tutti è l'uso dell'allocutivo *Ella* in luogo di *Lei* nel rivolgersi ad interlocutori/trici quali capi di Stato o di governo). Infine utilizzeremo (cfr. sez. 2.), a scopo comparativo, due corpora giornalistici: uno comprendente un anno del quotidiano francese *Le Monde* (17'895'009 tokens, relativi all'anno 2010) e un secondo costituito da articoli del quotidiano italiano *L'Adige* (20'343'148 tokens, 38'496 articoli relativi al periodo 1996-2006).

<sup>4</sup> L'approccio della *textométrie*, che ha come precursore la lessicometria, consiste nell'applicare calcoli statistici allo studio di corpora digitalizzati di testi. Tale approccio permette di articolare analisi quantitative e qualitative interrogando i contesti di occorrenza preferiti delle forme in un corpus specifico e permettendo così di osservare eventuali regolarità in quelle che si possono definire "proprietà combinatorie" (BLUMENTHAL 2002: 115) di una data forma con altre forme del lessico.

soggettività, allo scopo di verificare la relazione tra la loro posizione sulla scala e il loro uso in discorsi più o meno argomentativi. Infine, nella terza sezione, studieremo le associazioni specifiche tra diverse categorie di forme lessicali indicanti soggettività e alcuni connettivi che mostrano comportamenti simili o, al contrario, ben distinti, in relazione con i risultati presentati nelle prime due sezioni della ricerca. Analizzeremo così quali categorie contribuiscono al *ranking* del connettivo sulla scala di soggettività e al suo grado di rappresentazione a seconda del corpus.

## 2. Connettivi e soggettività

I nostri studi precedenti (ROSSARI/RICCI in stampa) hanno dimostrato come i connettivi possano veicolare più o meno fortemente un tratto di soggettività. Per individuare la presenza di tale tratto, si è inizialmente effettuata la ricerca di tutte le forme che presentassero con ogni connettivo – entro un intervallo di 5 *tokens* a sinistra e 5 a destra dello stesso – un’associazione statisticamente significativa; queste forme sono state estratte automaticamente grazie al software TXM (HEIDEN 2010) su corpora rappresentanti diversi generi testuali. Tra le forme così ottenute, sono state prese in considerazione tutte quelle che veicolano un’indicazione modale, ovvero un’indicazione che concerne l’atteggiamento della/del parlante rispetto alla propria enunciazione o al contenuto del proprio enunciato. Tutte le singole forme sono state raggruppate in un unico macro-tipo, che porta l’etichetta di *speaker’s involvement forms* (SIFs). In base a questo studio, abbiamo esteso il paradigma di connettivi e li abbiamo ordinati su una scala di soggettività utilizzando il calcolo del Log-Likelihood (LL)<sup>5</sup>, indicante associazione statisticamente significativa tra ognuno dei connettivi analizzati e la macrocategoria delle SIFs. Le figure 1. e 2. presentano le scale di soggettività in francese e in italiano. La ricerca è stata effettuata sui due corpora (enciclopedia e discorso politico) che, come già menzionato, hanno

---

<sup>5</sup>Questo calcolo esprime la probabilità che l’associazione tra due elementi non sia dovuta al caso; a partire da un punteggio soglia di 10,83, tale probabilità è superiore al 99,9%.



- quindi* – si ritrovano tra i quattro primi sulla scala; la loro stabilità nei due corpora costituisce un indizio per considerarli come codificanti convenzionalmente un alto grado di soggettività.
- differenze in funzione del genere testuale riscontrabili in entrambe le lingue: *par exemple / per/ad esempio* è molto più fortemente soggettivo nei corpora enciclopedici, in francese come italiano, rispetto ai discorsi presidenziali;
  - una differenza tra generi che risalta in francese: il connettivo *car*, ai primi posti sulla scala nel discorso enciclopedico, si ritrova in fondo alla scala nei discorsi presidenziali.

L'osservazione di queste tendenze sfocia nella necessità di un approfondimento in due direzioni. Da una parte, la stabilità osservata tra lingue e generi per le forme sopra menzionate invita a verificare in quale misura si faccia ricorso a queste forme in funzione della forza della dimensione argomentativa di un genere discorsivo. A tale scopo abbiamo aggiunto il corpus giornalistico menzionato nella nota 3, che ci servirà come *tertium comparationis*. È questo l'oggetto della sezione 2.

D'altra parte, sia le similitudini che le differenze che un connettivo può mostrare sulla scala in funzione del corpus considerato invitano a un approfondimento sulle specifiche categorie cui appartengono le forme che con i connettivi si associano in modo significativo. Ci si può in effetti domandare se un connettivo avente la stessa forza soggettiva nei corpora considerati si associ con le stesse tipologie di forme lessicali indicanti soggettività in un corpus rispetto ad un altro; ma anche se un connettivo la cui posizione sulla scala non è la stessa, ma varia in funzione del corpus riveli associazioni che potrebbero spiegare tali variazioni. Questa domanda sarà l'oggetto della sezione 3.

### **3. Grado di rappresentazione dei connettivi in diversi generi testuali**

Per verificare se esista una correlazione tra la soggettività del connettivo e il grado di rappresentazione (sovra- o sottorappresentazione) dello stesso in base alla natura più o meno argomentativa del genere considerato, abbiamo calcolato tale grado di rappresentazione su una

partizione per generi testuali dell'insieme dei nostri tre corpora. I grafici sottostanti forniscono i risultati di questo calcolo. Le soglie oltre le quali un connettivo si può considerare sovra- e sottorappresentato sono rispettivamente 2 e -2.

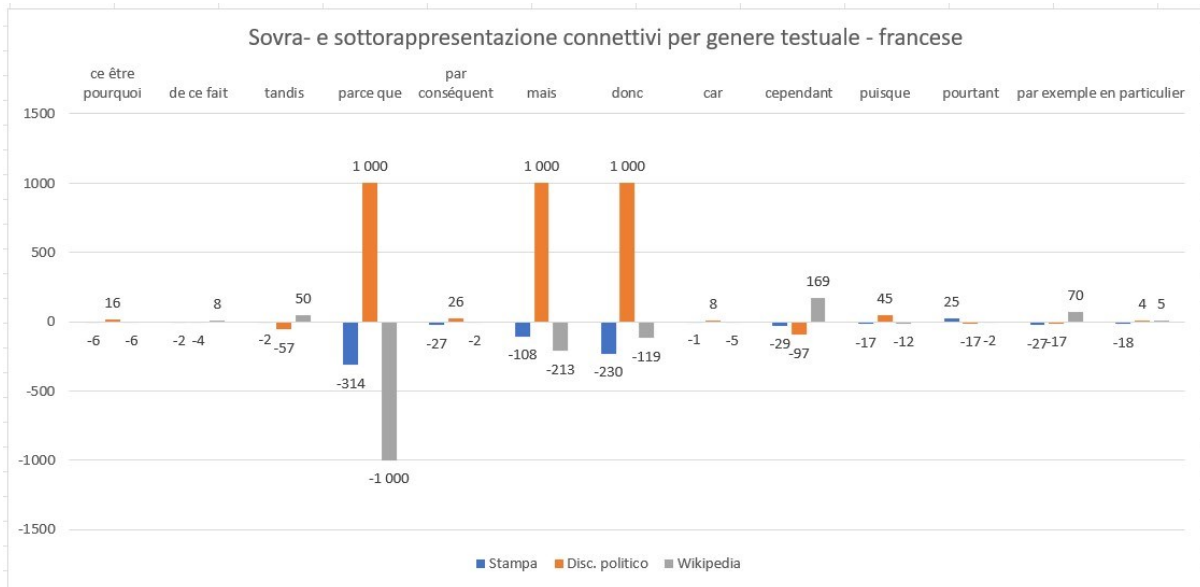


Fig. 3 - Grado di rappresentazione dei connettivi in tre generi testuali – francese

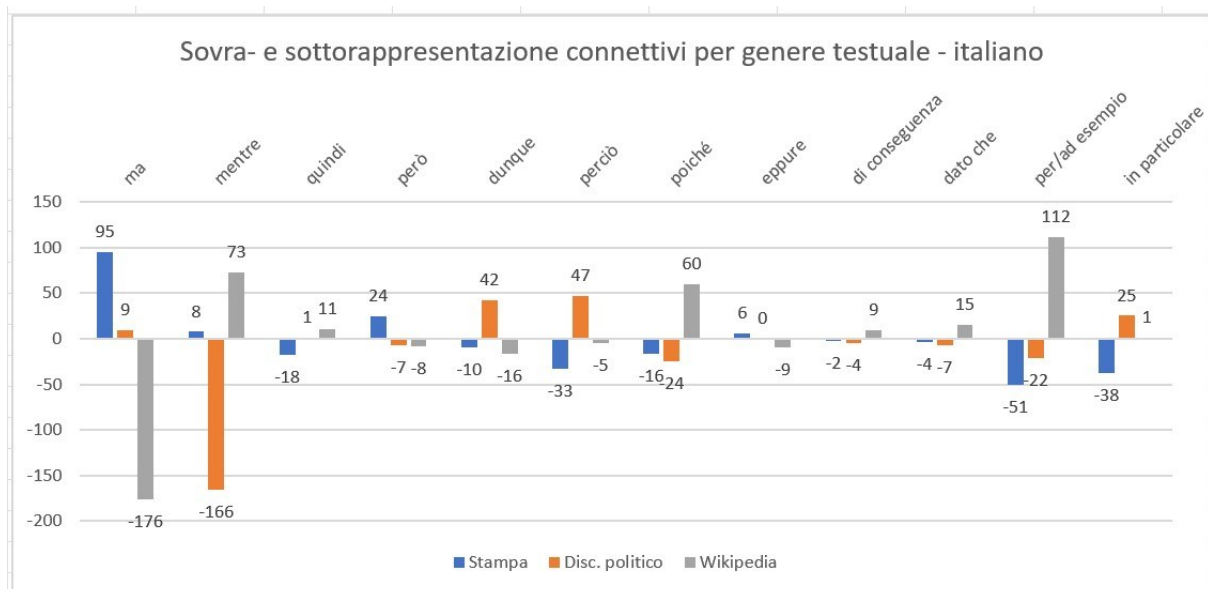


Fig. 4 - Grado di rappresentazione dei connettivi in tre generi testuali – italiano

In base alla posizione del connettivo sulle scale di soggettività, i grafici fanno rilevare tre fatti

salienti:

1. I connettivi che risultavano essere ai primi posti sulla scala di soggettività (*mais, donc, ma, però*) manifestano un comportamento totalmente in linea con quanto ci si potrebbe aspettare in funzione della natura del corpus considerato: sono cioè sovrarappresentati nel discorso politico e sottorappresentati nel discorso enciclopedico. Anche *par conséquent* e *di conseguenza* risultano in linea con quanto atteso: situati in basso sulla scala di soggettività, oppure addirittura non presenti, sono ambedue sovrarappresentati nel discorso enciclopedico, così come *de ce fait*. Questi risultati consolidano le ipotesi sul carattere convenzionale del tratto di soggettività, sia esso fortemente o debolmente presente nel connettivo.
2. Per quanto concerne invece le tendenze più inaspettate, si osserva che *quindi* (situato in cima alla scala di soggettività) risulta sovrarappresentato nel discorso dai toni più neutri. Occorre tuttavia sottolineare che *quindi* presenta, nel discorso enciclopedico, molteplici valori non argomentativi: può innanzitutto avere un valore che è strettamente di successione temporale, nel senso di *poi, successivamente*, come nell'esempio seguente, tratto dal nostro corpus di Wikipedia:

- (1) Un gruppo di invertebrati sono stati [...] esposti per 12 giorni alle radiazioni cosmiche [...] e *quindi* recuperati dopo il rientro della capsula sulla Terra.

Inoltre, in un numero ancora maggiore di casi (84 su un campione di 100 occorrenze esaminate individualmente), *quindi* segnala una relazione che si situa a metà tra segnale temporale e di consecuzione tra eventi, del tutto comparabile a *di conseguenza*, oppure un legame tra due modi di designare una stessa realtà, e non un'operazione inferenziale, tipica del discorso argomentativo. Esempi rispettivi di questi usi non argomentativi sono i seguenti:

- (2) Camille Desmoulins, mentre arringava la folla parigina alla rivolta, chiese ai manifestanti quale colore adottare come simbolo della rivoluzione francese [...]: i manifestanti risposero " Il verde! Il verde! Vogliamo delle coccarde verdi! ". Desmoulins colse *quindi* una foglia verde da terra e

se l'appuntò al cappello come segno distintivo dei rivoluzionari.

- (3) a nord, un altro stretto, l'Harlem River, la [= l'isola di Manhattan] divide dal Bronx e *quindi* dal resto del continente statunitense.

Nel corpus costituito da discorsi politici, *quindi* si presenta invece nella maggioranza dei casi con un valore chiaramente argomentativo di tipo inferenziale-deduttivo, come in (4):

- (4) È inutile dire quanto siamo convinti che viviamo in un mondo completamente diverso da quello anche di alcuni decenni fa: siamo immersi in un processo di globalizzazione che ha portato anche grandi benefici all'Estonia, ma la cui crisi in questo momento colpisce anche l'Estonia, e *quindi* dobbiamo pensare veramente a dare delle fondamenta più solide all'ordine mondiale, al sistema delle relazioni internazionali.

Questa diversità di valori di *quindi* contribuisce a rendere conto della sovrarappresentazione di questo connettivo nel corpus enciclopedico.

3. Per quanto riguarda i connettivi ai quali le scale di soggettività non ci permettono di attribuire un tratto più o meno soggettivo (dal momento che si posizionano in cima alla scala in un corpus e in fondo ad essa in un altro), osserviamo comunque che non sono in situazione di banalità quanto alla rappresentatività nei generi considerati: *par exemple* e *per/ad esempio* si ritrovano in sottorappresentazione nel discorso politico e in sovrarappresentazione nel discorso enciclopedico; il contrario si osserva per *car*.

L'analisi delle categorie di forme con le quali questi connettivi sono associati in modo specifico ci consentirà di fornire un'interpretazione in merito a questi diversi comportamenti.

#### **4. Analisi delle categorie di forme soggettive in relazione ai connettivi**

Alla luce di quanto osservato nelle sezioni 1. e 2., è necessario addentrarsi oltre per capire la sensibilità di un connettivo rispetto alle diverse categorie di SIFs. In effetti, considerare l'insieme indistinto di tutte le forme lessicali esprimenti soggettività non fornisce alcuna indicazione sulla possibilità che, all'interno di tale insieme, determinate tipologie di forme soggettive possano contribuire in misura maggiore o minore al tratto di soggettività di un

connettivo in un determinato corpus.

Riprendiamo qui le categorie di SIFs individuate In ROSSARI/RICCI (in stampa): si tratta di cinque diverse categorie morfo-sintattico-semantiche, che fanno trasparire in modi differenti la soggettività della/del parlante nel suo dire: (1) una categoria di forme che lasciano chi parla piuttosto sullo sfondo: si tratta degli aggettivi costruiti con i suffissi *-ible* e *-able* in francese e *-bile* in italiano che veicolano un'indicazione di possibilità o necessità; (2) una categoria di forme che, al contrario, pongono chi parla in primo piano: gli avverbi, tra i quali abbiamo preso in considerazione gli avverbi modali, valutativi e di enunciazione; infine, categorie che possono essere considerate intermedie: (3) verbi modali indicanti volontà, possibilità, necessità; (4) verbi che manifestano un'attività di pensiero; (5) *verba dicendi*.

Nei grafici sottostanti sono riportate le associazioni specifiche, ottenute tramite il calcolo LL, tra le cinque categorie di SIFs e i connettivi; abbiamo preso in considerazione sia i connettivi che fanno apparire convergenza tra le scale di soggettività e la loro rappresentazione nei corpus (*mais, donc, ma, però*), sia quelli che hanno presentato una divergenza (*quindi*), sia quelli per i quali le scale di soggettività non hanno permesso di far emergere specifiche tendenze (*ad esempio, par exemple e car*).

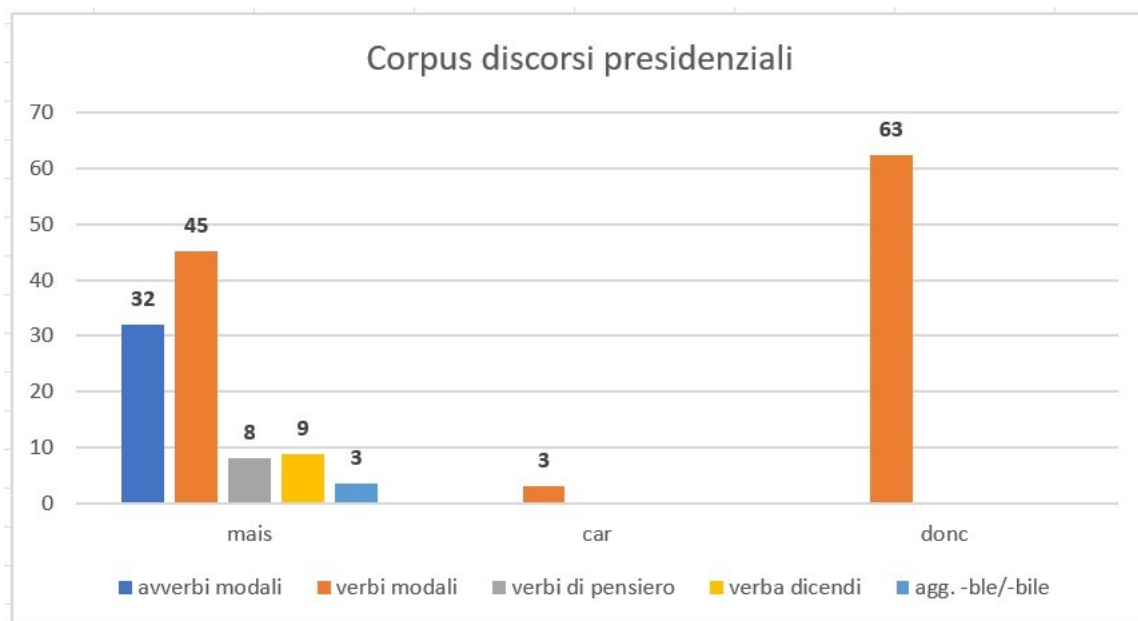


Fig. 5 - Corpus discorsi presidenziali francesi - associazioni specifiche tra categorie di forme soggettive e

connettivi

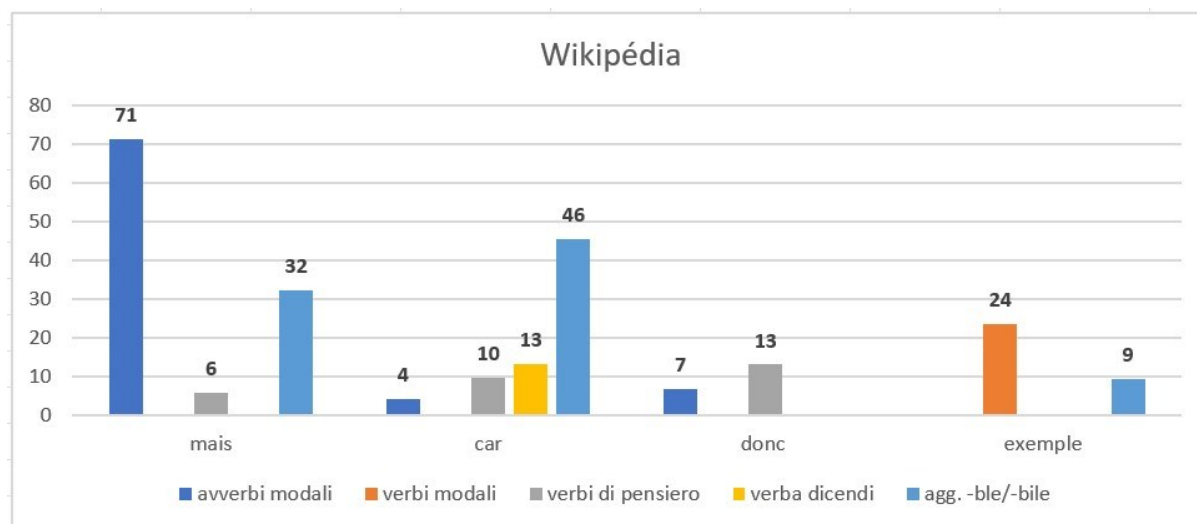


Fig. 6 - Wikipédia francese: associazioni specifiche tra categorie di forme soggettive e connettivi

La comparazione tra i due grafici fa emergere un contrasto concernente la categoria dei verbi modali: mentre *car* e *donc*, nel corpus dei discorsi presidenziali, devono la loro attrazione con le SIFs a questa unica categoria, si osserva il contrario nel corpus enciclopedico: né l'uno né l'altro vengono attratti dai verbi modali; *car* viene attratto da tutte le altre categorie e *donc* da due categorie. Anche *mais* attira i verbi modali solo nel corpus dei discorsi presidenziali e non in Wikipedia. *Par exemple* fa emergere il contrasto relativo alla categoria dei verbi modali in modo speculare: è l'unico connettivo associato a questa categoria nel corpus di Wikipedia, mentre non è associato ad alcuna categoria nel discorso politico, e per questo non appare nel grafico, il che indica che la sua attrazione con le SIFs non è legata ad alcuna in particolare tra queste categorie di forme. Questa tendenza non è in contraddizione con la dimensione argomentativamente neutra del discorso enciclopedico: si rileva in effetti che un verbo modale come *devoir*, nei casi in cui si combina con *par exemple*, serve innanzitutto ad enunciare una

regola che viene illustrata dal contenuto introdotto dal connettivo; e quindi, in questo tipo di configurazione, non vi è legame argomentativo tra l'esempio e l'esemplificato:

- (5) La reproduction entre individus d'une même espèce doit être possible en conditions naturelles et la progéniture doit être viable et féconde. Par exemple, le cheval et l'âne sont deux espèces interfécondes mais leurs hybrides (mulet, bardot) le sont rarement.

La riproduzione tra individui di una stessa specie deve essere possibile in condizioni naturali e la prole deve essere vitale e feconda. Per esempio, il cavallo e l'asino sono due specie interfeconde, ma i loro ibridi (mulo, bardotto) lo sono raramente.

Diversamente, quando *par exemple* viene usato con *devoir* nel discorso politico, l'enunciato introdotto dal connettivo serve a giustificare una possibile 'richiesta di fare' veicolata dal valore deontico del verbo.

- (6) Des règles particulières *devront* être fixées pour encourager la production des jeunes. *Par exemple*, 50 % des programmes musicaux et des chansons *devraient* être réservés à des programmes musicaux et à des chansons d'auteurs ou de chanteurs françaises ou français.

Regole specifiche dovranno essere fissate per incoraggiare la produzione giovanile. Per esempio, il 50% dei programmi musicali e delle canzoni dovrebbe essere riservato a programmi musicali e a canzoni di autori o cantanti francesi.

Più globalmente, il fatto che solo nel discorso politico la categoria dei verbi modali risalti mostra che la presenza della/del parlante si manifesta soprattutto tramite questa categoria di forme; nel discorso enciclopedico, questa presenza si manifesta invece principalmente attraverso la presenza di aggettivi in *-ible/-able*, che appaiono in associazione con tutti i connettivi e non appaiono mai nel discorso politico (se non con *mais*). Questo contrasto è in linea con il fatto che i verbi modali fanno maggiormente avvertire la presenza della/del parlante nel suo dire rispetto agli aggettivi modali; basti pensare a una sequenza quale: *les conditions de travail des ouvriers sont améliorables / le condizioni degli operai sono migliorabili*, che appare molto meno 'impregnata' della presenza di chi emette l'enunciato rispetto a: *nous*

*devons/pouvons améliorer les conditions des ouvriers / les conditions des ouvriers doivent être améliorées / dobbiamo/possiamo migliorare le condizioni degli operai / le condizioni degli operai devono essere migliorate.*

I grafici sottostanti mostrano i risultati per la stessa ricerca effettuata sui connettivi italiani già menzionati nelle sezioni 1. e 2. : *ma, però, quindi e ad esempio.*

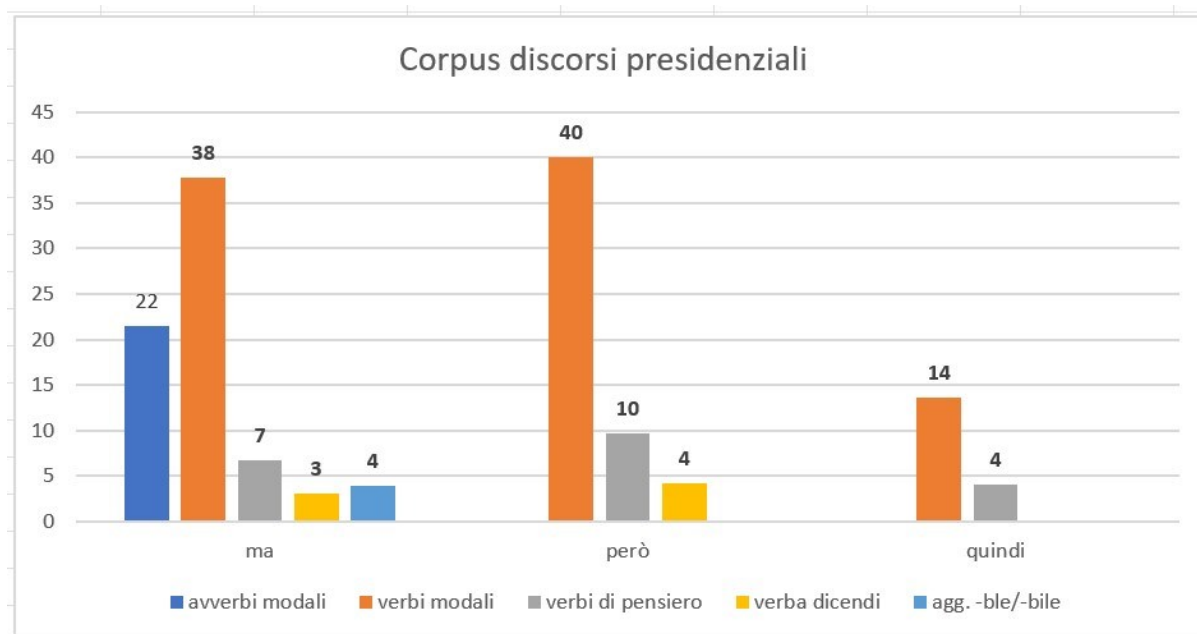


Fig. 7 - Corpus discorsi presidenziali italiani – associazioni specifiche tra categorie di forme soggettive e connettivi

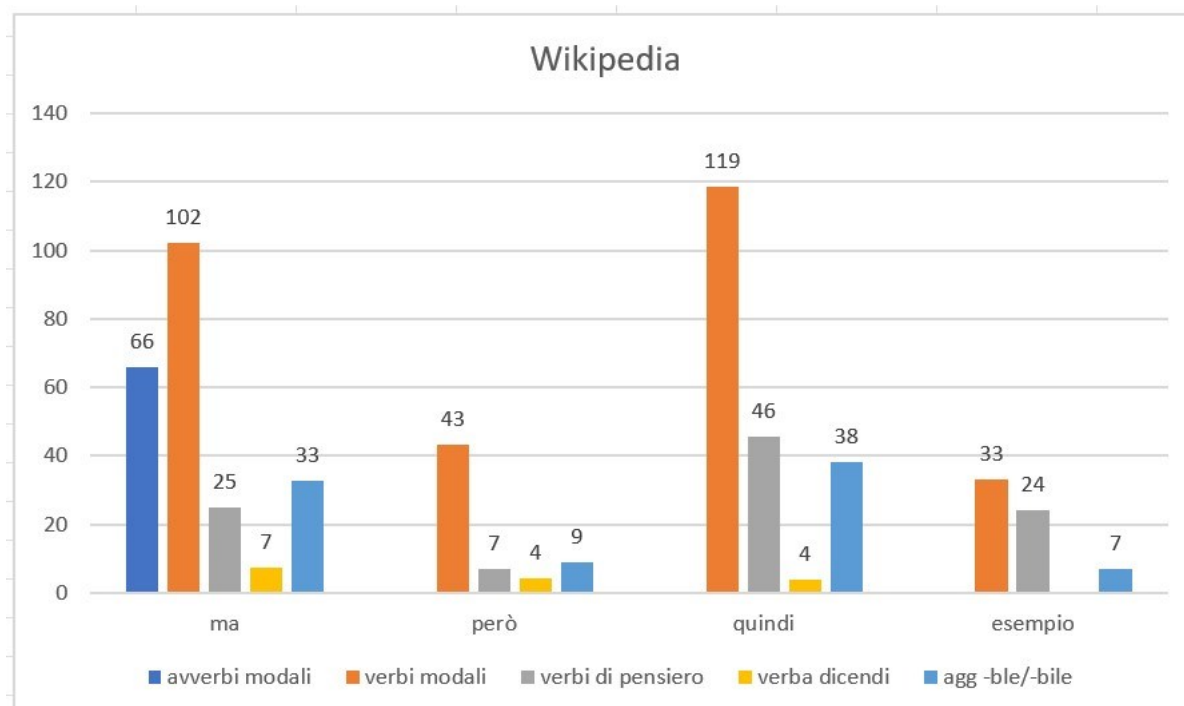


Fig. 8 - Wikipedia italiano – associazioni specifiche tra categorie di forme soggettive e connettivi

Sebbene si possa osservare minore complementarità tra i due corpora rispetto al francese, si rilevano alcuni contrasti netti. In primo luogo, si ritrova – come in francese – una preferenza per gli aggettivi con suffisso *-bile* nel corpus Wikipedia, il che conferma che Wikipedia favorisce l'uso di forme che veicolano la soggettività della/del parlante ponendola sullo sfondo. Si constata inoltre che *per/ad esempio* viene associato con alcune delle nostre 5 categorie soltanto nel discorso enciclopedico e mai nei discorsi presidenziali. Di contro, la categoria dei verbi modali, che in francese sembra particolarizzare il discorso politico e non quello enciclopedico, in italiano risulta rappresentata nei due corpora, si può dire anzi decisamente ben rappresentata, nella misura in cui la si ritrova associata a tutti i connettivi che mostrano associazioni specifiche. Interessante inoltre rilevare che la categoria degli avverbi modali caratterizza solo *ma*, e questo in entrambi i corpora.

In sintesi, confrontando le due lingue e i due corpora si osserva che:

1. Vi sono forme linguistiche proprie di un genere discorsivo, tanto in francese quanto in italiano: si tratta degli aggettivi di possibilità/necessità costruiti con il suffisso *-ible/able/-bile*, che particolarizzano il genere enciclopedico.
2. Vi sono connettivi che presentano un'associazione specifica con una o più categorie di SIFs unicamente in un genere testuale e non nell'altro, tanto in francese quanto in italiano: è il caso di *par exemple* e *per/ad esempio*, che ritroviamo associati ad alcune delle categorie soltanto in Wikipedia, e a nessuna nel corpus di discorsi presidenziali.
3. Vi sono differenze tra francese e italiano: le troviamo ad esempio con i verbi modali, che in francese – ma non in italiano – sono in gioco esclusivamente nel discorso politico.
4. In entrambe le lingue, vi è un connettivo che ha la proprietà di associarsi con il maggior numero di categorie in entrambi i corpora: si tratta di *mais/ma*.

In linea generale, il discorso enciclopedico appare, in entrambe le lingue, più ricco di quello politico per quanto riguarda le categorie di SIFs messe in gioco dai connettivi. Ciò è particolarmente vero per il francese, lingua nella quale il discorso politico fa rilevare due connettivi (*car* e *donc*) attratti esclusivamente dai verbi modali.

## **5. Conclusioni**

I metodi statistici utilizzati nell'analisi dei corpora hanno fornito diversi tipi di apporto alla comprensione del funzionamento della dimensione argomentativa di due generi discorsivi attraverso lo studio di un gruppo di connettivi.

1. La costituzione di scale di soggettività ci ha permesso di osservare le caratteristiche del funzionamento dei connettivi che appaiono costanti e non dipendenti dal genere testuale. Abbiamo visto come vi siano alcuni connettivi in ciascuna lingua che sembrano codificare convenzionalmente un tratto marcato di soggettività: *mais* e *donc*, *ma*, *però* e *quindi*. Specularmente, abbiamo osservato una stabilità per quanto riguarda *par conséquent* e *di conseguenza* – connettivi, questi, che hanno la caratteristica di codificare la soggettività in modo meno marcato.
2. Il calcolo del grado di rappresentazione dei connettivi all'interno dei diversi corpora, con l'introduzione di un *tertium comparationis*, conferma in parte la previsione secondo la quale i connettivi fortemente associati a forme veicolanti la soggettività (*mais*, *donc*, *ma*, *però*) sono in situazione di sottorappresentazione in un genere testuale che si vuole neutro (corpus enciclopedico); un caso, ricordiamo, che a un primo sguardo appare in controtendenza è quello di *quindi*, sovrarappresentato in Wikipedia; come si è visto si tratta però di una controtendenza solo apparente, alla luce degli usi maggiormente non argomentativi nel discorso enciclopedico.
3. Infine, l'analisi delle associazioni specifiche tra un sottogruppo di connettivi, scelti in funzione della loro proprietà di mostrare tendenze stabili o meno relativamente ai corpora interrogati, e le diverse categorie che compongono le SIFs, ci ha permesso di identificare

differenze che riguardano tanto i corpora quanto gli stessi connettivi: abbiamo visto infatti che gli aggettivi in *-ible/-able/-bile* caratterizzano il discorso enciclopedico nelle due lingue; e, per il francese, abbiamo visto invece che i verbi modali sono caratteristici del discorso politico. Per quanto riguarda i connettivi, abbiamo potuto riscontrare che il connettivo di esemplificazione (*par exemple, per/ad esempio*) sembra essere caratteristico del discorso enciclopedico, dato che è sovrarappresentato nel corpus Wikipedia ed è associato a un certo numero di categorie soltanto in quel corpus.

È, insomma, tramite l'incrocio delle diverse tipologie di metodi statistici che si riesce ad ottenere un'istantanea delle relazioni possibili tra l'uso delle forme che garantiscono l'articolazione argomentativa del testo e lo scopo stesso del testo, che ne determina il genere. È evidente che questo tipo di analisi dovrà essere proseguito estendendo via via il campo di ricerca, sia a livello delle forme linguistiche analizzate che a livello dei corpora considerati.

### **Indicazioni bibliografiche**

ANSCOMBRE 2002 = JEAN-CLAUDE ANSCOMBRE, *Mais/pourtant dans la contre-argumentation directe : raisonnement, genericité, et lexicque*, in «Linx», XLVI (2002), pp. 115-131.

ANSCOMBRE/DUCROT 1976 = JEAN-CLAUDE ANSCOMBRE/OSWALD DUCROT, *L'argumentation dans la langue*, in «Langages», XLII (1976), pp. 5-27.

ANSCOMBRE/DUCROT 1977 = JEAN-CLAUDE ANSCOMBRE/OSWALD DUCROT, *Deux mais en français ?* in «Lingua», XLIII (1977), pp. 23-40.

BLUMENTHAL 2002 = PETER BLUMENTHAL, *Profil combinatoire des noms. Synonymie distinctive et analyse contrastive*, in «Zeitschrift für französische Sprache und Literatur», CXII (2002), pp. 115-138.

DUCROT 1984 = OSWALD DUCROT, *Le mots du discours*, Paris, Minuit, 1984.

FERRARI 2022 = ANGELA FERRARI, *Messaggi da Palazzo federale. Osservazioni sulle*

*allocuzioni di Capodanno in lingua italiana dei Presidenti della Confederazione Svizzera* (2011-2021), in «Ricognizioni», IX (2022), pp. 115-139.

HEIDEN 2010 = SERGE HEIDEN, *The TXM Platform: Building Open-Source Textual Analysis Software Compatible with the TEI Encoding Scheme*, in *Proceedings of the 24th Pacific Asia Conference on Language, Information and Computation, Nov 2010, Sendai, Japan*, 2010, pp. 389-398.

LANDOLSI 2018 = HOUDA LANDOLSI, *L'exemplification et ses marqueurs*, Thèse de doctorat, Uppsala University, Disciplinary Domain of Humanities and Social Sciences, Faculty of Languages, Department of Modern Languages, Romance Languages, 2018.

LEBART/SALEM 1994 = LUDOVIC LEBART/ANDRÉ SALEM, *Statistique textuelle*, Paris, Dunod, 1994.

NØLKE *et al.* 2005 = HENNING NØLKE/KJERSTI FLØTTUM/COCO NORÉN, *ScaPoLine: La théorie scandinave de la polyphonie linguistique*, Paris, Éditions Kimé, 2005.

PINCEMIN 2011 = BÉNÉDICTE PINCEMIN, *Sémantique interprétative et textométrie – Version abrégée*, in «Corpus», X (2011), pp. 259-269.

RABATEL 2005 = ALAIN RABATEL, *Les postures énonciatives dans la co-construction dialogique des points de vue : co-énonciation, sur-énonciation, sous-énonciation*, in JACQUES BRES/PIERRE PATRICK HAILLET/SYLVIE MELLET/HENNING NØLKE/LAURENCE ROSIER (dir.), *Dialogisme, polyphonie: approches linguistiques*, Bruxelles, De Boeck Duculot, 2005, pp. 95-110.

RABATEL 2006 = ALAIN RABATEL, *Les auto-citations et leurs reformulations : des surassertions surénoncées ou sousénoncées*, «Travaux neuchâtelois de linguistique», LII (2006), pp. 71-84.

RASTIER 2008 = FRANÇOIS RASTIER, *Que cachent les “données textuelles” ?*, in SERGE HEIDEN/BÉNÉDICTE PINCEMIN (a c. di), *Actes des 9es Journées internationales d'Analyse statistique des Données Textuelles*, Lyon, Presses Universitaires de Lyon, 2008, I, pp. 13-

26.

ROSSARI/RICCI IN STAMPA = *The semantics and pragmatics of discourse markers*, in MAJ-BRITT

MOOSEGARD-HANSEN/JACQUELINE VISCONTI (a c. di) *Manual of Discourse Markers in*

*Romance*.