

GIULIA LOMBARDI¹

La leggibilità dei testi istituzionali italiani destinati agli stranieri

Abstract

The article aims to analyze linguistic complexity within a corpus of Italian institutional texts (ISTR) addressed to migrants and asylum seekers. Lexical, morpho-syntactic and syntactic features retrieved from a computational linguistic analysis in the above-mentioned texts are investigated and compared with the features of four other corpora: AMB and BUR_orig (Italian legal and administrative texts), 2PAR (texts excerpted from an Italian newspaper) and A2 (texts addressed to students of Italian as L2). The results reveal relevant differences only between the linguistic profile of AMB and BUR_orig and the sub-section of institutional and informative texts specifically created for migrants and asylum seekers' needs (such as brochures and advertising materials aimed at disseminating multiple regulations). The general linguistic gap between ISTR and 2PAR and A2 highlights an evident urgency to assess the readability of those texts in order to facilitate communication with the intended users.

1. Introduzione

L'italiano istituzionale (REI 2005) è una varietà dell'italiano comune che viene utilizzata principalmente nella redazione di testi scritti (moduli, manifesti, vademecum, normative, lettere, pubblicità) ascrivibili al dominio del diritto e dell'amministrazione, ma anche dell'informazione e della comunicazione. Ad accomunare questi testi sono l'elaborazione e la trasmissione da parte di un centro dotato di potere normativo e l'appartenenza al *milieu* della comunicazione pubblica, intesa come la «forma di interazione tra le amministrazioni pubbliche e la comunità in funzione dell'interesse generale per la garanzia e la tutela dei diritti della cittadinanza» (Vellutino 2018: 12). Si tratta di testi che forniscono informazioni specialistiche rilevanti per la comunità, comunicazioni di pubblica utilità e messaggi per la coesione sociale o per la promozione del brand istituzionale. Essendo testi fortemente vincolanti, è necessaria una piena comprensione da parte di tutti i cittadini, compresi coloro che si trovano in una situazione socio-culturale problematica o che hanno una scarsa conoscenza della lingua italiana. Il rischio derivante dalla mancata o parziale comprensione di questi testi da parte degli immigrati che si trovano in Italia è grande, non solo in ragione delle funzioni *regolative* della comunicazione istituzionale (non comprendere pienamente un documento potrebbe avere delle

¹ Università di Genova.

conseguenze importanti per il progetto migratorio degli utenti coinvolti), ma anche di quelle *educative* ed *inclusive*, poiché solo attraverso la piena conoscenza delle normative, delle procedure ma anche delle politiche di integrazione previste dal Paese di accoglienza è possibile attivare una vera partecipazione civica e comunitaria. Non a caso la stessa Unione Europea da anni impone ai Paesi membri l'attivazione di progetti interculturali mirati (tra i quali, ad esempio, servizi di mediazione linguistica e di traduzione e interpretariato) a supporto delle Pubbliche Amministrazioni e l'adozione da parte delle stesse di nuove strategie per la comunicazione pubblica (Council of Europe 2017).

Tuttavia, la situazione italiana è piuttosto problematica. Negli uffici pubblici si assiste spesso al fallimento della comunicazione tra istituzioni e cittadini stranieri per almeno due fattori (Rhazzali 2015): l'indisponibilità del personale (per sovraccarico di lavoro, frustrazione e formazione scarsa o nulla) e l'organizzazione pratica degli uffici pubblici (caratterizzata spesso da mancanza di sale di attesa, da orari di apertura ridotti e dalla collocazione degli uffici in zone non raggiungibili dai mezzi pubblici). A ciò si aggiunge l'estrema variabilità regionale e locale dell'offerta dei servizi di mediazione, nella migliore delle circostanze limitati ad alcune ore, a poche lingue e concentrati presso alcuni servizi pubblici a discapito di altri. Inoltre, l'incompletezza del quadro normativo italiano in relazione alle pratiche di intervento fa sì che la selezione dei traduttori e degli interpreti impiegati, per esempio, nei tribunali sia spesso arbitraria, limitata a poche lingue e applicata solo ad alcune fasi dell'iter giudiziario.

Tutto ciò contribuisce a straniare, depotenziare ed emarginare soprattutto gli stranieri di recente immigrazione, che, nonostante la scarsa conoscenza della lingua veicolare, devono sin da subito relazionarsi con diversi tipi di testi istituzionali (permessi di soggiorno, richieste, moduli e iscrizioni). Il presente contributo mira ad analizzare la leggibilità di testi istituzionali italiani destinati agli stranieri per verificare se, almeno dal punto di vista linguistico, siano rispettate le regole universali dei *plain languages* come richiesto dalla normativa europea e italiana.

2. I testi istituzionali italiani

2.1 Caratteristiche linguistiche dei testi istituzionali italiani

Il linguaggio istituzionale è impiegato in una tale varietà di forme testuali e stilistiche che alcuni linguisti hanno proposto di utilizzare la definizione al plurale (*italiani istituzionali*, Vellutino 2018).

Secondo Fortis (2005) e Mortara Garavelli (2001) si tratta di un linguaggio settoriale ma non specialistico perché, pur essendo legato a funzioni comunicative e contesti d'uso specifici, data l'estrema variabilità degli oggetti di comunicazione, non ha un'identità tecnico-specialistica definita, potendo includere termini derivanti da moltissimi altri linguaggi settoriali (medico, economico o politico, per esempio). Il diritto e le amministrazioni utilizzano il linguaggio istituzionale ad un grado elevato di formalità, normatività, complessità ed esclusività; al contrario, la

pubblicità e la comunicazione divulgativa sono caratterizzate da un uso informale, comune, meno restrittivo e specialistico dello stesso linguaggio (Lubello 2014).

Dal punto di vista pragmatico, i testi istituzionali – soprattutto quelli giuridico-amministrativi – sono caratterizzati per lo più da enunciati prescrittivi, atemporali e impersonali. La forza illocutiva di tali enunciati si concretizza nella disposizione gerarchica dei contenuti, nell'utilizzo accorto degli elementi di coesione e nel ricorso frequente a strategie di messa in evidenza del tema degli enunciati: per questa ragione, la morfosintassi è ricca di frasi dichiarative, imperative, causali e ipotetiche (Brunato 2014; Brunato & Venturi 2014; Lastrucci & Viana 1991; Lubello 2014; Venturi 2012). I verbi sono prevalentemente indicativi presenti o futuri deontici-iussivi, mentre il ricorso ai tempi passati narrativi è prevalentemente ristretto alla ricostruzione dei fatti inclusa nelle sentenze o in alcuni testi divulgativi. Nei testi ascrivibili al dominio normativo è escluso il ricorso ad avverbi valutativi (come *addirittura, persino, purtroppo, sfortunatamente*) ed è rara la coniugazione dei verbi alla prima persona o al modo imperativo, a cui viene generalmente preferito l'infinito. Per la stessa ragione è ampio il ricorso alla diatesi passiva, alle forme implicite di subordinazione, all'omissione del soggetto e al participio congiunto o assoluto. Anche gli aggettivi qualificativi sono rari, mentre abbondano i deittici e rimandi testuali interni. Se i principi di sinteticità e semplicità dovrebbero guidare l'organizzazione e la redazione di qualunque testo giuridico e amministrativo, non bisogna dimenticare che «il giurista perfetto scrittore dovrebbe essere conciso ma non involuto» (Bambi 2018: 47): in questi testi proliferano, invece, pronomi, nominalizzazioni, enclisi pronominali, omissioni dell'articolo e strutture preposizionali libere. Per questa ragione si raggiunge spesso un'elevata densità informativa, che si concretizza in frasi lunghe e complesse, con sintagmi nominali pesanti. La necessità di mitigare il valore assertivo degli enunciati si concretizza spesso in doppie negazioni, litoti, eufemismi, perifrasi, ellissi e metafore di cortesia. Per evidenziare il tema, si incontrano frequentemente inversioni nell'ordine non marcato dei costituenti del periodo (il verbo precede il soggetto, l'aggettivo precede il nome, la *coniunctio* relativa si inserisce tra l'articolo e il nome). Dal punto di vista lessicale (Gualdo & Telve 2011), poi, ai tecnicismi necessari si associano spesso tecnicismi collaterali, prassismi, aulicismi, arcaismi, forestierismi, abbreviazioni e formule andate cristallizzandosi con l'uso.

2.2 Contro il *burocratese*

Per le caratteristiche appena delineate, la lingua istituzionale dalla metà del secolo scorso è stata oggetto di numerose illustri critiche: basti pensare al ben noto *burocratese* di Calvino (Calvino 1965).

Da allora non sono mancate iniziative concrete per promuovere la redazione di testi istituzionali leggibili, comprensibili e chiari: manifesti (REI 2005), normative (come la legge n. 24 del 07/08/1990, l'articolo 8 della direttiva del Ministero per la funzione pubblica del 07/02/2002 e la direttiva sulla *Semplificazione del linguaggio delle pubbliche amministrazioni* del 2005), manuali tecnici e pubblicazioni (solo per citarne alcuni: il *Codice di stile delle comunicazioni scritte ad uso delle amministrazioni*

pubbliche del 1993; Cortelazzo & Pellegrino 2003; Fioritto 1997; Ittig & Accademia della Crusca 2011) e persino una petizione lanciata da Annamaria Testa²¹ nel 2015.

Le strategie suggerite convergono su alcuni capisaldi della scrittura chiara e leggibile:

- identificare l'interlocutore;
- selezionare accuratamente le informazioni essenziali da comunicare all'utenza individuata;
- rispettare la temporalità e organizzare in maniera gerarchica le informazioni in modo tale da dare priorità e rilevanza alle informazioni più importanti;
- controllare che ogni sezione del testo rispetti i requisiti base di coerenza e coesione;
- sciogliere il discorso ricorrendo a più frasi distinte ed esplicitando la relazione logica tra gli enunciati;
- limitare il ricorso alla nominalizzazione;
- scrivere frasi che non superino le 25 parole;
- preferire la forma attiva negli enunciati e soprattutto nella subordinazione;
- glossare, parafrasare e disambiguare i tecnicismi necessari;
- utilizzare il più possibile parole appartenenti al Vocabolario di Base (De Mauro 2000);
- evitare tecnicismi collaterali, parole rare, sigle e acronimi non trasparenti.

Il susseguirsi di tante iniziative in parte simili tra loro e convergenti è motivato da quella che Fortis (2005) ha definito «resistenza al cambiamento». Non si tratta di ridondanza: diversi fattori si sono frapposti e tuttora si frappongono all'applicazione delle normative e delle linee guida dei linguisti sopra citate, tanto che ancora nel 2020 la redazione delle varie autocertificazioni necessarie a giustificare gli spostamenti in periodo di *lockdown* ha scatenato non poche critiche e fraintendimenti da parte di cittadini e autorità (Cassese 2020). Tra le ragioni individuate da Fortis ritroviamo l'incompetenza, l'inerzia, la frustrazione e il conformismo dei funzionari incaricati della redazione dei documenti, ma anche il timore che avvicinare la lingua degli atti a quella dei cittadini possa inficiarne l'autorevolezza e limitare quella vaghezza spesso necessaria agli stessi impiegati per declinare ogni responsabilità personale nell'interpretazione dei testi.

Peraltro, nemmeno il ricorso ad una lingua semplificata e totalmente epurata da ogni riferimento tecnico-dotto è sempre e in assoluto preferibile: semplificare troppo ciò che è filosoficamente e teoricamente complesso rischia di imbrogliare e raggirare l'utente (Bambi 2018). Tuttavia, in alcuni casi specifici, l'applicazione di alcune delle norme sopra elencate può veramente fare la differenza tra l'inclusione o l'emarginazione di una larga fetta di utenza. Nel prossimo paragrafo cercheremo di verificare se la redazione dei documenti istituzionali destinati ai cittadini stranieri rispetti o meno tali indicazioni.

21. Vedi <http://chnge.it/rGMFVcLnGS>

3. *Analisi linguistico-computazionale dei testi istituzionali italiani destinati agli stranieri*

3.1 Metodologia del monitoraggio linguistico

Per realizzare l'analisi linguistico-computazionale dei testi istituzionali destinati agli adulti stranieri è stato necessario creare un corpus monotematico e sincronico che fosse il più possibile rappresentativo della varietà linguistica in oggetto. Il corpus, denominato ISTR, comprende:

- 72 atti amministrativi e moduli, ovvero notifiche, auto-certificazioni, iscrizioni, istanze, richieste e verbali (d'ora in poi *Att.a/Mod* nelle figure);
- 22 pronunce giurisdizionali, ovvero gli esiti delle Commissioni Territoriali o i giudizi dei tribunali incaricati di valutare le domande di asilo politico (d'ora in poi *Pr.Giur* nelle figure);
- 5 documenti informativi, ovvero vademecum e guide scritte a scopo divulgativo e realizzate espressamente per l'utenza straniera (d'ora in poi *Doc.Inf* nelle figure).

Il corpus comprende un totale di 142.597 *tokens*. La selezione dei testi e delle tre sotto-categorie è stata motivata dalla scelta di occuparsi nello specifico del linguaggio istituzionale dell'amministrazione, del diritto e della comunicazione utilizzato per la redazione dei testi più significativi per il progetto migratorio dei cittadini stranieri in Italia.

Per procedere alla profilatura del linguaggio istituzionale per gli stranieri e allo stesso tempo valutarne il grado di leggibilità, si è deciso di ricorrere a un monitoraggio comparativo di alcuni tratti morfosintattici, lessicali e sintattici tra ISTR e:

- AMB (Venturi 2011), che comprende testi legali e amministrativi per un totale di 1.627.757 *tokens*;
- BUR_orig (Brunato 2014), che comprende testi burocratici italiani per un totale di 62.208 *tokens*;
- 2PAR (Piemontese 1996), che comprende testi tratti dal mensile di lettura facilitata *DueParole*, per un totale di 72.987 *tokens*;
- A2 (Lombardi 2020), che comprende testi tratti dai principali manuali di italiano per stranieri espressamente dedicati al raggiungimento del livello A2 di conoscenza della lingua italiana L2, per un totale di 28.672 *tokens*.

La metodologia, già ampiamente sperimentata da Montemagni (2013), Dell'Orletta *et al.* (2014), Venturi (2011; 2012) e Brunato (2014), prevede l'analisi di alcuni tratti che, oltre ad essere rappresentativi delle varietà linguistiche dei vari corpora comparati, risultano predittivi del grado di difficoltà dei testi e consentono pertanto di verificare l'occorrenza delle caratteristiche

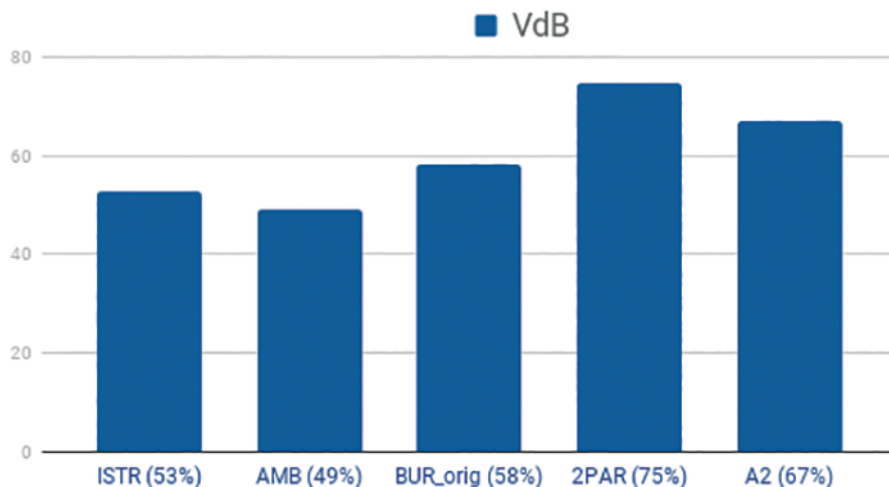
linguistiche suggerite dalle linee-guida per la redazione di testi amministrativi leggibili e chiari (vedi paragrafo precedente). La comparazione tra corpora qualitativamente diversi è finalizzata alla profilatura dei diversi linguaggi istituzionali rappresentati dalle sotto-sezioni del corpus ISTR. Lo scopo è quello di verificare se gli atti amministrativi, la modulistica, le pronunce giurisdizionali e i documenti informativi sulla normativa vigente in Italia in materia di immigrazione presentano tratti linguistici più vicini al linguaggio burocratico e normativo o ai *plain languages*. La scelta di includere un corpus contenente testi esplicitamente redatti per gli apprendenti L2 (A2) permette di discutere meglio il grado di difficoltà dei testi di ISTR in relazione all'utenza straniera. I limiti del metodo comparativo verranno discussi laddove ci si troverà nell'impossibilità di confrontare alcuni specifici tratti linguistici tipici di alcune strutture testuali (come i moduli).

Per l'analisi linguistico-computazionale è stata utilizzata la piattaforma per il trattamento automatico dell'italiano sviluppata dall'Istituto di Linguistica Computazionale Zampolli di Pisa, anch'essa già validata ed utilizzata nei lavori sopra citati, sia per la profilatura linguistica dei testi specialistici che per l'identificazione delle marche testuali di difficoltà (Dell'Orletta *et al.* 2011).

3.2 Caratteristiche lessicali dei testi istituzionali italiani per stranieri

La relativa limitatezza del corpus ISTR, che è giustificata dalla specificità dei testi in esso contenuti, non permette di ricavare valori significativi riguardo alla densità e alla *type/token ratio*. Tuttavia, l'analisi della percentuale di parole in termini di lemmi diversi appartenenti al Vocabolario di Base (VdB) individuato da De Mauro (fig. 1) rivela che i testi istituzionali destinati agli stranieri presentano valori in linea con quelli riportati da AMB e BUR_orig.

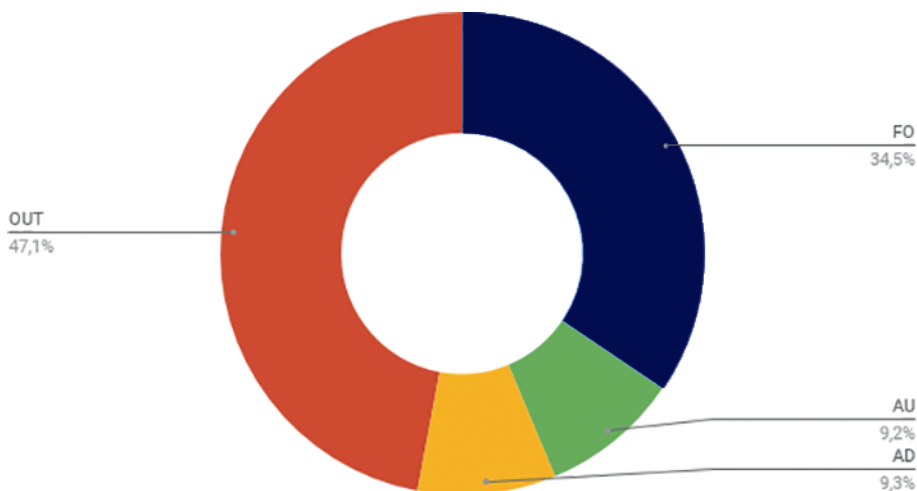
Figura 1 - Percentuale di parole in termini diversi appartenenti al Vocabolario di Base di De Mauro



Il VdB rappresenta le circa 7500 parole usate e comprese dalla maggior parte degli italofofoni: per definire quanto i testi in oggetto siano lessicalmente comprensibili per l'utenza straniera, occorre focalizzarci sulla sottocategoria del Lessico Fondamentale (FO) contenuta nel VdB, ovvero quelle circa 2000 parole comuni e frequenti che costituiscono anche il *lexical threshold* (Laufer 2010; Laufer & Aviad-Levitzky 2017) per la comprensione dei testi in lingua seconda.

Osservando la figura 2, si nota che all'interno di ISTR solo il 34,5% delle parole appartiene alla categoria del lessico Fondamentale (FO), mentre il 18% appartiene alle categorie ad Alto Uso (AU) e ad Alta Disponibilità (AD); sommando questo dato alla percentuale di parole estranee al VdB precedentemente discussa (OUT), potremmo ipotizzare che il 65,1% delle parole dei testi in oggetto sono sconosciute o poco note all'utenza straniera.

Figura 2 - Distribuzione del lessico di base in ISTR



È ipotizzabile che all'interno di questa categoria rientrino, accanto a numerosi vocaboli tecnici indispensabili (*comma, sussidiaria, status, ricongiungimento*), anche tecnicismi collaterali e preziosismi (*produrre duplice copia di, congiuntamente a, nonché, mendace, pena pecuniaria*). All'interno di ISTR, la percentuale di lessico fondamentale è così distribuita: 63,89 % negli atti amministrativi e nella modulistica, 68% nelle pronunce giurisdizionali e 72% nei documenti informativi. Questi ultimi sono i testi che sembrano formulati tenendo maggiormente in considerazione le esigenze dei destinatari individuati.

Se verificiamo la distribuzione percentuale di parole in termini di lemmi diversi appartenenti alla lista lessicale di livello A2 prevista dal Profilo della lingua italiana (Spinelli & Parizzi 2010), ricaviamo che solo il 21,14% delle parole contenute in ISTR appartiene alla categoria menzionata: in particolare, vi rientrano il 23,92% delle parole contenuti negli atti amministrativi e nella modulistica, il

23,33% delle parole contenute nei documenti informativi e il 16,17% delle parole contenute nelle pronunce giurisdizionali.

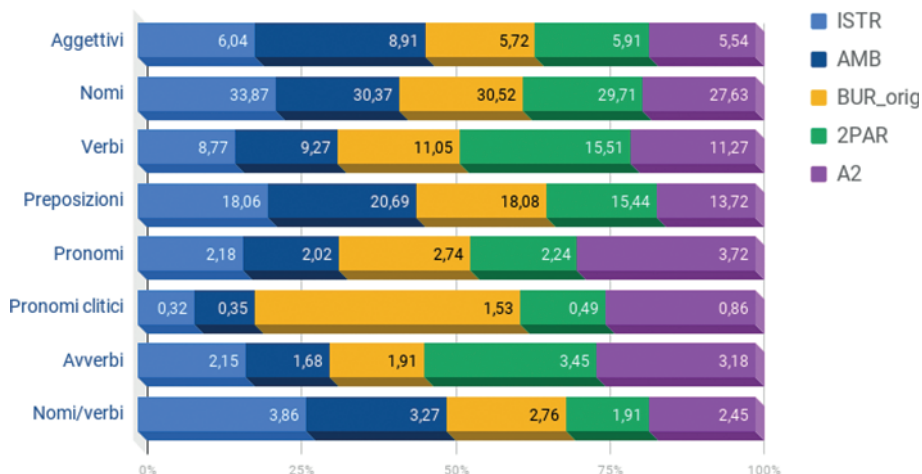
Ciò che colpisce di questi dati è che i testi inclusi nella categoria delle pronunce giurisdizionali sono in buona parte esiti di richieste di protezione internazionale inseriti in pratiche che generalmente si svolgono nel primo anno di ingresso dei richiedenti: durante questo spazio temporale (anche in considerazione della spesso scarsa scolarizzazione pregressa dei richiedenti) difficilmente viene superato il livello A2 di conoscenza della lingua italiana. Tale soglia risulta significativa anche perché coincide con il livello A2 di abilità linguistica richiesto dallo Stato italiano a tutti coloro che effettuano la domanda per il rilascio del permesso di soggiorno di lungo periodo.

3.3 Caratteristiche morfosintattiche dei testi istituzionali italiani per stranieri

L'analisi della distribuzione percentuale tra le categorie morfosintattiche primarie rappresenta una delle principali marche di stile (Biber 1993). Una percentuale maggiore di nomi caratterizzerebbe i testi informativi rispetto a quelli narrativi, mentre un elevato valore nel rapporto tra verbi e nomi sarebbe indice di elevata astrazione e densità informativa (Montemagni 2013).

I dati rappresentati nella figura 3 si presentano in linea con la letteratura, con ISTR che apparentemente presenta un'elevata percentuale di nomi e un elevato valore del rapporto tra nomi e verbi, superiori ad AMB e BUR_orig.

Figura 3 - Distribuzione percentuale tra le categorie morfosintattiche primarie

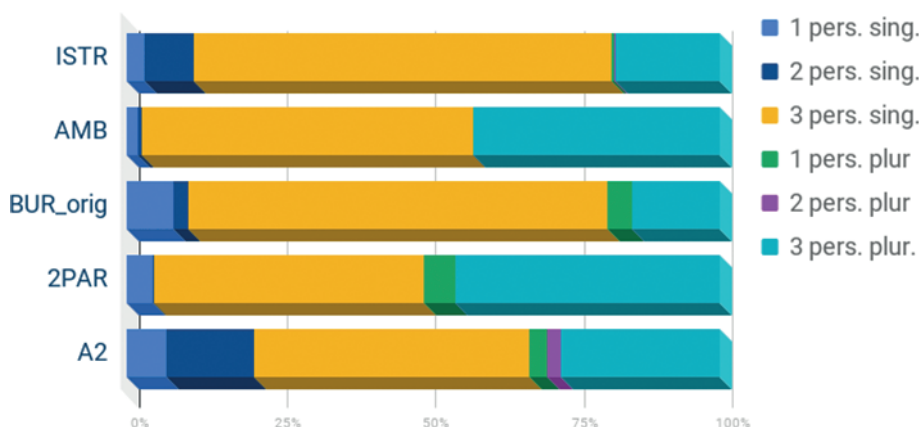


Tuttavia, ad influenzare particolarmente la *noun/verb ratio* è l'elevata percentuale di nomi presente negli atti amministrativi e nella modulistica (nomi/verbi= 5,06), dovuta alla prevalenza di moduli che per struttura testuale spesso non prevedono l'inserzione di verbi. Infatti, senza la modulistica, il valore di ISTR scende a 2,7. Quanto alla categoria delle pronunce giurisdizionali, il dato (2,75)

è in linea con i dati riportati da AMB ed è indice dell'elevata nominalizzazione presente in questo tipo di testi. Il principale problema per l'utenza straniera potrebbe essere costituito dall'astrattezza che deriva da queste scelte stilistiche, considerato che gli studenti di livello A2, in base alle tappe di acquisizione linguistica (Bernini 2003; Lo Duca 2007), generalmente conoscono e utilizzano principalmente vocaboli concreti e di uso comune. I dati relativi ai documenti informativi scritti per l'utenza straniera (2,18) mostrano nuovamente un'attenzione maggiore alla leggibilità, con la ricerca di una scrittura il più possibile sciolta e discorsiva.

Osservando la distribuzione percentuale della coniugazione verbale, riportata in figura 4, in linea con quanto prospettato dalla letteratura sul linguaggio istituzionale, emerge come BUR_orig, AMB e ISTR siano caratterizzati da un ricorso massiccio alla terza persona singolare, probabile spia di quella tendenza alla spersonalizzazione descritta nel § 2.

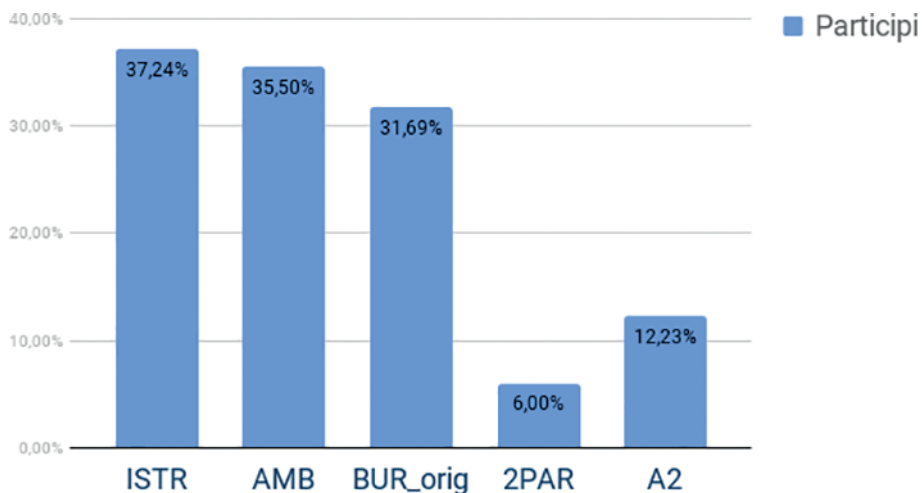
Figura 4 - Distribuzione percentuale della coniugazione verbale



Rispetto a BUR_orig (0,84) e AMB (0,78%), ISTR mostra una percentuale media maggiore di verbi alla seconda persona singolare (7,14%): mentre, però, le pronunce giurisdizionali (con l'1,31%) si attestano su valori simili al linguaggio normativo, gli atti amministrativi e la modulistica per stranieri (con il 6,6%) presentano un dato sensibilmente più elevato di BUR_orig. Tuttavia, a risultare particolarmente elevata (43,33%) è la percentuale dei verbi alla seconda persona singolare contenuti nei documenti informativi. Si tratta di un dato molto positivo nella valutazione della leggibilità dei testi per l'utenza straniera: rivolgersi direttamente all'interlocutore non solo risulta efficace e funzionale dal punto di vista comunicativo, ma è anche un uso linguistico estremamente familiare agli stranieri, considerato che gran parte dei manuali relativi al livello A2 propongono soprattutto forme testuali dialogiche. A riprova di ciò, il valore suddetto si attesta sul 13% per il corpus A2.

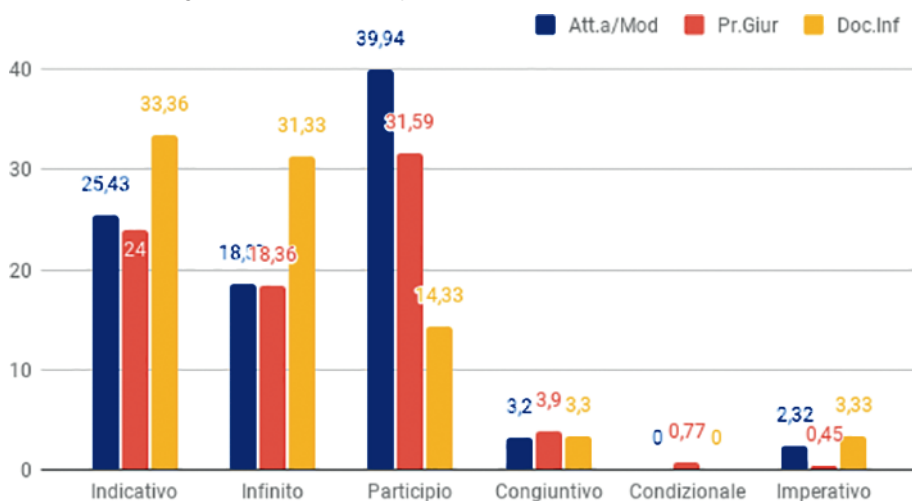
Per quanto riguarda la distribuzione percentuale dei participi non inclusi nei tempi composti (fig. 5), indice di nominalizzazione, di personalizzazione e di ricorso a forme implicite di subordinazione, si può osservare come il valore di ISTR (37,24%) sia vicino a quello di AMB (35,5%) e BUR_orig (31,69%), discostandosi invece dai testi di DueParole (6%) e dei manuali di italiano L2 di livello A2 (12,23%).

Figura 5 - Distribuzione dei participi non inclusi in passato prossimo o trapassato prossimo



La figura 6 mostra come siano soprattutto gli atti amministrativi con la modulistica (39,94%) e le pronunce giurisdizionali (31,59%) a contribuire al raggiungimento di un numero così elevato.

Figura 6 - Distribuzione percentuale dei modi verbali in ISTR



I documenti informativi, attestandosi sul 14,33%, presentano nuovamente un dato in linea con il profilo linguistico dei testi semplificati. Analogamente, prevalgono l'uso dell'indicativo e dell'infinito, possibile spia di un ricorso maggiore a frasi infinitive e dichiarative e a subordinate esplicite. L'occorrenza di una buona percentuale di imperativi è probabilmente correlata al dato relativo alla coniugazione verbale, come conferma di una generale tendenza all'utilizzo della seconda persona singolare.

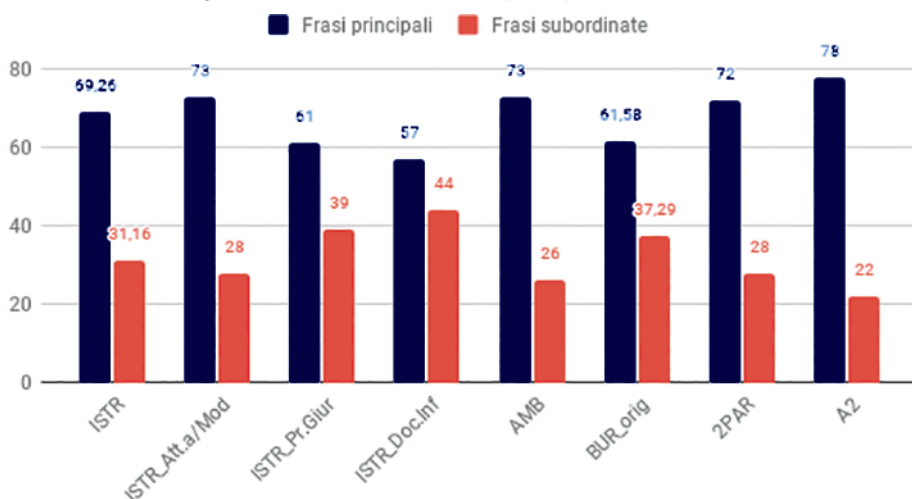
3.4 Caratteristiche sintattiche dei testi istituzionali italiani per stranieri

Uno dei dati che maggiormente correla con la leggibilità dei testi è la lunghezza media delle frasi, calcolata come numero medio di *tokens* per frase; non a caso, le indicazioni dei linguisti a proposito della riscrittura dei testi istituzionali fanno riferimento a una lunghezza massima di 25 parole.

I dati raccolti dall'analisi linguistico-computazionale rivelano che mediamente in ISTR si hanno 36,24 *tokens* per frase, contro il 26,72% di BUR_orig e il 26% AMB. All'interno di ISTR sono le pronunce giurisdizionali a presentare gli enunciati più lunghi, con una media di 38,27% di parole per frase. I documenti informativi, invece, presentano un dato (26,6%) in linea con AMB e BUR_orig e nemmeno troppo lontano da 2PAR (18,67%) e A2 (12,68).

Per quanto riguarda l'analisi della distribuzione percentuale media delle frasi subordinate e principali, riassunta nella figura 7, possiamo concludere che i dati di ISTR si collocano idealmente tra quelli di AMB e di BUR_orig, non discostandosi troppo nemmeno da 2PAR.

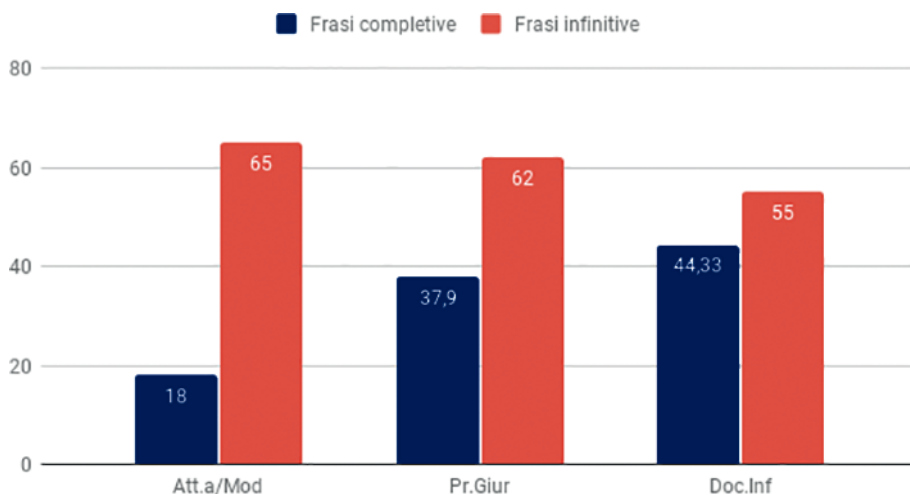
Figura 7 - Distribuzione di frasi principali e subordinate



Ciò che colpisce, in particolare, è l'elevata percentuale di frasi subordinate presenti nei documenti informativi destinati agli stranieri. Sembra un dato discordante con quanto discusso finora, ma per avere un quadro più chiaro della sin-

tassi dei testi dobbiamo preliminarmente ricordare che la subordinazione non è necessariamente sinonimo di complessità: ciò che è importante è il ricorso a frasi esplicite che ben segnalino il nesso logico tra gli enunciati (Bambi 2018). Osservando la distribuzione interna a ISTR (fig. 8) si può notare come i documenti informativi siano caratterizzati da un numero elevato di completeive e da un numero inferiore di infinitive.

Figura 8 - Distribuzione delle frasi infinitive e completeive in ISTR



Questo dato sembra mostrare una preferenza dei redattori per l'esplicitazione della subordinazione e per la disambiguazione dei soggetti degli enunciati, come emerge dal seguente esempio tratto dal *Vademecum sulla richiesta di protezione internazionale*:

- (1) Se sei privo di mezzi di sussistenza e non sei accolto in un centro di accoglienza governativo, puoi chiedere alla Questura che attivi la Prefettura per l'accoglienza in uno dei centri predisposti.

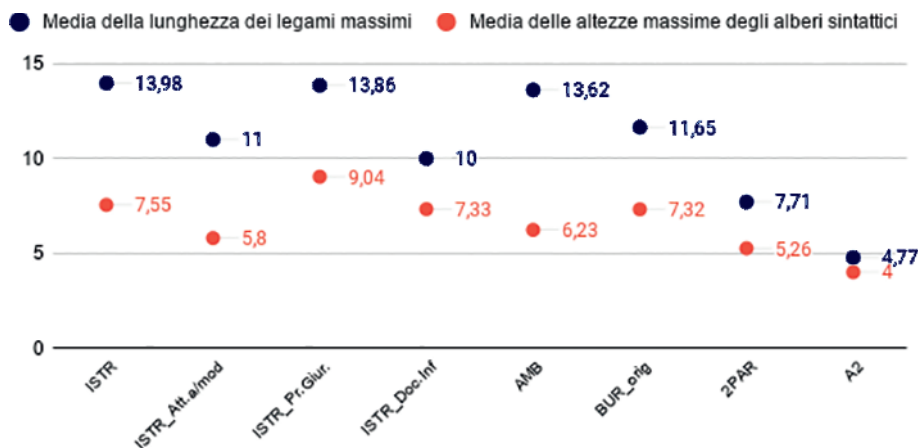
La presenza relativamente ridotta di frasi infinitive, invece, se correlata ai dati sulla distribuzione dei modi verbali riportata nella figura 6, rivela che probabilmente una buona parte delle occorrenze di infiniti precedentemente segnalata è da ascrivere a forme negative di imperativo.

Per indagare ulteriormente le caratteristiche linguistiche dei testi destinati agli stranieri e allo stesso tempo valutarne la leggibilità, occorre discutere ancora due dati fondamentali: la distanza media massima tra testa e dipendente nelle relazioni di dipendenza (calcolata come media della lunghezza dei legami di dipendenza massimi) e i livelli di incassamento gerarchico delle frasi (ovvero la media delle altezze massime degli alberi sintattici).

Come risulta dalla figura 9, in ISTR mediamente occorrono 13 parole tra testa e dipendente, un dato in linea con quanto emerge per AMB, ma molto lontano dai valori di 2PAR e A2. In altri termini, anche i dati sulla complessità sin-

tattica mostrano che ci sono maggiori analogie tra ISTR e i testi amministrativi e burocratici rispetto ai testi di facile leggibilità e a quelli pensati appositamente per gli stranieri con un livello di conoscenza dell'italiano inferiore all'A2.

Figura 9 - *Analisi della complessità sintattica*



Come è noto, più il nesso è distanziato, più risulta difficile per il lettore identificare la relazione tra testa e dipendente (Lumbelli 1989). Un numero medio elevato di parole correla spesso con la presenza di aggiunte relativizzanti e problematizzati, che costituiscono grossi ostacoli cognitivi alla comprensione dei testi. Si legga ad esempio la frase tratta dalla sentenza n. 2985/2016 del Tribunale di Perugia:

- (2) *L'essersi* il ricorrente successivamente *dissociato* dalla confraternita e per questo motivo avendo subito reiterate intimidazioni e minacce di morte nonostante l'avvenuto cambiamento per ben due volte della propria residenza in altre città della Nigeria, diversamente da quanto ritenuto dalla Commissione nel provvedimento impugnato, *costituisce* un fondato motivo per riconoscere al ricorrente la protezione umanitaria.

In linea con quanto discusso, anche l'analisi del numero medio massimo di relazioni consecutive tra foglia e radice verbale all'interno dell'albero sintattico colloca ISTR (e soprattutto le pronunce giurisdizionali) al primo posto per quanto riguarda l'incassamento gerarchico. Esemplificativa risulta la frase tratta dalla sentenza RG 2284/2017 del Tribunale di Genova:

- (3) In assenza di elementi di giudizio specifici relativamente alle caratteristiche di tali provvidenze asseritamente escluse dal principio della parità di trattamento (quali, requisiti previsti per l'erogazione di detti benefici; modalità di erogazione degli stessi ed altro) questo Giudice non ha ragione di porsi il dubbio rappresentato dal quesito formulato dall'Istituto.

In entrambi i casi, i documenti informativi si attestano su numeri inferiori rispetto ad AMB e BUR_orig ma ancora lontani da quelli caratterizzanti i testi semplificati.

4. Conclusioni

Le metodologie computazionali per il trattamento automatico del testo dimostrano che complessivamente i testi istituzionali destinati agli stranieri presentano tratti lessicali e morfo-sintattici più vicini al profilo linguistico dei testi burocratici e normativi che a quello dei *plain languages*. Tuttavia, all'interno del corpus ISTR esistono differenze significative tra le varie sottocategorie. I documenti informativi, scritti per divulgare norme e procedure, e indirizzati espressamente all'utenza, mostrano diverse caratteristiche linguistiche in comune con i testi ad alta leggibilità e in linea con le strategie suggerite dalla letteratura sulla semplificazione dei linguaggi istituzionali: frasi più brevi, meno participi, più imperativi e verbi coniugati alla seconda persona, subordinazione esplicita, percentuale elevata di parole appartenenti al VdB, *noun/verb ratio* limitata, limitata densità informativa. Gli atti amministrativi, la modulistica e le pronunce giurisdizionali si rivelano invece molto più complessi, presentando molti tratti in comune con AMB e BUR_orig, spesso in percentuale ancora maggiore: si pensi ad esempio ai livelli di incassamento gerarchico, alla lunghezza degli enunciati e alla *noun/verb ratio*. È vero che non si tratta di testi specificamente divulgativi, ma sono pur sempre documenti destinati a un'utenza precisa: comunicare a un richiedente asilo che

- (4) alla luce di quanto detto, la Corte ritiene che, per la sussistenza di una tale situazione di insicurezza, il xxxxxx, che è lontano dal Paese d'origine dall'aprile 2014, potrebbe subire ripercussioni dannose in caso di rimpatrio e ciò induce a riconoscergli, in riforma dell'impugnata ordinanza, un permesso di soggiorno per ragioni umanitarie (Sentenza n. 1036/2017 pubbl. il 09/03/2017 RG n. 154/2016)

è quantomeno poco efficace. Occorre senz'altro mettere gli stranieri che si trovano nel nostro Paese nelle condizioni non solo di capire e conoscere le leggi e le procedure che li riguardano, ma anche di compilare e comprendere autonomamente i documenti e gli atti processuali che fanno parte di quelle procedure.

Semplificare la redazione dei testi destinati agli stranieri è urgente e necessario, ma occorre ricordare che aumentare la leggibilità di tutte le categorie di testi gioverebbe in generale all'intera categoria *istituzionale*: basti pensare che delle 133 ordinanze d'irricevibilità (al 2018) relative alle domande di pronuncia pregiudiziale presentate alla Corte di Giustizia dell'Unione Europea, il 30% erano italiane e, come è stato evidenziato (Visconti 2018), i testi sono stati rigettati non per la lacunosità delle informazioni, ma per la prolissità e l'oscurità del linguaggio da cui erano composte.

Bibliografia

- BAMBI F. (2018), Per un breviario di buona scrittura giuridica, in *Quaderni Fiorentini*, 47: 43-53.
- BERNINI G. (2003), Come si imparano le parole. Osservazioni sull'acquisizione del lessico, in L2 in *Itals*, I(2): 23-47.
- BIBER D. (1993), Using Register-Diversified Corpora for General Language Studies, in *Computational Linguistics*, 19(2): 219-241.
- BRUNATO D. (2014), Complessità necessaria o stereotipi del “burocratese”? Un'indagine sulla leggibilità del linguaggio amministrativo da una prospettiva linguistico-computazionale, in *La lingua variabile nei testi letterari, artistici e funzionali contemporanei: analisi, interpretazione, traduzione. Atti del XIII Congresso della SILFI*, Palermo: 32-35.
- BRUNATO D. & VENTURI G. (2014), Le tecnologie linguistico-computazionali nella misura della leggibilità di testi giuridici, in *Informatica e Diritto*, 23(1): 111-142.
- CALVINO I. (3 febbraio 1965), L'antilingua, in *Il Giorno*.
- CASSESE S. (1993), *Codice di Stile delle comunicazioni scritte ad uso delle amministrazioni pubbliche*, Presidenza del Consiglio dei Ministri, Roma.
- CASSESE S. (2020), *Coronavirus, il dovere di essere chiari*, https://www.corriere.it/editoriali/20_marzo_23/dovere-essere-chiari-b5b36828-6d39-11ea-ba71-0c6303b9bf2d.shtml?fbclid=IwAR0Z87d-R8EHPenSMSL08Xc8-IorOIgozLTQmBy05BodGLu9J_-QUYMkQ2g.
- CORTELAZZO M. & PELLEGRINO F. (2003), *Guida alla scrittura istituzionale*, Laterza, Roma.
- COUNCIL OF EUROPE (2017), *Adult migrants: Integration and education. Extracts from conventions, recommendations, resolutions, and reports*, <https://rm.coe.int/1680306f0b>.
- DE MAURO T. (2000), *Grande dizionario dell'italiano dell'uso (GRADIT)*, UTET, Torino.
- DELL'ORLETTA F., MONTEMAGNI S., VENTURI G. & MORUZZI G. (2011), READ - IT: Assessing Readability of Italian Texts with a View to Text Simplification, in *Proceedings of the 2nd Workshop on Speech and Language Processing for Assistive Technologies (SLPAT 2011)*: 73-83.
- DELL'ORLETTA F., MONTEMAGNI S. & VENTURI G. (2014), Assessing document and sentence readability in less resourced languages and across textual genres, in *Recent Advances in Automatic Readability Assessment and Text Simplification. Special Issue of the International Journal of Applied Linguistics*, 165(2): 163-193.
- FIORITTO A. (1997), *Manuale di stile*, Il Mulino, Bologna.
- FORTIS D. (2005), Il linguaggio amministrativo italiano, in *Revista de Llengua i Dret*, 43: 47-116.
- GUALDO R. & TELVE S. (2011), *Linguaggi specialistici dell'italiano*, Carocci, Roma.
- ITTIG & ACCADEMIA DELLA CRUSCA (2011), *Guida alla redazione degli atti amministrativi*, ITTIG, Firenze

- LASTRUCCI E. & VIANA L. (1991), La struttura del linguaggio normativo e la comprensione dei testi di natura giuridico-burocratica, in *Scuola e Società*, 12: 531-554.
- LAUFER B. (2010), Lexical threshold revisited: Lexical text coverage, learners vocabulary size and reading comprehension, in *Reading in a Foreign Language*, 22(1): 15-30.
- LAUFER B. & AVIAD-LEVITZKY T. (2017), What Type of Vocabulary Knowledge Predicts Reading Comprehension: Word Meaning Recall or Word Meaning Recognition?, in *Modern Language Journal*, 101(4): 729-741.
- LO DUCA M. G. (2007), Quante e quali parole nell'insegnamento dell'italiano L2? Riflessioni in margine alla costruzione di un Sillabo, in PISTOLESI E. (a cura di), *Lingua, scuola e società. I nuovi bisogni comunicativi nelle classi multiculturali*, Istituto Gramsci, Trieste: 135-150.
- LUBELLO S. (2014), *Il linguaggio burocratico*, Carocci, Roma.
- LUMBELLI L. (1989), *Fenomenologia dello scrivere chiaro*, Editori Riuniti, Roma.
- MONTEMAGNI S. (2013), Tecnologie linguistico-computazionali e monitoraggio della lingua italiana, in *Studi Italiani Di Linguistica Teorica e Applicata (SILTA)*, 1: 145-172.
- MORTARA GARAVELLI B. (2001), *Le parole e la giustizia*, Einaudi, Torino.
- REI (2005), *Parole chiare per tutti*, http://ec.europa.eu/translation/italian/rei/about/documents/manifesto_italiano_istituzionale_qualita_it.pdf.
- RHAZZALI M. K. (2015), *Comunicazione interculturale e sfera pubblica*, Carocci, Roma.
- SPINELLI B. & PARIZZI F. (2010), *Profilo della lingua italiana*, La Nuova Italia, Firenze.
- VELLUTINO D. (2018), *L'italiano istituzionale per la comunicazione pubblica*, Il Mulino, Bologna.
- VENTURI G. (2011), *Lingua e diritto: una prospettiva linguistico-computazionale*, tesi di laurea.
- VENTURI G. (2012), Investigating legal language peculiarities across different types of Italian legal texts: an NLP-based approach, in *IALF Porto*: 138-156.
- VISCONTI J. (2018), La chiarezza tra superfluo e necessario, in *Quaderni fiorentini*, 47: 15-20.