

Regolamenti nazionali ed europei in italiano a confronto: analisi linguistica qualitativa e quantitativa di un corpus*

ELENA TOMBESI

Università del Piemonte Orientale, elena.tombesi@uniupo.it

ABSTRACT

The aim of this paper is to demonstrate the greater linguistic simplicity of European regulations compared to the one found in Italian regulations, mainly because of the well-known universals of translation (simplification, explicitation, normalization and levelling out), and because of precise graphic choices. Through linguistic analyses carried out with the help of recent computational systems, it has been possible to monitor the structure of "Italian translations" in relation to the Italian used in national regulations, with a particular focus on morphosyntactic and syntactic levels, which usually constitute the main issue with this type of texts. In order to facilitate and standardize the interpretation of qualitative and quantitative analyses, each corpus ("corpus IT" and "corpus UE") consists of 25 regulations currently in force, issued over the last decade and belonging to the same categories of topics: employment, environment, energy, social policies and market. Despite some limitations due to automatic analysis of texts, the goals achieved make it possible to highlight the still unexplored potential for the adoption of these instruments in this field of study. This could therefore represent the starting point for a series of future developments centered on the monitoring of relevant linguistic features and possible linguistic interferences.

KEYWORDS

Legal Language, Universals of Translation, Linguistic Variation, Corpus Linguistics

* Questo articolo è la rielaborazione della tesi di laurea magistrale dell'autrice, discussa a novembre 2018 a conclusione del corso di Italianistica, Culture letterarie europee, Scienze linguistiche dell'Università di Bologna sotto la direzione del professor Matteo Viale.

Agli inizi degli anni novanta, sistemi computazionali sempre più evoluti hanno portato al progresso dell'analisi linguistica rivolta alla lingua del diritto, in seguito all'accresciuta consapevolezza di dover impiegare maggiore cura nella redazione dei testi normativi e amministrativi affinché possano risultare più chiari e comprensibili a tutti. Dunque, se la lingua del diritto è destinata a dover comunicare efficacemente con i cittadini, è necessario che essa risulti chiara nei contenuti e nella strutturazione, precisa tanto da non permettere equivoci, semplice nelle scelte lessicali e sintattiche, economica poiché deve contenere tutto il necessario informativo, ma in maniera sufficiente¹.

Mentre le prime ricerche automatiche del Linguaggio erano volte all'analisi dei linguaggi altamente specialistici, come quello biomedico², solo qualche decennio dopo viene indirizzata l'attenzione anche a testi non nettamente separati da quello comune, come quello giuridico, un particolare tipo di linguaggio specialistico (Ondelli 2007, Gualdo 2011, Dell'Anna 2008): si tratta dello studio di Charrow e colleghi che «focalizzano l'attenzione delle loro analisi i) sui fattori storici, sociologici, politici, ecc., che hanno determinato la differenza tra il "legal sublanguage" e la lingua comune e ii) sulle possibilità di rendere in futuro tale linguaggio più accessibile ai non addetti ai lavori» (Venturi 2011: 54).

Le prime attività che hanno come scopo l'informatizzazione del processo di produzione normativa (legimatica) (Mercatali 1995) sono finalizzate alla progettazione di metodi e tecniche informatiche perlopiù rivolte alla correzione ortografica dei testi, ai controlli relativi alla loro leggibilità «e alle tecniche di disambiguazione appoggiate su approcci di intelligenza artificiale» (Venturi 2011: 31). Tuttavia, gli strumenti di analisi automatica del testo allora a disposizione non permettevano un'analisi puntuale su quali fossero le caratteristiche linguistiche realmente indicative della minor/maggior difficoltà dei testi giuridici. È così che si sono sviluppati, più recentemente, ricerche informatiche rivolte allo sviluppo di strumenti evoluti di controllo e gestione del testo normativo redatto: «ciò è reso possibile grazie a metodi di gestione della documentazione giuridica basati su tecniche e strumenti di Trattamento Automatico del Linguaggio» (*ibid.*).

Questa ricerca prende le mosse da un precedente studio condotto da Cortelazzo (Cortelazzo 2013) incentrato su un confronto quantitativo tra leggi italiane e direttive europee. È dunque sulla base di queste premesse che si sviluppa lo studio seguente, col quale vengono messi a confronto testi giuridici nazionali ed europei, combinando a una più tradizionale analisi linguistica di

¹ Tra i contributi più noti, segnalo: Fioritto (1997), Cortelazzo, Pellegrino, Viale (1999).

² Il progetto considerato pioniere in questo ambito è del 1965 ed è intitolato *Linguistic String Project*.

tipo qualitativo, un'indagine quantitativa condotta tramite i software Read-it³ e Profiling-ud⁴. Entrambi gli strumenti sono stati sviluppati dall'Istituto di Linguistica Computazionale "Antonio Zampolli" (ILC-CNR) di Pisa e sono grado di combinare l'analisi dei tratti linguistici lessicali con quelli relativi al livello sintattico e morfosintattico, vero nodo della complessità linguistica di questo genere di testi. L'elaborazione software di dati quantitativi prelevati è basata su due corpora da me appositamente costituiti e sufficientemente ampi (per maggiori informazioni e approfondimenti si veda Bolasco 2005: 1-37).

A causa della prassi che sembra essersi imposta all'interno delle Istituzioni europee, i testi europei, frutto perlopiù di traduzioni dalla lingua inglese, risultano più chiari e semplici da leggere rispetto ai testi nazionali; l'obiettivo è dunque quello di dimostrare, partendo dalle teorizzazioni promosse da Mona Baker circa gli universali traduttivi (Baker 1996)⁵, la maggior semplicità dei regolamenti europei rispetto a quelli nazionali analizzando i fenomeni linguistici rilevabili in confronto, laddove possibile, ai quattro universali traduttivi (esplicitazione, semplificazione, normalizzazione e convergenza) e dedicando ad ognuno un paragrafo.

1. LA LEGGE DELL'INTERFERENZA E GLI UNIVERSALI TRADUTTIVI

Secondo la legge di interferenza teorizzata da Toury nel 1995 all'interno del volume *Descriptive Translation Studies and Beyond*, i fenomeni linguistici tipici della lingua di partenza tendono ad essere trasferiti al testo di destinazione. Tuttavia occorre tenere a mente che la misura in cui tale interferenza viene realizzata dipende sia dall'esperienza professionale del traduttore e sia dalle condizioni socioculturali in cui una traduzione viene prodotta. In altre parole, se i traduttori sono degli esperti (proprio come quelli che operano all'interno delle istituzioni europee), essi tendono ad essere meno influenzati dalla lingua del testo sorgente rispetto a quanto avviene nel caso di traduttori non professionisti. Ovviamente anche il prestigio legato al tipo di testo e al contesto in cui esso viene prodotto modifica il funzionamento di questa legge. La traduzione tecnica, ad esempio derivante da un linguaggio settoriale,

³ Cfr: Dell'Orletta F. (2011) (ed.) "READ-IT: assessing readability of Italian texts with a view to text simplification", in *SLPAT '11 Proceedings of the Second Workshop on Speech and Language Processing for Assistive Technologies*, Edimburgo, Association for Computational Linguistics, pp. 73-83.

⁴ Brunato D. (2020) (ed.) "Profiling-UD: a Tool for Linguistic Profiling of Texts", in *Proceedings of 12th Edition of International Conference on Language Resources and Evaluation (LREC 2020)*, Marseille, France, pp. 7145-7151.

⁵ Tali universali, teorizzati appunto da Baker, sono fenomeni costanti che caratterizzano l'agire del traduttore e che si ripercuotono inevitabilmente sull'assetto del testo di arrivo.

può mostrare un livello di interferenza più basso rispetto a quello stimato all'interno di traduzione letteraria (Baker & Saldanha 2011: 307).

Questo è il motivo per cui, se in parte è fondamentale tenere conto della legge d'interferenza nei risultati ottenuti nelle successive ricerche, d'altra parte sarà possibile svolgere diverse analisi sui testi indipendentemente dall'influsso della lingua di partenza. Sarà proprio Mona Baker nel 1996 che, distinguendo i fenomeni che possono essere attribuiti al contatto tra due lingue specifiche e i fenomeni che invece discendono dal processo traduttivo, offrirà definizioni esaustive delle tendenze linguistiche che danno luogo ai cosiddetti *Translation universals* (Baker 1996): esplicitazione, semplificazione, normalizzazione e convergenza. Ciò vuol dire che i testi tradotti, rispetto alla lingua di arrivo, tendono ad essere più espliciti e meno ambigui, maggiormente semplificati in particolare nella scelta lessicale, a rispettare la struttura grammaticale esagerando le caratteristiche della lingua di destinazione e ad evitare costruzioni "marcate". Baker inoltre, con il termine *levelling out* (o convergenza) suggerisce che la traduzione possa implicare una sorta di "livellamento" dei testi tradotti tali da evitare espressioni stilisticamente marcate.

2. IL CORPUS ESAMINATO

Il corpus da me compilato è costituito da 50 testi giuridici in formato testo (.txt) divisi in due subcorpora chiamati rispettivamente corpus IT (testi estratti da Normattiva⁶) e corpus UE (testi estratti da EUR-Lex⁷). Il corpus IT è costituito da i seguenti tipi di testo: regolamenti governativi (DPR) e regolamenti ministeriali (DM, DPCM, decreti interministeriali), mentre il corpus UE esclusivamente da regolamenti europei.

L'attenzione sulla scelta dei rispettivi corpora, si è rivolta verso un insieme di testi simili a livello strutturale e quindi più facilmente comparabili. Secondo le indicazioni offerte dalla *Guida pratica comune del Parlamento europeo, del Consiglio e della Commissione per la redazione dei testi legislativi dell'Unione europea* (Guida 2015) i regolamenti comunitari sono atti normativi coercitivi, al punto che «il modo in cui si redige l'atto deve corrispondere alla natura vincolante o non vincolante del testo» (Guida 2015: 12). Inoltre i regolamenti comunitari (fonti derivate⁸ allo stesso modo dei regolamenti dell'esecutivo)

⁶ Normattiva, portale attivo dal 2010 per la raccolta della normativa vigente. Disponibile su: <https://www.normattiva.it> (ultimo accesso: 30 luglio 2020).

⁷ EUR-Lex, sito fondato nel 2001 per la consultazione online di tutti i testi giuridici emanati dall'UE. Disponibile online all'indirizzo: <https://eur-lex.europa.eu/homepage.html?locale=it> (ultimo accesso il 7 agosto 2020).

⁸ I regolamenti sono fonti derivate dai Trattati. Per maggiori informazioni si veda Barbera & Fusaro (2012: 70-78).

sono direttamente applicabili e integralmente obbligatori, per cui le loro norme devono essere redatte in modo tale che ogni cittadino interessato non abbia dubbi circa i diritti e gli obblighi da essi stabiliti. A differenza di quanto accade con le direttive o in qualche caso, con le decisioni, nei regolamenti comunitari non si fa mai riferimento ad autorità nazionali interposte, ma ai cittadini (a parte il caso in cui nell'atto sia previsto l'intervento complementare degli Stati membri).

Per quanto concerne invece la gerarchia delle fonti nazionali, la Costituzione stabilisce che lo Stato ha competenza regolamentare (subordinata alle leggi e per questo al di sotto delle fonti primarie e costituzionali) nelle sole materie statali, per cui i testi inclusi nel corpus derivano esclusivamente da leggi precedenti e hanno lo scopo a volte di specificare il contesto esecutivo della legge, a volte di attuarlo e integrarlo, altre di delegificarlo (Guida 2015: 111).

La ricerca proposta si basa sull'analisi di un corpus da me appositamente costruito, comprensivo di 50 regolamenti (500.000 token) emanati durante l'ultimo decennio (dal 2008 al 2018). Ciascuna delle due sezioni del corpus (corpus IT e UE) è formata da 25 regolamenti l'una e, al fine di renderli ancor più omogenei, ho disposto che ognuna fosse costituita dai medesimi ambiti di riferimento e perciò la selezione dei testi è stata scelta in base a cinque macrocategorie argomentali: occupazione, ambiente, energia, politiche sociali e mercato ognuna comprendente a sua volta cinque regolamenti.

3. REGOLAMENTI NAZIONALI ED EUROPEI A CONFRONTO

3.1 DISPOSIZIONE GENERALE

Il nodo della complessità della lingua giuridica risiede nella struttura morfologica e sintattica. I testi giuridici presentano una costruzione molto elaborata, costituita da frasi lunghe, ricorsivamente incassate e regolate quindi da diversi gradi di subordinazione (Viale 2014). Tendenza tipica della testualità normativa è l'utilizzo in un'unica frase, (generalmente corrispondente al comma), di diverse unità informative, sia pure collegate tra di loro: il rischio è quello di assemblare troppe informazioni in un solo periodo. Gli strumenti attraverso i quali un contenuto nozionale complesso viene espresso in una sola frase sono: il frequente ricorso delle proposizioni implicite, le nominalizzazioni tramite *nomina actionis* (cioè l'uso di sostantivi al posto di verbi per esprimere intere proposizioni) ed infine gli incisi, ovvero frasi che, tra parentesi, virgole o trattini, contribuiscono a dare informazione aggiuntive al testo centrale. Vanno poi nella direzione di un aumento della complessità della lingua giuridica le locuzioni preposizionali e congiunzionali complesse, la presenza di connettivi di scarso uso nella lingua comune (p.es.

nonché, allorché, a fronte di, altresì, circa, ancorché), costruzioni a verbi supporto⁹, alto uso della costruzione passiva.

In ogni caso, anche dopo una semplice visione d'insieme, è facile accorgersi di come, nei regolamenti europei, si respiri un'aria sensibilmente diversa rispetto a quella dei regolamenti nazionali, anche solo da un punto di vista meramente grafico. I regolamenti europei appaiono globalmente più semplici nella struttura dell'articolato (in parte anche a causa del frequente ricorso degli elenchi puntati), e più chiari da consultare sia grazie all'utilizzo di note inserite a piè di pagina, che illustrano i vari rinvii alle norme (Cortelazzo & Pellegrino 2003: 67) e sia per il tipo di organizzazione testuale. Infatti, i documenti ufficiali pubblicati sulla Gazzetta¹⁰ vengono impaginati utilizzando un layout con due colonne di testo rimasto pressoché invariato dal 1874, periodo durante il quale veniva ancora pubblicata la Gazzetta del Regno d'Italia¹¹, che riduce significativamente gli spazi bianchi disponibili e rende più difficile la comprensione. Invece, nei regolamenti europei¹², troviamo una formattazione decisamente più chiara, nonostante siano presenti due diversi layout di impaginazione (sia con colonna singola che a doppia colonna) che, grazie a un font più leggero¹³ e a un utilizzo a più ampio respiro degli spazi bianchi, contribuisce a una migliore percezione visiva globale.

Oltre all'organizzazione grafica, un'ulteriore differenziazione è visibile già in quel che riguarda la struttura dell'articolato e i rispettivi commi. Nei regolamenti nazionali, gli articoli sono costituiti da commi numerati nei quali la condensazione sintattica trova la sua espressione più piena. Essa prevede che ci sia una corrispondenza tra capoverso e periodo, inteso come unità che va da punto fermo a punto fermo. Nelle leggi, come nei regolamenti, il comma presenta, tranne poche eccezioni, una sola frase che racchiude più informazioni,

⁹ Ad esempio: *porre in essere, dare attuazione, dare inizio, trovare applicazione*. Per maggiori informazioni si veda Masini (2009: 81).

¹⁰ La Gazzetta Ufficiale della Repubblica Italiana (GU) rappresenta la fonte ufficiale di conoscenza delle norme in vigore in Italia. Principale strumento di diffusione, informazione e ufficializzazione di testi legislativi, atti pubblici e privati, edita dall'Istituto Poligrafico e Zecca dello Stato e pubblicata in collaborazione con il Ministero della Giustizia, il quale provvede alla direzione e redazione della stessa.

¹¹ Il giornale ufficiale per la pubblicazione di leggi e decreti del Regno d'Italia dal 1860 al 1946, il cui archivio è consultabile gratuitamente online all'indirizzo: <http://augusto.agid.gov.it> (ultimo accesso 1'8/09/2018).

¹² Pubblicati su "La Gazzetta ufficiale dell'Unione europea" (GUUE), fonte del diritto europeo, costituita da un periodico pubblicato quotidianamente in tutte le lingue ufficiali dell'Unione.

¹³ Il font utilizzato nelle pubblicazioni ufficiali dei regolamenti europei è uno speciale font realizzato appositamente per l'Unione europea, chiamato *EU Albertina*. Le uniche eccezioni presenti sono riscontrabili nell'utilizzo del font *Helvetica*, impiegato nei form, e in alcuni fac-simile di lettere di accordo.

ma omogenee tra di loro (Guida 2011: 24). Invece, nei regolamenti europei, il comma risulta essere l'elemento non autonomo di un articolo o di un paragrafo complesso, nel senso che esso non racchiude l'intero contenuto informativo espresso dal periodo. Inevitabilmente, il risultato è quello di mostrare il contenuto dell'articolo attraverso l'utilizzo di più commi (e quindi più periodi), che nel complesso risultano brevi, chiari e forniti di un unico messaggio contenutistico (nei nostri regolamenti, lo stesso contenuto sarebbe espresso da un solo comma).

A dimostrazione di quanto detto, il dato relativo alla percentuale dei punti fermi di fine periodo nei due corpora (3% nel corpus UE, rispetto al 2% riscontrato in quello IT), ci dice che è presente una maggioranza di periodi nel corpus UE rispetto all'altro, mentre il dato relativo alla media di token per proposizione è di 17,5 per il corpus IT, contro i 16,7 del corpus UE.

3.2 ANALISI SINTATTICHE

Abbiamo già detto come questa lingua speciale sia caratterizzata dalla tendenza alla costruzione di strutture periodali nominalizzate (nelle quali la funzione predicativa è svolta dal nome) e che, a seconda dei casi, sostituiscono vere e proprie subordinate finali, causali, temporali, condizionali, ecc., fino al raggiungimento di strutture periodali lunghe e complicate, basate o su costruzioni ipotattiche ripetutamente incassate tra loro o da numerose proposizioni coordinate, le quali, troppo spesso interrotte dal frequente uso di incisi e di rinvii a precedenti norme, inducono a perdere il filo del discorso.

A questo proposito, il sistema computazionale conta, per entrambi i corpora, una percentuale pari a circa il 30% di proposizioni principali e il 70% di proposizioni subordinate. Tuttavia, l'esito delle analisi quantitative condotte sia sul tipo di proposizione presente (principale o subordinata), sia sull'analisi del tipo di congiunzione (coordinante o subordinante) risulta fuorviante e incapace di darci informazioni dettagliate sulla costituzione della struttura interna del periodo (paratattica o ipotattica). Un esame di tipo qualitativo condotto su ogni periodo, sembrerebbe dimostrare una prevalenza di strutture ipotattiche all'interno del corpus IT (in linea con quanto si legge in letteratura), ma d'altra parte, non è nemmeno la maggior quantità di proposizioni coordinate ad essere indice di una maggiore semplicità testuale. Come è stato dimostrato ampiamente da Garavelli (Garavelli 2003: 3-19), non è necessariamente la quantità di subordinate a determinare la complessità del testo, cosa che è da attribuire, invece, alla loro articolazione interna. Garavelli sostiene che non sempre un discorso costruito paratatticamente sia più semplice o più comprensibile di un discorso costruito ipotatticamente. Tutto questo perché «c'è un fatto che giustifica l'impiego di strutture ipotattiche: la loro attitudine a rendere comprensibile l'ordine gerarchico dei pezzi che compongono un

ragionamento» (*ibid.*). Il rischio a volte è quello di preferire l'utilizzo di complicate strutture coordinate tra loro, rese esplicite anche attraverso l'uso di congiunzioni coordinanti ed intervallate da lunghi incisi.

Un altro parametro sintattico interessante riguarda la tipologia interna delle subordinate, differenziate a partire dal modo verbale, indefinito per le proposizioni implicite e finito per le esplicite, come evidenziato nella Tabella 1 e nel relativo Grafico 1.

Tabella 1 – Numero di proposizioni implicite ed esplicite per corpus

	PROPOSIZIONI IMPLICITE	PROPOSIZIONI ESPLICITE
Corpus IT	81%	19%
Corpus UE	71%	29%

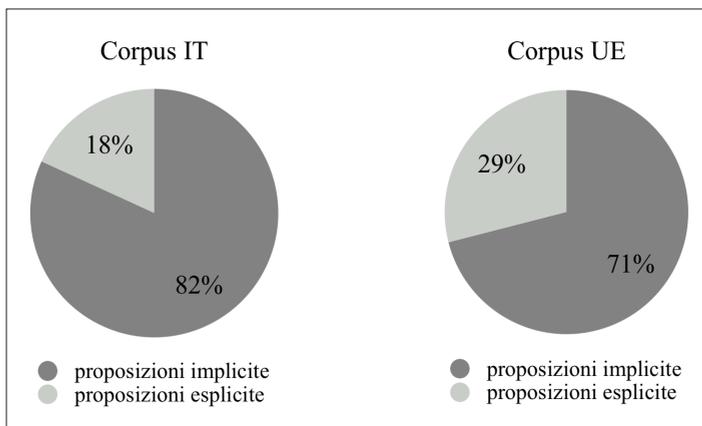


Grafico 1 – Areogramma relativo alla Tabella 1

I dati mostrano, per entrambi i corpora, una maggior percentuale di proposizioni implicite su quelle esplicite, anche se la percentuale si alza notevolmente nel corpus di regolamenti italiani.

Per quanto riguarda l'uso delle forme nominalizzate tramite *nomina actionis* (le quali, in mancanza di verbi, permettono di inglobare in un'unica frase più proposizioni), le analisi condotte non solo evidenziano una maggior percentuale di sostantivi rispetto ai verbi in entrambi i corpora (in linea con la tendenza ad usare più sostantivi nella lingua scritta rispetto a quella parlata; Montemagni

2013: 13), ma anche una maggiore profondità delle catene preposizionali a testa nominale nei regolamenti nazionali: si conta una lunghezza pari a 1,6 nei regolamenti nazionali contro l'1,5 riscontrata all'interno dei regolamenti europei. Tuttavia, il software per l'analisi dei testi non è in grado di fornire valori specifici relativi alla percentuale di forme nominalizzate che sostituiscono intere proposizioni. Per monitorare questo tratto linguistico è possibile svolgere, quindi, solo un'analisi di tipo qualitativo.

Come detto precedentemente, le strutture periodali complesse, tipiche della lingua giuridica, sono caratterizzate non solo da un alto livello di incassamento in combinazione, ma anche dalla concomitanza con più proposizioni coordinate tra loro. Tutto questo è in grado di generare strutture periodali lunghe e particolarmente complicate per le operazioni di lettura e di comprensione del testo. La frase unica e la costruzione dei periodi secondo il meccanismo delle scatole cinesi obbligano le proposizioni e i vari incisi ad essere separati tra loro da un alto numero di virgole. Anche per questo valore, il corpus IT conta un numero maggiore di queste, precisamente 1669 virgole in più rispetto all'altro corpus (ovvero, una media di 67 virgole per regolamento). In ogni caso, è possibile notare, nel complesso, una diversa conformazione nella stesura dei due tipi di regolamenti.

Nella Tabella 2 riporto le analisi della struttura sintattica e dei rapporti di dipendenza.

Tabella 2 – Struttura sintattica e rapporti di dipendenza

ANALISI STRUTTURA SINTATTICA	CORPUS IT	CORPUS UE
Media di token per periodo	24,08	19,02
Media di token per proposizione	17,50	16,70
Media di teste verbali per periodo	1,90	2,10
Media lunghezza dei <i>link</i> massimi della frase	9,90	7,60
Media delle altezze massime degli alberi sintattici	5,60	5,30
Media catene preposizionali per proposizione	2,00	1,00

La “media di token per periodo”, ovvero il numero delle parole che sono contenute tra un capoverso e il punto, e la “media di token per proposizione” dimostrano, già a prima vista, che i periodi dei regolamenti nazionali risultino essere più lunghi degli altri, anche se questo dato non ci dice ancora niente

sull'articolazione delle diverse proposizioni all'interno del periodo. Il valore immediatamente sotto dimostra un maggior numero di teste verbali, ossia di proposizioni per periodo.

Per comprendere al meglio la struttura dell'albero sintattico e le relazioni di dipendenza dei diversi documenti, abbiamo bisogno di valori più specifici. La "media della lunghezza dei *link* massimi" misura, per ogni frase, le lunghezze massime (quindi le lunghezze maggiori) delle relazioni di dipendenza all'interno di un sintagma¹⁴. Anche in questo caso, l'analisi condotta riporta un valore di media maggiore nel corpus italiano, dimostrando una maggiore inaccessibilità ai rapporti esistenti tra le parole. È infatti noto che la contiguità semantica e sintattica degli elementi sintagmatici, permette una loro più immediata recuperabilità. Questa "lunghezza" viene calcolata come la distanza in parole (token) tra la testa sintattica e il suo dipendente e rappresenta dunque un fattore di complessità ampiamente riconosciuto nella letteratura linguistica e linguistico-computazionale. Nella frase che riporto come esempio, la lunghezza massima di dipendenza è quella che si crea tra il verbo principale (testa sintattica) e il soggetto (primo dipendente). Questa lunghezza è di 36 token (inclusa la punteggiatura).

I *candidati* dei concorsi per il reclutamento e per l'accesso ai ruoli del personale delle Forze armate, del personale delle Forze di polizia ad ordinamento militare e civile e del Corpo nazionale dei vigili del fuoco *devono rientrare* entro i valori limite di ciascuno dei parametri fisici indicati nella tabella in allegato "A".

Ancora, attraverso la "media delle altezze massime degli alberi per proposizione", si misura la complessità del testo a partire dai livelli di incassamento gerarchico: infatti, in presenza di più di una proposizione subordinata all'interno dello stesso periodo, è decisivo ricostruire quale tipo di rapporto sussista tra di esse, cioè se siano ricorsivamente incassate l'una all'interno dell'altra. L'indicazione dei livelli di incassamento gerarchico all'interno della struttura sintattica della frase, può essere ricostruita a partire dall'altezza massima dell'albero. Questa altezza misura la massima distanza che intercorre tra una foglia (rappresentata da token del testo senza dipendenti) e la radice dell'albero, espressa come numero di archi (ovvero relazioni di dipendenza) attraversati nel cammino foglia-radice.

Nella frase che segue, la massima distanza che intercorre tra una foglia e la radice dell'albero sintattico è quella tra il token "indicate" e la radice "hanno":

¹⁴ I sintagmi sono costruiti attorno a una testa, attorno alla quale vengono classificati e da cui prendono il nome. Per "testa" si intende la classe di parole che rappresenta il minimo elemento che da solo possa costituire un sintagma (es: nel sintagma "la copertina blu", la "testa" è copertina e nel complesso la costruzione prende il nome di "sintagma nominale". Per un maggiore approfondimento si veda Berruto & Cerruti (2015: 135-138).

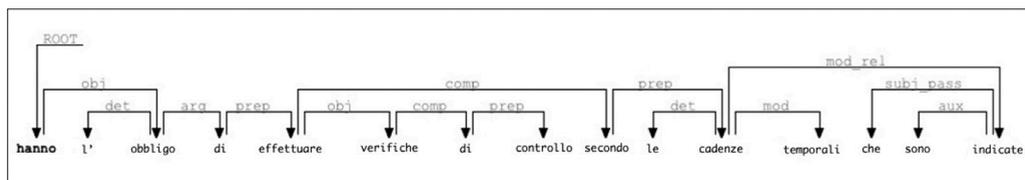


Figura 1 – Rappresentazione grafica dei rapporti di incassamento gerarchici

Gli enti e i privati [...] hanno l'obbligo di effettuare verifiche di controllo secondo le cadenze temporali che sono indicate.

La massima distanza conta 6 relazioni di dipendenza di tipo “obj”, “arg”, “prep”, “comp”, “prep”, “mod rel”¹⁵.

Infine, la “media di catene preposizionali per frase”, calcola il valore medio delle ‘catene’ preposizionali estratte dal corpus per tutte le proposizioni del documento. Una catena preposizionale viene calcolata come il numero di complementi preposizionali dipendenti da uno stesso sostantivo.

4. UNIVERSALI TRADUTTIVI

4.1 ESPLICITAZIONE

Secondo Baker, i traduttori dimostrerebbero «an overall tendency to spell things out rather than leave them implicit in translation» (Baker 1996: 180). L'esplicitazione rappresenta, in sostanza, la modalità linguistica con la quale si rendono esplicite all'interno del testo di destinazione le informazioni che sono invece implicite nel testo di partenza. Tale concetto è stato introdotto per la prima volta da Vinay e Darbelnet all'interno del *Routledge Encyclopedia of Translation Studies* (Baker & Saldanha 2011: 104), i quali lo hanno descritto come pratica obbligatoria, se dettata da motivi strettamente grammaticali¹⁶ o come facoltativa, se dettata da preferenze stilistiche (tra queste: l'uso dei connettivi e l'uso di subordinate relative al posto di lunghe catene nominali).

Ovviamente, per ciò che interessa in questo studio, sarà preso in considerazione solo l'ultimo caso citato, ossia quello legato ai fenomeni dell'esplicitazione opzionale. Nel dettaglio, al fine di sottolineare la diversa natura dei rispettivi

¹⁵ Le sigle identificano i seguenti elementi: “obj” = oggetto diretto, “arg” = argomento, “prep” = preposizione, “compl” = complemento, “mod rel” = proposizione relativa.

¹⁶ Il russo, ad esempio, non possiede articoli determinativi per cui, nella la fase di traduzione dal russo all'italiano, il traduttore è costretto ad inserirli.

regolamenti, mi occuperò del monitoraggio delle forme implicite ed esplicite, osservando il modo verbale della proposizione dipendente e della diversa presenza dei soggetti espressi nei rispettivi corpora. Successivamente mi dedicherò all'analisi delle congiunzioni subordinanti.

4.2 L'UTILIZZO DELLE FORME IMPLICITE COME «COSTRUTTI SINTETICI»¹⁷

Tutti i testi giuridici analizzati condividono l'utilizzo di una sostanziale maggioranza di subordinate implicite rispetto a quelle esplicite, le quali contemplano l'uso del verbo al modo finito preceduto da congiunzione subordinante¹⁸. Il corpus UE conta il 28% di subordinate esplicite e il 72% di implicite sul totale di subordinate, mentre quello italiano conta il 26% di esplicite e il 74% di implicite.

Nel Grafico 2 si tiene conto della struttura e dell'articolazione interna delle subordinate implicite (modo participio, infinito e gerundio), dato fondamentale al fine del monitoraggio sintattico dei corrispettivi corpora.

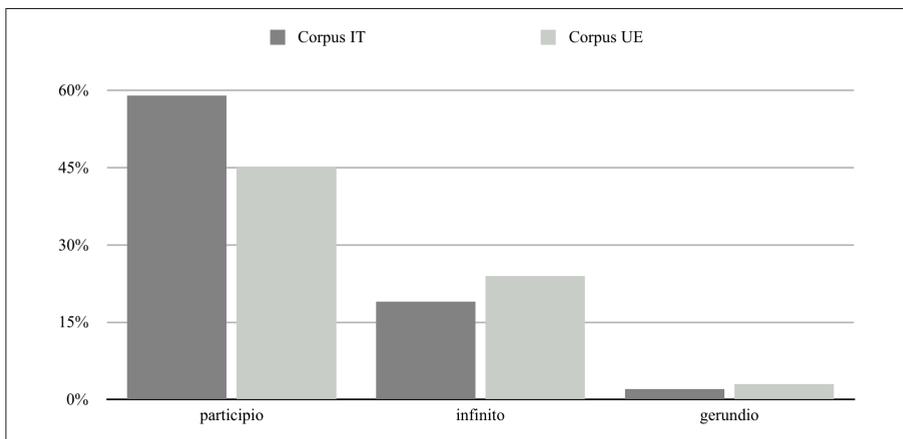


Grafico 2 – Struttura e articolazione interna delle subordinate implicite

Come nota Garavelli, una delle caratteristiche più salienti dei testi giuridico-amministrativi è l'uso massiccio di «frasi ridotte participiali» (Mortara Garavelli 2001: 156), in conformità con una notevole propensione a forme

¹⁷ Concetto teorizzato da Mortara Garavelli (2011: 156).

¹⁸ Per un approfondimento su i tipi di subordinazione si veda Serianni (2010: 545-632).

sintetiche che racchiudono la sovraestensione dell'infinito, i participi verbali ed infine l'enclisi del -si con l'infinito retto da un verbo modale. Quest'ultima costruzione nell'intero corpus compare poco e, salvo errori di computo, si trova più spesso nei regolamenti UE (perlopiù come enclisi del verbo *applicare*) probabilmente a causa della maggior tendenza dei traduttori al conservatorismo linguistico. Ritornando all'uso dei participi, bisogna sottolineare il fatto che essi sono una caratteristica da non sottovalutare nella lingua giuridica (Viale 2014), sia quando svolgono la funzione verbale e usati quindi per una subordinazione di tipo implicito, sia quando vengono usati con funzione nominale. Al contrario, la *Guida per la redazione degli atti amministrativi* invita espressamente ad «evitare i costrutti sintetici come [...] le forme implicite del verbo, come gerundi o participi, quando potrebbero essere usate le forme esplicite» (Guida 2011: 23). Anche il gerundio è usato nei regolamenti, ma in percentuale molto più bassa e prevalentemente con funzione semantica di modo (si conta una maggioranza di verbi principali al gerundio nel corpus europeo 5%, rispetto al 3% riscontrato in quello italiano). Esso «causa una serie di problemi al lettore» (Cortelazzo & Pellegrino 2003: 82). Prima di tutto, non rende visibile il soggetto della frase e questo può provocare casi di incertezza su chi esso sia (nell'esempio sotto, *considerando* non chiarisce il soggetto depositante). In secondo luogo, esso permette la condensazione in un'unica frase del contenuto di più frasi oltre a lasciare incerto (in alcuni casi) il tipo di legame che le lega alla reggente (legame ipotetico, strumentale, modale, temporale ecc.):

Il valore massimo della trasmittanza (U) [...] *considerando* le parti trasparenti e/o opache che le compongono, deve rispettare [...] i limiti riportati nelle tabelle 4.a e 4.b al punto 4 dell'allegato C al decreto legislativo

Gli esempi che seguono hanno lo scopo di mostrare l'esteso utilizzo di subordinate implicite ripetutamente incassate tra loro all'interno dei regolamenti nazionali. Nell'articolo 1, comma 1 del regolamento n. 392/1987 (da me riportato a solo titolo esemplificativo, in quanto non facente parte del corpus) leggiamo:

Le amministrazioni dello Stato anche ad ordinamento autonomo e gli enti pubblici non economici a carattere nazionale e quelli che *svolgono* attività in una o più regioni [...] *sono tenute ad osservare* [...] le modalità di cui agli articoli seguenti, nell'assunzione di personale *appartenente* ai comparti di contrattazione collettiva, *determinati* dal decreto del Presidente della Repubblica 5 marzo 1986, n. 68, *da inquadrare* in profili professionali *ascritti* a livelli retributivo-funzionali che *richiedono* il solo requisito della scuola dell'obbligo e *da adibire* a mansioni per le quali non sia previsto un titolo professionale nella declaratoria dei profili professionali

e ancora, nelle note dell'articolo 3 del regolamento n. 29/2015 sulle disposizioni in materia di liquidazione del TFR:

Il divieto di iniziare o proseguire le azioni cautelari o esecutive di cui al terzo comma *può essere richiesto* dall'imprenditore [...], *depositando* presso il tribunale

competente ai sensi dell'art. 9 la documentazione di cui all'art. 161, primo e secondo comma, lettere a), b), c) e d) e una proposta di accordo *corredata* da una dichiarazione dell'imprenditore, *avente* valore di autocertificazione, *attestante* che sulla proposta *sono in corso* trattative con i creditori che *rappresentano* almeno il sessanta per cento dei crediti e da una dichiarazione del professionista *avente* i requisiti di cui all'art. 67

ed infine all'art. 2, primo comma del regolamento n. 177/2011 sulle norme per la qualificazione delle imprese e dei lavoratori autonomi operanti in ambienti sospetti di inquinamento:

avvenuta effettuazione di attività di informazione e formazione di tutto il personale, ivi *compreso* il datore di lavoro ove *impiegato* per attività lavorative in ambienti sospetti di inquinamento o confinati, specificamente *mirato* alla conoscenza dei fattori di rischio propri di tali attività

Negli esempi riportati troviamo casi di participi passati (p. es. *determinati*, *mirato*, *ecc.*) e participi presenti (p. es. *appartenente*, *avente*, *ecc.*), tutti con funzione verbale. Sicuramente rispetto al modo gerundio, con l'utilizzo del participio e per via della sua natura aggettivale, è ancora possibile il recupero del soggetto. Nel caso preso in oggetto, è ovvio che il riferimento ad *appartenente* sia il più vicino *personale* piuttosto che *modalità*. Ma la questione non è sempre così chiara. Nell'esempio sotto riportato:

e da una dichiarazione del professionista *avente* i requisiti di cui all'art. 67

avente a chi si riferisce? alla dichiarazione o al professionista? Questi sono casi che sarebbero stati facilmente evitati nel caso in cui il participio presente fosse stato reso, come detto precedentemente, con una subordinata esplicita relativa (introdotta da *la quale* o *il quale* abbia).

In generale, può apparire un paradosso il fatto che si possa tacciare di sintesi una lingua ben nota per la costante ripetizione di elementi sempre uguali a sé stessi. Eppure la compressione subordinativa è tipica della lingua giuridica, che non ama esplicitare i verbi reggenti delle sue subordinate, quanto piuttosto usare la loro versione indefinita. Le massime conseguenze di ciò vengono esplicitate dall'ultimo esempio riportato, in cui al sintagma verbale *specificatamente mirato*, è difficile attribuire un riferimento immediato (*personale* o *datore di lavoro*), soprattutto per il fatto che l'elemento semanticamente più vicino al sintagma in oggetto sia *attività di informazione e formazione* (che comunque avrebbe voluto il participio al genere femminile e non maschile). In ogni caso, dal grafico 4.2 emerge il maggior utilizzo di forme al participio nel corpus dei regolamenti nazionali (51%) rispetto a quello dei regolamenti europei (42%).

È bene rivolgere le ultime considerazioni sulle forme implicite all'uso del modo infinito. L'intento è quello di trovare una conferma in quanto osservato da Garavelli quando, tra i più vistosi «costrutti sintetici» (Mortara Garavelli 2001: 156) dei testi giuridici, cita le complete con infinito del tipo:

o l'interessato chiede di *applicare* (= che sia applicata) la legislazione

Per quanto riguarda il corpus UE si può notare, in generale, l'utilizzo di numerose subordinate circostanziali con l'infinito. Osserviamo gli esempi che seguono:

- 1) *Nel cercare* un equilibrio tra l'obiettivo della tutela dei diritti di proprietà intellettuale e le libertà fondamentali [...], si dovrebbe tenere conto della giurisprudenza della Corte di Giustizia dell'Unione europea.
- 2) nell'assunzione di personale [...] *da inquadrare* in profili professionali.
- 3) mettere direttamente in pericolo la sicurezza stradale in misura tale *da determinare* una decisione di fermo del veicolo.

Ovviamente anche per questi casi l'utilizzo dell'infinito al posto di una subordinata esplicita finale (esempio 1), relativa (esempio 2) o consecutiva (esempio 3) risponde in parte all'esigenza della compressione sintattica «che è un tendere alla sintesi, [...] sintesi strutturale, che non vuol dire eliminazione della prolissità su altri piani dell'espressione, e sul livello dell'organizzazione (o forma) del contenuto» (Mortara Garavelli 2001: 162) e in parte a motivi di semplicità espositiva.

Come mostrato dai commi riportati nelle pagine precedenti, nella frequente forma “da + infinito” è presente un vero e proprio calco dalla forma latina della perifrastica con gerundivo¹⁹, un modo di esprimere l'obbligatorietà di un'azione.

Le analisi sulle percentuali di verbi principali all'infinito mostrano una percentuale maggiore di verbi principali all'infinito (esclusi modali e ausiliari) nel corpus UE 24%, contro il 19% del corpus IT. La differenza numerica del valore percentuale è data dalla combinazione di diversi fattori. In primo luogo dal fatto che nei regolamenti europei, l'alto numero di elenchi puntati determini, in qualche caso, un maggior utilizzo dell'infinito in posizione iniziale e, in secondo luogo, per il maggiore uso di congiunzioni, preposizioni o locuzioni preposizionali che richiedano obbligatoriamente l'infinito (es. *al fine di, allo scopo di, prima di, dopo, ecc.*). Anche la maggior presenza di verbi modali aumenta la percentuale considerata, dal momento che questi particolari verbi reggono obbligatoriamente l'infinito (es. *può scegliere*). Infine, ma non meno importante, tale valore potrebbe dipendere dal tipo di subordinata riscontrata. Infatti, il software è in grado di tenere conto sia del numero di subordinate completive (introdotte dal *che*) sia di subordinate circostanziali con l'infinito (finali, causali, consecutive, ecc.) dipendenti da una testa verbale ed introdotte da una preposizione. Il programma ha individuando una percentuale più alta di proposizioni con l'infinito nel corpus UE (1%) rispetto a quella del corpus IT (0,6%). Ne è un esempio il periodo riportato nella rappresentazione²⁰ di Figura 2.

¹⁹ Il gerundivo esprime un obbligo in forma passiva quando ha funzione di predicato.

²⁰ Questa rappresentazione è stata elaborata attraverso il software *Lingua: Linguistic Annotation Pipeline*.

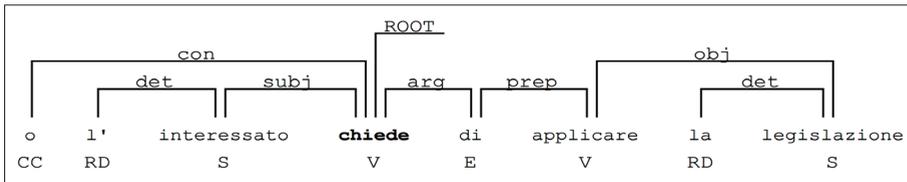


Figura 2 – Rappresentazione grafica dell’annotazione sintattica a dipendenza

Come si può vedere, il programma identifica, a partire dalla radice verbale *chiede* (ROOT), l’arco di dipendenza “arg”, che monitora la completiva implicita introdotta da preposizione *di* (indicata dalla lettera “E”) che regge il verbo all’infinito *applicare* (indicato invece con “V”).

In generale, per motivi di sinteticità, la lingua giuridica tende a costruire le subordinate complete con l’infinito. Invece, è una caratteristica propria dei regolamenti europei, quella di utilizzare preferibilmente circostanziali con l’infinito introdotto da preposizioni, quando possibile, al posto del participio e delle forme nominalizzate che appesantiscono di molto la struttura della frase, distogliendo il lettore dalla comprensione del nucleo centrale della frase, il verbo. Osserviamo un altro esempio tratto da un regolamento europeo:

Il fornitore ha il diritto *di chiedere* all’abbonato *di fornire* (= che fornisca) le informazioni necessarie *per determinare* lo Stato membro di residenza

Nell’esempio sopra riportato, mentre il primo uso dell’infinito (*di chiedere*) è obbligatorio a causa dell’identità tra soggetto della reggente e della dipendente, il secondo uso dell’infinito (*di fornire*), non è sintatticamente richiesto dal contesto (come si vede, in sostituzione poteva essere usato il congiuntivo), a dimostrazione della preferenza accordata per questo modo verbale. L’ultimo uso dell’infinito (*per determinare*), costituisce una proposizione subordinata finale.

4.3 CONGIUNZIONI FRASALI E SUBORDINATE ESPLICITE

Congiunzioni, avverbi e locuzioni servono ad evidenziare i rapporti logici tra le diverse frasi, ed è per questo che vengono complessivamente chiamati connettivi. Un appropriato uso dei connettivi contribuisce sia alla coerenza tematica del testo, assicurando la progressione logica delle informazioni, sia alla coesione testuale, poiché in grado di rendere il testo compatto. Ciò vuol dire inevitabilmente che un testo nel quale i legami logici sono esplicitati è più facile di uno in cui non lo sono. Per quanto riguarda lo studio sui connettivi, le numerose ricerche di linguistica testuale condotte sui corpora hanno dato un contributo fondamentale. Molto importante per il riconoscimento della

specificità della connessione testuale, è stato l’inserimento della categoria delle congiunzioni testuali nell’edizione del 1997 del DISC, il dizionario curato da Sabatini & Coletti (2008), categoria di congiunzioni con una propria specificità.

Di questo vasto insieme, mi occuperò esclusivamente del caso in cui i connettivi esplicitano un tipo di «relazione tra eventi»²¹ rappresentati nel testo (es: relazioni di tempo, fine, comparazione, ecc.), concentrandomi in particolare sull’uso delle congiunzioni frasali subordinanti e sul loro rapporto con le forme di dipendenza esplicita ed implicita. Il mio obiettivo è quello di osservare a livello microscopico la quantità e la qualità delle subordinate presenti nei rispettivi corpora e di poterne trarre delle conseguenze a livello di complessità testuale.

Il software utilizzato per le analisi mi ha permesso di individuare la percentuale di congiunzioni coordinanti e subordinanti sul numero di token totali²². Nella Tabella 3 mi limito a riportare la percentuale di congiunzioni coordinanti e subordinanti, la percentuale di subordinate esplicite su ogni corpus e la relativa media di subordinate.

Tabella 3 – Valori relativi alle congiunzioni coordinanti/subordinanti e alle subordinate esplicite

	CORPUS IT %	CORPUS UE %
Congiunzioni coordinanti (%)	3,26	3,42
Congiunzioni subordinanti (%)	0,50	0,60

La percentuale subordinate esplicite per regolamento, cioè il numero medio di proposizioni subordinate e introdotte da congiunzione subordinante, risulta essere quasi la medesima. Questo dato può essere spiegato solo alla luce di una successiva analisi qualitativa: all’interno del corpus IT, sulle 900 congiunzioni

²¹ Gli altri due tipi di relazioni descritte sono: le relazioni di collocazione testuale e le relazioni logiche di organizzazione testuale. Cortelazzo & Pellegrino (2003: 50).

²² A questo punto si potrebbe pensare di giungere a preliminari informazioni sul rapporto tra paratassi e ipotassi all’interno dei periodi analizzati. In realtà, questo non è possibile stabilirlo, se non attraverso una accurata analisi qualitativa condotta sui due corpora. Infatti, in primo luogo, le subordinate esplicite non contengono ovviamente tutto l’insieme di subordinate presenti nel testo (cioè la percentuale di congiunzioni subordinanti non esaurisce il conteggio di tutte le subordinate), ed in secondo luogo, le congiunzioni coordinanti non vengono calcolate solo quando connettono due proposizioni (principali o subordinate) di pari livello, ma anche in casi di collegamento tra due aggettivi, sostantivi, ecc. (come ad esempio: *bello e brutto, libri e quaderni*). Quindi, i valori trovati non permettono di calcolare eventuali rapporti paratattici/ipotattici tra le frasi.

subordinanti che introducono proposizioni esplicite, sono attestati 442 “qualora” e 266 “se”, pari al 79% di tutte le congiunzioni.

La Tabella 4 mostra la frequenza assoluta, per ogni corpora, delle principali congiunzioni subordinanti individuate nel testo (solo se introducono proposizioni di modo finito²³).

Tabella 4 – Frequenza assoluta delle principali congiunzioni subordinanti

CONGIUNZIONI FRASALI	CORPUS IT	CORPUS UE
CAUSALI		
poiché	1	41
perché	0	5
dal momento che	0	5
FINALI		
affinché/perché	14	50
CONCESSIVE		
sebbene	1	4
benché	2	2
anche se	7	13
nonostante	1	5
TEMPORALI		
allorché	0	23
mentre	7	9
finché	2	12
prima che	0	11
dopo che	2	6
CONDIZIONALI		
quando (temporale e condizionale)	21	109
qualora	442	163
purché	27	51
se	266	645
a condizione che	38	43

²³ Non sono state prese in considerazione quindi le congiunzioni o le locuzioni preposizionali che necessitano di un modo indefinito, come: *pur, al fine di, di, per, da, prima di, pur, pur di, in caso di, ecc.* Inoltre, nel caso di congiunzioni, come *sebbene* o *benché*, che possono introdurre proposizioni implicite o esplicite, il calcolo è stato fatto solo su quest'ultime.

nel caso in cui	64	56
nella misura in cui	0	17
ove	115	58
che	1.301	2.796

È interessante considerare l'utilizzo, in ognuno dei due corpora, delle congiunzioni classificate per tipo. Come si può vedere dalla tabella sopra, le prime dodici congiunzioni frasali, che introducono rispettivamente le subordinate esplicite causali, finali²⁴, concessive e temporali, sono rarissime nel corpus italiano (37 occorrenze, con l'esclusione della congiunzione *quando*, che talvolta può avere valore condizionale), rispetto a quelle presenti nel corpus europeo.

Si nota invece una maggior presenza di subordinate condizionali in entrambi i corpora, nonostante queste siano anche espresse tramite nominalizzazione (p. es. *in caso di risoluzione del rapporto di lavoro*). Probabilmente il più alto uso rispetto ad altre congiunzioni, può essere attribuito al fatto che, al fine di garantire maggior chiarezza nella connessione tra «una fattispecie e il relativo effetto giuridico» (Pattaro, Sartor & Capelli 1995: 35) possa risultare più semplice (e meno ambiguo) il legame condizionale esplicito tra le proposizioni in oggetto invece di quello implicito²⁵. Inoltre, di protasi introdotte dalla congiunzione *se*, si contano maggiori occorrenze nel corpus UE. Questa congiunzione è l'unica che, tra le altre, necessita di un verbo al modo indicativo (al posto del più difficile modo congiuntivo) in linea anche con i suggerimenti offerti da diverse indicazioni per la redazione degli atti amministrativi²⁶. D'altra parte, il corpus IT mostra un maggior utilizzo della congiunzione *qualora*, la quale richiede necessariamente il modo congiuntivo.

La congiunzione *ove* quando ha una sfumatura ipotetica traducibile con *nel caso in cui*, *qualora*, o quando funge da relativo col significato di *dove*, compare maggiormente nel corpus IT, ma bisogna tener conto che, per il suo scarso utilizzo nella lingua comune e per la sua poca chiarezza, sarebbe meglio venisse sostituito da altre congiunzioni che risultino meno ambigue.

²⁴ La finale è una delle subordinate la cui forma implicita è indubbiamente più semplice di quella esplicita. Nei regolamenti da me esaminati, il caso più riscontrato di congiunzione finale che introduce una subordinata esplicita è *affinché* (il *perché* introduce solo una subordinata finale nel corpus UE).

²⁵ Una subordinata condizionale implicita può essere resa o con il gerundio, o con l'infinito preceduto da *a* (p. es. *ti dichiareresti sconfitto, lasciando il campo; a sgridarlo non otterrai i risultati sperati*).

²⁶ Solo per citarne alcuni: Cortelazzo & Pellegrino (2003), Guida (2011), oppure Guida *alla redazione dei testi normativi* (Circolare del 2 maggio 2001, n. 1/1.1.26/10888/9.92).

Un'ultima considerazione deve essere fatta sulla congiunzione *che*; per essere precisi, nonostante di essa se ne faccia un ampio uso in entrambi i corpora, la sua presenza è di gran lunga superiore nei regolamenti europei e ciò deve essere attribuito a due circostanze concomitanti:

- 1) Il maggior ricorso, durante la fase di traduzione, al meccanismo dell'esplicitazione. I participi (con funzione verbale) tendono ad essere sostituiti nelle corrispettive subordinate relative (in genere introdotte dal pronome relativo *che* quando ha funzione di soggetto).
- 2) La molteplicità di funzioni della congiunzione *che*. Questa congiunzione può avere, a seconda dei casi, valori diversi di tipo: completivo (soggettivo, oggettivo, dichiarativo) e consecutivo, oltre che a presentarsi in un gran numero di locuzioni congiuntive (*a condizione che, nel caso che, dato che, prima che, ecc.*).

A conclusione di quanto detto fin qui relativamente all'analisi delle caratteristiche sintattiche e morfosintattiche dei testi, si può affermare che, sebbene la forma di subordinazione implicita rispetto all'esplicita venga maggiormente utilizzata in entrambi i corpora, per esaminare la vera complessità del testo abbiamo bisogno di ulteriori dati relativi all'articolazione interna della proposizione. Ecco perché è stato necessario svolgere un'analisi più dettagliata su come vengono presentate le diverse subordinate al fine di monitorare in maniera completa la conformazione interna dei periodi nei rispettivi corpus. Infatti, un'analisi qualitativa condotta sulle preposizioni è in grado di precisare l'analogia funzionale che queste instaurano con le congiunzioni subordinanti appena analizzate: se quest'ultime introducono una subordinata esplicita, le prime ne introducono una implicita, oltre al fatto di raccordare i costituenti della proposizione. Ciò che si ricava è che, nei regolamenti europei, l'infinito è presente oltre che nei diversi tipi di complete (oggettive, soggettive e dichiarative), anche per le subordinate circostanziali di tipo finale, temporale e consecutivo (introdotte dalle rispettive preposizioni o locuzioni) mentre, nei regolamenti nazionali, si preferisce ricorrere generalmente alla nominalizzazione per quanto riguarda le subordinate finali e ai costrutti col participio passato per le subordinate temporali.

Negli esempi 1, 2 e 3 riporto casi di frasi nominalizzate estratte dal Corpus IT in sostituzione di intere subordinate (finali e ipotetiche), costituite da sostantivi deverbali terminanti in *-mento*, *-zione* e *-io* ed introdotte da preposizioni o locuzioni preposizionali:

- 1) le domande *per l'ottenimento* dell'autorizzazione unica ambientale
- 2) e fanno eseguire la procedura *di valutazione* della conformità di cui all'articolo
- 3) *in caso di violazione* degli obblighi derivanti dal presente codice

Se sostituissimo i sostantivi deverbali in grassetto con i corrispettivi verbi transitivi (*ottenere, valutare, violare*), ci accorgeremmo subito che queste stesse proposizioni risultano immediatamente più semplici per la fase di lettura del testo. Come abbiamo già evidenziato, la nominalizzazione complica la comprensione della frase, affatica la lettura, soprattutto perché, mentre i verbi rendono dinamica la comunicazione, i nomi la rendono maggiormente statica e astratta.

4.4 ESPRESSIONE DEL SOGGETTO E SPERSONALIZZAZIONI

L'italiano è una lingua a soggetto nullo (*o pro-drop*), nella quale la sua ripetizione ha funzione semantica di ripresa del soggetto (ad esempio in frasi molto lunghe e non sequenziali), ma non sintattica. L'inglese invece pretende sempre l'espressione del soggetto.

Durante la fase di traduzione dall'inglese all'italiano, potrebbe attivarsi maggiormente la tendenza a conservare forme soggetto anche là dove in originale non si sarebbero mantenute (Cardinaletti & Garzone 2012: 35-53), per il carattere di conservatorismo²⁷ tipico del processo di traduzione. Questo dimostra la superiore percentuale di soggetti espressi nel corpus UE, sia quando il verbo ha forma attiva, sia quando ce l'ha passiva, come evidenziato dalla Tabella 5.

Tabella 5 – Percentuali e media dei soggetti espressi

	CORPUS IT %	CORPUS UE %
Radici verbali con soggetto espresso	59,00	75,00
Relazione di dipendenza: verbo attivo-soggetto	1,00	2,00
Relazione di dipendenza: verbo passivo-soggetto	0,97	0,92
Media dei pronomi personali di terza persona per testo	7,16	26,36

I valori che si riferiscono al calcolo delle radici verbali con soggetto espresso mostrano una netta minoranza di queste nel corpus IT, fatto attribuibile, in parte, ad un alto uso di subordinate implicite caratterizzate dall'assenza di un soggetto esplicito (nel paragrafo 6 dedicato al fenomeno del conservatorismo linguistico, esaminerò, inoltre, la posizione degli elementi all'interno della frase (soggetto, verbo, complemento)).

²⁷ Questo universale traduttivo verrà esaminato più nello specifico all'interno del paragrafo 6.

I dati successivi esprimono, invece, la percentuale di relazioni di dipendenza che vengono a crearsi rispettivamente tra una testa verbale attiva e il suo soggetto (relazione di dipendenza: verbo attivo-soggetto) e una testa verbale passiva e il suo soggetto (relazione di dipendenza: verbo passivo-soggetto). Quindi, queste ultime misurazioni, non ci forniscono la percentuale di verbi attivi/passivi presenti in ogni corpus, ma il numero delle relazioni di dipendenza instaurate tra soggetto e verbo attivo/passivo. L'ultimo valore riposta la media dei pronomi personali soggetto di terza persona (singolare e plurale) presenti nei rispettivi regolamenti: l'utilizzo quasi quadruplicato dei pronomi all'interno del corpus UE è in linea con la tendenza del traduttore a esplicitare maggiormente il soggetto della frase rispetto a ciò che farebbe nella lingua originale (in particolar modo dall'inglese, francese o in generale, da una lingua a soggetto obbligatorio; D'Achille 2010: 126).

Per quanto riguarda il fenomeno della spersonalizzazione, sia la forma passiva e sia la costruzione impersonale «mirano a occultare, o a mettere in secondo piano, l'agente, cioè colui che compie le azioni indicate negli enunciati» (Cortelazzo & Pellegrino 2003: 100). D'altronde, come ci dice già il nome, la costruzione impersonale con l'uso del *si* (proclitico)²⁸ per i soli verbi che ammettono il soggetto, è la primaria forma di spersonalizzazione, divenendo, in assenza di complemento oggetto, il soggetto della frase.

Alla luce di queste considerazioni, i dati relativi alla presenza del *si* preverbale (tenendo conto del fatto che talvolta il clitico *si* ha valore riflessivo quando rappresenta un complemento diretto o indiretto del verbo; Salvi 2001: 115) dimostrano che nel corpus IT il clitico *si* introduce un numero maggiore di verbi (0,32%), incrementando il senso di oscurità dell'agente che compie l'azione espressa nell'enunciato, anche se nel complesso i valori di entrambi i corpora non distano molto l'uno dall'altro (0,27% per il corpus UE).

5. SEMPLIFICAZIONE

Secondo Baker, durante la fase di traduzione vi è la tendenza inconscia da parte del traduttore a semplificare la lingua di arrivo (Baker 1996: 176) e a renderla così più facile al lettore. Questa tendenza conduce alla selezione di un unico significato del testo, bloccando quindi altre interpretazioni. In questo modo si aumenta dunque non solo la semplificazione, ma anche l'esplicitazione del testo, risolvendo possibili casi di ambiguità. La semplificazione costituisce quindi un'altra fondamentale caratteristica del processo traduttivo.

A sostegno di ciò, riporto alcuni valori che dimostrino una maggior tendenza alla semplificazione e alla standardizzazione lessicale nei regolamenti UE rispetto al corpus IT, anche se le maggiori complessità linguistiche e le

²⁸ La costruzione di un verbo reggente unito al *-si* enclitico, nei regolamenti da me esaminati, è presente solo nel costrutto *trattasi*, forma ormai stereotipata.

differenze tra i due corpora risiedono altrove (cioè nel livello morfosintattico e sintattico, osservato già in parte nel paragrafo precedente).

Prima di considerare i dati dell'indice di *ricchezza lessicale*, ricordo quelli relativi alla *Lunghezza media del periodo* (calcolata in base al numero di token che la costituiscono): 19,02 per il corpus UE, contro il 24,08 per quello IT. Anche se la lunghezza media delle frasi è evidentemente più lunga nel corpus IT, questi valori da soli, non sono in grado di dirci molto sulla effettiva semplicità/complessità dei testi, poiché non tengono per niente in considerazione l'articolazione interna del periodo. Infatti, come ricordato nel capitolo precedente, la complessità sintattica di un testo è data dalla combinazione di più fattori in relazione tra loro (ad esempio i ripetuti incassamenti)²⁹.

5.1 INDICE DI RICCHEZZA LESSICALE: CONFRONTO TRA I CORPORA

Un'interessante panoramica sulle caratteristiche linguistiche del corpus e dei rispettivi regolamenti proviene dai dati quantitativi relativi alla ricchezza lessicale. In realtà, rispetto ad altri tipi di analisi linguistiche, le principali differenze sull'assetto dell'italiano tra i testi normativi nazionali e quelli europei frutto di traduzione risiedono, come avremo modo di vedere, non tanto sugli aspetti lessicali, ma su quelli morfologici e sintattici.

Prima di passare in rassegna le analisi relative alla ricchezza lessicale, occorre evidenziare la distribuzione del lessico dei diversi corpora in base alle parti del discorso individuate dalla grammatica tradizionale.

Il Grafico 3 riporta la ripartizione per categorie grammaticali sul totale delle occorrenze.

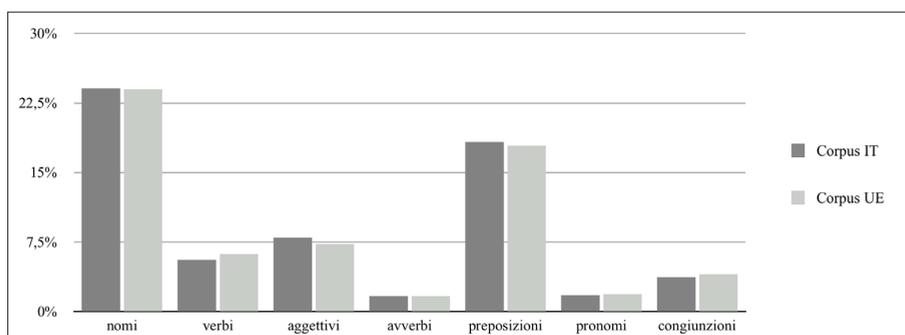


Grafico 3 – Ripartizione per categorie grammaticali sul totale delle occorrenze

²⁹ A questo proposito è giusto fornire un chiarimento. È noto come esista un ulteriore indicatore di leggibilità del testo, l'indice *Gulpease*. Ma questo valore, approssima la complessità linguistica del testo attraverso due parametri (la lunghezza della parola e la lunghezza della frase) anche fin troppo superficiali che hanno una scarsa implicazione nei processi cognitivi di comprensione del testo.

Nel complesso, non emergono differenze significative nella distribuzione delle parti del discorso tra i due corpora esaminati. Infatti, come osservato da Montemagni in un recente studio (Montemagni 2013: 15), testi ad alto contenuto informativo hanno una minore percentuale di verbi rispetto ai sostantivi. In ogni caso, la percentuale più bassa di verbi nel corpus IT è in linea con le considerazioni precedenti relative ad una maggiore condensazione sintattica tramite nominalizzazione che determina, a sua volta, una maggior percentuale di preposizioni.

Vediamo ora i valori relativi alla ricchezza lessicale.

Nella tabella 6, la seconda colonna riporta il rapporto tra il numero di occorrenze (N) e il numero di parole tipo (V)³⁰, valore conosciuto anche come *Type/Token Ratio* (d'ora in avanti TTR). Inoltre, dato che l'indice TTR è sensibile alla lunghezza del testo, esso è stato calcolato nei diversi corpora su porzioni di testo della stessa lunghezza³¹. È vero che il dato conferma una percentuale di ricchezza lessicale leggermente inferiore nel corpus dei regolamenti tradotti (in linea con quanto sostenuto da Baker), ma il risultato non lo si può certo considerare statisticamente significativo. Inoltre, Baker considera una minore densità lessicale (intesa come il rapporto tra parole piene e vuote) un indice di maggiore semplificazione, ridondanza e facilità del testo tradotto (Baker 1996: 183). I dati da me riportati mostrano, in generale, una densità lessicale (pari a 0,5) più alta rispetto a quanto misura normalmente nella lingua parlata (in cui l'indice normalmente si aggira attorno allo 0,3-0,4) ma non evidenziano una sostanziale differenza numerica tra i due corpora. Leggermente più elevata risulta essere, nel corpus IT, la percentuale di *hapax*, ovvero di parole che occorrono una sola volta nell'intero corpus. In linea con la tendenza della semplificazione traduttiva, un maggior tasso di *hapax* significa testo meno ripetitivo e più variato al suo interno (Cortelazzo 2013: 60). In ogni caso, ciò che emerge è che la differenza principale tra i due corpora esaminati risiede altrove (nel livello sintattico e morfosintattico della lingua e non in quello lessicale).

Tabella 6 – Valori relativi alla ricchezza lessicale

	V/N (LEMMI)	DENSITÀ LESSICALE (DL)	% HAPAX
Corpus IT	0,55	0,51	35,905%
Corpus UE	0,45	0,54	35,879%

³⁰ Preciso che con “V” si fa riferimento al numero di lemmi diversi presenti e non alle forme grafiche diverse: *prescrive* e *prescrivono*, quindi, vengono considerate dal software come due manifestazioni dello stesso lemma e quindi calcolate come una sola parola tipo.

³¹ L'indice è stato calcolato su 200 parole per ogni corpus.

Gli ultimi valori che intendo riportare nei Grafici 4 e 5 sono quelli relativi alla percentuale di parole appartenenti al *Vocabolario di Base della lingua italiana* (d'ora in avanti VdB) di Tullio de Mauro e alla distribuzione interna di questo lessico.

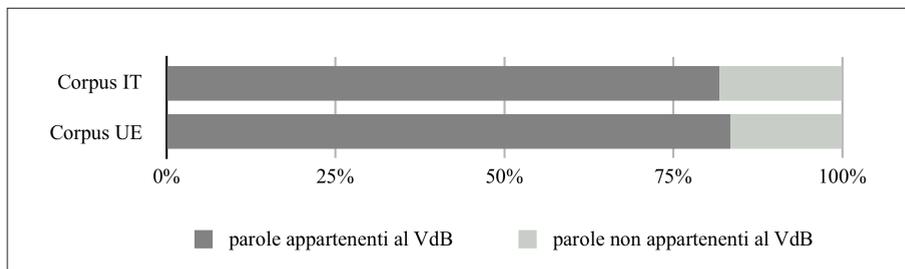


Grafico 4 – Percentuale di parole appartenenti al VdB

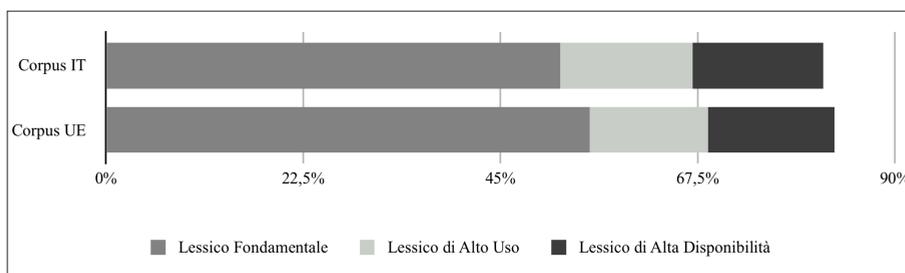


Grafico 5 – Distribuzione interna del lessico VdB

I grafici mostrano un'alta percentuale di parole appartenenti al VdB per entrambi i corpora, probabilmente dovuto in parte alla tendenza alle «ridefinizioni» (Mortara Garavelli 2001: 11), cioè a riutilizzare in modo specialistico, termini della lingua comune. Il corpus UE è caratterizzato da una più alta percentuale di parole appartenenti al VdB (83,4%) rispetto al corpus IT (81,8%) e una percentuale più di parole che appartengono al Lessico Fondamentale: 55,2%, contro il 51,8% riscontrato nel corpus IT.

6. NORMALIZZAZIONE (O CONSERVATORISMO)

L'universale traduttivo individuato dalla Baker con la parola inglese *normalism* (Baker 1996: 175), sempre all'interno del saggio *Corpus-based Translation Studies: The Challenges that Lie Ahead*, esprime la tendenza del traduttore non solo a

ricalcare, ma addirittura esagerare le strutture tipiche della lingua di arrivo. Potrebbe risultare naturale credere che attraverso testi tradotti si diffondano maggiormente parole o costrutti della lingua di partenza. Forse questo fenomeno avviene con certe «traduzioni invisibili» (Grasso 2007: 24), frutto di traduttori improvvisati o comunque privi di formazione. Al contrario, se a tradurre sono traduttori preparati, le traduzioni diventano un luogo di tutela della tradizione della lingua d'arrivo.

In altre parole, questo vuol dire che il traduttore professionista tende a conservare le strutture standard della lingua e a fare, per esempio, minor uso di frasi marcate tipiche del linguaggio parlato³² e ad utilizzare costruzioni “regolari” della lingua d'arrivo, che nel caso dell'italiano ha l'ordine lineare del tipo SVO con la proposizione principale anteposta a quella subordinata. Questo preciso ordine dei costituenti nella frase implica una conseguente strutturazione dell'informazione veicolata: il tema (o dato), cioè l'entità intorno alla quale si predica qualcosa, viene posta prima del rema (o nuovo), ossia la predicazione vera e propria che viene fatta (Berruto & Cerruti 2015: 157).

In genere, le indicazioni per la redazione dei testi normativi raccomandano lo schema soggetto-verbo-complemento (SVO) (Cassese 1992: 307-330). A questo proposito, mi è stato possibile condurre alcune analisi computazionali relative a questo parametro, a dimostrazione del fatto che durante la fase di traduzione si tende maggiormente alla normalizzazione sintattica.

Gli istogrammi riportati al Grafico 6 al mostrano per prima cosa la distribuzione posizionale del soggetto frasale rispetto al verbo, mentre quelli del

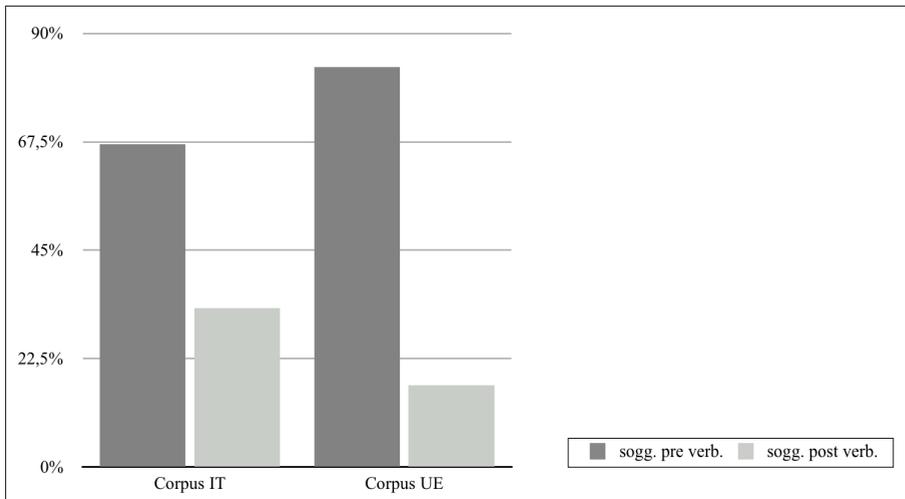


Grafico 6 – Distribuzione posizionale del soggetto frasale rispetto al verbo

³² Si fa riferimento a: dislocazioni, frasi scisse, “che polivalente”, concordanze a senso, false partenze, autocorrezioni, ecc. Per un più ampio approfondimento si veda Roggia (2006: 222).

Grafico 7 dimostrano la posizione delle proposizioni subordinate rispetto alle principali. Per quanto riguarda le informazioni rappresentate dall'istogramma di sinistra, nonostante la maggioranza in entrambi i corpora di proposizioni "non marcate" con soggetto anteposto al verbo (perlomeno per quanto riguarda gli enunciati monoproposizionali), nel corpus IT è presente una più alta concentrazione di soggetti post-verbali rispetto a quelli presenti nel corpus UE, dato in linea con la tendenza al rispetto dell'universale traduttivo del conservatorismo, ma che potrebbe dipendere anche dalla legge dell'interferenza tra la lingua di partenza e quella di arrivo: se le traduzioni nelle Istituzioni europee derivano perlopiù dalla lingua inglese, non è del tutto sbagliato pensare che il maggior utilizzo del soggetto preverbale nel corpus UE sia motivato dal rispetto «dell'ordine sintattico obbligatorio in inglese per le proposizioni affermative» (Cardinaletti & Garzone 2012: 44).

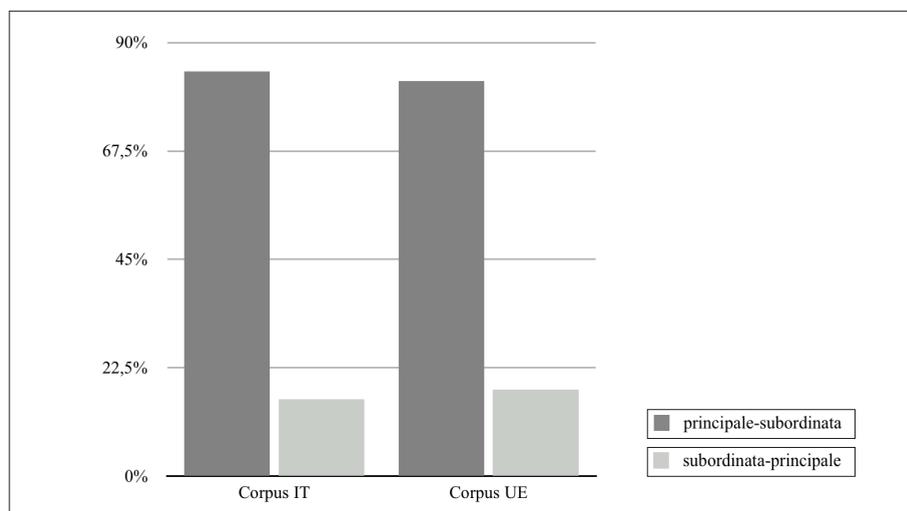


Grafico 7 – Posizione delle proposizioni subordinate rispetto alle principali

In entrambi i corpora ricorre maggiormente l'ordine principale-subordinata al posto di quello subordinata-principale (% superiore nel corpus UE). Dato che, per la maggioranza dei casi, nella distribuzione dell'informazione della frase, il *tema* precede la parte di informazione nuova (*rema*), l'anteposizione della subordinata circostanziale temporale, causale, condizionale è giustificata al fine di «costituire lo *sfondo* (*setting*) della parte rematica e nuova» (Mortara Garavelli 2001: 88).

Per quanto riguarda invece l'analisi morfologica condotta sui verbi, analisi che in varia misura può essere considerata indicatrice di una maggiore/minore

predisposizione al conservatorismo linguistico, riporto nella Tabella 7 i valori percentuali ottenuti per quanto riguarda rispettivamente i modi e i tempi verbali riscontrati.

Tabella 7 – Percentuale relativa a modi e tempi verbali riscontrati in ogni corpora

MODO VERBALE	CORPUS IT %	CORPUS UE %	TEMPO VERBALE	CORPUS IT %	CORPUS UE %
Indicativo	80,54	90,40	Presente	96,85	97,55
Congiuntivo	16,84	8,02	Futuro	2,05	1,33
Condizionale	0,24	0,49	Passato	1,10	1,12

Come è ovvio, i risultati quantitativi mostrano delle particolari tendenze sull'uso dei verbi nella lingua giuridica in rapporto, ad esempio, allo studio condotto sui testi giornalistici (Ondelli e Viale 2010). Questo perché la lingua giuridica fa prevalentemente uso del modo indicativo sia in frasi reggenti che in dipendenti mentre il congiuntivo nelle frasi principali è del tutto assente, il che vuol dire che mancano interrogative dirette dubitative, principali ottative, esortative, permissive. Come descritto in letteratura (Ondelli e Viale 2010: 116), i risultati ottenuti confermano una percentuale di verbi all'indicativo (al tempo presente non-deittico) superiore a tutti gli altri modi verbali e con un tasso maggiore nel corpus UE (90,4%) rispetto a quello IT (80,54%), probabilmente per via del maggior numero di proposizioni subordinate all'indicativo.

In uno studio sulle caratteristiche dell'italiano contemporaneo e in particolare sulle analisi delle variazioni linguistiche rispetto all'asse diamesico, D'Achille (2010: 196) dimostra come si verifichi oggi una maggiore tendenza nella produzione scritta all'uso di più modi e tempi verbali rispetto a quanto accade nell'elaborazione orale. Di conseguenza possiamo dedurre che nei regolamenti europei, dove la propensione al conservatorismo linguistico è maggiore, si sarebbe potuto riscontrare un utilizzo più frequente del congiuntivo in frasi dipendenti (rispetto al modo indicativo) per un atteggiamento di ipercorrettismo³³, anche quando esso non è strettamente richiesto dal contesto. Questo fatto però non si verifica: nel corpus UE il congiuntivo è utilizzato solo in una piccola percentuale, mentre nell'altro corpus il congiuntivo è ampiamente usato, in parte a causa dell'uso di congiunzioni che richiedono obbligatoriamente questo modo verbale (cfr. *qualora*, par. 4.3).

Discorso a parte va riservato invece sul condizionale. Questo modo verbale, per via della imperatività delle norme giuridiche, nel corpus IT non compare

³³ Per un più ampio approfondimento si veda Cardinaletti e Garzone (2012: 48-51).

mai né in frasi indipendenti né nell'apodosi di periodi ipotetici e salvo errori di conteggio, esso compare perlopiù come verbo modale (*potere, dovere*) e come ausiliare. All'interno del corpus UE la situazione è alquanto diversa. Infatti, in questi regolamenti, il modo condizionale è usato circa 755 volte e soprattutto nel sintagma formato da "verbo modale al condizionale + infinito" (p. es. *potrebbe ottenere*) o come ausiliare in una forma condizionale passata. Nello specifico, questo modo verbale compare fuori dall'articolato, all'interno della premessa iniziale, probabilmente per via del carattere motivazionale di quest'ultima. Ricordo che, in inglese, per le norme si usa il modale *shall* seguito dal verbo all'infinito (costrutto che in genere viene tradotto con il presente indicativo), mentre viene utilizzato *should* o *must* quando la frase acquisisce una sfumatura attenuativa, delineando azioni non prescrittive (in italiano viene tradotto con un condizionale) (Biel 2014: 341). Dato che è soprattutto in quest'ultima forma che compare nel preambolo iniziale dei regolamenti scritti in inglese, è ovvio che l'alto tasso di verbi al condizionale dipenda dalla legge dell'interferenza linguistica.

Oltre a ciò, il condizionale compare frequentemente in proposizioni principali come si vede dall'esempio che segue:

- 1) La creazione di un programma unico per l'istruzione, la formazione, la gioventù e lo sport *dovrebbe* condurre a considerevoli semplificazioni

sia in numerose proposizioni relative:

- 2) L'ispettore osserva le regolamentazioni [...] ed evita, per quanto possibile, di interferire con le operazioni di pesca e con lo stivaggio del pescato nonché di compiere azioni che *potrebbero* pregiudicare la qualità delle catture a bordo

ed eccezionalmente, anche nell'apodosi di un periodo ipotetico:

- 3) Se detti fornitori si avvalgono di tale facoltà, *dovrebbero* ottemperare agli stessi obblighi imposti dal presente regolamento
- 4) Qualora emergano nuovi elementi, l'esenzione *dovrebbe* essere riesaminata di conseguenza

Un caso di morfologia verbale "in crescita" per il quale è stata anche postulata l'ipotesi di una possibile influenza della lingua inglese sull'italiano è la perifrasi progressiva "stare + gerundio", stimata in espansione nell'italiano contemporaneo³⁴. Ciò che ho potuto constatare è che, all'interno dei regolamenti europei e nazionali, questa forma è pressoché assente. Salvo errori di calcolo, si conta una sola occorrenza della forma perifrastica progressiva nel corpus IT (p. es. *stanno generando*), e quattro occorrenze in quello UE (p. es. *sta preparando*,

³⁴ Per un più ampio approfondimento si veda Ondelli & Viale (2010: 36-37).

sta emergendo, stanno diminuendo, stia agendo), probabilmente per via di una leggera influenza della lingua inglese.

7. CONVERGENZA

L'esame dell'ultimo universale traduttivo è il più complesso da dimostrare. Questo perché il fenomeno della convergenza linguistica (o *levelling out*) (Baker 1996:177) consiste di fatto in un processo di livellamento dei fenomeni linguistici dei testi tradotti tali da evitare espressioni stilisticamente marcate. Questo fenomeno però, sostiene la Baker, è determinato da una più alta similitudine in termini di caratteristiche linguistiche, quali ad esempio la densità lessicale, il rapporto *type/token* e la lunghezza media della del periodo (in termini di token), rispetto ai singoli testi scritti in originale, ma ciò non dipende in nessun modo dalle lingue prese in considerazione. Insomma, la convergenza linguistica deve essere misurata all'interno di un corpus di testi tradotti e consiste nella tendenza di questi testi a risultare più omogenei tra loro e a gravitare maggiormente verso il centro di un *continuum*³⁵.

Attraverso il grafico chiamato diagramma a scatola e baffi (o *box-plot*³⁶) mi è stato possibile realizzare una rappresentazione grafica per ogni tipo di tratto linguistico esaminato (densità lessicale, rapporto *type/token* e numero medio di token per periodo).

Come dimostrato dai Grafici 8, 9 e 10, per tutte e tre le misure linguistiche esaminate, il fenomeno della convergenza è maggiormente visibile all'interno

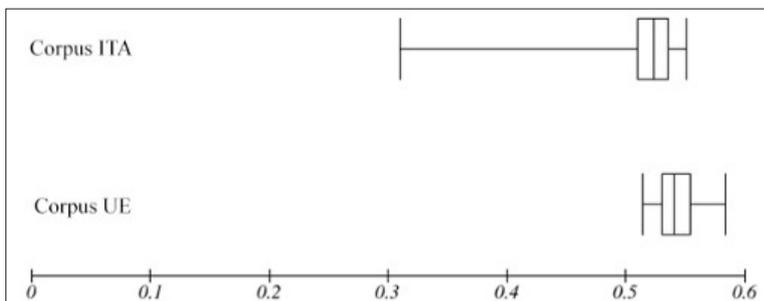


Grafico 8 – Densità lessicale

³⁵ Insieme di varietà linguistiche non separate da confini netti, ma aventi punti di contatto e di sovrapposizione tali da determinare il passaggio graduale dell'una nell'altra.

³⁶ All'interno di questa rappresentazione, la scatola ci dà le informazioni sulla variabilità dei valori, la mediana al suo interno invece sulla tendenza centrale. Dalla scatola si diramano due segmenti, i baffi, dai cui estremi vengono delimitati il minimo e dal massimo dei valori riscontrati nel testo. Ne consegue che, se il diagramma è più concentrato, i valori rappresentati sono più vicini e quindi le caratteristiche linguistiche più simili tra loro.

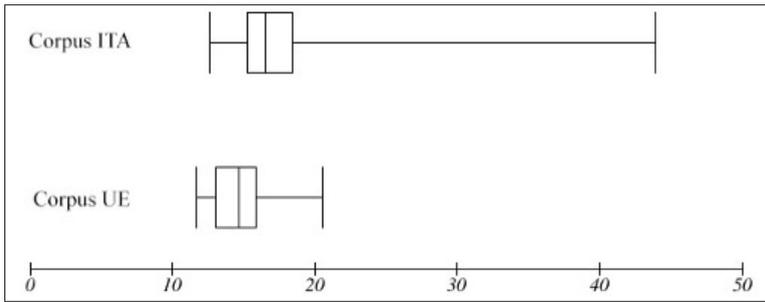


Grafico 9 – TTR

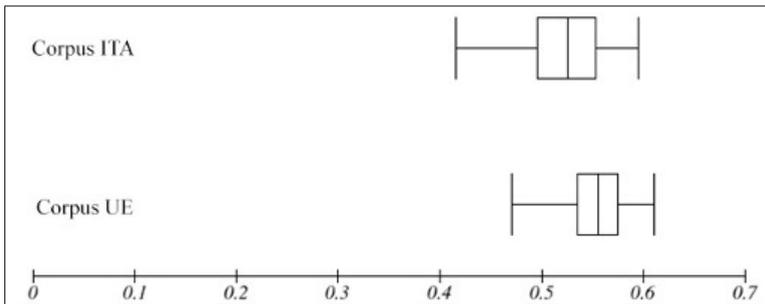


Grafico 10 – Numero medio di token per periodo

dei testi tradotti, in linea con quanto sostenuto in letteratura. Inoltre, in tutti e tre i grafici, la “scatola” nei diagrammi di riferimento del corpus UE, risulta essere ogni volta più centrata rispetto al termine dei “baffi” in rapporto a quella dell’altro corpus, implicando, ancora una volta, la direzione di una via di mezzo tra due valori estremi.

CONCLUSIONI

L’intero insieme delle analisi svolte sui testi ci permette di stabilire non solo quale sia l’assetto “dell’italiano delle traduzioni” presente nei regolamenti europei in rapporto a quello riscontrabile nei regolamenti nazionali, ma anche di rispondere alla domanda posta all’inizio, ovvero se quella sorta di “lingua franca” utilizzata all’interno delle Istituzioni europee, l’euroletto, veicolo funzionale per la comunicazione sovranazionale sia più semplice di quella utilizzata nei regolamenti nazionali. Ora che è possibile avere un quadro completo delle ricerche svolte, sappiamo che la risposta è: sì; tanto che è possibile osservare come, in generale, i testi europei tradotti risultino più semplici da leggere anche per una persona non esperta in materia. Ovviamente questo è

dettato, in parte, dalla spinta semplificativa derivante del lavoro di traduzione e, in parte, da un accurato esame nel tentativo di maggior chiarezza espositiva da parte giuristi-linguisti delle Istituzioni dell'Unione europea (Guida 2015: 10-11).

È noto come la vera complessità della lingua giuridica risieda maggiormente nel livello morfosintattico e sintattico piuttosto che in quello lessicale (come emerge dalle analisi relative alla ricchezza lessicale). In effetti, il nodo della difficoltà è riassumibile nell'utilizzo di una serie di elementi linguistici, più diffusi nel corpus IT, che hanno lo scopo di sintetizzare in meno parole il contenuto informativo altrimenti reso da più proposizioni, rendendo così il testo finale più denso di informazioni e di più difficile comprensione (svolgono questo scopo i participi passati e presenti, i gerundi e le nominalizzazioni).

D'altra parte, occorre specificare che mentre per alcune analisi non è stato considerato l'influsso della lingua di partenza in linea con la teoria degli universali, altri risultati risentano, al contrario, dell'interferenza con la lingua di partenza e che abbiamo visto essere perlopiù l'inglese. Infatti, per quanto concerne la struttura sintattica, si nota una maggiore tendenza sia all'utilizzo di costruzioni "non marcate" del tipo SVO con soggetti pre-verbali, sia di pronomi personali soggetto di terza persona, in quei regolamenti che fanno parte del corpus UE, ma tale comportamento, più che per un fattore di conservatorismo, rappresenta un fenomeno di interferenza linguistica dall'inglese dato che, in questa lingua, il soggetto è sia obbligatoriamente espresso, che anteposto al verbo nelle frasi affermative.

Per concludere, l'obiettivo di monitorare nel dettaglio la lingua giuridica e di affrontare congiuntamente il tema della complessità di quest'ultima attraverso l'indagine linguistica dei fenomeni che la rendono tale, sembra essere stato raggiunto. Sebbene consci di taluni limiti posti nell'analisi automatica dei testi tramite software specifici (un esempio su tutti: la funzione svolta dalle nominalizzazioni tramite *nomina actionis* non può essere svolto se non dopo un'attenta analisi qualitativa), i traguardi conseguiti permettono di evidenziare le potenzialità nell'adozione di strumenti di Trattamento Automatico del Linguaggio ancora non pienamente esplorate per quest'ambito di studi e dovrebbero quindi rappresentare il punto di partenza per una serie di sviluppi futuri incentrati a monitorare tutti i tratti linguistici rilevanti e a verificare gli influssi linguistici dei testi normativi italiani sui traduttori europei e viceversa³⁷.

La Tabella 8 sintetizza i valori delle diverse analisi che sono stati ottenuti per ognuno dei due corpora esaminati, al fine di mostrare più chiaramente il quadro complessivo dei tratti linguistici individuati:

³⁷ Sulla linea tracciata da Cortelazzo (2013: 57-65).

Tabella 8 – Schema riassuntivo dei tratti linguistici analizzati all’interno dei due corpora

TRATTI LINGUISTICI ANALIZZATI	CORPUS IT	CORPUS UE
Media di token per periodo	24,08	19,02
Media di token per proposizione	17,50	16,70
Media teste verbali per periodo	1,90	2,10
Media della lunghezza dei <i>link</i> massimi della frase	9,90	7,60
Media delle altezze massime degli alberi sintattici	5,60	5,30
Media catene preposizionali per proposizione	2,00	1,00
Proposizioni esplicite	19%	29%
Proposizioni implicite	81%	71%
Congiunzioni coordinanti	3,26%	3,42%
Congiunzioni subordinanti	0,5%	0,6%
Modo participio	59%	45%
Modo infinito	19%	24%
Modo gerundio	2%	3%
Modo indicativo	80,54%	90,4%
Modo congiuntivo	16,84%	8,02%
Modo condizionale	0,24%	0,49%
Tempo presente	96,85%	97,55%
Tempo futuro	2,05%	1,33%
Tempo passato	1,1%	1,12%
Soggetto espresso	59%	75%
Media dei pronomi personali di terza persona per testo	7,16	26,36
Verbi introdotti dal clitico <i>si</i>	0,32%	0,27%
Ricchezza lessicale (TTR)	0,55	0,45
Densità lessicale	0,51	0,54
Hapax	35,905%	35,879%
Parole appartenenti al VdB	81,8%	83,4%
Soggetto pre-verbale	67%	83%
Soggetto post-verbale	33%%	17%
Rapporto principale-subordinata	84%	82%
Rapporto subordinata-principale	16%	18%

- Attardi G. & Dell'Orletta F. (2009) "Reverse Revision and Linear Tree Combination for Dependency Parsing" in *Proceedings of NAACL HLT 2009 – North American Chapter of the Association for Computational Linguistics – Human Language Technologies*, Boulder, Colorado, Association for Computational Linguistics, pp. 261-264.
- Attardi G., Dell'Orletta F., Simi M. & Turian J. (2009) "Accurate Dependency Parsing with a Stacked Multilayer Perceptron", in *Proceedings of EVALITA 2009 – Evaluation of NLP and Speech Tools for Italian*, Reggio Emilia.
- Baker M. (1996) "Corpus-based Translation Studies: The Challenges that Lie Ahead", in *Terminology, LSP and Translation. Studies in Language Engineering in Honour of Juan C. Sager*. Ed. by H. Somers, Amsterdam, J. Benjamin, pp. 175-186.
- Baker M. & Saldanha G. (2011) (ed.) *Routledge Encyclopedia of Translation Studies*, London and New York, Routledge.
- Barbera A. & Fusaro C. (2012) *Corso di diritto pubblico*, Bologna, Il Mulino.
- Berruto G. & Cerruti M. (2015) *La linguistica. Un corso introduttivo*, Torino, UTET Università.
- Biel Ł. (2014) "The Textual Fit of Translated EU Law: a Corpus-based Study of Deontic Modality", in *The Translator*, Vol. 20, No. 3, Taylor & Francis, pp. 332-355.
- Bolasco S. (2005) "Statistica testuale e text mining: alcuni paradigmi applicativi", in *Quaderni di statistica*, Vol. 7, Liguori Editore, pp. 1-37.
- Cardinaletti A. & Garzone G. (2012) *L'italiano delle traduzioni*, Milano, FrancoAngeli.
- Cassese S. (1992) "Introduzione allo studio della normazione", in *Rivista trimestrale di diritto pubblico*, No. 2, Milano, Giuffrè Editore, pp. 307-330.
- Cortelazzo M. A., Pellegrino F. & Viale M. (1999) *Semplificazione del linguaggio amministrativo. Esempi di scrittura per le comunicazioni ai cittadini*, Padova, Comune di Padova.
- Cortelazzo M. A. & Pellegrino F. (2003) *Guida alla scrittura istituzionale*, Bari, Laterza.
- Cortelazzo M. A. (2013) "Leggi italiane e direttive europee a confronto", in *Realizzazioni testuali ibride in contesto europeo. Lingue dell'UE e lingue nazionali a confronto*. A cura di S. Ondelli, Trieste, EUT Edizioni Università di Trieste, pp. 57-66.
- D'Achille P. (2010) *L'italiano contemporaneo*, Bologna, Il Mulino.
- Dell'Anna M. V. (2008) "Il lessico giuridico italiano. Proposta di descrizione", in *Lingua Nostra*, Vol. 69, No. 3-4, pp. 98-110.
- Dell'Orletta F. (2009) "Ensemble System for Part-of-Speech Tagging", in *Proceedings of EVALITA 2009 – Evaluation of NLP and Speech Tools for Italian*, Reggio Emilia.
- Dell'Orletta F., Montemagni S. & Venturi G. (2011) "READ-IT: Assessing Readability of Italian Texts with a View to Text Simplification", in *SLPAT '11 – SLPAT '11 Proceedings of the Second Workshop on Speech and Language Processing for Assistive Technologies*, Edinburgh,

- Association for Computational Linguistics, pp. 73-83.
- Grasso D. E. (2007) *Innovazioni sintattiche in italiano (alla luce della nozione di calco)*, Thèse de doctorat, Université de Genève, no. L. 629. Disponibile online all'indirizzo <https://archive-ouverte.unige.ch/unige:475/THESIS> (ultimo accesso il 10/11/2018).
- Gualdo R. (2011), "Il linguaggio del diritto", in R. Gualdo & S. Telve, *Linguaggi specialistici dell'italiano*, Roma, Carocci, pp. 411-477.
- Guida (2011) *Guida alla redazione degli atti amministrativi. Regole e suggerimenti*. A cura del gruppo di lavoro promosso da Istituto di teoria e tecniche dell'informazione giuridica e Accademia della Crusca, Firenze, Ittig-Cnr. Disponibile online all'indirizzo <http://www.ittig.cnr.it/Ricerca/Testi/GuidaAttiAmministrativi.pdf> (ultimo accesso il 10/11/2018).
- Guida (2015) *Guida pratica comune del Parlamento europeo, del Consiglio e della Commissione per la redazione dei testi legislativi dell'Unione europea*, Lussemburgo, Ufficio delle pubblicazioni dell'Unione europea. Disponibile online all'indirizzo <https://eur-lex.europa.eu/content/techleg/IT-guida-pratica-alla-redazione-di-testi-legislativi.pdf> (ultimo accesso il 10/11/2018).
- Masini F. (2009) "Combinazioni di parole e parole sintagmatiche", in *Spazi linguistici. Studi in onore di Raffaele Simone*. A cura di E. Lombardi Vallauri & L. Mereu, Roma, Bulzoni, pp. 191-209.
- Mercatali P. (1995) "Legimatica e redazione delle leggi" in *Legimatica. Informatica per legiferare*. A cura di C. Biagioli, P. Mercatali & G. Sartor, Napoli, ESI, pp. 37-74.
- Montemagni S. (2013) "Tecnologie linguistiche-computazionali e monitoraggio della lingua italiana", in *Studi Italiani di Linguistica Teorica e Applicata (SILTA)*, Anno XLII, No. 1, Pisa, Pacini Editore, pp. 145-172.
- Mortara Garavelli B. (2001) *Le parole e la giustizia. Divagazioni grammaticali e retoriche su testi giuridici italiani*, Torino, Einaudi.
- Mortara Garavelli B. (2003) "Strutture testuali e stereotipi nel linguaggio forense", in *La lingua, la legge, la professione forense. Atti del convegno Accademia della Crusca*. A cura di A. Mariani Marini, Milano, Giuffrè Editore, pp. 3-19.
- Ondelli S. (2020), *L'italiano delle traduzioni*, Roma, Carocci Editore.
- Ondelli S. (2007) *La lingua del diritto: proposta di classificazione di una varietà dell'italiano*, Roma, Aracne.
- Ondelli S. & Viale M. (2010) "L'assetto dell'italiano delle traduzioni in un corpus giornalistico. Aspetti qualitativi e quantitativi", in *Rivista internazionale di tecnica della traduzione*, Vol. 12, pp. 1-62.
- Pattaro E., Sartor G. & Capelli A. (1995) "Norma": manuale per la redazione dei testi normativi, Bologna, CIRSIFID. Disponibile online all'indirizzo: <http://www.maldura.unipd.it/buro/manuali/norma.pdf> (ultimo accesso il 10/11/2018).
- Roggia C. E. (2006) "Costruzioni marcate tra scritto e parlato: la frase scissa" in *Parole frasi testi tra scritto e parlato*. A cura di A. Ferrari, Lugano, Cenobio, LV/3, pp. 222-230.
- Sabatini F. & Coletti V. (2008) *Dizionario della Lingua Italiana*, Firenze, Giunti. Disponibile online all'indirizzo: http://dizionari.corriere.it/dizionario_italiano/ (ultimo accesso il 10/11/2018).
- Salvi G. (2011) "La frase semplice", in *Grande grammatica italiana di consultazione. Vol. I. La Frase. I sintagmi nominale e preposizionale*. A cura di L. Renzi, G. Salvi & A. Cardinaletti, Bologna, Il Mulino.
- Venturi G. (2011) *Lingua e diritto: una prospettiva linguistico-computazionale*, Tesi di dottorato, Università di Torino. Disponibile online all'indirizzo: http://www.italianlp.it/wp-content/uploads/2013/10/Dottorato_linguaEdiritto_rev6_nuovaIntro.pdf (ultimo accesso il 10/11/2018).
- Viale M. (2014) "Quale italiano per le leggi?", in *Nel labirinto del burocrata, Speciale Treccani Lingua Italiana*. A cura di S. Novelli. Disponibile online all'indirizzo: www.treccani.it/lingua_italiana/speciali/burocrata/Viale.html (ultimo accesso il 13/11/2018).