

Costruzione e analisi stilometrica di un *corpus* parallelo di traduzioni giapponesi del *Trionfo della morte*

GABRIELE CAMILLERI

1. Introduzione

Questo studio mira ad analizzare l'evoluzione stilistica nelle traduzioni giapponesi del *Trionfo della morte* di Gabriele D'Annunzio, pubblicate tra il 1913 e il 2010. Le traduzioni sono state codificate secondo lo standard TEI-XML suddivise in cinque sotto-*corpora*: narrazione, dialoghi del protagonista, dialoghi dell'eroina, monologhi interiori e corrispondenza epistolare. Utilizzando in primo luogo un set di 20 metriche testuali, la metodologia adottata ha incluso l'analisi delle componenti principali (PCA) per la riduzione della dimensionalità, seguita da un *clustering* gerarchico e dall'analisi della varianza (ANOVA), per identificare le caratteristiche distintive di ciascun gruppo. Tramite l'analisi di Zeta, è stato inoltre individuato e analizzato un secondo set di 20 trigrammi di *token* per individuare i pattern lessicali distintivi in ogni sotto-*corpus*. I risultati ottenuti evidenziano gli elementi che distinguono le traduzioni antecedenti e successive al secondo dopoguerra, fornendo così una base per ulteriori studi comparativi.

2. Precedenti Ricerche

2.1. Il “traduzione” , stilometria e linguistica dei *corpora*

Il “traduzione” (*translationese*), termine che si riferisce all'insieme di caratteristiche linguistiche osservate con maggiore frequenza nei testi tradotti rispetto ai testi prodotti nella stessa lingua di arrivo, è da lungo tempo oggetto di studio della linguistica dei *corpora*. Sebbene gli studi iniziali si siano principalmente concentrati sull'identificazione di potenziali “universali della traduzione” (Baker, 1993; Mauranen *et al.*, 2004; Mauranen, 2008), come la semplificazione o l'esplicitazione, le ricerche successive hanno esplorato anche la possibile influenza che la lingua di partenza o la lingua di arrivo possono esercitare sulle scelte traduttive (Chesterman, 2004). Oltre a considerare le differenze tra linguaggio di traduzione e linguaggio non-traduttivo, l'utilizzo di tecniche quantitative e stilometriche sono largamente impiegate anche per confrontare lo stile dei singoli traduttori, sia prendendo in esa-

me traduzioni contemporanee (Wang *et al.*, 2012; Ji *et al.*, 2012; Lynch *et al.*, 2018) che prodotte a distanza di tempo (Burrows, 2002; Rybicki, 2006).

La ricerca sul “traduzionese” in giapponese ha suggerito come sue caratteristiche distintive, in linea di massima, un uso più frequente di pronomi personali e dimostrativi, di prestiti linguistici, soggetti espliciti, nomi astratti o inanimati, aventi funzione di soggetto con verbi transitivi, e una maggiore lunghezza dei paragrafi¹. Tuttavia, la letteratura in merito è principalmente di natura qualitativa, in parte a causa della mancanza di *corpora* dedicati. Si possono citare due studi che utilizzano un approccio basato sui *corpora* per analizzare il linguaggio di traduzione giapponese: Fukuchi Meldrum (2009) mette a confronto un *corpus* di diciotto traduzioni e diciotto testi in lingua nativa, pubblicati tra il 1980 e il 2006, utilizzando metriche quali la frequenza di pronomi di terza persona, di prestiti linguistici in *katakana*, di caratteristiche del linguaggio femminile, la lunghezza dei paragrafi e l’uso di referenti inanimati come agenti di verbi intransitivi. Konishi (2016) analizza lo stile di sei traduzioni della metà dell’era Meiji (1888-1892) e di due romanzi giapponesi di periodo più tardo (1906-1916), comparando la lunghezza media della frase, il MVR (*Modifier Verb Ratio*) e la frequenza di parti del discorso.

2.2. *Trionfo della morte* in Giappone

La tabella 1 riassume in ordine cronologico le principali traduzioni del *Trionfo della morte* in giapponese. Tali dati sono stati ottenuti, sia facendo riferimento a quanto già sintetizzato da Yonekawa (1991, pp. 220-221) e Hirayama (2011, pp. 126-128), sia mediante un successivo confronto con il *database* della Biblioteca della Dieta nazionale del Giappone².

Anno di Pubblicazione	Traduttore	Casa Editrice	Note
1901	Ueda Bin	Bunyūkan	Riadattamenti del quinto e sesto libro.
1913	Ikuta Chōkō	Shinchōsha	
1913	Ishikawa Gian	Dainippontoshō	
1914	Ikuta Chōkō	Seinengakujutsusha	Traduzione ridotta.
1914	Hinotsuki Akinori	Akagishōsho	Traduzione ridotta.
1921	Mikami Otokichi	Tōkasha	Pubblicati solo primo, secondo, terzo e quarto libro.
1922	Autore sconosciuto	Bunkenshoin	Traduzione ridotta.

¹ Si faccia riferimento a Fukuchi Meldrum (2009) e Wakabayashi (2009).

² <https://ndlsearch.ndl.go.jp/>

1923	Rikishi Heizō	Kamigatayashuppanbu	Traduzione ridotta.
1924	Saitō Ryūtarō	Shunyōdōshoten	
1927	Yaguchi Tatsu	Chōbunkaku	
1928	Ikuta Chōkō	Shinchōsha	Nuova traduzione.
1939	Harada Kenji	Kaizōsha	Pubblicati solo primo e secondo libro.
1951	Shioda Takeshi	Kadokawashoten	
1958	Iwasaki Junkō	Kawadeshobōshinsha	
1961-63	Nogami Soichi	Iwanamishoten	
2010	Waki Isao	Shōraisha	

Tabella 1.

Cronologia delle traduzioni giapponesi del *Trionfo della morte*

Ueda Bin è il primo autore ad introdurre il *Trionfo della morte* attraverso le traduzioni di due frammenti del quinto e sesto libro, *Enjo monogatari* (Storia di una donna seducete, pubblicata per la prima volta nella rivista *Teikoku bungaku* nel 1900) e *Gakusei* (Voce musicale), incluse all'interno della raccolta *Miotsukushi* (Marcatori di canale) nel 1901. Questi primi contatti con l'opera dannunziana sono riconosciuti come elementi che hanno aiutato a diffondere la poetica decadentista, insieme alla figura del letterato-esteta e quella della *femme fatale*, tra l'intelligenza giapponese (Hiraishi, 2023, pp. 203-204). Inoltre, per merito dell'influenza di Ueda, parte della sua fortuna nel primo dopoguerra è dovuta infatti anche all'interesse creatosi attorno alla vita privata dell'autore, soprattutto in seguito al suo "esilio" parigino nel 1910 (Yonekawa, 1991, p. 221) e al presunto coinvolgimento dell'opera nel tentativo di doppio suicidio di Morita Sōhei e Hiratsuka Haruko nel 1908, i quali si sarebbero verosimilmente ispirati al finale del romanzo. Tale evento fornirà da modello sia per la novella *Baien* (Fumo, 1909) di Morita, che per il romanzo *Sorekara* (E poi, 1910) di Natsume Sōseki, in cui si possono ritrovare i sopracitati elementi, figure e sensibilità tipiche dello stile dannunziano (Hiraishi, 2023, pp. 205-219). Tuttavia, le traduzioni di D'Annunzio in questo primo periodo sono mediate da fonti francesi, inglesi e tedesche, come è il caso di Ueda, Ishikawa e Ikuta (Hirayama 2011, p. 128; Hiraishi, 2023, p. 202): solo nel secondo dopoguerra saranno pubblicate traduzioni dirette dall'italiano. Hirayama (2011, pp. 115-118) considera brevemente le omissioni nelle ultime due sezioni del romanzo nella traduzione del 1928 di Ikuta, confronta queste sezioni con quelle nella versione di Soichi Nogami del 1961 e nella traduzione inglese di Georgina Harding del 1898 – a cui Ikuta ha fatto riferimento, insieme a quelle di Hornblow in inglese del 1897 e di Gagliardi in tedesco del 1904 – e mostra così le parti in cui Ikuta riprende direttamente dalla Harding e l'effetto di tale "traduzione a staffetta".

Nonostante la longeva fortuna dell'opera in Giappone, studi comparitivi sullo

stile delle diverse traduzioni, sia sul piano qualitativo che quantitativo, rimangono sporadici. A tal proposito, questa analisi preliminare, facendo riferimento alla precedente citata letteratura di ambito stilometrico e della linguistica dei *corpora*, intende individuare gli elementi che differenziano e accomunano lo stile dei traduttori dell'opera dannunziana. I risultati così ottenuti potranno potenzialmente fornire la base per ulteriori studi sul “traduzione” giapponese con *corpora* comparativi.

3. Metodologia

3.1. Costruzione del *corpus*

Per questo studio pilota sono state selezionate otto traduzioni complete del *Trionfo della morte*, pubblicate tra il 1913 e il 2010, basandosi su criteri di disponibilità e copertura testuale. Le traduzioni utilizzate sono le seguenti: Ishikawa (1913), Mikami (1921), Yaguchi (1927), Ikuta (1928), Harada (1939), Iwasaki (1958), Nogami (1961) e Waki (2010). Gli interi primi due libri, *Il passato* e *La casa paterna*, sono stati scelti come oggetto dello studio.

I testi sono stati digitalizzati e codificati in TEI-XML, uno standard per la rappresentazione digitale dei testi umanistici. La segmentazione è stata effettuata a livello di frase (<s>), e i principali elementi testuali sono stati annotati come segue:

- <p> per i paragrafi
- <said> per i dialoghi e monologhi interiori, con l'attributo @aloud per distinguere i primi (true) dai secondi (false)
- <q> per la comunicazione scritta, come la corrispondenza epistolare tra Giorgio e Ippolita.

I parlanti e gli interlocutori, se presenti, sono stati identificati tramite gli attributi @who e @toWhom. L'immagine sottostante riporta un esempio di tale codifica:

```
<p>
  <s>ジョルジオは羣集の動静を見廻してから、</s>
</p>
<p>
  <said aloud = "true" who="ジョルジオ" toWhom="イッポリタ">
    『
      <s>誰か高處から墜落ちたのだらう。</s>
    』
  </said>
</p>
```

Figura 1

Esempio di codifica utilizzata nel corpus.

Sono stati così individuati cinque diversi sotto-*corpora* per ogni traduzione: narrazione, i dialoghi del protagonista Giorgio, i dialoghi dell'eroina Ippolita, i monologhi interiori di Giorgio e la corrispondenza epistolare di Giorgio a Ippolita.

3.2. Analisi stilometrica

3.2.1. Tokenizzazione dei testi

I testi sono stati tokenizzati utilizzando **UniDic2UD** (Yasuoka, 2020), un *tokenizer* basato sulle *Universal Dependencies*, che supporta l'analisi morfologica, la lemmatizzazione e l'annotazione delle parti del discorso. Per gestire la presenza di *kyūjitai* e *kana* arcaici nelle traduzioni antecedenti il secondo dopoguerra, è stato utilizzato il dizionario *Kingendai kōgo shōsetsu* (Ogiso *et al.*, 2013), mentre per i testi più recenti è stato impiegato *Gendai kakikotoba* (Den *et al.*, 2007). Dopo la tokenizzazione, sono state calcolate metriche testuali comuni a tutti i sotto-*corpora*, insieme alle frequenze relative di un set di trigrammi rappresentativi.

3.2.2. Selezione e calcolo delle metriche testuali

Negli studi di traduzione basati sui *corpora*, pur non essendoci un consenso universale su quali caratteristiche esaminare, si prendono di solito in esame la frequenza di parti del discorso o metriche quali la lunghezza della frase e la diversità del vocabolario. Il seguente set di 22 caratteristiche lessicali, sintattiche e relative ai caratteri è stato quindi calcolato ed estratto per ogni documento:

Metriche Testuali
Lunghezza media della frase.
Rapporto <i>type-token</i> standardizzato.
Densità lessicale media della frase.
Dipendenza a distanza media della frase.
Frequenza relativa degli aggettivi.
Frequenza relativa delle apposizioni.
Frequenza relativa degli avverbi.
Frequenza relativa dei verbi ausiliari.
Frequenza relativa congiunzioni coordinanti.
Frequenza relativa dei determinatori.

Frequenza relativa delle interiezioni.
Frequenza relativa dei nomi comuni.
Frequenza relativa dei numerali.
Frequenza relativa delle particelle.
Frequenza relativa dei pronomi.
Frequenza relativa dei nomi propri.
Frequenza relativa della punteggiatura.
Frequenza relativa delle congiunzioni subordinanti.
Frequenza relativa dei verbi principali.
Proporzione di <i>hiragana</i> .
Proporzione di <i>kanji</i> .
Proporzione di <i>katakana</i> .

Tabella 2. Elenco delle metriche testuali utilizzate

La lunghezza media della frase è calcolata adottando un metodo simile a quello usato in Konishi (2016, pp. 46-47), attraverso il rapporto tra il numero totale di *bunsetsu* e il numero totale di frasi in un sotto-*corpus*. UniDic utilizza le *tan tan'i* (“unità corte”, i più piccoli elementi linguistici aventi significato in giapponese) come unità di base. Tuttavia, poiché si tratta di elementi testuali più simili al morfema che alla parola, non catturano a pieno la dimensione sintattica di una frase (Omura *et al.*, 2021). I *bunsetsu* sono invece definiti come clausole composte da una singola parola indipendente e da una o più parole dipendenti (ad esempio, ausiliari o particelle).

Il rapporto *type-token* è calcolato come la media del rapporto tra i diversi tipi di parole su una finestra di 1.000 *token*, consentendo così di tenere conto della differenza di lunghezza nei testi di alcuni sotto-*corpora*. La densità lessicale media è misurata come la media del numero totale di *token* lessicali (nomi, aggettivi, verbi, avverbi e numerali) rispetto al numero totale di *token* in una frase. La dipendenza a distanza è definita come la distanza posizionale tra una testa *x* e il suo dipendente *y*, utilizzando una formula proposta in Liu, Hudson e Feng (2009, p. 166).

Alcune classi di parole devono essere ulteriormente specificate: sebbene il giapponese non abbia articoli o determinatori come classe, un piccolo numero di elementi adnominali con una funzione simile (*kono* “questo”, *sono* “quello”, *ano* “quello”, *dono* “quale”) sono classificati in tale maniera. La classe delle apposizioni comprende le particelle post posizionali di caso e avverbiali, mentre la classe delle “particelle” riguarda principalmente quelle post posizionali di fine frase, usate per esprimere il tono e l'affetto.

3.2.3. Selezione e calcolo dei trigrammi

Per identificare le espressioni caratteristiche dei vari sotto-*corpora*, sono stati estratti trigrammi di *token* da ciascuna traduzione, senza applicare filtri preliminari. Dopo la tokenizzazione, si è proceduto alla standardizzazione dell'ortografia di *kana* e *kanji*.

Per selezionare i trigrammi distintivi, è stata utilizzata l'analisi di Zeta (Burrows, 2007), un metodo per misurare la prominenza di una caratteristica confrontando due gruppi di testi. In questo studio, il valore di Zeta è stato calcolato iterativamente, confrontando ogni traduzione con tutte le altre all'interno di ciascun sotto-*corpus*. Questo approccio permette di evidenziare le espressioni distintive di ciascun documento, riducendo l'influenza degli elementi più frequenti nel *corpus*.

Per ogni categoria, i 20 trigrammi più caratteristici sono stati classificati e selezionati sulla base del loro valore di Zeta (elencati in ordine decrescente nella tabella 3).

	Narrazione	Dialoghi (Giorgio)	Dialoghi (Ippolita)	Monologhi interiori	Epistole
1.	て_い_た <i>te-i-ta</i>	て_い_た <i>te-i-ta</i>	まし_た_。 <i>mashi-ta-</i>	の_だ_。 <i>no-da-</i>	て_いる_。 <i>te-iru-</i>
2.	い_た_。 <i>i-ta-</i>	まし_た_。 <i>mashi-ta-</i>	でし_た_。 <i>deshi-ta.</i>	だろう_か_? <i>darō-ka-?</i>	まし_た_。 <i>mashi-ta-</i>
3.	だつ_た_。 <i>dat-ta</i>	の_です_。 <i>no-desu-</i>	た_わ_。 <i>ta-wa-</i>	て_いる_の <i>te-iru-no</i>	の_です_。 <i>no-desu-</i>
4.	あつ_た_。 <i>at-ta-</i>	でし_た_。 <i>deshi-ta-</i>	た_の_。 <i>ta-no-</i>	の_だろう_か <i>no-darō-ka</i>	、_きみ_の <i>,-kimi-no</i>
5.	で_あつ_た <i>de-at-tta</i>	た_ん_だ <i>ta-n-da</i>	て_い_た <i>te.i-ta</i>	の_か_? <i>no-ka-?</i>	ぼく_は_、 <i>boku-wa-</i>
6.	彼_は_、 <i>kare-wa-</i>	い_た_。 <i>i-ta-</i>	ませ_ん_わ <i>mase-n-wa</i>	いる_の_だ <i>iru-no-da</i>	、_ぼく_の <i>,-boku-no</i>
7.	の_だつ_た <i>no-dat-ta</i>	ん_だ_。 <i>n-da-</i>	ませ_ん_でし <i>mase-n-deshi</i>	だろう_か_。 <i>darō-ka-</i>	、_ぼく_は <i>,-boku-wa</i>
8.	で_ある_。 <i>de-aru-</i>	、_わたし_は <i>,-watashi-wa</i>	まし_た_わ <i>mashi-ta-wa</i>	自分_は_、 <i>jibun-wa-</i>	、_あな た_の <i>,-anata-no</i>

9.	て_いる_。 <i>te-iru-</i>	、_ぼく_は <i>,-boku-wa</i>	ん_で_し_た <i>n-deshi-ta</i>	も_し_れ_ない <i>mo-shire-nai</i>	の_だ_。 <i>no-da-</i>
10.	の_中_に <i>no-naka-ni</i>	い_まし_た <i>i-mashi-ta</i>	の_よ_。 <i>no-yo-</i>	、_わたし_は <i>,-watashi-wa</i>	わたし_は_、 <i>watashi-wa-</i>
11.	よう_に_、 <i>yō-ni-</i>	きみ_は_ぼく <i>kimi-wa-boku</i>	た_の_よ <i>ta-no-yo</i>	、_自分_は <i>,-jibun-wa</i>	の_中_に <i>no-naka-ni</i>
12.	て_来_た <i>te-ki-ta</i>	、_きみ_の <i>,-kimi-no</i>	だ_つ_た_わ <i>dat-ta-wa</i>	て_いる_。 <i>te-iru-</i>	が_、_ぼく <i>ga-,-boku</i>
13.	そう_し_た <i>sō-shi-ta</i>	か_い_？ <i>ka-i-?</i>	わ_ね_。 <i>wa-ne-</i>	か_も_し_れ <i>ka-mo-shire</i>	し_て_いる <i>shi-te-iru</i>
14.	いつ_た_。 <i>it-ta.</i>	、_きみ_は <i>,-kimi-wa</i>	です_もの_。 <i>desu-mono-</i>	、_自分_の <i>,-jibun-no</i>	に_は_、 <i>ni-wa-</i>
15.	の_で_ある <i>no-de-aru</i>	わたし_は_あ なた <i>watashi-wa -anata</i>	い_まし_た <i>i-mashi-ta</i>	に_な_つ_て <i>ni-nat-te</i>	、_わた し_は <i>,-watashi-wa</i>
16.	言_つ_た_、 <i>it-ta,</i>	ぼく_は_、 <i>boku-wa-</i>	です_よ_。 <i>desu-yo-</i>	で_ある_。 <i>de-aru</i>	て_いる_の <i>te-iru-no</i>
17.	来_た_。 <i>ki-ta-</i>	て_い_まし <i>te-i-mashi</i>	ます_わ_。 <i>masu-wa-</i>	て_い_た <i>te-i-ta</i>	ぼく_に_は <i>boku-ni-wa</i>
18.	た_の_で <i>ta-no-de</i>	、_ぼく_の <i>,-boku-no</i>	に_は_、 <i>ni-wa-</i>	で_あ_ろ_う_か <i>de-arō-ka</i>	わた し_に_は <i>wata- shi-ni-wa</i>
19.	の_よう_に <i>no-yō-ni</i>	の_中_に <i>no-naka-ni</i>	まし_た_の <i>mashi-ta-no</i>	た_の_だ <i>ta-no-da</i>	て_い_た <i>te-i-ta</i>
20.	し_た_。 <i>shi-ta-</i>	だ_つ_た_。 <i>dat-ta-</i>	、_わたし_の <i>,-watashi-no</i>	いる_の_だ ろ_う <i>iru-no-darō</i>	の_こ_と_を <i>no-koto-o</i>

Tabella 3. I 20 trigrammi più caratteristici per ogni sotto-corpus.

3.2.4. Analisi delle componenti principali e *clustering* gerarchico

In primo luogo, l'Analisi delle Componenti Principali (PCA) è stata impiegata per ridurre la complessità dei dati, individuando le variabili più rilevanti. Successivamente, il *clustering* gerarchico è stato utilizzato per raggruppare le traduzioni in base alla loro similarità stilistica.

I risultati di questa analisi sono rappresentati con un dendrogramma, dove l'altezza dei rami indica la dissimilarità tra i gruppi individuati. Il numero di *cluster* è stato determinato principalmente ricorrendo a una metrica quantitativa, il coefficiente di Silhouette (*s*), che misura quanto ogni testo è simile al proprio *cluster* rispetto agli altri. Il punteggio varia da -1 a 1, dove un valore superiore al 0.4 indica che il punto è ben abbinato al proprio *cluster*.

Infine, per identificare gli elementi distintivi di ciascun *cluster*, è stata applicata un'analisi della varianza (ANOVA), che ha permesso di individuare le caratteristiche stilistiche più rilevanti per ogni gruppo di traduzioni (con $p < .05$).

4. Risultati

4.1. Sotto-corpora narrativi

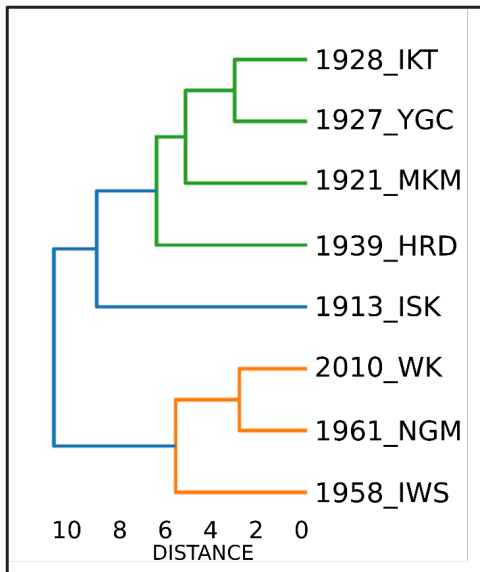


Figura 2.
Clustering gerarchico (narrazione, metriche testuali)

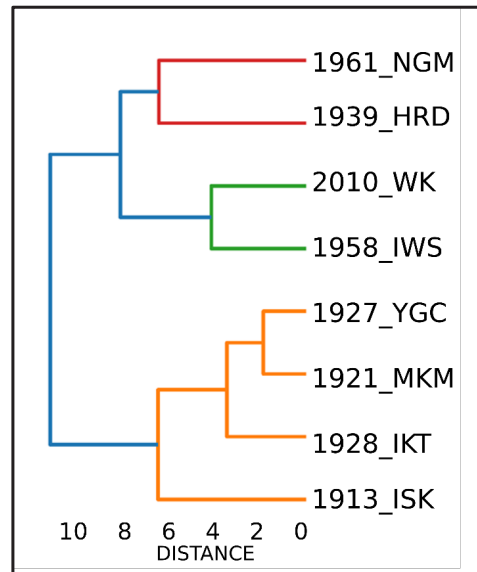


Figura 3.
Clustering gerarchico (narrazione, metriche trigrammi)

Entrambe le soluzioni di *clustering* ottenute mostrano una netta differenziazione diacronica. Nel caso delle metriche testuali, si possono individuare due gruppi, con il secondo dopoguerra come spartiacque. La caratteristica principale che differenzia i due gruppi è una maggiore diversità lessicale nelle opere antecedenti il dopoguerra, evidenziata da un rapporto *type-token* più alto (0.38) e da una maggiore proporzione di *kanji* (32%), elementi che contribuiscono alla densità informativa del testo. Inoltre, il *cluster* delle opere più recenti presenta una frequenza più elevata di nomi propri, oltre il doppio rispetto al gruppo opposto. Questo fattore suggerisce una tendenza delle traduzioni del dopoguerra a esplicitare i soggetti verbali, una caratteristica già documentata nella ricerca sul “traduzionese” giapponese. Due dei supposti “universali” del traduzionese sono la semplificazione e l’esplicazione (Baker, 1993): per quanto riguarda quest’ultima, Pápai (2008) osserva che essa è spesso associata a una ridotta varianza lessicale. Tuttavia, per verificare se effettivamente le traduzioni più recenti mostrino una maggiore propensione all’esplicitazione, sarà necessario un confronto con un *corpus* composto da testi non tradotti.

Nel *clustering* basato sui trigrammi, le opere meno recenti (1913, 1921, 1927, 1928) formano un *cluster*, mentre le restanti quattro si dividono in due gruppi: uno composto da Iwasaki (1958) e Waki (2010) e un altro da Harada (1939) e Nogami (1961). L’elemento più discriminante tra questi due *cluster* è l’uso della copula *da* nella sua forma passata *datta*, inclusa la costruzione composita con la particella *no* (*no-datta*). Sebbene vi siano pareri contrastanti sul ruolo di *no-da*, è generalmente accettato che questa costruzione esprima la soggettività del parlante, trasformando l’enunciato in una spiegazione, commento, o interpretazione personale. Ciò la rende tipica del parlato e della scrittura in prima persona. Tuttavia, quando la costruzione appare nella forma passata (*no-datta*), il punto di vista del parlante si sposta a un tempo anteriore, generando un effetto di narrazione retrospettiva tipico dei testi in terza persona (Noda, 1997; Tanomura, 2002).

In ambito traduttivo, Kaneko (2018) ha analizzato dieci traduzioni giapponesi del racconto *Dušečka* di Čechov (1899), pubblicate tra il 1934 e il 2010. Nonostante la frequenza di *no-da* nelle parti narrative sia simile tra le traduzioni, la sua distribuzione non è uniforme e la correlazione tra il suo uso e le caratteristiche del testo originale non è risultata chiara. Un’analisi comparativa delle traduzioni in ciascun *cluster* mostra un fenomeno simile: la presenza della costruzione *no-datta* nelle traduzioni della stessa frase è limitata, attestandosi intorno al 10% nelle versioni di Waki (2010) e Iwasaki (1958) e al 7% in quelle di Harada (1939) e Nogami (1961). Inoltre, nessuna occorrenza è condivisa da tutti e quattro i traduttori. Sebbene, come osservato da Kaneko (2018), vi sia un’ampia variabilità nelle scelte traduttive, un’analisi qualitativa conferma che il suo uso rispecchia generalmente le funzioni descritte in letteratura. Se si considera infatti la situazione narrativa che viene tradotta, si tratta principalmente di affermazioni, giudizi o commenti da parte della voce narrante. Di seguito sono riportate due occorrenze differenti, ma in entrambe si può notare, nell’originale, l’accentuazione della voce del narratore onnisciente:

Ed egli era legato a quella suprema tristezza, forse fino alla morte.

Soshite, kare-wa sō-shita saikō-no hiai-ni, osoraku-wa shinuru-toki-made shibaritsu-ke-rare-te-iru-no-datta. (Harada, 1939)

Egli dava al suo spirito attitudini arteficiose e irreparabili.

Tsumari, kare-ni-wa jibun-no seishin-wo amari-ni-mo konemawashi, shūshū-tsuka-naku-shite-shimau-seiheki-ga aru-no-datta. (Waki, 2010)

Un ulteriore elemento discriminante tra i *cluster* riguarda l'aspetto e il tempo narrativo. Nei gruppi composti dalle traduzioni più recenti, l'uso della forma progressiva *-te-iru* al presente è marginale (attorno allo 0.03%), mentre la forma passata *-te-ita* appare con frequenza almeno doppia rispetto al *cluster* più antichi. Riguardo al suo uso nella narrativa giapponese, Kudō (1995) osserva che *-te-ita* è generalmente impiegato per far avanzare la narrazione, mentre *-te-iru* è più frequentemente utilizzato per portare in primo piano l'esperienza soggettiva di un personaggio nel presente narrativo. Uno studio recente di Lee (2024) ha confrontato l'occorrenza delle due forme in un *corpus* di romanzi di periodo *Meiji* (1887-1909) e in uno del dopoguerra (1948-1980), evidenziando che l'uso di *-te-iru* è molto più frequente nei romanzi *Meiji*, mentre il suo uso limitato *corpus* più recente si allinea con le osservazioni di Kudō.

Sebbene in questo studio il confronto riguardi traduzioni dell'epoca Taishō e dell'inizio dell'era Shōwa, un ulteriore confronto del rapporto tra le frasi narrative in *-te-iru* e quelle in *-te-ita* suggerisce una differenza stilistica simile: nei *cluster* più antichi, tale rapporto è al massimo di 1:1.7, mentre nelle traduzioni più recenti varia da un minimo di 1:8 (in Iwasaki) a un massimo di 1:63 (in Waki). Pertanto, anche in questo gruppo si osserva l'influenza di strategie narrative più vicine a modelli stilistici meno moderni.

4.2. Sotto-corpora dialogici (Giorgio)

Per quanto riguarda i sotto-corpora dei dialoghi del protagonista Giorgio, la divisione suggerita dal *clustering* effettuato con le metriche testuali separa le traduzioni di Ishikawa (1913), Yaguchi (1927) e Harada (1939) dalle restanti. Tale risultato sembra dovuto alla minore complessità sintattica di tali opere: la lunghezza media della frase e la dipendenza a distanza risultano inferiori rispetto agli altri testi, portando a un ritmo dialogico più frammentato. In media, le frasi nei dialoghi di queste traduzioni sono più brevi di circa un *bunsetsu*, suggerendo una tendenza a segmentare le informazioni in unità più semplici. Anche la frequenza delle congiunzioni subordinanti è inferiore rispetto all'altro *cluster*, indicando un possibile minore grado di ipotassi nei dialoghi. Tuttavia, il valore silhouette per questo raggruppamento è il più basso tra tutti quelli ottenuti (0.19), suggerendo che i due *cluster* individuati non presentano differenze sostanziali.

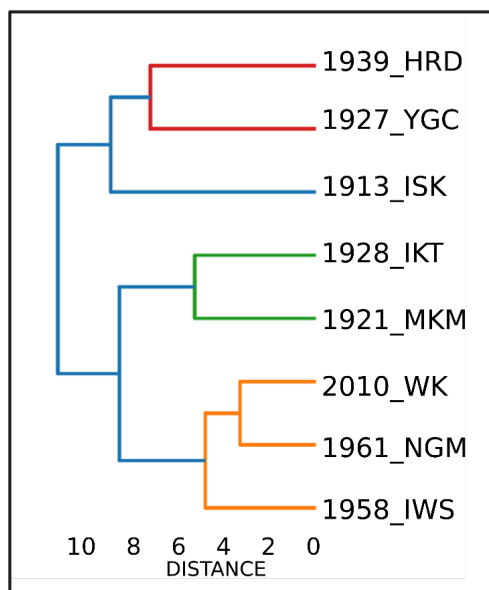


Figura 4.
Clustering gerarchico
(Giorgio, metriche testuali)

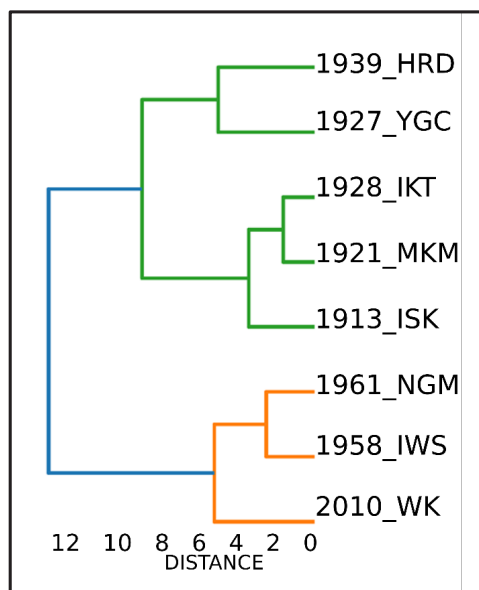


Figura 5.
Clustering gerarchico
(Giorgio, metriche trigrammi)

Il *clustering* con i trigrammi, invece, propone una soluzione a tre *cluster* più marcata, con un valore di silhouette relativamente elevato (0.47), che organizza i gruppi principalmente in base alla diacronia: il primo gruppo è composto dalle traduzioni di Iwasaki (1958), Nogami (1961) e Waki (2010); il secondo da Ishikawa (1913), Mikami (1921) e Ikuta (1928); il terzo da Yaguchi (1927) e Harada (1939). Ancora una volta, le traduzioni del dopoguerra si distinguono, come nel sotto-corpus narrativo, per il frequente uso dell'ausiliare *-te-ita* e del costrutto *no-da* (inclusa la forma contratta *n-da*, maggiormente tipica del parlato). Parallelamente, risultano discriminanti i trigrammi contenenti il pronome di seconda persona *kimi*, come in *kimi-wa-boku* o *,-kimi-no*, che sono totalmente assenti negli altri due *cluster*.

L'uso di determinati elementi linguistici nei dialoghi delle opere di finzione, come pronomi di prima e seconda persona, forme ausiliari e particelle finali, è stato concettualizzato da Kinsui (2003; 2014) sotto il termine *yakuwarigo* ("linguaggio di ruolo"). Questo fenomeno si riferisce all'impiego sistematico di elementi linguistici per caratterizzare i personaggi in base a stereotipi sociali, piuttosto che riflettere l'uso reale della lingua. Per quanto riguarda il genere, la ricerca sulla linguistica giapponese ha tentato di categorizzare le forme linguistiche "maschili" e "femminili". Shibamoto-Smith (2004) distingue i pronomi di seconda persona in base al loro grado di formalità: *kisama*, *ore* e *omae* sono informali, *boku* e *kimi* sono relativamente neutri, mentre *anata*, *watashi* e *watakushi* sono estremamente formali. Kinsui (2023) approfondisce questa distinzione introducendo il concetto di *face risk*, ovvero

il grado di tensione sociale legato alla relazione tra parlante e interlocutore: *boku*, *kimi*, *kisama*, *omae* e *ore* sono usati in situazioni di *face risk* medio-basso, mentre *anata*, *watakushi* e *watashi* sono tipici di situazioni ad alta tensione sociale.

Analizzando qualitativamente l'uso di questi pronomi nei dialoghi di Giorgio, emerge una chiara distinzione tra i *cluster*. Le traduzioni più recenti (1958, 1961, 2010) privilegiano l'uso della coppia *boku/kimi*, mentre nelle traduzioni più antiche prevalgono *watashi/anata*, con la sola eccezione di Yaguchi (1927), che impiega *boku* per la prima persona. La traduzione di Ishikawa (1913) si distingue da tutte le altre per una maggiore varietà nei pronomi: Giorgio si riferisce a sé stesso con *ore* e agli interlocutori con *omae*, sostituendoli con *watashi* e *anata* solo quando parla con familiari più anziani.

Questa tendenza indica una progressiva semplificazione stilistica nelle traduzioni più recenti, con un generale spostamento verso un tono più colloquiale e meno formale, indipendentemente dal contesto interlocutorio. Tale fenomeno è supportato anche dalla distribuzione delle forme verbali: i trigrammi contenenti la copula *-desu* e l'ausiliare *-masu*, tipici di un registro formale, sono praticamente assenti nelle traduzioni del dopoguerra, mentre vanno a caratterizzare *cluster* più antichi volte maggiore (in particolare nelle traduzioni di Yaguchi e Harada, loro frequenza media è quasi quattro volte rispetto al secondo *cluster*).

4.3. Sotto-corpora dialogici (Ippolita)

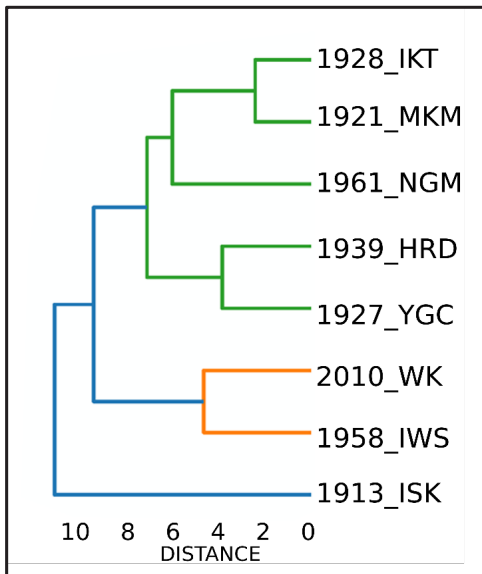


Figura 6.
Clustering gerarchico
(Ippolita, metriche testuali)

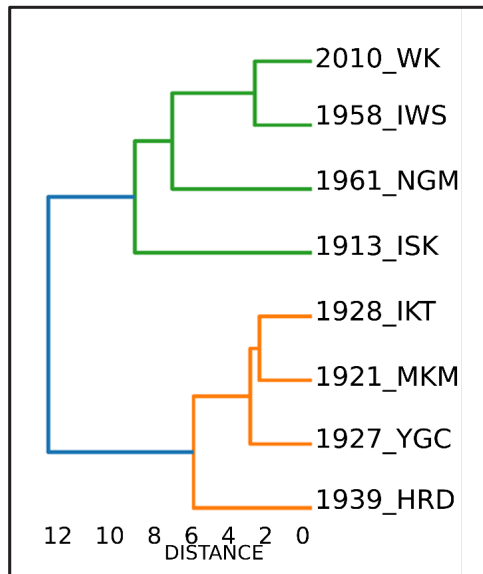


Figura 7.
Clustering gerarchico
(Ippolita, metriche trigrammi)

Nell'analisi dei dialoghi di Ippolita, le soluzioni di *clustering* ottenute mostrano analogie con quanto osservato nel sotto-corpus dei dialoghi del protagonista. A livello di metriche testuali, la traduzione di Ishikawa (1913) si distingue nettamente da tutte le altre per una maggiore varietà sul piano lessicale, con valori in media più alti del 4% sia per il rapporto *type-token* che per la densità lessicale. Tentativi di raggruppamento con tre o quattro *cluster* non hanno prodotto esiti significativi, suggerendo che, sotto questo aspetto, le traduzioni non presentano differenze sostanziali.

Nel *clustering* basato sui trigrammi, invece, i risultati non evidenziano una netta divisione diacronica. Il primo *cluster* è composto da Mikami (1921), Yaguchi (1927), Ikuta (1928) e Harada (1939), mentre Ishikawa (1913) si raggruppa con le traduzioni del dopoguerra. Anche in questo sotto-corpus, la presenza di forme linguistiche tipiche di un registro formale – l'ausiliare *-masu* e la copula *-desu* – risulta essere un elemento chiave nell'organizzazione dei *cluster*: pur non completamente assenti nel secondo gruppo, questi trigrammi sono mediamente dieci volte più frequenti nel primo. Si conferma quindi la tendenza delle traduzioni del dopoguerra a adottare un registro più informale nei dialoghi dei personaggi, senza evidenti differenze sul piano quantitativo.

Un ulteriore fattore discriminante tra i due *cluster* è dato dalla frequenza di costrutti formati dall'ausiliare *-ta* seguito da particelle finali come *-wa*, *-no* e *-noyo*. Queste particelle sono elementi caratteristici dell'*onna-kotoba* (linguaggio femminile), che, sebbene cadute in disuso nel parlato contemporaneo, sono frequentemente impiegate nelle opere di finzione per rappresentare un registro stereotipato del parlato femminile (Kinsui, 2003; Kinsui, 2014; Nakamura, 2013; Shibamoto-Smith, 2004). L'uso dell'*onna-kotoba* è considerato un tratto tipico del "traduzione-nese" giapponese: studi qualitativi mostrano che, mentre nelle opere autoctone pubblicate negli ultimi anni si osserva una maggiore varietà nelle strategie linguistiche adottate per i personaggi femminili, nelle traduzioni si continua a ricorrere frequentemente a questi elementi indipendentemente dal carattere o dal ruolo del personaggio (Nakamura, 2013; Furukawa, 2013). L'analisi quantitativa di Fukuchi Meldrum (2009) non ha confermato in modo definitivo un uso eccessivo dell'*onna-kotoba* nelle traduzioni, ma ha evidenziato come il genere del traduttore possa influenzare le scelte linguistiche.

In questo caso specifico, le tre traduzioni più recenti, insieme a quella di Ishikawa, mostrano una frequenza significativamente più alta di trigrammi contenenti queste particelle finali, con valori medi circa quattro volte superiori rispetto al primo *cluster*. Se da un lato queste caratteristiche sono presenti in tutti i testi analizzati, e non si osserva la stessa variazione nell'uso dei pronomi registrata nei dialoghi del protagonista, si può ipotizzare che le traduzioni del dopoguerra tendano a enfatizzare lo stereotipo della parlata femminile. Questa ipotesi, in linea con quanto riportato dalla letteratura qualitativa, andrebbe verificata con un confronto diretto tra testi tradotti e opere originali giapponesi pubblicate nello stesso periodo.

4.4. Sotto-corpora dei monologhi interiori

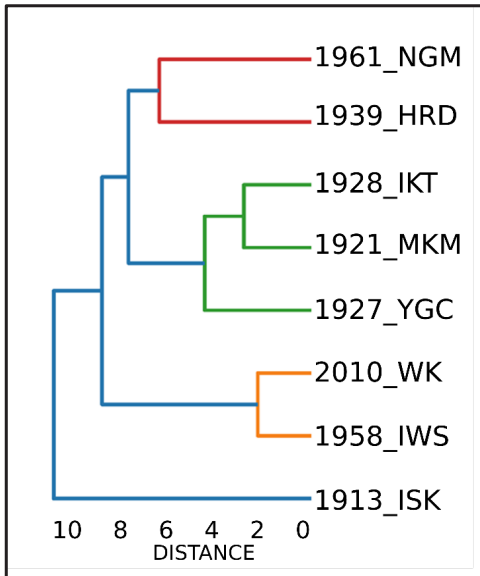


Figura 8.
Clustering gerarchico
(monologhi interiori, metriche testuali)

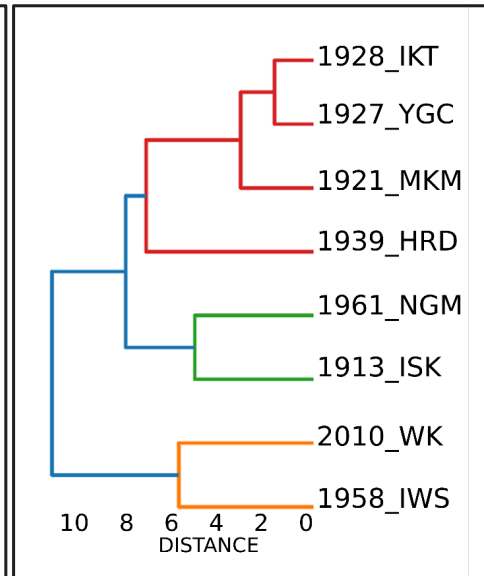


Figura 9.
Clustering gerarchico
(monologhi interiori, metriche trigrammi)

Nell'analisi dei monologhi interiori di Giorgio mediante le metriche testuali, per ottenere una visione più dettagliata delle differenze stilistiche, è stata utilizzata una soluzione a quattro *cluster*. I coefficienti di silhouette per due *cluster* (0.31) e per quattro (0.27) risultano relativamente simili, permettendo così una visione più granulare dei dati.

Il primo *cluster* comprende le traduzioni di Iwasaki (1958) e Waki (2010), il secondo quelle di Mikami (1921), Yaguchi (1927) e Ikuta (1928), il terzo Harada (1939) e Nogami (1961), mentre la traduzione di Ishikawa (1913) si distingue nettamente dalle altre, costituendo un *cluster* a sé stante. Tale separazione è motivata dalla sua maggiore diversità lessicale, con i valori più elevati sia per il rapporto *type-token* (0.39) che per la densità lessicale (0.38), confermando una tendenza stilistica costante in tutte le parti della narrazione. Il secondo *cluster*, pur non raggiungendo i livelli di Ishikawa, mantiene un livello relativamente elevato di varietà del lessico (0.36). Come già osservato nel sotto-corpus narrativo, anche qui le traduzioni più recenti mostrano una tendenza alla semplificazione sintattica e lessicale. Una peculiarità del *cluster* di Harada e Nogami è la più alta frequenza relativa di pronomi, superiore di oltre il doppio rispetto a Iwasaki, Waki, e Ishikawa. Questa caratteristica suggerisce una diversa gestione dell'elemento referenziale nella rappresentazione dei monologhi interiori, che potrebbe essere connessa a una maggiore esplicitazione del soggetto in tali testi.

L'analisi con i trigrammi fornisce risultati ancora più significativi. In questo caso, le traduzioni di Iwasaki e Waki si separano nettamente da tutte le altre, formando un *cluster* caratterizzato da un uso elevato della forma congetturale *-darō* (*darō-ka-?*, *iru-no-darō*), del pronome riflessivo *jibun* e della struttura *te-iru-no*, suggerendo una strategia traduttiva simile nella rappresentazione del monologo interiore. A questo stadio dell'analisi non è possibile confermare se Waki si sia basato sulla traduzione di Iwasaki nella stesura della propria versione, ma la notevole vicinanza tra le due opere in termini di scelte lessicali è stata osservata anche nei sotto-*corpora* narrativo e dei dialoghi di Ippolita.

4.5. Sotto-*corpora* delle epistole

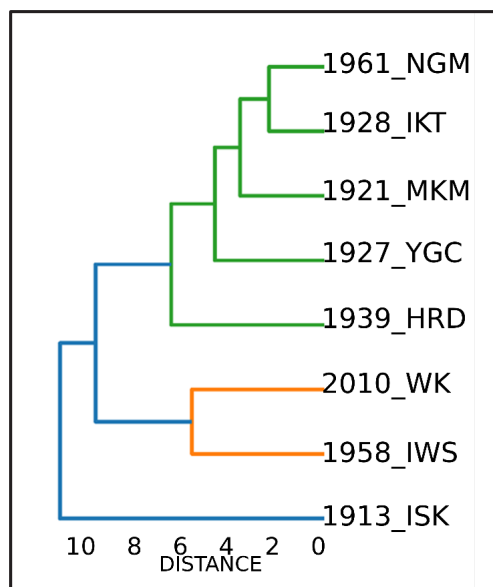


Figura 10.
Clustering gerarchico
(Epistole, metriche testuali)

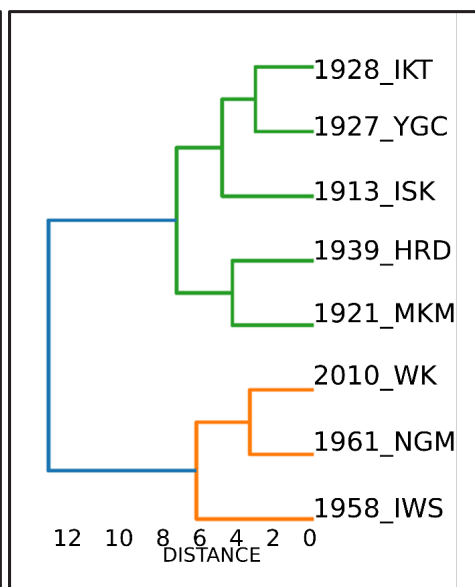


Figura 11.
Clustering gerarchico
(Epistole, metriche trigrammi)

Infine, se si analizzano i risultati inerenti alla corrispondenza epistolare di Giorgio, sul piano delle metriche testuali, la traduzione di Ishikawa (1913) si distingue nettamente dalle altre, formando un *cluster* a sé stante. Come già osservato nel sotto-*corpus* narrativo, essa presenta una maggiore varietà lessicale, con un rapporto *type-token* più elevato (0.39) e una densità lessicale superiore (0.38).

Per quanto riguarda la distribuzione dei trigrammi, i risultati mostrano raggruppamenti e caratteristiche simili a quelli già osservati nei dialoghi di Giorgio. Le

traduzioni di Iwasaki (1958), Nogami (1961) e Waki (2010) si distanziano dal resto, nuovamente per la differente distribuzione dei pronomi di seconda persona, con un uso più frequente di *kimi* e *boku* rispetto ad *anata*. Questa tendenza conferma la predilezione delle traduzioni più recenti per un tono più informale e colloquiale nei dialoghi del protagonista, una scelta stilistica che si riflette anche nella sua corrispondenza epistolare. In altre parole, la modalità di comunicazione epistolare non presenta differenze significative rispetto ai dialoghi dello stesso personaggio, suggerendo una certa coerenza nella rappresentazione della sua voce narrativa tra contesti comunicativi differenti.

5. Conclusioni, limiti e futuri sviluppi

Le analisi condotte in questo studio hanno evidenziato quantitativamente come le traduzioni giapponesi del *Trionfo della morte* si assomiglino e si differenzino sul piano stilistico. In generale, le metriche testuali hanno mostrato una capacità limitata di individuare *cluster* nettamente distinti, suggerendo che le differenze morfosintattiche tra le traduzioni non siano sempre marcate. Tuttavia, nei monologhi interiori e nella narrazione, alcuni parametri – come la densità lessicale e la varietà del vocabolario – hanno evidenziato tendenze coerenti con quelle riscontrate nell'analisi dei trigrammi. Nella maggior parte dei sotto-*corpora*, emerge una chiara divisione diacronica tra le traduzioni antecedenti e quelle successive al dopoguerra. Sul piano narrativo, i due gruppi si distinguono sia per la complessità sintattica, sia per determinate scelte lessicali, quali l'uso della copula e di specifici ausiliari. La distribuzione e l'uso di tali caratteristiche appaiono inoltre coerenti con la letteratura esistente, sia qualitativa che quantitativa.

Nei sotto-*corpora* dialogici, il grado di formalità emerge come un elemento chiave nella formazione dei *cluster*, con le traduzioni più recenti che tendono all'uso di un registro più colloquiale. La corrispondenza epistolare ha mostrato tendenze simili a quelle osservate nei dialoghi, suggerendo una coerenza nello stile adottato per la resa della voce dei personaggi principali.

Le traduzioni di Iwasaki (1958) e Waki (2010) si sono dimostrate particolarmente affini, risultando raggruppate insieme in tre sotto-*corpora* su cinque, non solo per la somiglianza lessicale, ma anche per la distribuzione di costrutti stilisticamente marcati, come l'uso della copula *no-da* e delle forme verbali congetturali. La traduzione di Ishikawa (1913), invece, si distingue come la più idiosincratica, separandosi completamente dalle altre in numerose soluzioni di *clustering* e mostrando talvolta caratteristiche più simili a traduzioni lontane nel tempo.

Tuttavia, le considerazioni tratte da questa analisi riguardano esclusivamente il confronto tra lo stile dei diversi traduttori. Per verificare se le differenze diacroniche riflettano effettivamente un'evoluzione nel linguaggio della traduzione, sarà necessario effettuare confronti con *corpora* comparativi, composti sia da traduzioni che da

testi originali in giapponese prodotti nello stesso arco temporale. Inoltre, per quanto riguarda il metodo di selezione dei trigrammi, l'analisi di Zeta tende a individuare parole a frequenza relativamente alta, il che potrebbe non permettere di cogliere tutte le sfumature stilistiche. Un'analisi basata su trigrammi a frequenza media o l'utilizzo di tecniche come la *Tf-Idf* (*Term Frequency-Inverse Document Frequency*) potrebbe fornire prospettive differenti sulle strategie adottate dai vari traduttori.

Riferimenti bibliografici

- Baker, Mona (1993). "Corpus Linguistics and Translation Studies: Implications and Applications." In Baker, Mona; Francis, Gill; Tognini-Bonelli, Elena (a cura di). *Text and Technology: In Honour of John Sinclair*. Amsterdam: John Benjamins, pp. 233-250.
- Burrows, John (2002). "The Englishing of Juvenal: Computational Stylistics and Translated Texts". *Style*, 36 (4), pp. 677-699.
- Burrows, John (2007). "All the Way Through: Testing for Authorship in Different Frequency Strata". *Literary and Linguistic Computing*, 22 (1), pp. 27-47.
- Chesterman, Andrew (2004). "Beyond the particular". In Mauranen, Anna; Kujamäki Pekka (a cura di), *Translation universals: Do they exist?*. Amsterdam: John Benjamins, pp. 33-50.
- Den, Yasuharu; Ogiso, Toshinobu; Ogura, Hideki, Yamada, Atsushi; Menematsu, Nobuaki; Uchimoto, Kiyotaka, Koiso; Hanae (2007). "Kopasu nihongogaku no tame no gengo shigen: keitaisokaiseki yō denshika jisho no kaihatsu to sono ōyō". *Nihongo kagaku*, 22, pp. 101-123.
- Fukuchi Meldrum, Yukari (2009). "Translationese in Japanese Literary Translation". *Traduction, Terminologie, Redaction*, 22 (1), pp. 93-118.
- Furukawa, Hiroko. (2013). "Onna-kotoba to honyaku - risō no onna-rashisa he no bunka-nai honyaku". *Tsuyaku-honyaku-kenkyū*, 17, pp. 93-111.
- Hiraishi, Noriko (2023). "Fin-de-Meiji as Fin-de-siècle: D'Annunzio and Japanese Literature". In Segnini, Elisa; Subialka, Michael (a cura di). *Gabriele D'Annunzio and World Literature: Multilingualism, Translation, Reception*. Edinburgh: Edinburgh University Press, pp. 201-220.
- Hirayama, Jōji (2011). *Danuntsuio to nihon kindai bungaku*. Tōkyō: Shironsha.
- Ji, Meng; Oakes, Michael P. (2012). "A Corpus Study of Early English Translations of Cao Xueqin's *Hongloumeng*". In Oakes, Michael P.; Ji, Meng (a cura di). *Quantitative Methods in Corpus-based Translation Studies*. Amsterdam: John Benjamins, pp. 177-208.
- Kaneko, Yuriko (2018). "Asupekuto no kankyū to modariti no nōtan: noda-bun o meguru nichiro taishō-kenkyū". *Kōbe gaidai ronsō*, 69 (1), pp. 81-93.
- Kinsui, Satoshi (2003). *Vācharu nihongo: yakuwarigo no nazo*. Tōkyō: Iwanami Shoten.
- Kinsui, Satoshi (2014). *Yakuwarigo no shōjiten*. Tōkyō: Kenkyūsha.
- Kinsui, Satoshi (2023). "Yakuwarigo no jendā to pawā". *Shakaigengogaku*, 26 (1), pp. 37-48.
- Kudō, Mayumi (1995). *Asupekuto tensu taikai to tekusuto - gendai nihongo no jikan no hyōgen* -. Tōkyō: Hitsuji Shobō.
- Konishi, Hikari (2016). "Kindai kōgobu honyaku shōsetsu kōpasu kochiku no gaiyō to keiryō bunseki". *NINJAL Research Papers*, 11, pp. 37-61.

- Lee, Gyomin (2024). “Ji no bun deno jōtai hyōgen no tensu keishiki no shiyō yōsō - meiji-ki to sengo no sanninshō-shōsetsu wo taishō toshite —”. *Nihongo kyōiku*, 109, pp. 89-107.
- Liu, Haitao, Hudson, Richard; Feng, Zhiwei (2009). “Using a Chinese Treebank to Measure Dependency distance”. *Corpus Linguistics and Linguistic Theory*, 5 (2), pp. 161-174.
- Lynch, Gerald; Vogel, Carl (2018). “The Translator’s Visibility: Detecting Translatorial Fingerprints in Contemporaneous Parallel Translations”. *Computer Speech & Language*, 52, pp. 79-104.
- Mauranen, Anna (2008). “Universals Tendencies in Translation”. In Anderman, Gunill; Rogers, Margaret. (a cura di). *Incorporating Corpora: the linguist and the translator*. Bristol: Multilingual Matters, pp. 32-48.
- Mauranen, Anna; Kujamäki, Pekka (2004) (a cura di). *Translation Universals: Do they exist?*. Amsterdam: John Benjamins.
- Nakamura, Momoko (2013). *Hon'yaku ga tsukuru nihongo: hiroin wa [onna kotoba] o hanashitsuzukeru*. Tōkyō: Hakutakusha.
- Noda, Harumi (1997). *[No(da)] no kinōbungaku*. Tōkyō: Kuroshio Shuppan.
- Ogiso, Toshinobu; Komachi, Mamoru; Matsumoto, Yuji. (2013). “Morphological Analysis of Historical Japanese Text”. *Journal of Natural Language Processing*, 20 (5), pp. 727-748.
- Omura, Mai; Wakasa, Aya; Asahara, Masayuki (2021). “Word Delimitation Issues in UD Japanese”. *Proceedings of the Fifth Workshop on Universal Dependencies (UDW, SyntaxFest 2021)*, pp. 142-150.
- Pápai, Vilma (2004). “Explicitation - A Universal of Translated Text?”. In Mauranen, Anna; Kujamäki Pekka (a cura di), *Translation universals: Do they exist?*. Amsterdam: John Benjamins, pp. 143-164.
- Rybicki, Jan (2006). “Burrowing into Translation: Character Idiolects in Henryk Sienkiewicz’s Trilogy and its two English Translations”. *Literary and Linguistic Computing*, 21 (1), pp. 91-103.
- Shibamoto-Smith, Janet S. (2004). “Language and gender in the (hetero) romance: ‘Reading’ the ideal hero/ine through lovers’ dialogue in Japanese romance fiction”. *Japanese Language, Gender, and Ideology: Cultural Models and Real People*, pp. 113-130
- Tanomura, Tadaharu (2002). *Gendai nihongo no bupō 1 - [noda] no imi to yōhō*. Ōsaka: Izumi Shoin.
- Wakabayashi, Judy (2009). “Translational Japanese: a Transformative Strangeness within”. *Portal: Journal of Multidisciplinary International Studies*, 6 (1), pp. 1-20.
- Wang, Qing; Li, Defeng (2012). “Looking for Translator’s Fingerprints: A Corpus-Based Study on Chinese translations of *Ulysses*”. *Literary and Linguistic Computing*, 27, pp. 81-93.
- Yasuoka, Kōichi (2020). “Ketaisokaiseikibu no tsukekae ni kindai nihongo (kyūji kyūkana) no kakariuke kaiseki”. *Jōhō shori gakkai kenkyū hōkoku: kenkyū hōkoku jinbungaku to konpyūta (CH)*, 3, pp. 1-8.
- Yonekawa, Ryōfu (1991). “Itaria bungaku yawa / Danuntsuio o meguru 9-shō -3- [shi no shōri]kakuyō honyaku Hon'yaku no henshen”. *Gendaishi teichō*, 34 (5), pp. 220-223.