

Nome: *Computational Literary Criticism*. Nascita: 1987.  
Genitore: John F. Burrows (Ostetrica: Narratologia)

Pietro Mazzarisi

Università degli studi di Trieste

([pietro.mazzarisi@units.it](mailto:pietro.mazzarisi@units.it))

#### Abstract

L'articolo riconsidera le origini della *Computational Literary Criticism* (CLC) proponendone l'anticipazione dell'atto fondativo al 1987, anno in cui John F. Burrows pubblica *Word-Patterns and Story-Shapes*. Benché tradizionalmente se ne attribuisca il primato al *Distant Reading* di Franco Moretti (2000), il contributo di Burrows è la prima consapevole proposta di applicazione di metodi computazionali a fini di critica e teoria letterarie. Inoltre, l'approccio narratologico adottato e basato sulla distinzione tra diegesi, mimesi e discorso indiretto libero, evidenzia una precoce cornice teorica, trascurata successivamente. L'articolo ricostruisce contesto, apporto di Burrows e ragioni della mancata ricezione, avanzando una rilettura della genealogia della CLC alla luce di questo pionieristico contributo.

---

#### Parole chiave

CLC, Distant Reading, Burrows, 1987, Moretti

---

#### DOI

<https://doi.org/10.58015/2036-2293/797>

---

#### Diritto d'autore

Questo lavoro è fornito con la licenza *Creative Commons Attribuzione - Non commerciale - Condividi allo stesso modo 4.0 Internazionale*: <https://creativecommons.org/licenses/by-nc-sa/4.0/>. Gli autori mantengono il diritto d'autore sui propri articoli e materiali supplementari e mantengono il diritto di pubblicazione senza restrizioni.

---

## 1. Introduzione

Lo studio digitale della letteratura con fini di critica e teoria letterarie, inevitabilmente, individua le proprie origini nel *Distant Reading* di Franco Moretti. La ricerca qui presentata mira a rivalutare e far emergere, invece, l'indiscutibile primato e accortezza teorico-metodologica di John Frederick Burrows, uno dei pionieri del *Literary Computing*, successivamente fluito sotto l'etichetta più "iperonimica" di *Digital Humanities* (DH)<sup>1</sup>. Grazie alla creazione di una opposizione concettuale con il *Close Reading* (elaborato dalla scuola formalista americana del *New Criticism* degli anni '30 del Novecento)<sup>2</sup> e grazie all'efficace sintesi offerta da Moretti per lo studio computativo con fini teorico-letterari, la definizione di *Distant Reading* data a questa finalità della *Computational Literary Criticism*<sup>3</sup> nel 2000<sup>4</sup> rimane, e rimarrà per sempre, indicativa della nuova rivisitazione metodologica degli studi di teoria letteraria, benché la prima paternità vada attribuita a Burrows, sottraendo così al silenzio un precedente che, per coerenza metodologica e ambizione teorica, può essere considerato fondativo.

L'obiettivo di questo articolo è dunque ricontestualizzare nascita e prima proposta della CLC con fini teorico-letterari al 1987. In questo anno, mentre era a capo del dipartimento di inglese dell'università di Newcastle (Australia), Burrows pubblica la monografia *Computation into Criticism: A Study of Jane Austen's Novels and an Experiment in Method*<sup>5</sup>, un lavoro che ha attirato critiche ed elogi, al punto che lo stesso Moretti e studi successivi di DR ne hanno più volte lodati i meriti, ma sempre esprimendo gli elogi, appunto, alla sola monografia. Il testo che il presente articolo vuole mettere risalto è però un altro, sempre dello stesso 1987, ma apparso (successivamente) nelle pagine della rivista *Literary and Linguistic Computing* dove Burrows pubblica il saggio *Word-Patterns and Story-Shapes: The Statistical Analysis of Narrative Style*<sup>6</sup>, da lui stesso definito metodologicamente più avanzato rispetto alla monografia poco prima data alle stampe (cfr. paragrafo 3). Tale saggio, nonostante la metodologia e la cornice teorica siano state elaborate in prospettiva di attribuzione autoriale, va di diritto collocato all'origine della CLC con fini teorico-letterari, stante l'impiego di un approccio narratologico e l'importanza e la versatilità in questa prospettiva, peraltro subito colta e ravvisata dallo stesso Burrows.

Con questo obiettivo, nei paragrafi che seguono, si presenta uno stato dell'arte in materia, una disamina dell'apporto di Burrows dato nel 1987, le lodi che, da Moretti in poi, sono state riservate alla monografia dello studioso australiano (e le, apparenti, assenze di consapevolezza del suddetto saggio) e l'allora, come ora, insoluta questione riguardante la trattazione del *discorso indiretto libero*<sup>7</sup> seguite, infine, dalle conclusioni dell'articolo.

---

<sup>1</sup> D'ora in poi in acronimo DH.

<sup>2</sup> Il termine appare per la prima volta in Ivor Armstrong Richards, *Practical Criticism. A Study of Literature Judgement*, London, Kegan Paul, Trench, Trubner & Co. LTD, 1929.

<sup>3</sup> D'ora in poi in acronimo, rispettivamente, DR e CLC. Altre nominazioni sono *Computational Literary Studies* (CLS), *Digital Literary Studies* (DLS) e, più recentemente, *Literary Mathematics*.

<sup>4</sup> Il termine appare per la prima volta in questo passaggio: «[...] we know how to read texts, now let's learn how not to read them. Distant reading [...] is a condition of knowledge: it allows you to focus on units that are much smaller or much larger than the text: devices, themes, tropes — or genres and systems. And if, between the very small and the very large, the text itself disappears, well, it is one of those cases when one can justifiably say, Less is more», Franco Moretti, *Conjectures on World Literature*, «New Left Review», 1, 2000, <<https://newleftreview.org/issues/iii1/articles/franco-moretti-conjectures-on-world-literature>> (Consultato: 10 novembre 2025). Nello stesso anno se ne ipotizza l'impiego anche in prospettiva di conoscenza archivistica in Franco Moretti, *The Slaughterhouse of Literature*, «Modern Language Quarterly», 61, 1, 2000, pp. 207–227.

<sup>5</sup> John F. Burrows, *Computation into Criticism: A Study of Jane Austen's Novels and an Experiment in Method*, Oxford, Clarendon Press, 1987.

<sup>6</sup> John F. Burrows, *Word-Patterns and Story-Shapes: the Statistical Analysis of Narrative Style*, «Literary and Linguistic Computing», 2, 2, 1987, pp. 61–70.

<sup>7</sup> D'ora in poi in acronimo DIL.

## 2. La Computational Literary Criticism pre e post Moretti

L'avvio di studi in direzione di ciò che comunemente oggi si definisce DR viene generalmente fissato al diciannovesimo secolo e spesso associato ad approcci sociologici per via del ricorso a metodologie quantitative basate sui dati<sup>8</sup>. Sul finire dell'Ottocento, attornati dal clima storico-scientifico del Positivismo, si individuano in un articolo di Mendenhall e in una monografia di Sherman i primi due esperimenti in direzione di interpretazioni numeriche e quantitative dei testi letterari<sup>9</sup>.

Nella prima metà del Novecento, gli studi formali sulla letteratura trovano una più strutturata elaborazione grazie a quello che comunemente viene definito il formalismo russo («scuola morfologica» secondo i propri appartenenti), un indirizzo di studi costituito da un gruppo di teorici che si riuniscono nell'OPAJAZ di Pietrogrado (successivamente San Pietroburgo) e nel Circolo linguistico di Mosca<sup>10</sup>. Bisogna comunque superare il secondo conflitto mondiale e arrivare agli anni Sessanta per individuare un percorso di evoluzione nella direzione del contemporaneo DR, alimentato dalla nascita dei primi calcolatori e coerentemente spinto da istanze sociologiche. Nel 1961 Raymond Williams presenta un lavoro interdisciplinare tra sociologia e critica letteraria con il libro *The Long Revolution*<sup>11</sup> in cui propone una indagine che, partendo da aspetti socio-economici ed estetici, intende dimostrare come alla rivoluzione industriale e democratica si sia accostata una rivoluzione culturale, attestata dalla crescita della stampa popolare e della platea di lettori e lettrici esperita nei paesi occidentali industrializzati e democratici. Una crescita che sembra anticipare lo *slaughterhouse* di Moretti quando afferma che il romanzo (dell'Ottocento nel caso di Williams) è una terra, nella sua vastità, fundamentalmente sconosciuta, a causa della dispersione dei testi negli archivi (o altrove) e dei limiti insiti nella durata della vita umana di poterli leggere tutti. Nella decade successiva, alle spinte promotrici se ne alternano di conservatrici, per esempio Lotman rivendica l'irriducibilità del testo artistico-letterario al solo livello di analisi linguistica<sup>12</sup> e Wittig e Beatie, più duramente, ribadiscono l'estraneità della critica e della teoria letterarie con concetti come la statistica e con metodi come quelli informatizzati: la prima richiamandosi agli assunti del post-strutturalismo e alla caduta di schemi categorizzanti<sup>13</sup>; il secondo<sup>14</sup> rifacendosi al noto lavoro di Snow sulle due culture<sup>15</sup> che tutt'oggi, in alcuni ambienti, rimane antonomastico di una rigida linea di demarcazione tra le scienze dure e gli studi umanistici. Nel 1984 compare la monografia di Radway in cui la critica prende in analisi venti romanzi rosa<sup>16</sup> impiegando una metodologia che segna un ulteriore passo in avanti in direzione sociologica (qui in combinazione con i contributi marxisti e l'antropologia) e per questo viene vista, da alcuni/e studiosi/e di DR, come all'origine del nuovo corso<sup>17</sup>.

<sup>8</sup> Per uno stato dell'arte più recente della CLC si veda: Chris Beausang, *A Brief History of the Theory and Practice of Computational Literary Criticism* (1963-2020), «Magazén», 1, 2, 2020, pp. 181-201. <<http://doi.org/10.30687/mag/2724-3923/2020/02/002>> (Consultato: 10 novembre 2025).

<sup>9</sup> Thomas C. Mendenhall, *The Characteristic Curves of Composition*, «Science», 9, 214, 1887, pp. 237-246; Lucius A. Sherman, *Analytics of Literature: A Manual for the Objective Study of English Prose and Poetry*, Boston, Ginn, 1893.

<sup>10</sup> Tzvetan Todorov, a cura di, *I formalisti russi. Teoria della letteratura e metodo critico*, Torino, Einaudi, 1968.

<sup>11</sup> Raymond Williams, *The Long Revolution*, London, Chatto & Windus, 1961.

<sup>12</sup> Jurij Mikhajlovič Lotman, *Struktura khudozhestvennogo teksta*, Moskva, Iskusstvo, 1970.

<sup>13</sup> Susan Wittig, *The Computer and the Concept of Text*, «Computers and the Humanities», 11, 4, 1977, pp. 211-215.

<sup>14</sup> Bruce A. Beatie, *Measurement and the Study of Literature*, «Computers and the Humanities», 13, 3, 1979, pp. 185-194.

<sup>15</sup> Charles Peirce Snow, *The Two Cultures and the Scientific Revolution*, London, Cambridge University Press, 1959.

<sup>16</sup> Janice Radway, *Reading the Romance: Women, Patriarchy, and Popular Literature*, Chapel Hill, University of North Carolina Press, 1984.

<sup>17</sup> Ted Underwood, *A Genealogy of Distant reading*, «Digital Humanities Quarterly», 11, 2, 2017, <<http://www.digitalhumanities.org/dhq/vol/11/2/000317/000317.html>> (Consultato: 10 novembre 2025).

Con interviste e questionari, lo studio di Radway, che ha anche il pregio di essere una pietra miliare all'interno degli studi femministi, offre dimensioni quantitative lontane dalle dimensioni di aggregati odierne, eppure utili a inquadrare domande complesse su come vengono rese le figure femminili (siano esse protagoniste o no) e le controparti maschili e quali caratteristiche inciderebbero sulla popolarità dei testi letterari principalmente orientati ai pubblici femminili (com'è nel caso di quelli rosa). L'anno successivo, nel 1985, la monografia di Cesare Segre, *Avviamento all'analisi del testo letterario*<sup>18</sup>, esprime forti dubbi sulla possibilità che l'analisi stilistica possa utilmente avvantaggiare la profonda interpretazione tematica dei testi artistico-letterari.

Tralasciando qui l'apporto di Burrows a cui è dedicato il paragrafo 3, agli albori del nuovo millennio avviene la nascita del DR di Franco Moretti (cfr. paragrafo 1), impiegato dall'autore per ipotizzare la possibilità di non leggere e allo stesso tempo accedere alle caratteristiche di tutta la produzione letteraria finita nel dimenticatoio degli archivi<sup>19</sup>. La stessa definizione del termine prevedeva già una dimensione teorica da rintracciare nella contrapposizione creata con il *Close Reading*<sup>20</sup>, ma l'affascinante idea (per le potenzialità e conseguenze conoscitive) di una scansione integrale delle incognite archiviali risulta ancorché impervia per le effettive possibilità di realizzazione. L'iniziale obiettivo di confronto con il *great unread* viene dunque avviato in maniere differenti: Rosen, per esempio, desume dalle difficoltà, costi e tempi di digitalizzazione di tutti gli archivi il fatto che la totalità archiviale non sia di per sé un obiettivo da perseguire in termini assoluti<sup>21</sup>; più pragmaticamente, Bode dismette l'idea dell'impresa sottolineando che non sussiste coincidenza tra canoni e corpora e, dunque, la non completezza di un corpus non ne implica automaticamente la scarsa rilevanza rispetto al periodo storico preso in analisi<sup>22</sup>. L'affacciarsi del DR di Moretti nella già travagliata dialettica sul ruolo dell'informatica negli studi letterari acuisce le contrapposizioni fra gli schieramenti: McGann nel 2001 rispolvera l'immagine sovietica di un radioso futuro con tante promesse e poca concretezza<sup>23</sup>; McCarty sottolinea nella mancanza di teoria il punto più debole della CLC<sup>24</sup>; Fish traccia una netta linea tra i due mondi, dove il lavoro critico tradizionale è orientato verso il vero, il rilevante, il serio e l'interpretazione mentre la CLC sarebbe orientata verso il falso, il rumore e il gioco<sup>25</sup>; demarcazione ribadita nel 2012 anche da Marche con un titolo molto più eloquente (*Literature is not Data: Against Digital Humanities*)<sup>26</sup>; nel 2013, Gooding scorge nella CLC la tendenza a generalizzare e pertanto

<sup>18</sup> Cesare Segre, *Avviamento all'analisi del testo letterario*, Torino, Einaudi, 1985.

<sup>19</sup> Moretti, *The Slaughterhouse of Literature*, cit.

<sup>20</sup> Franco Moretti, *Distant reading*, London, Verso, 2013.

<sup>21</sup> Jeremy Rosen, *Combining Close and Distant, or the Utility of Genre Analysis: A Response to Matthew Wilkens's 'Contemporary Fiction by the Numbers'*, «Post» 45, 3, 2011, <<https://post45.org/2011/12/combining-close-and-distant-or-the-utility-of-genre-analysis-a-response-to-matthew-wilkenss-contemporary-fiction-by-the-numbers/>> (Consultato: 10 novembre 2025).

<sup>22</sup> Katherine Bode, *Reading by Numbers: Calibrating the Literary Field*, London, Anthem, 2014; Katherine Bode, *The Equivalence of 'Close' and 'Distant' Reading: Or, Towards a New Object for Data-Rich Literary History*, 2017, «Modern Language Quarterly», 78, 1, 2017, pp. 77–106.

<sup>23</sup> Poiché l'uso della tecnologia digitale con i suoi strumenti non ha ancora esteso le modalità in cui le opere artistico-letterarie vengono interpretate dai critici, Jerome McGann, *Radiant Textuality: Literature after the World Wide Web*, New York, Palgrave Macmillan, 2001.

<sup>24</sup> Peralto accompagnata da generale impreparazione, motivi per cui l'informatica letteraria non sarebbe in grado di trasformare in prove gli indizi forniti, anche a causa della mole di dati, ingestibile da chi appartiene alla generazione X e precedenti, Willard McCarty, *Literary enquiry and experimental method: What has happened? What might?*, in *Storia della scienza e linguistica computazionale: sconfinamenti possibili*, a cura di Liborio Dibattista, Franco Angeli, Milano, 2009, pp. 32–54.

<sup>25</sup> Stanley Fish, *Mind Your P's and B's: The Digital Humanities and Interpretation*, «Opinionator, New York Times», 23 gennaio 2012, <<http://opinionator.blogs.nytimes.com/2012/01/23/mind-your-ps-and-bs-the-digital-humanities-and-interpretation>> (Consultato: 10 novembre 2025).

<sup>26</sup> Stephen Marche, *Literature is not Data: Against Digital Humanities*, «Los Angeles Review of Books», 28 ottobre 2012, <<http://lareviewofbooks.org/essay/literature-is-not-data-against-digital-humanities#>> (Consultato: 10 novembre 2025).

ne diagnostica una cronica sofferenza di assenza di riflessività<sup>27</sup>; sempre nel 2013 Eysers sostiene che l'analisi letteraria non si basa né si può basare su dati oggettivi né su dati empirici di carattere quantitativo<sup>28</sup>; infine, nel 2014 Golumbia pronuncia il *requiem* disciplinare a causa di invasione barbarica<sup>29</sup> e due anni dopo, insieme ad Allington e Brouillette sul *Los Angeles Review of Books*, denuncia lo spargimento di sale capitalistico sul desolato terreno della critica letteraria classica<sup>30</sup>.

Tutte queste reazioni (ri)viste oggi, mentre i corsi di DH segnano distinzione e successo nell'offerta formativa degli atenei, sembrano più che altro rientrare nella millenaria contrapposizione genitori/prole. Fortunatamente, avranno forse complicato ma di certo non hanno fermato le ricerche, le quali intanto non hanno perso di vista l'obiettivo "sociologico" del DR: con Heuser e Le-Khac dimostrando come nei romanzi britannici dell'Ottocento avvenga un cambio nelle descrizioni, che da più astratte vanno sempre più concretizzandosi<sup>31</sup>; con Cordell e l'analisi delle ristampe e circolazioni di quotidiani e periodici del periodo che precede la guerra civile americana per ottenere una diversa prospettiva sulla produzione letteraria pre-bellica<sup>32</sup>; con Wilkens e i nessi tra letteratura, spazio, tempo e geografia nelle opere finzionali nel periodo bellico civile americano<sup>33</sup>; con Klein e lo studio su come, nel post-guerra civile, negli archivi vengano silenziate le voci della schiavitù afro-americana<sup>34</sup>; con Jockers e Kirilloff e l'indagine sulle questioni di genere e l'agentività dei personaggi nel romanzo dell'Ottocento<sup>35</sup>; con la modellazione degli argomenti (*Topic Modeling*), termine ombrello per algoritmi di diversa natura, tra cui il più conosciuto, il *Latent Dirichlet Allocation* che rimanda alla teoria bayesiana della probabilità<sup>36</sup> e che applicato su un testo o su un corpus si basa su tecniche di *Text Mining* non supervisionate<sup>37</sup> per trovare *token* lessicali caratterizzati da alte frequenze di co-occorrenze da cui emergono *pattern* intratestuali e intertestuali. Molti studi che segnano e compongono il post-Moretti del DR si sono interrogati, hanno fatto ricorso o hanno fatto didattica con la modellazione degli argomenti: Jockers, uno dei co-fondatori dello Stanford Literary Lab, vi ha fatto ricorso e l'ha esposta in *Macroanalysis: Digital Methods and Literary History*<sup>38</sup>, insegnata in *Text Analysis with R* for

<sup>27</sup> Paul Gooding, *Mass Digitization and the Garbage Dump: The Conflicting Needs of Quantitative and Qualitative Methods*, «Literary and Linguistic Computing», 28, 3, 2013, pp. 425–431.

<sup>28</sup> Tom Eysers, *The Perils of the "Digital Humanities": New Positivism and the Fate of Literary Theory*, «Postmodern Culture», 23, 2, 2013, <<http://www.pomoculture.org/2015/07/08/the-perils-of-the-digital-humanities-new-positivism-and-the-fate-of-literary-theory/>> (Consultato: 10 novembre 2025).

<sup>29</sup> David Golumbia, *Death of a Discipline*, «Differences», 25, 1, 2014, pp. 156–176.

<sup>30</sup> Daniel Allington, David Golumbia e Sarah Brouillette, *Neoliberal Tools (and Archives): A Political History of Digital Humanities*, «Los Angeles Review of Books», 01 maggio 2016, <<https://lareviewofbooks.org/article/neoliberal-tools-archives-political-history-digital-humanities/>> (Consultato: 10 novembre 2025).

<sup>31</sup> Ryan Heuser e Long Le-Khac, *A Quantitative Literary History of 2,958 Nineteenth-Century British Novels: The Semantic Cohort Method*, «Stanford Literary Lab Pamphlets», 4, May, 2012, <<https://litlab.stanford.edu/LiteraryLabPamphlet4.pdf>> (Consultato: 10 novembre 2025).

<sup>32</sup> Ryan Cordell, *Reprinting, Circulation, and the Network Author in Antebellum Newspapers*, «American Literary History», 27, 3, 2015, pp. 417–445.

<sup>33</sup> Matthew Wilkens, *The Geographic Imagination of Civil War-Era American Fiction*, «American Literary History», 25, 4, 2013, pp. 803–840.

<sup>34</sup> Lauren F. Klein, *The Image of Absence: Archival Silence, Data Visualization, and James Hemings*, «American Literature», 85, 4, 2013, pp. 661–688.

<sup>35</sup> Matthew L. Jockers e Gabi Kirilloff, *Gender and Character Agency in the 19th Century Novel*, «Cultural Analytics», 2016, <<https://doi.org/10.22148/16.010>> (Consultato: 10 novembre 2025).

<sup>36</sup> David M. Blei, Andrew Y. Ng e Michael I. Jordan, *Latent Dirichlet Allocation*, «Journal of Machine Learning Research», 3, 2003, pp. 993–1022.

<sup>37</sup> David M. Blei, *Probabilistic Topic Models*, «Communications of the ACM», 55, 4, 2012, pp. 77–84, <<https://doi.org/10.1145/2133806.2133826>> (Consultato: 10 novembre 2025).

<sup>38</sup> Matthew L. Jockers, *Macroanalysis: Digital Methods and Literary History*, Urbana, University of Illinois Press, 2013.

*Students of Literature*<sup>39</sup> e impiegata, con Archer, per decifrare un presunto codice sottostante la popolarità editoriale dei best seller<sup>40</sup>; Piper l'ha usata sofisticandola<sup>41</sup>; Underwood ne ha fatto oggetto di riflessioni critiche<sup>42</sup> e impieghi sperimentali<sup>43</sup>.

Infine, sul versante di un ricomponimento teorico con la critica classica, Piper<sup>44</sup> ha sottolineato il carattere circolare e dunque ermeneutico dei passaggi da micro-analisi a macro-analisi e di nuovo a micro analisi come base di un *mixed method* e, con l'obiettivo di individuare procedure condivise con la critica classica, Gavin ha intravisto nel metodo di Empson di disambiguare la tipica condensazione semantica nei testi poetici un possibile collegamento e campo comune con la semantica vettoriale e nei metadati dei testi letterari (intesi in senso lato) un possibile collegamento teorico con la semantica distribuzionale<sup>45</sup>.

### 3. Origine, primato e approccio (narratologico) di Burrows

La produzione scientifica di Burrows è principalmente legata agli studi di attribuzione autoriale. In questo campo<sup>46</sup>, i primi esperimenti che si possano avvicinare al concetto moderno di stilometria andrebbero collocati alla fine del Settecento, quando Frederick Mosteller e David Wallace basarono l'attribuzione dell'autorialità ricorrendo alla somiglianza d'impiego delle parole funzionali da parte di autori e autrici reali e pseudonimi (articoli, congiunzioni, preposizioni ecc.<sup>47</sup>; su questo primato, però, i pareri non sono unanimi, essendoci anche chi posticipa i primi tentativi di lettura a distanza all'Ottocento<sup>48</sup>).

Volgendo anche qui agli anni Sessanta del Novecento, un buon punto di osservazione viene offerto nelle pagine della rivista *Computers and the Humanities*, dove tra il 1966 e il 1969 si consuma una delle prime polemiche, con posizioni interessanti poiché spesso si riproporranno in seguito. Louis Tonko Milic in un articolo del 1966<sup>49</sup> avverte la necessità e l'utilità di una sinergia per l'attività critico-letteraria tra l'approccio intuitivo che spesso ha distinto tale attività - coadiuvato certo da profonda conoscenza e competenze ma la cui rinuncia, per inciso, è vista sfavorevolmente da parte di quello schieramento che può definirsi più tradizionalista - e l'enumerazione, intanto messa in risalto dalla diffusione degli elaboratori. L'anno successivo, nella monografia pubblicata sullo stile di Jonathan Swift<sup>50</sup>, l'autore croato naturalizzato statunitense lamenta che i termini impiegati nell'attività critica classica risultano scaturire spesso da impressioni e pertanto ribadisce l'auspicio di ancorare tale attività a dati quanto più possibile oggettivi. Con questo scopo, Milic individua nella sintassi il livello linguistico più adeguato e promettente per avviare degli studi quantitativi indirizzati a sondare i pattern che,

---

<sup>39</sup> Matthew L. Jockers, *Text Analysis with R for Students of Literature*, Berlin, Springer, 2014.

<sup>40</sup> Matthew L. Jockers e Jodie Archer, *The Bestseller Code. Anatomy of the Blockbuster Novel*, New York, St. Martin's Press, 2016.

<sup>41</sup> Andrew Piper, *Enumerations. Data and Literary Studies*, Chicago, University of Chicago Press, 2018.

<sup>42</sup> Ted Underwood, *What kinds of 'topics' does topic modeling actually produce?*, «The Stone and the Shell», 01 aprile 2012, <<http://tedunderwood.com/2012/04/01/what-kinds-of-topics-does-topic-modeling-actually-produce/>> (Consultato: 10 novembre 2025).

<sup>43</sup> Ted Underwood, *Distant Horizons. Digital Evidence and Literary Change*, Chicago, University of Chicago Press, 2019.

<sup>44</sup> Andrew Piper, *Novel Devotions: Conversional Reading, Computational Modeling, and the Modern Novel*, «New Literary History», 46, 1, 2015, pp. 63–98.

<sup>45</sup> Michael Gavin, *Vector Semantics, William Empson, and the Study of Ambiguity*, «Critical Inquiry», 44, 4, 2018, pp. 641–673; Michael Gavin, *Literary Mathematics. Quantitative Theory for Textual Studies*, Stanford, Stanford University Press, 2022.

<sup>46</sup> Anche in questo caso per uno stato dell'arte più recente si veda: Chris Beausang, *A Brief History of the Theory and Practice of Computational Literary Criticism (1963-2020)*, cit.

<sup>47</sup> David I. Holmes, *The Evolution of Stylometry in Humanities Scholarship*, «Literary and Linguistic Computing», 13, 3, 1998, pp. 111–117.

<sup>48</sup> Jack Grieve, *Quantitative Authorship Attribution: An Evaluation of Techniques*, «Literary and Linguistic Computing», 22, 3, 2007, pp. 251–270.

<sup>49</sup> Louis T. Milic, *The Next Step*, «Computers and the Humanities», 1, 1, 1966, pp. 3–6.

<sup>50</sup> Louis T. Milic, *A Quantitative Approach to the Style of Jonathan Swift*, 1967, Berlin, Mouton.

diversamente, andrebbero persi laddove si ricorresse agli approcci soggettivi e qualitativi. Va sottolineato che le premesse armonizzanti di Milic trovano puntuale corrispondenza con l'approccio proposto, poiché il suo studio sintattico sulle opere di Swift prevede una successiva fase di lettura ravvicinata e può, dunque, essere visto come un precursore di quelli che successivamente verranno chiamati Mixed Methods<sup>51</sup>, ovvero metodi che cercano di trovare un punto di incontro tra il tradizionale e l'informatico. Un paio di anni dopo, nel 1969, anche la reazione critica di Mesthene anticipa, in parte, una delle debolezze che oggi vengono imputate, più in generale, agli approcci computazionali: mancanza di obiettività e imparzialità determinata dal fatto che questi approcci non sono scevri di pregiudizi spesso imputabili agli studiosi e alle studiose che li hanno elaborati<sup>52</sup>. Dall'inizio degli anni Settanta, affiorano primi lavori e ricerche su aspetti che dopo, sempre più, diverranno centrali in questo tipo di indagini e verranno trattati con sempre più maggiore precisione: *Part-Of-Speech* (POS), *Cluster Analysis*, *Principal Component Analysis* (PCA) e *Function Words*. La rivista *Computers and the Humanities* continua a essere il punto di incontro (e talvolta scontro). Nelle sue pagine vengono pubblicati: nel 1971 l'articolo *Style, Precept, Personality* di Cluett, uno dei primi esempi di attenzione posta sulle parti del discorso<sup>53</sup>; nel 1973 l'articolo *On the Distinction Between a Novel and a Romance* di Brainerd, uno dei primi lavori sulle analisi dei cluster<sup>54</sup>; nello stesso anno, l'articolo *An Application of Principal Component Analysis to the Works of Molière* di Sainte-Marie, Robillard & Bratley è una delle prime applicazioni dell'analisi dei componenti<sup>55</sup>; nel 1975 l'articolo *The Use of Function Word Frequencies as Indicators of Style* di Damerau<sup>56</sup> non può essere indicato come uno dei primi esempi di indagine sulle parole funzionali. Damerau infatti prosegue la tradizione di studi sulle parole funzionali che si è vista sorgere già alla fine del Settecento, ma conferisce a essa una cornice teoricamente basata che però risente di assunti behavioristi degli anni precedenti. La cornice elaborata da William J. Paisley circa un decennio prima<sup>57</sup>, ipotizza la presenza di tracce autoriali (dalla scrittura alla pittura) nelle abitudini di codifica minori (*minor encoding habits*) le quali, secondo l'autore, «lie at an opposite pole from complex constructions on the continua of deliberation and self-consciousness»<sup>58</sup>. Ne esclude l'inconscio come origine fissandola piuttosto nell'età scolare, o comunque in periodo formativo, grazie ai rinforzi che per esempio uno scrittore o una scrittrice ha ricevuto da giovane sull'impiego delle preposizioni<sup>59</sup>. Inserendo metodologicamente test quali il *chi-square*, il Fisher e il Mann-Whitney all'interno dei suddetti assunti teorici tratti da Paisley, Damerau inquadra le parole funzionali come le migliori prove per attribuire l'autorialità.

<sup>51</sup> Un esempio è il concetto di *Scalable reading*, si veda Martin Mueller, *Shakespeare His Contemporaries: collaborative curation and exploration of Early Modern drama in a digital environment*, «Digital Humanities Quarterly», 8, 3, 2014, §§ 4, 31, 32, <<http://www.digitalhumanities.org/dhq/vol/8/3/000183/000183.html>> (Consultato: 10 novembre 2025).

<sup>52</sup> Emmanuel G. Mesthene, *Technology and Humanistic Values*, «Computers and the Humanities», 4, 1, 1969, pp. 1–10. È questo un aspetto etico oggi dibattuto nella modellazione, poiché essa ripropone i percorsi a cui è stata allenata e dunque se viene prima allenata e successivamente impiegata da una banca che nella cessione dei mutui ha uno storico di respingimenti di domande da parte di soggetti appartenenti a minoranze etniche o a etnie emarginate in una data società, il modello riproporrà lo stesso schema. Si veda per esempio Michelle Seng Ah Lee e Luciano Floridi, *Algorithmic Fairness in Mortgage Lending: from Absolute Conditions to Relational Trade-offs*, «Minds and Machines», 31, 2021, pp. 165–191. Si presume che questo aspetto verrà superato con il passaggio dall'IA generativa all'IA generale.

<sup>53</sup> Robert Cluett, *Style, Precept, Personality*, «Computers and the Humanities», 5, 5, 1971, pp. 257–277.

<sup>54</sup> Barron Brainerd, *On the Distinction Between a Novel and a Romance*, «Computers and the Humanities», 7, 5, 1973, pp. 259–270.

<sup>55</sup> Paule Sainte-Marie, Pierre Robillard e Paul Bratley, *An Application of Principal Component Analysis to the Works of Molière*, «Computers and the Humanities», 7, 3, 1973, pp. 131–137.

<sup>56</sup> Fred J. Damerau, *The Use of Function Word Frequencies as Indicators of Style*, «Computers and the Humanities», 9, 6, 1975, pp. 271–280.

<sup>57</sup> William J. Paisley, *Identifying the unknown communicator in painting, literature and music: The Significance of Minor Encoding Habits*, «Journal of Communication», 14, 4, 1964, pp. 219–237.

<sup>58</sup> Ivi, p. 235–236.

<sup>59</sup> Ivi, p. 236.

In questo periodo, emergono critiche esterne ed interne. Sul fronte interno, all'assunto teorico behaviorista di Damerau di per sé abbastanza superato dalle prospettive cognitive quando pubblica il suo lavoro nel 1975, già nello stesso anno e poi nel 1979 il già citato Brainerd è tra i primi a notare la non adeguatezza della tipologia testuale degli scritti artistici come banchi di analisi per il *chi-square*<sup>60</sup>. Mentre sul fronte esterno, nel 1979 arriva la critica forte di Fish<sup>61</sup> articolata su due aspetti: la natura interpretativa dei modelli formali e la conseguente arbitrarietà tra la scelta di un modello formale al posto di un altro per la successiva attività interpretativa. A differenza di altre prese di posizione avverse alle pratiche di attribuzione autoriale con metodi statistici e informatizzati, le critiche di Fish hanno una chiara base ontologica (piuttosto che ideologica) e per questo motivo sono estendibili anche al campo del DR.

Infine, la successiva decade degli anni '80 è decisiva in ambito di attribuzione autoriale (e per le sue prospettive future di evoluzione) con la comparsa, appunto, dei primi lavori di John Burrows. Nel 1986 Burrows pubblica la ricerca *Modal Verbs and Moral Principles*<sup>62</sup> in cui indirizza la propria attenzione sulle percentuali d'impiego dei verbi modali nelle opere di Jane Austen. L'anno dopo, nel 1987, in *Word-Patterns and Story-Shapes: the Statistical Analysis of Narrative Style*<sup>63</sup> si concentra sull'analisi dei componenti (*Principal Component Analysis*, PCA) seguendo il solco tracciato, a partire dall'inizio degli anni '70, da Sainte-Marie, Robillard e Bratley<sup>64</sup> e focalizzandosi sugli elenchi delle parole funzionali, da cui deriverà la sua personale applicazione del metodo PCA, all'origine di quello che, all'inizio degli anni Duemila, diventerà il metodo Delta<sup>65</sup> (il metodo Delta rimane tutt'oggi centrale nel dibattito e nella disciplina dell'attribuzione autoriale, malgrado sia stato oggetto di critiche e numerose proposte di revisioni avanzate negli anni<sup>66</sup>).

Ma proprio il saggio del 1987 è un lavoro con una cornice teorica e conseguenze in direzione di una CLC con fini di teoria letteraria che non hanno incontrato ricezione né riconoscimento nella storia del DR. In questo saggio, l'autore propone una partizione dei testi letterari in diverse componenti: dialoghi (ovvero la mimesi), 'narrazioni pure' che nella dicitura di questo articolo si riferiscono alla diegesi senza il DIL e, appunto, il DIL. Questa partizione - impiegata appunto ai fini dell'attribuzione dell'autorialità - non è stata conservata all'interno dei lavori che sono seguiti da parte di Burrows o in altre ricerche sull'attribuzione dell'autorialità né è stata ereditata negli studi di DR e le ipotesi sul perché ciò non sia avvenuto sono presentate nel paragrafo 4. Intanto, in questa direzione va notato che l'invalso negli studi di DR sembra essere rimasto immutato fin dall'inizio a oggi: ovvero analizzare centinaia e migliaia di testi di prosa letteraria senza distinguere le componenti essenziali di cui la maggior parte si compone, la diegesi e la mimesi (si vedano, per esempio, i titoli dal 2000 in poi citati nel paragrafo 2).

<sup>60</sup> Barron Brainerd, *Statistical Analysis of Lexical Data Using Chi-squared and Related Distributions*, «Computers and the Humanities», 9, 4, 1975, pp. 161-178; Barron Brainerd, *Pronouns and Genre in Shakespeare's Drama*, «Computers and the Humanities», 13, 1, 1979, pp. 3-16.

<sup>61</sup> Stanley E. Fish, *What is Stylistics and Why Are They Saying Such Terrible Things about It? Part II*, «Boundary 2», 8, 1, 1979, pp. 129-146.

<sup>62</sup> John F. Burrows, *Modal Verbs and Moral Principles*, «Literary and Linguistic Computing», 1, 1, 1986, pp. 9-23.

<sup>63</sup> Burrows, *Word-Patterns and Story-Shapes: the Statistical Analysis of Narrative Style*, cit.

<sup>64</sup> Paule Sainte-Marie, Pierre Robillard e Paul Bratley, *An Application of Principal Component Analysis to the Works of Molière*, cit.

<sup>65</sup> John F. Burrows, 'Delta', «Literary and Linguistic Computing», 17, 3, 2002, pp. 267-287. È il primo metodo che permette di superare il 'gioco chiuso' all'interno dell'attribuzione dell'autorialità che fino ad allora ne aveva condizionato gli sviluppi (per "gioco chiuso" si intende la presenza di due o tre identità indiziate per una certa opera priva di autorialità, una condizione dettata dall'incapacità dei metodi precedenti di includere più identità autoriali).

<sup>66</sup> Parole funzionali e PCA nell'attribuzione dell'autorialità hanno continuato a rivelarsi efficaci negli anni Novanta del Novecento, cfr. David H. Craig, *Plural Pronouns in Roman Plays by Shakespeare and Jonson*, «Literary and Linguistic Computing», 6, 3, 1991, pp. 180-186; Emily K. Tse, Fiona J. Tweedie e Bernard D. Frischer, *Unravelling the Purple Thread: Function Word Variability and the Scriptorum Historiae Augustae*, «Literary and Linguistic Computing», 13, 3, 1998, pp. 141-149.

Burrows, invece, pone la riflessione su questi aspetti come base iniziale nel suo lavoro: «Prose fiction, arguably, is the most mixed of literary forms [...]. The language of its dialogue and that of its narrative usually differ from each other in some obvious and many less obvious ways»<sup>67</sup>. Nella monografia del 1987, l'autore si era occupato di due dimensioni: il linguaggio dei personaggi (i dialoghi, la mimesi) e il DIL. Nell'introduzione del saggio del 1987 dichiara che esso rappresenta un passo metodologicamente in avanti rispetto alla monografia dello stesso anno:

So far as they bear on the language of characterization, whether in dialogue or in the narrative rendering of the characters' ideas, they are discussed at some length in my recent book, *Computation into Criticism: A Study of Jane Austen's Novels and an Experiment in Method* (Oxford, Clarendon Press, 1987). The present paper carries the argument somewhat further into the territory of 'pure narrative' and tries to show that statistical analysis can cast new light on some important determinants of narrative style<sup>68</sup>.

Infatti, le analisi presentate nel saggio aggiungono quest'ulteriore dimensione da egli definita *pure narrative* e che coincide con la definizione platonica della diegesi<sup>69</sup> nonché coincide con quella che, negli studi narratologici condotti da Genette nel decennio precedente, era stata denominata *récit*, all'interno della distinzione teorico-terminologica rispetto a *discours*, *histoire* e *narration* delineata in *Figures III*:

Je propose, sans insister sur les raisons d'ailleurs évidentes du choix des termes, de nommer histoire le signifié ou contenu narratif (même si ce contenu se trouve être, en l'occurrence, d'une faible intensité dramatique ou teneur événementielle), récit proprement dit le signifiant, énoncé, discours ou texte narratif lui-même, et narration l'acte narratif producteur et, par extension, l'ensemble de la situation réelle ou fictive dans laquelle il prend place [...]. Récit et narration se passent de justification. Pour histoire, et malgré un inconvénient évident, j'invoquerai l'usage courant (on dit : " raconter une histoire "), et un usage technique, certes plus restreint, mais assez bien admis depuis que Tzvetan Todorov a proposé de distinguer le " récit comme discours " (sens 1) et le " récit comme histoire " (sens 2). J'emploierai encore dans le même sens le terme diégèse, qui nous vient des théoriciens du récit cinématographique<sup>70</sup>.

Burrows in questa ricerca computerizzata presenta dunque una evoluzione di quanto fino a lì concepito nella monografia, introducendo la divisione e analisi della prosa letteraria non più in due dimensioni ma in tre, che, nella sua terminologia, vengono rispettivamente definite: *pure narrative* (narrazione pura, diegesi priva di DIL), *dialogue* (dialoghi dei personaggi, mimesi) e *character narrative* (DIL). Gli obiettivi sono una comparazione intra-autoriale tra i sei romanzi di Jane Austen - *Sense and sensibility*, 1811; *Pride and Prejudice*, 1813; *Mansfield Park*, 1814; *Emma*, 1815; *Northanger Abbey*, 1818 (scritto nel 1803); *Persuasion*, 1818 - e una comparazione inter-autoriale tra questi e alcuni romanzi di Henry James (*The Awkward Age*, 1899), Edward Morgan Forster (*Howards End*, 1910), Georgette Heyer (*Frederica*, 1965) e l'anonima autorialità moderna che ha completato il frammento di *Sandition* della stessa Austen. Ricorrendo a questa partizione dei testi letterari in tre dimensioni e basando le prove sulle differenze di *frequency-profile* dei trenta tipi di parole più comuni, lo studio di Burrows mira a raggiungere l'obiettivo di distinguere lo stile narrativo in autori e autrici diverse, ottenendo, nel circoscritto caso

<sup>67</sup> Burrows, *Word-Patterns and Story-Shapes: the Statistical Analysis of Narrative Style*, cit., p. 61.

<sup>68</sup> Ivi, pp. 61-62.

<sup>69</sup> Platone, *La Repubblica*, a cura di Giuseppe Lozza, Milano, Mondadori, 1990, p. 199 *et passim* (Libro III, 392d *et passim*).

<sup>70</sup> Gérard Genette, *Figures III*, Paris, Seuil, 1972, p. 72. Nell'indice finale sull'impiego tecnico dei termini teorici discussi nel libro, la diegesi viene definita come l'universo spazio-temporale entro cui va a situarsi il racconto (*récit*) nel suo secondo senso, ivi, p. 280.

di Austen, una distinzione sull'alternarsi delle fasi nello stile narrativo di una singola autorialità<sup>71</sup>.

Il metodo elaborato da Burrows presenta però un punto debole, riconosciuto dallo stesso autore, per quel che riguarda il confine tra la diegesi (la *pure narrative*) e il DIL (la *character narrative*). Propone che la distinzione sia necessaria, ma se da una parte la distinzione tra la diegesi (*pure narrative*) e la mimesi (*dialogue*) è formalmente espressa nella maggior parte delle opere di prosa letteraria con il ricorso ai segni di interpunzione (solitamente ricorrendo alle virgolette alte o basse, ai trattini, ai due punti ecc.), dall'altra parte la distinzione tra la *pure narrative* e la *character narrative* (nella sua definizione) presuppone un intervento e un giudizio umano che, secondo gli assiomi statistici, inficiano e non garantiscono la validità dei dati raccolti. Infatti, affinché vi sia validità statistica, il campione su cui si conducono le indagini non deve essere stato preventivamente alterato o modificato secondo principi di parzialità umana, pena appunto la non più rappresentabilità del campione.

Su questo punto il saggio di Burrows presenta, dunque, una debolezza, benché l'autore si appelli a una presunta solidità in Austen nell'impiego del DIL che, però, egli stesso riconosce arbitraria poiché derivata dal proprio giudizio personale:

The idea that 'pure narrative' and 'character narrative' (as I call them) should be distinguished is widely recognized. The subtlety and exactitude with which Jane Austen manages such instruments as 'free indirect discourse' make a firm, if complex, ground for the exercise of critical judgement. Yet the actual distinctions on which the following evidence rests do rely on my judgement, contestable in every instance, of those moments in each text at which a transition from one mode of narrative to the other should be marked<sup>72</sup>.

Anche se largamente accettata, la distinzione tra diegesi e DIL rimanda (prepotentemente in ambito di CLC e/o DR) alla necessaria delineazione dei confini formali che, appunto, separano il DIL dalla diegesi. E in realtà, come si approfondisce nel paragrafo 4, sia la stessa distinzione formale sia la stessa assegnazione dell'etichetta di DIL a passaggi diegetici narrativi o a passaggi di discorso riportato rimane non univocamente definibile. Malgrado ciò, va riconosciuto che Burrows non aveva altri strumenti se non quelli manuali per condurre una azione del genere all'interno del corpus su cui stava lavorando, che anche per questo motivo il corpus era ridotto alla dimensione di circa una decina di testi e che, all'epoca, il potenziale di calcolo dei computer era ben diverso da quello attuale.

Inoltre, e qui è l'obiettivo del presente articolo, va riconosciuto che il saggio di Burrows è stato il primo contributo a intravedere le potenzialità di una critica e teoria letterarie con basi computazionali e a ribadire il carattere prettamente inferenziale che, di conseguenza, avrebbe adottato. Tredici anni prima di Moretti, Burrows aveva proposto l'idea di una analisi letteraria computazionale che andasse a toccare le tematiche più rilevanti per la critica e la teoria letterarie. Con largo anticipo, aveva riconosciuto come la sua metodologia travalicasse gli ambiti dell'attribuzione autoriale e, pertanto, ne aveva proposto l'impiego come una nuova modalità per informare il giudizio critico, teorico, letterario:

The method of analysis employed in this paper can, finally, be brought to bear not only on questions of authorship and chronological change but on questions nearer to the leading interests of literary critics and narrative theorists [...]. The appropriateness of this method of analysis for the purposes of literary criticism [...] rests on the extent to which there is a worthwhile correspondence between the evidence offered and the inferences drawn from it and on the extent to which those inferences seem persuasive [...]. That is not

---

<sup>71</sup> Burrows, *Word-Patterns and Story-Shapes: the Statistical Analysis of Narrative Style*, cit., pp. 67-68.

<sup>72</sup> Ivi, p. 62.

to propose a substitute for critical judgement but only to suggest a new way of informing it<sup>73</sup>.

Infine, in questo primato, Burrows aveva altresì abbozzato e adottato una cornice teorica di riferimento con forti radici classiche, la distinzione diegetica-mimetica, con autorevoli impieghi e approfondimenti novecenteschi, dallo *skaz* di Ejchenbaum alla polifonia di Bachtin e agli studi narratologici di Genette, e con prospettive di sviluppo di cui non c'è traccia negli studi di DR, benché l'aspetto su cui più spesso sono stati e sono tutt'oggi criticati consista proprio nell'assenza di quadri teorici da accompagnare agli studi di critica e teoria letterarie condotti con metodologie e tecniche computazionali.

#### 4. Ricezione e discorso indiretto libero

Perché a questa pionieristica ricerca di Burrows non sono seguiti studi nella direzione tracciata né espliciti riconoscimenti del primato? È possibile che abbiano giocato diversi fattori, per i quali si avanzano altrettante ipotesi. Una prima ipotesi può rimandare al fatto che, nello stesso 1987, prima del saggio è stata pubblicata la monografia che ha attirato critiche ed elogi, al punto da aver oscurato il saggio pubblicato successivamente nello stesso anno. Una seconda ipotesi può ricollegarsi all'efficacia e alla popolarità riscosse nell'attribuzione autoriale dal metodo Delta che hanno fatto apparire, anche allo stesso autore, la metodologia impiegata nel saggio del 1987 obsoleta condannandone la fortuna non solo in termini di attribuzione autoriale ma *in toto*. Un'ultima ipotesi si può formulare in termini di merito, ovvero vedendo nella debolezza metodologica del saggio del 1987 (per quanto riguarda la delineazione e separazione formale tra diegesi e DIL) la sua sfortunata ricezione.

Per quanto riguarda la prima ipotesi, Moretti almeno due volte ha elogiato il lavoro di Burrows, ma si è sempre riferito alla monografia del 1987 e non al saggio dello stesso anno; nelle bibliografie di Moretti, infatti, il saggio non è rinvenibile e gli stessi elogi hanno sempre trovato 'sedi secondarie'. La prima volta è avvenuto nel 2006 quando Moretti ha postato online una delle sue risposte nel dibattito che è seguito alla pubblicazione del libro *Graphs, Maps, Trees. Abstract Models for a Literary History*<sup>74</sup> (risposta poi confluita nella curatela di Goodwin e Holbo). In questo post online ha affermato che la monografia di Burrows è il miglior esempio di analisi quantitativa mai condotta e dalla quale, però, non sono sortiti effetti incoraggianti a proseguire in quel solco:

The field of, loosely speaking, science and literature is full of false starts. The best example of quantitative analysis ever done—Burrows' multivariate analysis of Austen's style—was published 20 years ago, and has had, if I'm not mistaken, hardly any effects. (No one has mentioned it in the course of this discussion either.) We're all working uphill, and I'm not sure it's going to change soon<sup>75</sup>.

Nel 2013, di nuovo Moretti dichiara se stesso e tutto il campo del DR in debito con Burrows, ma lo fa in una nota a piè di pagina (della sua famosa monografia *Distant Reading*): «The model here remains John Burrows' analysis of Austen's characters' styles in Computation into Criticism; that he did it twenty years ago, without the help of today's technology, puts us all to shame»<sup>76</sup>.

---

<sup>73</sup> Ivi, pp. 69-70.

<sup>74</sup> Franco Moretti, *Graphs, Maps, Trees. Abstract Models for a Literary History*, London, Verso, 2005.

<sup>75</sup> Franco Moretti, *Moretti Responds (II)*, in *Reading Graphs, Maps, Trees. Responses to Franco Moretti*, a cura di Jonathan Goodwin e John Holbo, Anderson, Parlor Press, 2011, p. 73. Originariamente postato il 15 gennaio 2006 su un sito che viene riportato nella risposta pubblicata in questa curatela, ma che non sembra più essere attivo: <[http://www.thevalve.org/go/valve/article/moretti\\_responds\\_ii/](http://www.thevalve.org/go/valve/article/moretti_responds_ii/)> (Consultato: 10 novembre 2025).

<sup>76</sup> Moretti, *Distant Reading*, cit., p. 207.

Nello stesso 2013, anche Jockers riconosce i traguardi fondamentali di Burrows spendendo parole di elogio per la monografia del 1987. Neanche Jockers però sembra essere a conoscenza del saggio dello stesso anno, neanche nelle sue bibliografie ve n'è traccia:

Arguments like those made by Burrows have been, and perhaps remain, underappreciated in contemporary literary discourse precisely because they are, or appear to be, definitive statements. As "findings," not "interpretations," they have about them a deceptive simplicity, a simplicity or finality that appears to render them "uninteresting" to scholars conditioned to reject the idea of a closed argument<sup>77</sup>.

*Word-Patterns and Story-Shapes: the Statistical Analysis of Narrative Style*, per quanto fin'ora è stato trovato, non sembra comparire neanche in altri studi di DR. Compare, nel 2009, in uno studio di un noto gruppo di ricercatori israeliani su questioni di attribuzione autoriale, dove il riferimento si limita appunto a quell'ambito<sup>78</sup>. È possibile che il saggio compaia citato altrove, ma fino ad oggi la metodologia e l'impianto teorico impiegati in questo saggio non hanno avuto ricezione né sono stati successivamente sviluppati negli studi di DR. Parimenti a Burrows non era stato riconosciuto il primato di aver elaborato e proposto un metodo computazionale per la critica e la teoria letterarie.

Sul versante del DIL bisogna, invece, partire da considerazioni più prettamente linguistiche per poterne, poi, valutare conseguenze e ricadute in ordine di dispositivo narrativo e il saggio *Word-Patterns and Story-Shapes* può essere considerato come un indizio dell'impossibilità di svolgere una ricerca automatizzata su questa particolare tipologia di discorso riportato; impossibilità (attuale) che gli studi linguistici anteriori e gli studi letterari successivi sembrano confermare.

Il DIL rientra, infatti, nella categoria del discorso riportato - insieme al discorso diretto, al discorso indiretto e ad altre forme di discorso riportato - e i principali studi linguistici italiani su questo aspetto (adottato, appunto, come espediente narrativo soprattutto a partire dall'Ottocento) non ne presentano delimitazioni univoche. In Herczeg<sup>79</sup>, Cane<sup>80</sup>, Mortara Garavelli<sup>81</sup> e Calaresu<sup>82</sup> non è possibile trarre definizioni, implicazioni o esclusioni formalmente circostanziate e utili a operazionalizzarlo (secondo la definizione di *operationalizing* che si è imposta nel campo con il DR di Moretti<sup>83</sup>).

Per Mortara Garavelli, per esempio, il DIL è una condizione pragmatica data dall'«incontro-assimilazione tra diegesi e mimesi»<sup>84</sup>. Tale condizione è rilevabile soprattutto laddove si incontrino i due centri deittici all'origine della citazione - il centro

---

<sup>77</sup> Jockers, *Macroanalysis: Digital Methods and Literary History*, cit., p. 30.

<sup>78</sup> Moshe Koppel, Jonathan Schler e Shlomo Argamon, *Computational Methods in Authorship Attribution*, «Journal of the American Society for Information Science and Technology», 60, 1, 2009, pp. 9-26.

<sup>79</sup> Giulio Herczeg, *Lo stile indiretto libero in italiano*, Firenze, Sansoni, 1963.

<sup>80</sup> Eleonora Cane, *Il discorso indiretto libero nella narrativa italiana del Novecento*, Roma, Silva, 1969.

<sup>81</sup> Bice Mortara Garavelli, *La parola d'altri*, Palermo, Sellerio, 1985; Bice Mortara Garavelli, *Il discorso riportato*, in *Grande grammatica italiana di consultazione*, a cura di Lorenzo Renzi, Giampaolo Salvi e Anna Cardinaletti, Bologna, il Mulino, vol. 3, 1995, pp. 427-468.

<sup>82</sup> Emilia Calaresu, *Il discorso riportato. Una prospettiva testuale*, Modena, Il Fiorino, 2000; Emilia Calaresu, *Testuali parole. La dimensione pragmatica e testuale del discorso riportato*, Milano, Franco Angeli, 2004.

<sup>83</sup> «The operational approach refers specifically to concepts, and in a very specific way: it describes the process whereby concepts are transformed into a series of operations—which, in their turn, allow to measure all sorts of objects. Operationalizing means building a bridge from concepts to measurement, and then to the world. In our case: from the concepts of literary theory, through some form of quantification, to literary texts», Franco Moretti, «Operationalizing»: or, the function of measurement in modern literary theory, «Pamphlets of the Stanford Literary Lab», 6, December, 2013, p. 1, <<https://litlab.stanford.edu/assets/pdf/LiteraryLabPamphlet6.pdf>> (Consultato: 10 novembre 2025).

<sup>84</sup> Mortara Garavelli, *La parola d'altri*, cit., p. 105.

deittico originario di un enunciato e il centro deittico del contesto citante - e possono essere assunti come suoi segnali di riconoscimento i fenomeni intonativi, morfosintattici, lessicali, testuali e contestuali attraverso i quali si segnala, ma

l'impressione che se ne ritrae è che nessuno basti singolarmente a definire la specificità fenomeno [...]: né, d'altra parte, sembra sufficiente mettere insieme i tratti rilevabili sui vari piani per avere un identikit completo - e omogeneo - di questo tipo di discorso<sup>85</sup>.

Inoltre, il livello lessicale può concorrere a intorbidire un quadro già di per sé sfumato, sfocato e privo, formalmente, di indicatori univoci, siano essi delimitanti, implicanti o escludenti. Infatti, relativamente al lessico, quanta più similarità intercorre tra la lingua impiegata nella diegesi e la lingua impiegata nella mimesi, tanta maggiore sarà la difficoltà nel riconoscere la mimesi all'interno del contesto diegetico<sup>86</sup>. L'assenza di segnali linguistici definiti, al momento, rende impossibile operationalizzare il DIL, poiché è, di conseguenza, impossibile tradurlo nel linguaggio della programmazione informatica.

Un'altra conferma sembra arrivare oltre e dopo Burrows. Infatti, anche Moretti aveva tentato una trattazione automatica della diegesi con finalità di critica e teoria letterarie nel suo pamphlet del 2013, *"Operationalizing": or, the function of measurement in modern literary theory*<sup>87</sup>. Dopo aver esposto il percorso di operazioni che si rende necessario quando si tenta di inserire una metodologia computazionale entro una cornice teorico-letteraria, Moretti tenta di calcolare la quantità di spazio narrativo che viene assegnata a ogni personaggio all'interno dei testi di prosa letteraria. Con questo obiettivo si affida al concetto teorico di *character-space*, ovvero lo spazio del personaggio elaborato da Alex Woloch. Presto si rende conto, però, che il concetto teorico di Woloch è non operationalizzabile. Citando il seguente passaggio di *Pride and Prejudice*, «il signor Bennet era un tale imprevedibile miscuglio di ingegno, sarcasmo, riservatezza e capriccio, che alla moglie non erano bastati ventitré anni per capire il suo temperamento»<sup>88</sup>, si palesa chiaramente come il concetto di Woloch travalichi la sola mimesi (quanto dialogo viene prodotto da un personaggio) e si estenda anche alla diegesi, quanto si dice su quel personaggio (sollevando inoltre problemi sul come trattare e definire i confini formali quando un narratore descrive cosa un personaggio pensa di un altro personaggio<sup>89</sup>). Con l'affinarsi sempre più dell'intelligenza artificiale generativa oppure con l'ipotizzato passaggio all'intelligenza generale sarà magari possibile, un domani, contemplare casi come quest'ultimo, nonché tutti i parametri che si possono intersecare nella resa del DIL per una sua individuazione automatizzata (studi più recenti al riguardo si indirizzano in questo senso<sup>90</sup>).

---

<sup>85</sup> *Ibidem*.

<sup>86</sup> Ivi, p. 120 *et passim*.

<sup>87</sup> Moretti, *"Operationalizing": or, the function of measurement in modern literary theory*, cit.

<sup>88</sup> Jane Austen, *Orgoglio e pregiudizio*, Milano, Sperling, 2002, pp. 5-6.

<sup>89</sup> In un caso del genere è già difficile per un essere umano stabilire i confini formali in cui termina lo spazio del signor Bennet e inizia quello della moglie. E se un essere umano non è in grado di indicare i confini formali su cui il calcolatore deve di conseguenza computare, la ricerca non è automatizzabile. Infatti, Moretti, di conseguenza, restringe il campo d'analisi ai testi teatrali: «Plays are easier in this respect; as there are no ambiguities in how words are distributed among the various speakers, character-space turns smoothly into "word-space"—"the number of words allocated to a particular character"—and, by counting the words each character utters, we can determine how much textual space it occupies», Moretti, *"Operationalizing": or, the function of measurement in modern literary theory*, cit., p. 2.

<sup>90</sup> Kristiina Taivalkoski-Shilov, *Free Indirect Discourse: An Insurmountable Challenge for Literary MT Systems?*, in *Proceedings of the Qualities of Literary Machine Translation*, Dublin, 10 novembre, a cura di James Hadley et al., European Association for Machine Translation, 2019, pp. 35-39, <<https://aclanthology.org/W19-7305.pdf>> (Consultato: 10 novembre 2025); Andrew Piper e Sunyam Bagga, *Using Large Language Models for Understanding Narrative Discourse*, in *Proceedings of the The 6th Workshop on Narrative Understanding*, 15 novembre, a cura di Yash Kumar Lal et al., Association for Computational Linguistics, 2024, pp. 37-46, <<https://doi.org/10.18653/v1/2024.wnu-1.4>> (Consultato: 10 novembre 2025); Aurora Argenzio,

## 5. Conclusioni

Il presente articolo ha ricollocato storicamente e teoricamente la nascita della *Computational Literary Criticism* nei suoi impieghi ai fini di critica e teoria letterarie, attribuendone la primogenitura non alla svolta concettuale del *Distant Reading* di Franco Moretti, bensì al saggio del 1987 di John F. Burrows, *Word-Patterns and Story-Shapes*. Tale riconfigurazione genealogica non sminuisce la portata propulsiva del contributo di Franco Moretti né ignora la fortuna che esso ha avuto nel promuovere e diffondere un nuovo approccio agli studi letterari. Piuttosto, intende sottrarre al silenzio un precedente che, per coerenza metodologica e ambizione teorica, può essere considerato fondativo.

Burrows, operando nel contesto dell'attribuzione autoriale, elabora un impianto teorico che si radica nella tradizione narratologica e propone una divisione del testo letterario in componenti strutturali (diegesi, mimesi, discorso indiretto libero) che apre a riflessioni critiche oltre la mera misurazione stilistica. Il suo saggio *Word-Patterns and Story-Shapes: the Statistical Analysis of Narrative Style* non si limita a mostrare l'efficacia della statistica nell'analisi narrativa, ma avanza l'idea che l'inferenza quantitativa possa alimentare il giudizio critico in modo non sostitutivo ma complementare. Questo risulta essere un nucleo e passaggio metodologico – dalla computazione all'argomentazione critica – che costituisce, a tutti gli effetti, una proposta di teoria letteraria computazionale *ante litteram*.

Le ragioni della mancata ricezione di questo saggio sono imputabili a fattori diversi, tra cui l'ombra lunga della monografia pubblicata nello stesso anno, la fortuna del metodo Delta e, non da ultimo, i limiti operativi legati alla definizione e all'individuazione del discorso indiretto libero. Tuttavia, queste fragilità non tolgono valore al tentativo di Burrows, che rimane un esempio precocemente lucido di convergenza tra formalismo narratologico e metodo computazionale. Riconoscere questo primato non ha solo valore storico: significa ripensare le fondamenta teoriche della CLC, aprendo spazi per una riflessione più consapevole sull'intersezione tra metodo, teoria e interpretazione negli studi letterari digitali.

---

Fabio Ciotti e Anna C. Corradino, *Usare i Large Language Model per l'analisi del testo narrativo: strategie di prompt engineering per il riconoscimento del discorso indiretto libero nella narrativa italiana 1830-1930*, in *Diversità, Equità e Inclusione: Sfide e Opportunità per l'Informatica Umanistica nell'Era dell'Intelligenza Artificiale*, 11-13 giugno, a cura di Simone Rebor, Marco Rospocher e Stefano Bazzaco, Università di Verona, Atti del XIV Convegno Annuale, 2025, pp. 349-356, <<https://amsacta.unibo.it/id/eprint/8380>> (Consultato: 10 novembre 2025).