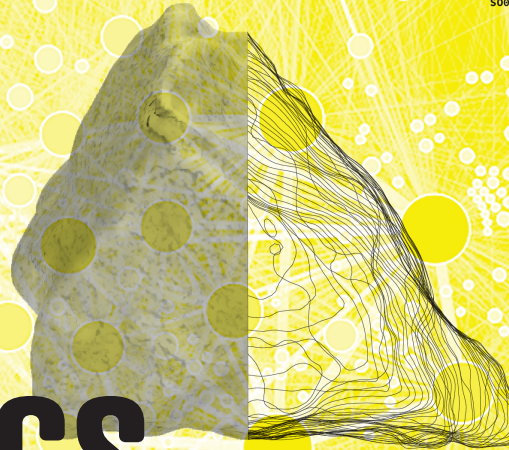
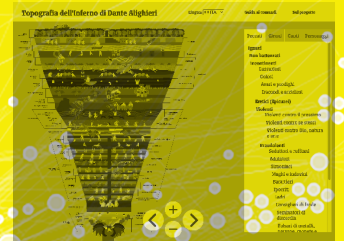


7.18

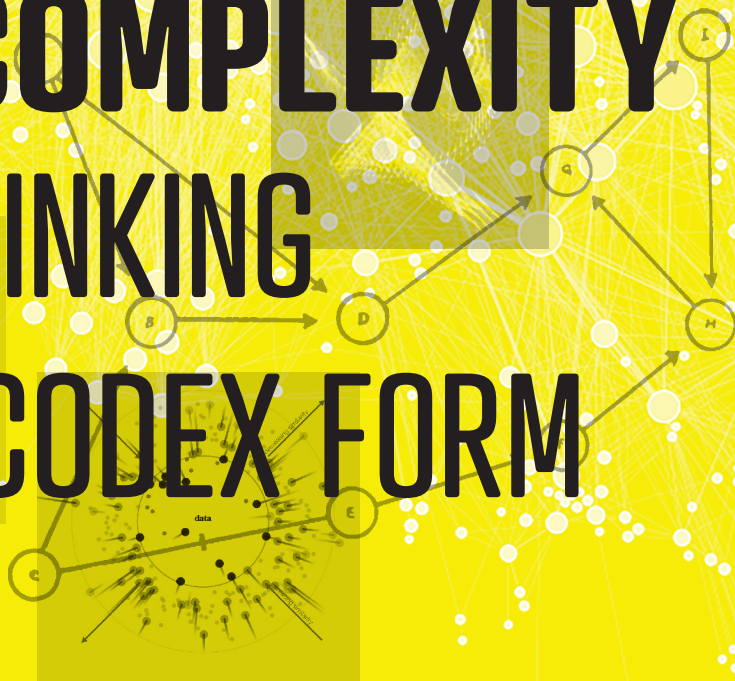
19° 38' N 149° 08' 21" E



5091



WRITINGS OF COMPLEXITY RETHINKING THE CODEX FORM



Progetto Grafico

N. 41, V. 22, Dicembre • December 2025

International Journal
of Communication Design

Semestrale pubblicato da AIAP,
Associazione italiana design
della comunicazione visiva
*Half-yearly published by AIAP,
the Italian Association of Visual
Communication Design*

> pgjournal.aiap.it

ISSN print: 1824-1301

Registrazione del Tribunale di Milano
n. 709 del 19/10/1991. Periodico
depositato presso il Registro Pubblico
Generale delle Opere Protette.
*Milan Court Registration No. 709 of
October 19, 1991. Periodical filed with the
General Public Register of Protected Works.*

Progetto Grafico adotta il sistema
di revisione del double-blind peer review.
*Progetto Grafico adopts a double-blind peer
review system.*

INDICIZZAZIONE

INDEXING

Progetto Grafico è stata inclusa nella
lista ANVUR delle riviste di classe A
per l'area O8 e i settori O8/C1, O8/D1,
O8/E1, O8/F1.

*Progetto Grafico has been included in the
Italian ANVUR list of Class A Journals
for area O8 and sectors O8/C1, O8/D1,
O8/E1, O8/F1.*

Il n. 41 è stato stampato in Italia
da PressUp, Nepi (VT) nel mese
di gennaio 2026

*Issue 41 was printed by PressUp, Nepi
(VT), Italy, in January 2026*

EDITORE

PUBLISHER

AIAP
Associazione italiana design
della comunicazione visiva
via Amilcare Ponchielli, 3
20129 Milano
+39 02 29520590
> aiap@aiap.it
> www.aiap.it

AIAP



CONSIGLIO DIRETTIVO AIAP 2025–2028

AIAP BOARD 2025–2028

PRESIDENTE

PRESIDENT

Francesco E. Guida

VICE PRESIDENTESSA

VICE PRESIDENT

Fabiana Ielacqua

SEGRETARIA GENERALE

GENERAL SECRETARY

Ilaria Montanari

CONSIGLIERI

BOARD MEMBERS

Isabella Battilani

Matteo Carboni

Gaetano Grizzanti

Maria Loreta Pagnani

COLLEGIO DEI PROBIVIRI

PANEL OF ARBITRATORS

Laura Bortoloni *Presidente* *President*

Simonetta Scala *Segretaria* *Secretary*

Stefano Tonti *Past President*

Giangiorgio Fuga

Claudio Madella

REVISORE DEI CONTI

AUDITOR

Dario Carta

SEGRETARIA E AMMINISTRAZIONE

SECRETARIAT AND ADMINISTRATION

Elena Panzeri

PAST PRESIDENT

PAST PRESIDENT

Marco Tortoioli Ricci

CENTRO DI DOCUMENTAZIONE SUL PROGETTO GRAFICO DI AIAP

AIAP GRAPHIC DESIGN DOCUMENTATION CENTRE

> www.aiap.it/cdpg/

RESPONSABILE ARCHIVIO, RICERCHE E BIBLIOTECA

ARCHIVE, RESEARCH AND LIBRARY MANAGER

Lorenzo Grazzani

> biblioteca@aiap.it

DIRETTORE SCIENTIFICO & RESPONSABILE

SCIENTIFIC & MANAGING DIRECTOR

Carlo Martino *Sapienza Università di Roma*

COMITATO SCIENTIFICO INTERNAZIONALE

INTERNATIONAL SCIENTIFIC COMMITTEE

José Manuel Allard *Pontificia Universidad Católica de Chile*

Andreu Balius *EINA, Centre Universitari de Disseny i Art de Barcelona*

Helena Barbosa *Universidade de Aveiro*

Letizia Bollini *Libera Università di Bolzano*

Mauro Bubbico *Abadir Accademia di Design e Comunicazione Visiva*

Valeria Bucchetti *Politecnico di Milano*

Fiorella Bulegato *Università Iuav di Venezia*

Paolo Ciuccarelli *Northeastern University*

Vincenzo Cristallo *Politecnico di Bari*

Federica Dal Falco *Sapienza Università di Roma*

Davide Fornari *ECAL/Haute école d'art et de design de Lausanne*

Rossana Gaddi *Università degli Studi "G. d'Annunzio" Chieti-Pescara*

Stuart Medley *Edith Cowan University*

Francesco Monterosso *Università degli Studi di Palermo*

Matteo Moretti *Università degli Studi di Sassari*

Luciano Perondi *Università Iuav di Venezia*

Daniela Piscitelli *Università degli Studi della Campania "Luigi Vanvitelli"*

Emanuele Quinz *Université Paris 8 Vincennes-Saint-Denis*

Chiara Lorenza Remondino *Politecnico di Torino*

Elisabeth Resnick *Massachusetts College of Art and Design*

Fiona Ross *University of Reading*

Dario Russo *Università degli Studi di Palermo*

Gianni Sinni *Università Iuav di Venezia*

Michael Stoll *Technische Hochschule Augsburg*

Davide Turrini *Università degli Studi di Firenze*

Carlo Vinti *Università degli Studi di Camerino*

DIRETTORI DEL COMITATO EDITORIALE

EDITORS-IN-CHIEF

Alessio Caccamo *Sapienza Università di Roma*

Vincenzo Maselli *Sapienza Università di Roma*

COMITATO EDITORIALE INTERNAZIONALE

INTERNATIONAL EDITORIAL BOARD

Roberta Angari *Università degli Studi della Campania "Luigi Vanvitelli"*

Carlotta Belluzzi Mus *Sapienza Università di Roma*

Laura Bortoloni *Università degli Studi di Ferrara*

Josefina Bravo *University of Reading*

Fabiana Candida *Sapienza Università di Roma*

Dario Carta *CFP Bauer Milano*

Francesca Casnati *Politecnico di Milano*

Leonardo Gómez Haro *Universidad Politécnica de Valencia*

Pilar Molina *Pontificia Universidad Católica de Chile*

María Griñán Montealegre *Universidad de Murcia*

Cristina Marino *Università degli Studi di Parma*

Fabiana Marotta *Università degli Studi di Napoli "Federico II"*

Chris Nuss *University of Birmingham*

Giulia Panadisi *Università degli Studi "G. d'Annunzio" Chieti-Pescara*

Dario Rodighiero *Universiteit Groningen*

Francesca Scalisi *Università degli Studi di Palermo*

Anna Turco *Sapienza Università di Roma*

MAIL DI CONTATTO & SOCIAL MEDIA

CONTACT MAILS & SOCIAL MEDIA

Director > director.pgjournal@aiap.it

Editorial > editors.pgjournal@aiap.it

Instagram @progetto_grafico_journal

LinkedIn @Progetto Grafico Journal

PROGETTO GRAFICO EDITORIALE

EDITORIAL DESIGN

Anna Turco

IMPAGINAZIONE

EDITING

Alessio Caccamo, Vincenzo Maselli, Roberta Sacco

COPERTINA

COVER

Si ringrazia Mauro Bubbico per aver progettato e donato gratuitamente la copertina del numero 41 di Progetto Grafico.

We thank Mauro Bubbico for designing and generously donating the cover of issue 41 of Progetto Grafico

CARATTERI TIPOGRAFICI

TYPEFACE

Calvino by Andrea Tartarelli · Zetafonts

Atrament by Tomáš Brousil · Suitcase Type Foundry

PER LE ATTIVITÀ SVOLTE NEL 2025 RELATIVE AL DOUBLE-BLIND PEER REVIEW PROCESS, SI RINGRAZIANO I SEGUENTI REVISORI
AS CONCERN THE DOUBLE-BLIND PEER REVIEW PROCESS DONE IN 2025
WE WOULD THANKS THE FOLLOWING REFEREES

Emanuela Bonini Lessing *Università Iuav di Venezia*
Lisa Bergerheimer *Offenbach University of Art and Design*
Alessia Brischetto *Università degli Studi di Firenze*
Daniela Calabi *Politecnico di Milano*
Gianluca Camillini *Libera Università di Bolzano*
Susanna Cerri *Università degli Studi di Firenze*
Marcello Costa *Università degli Studi di Palermo*
Andrea Di Salvo *Politecnico di Torino*
Cinzia Ferrara *Università degli Studi di Palermo*
Irene Fiesoli *Università degli Studi di Firenze*
Laura Giraldi *Università degli Studi di Firenze*
Tommaso Guariento *Università Ca' Foscari Venezia*
Francesco E. Guida *Politecnico di Milano*
Ilaria Mariani *Politecnico di Milano*
Raffaella Massacesi *Università degli Studi di Chieti-Pescara*
Federico Oppedisano *Università di Camerino*
Pietro Nunziante *Università degli Studi di Napoli Federico II*
Jonathan Pierini *Università degli Studi di Urbino Carlo Bo*
Leonardo Romei *Sapienza Università di Roma*
Paolo Tamborrini *Università degli studi di Parma*
Umberto Tolino *Politecnico di Milano*

DIRITTI

COPYRIGHTS

La rivista è pubblicata in open access. Tutto il materiale scritto dai collaboratori è disponibile sotto la licenza Creative Commons Attribuzione-NonCommerciale - Condividi allo stesso modo 4.0. Significa che può essere riprodotto a patto di citare Progetto Grafico, di non usarlo per fini commerciali e di condividerlo con la stessa licenza. *This is an open access publication. All material written by the contributors is available under Creative Commons license Attribution-NonCommercial- Share Alike 4.0 International. It can be reproduced as long as you mention Progetto grafico, do not use it for commercial purposes and share it with the same license.*



Le immagini utilizzate in Progetto Grafico rispondono alla pratica del fair use (Copyright Act 17 U.S.C. 107) recepita per l'Italia dall'articolo 70 della Legge sul Diritto d'autore che ne consente l'uso a fini di critica, insegnamento e ricerca scientifica a scopi non commerciali. *The images used in Progetto Grafico comply with fair use practices (Copyright Act 17 U.S.C. 107), implemented in Italy by Article 70 of the Copyright Law, which allows their use for criticism, teaching, and scientific research for non-commercial purposes.*

RINGRAZIAMENTI

AKNWOLEDGEMENTS

Progetto Grafico ringrazia Zetafonts per aver gentilmente concesso l'uso gratuito di uno dei suoi caratteri tipografici per la realizzazione di questa rivista. *Progetto Grafico thanks Zetafonts for kindly providing the complimentary use of one Calvino typefaces for this journal.*

ZETAFONTS™

CREATIVITÀ E CULTURA NELL'EPOCA DELL'AI GENERATIVA

IL RUOLO DELLA SPECIFICITÀ CULTURALE NEL PROGETTO DI CONTENUTI GENERATI DA INTELLIGENZE ARTIFICIALI

4


244 – 263

Francesco Burlando

 [0000-0001-5535-8382](https://orcid.org/0000-0001-5535-8382)


Università degli Studi di Genova
francesco.burlando@unige.it

Boyu Chen

 [0009-0008-0009-180X](https://orcid.org/0009-0008-0009-180X)

Università degli Studi della Campania Luigi Vanvitelli
boyu.chen@unicampania.it

Niccolò Casiddu

 [0000-0002-5010-038X](https://orcid.org/0000-0002-5010-038X)

Università degli Studi di Genova
casiddu@unige.it

Lo studio esplora il rapporto tra lingua, cultura e intelligenza artificiale generativa attraverso un confronto qualitativo tra modelli text-to-image multilingue. Partendo da tre brani letterari originariamente scritti in inglese, italiano e cinese, è stato progettato un esperimento volto ad analizzare l'impatto della lingua d'origine dell'utente nella produzione di contenuti visivi attraverso software di AI, nonché l'impatto della lingua con cui questi ultimi sono stati addestrati. I testi sono stati trasformati in prompt utilizzando sistemi linguistici differenti, e le immagini generate sono state successivamente confrontate in base a criteri qualitativi. Lo studio adotta un approccio esplorativo e si concentra su aspetti come la resa semantica, la coerenza stilistica e la presenza di elementi culturali impliciti o espliciti. L'obiettivo è quello di sondare la possibilità che le AI generative non siano strumenti neutrali destinati ad appiattire culturalmente i contenuti generati in futuro, ma che possano invece riflettere, o distorcere, le specificità cognitive e simboliche legate alla lingua e alla cultura con cui interagiscono.

Introduzione

Il linguaggio rappresenta uno dei principali strumenti attraverso cui l'essere umano traduce le proprie immagini mentali in contenuti condivisibili foneticamente o graficamente, rendendo possibile la comunicazione e la cooperazione tra individui.

245

Non a caso, lo sviluppo di un linguaggio articolato è ritenuto da numerosi studiosi uno dei fattori chiave che ha consentito all'*Homo Sapiens* di prevalere sulle altre specie umane e animali, grazie alla capacità di costruire narrazioni complesse e condividere informazioni astratte (Rizzolatti & Sinigaglia, 2006). Il linguaggio ha inoltre permesso di cooperare su larga scala, fondando credenze comuni e strutture sociali basate su entità intangibili, come religioni, ideologie e istituzioni, capaci di unire gruppi molto estesi di individui (Harari, 2015). Nel corso della storia, le civiltà hanno elaborato forme di linguaggio estremamente differenti tra loro che variano non solo sotto il profilo fonetico e grammaticale, ma anche dal punto di vista cognitivo e concettuale che ne hanno plasmato la cultura.

La linguistica tipologica riconosce l'esistenza di numerose famiglie linguistiche, che condividono la stessa lingua madre originaria (Campbell, 2018). Tali sistemi appaiono profondamente diversi tra loro sia nella struttura grammaticale che nei principi di codifica: basti pensare alla differenza tra lingue Indoeuropee, composte da caratteri alfabetici come l'italiano o l'inglese, le lingue Semitico-camitiche agglutinanti come il turco, o lingue Sino-tibetane, caratterizzate da una natura logografica, come il cinese

che si basa su pittogrammi e caratteri ideografici piuttosto che su un alfabeto fonetico (Comrie, 1989; DeFrancis, 1989). È stato ampiamente dimostrato che il linguaggio influenza la percezione e il pensiero, secondo l'ipotesi di relativismo linguistico (Whorf, 2012): l'esempio spesso citato delle molteplici parole presenti nelle lingue inuit per indicare la neve è certamente iperbolico e tuttavia testimonia la realtà scientifica secondo cui il lessico riflette e allo stesso tempo modella l'esperienza individuale del mondo (Everett, 2023). Nell'era contemporanea, la tecnologia digitale sta rimodellando profondamente le dinamiche comunicative. L'interazione quotidiana con macchine, interfacce conversazionali e Intelligenze Artificiali (AI) sta trasformando la funzione e la forma del linguaggio umano. La comunicazione scritta si è semplificata, ridotta a messaggi brevi, codificati ed efficienti come abbreviazioni, tag o emoji, in parte per adattarsi ai vincoli dei media digitali, in parte per rispondere agli algoritmi di interpretazione delle AI. Con l'avvento dell'AI generativa, emerge una nuova forma di pressione evolutiva: l'umano sembra orientato a modulare il proprio linguaggio per renderlo interpretabile dall'AI, potenzialmente invertendo il paradigma in cui la tecnologia si adatta all'uomo (Wang et al., 2025). In questa prospettiva, ci si può chiedere se il linguaggio umano non stia forse evolvendo per diventare più conforme alle esigenze computazionali piuttosto che alle proprie capacità espressive, ponendo una sfida evolucionistica inedita in cui l'AI, più veloce, efficiente e articolata nella produzione linguistica, potrebbe assumere un ruolo dominante come già è successo all'*Homo Sapiens* in passato. Questo scenario distopico e per ora del tutto speculativo solleva tuttavia questioni critiche concrete e attuali sul rischio di perdita della diversità linguistica e

culturale: in un mondo globalizzato e mediato sempre più dall'intelligenza artificiale, il rischio vero è quello di un appiattimento comunicativo universale, in cui vengono smarrite le sfumature culturali e le specificità locali, con gravi conseguenze sul piano creativo. La creatività personale e specifica che nasce anche dalla cultura individuale rischia di cedere il passo a una standardizzazione dei contenuti sempre più evidente. Tuttavia, seguendo il ragionamento opposto è possibile ipotizzare che la produzione di contenuti attraverso l'AI mantenga tratti distintivi culturali proprio grazie al fatto che le persone che interagiscono con i sistemi generativi portano con sé le specificità cognitive date da una lingua madre, una visione del mondo e una storia culturale.

Parallelamente, anche i modelli stessi di AI riflettono i dataset con cui sono stati addestrati, e sembra perciò logico ipotizzare che la lingua con cui sono composti tali dataset possa costituire una componente importante nell'influenza del funzionamento delle diverse AI. Da ciò si può quindi ipotizzare che la creatività di contenuti visivi, e non solo, assistita dall'AI possa manifestare stili riconoscibili, quasi scuole estetiche o narrative, riconducibili a specifici contesti culturali. Si potrebbe dunque parlare in futuro di un barocco digitale italiano o di una scuola fiamminga dell'AI? Estremizzando questo concetto è ipotizzabile che sopravvivano, o che si rigenerino in modo innovativo, tratti culturali identificabili capaci di generare, ad esempio, una nuova forma del *Made in Italy* creativo, mediato ma non neutralizzato dalla componente artificiale.

246

Metodologia

Per indagare l'ipotesi che l'intelligenza artificiale generativa possa esprimere, attraverso la produzione di immagini, tracce della cultura e della lingua in cui è stata addestrata, si è deciso di progettare un esperimento comparativo fondato sull'uso di due modelli avanzati di AI per la generazione di contenuti visivi. L'obiettivo consiste nell'osservare se la lingua originaria di ciascuna intelligenza artificiale, intesa come lingua prevalente nel dataset di training, potesse in qualche modo influenzare la resa visiva delle immagini prodotte a partire da input testuali letterari (Lindsey et al., 2025).

Per farlo, sono stati selezionati due modelli di intelligenza artificiale generativa text-to-image tra i più rappresentativi del panorama attuale. Il primo è Midjourney, un sistema sviluppato negli Stati Uniti in lingua inglese, largamente adottato in ambiti creativi, pubblicitari e artistici per la qualità estetica delle immagini generate a partire da prompt testuali

in lingua inglese. Il secondo è JiMeng (即梦) AI, un modello di generazione di immagini sviluppato in Cina e ottimizzato per l'elaborazione di prompt in lingua cinese. JiMeng è stato selezionato in quanto rappresentativo di un approccio tecnologico e culturale differente, formato su dataset prevalentemente in lingua cinese.

La scelta delle due intelligenze artificiali risponde a un criterio ben definito: l'inglese e il cinese sono attualmente le due lingue più diffuse nel panorama dell'intelligenza artificiale, non solo per numero di utenti e comunità di ricerca, ma anche per ampiezza e profondità dei dataset su cui i modelli vengono addestrati (Jiang et al., 2025). Utilizzare due AI consolidate, radicate in queste due tradizioni linguistiche, ci ha consentito di lavorare con strumenti robusti, già ampiamente impiegati nel contesto produttivo e accademico.

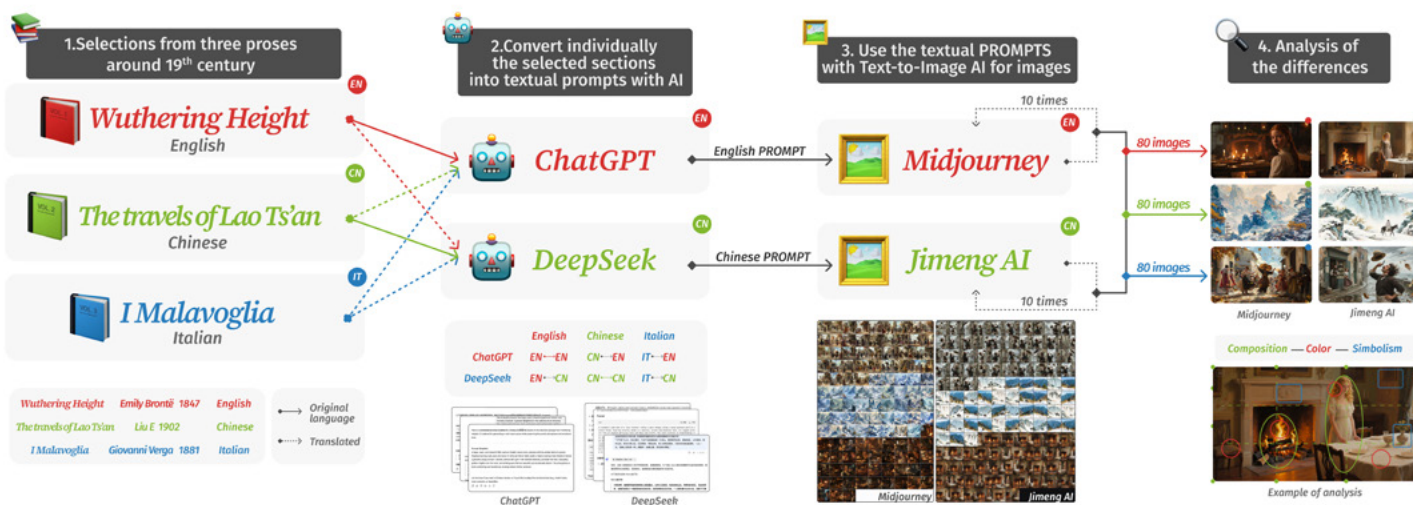
La fase successiva ha previsto la definizione dei materiali testuali da sottoporre ai due modelli. Si è deciso di utilizzare brani in prosa, tratti da opere letterarie riconosciute, in quanto la prosa consente una rappresentazione del linguaggio più naturale e meno artificiosa rispetto alla poesia, dove l'uso di figure retoriche e licenze stilistiche potrebbe falsare il confronto. Inoltre, la prosa letteraria offre descrizioni dettagliate e ricche di immagini mentali, rappresentando un ottimo ponte tra testo e visualità. Al fine di analizzare le immagini generate e presentare le relative osservazioni, è stato determinato a priori un apparato teorico di valutazione strutturato sulla base di tre elementi principali: composizione, colore e simbolismo. L'obiettivo risiede nell'identificare le principali differenze formali nelle immagini generate per ipotizzare possibili differenze nei processi generativi con cui ciascuna AI traduce ed elabora gli elementi. La prima dimensione, la composizione degli elementi dell'immagine, indaga come sia stata tradotta in rappresentazione la relazione narrativa tra i vari elementi presenti nel testo originale. Nello specifico, si determinano la gerarchia degli elementi primari e secondari, la prospettiva e l'organizzazione spaziale (Gombrich et al., 2002). In stretta relazione con la precedente, la seconda dimensione, il colore, oltre a possedere una funzione esteticamente arricchente, si rivela essenziale per veicolare le emozioni del brano. Con riferimento a Itten (1961), l'analisi su questo livello si concentra su come tonalità, luminosità e saturazione siano funzionali ad evocare l'apparato emozionale. La terza dimensione, infine, è quella del simbolismo che permette di indagare la profondità culturale e semantica delle immagini. Seguendo la teoria di Barthes (1977), l'analisi viene fatta a un primo livello di denotazione, legato alla semplice rappresentazione dell'oggetto, e a quello della connotazione, in cui emergono codici

	Inglese	Italiano	Cinese
ChatGPT-Mi djourny	A warm Victorian interior, large cozy room glowing with firelight, burning coal, peat, and wood, rustic wooden table set for a plentiful evening meal, candlelit atmosphere -- in the foreground, a delicate young woman, slender figure, golden flaxen ringlets cascading down her neck, fair skin, small refined features, exquisite face, soft lighting, romantic realism, oil painting style, cinematic composition, 19th century English countryside --v 6 --ar 16:9 --style raw	A windy village street in 19th-century Sicily, women's skirts flying in the breeze, dry leaves swirling through the air, a flamboyant barber holding a razor mid-air while gripping a client's nose, turning to flirt with passersby, his hair shiny and curled like silk. Nearby, a grumpy apothecary stands under a huge floppy hat like an upside-down umbrella, pretending to have a serious talk with the town secretary to avoid going to church, chuckling behind his thick beard and winking at laughing girls splashing through puddles — lively, cinematic scene, Italian realism, warm color palette, Verga-inspired mood, expressive faces, humorous detail --v 6 --ar 16:9	A traveler riding a donkey enters a mountain pass, surrounded by towering mountains like a giant screen, with layered earth and stone, dense mixed trees. After heavy snow, blue rocks contrast with white snow, yellow tree branches, clusters of green pines and cypresses like moss in a Chinese painting. Serene and joyful mountain landscape, traditional Chinese style, late Qing dynasty atmosphere, detailed, scenic, poetic, snow-covered landscape --v 6 --ar 16:9
DeepSeek-Ji meng	温暖欢快的房间，巨大的壁炉燃烧着煤、泥炭和木材，火光跃动。餐桌旁站着年轻的女主人：苗条如少女，身形优美。精致的脸庞，小巧白皙的五官，淡金色卷发垂落纤颈。眼眸迷人，气质忧郁却令人倾倒。古典油画风格，暖色调，柔焦。	19世纪西西里海滨小镇街景，狂风卷起枯叶和女人裙摆。前景：油腻卷发的理发师万尼·皮祖托单手持剃刀悬空，另一手捏顾客鼻子使其仰头，自己扭身回望路人，叉腰姿势夸张，头发丝缎般油亮反光。中景：药铺老板戴滑稽宽檐帽（形似头顶雨伞）斜倚店门，与秘书堂·西尔维斯特罗假意交谈，浓密络腮胡里藏着窃笑，偷瞄水洼里蹦跳的年轻姑娘。背景石砌建筑斑驳，湿漉漉石板路映出阴郁天空，风势强劲。写实油画风格，维托里奥·德·西卡电影色调，细节聚焦市井生活荒诞感	雪后冬日山景，高耸屏风状山脉迎面矗立，土石相间，树木丛生。青色岩石，白色积雪，黄色枯枝，成簇绿色松柏点缀其间如苔点。一人骑驴行进山间，欣赏风景

①

I prompt utilizzati per la generazione text-to-image.

The prompt used for the text-to-image generation.



② Il processo attraverso cui è stato svolto il lavoro di elaborazione dei testi, traduzione in prompt, generazione di immagini e analisi dei risultati ottenuti.
The process through which the work of text processing, translation into prompts, image generation, and analysis of the results obtained was carried out.

culturali (e.g. stili architettonici, abiti, e comportamenti della figura etc), che richiamano valori simbolici.

In seguito alla definizione dell'apparato teorico di base, sono stati selezionati tre brani letterari in tre lingue diverse: inglese, italiano e cinese. La scelta delle lingue risponde a un duplice obiettivo: da un lato, come già detto, confrontare modelli addestrati su lingue profondamente diverse tra loro; dall'altro, esplorare il ruolo della lingua dell'interlocutore nella traduzione di immagini mentali in immagini artificiali. L'inglese è stato scelto non solo per il ruolo centrale che ricopre nella comunicazione globale e nella tecnologia, ma anche perché rappresenta la lingua più utilizzata per interagire con i modelli generativi, specie nel mondo occidentale. Il cinese, come anticipato, è stato incluso per via della sua specificità linguistica e culturale.

Essendo una lingua logografica non alfabetica, i suoi caratteri non rappresentano suoni, bensì concetti o oggetti (Liang, 2022), e questa modalità di codifica risulta particolarmente interessante in un processo che esplora il rapporto tra significato e rappresentazione visiva.

Infine, l'italiano è stato selezionato per il suo valore culturale e per il tipo di linguaggio che tradizionalmente adotta nella forma scritta: una lingua dalla sintassi complessa, ricca di subordinate e sfumature, con una forte connotazione lirica e narrativa, soprattutto nel contesto ottocentesco. Per ciascuna lingua è stato identificato un brano tratto da un autore del XIX secolo, al fine di mantenere una certa coerenza temporale e letteraria. Sono stati identificati autori particolarmente descrittivi ed individuati brani che contenessero descrizioni di scene molto accurate e ricche di dettagli.

Si è prestata attenzione nell'individuare estratti che contenessero descrizioni riconducibili a scene che potessero rientrare in un'unica "inquadratura" per ottimizzare la riuscita del processo di conversione da testo a immagine. I brani vengono esposti di seguito in lingua originale.

Per l'inglese, è stato scelto un passo da *Cime Tempestose* (1847) di Emily Brontë, che descrive un ambiente domestico con toni caldi e vibranti.

We at length arrived in the huge, warm, cheerful apartment where I was formerly received. It glowed delightfully in the radiance of an immense fire, compounded of coal, peat, and wood; and near the table, laid for a plentiful evening meal, I was pleased to observe the 'missis,' an individual whose existence I

had never previously suspected. [...] She was slender, and apparently scarcely past girlhood: an admirable form, and the most exquisite little face that I have ever had the pleasure of beholding; small features, very fair; flaxen ringlets, or rather golden, hanging loose on her delicate neck; and eyes, had they been agreeable in expression, that would have been irresistible. (Bronte, 1847)

Per l'italiano, è stato selezionato un estratto da *I Malavoglia* (1881) di Giovanni Verga, con una scena di vita quotidiana ricca di dettagli sensoriali e sociali.

Il vento faceva volare le gonnelle e le foglie secche, sicché Vanni Pizzuto col rasoio in aria, teneva pel naso quelli a cui faceva la barba, per voltarsi a guardare chi passava, e si metteva il pugno sul fianco, coi capelli arricciati e lustri come la seta; e lo speciale se ne stava sull'uscio della sua bottega, sotto quel cappellaccio che sembrava avesse il paracqua in testa, fingendo aver discorsi grossi con don Silvestro il segretario, perché sua moglie non lo mandasse in chiesa per forza; e rideva del sotterfugio, fra i peli della barbona, ammiccando alle ragazze che sgambettavano nelle pozzanghere. (Verga, 1881)

Per il cinese, si è optato per un passo tratto da *I viaggi di Lao Ts'an* (1902) di Liu E, romanzo esemplare della narrativa cinese tardo-imperiale, che unisce introspezione, critica sociale e descrizioni naturalistiche.

子平进了山口，抬头看时，只见不远前面就是一片高山，像架屏风似的，迎面竖起，土石相间，树木丛杂。却当大雪之后，石是青的，雪是白的，树上枝条是黄的，又有许多松柏是绿的，一丛一丛，如画上点的苔一样。骑着驴，玩着山景，实在快乐得极。(Liu, 1902)①

① Entrando nel passo di montagna, Ziping alzò lo sguardo e osservò una catena di alte vette che si ergevano davanti a lui come una barriera, con i loro pendii rocciosi intervallati da alberi intricati. Dopo la recente nevicata le pietre sono di qīng*, la neve appariva di un bianco puro e i rami degli alberi di un giallo. Mentre i gruppi di pini sempreverdi e ginepri mantenevano un verde, si intrecciano nel paesaggio come i puntini del muschio in un dipinto tradizionale cinese. Cavalcando il suo asino e ammirando il paesaggio montano, provava un'immensa gioia.

La scelta di testi letterari, anziché descrizioni artificiali o prompt sintetici, ha permesso di simulare un processo creativo più autentico, vicino al modo in cui esseri umani (lettori, illustratori, progettisti) tradurrebbero immagini mentali in rappresentazioni concrete. Inoltre, l'utilizzo parallelo delle due AI su tutti i testi – inglese, italiano e cinese – consente di verificare non solo le influenze della lingua del testo, ma anche quelle dovute al substrato culturale dell'intelligenza artificiale che genera le immagini.

La metodologia proposta consente dunque di esplorare in modo sistematico l'eventuale correlazione tra lingua, cultura e output visivo nei sistemi di intelligenza artificiale generativa. Il confronto tra i risultati prodotti dalle due AI a partire dagli stessi testi, e l'analisi delle immagini generate, costituiranno la base per una valutazione qualitativa e semiotica, finalizzata a comprendere se e in che modo la lingua e la cultura del training abbia un impatto stilistico, iconografico o concettuale sui contenuti prodotti.

Procedimento

Il procedimento utilizzato è stato selezionato per rispecchiare nel modo più fedele il processo che un utente farebbe per generare una figura a partire da una sua immagine mentale. Il processo parte quindi da un'immagine mentale generata nella lingua madre dell'utente (Boroditsky, 2011), che equivale ai brani selezionati.

Poiché sempre più spesso gli utenti che intendono far affidamento a un AI generativa text-to-image utilizzano precedentemente un software text-to-text per generare il prompt da utilizzare in seguito (Schulhoff et al., 2024), e questo passaggio di traduzione che serve generalmente per rielaborare un pensiero in una stringa di testo più facilmente digeribile da un software si rivela enormemente utile nel nostro caso in quanto ovviamente i testi scelti non sono stati scritti con l'intento di essere sottoposti a un AI, ogni brano è stato quindi sottoposto in parallelo a ChatGPT e Deepseek attraverso il seguente prompt:

This is the content from [titolo del volume], you need to help me translate it into a [english or chinese] prompt for submitting it to [Midjourney OR Ji Meng].

Il testo viene quindi fornito in lingua originale e, perciò: il testo nativo in lingua inglese verrà mantenuto in inglese da ChatGPT e quindi sottoposto a Midjourney e tradotto in Cinese da Deepseek e quindi sottoposto a JiMeng; il testo nativo in lingua cinese verrà tradotto in inglese da ChatGPT e quindi sottoposto a Midjourney e mantenuto in Cinese da Deepseek e quindi sottoposto

a JiMeng; il testo nativo in lingua italiana verrà tradotto in inglese da ChatGPT e quindi sottoposto a Midjourney e tradotto in Cinese da Deepseek e quindi sottoposto a JiMeng.

La seconda fase prevede di sottoporre i prompt generati a Midjourney e JiMeng. Ogni prompt corrisponde fedelmente all'output ottenuto dalla prima fase. I prompt in lingua inglese sono stati sottoposti a Midjourney e i prompt in lingua cinese sono stati sottoposti a JiMeng. Per ogni prompt ciascun software elabora quattro immagini; questo specifico passaggio è dovuto alle modalità di funzionamento dei software di AI Generativa. A questo punto, per avere un campione più significativo, lo stesso processo è stato iterato dieci volte per ciascun prompt, arrivando a generare quindi, per ogni brano inizialmente selezionato, quaranta immagini tramite Midjourney (MJ) che derivano da un prompt in lingua inglese e quaranta immagini tramite Ji Meng AI (JM) che derivano da un prompt in lingua cinese ①. Dopo aver generato un totale di 240 immagini (Burlando et al., 2025), si è proceduto a mettere a confronto i risultati per analizzare quali differenze vi siano tra le immagini che sono originate da brani diversi e le immagini che sono originate da software diversi. Nello specifico, l'obiettivo dello studio non è tanto di indagare se un software sia meglio dell'altro nella fedeltà semantica al testo o nella ricchezza e coerenza visiva, quanto se siano riscontrabili segni culturali e simbolici come una culturalizzazione dell'immagine, o una coerenza e complessità iconografica, o ancora se vi siano estetiche ricorrenti ②.

Discussione e conclusioni

Dagli output prodotti si possono evincere alcuni spunti molto interessanti di analisi che sembrano suggerire che vi sia una differenza tra le immagini generate a seconda della lingua di origine e del software impiegato. Partendo dal testo di Emily Bronte, è possibile notare come nelle immagini generate da MJ vi sia una variabilità molto maggiore rispetto a quelle generate da JM. Nelle seconde, sembra sostanzialmente di assistere sempre alla stessa scena in cui i pesi tra i vari elementi compositivi sono sempre gli stessi e i cambiamenti si limitano alla posizione e all'orientamento del corpo della ragazza. Al contrario, MJ genera immagini che presentano un'altissima variabilità tanto nella scena ritratta quanto nel peso degli elementi compositivi, come nel caso della donna che il più delle volte rappresenta l'elemento principale dell'immagine con sguardo sempre fisso in camera quasi a voler rompere la quarta parete mentre, in alcuni casi, non è neanche presente nella scena. Se questi elementi potrebbero essere dettati da una differenza nel training dell'AI che non ha però diretto riscontro nella lingua e nella cultura

di provenienza del software specifico, ci sono ulteriori elementi di contrasto che appaiono più interessanti. Ad esempio, nelle immagini prodotte da MJ troviamo quasi sempre una presenza importante di candele disposte su candelabri di tutti i tipi suggerendo che la scena sia ambientata durante la sera, nonostante le candele non manchino anche nelle poche immagini che mostrano delle finestre da cui proviene chiaramente la luce del giorno. Al contrario, nelle immagini generate da JM non vi sono quasi mai elementi di illuminazione artificiale e la scena è sempre ambientata in un contesto diurno. Nelle poche immagini in cui si possono scorgere delle candele, sembra piuttosto trattarsi delle classiche lampade elettriche che richiamano la forma della candela. Questo è certamente dovuto al fatto che, avendo dato indicazione sul testo di origine del brano, ChatGPT si premura di fornire a MJ una contestualizzazione del periodo storico e dello stile dell'epoca, elementi completamente assenti nel prompt generato da DeepSeek.

Poiché è ipotizzabile che DeepSeek conosca la produzione della Bronte, non resta che domandarsi se vi sia un'influenza culturale generata dai dataset con cui le due AI sono state addestrate ③④.

251

Per quanto riguarda il brano di Liu E, si può dire lo stesso riguardo alla prima parte dell'analisi precedente: le immagini generate da JM sono sostanzialmente due scene ripetute in diverse varianti: una in stile fotorealistico, l'altra nel classico stile di un dipinto tradizionale cinese che si privilegia la creazione dell' 'yijing' (意境), ovvero un'atmosfera poetica e simbolica, piuttosto che la mera riproduzione realistica del mondo visibile (Xie, 2023). Tuttavia, all'interno delle immagini riconducibili a una delle due scene troviamo grandissima coerenza compositiva. Le immagini generate da MJ presentano nuovamente una grande variabilità, tanto nello stile compositivo quanto nei pesi degli elementi che compongono la scena.

Pur avendo tutte un'estetica riconducibile alla cultura cinese, la difficoltà di inquadrare precisamente uno stile specifico potrebbe essere sintomo di una carenza culturale da parte dell'AI in questione. Inoltre la lingua cinese presenta una percezione e una rappresentazione del colore più fine e precisa rispetto ad altre lingue. Di conseguenza, la distinzione tra le diverse tonalità è più approfondita. Ad esempio, il termine '青' (Qīng) indica una tonalità concettuale dei colori intermedi tra nero, verde, blu, variando dalle situazioni (Tao & Wong, 2019). Il brano selezionato di Liu presenta un ricco repertorio cromatico riferito a elementi naturali e oggetti concreti. Tuttavia, nel prompt generato automaticamente

da ChatGPT, il colore *qīng* è stato classificato in modo univoco come blu, una semplificazione che ha condizionato le fasi successive della generazione visiva tramite intelligenza artificiale, con conseguente perdita di coerenza estetica e semantica nelle immagini prodotte. Al contrario, DeepSeek ha dimostrato una maggiore sensibilità culturale, riuscendo a preservare l'ambiguità produttiva e simbolica del termine *qīng*, e a inserirlo con coerenza contestuale nelle immagini finali generate da JM ⑤⑥. Tale fedeltà riflette profondamente le differenze tra sistemi linguistici nel categorizzare la percezione visiva. Infatti, come dimostrato da Kay & Regier (2006), le categorie lessicali dei colori influenzano direttamente la precisione e la rapidità nei compiti cognitivi, rivelando un'integrazione adattiva tra linguaggio e percezione.

Le immagini meno rispondenti al testo sono certamente, in entrambe le versioni, quelle tratte dal romanzo di Verga. Se si può riconoscere che il linguaggio utilizzato dallo scrittore non abbia facilitato la traduzione in prompt ad opera di ChatGPT e DeepSeek, non si può negare che mancano molti elementi e che il quadro complessivo non restituisce la scena narrata originariamente. A ottenere un risultato peggiore è MJ, nonostante ci si potesse aspettare che l'opera di Verga potesse essere piuttosto sconosciuta da Deepseek. In realtà, il software cinese inserisce nel prompt la vicinanza al mare, che ritroviamo in tutte le immagini, nonostante l'autore non ne faccia menzione. Questo può essere frutto di una casualità, oppure dimostra una conoscenza del territorio in cui il romanzo è ambientato. ChatGPT, contrariamente, inserisce nel prompt l'indicazione di una piazza, elemento che comunque non viene riscontrato spesso visto che la maggior parte delle scene realizzate sono ambientate in una strada. In conclusione, si può ipotizzare che la maggior distanza linguistica e culturale tra l'autore e i software utilizzati abbia contribuito a intaccare la qualità dei risultati ottenuti ⑦⑧.

Limiti e prospettive future

La presente sperimentazione presenta alcune limitazioni metodologiche che è importante esplicitare. Una prima criticità riguarda la scelta dei testi letterari. Per garantire una certa coerenza stilistica e cronologica, si è deciso di selezionare brani di autori attivi nel XIX secolo. Tuttavia, se da un lato è plausibile considerare la produzione letteraria italiana e inglese di quel periodo come culturalmente e stilisticamente comparabile, dall'altro risulta più problematico stabilire un equivalente diretto nel contesto cinese. Nonostante ciò, si è optato per mantenere l'omogeneità temporale come criterio, dal momento che una selezione basata su altri criteri avrebbe comportato arbitrarietà analoghe in assenza di un criterio oggettivamente più solido.



③

Immagine tratta dal brano di Cime
Tempestose realizzata da MidJourney a
partire dal prompt in lingua inglese.
*Image taken from the part from Wuthering
Heights created by MidJourney based on the
prompt in English.*



Immagine tratta dal brano di Cime
Tempestose generata da JiMeng a partire dal
prompt in lingua cinese.
*Image taken from the part from Wuthering Heights
created by Jimeng based on the prompt in Chinese.*

Un secondo elemento da considerare riguarda l'asimmetria tra le lingue coinvolte nella sperimentazione. A differenza dell'inglese e del cinese, per i quali esistono AI generative text-to-image addestrate su grandi dataset in lingua nativa, non è attualmente disponibile una tecnologia analoga per l'italiano. Di conseguenza, il testo letterario in italiano è stato sottoposto a un passaggio di traduzione in entrambi i casi, mentre i testi inglesi e cinesi vengono tradotti solo in uno dei due casi. Questa condizione introduce un potenziale squilibrio, ma riflette anche una situazione realistica: è infatti improbabile, almeno nel medio periodo, che emergano AI generative addestrate in modo paritario su tutte le lingue del mondo. In questo senso, la necessità di tradurre dalla propria lingua madre a quella nativa del modello AI rappresenta una dinamica già oggi diffusa, e destinata a restare centrale nel modo in cui gli utenti interagiranno con le tecnologie generative. Inoltre, è doveroso ammettere che anche il solo passaggio di traduzione da testo a prompt è, per quanto irrinunciabile, in qualche modo deleterio rispetto al valore linguistico e artistico della produzione originale in prosa.

254

Un ulteriore elemento limite risiede nell'approccio qualitativo adottato per l'analisi dei risultati. Benché sia stato definito a priori un apparato teorico di valutazione, quest'ultimo è volto solo a individuare gli elementi compositivi principali su cui sarebbe stata svolta l'analisi in seguito. Questa modalità permette una lettura ricca e interpretativa delle immagini generate, ma non si basa su una griglia di valutazione quantitativa costruita a priori su criteri specifici e standardizzati. In altre parole, i criteri osservati emergono nel corso dell'analisi in modo qualitativo, rendendo difficile una replicabilità comparativa dello studio. Tuttavia, questo limite è in parte giustificato dalla natura esplorativa del lavoro, che ha come obiettivo principale quello di sondare la fattibilità e la rilevanza di uno studio comparativo tra lingua, cultura e AI generativa.

I recenti studi di Lindsey et al. (2025) evidenziano come un approccio esplorativo bottom-up possa essere preferibile in simili contesti ancora poco esplorati per non limitare il campo di indagine e favorire lo sviluppo delle ricerche future. In prospettiva, i risultati emersi aprono la strada a studi più sistematici e controllati, nei quali l'indagine possa essere condotta su un numero maggiore di testi e con una griglia di valutazione definita ex ante, basata su parametri linguistici, semiotici e visuali condivisi. Una volta

verificata l'ipotesi che le AI generative riflettano, in modo diretto o indiretto, specificità culturali apprese durante il training, sarà possibile passare a una seconda fase della ricerca: non solo riconoscere la presenza di queste influenze, ma comprenderne la natura e il funzionamento, interrogandosi su come tali differenze possano incidere sull'universalità, la rappresentatività e l'equità dei contenuti prodotti dai modelli.

CREATIVITY AND CULTURE IN THE AGE OF GENERATIVE AI THE ROLE OF CULTURAL SPECIFICITY IN THE DESIGN OF AI-GENERATED CONTENT

Artificial Intelligence, Prompt design, Generative AI, language, human-AI interaction

Abstract

The paper explores the relationship between language, culture, and generative artificial intelligence through a qualitative comparison of multilingual text-to-image models. Starting from three literary passages originally written in English, Italian, and Chinese, a trial was designed in order to analyze the impact of the user's native language on the production of visual content through AI software, as well as the impact of the language with which the latter were trained. The texts were transformed into prompts using different linguistic systems, and the generated images were then compared based on qualitative criteria.

The study takes an exploratory approach and focuses on aspects such as semantic rendering, stylistic consistency, and the presence of implicit or explicit cultural elements. The aim is to explore the possibility that generative AI is not a neutral tool destined to culturally flatten the content generated in the future, but that it may instead reflect, or distort, the cognitive and symbolic specificities linked to the language and culture with which it interacts.

Introduction

Language is one of the primary instruments through which humans translate their internal visual representations into content that can be both phonetically and graphically shared, thereby facilitating communication and cooperation among individuals.

For this reason, the development of articulate language is considered by many scholars to be one of the key factors that allowed *Homo Sapiens* to prevail over other human and animal species (Rizzolatti & Sinigaglia, 2006). This ability to construct complex narratives and share abstract information has been demonstrated by *Homo Sapiens* since the earliest stages of their evolution. Language has also enabled large-scale cooperation, establishing common beliefs and social structures based on intangible entities, such as religions, ideologies, and institutions, capable of setting up very large groups of individuals (Harari, 2015).

Throughout history, civilisations have developed extremely divergent forms of language, which vary not only in terms of phonetics and grammar, but also in terms of cognition and concept. These divergent forms of language have had a profound impact on the cultures to which they are associated. Typological linguistics acknowledges the existence of numerous language families that share the same original mother language (Campbell, 2018). These systems appear profoundly different from each other in both grammatical structure and coding principles, considering the difference between Indo-European languages, composed of alphabetic characters such as Italian or English, Semitic-Hamitic agglutinative languages such as Turkish, or Sino-Tibetan languages, characterised by a logographic nature, such as Chinese, which is based on pictograms and ideographic characters rather than a phonetic alphabet (Comrie, 1989; DeFrancis, 1989).

It has been widely demonstrated that language influences perception and thought, according to the linguistic relativity hypothesis (Whorf, 2012): the often cited example of the multiple words existing in Inuit languages to refer to snow is certainly hyperbolic, yet it bears witness to the scientific reality that vocabulary both reflects and shapes individual experience of the world (Everett, 2023).

In the contemporary era, digital technology is deeply reshaping communicative dynamics. Daily interaction with machines, conversational interfaces, and Artificial Intelligence (AI) is transforming the function and form of human language. Written communication has been simplified, reduced to short, coded, and efficient messages such as abbreviations, tags, or emojis, partially to adapt to the constraints of digital media, and partially to respond to AI interpretation algorithms. The advent of generative AI has given rise to a novel form of evolutionary pressure, whereby

humans appear to be exhibiting a propensity to modify their language in order to make it more interpretable by AI. This phenomenon has the potential to invert the established paradigm, by which technology has historically adapted to the needs of humans (Wang et al., 2025). From this perspective, it can be wondered whether human language is evolving to become more compliant with computational requirements rather than its own expressive capabilities, posing an unprecedented evolutionary challenge in which AI, which is faster, more efficient, and more articulate in linguistic production, could assume a dominant role as *Homo Sapiens* did in the past.

This dystopian scenario, which is entirely speculative for now, nevertheless raises concrete and actual critiques about the risk of losing linguistic and cultural diversity: in a globalized world increasingly mediated by Artificial Intelligence, the real risk is to move towards a universal communicative flattening, in which cultural nuances and local specificities are erased, with serious consequences in terms of creativity. Personal and specific originality, which is also born from individual culture, risks giving way to an increasingly evident standardization of content. However, following the opposite line of reasoning, it is possible to assume that the production of content through AI maintains distinctive cultural traits precisely because the people who interact with generative systems bring with them the cognitive specificities given by a mother language, a personal vision, and a cultural history. Likewise, AI models themselves reflect the datasets with which they have been trained, and it therefore seems logical to assume that the language in which these datasets are composed may be an important factor influencing the functioning of different AIs. From this, it can be hypothesized that the creativity of visual content - and not only that - assisted by AI may manifest recognizable styles, almost aesthetic or narrative schools, attributable to specific cultural contexts. Thus, it can be envisioned the emergence of an Italian digital Baroque or a Flemish school of AI. Taken to its logical extreme, this reasoning suggests that identifiable cultural traits may not only persist but also be reconfigured in innovative forms, potentially generating new expressions of a *Made in Italy* creativity, mediated, yet not neutralised by the artificial component.

Methodology

To investigate the hypothesis that generative AI can express, through the production of images, traces of the culture and language on which it was trained, we decided to design a comparative experiment based on the use of two advanced AI models for the generation of visual content. The objective of this study is to observe whether the original language of each artificial intelligence, understood as the predominant language

in the training dataset, could in some way influence the visual rendering of the images produced from literary textual inputs (Lindsey et al., 2025).

To this end, two of the most representative text-to-image generative AI models currently available were selected for investigation. The first of these is Midjourney, a system developed in the United States, which has gained wide popularity in creative, advertising and artistic fields due to the aesthetic quality of the images generated from English text prompts. The second is JiMeng (即梦) AI, an image generation model developed in China and optimised for processing prompts in Chinese language.

The selection of these two distinct Artificial Intelligences is informed by a clearly delineated criterion: in the contemporary context, English and Chinese are the two most prevalent languages within the domain of AI, not only in terms of the magnitude of their user base and research communities, but also in the comprehensive scope and intricate depth of the datasets on which these models are trained (Jiang et al., 2025). The use of two established AI models, grounded in these two linguistic traditions, has permitted the development of robust tools that are already extensively employed in both production and academic domains. The next step involved defining the textual materials to be submitted to the two models. It was decided to use prose passages taken from recognized literary works, since prose allows for a more natural and less artificial representation of language than poetry, where the use of figure of speech and stylistic licenses could distort the comparison. In addition, literary prose offers detailed descriptions rich in mental images, representing an excellent bridge between text and visuality.

In order to analyze the generated images and present the related observations, an a priori theoretical evaluation framework was established based on three main elements: composition, color, and symbolism. The objective lies to identify the main formal differences in the generated images for hypothesizing possible differences in the generative processes in which each AI translates and processes the elements. The first dimension, the composition of the image elements, investigates how the narrative relationship between the various elements in the original text has been translated into representation. Specifically, the hierarchy of primary and secondary elements,

perspective, and spatial organization are determined (Gombrich et al., 2002). Closely related to the previous one, the second dimension, color, in addition to having an aesthetically enriching function, proves essential for conveying the emotions of the composition. With reference to Itten (1961), the analysis at this level focuses on how hue, brightness, and saturation are functional in evoking the emotional apparatus. Finally, the third dimension is symbolism, which allows us to investigate the cultural and semantic depth of images. Following Barthes' theory (1977), the analysis is carried out at a first level of denotation, linked to the simple representation of the object, and at the level of connotation, where cultural codes emerge (e.g., architectural styles, clothing, and behaviors of the figure, etc.), which recall symbolic values. In accordance with the definition of the basic theoretical framework, three literary passages were selected in three different languages: English, Italian, and Chinese.

The choice of languages responds to a twofold objective: at first, as already mentioned, to compare models trained on languages that are profoundly different from each other. Moreover, to explore the role of the interlocutor's language in the translation of mental images into artificial ones. English was chosen not only for its central role in global communication and technology, but also because it is the most widely used language for interacting with generative models, especially in the Western world. Chinese, as anticipated, was included because of its linguistic and cultural specificity. Being a logographic, non-alphabetic language, its characters do not represent sounds, but rather concepts or objects (Liang, 2022), and this mode of encoding is particularly interesting in a process that explores the relationship between meaning and visual representation. Finally, Italian was selected for its cultural value and for the type of language it traditionally adopts in written form: a language with complex syntax, rich in subordinate clauses and nuances, with a strong lyrical and narrative connotation, especially in the nineteenth-century context. For each language, it was identified a prose composition from a 19th-century author in order to maintain a certain temporal and literary consistency. Particularly descriptive authors were identified and passages containing very accurate and detailed descriptions of scenes were selected. An extra attention was made to identify extracts, containing descriptions that could be traced back to scenes that could be accommodated within a single "frame", with a view to optimizing the success of the text-to-image conversion process. The passages are presented below in their original language. For English, a part from *Wuthering Heights* (1847) by Emily Brontë was chosen, which describes a domestic setting with warm and vibrant tones.



⑤

Immagine tratta dal brano de I viaggi di Lao Ts'an realizzata da MidJourney a partire dal prompt in lingua inglese.

Image taken from the part from The Travels of Lao Ts'an created by MidJourney based on the prompt in English.

⑥

Immagine tratta dal brano de I viaggi di Lao Ts'an realizzata da JiMeng a partire dal prompt in lingua cinese.

Image taken from the part from The Travels of Lao Ts'an created by JiMeng based on the prompt in Chinese.

We at length arrived in the huge, warm, cheerful apartment where I was formerly received. It glowed delightfully in the radiance of an immense fire, compounded of coal, peat, and wood; and near the table, laid for a plentiful evening meal, I was pleased to observe the 'missis,' an individual whose existence I had never previously suspected. [...] She was slender, and apparently scarcely past girlhood: an admirable form, and the most exquisite little face that I have ever had the pleasure of beholding; small features, very fair; flaxen ringlets, or rather golden, hanging loose on her delicate neck; and eyes, had they been agreeable in expression, that would have been irresistible (Bronte, 1847).

For Italian, a part from Giovanni Verga's *I Malavoglia* (1881) was selected, featuring a scene from everyday life rich in sensory and social details.

258

Il vento faceva volare le gonnelle e le foglie secche, sicché Vanni Pizzuto col rasoio in aria, teneva pel naso quelli a cui faceva la barba, per voltarsi a guardare chi passava, e si metteva il pugno sul fianco, coi capelli arricciati e lustri come la seta; e lo speciale se ne stava sull'uscio della sua bottega, sotto quel cappellaccio che sembrava avesse il paracqua in testa, fingendo aver discorsi grossi con don Silvestro il segretario, perché sua moglie non lo mandasse in chiesa per forza; e rideva del sotterfugio, fra i peli della barbona, ammiccando alle ragazze che sgambettavano nelle pozzanghere. (Verga, 1881) ❶

For Chinese, a part was chosen from Liu E's *The Travels of Lao Ts'an* (1902), an exemplary novel of late imperial Chinese fiction that combines introspection, social criticism, and naturalistic descriptions.

子平进了山口，抬头看时，只见不远前面就是一片高山，像架屏风似的，迎面竖起，土石相间，树木丛杂。却当大雪之后，石是青的，雪是白的，树上枝条是黄的，又有许多松柏是绿的，一丛一丛，如画上点的苔一样。骑着驴，玩着山景，实在快乐得极。(Liu, 1902) ❷

The choice of literary texts, rather than artificial descriptions or synthetic prompts, allowed us to simulate a more authentic creative process, closer to the way humans (readers, illustrators, designers) would translate mental images into concrete representations. Furthermore, the parallel use of the two AIs on all texts, English, Italian, and Chinese, allows us to verify not only the influences of the language of the text, but also those due to the cultural substrate of the artificial intelligence that generates the images.

The proposed methodology therefore allows us to systematically explore the possible correlation between language, culture, and visual output in generative artificial intelligence systems. The comparison between the results produced by the two AIs from the same texts and the analysis of the generated images will form the basis for a qualitative and semiotic evaluation, aimed at understanding whether and how the language and culture of the training has a stylistic, iconographic, or conceptual impact on the content produced.

Procedure

The procedure used was selected to reflect as closely as possible the process that a user would follow to generate a figure from a mental image.

The process therefore starts from a mental image generated in the user's native language (Boroditsky, 2011), which is equivalent to the selected passages. Users who intend to rely on generative text-to-image AI increasingly use text-to-text software to generate the prompt to be used later (Schulhoff et al., 2024), and this translation step, which generally serves to rework a thought into a string of text that is more easily processable by the software, proves to be enormously useful in our case as the chosen texts were obviously not written with the intention of being submitted to an AI. Each passage was therefore submitted in parallel to ChatGPT and Deepseek using the following prompt:

❶

The wind blew up skirts and dry leaves, so Vanni Pizzuto, with his razor in the air, held those he was shaving by the nose, turning to look at those passing by, and put his fist on his hip, his hair curly and shiny like silk; and the apothecary stood at the door of his shop, under that big hat that looked like he had an umbrella on his head, pretending to have important conversations with Don Silvestro the secretary, so that his wife wouldn't force him to go to church; and he laughed at the subterfuge, among the hairs of his beard, winking at the girls who were splashing in the puddles.

❷

Entering the mountain pass, Ziping looked up and saw a chain of high peaks rising before him like a barrier, their rocky slopes interspersed with tangled trees. After the recent snowfall, the stones were qīng, the snow appeared pure white, and the tree branches were yellow. While the clusters of evergreen pines and junipers remained green, they intertwined in the landscape like dots of moss in a traditional Chinese painting. Riding his donkey and admiring the mountain scenery, he felt immense joy.

This is the content from [titolo del volume], you need to help me translate it into a [english or chinese] prompt for submitting it to [Midjourney OR Ji Meng].

The text is therefore provided in its original language and, therefore: the native English text will be kept in English by ChatGPT and then submitted to Midjourney and translated into Chinese by Deepseek and then submitted to JiMeng; the native Chinese text will be translated into English by ChatGPT and then submitted to Midjourney and kept in Chinese by Deepseek and then submitted to JiMeng; the native Italian text will be translated into English by ChatGPT and then submitted to Midjourney and translated into Chinese by Deepseek and then submitted to JiMeng.

The second phase involves submitting the generated prompts to the AIs. Each prompt corresponds exactly to the output obtained in the first phase. The English prompts were submitted to Midjourney and the Chinese prompts were submitted to JiMeng. For each prompt, each software program processes four images; this specific step is due to the way Generative AI software works. At this point, in order to obtain a more meaningful sample, the same process was repeated ten times for each prompt, thus generating, for each initially selected composition, forty images via Midjourney (MJ) derived from an English-language prompt and forty images via Ji Meng AI (JM) derived from a Chinese-language prompt ①.

After generating a total of 240 images (Burlando et al., 2025), the results were compared to analyze the differences between images originated from different passages and images originated from different software. Specifically, the aim of the study is not to investigate whether one software program is better than another in terms of semantic fidelity to the text or visual richness and consistency, but rather whether cultural and symbolic signs can be found, such as a culturalization of the image, or iconographic consistency and complexity, or even whether there are recurring aesthetics ②.

Discussion and conclusions

The outputs produced highlight some very interesting points for analysis, which seem to suggest that there is a difference between the images generated depending on the source language and the software used. Starting with Emily Bronte's text, it can be seen that the images generated by MJ show much greater variability than those generated by JM. In the latter, it seems that the observer is essentially witnessing the same scene, in

which the weights of the various compositional elements are always the same and the changes are limited to the position and orientation of the girl's body. In contrast, MJ generates images that show a very high degree of variability both in the scene portrayed and in the weight of the compositional elements, as in the case of the woman who most often represents the main element of the image, always staring at the camera as if wanting to break the fourth wall, while in some cases she is not even present in the scene. While these elements could be dictated by a difference in AI training that is not directly reflected in the language and culture of origin of the specific software, there are other contrasting elements that appear more interesting. For example, in the images produced by MJ, we almost always find a significant presence of candles arranged on candlesticks of all kinds, suggesting that the scene is set in the evening, even though candles are also present in the few images that show windows from which daylight clearly shines in. On the contrary, in the images generated by JM, there are almost never any elements of artificial lighting, and the scene is always set in a daytime context. In the few images where candles can be seen, they appear to be classic electric lamps that resemble the shape of candles. This is certainly due to the fact that, having provided information on the source text of the passage, ChatGPT takes care to provide MJ with a contextualization of the historical period and style of the era, elements that are completely absent in the prompt generated by DeepSeek. Since it is conceivable that DeepSeek is familiar with Bronte's work, one can only wonder whether there is a cultural influence generated by the datasets with which the two AIs were trained ③④.

As for Liu E's piece, the same can be said about the first part of the previous analysis: the images generated by JM are essentially two scenes repeated in different variations: one in a photorealistic style, the other in the classic style of a traditional Chinese painting that favors the creation of 'yijing' (意境), or a poetic and symbolic atmosphere, rather than the mere realistic reproduction of the visible world (Xie, 2023). However, within the images attributable to one of the two scenes, we find great compositional consistency.

The images generated by MJ again show great variability, both in compositional style and in the weighting of the elements that make up the scene. Although they all have an aesthetic attributable to Chinese culture, the difficulty in precisely framing a specific style could be symptomatic of a cultural deficiency on the part of the AI in question. Furthermore, the Chinese language has a more refined and precise perception and representation of color if compared to other languages. As a result, the distinction between different tones is more profound. For example, the term '青' (Qīng) indicates a conceptual shade of colors between black, green, and blue, varying

according to the situation (Tao & Wong, 2019). Liu's selected passage presents a rich chromatic repertoire referring to natural elements and concrete objects. Yet, in the prompt automatically generated by ChatGPT, the color *qing* was uniquely classified as blue, a simplification that influenced the subsequent stages of visual generation through artificial intelligence, resulting in a loss of aesthetic and semantic consistency in the images produced. In contrast, DeepSeek demonstrated greater cultural sensitivity, managing to preserve the productive and symbolic ambiguity of the term *qing* and to insert it with contextual consistency into the final images generated by JM ©©. This fidelity deeply reflects the differences between linguistic systems in categorizing visual perception. In fact, as demonstrated by Kay & Regier (2006), lexical categories of colors directly influence the accuracy and speed of cognitive tasks, revealing an adaptive integration between language and perception.

The images that are least consistent with the text are certainly, in both versions, those taken from Verga's novel. While it can be acknowledged that the language used by the writer did not facilitate translation into prompts by ChatGPT and DeepSeek, it cannot be denied that many elements are missing and that the overall picture does not reflect the scene originally narrated. MJ achieves the worst result, although it could be expected that Verga's work would be rather unknown to Deepseek. In fact, the Chinese software includes proximity to the sea in the prompt, which we find in all the images, even though the author makes no mention of it. This may be a coincidence, or it may demonstrate knowledge of the area in which the novel is set. ChatGPT, on the other hand, includes a reference to a square in the prompt, an element that is not often found, as most of the scenes are set on a street. In conclusion, it can be assumed that the greater linguistic and cultural distance between the author and the software used has contributed to the quality of the results obtained ©©.

260

Limitations and future perspectives

The work has certain methodological limitations that are of significance and must be highlighted. The initial issue concerns the selection of literary texts. In order to ensure stylistic and temporal consistency, it was decided that pieces would be selected by authors who were active during the 19th century. While it is plausible to consider Italian and English literary production from that period as culturally and stylistically comparable, it

is more difficult to establish a direct equivalent in the Chinese context. Notwithstanding this, the decision was taken to maintain temporal homogeneity as a criterion, since a selection based on other criteria would have entailed similar arbitrariness in the absence of a more objective criterion.

A secondary element that merits consideration is the asymmetry between the languages employed in the experiment. In contrast to the English and Chinese languages, for which there are generative text-to-image artificial intelligence (AI) systems that have been trained on large datasets in the respective native languages, no analogous technology is currently available for Italian. Consequently, Italian literary text has been translated in both cases, while English and Chinese texts are translated in only one of the two cases. This condition gives rise to a potential imbalance, yet it also reflects a realistic situation: it is very unlikely, at least in the medium term, that generative AIs trained equally in all the world's languages will emerge. In this sense, the need to translate from one's mother tongue to the native language of the AI model is already a widespread dynamic and is destined to remain central to the way users interact with generative technologies. Moreover, even the process of translating from text to prompt, while essential, can in some respects compromise the linguistic and artistic value of the original prose.

A further limitation of the study is the qualitative approach adopted for the analysis of the results. Although an a priori theoretical evaluation framework was defined, it was only intended to identify the main compositional elements on which the analysis would subsequently be based. This methodological approach permits a rich and interpretative reading of the images generated, but is not based on a quantitative evaluation grid constructed an a priori on specific and standardized criteria. In other words, the criteria observed emerge during the analysis in a qualitative manner, making it difficult to replicate the study comparatively. However, this limitation is partly justified by the exploratory nature of the work, whose primary objective is to investigate the feasibility and relevance of a comparative study between language, culture, and generative AI. Recent studies by Lindsey et al. (2025) highlight how a bottom-up exploratory approach may be preferable in such contexts, which are still relatively unexplored, in order not to limit the field of investigation and to encourage the development of future research in which these results will pave the way for more systematic and controlled studies.

These will involve investigating a larger number of texts and will also use an ex ante evaluation grid, which could be based on shared linguistic, semiotic and



⑦
Immagine tratta dal brano de I Malavoglia
realizzata da Midjourney a partire dal prompt
in lingua inglese.
*Image taken from the part from I Malavoglia
created by Midjourney based on the English prompt.*

⑧
Immagine tratta dal brano de I Malavoglia
realizzata da JiMeng a partire dal prompt in
lingua cinese.
*Image taken from the part from I Malavoglia
created by JiMeng based on the prompt in Chinese.*

visual parameters. Subsequent to the verification of the hypothesis that generative AIs directly or indirectly reflect cultural specificities learned during training, it will be possible to progress to a second phase of research. This will entail not only the recognition of the presence of these influences, but also an understanding of their nature and functioning. Furthermore, it will be necessary to question how these differences may affect the universality, representativeness, and fairness of the content produced by the models.

REFERENCES

- Arcodia, G. F., & Basciano, B. (2018). *The construction morphology analysis of Chinese word formation*. In *The construction of words: Advances in Construction Morphology* (pp. 219-253).
- Barthes, R. (1977). *Image-music-text*. Fontana Press.
- Boroditsky, L. (2011). How language shapes thought. *Scientific American*, 304(2), 62-65. <http://www.jstor.org/stable/26002395>
- Burlando, F., Chen, B., & Casiddu, N. (2025). *Dataset for the "Creativity and culture in the age of generative AI. The role of cultural specificity in the design of AI-generated content" paper* [Data set]. Zenodo. <https://doi.org/10.5281/zenodo.15761085>
- Campbell, L. (2018). How many language families are there in the world? *Anuario del Seminario de Filología Vasca "Julio de Urquijo"*, 52(1-2), 133-152. <https://doi.org/10.1387/asju.20195>
- Comrie, B. (1989). *Language universals and linguistic typology: Syntax and morphology*. University of Chicago Press.
- DeFrancis, J. (1989). *Visible speech: The diverse oneness of writing systems*. University of Hawaii Press.
- Emily, B. (1847). *Wuthering Heights*. Thomas Cautley Newby.
- Everett, C. (2023). *A myriad of tongues: How languages reveal differences in how we think*. Harvard University Press.
- Gombrich, E. H., Hochberg, J., Black, M., & Fontana, L. (2002). *Arte, percezione e realtà. Come pensiamo le immagini*. Einaudi.
- Harari, Y. N. (2015). *Homo Deus: A brief history of tomorrow*. Ballantine Books.
- Itten, J. (1961). *The art of color: The subjective experience and objective rationale of color*. Wiley.
- Jiang, Q., Gao, Z., & Karniadakis, G. E. (2025). DeepSeek vs. ChatGPT vs. Claude: A comparative study for scientific computing and scientific machine learning tasks. *Theoretical and Applied Mechanics Letters*, 15(3), 100583. <https://doi.org/10.1016/j.taml.2025.100583>
- Kay, P., & Regier, T. (2006). Language, thought and color: Recent developments. *Trends in Cognitive Sciences*, 10(2), 51-54. <https://doi.org/10.1016/j.tics.2005.12.007>
- Lindsey, J., Gurnee, W., Ameisen, E., Chen, B., Pearce, A., Turner, & Batson, J. (2025). *On the biology of a large language model*. Transformer Circuits Thread.
- Rizzolatti, G., & Sinigaglia, C. (2006). *So quel che fai: Il cervello che agisce e i neuroni specchio*. Cortina.
- Schulhoff, S., Ilie, M., Balepur, N., Kahadze, K., Liu, A., Si, & Resnik, P. (2024). *The prompt report: A systematic survey of prompting techniques*. arXiv:2406.06608.
- Tan, S. (ca. 1896). *Lao Tan you ji* [Travels of Old Tan]. Jiangnan Shuju.
- Tao, J., & Wong, J. (2019). The confounding Mandarin colour term "qīng": Green, blue, black or all of the above and more? In *Studies in ethnopragmatics, cultural semantics, and intercultural communication: Minimal English (and Beyond)* (pp. 95-116). Springer Singapore. https://doi.org/10.1007/978-981-32-9979-5_6
- Ting, L. (2022). Towards a semiotics of Chinese characters. *Signs and Media*, 1(2), 111-141. <https://doi.org/10.1163/25900323-12340017>
- Verga, G. (1881). *I Malavoglia*. Treves.
- Wang, X., Wang, Y., & Qiu, X. (2025). How to talk to AI: The role of preset prompt language styles in shaping conversational experience. *International Journal of Human-Computer Interaction*, 41(12), 7763-7778. <https://doi.org/10.1080/10447318.2024.2400398>
- Whorf, B. L. (2012). *Language, thought, and reality: Selected writings of Benjamin Lee Whorf*. MIT Press.
- Xie, P. (2023). Research on the aesthetic value and artistic characteristics in Chinese painting. *Art and Performance Letters*, 4, 51-55. <http://dx.doi.org/10.23977/artpl.2023.040410>

BIO

Francesco Burlando

Designer e PhD, è assegnista di ricerca presso il Dipartimento di Architettura e Design dell'Università di Genova (Italia) e Visiting Professor presso la Beijing University of Chemical Technology. La sua ricerca si concentra sullo studio e l'applicazione delle tecnologie, con particolare riferimento all'intelligenza artificiale e alla robotica, attraverso un approccio more-than-human.

Designer and PhD, is a Research Fellow at the Department of Architecture and Design, University of Genoa (Italy) and visiting professor at the Beijing University of Chemical Technology. His research focuses on the study and application of technologies, with particular reference to AI and robotics, through a more-than-human-centered approach.

Boyu Chen

Designer, è dottorando nell'ambito del Programma di Dottorato Nazionale italiano. Studia presso il Dipartimento di Architettura e Disegno Industriale dell'Università della Campania Luigi Vanvitelli e il Dipartimento di Architettura e Design dell'Università di Genova. La sua ricerca si concentra principalmente sul Design Inclusivo, in particolare sulla navigation experience e sulla walkability.

Designer, is a Ph.D. candidate in the National Doctoral Program of Italy. He studies at the Department of Architecture and Industrial Design at the University of Campania Luigi Vanvitelli and the Department of Architecture and Design at the University of Genoa. His research primarily focuses on Inclusive Design, particularly navigation experience and walkability.

Niccolò Casiddu

Professore Ordinario presso il Dipartimento di Architettura e Design dell'Università di Genova (Italia), nonché Visiting Professor presso la Beijing University of Chemical Technology (Cina). La sua ricerca si concentra sull'accessibilità e l'inclusione in ambienti e prodotti, in particolare sulla facilitazione dell'interazione tra tecnologia e utenti.

Full professor of the Department of Architecture and Design, Polytechnic School of the University of Genoa (Italy), as well as a visiting Professor at the Beijing University of Chemical Technology (China). His research focuses on accessibility and inclusion in environments and products, particularly on facilitating the interaction between technology and users.

ACKNOWLEDGEMENTS

Benché il paper sia il risultato del lavoro congiunto di tutti gli autori, si prega di considerare più nello specifico quanto segue. I paragrafi "Introduzione", "Metodologia" e "Limiti e prospettive future" sono da attribuirsi a F.Burlando, i paragrafi "Procedimento" e "Discussione e conclusioni" a B.Chen. Il lavoro è stato svolto sotto la supervisione di N.Casiddu.

Although the paper is the result of the joint work of all the authors, please consider the following details. The sections "Introduction," "Methodology," and "Limitations and future perspectives" are attributable to F. Burlando, while the sections 'Procedure' and "Discussion and conclusions" are attributable to B. Chen. The work was carried out under the supervision of N. Casiddu.