

## Docenti cinesi di italiano LS e valutazione di task scritti: tra correttezza formale e adeguatezza funzionale

Foreign language teaching in Chinese universities is described as predominantly teacher-centred, memorization-based and focused on the development of formal accuracy rather than of communicative competence. In the present study, six expert teachers of Italian as foreign language were trained to the use of two different rating scales for assessing writing skills of 156 Chinese university students, learners of Italian. A corpus of descriptions, narrations and argumentations was collected and all the texts produced were evaluated using both the parameters of a Traditional scale (language use, morphosyntactic correctness, vocabulary, spelling and punctuation) and those of a Functional Adequacy (FA) scale (content, task requirements, comprehensibility and coherence and cohesion). First results confirm that the FA scale is a reliable tool for assessing the ease of use and the appropriateness of written productions in foreign language, independently from their morphosyntactic correctness. This evaluation tool can be adopted even by Chinese teachers when adequately trained.

### 1. *Introduzione*

L'insegnamento della lingua italiana nelle Università e negli Istituti di Cultura in Cina inizia negli anni Cinquanta del secolo scorso e vive fasi alterne, che rispecchiano l'andamento delle relazioni diplomatiche con l'Italia (Oneto, 1998). È dal 2006, con l'avvio del Progetto Marco Polo, che cresce esponenzialmente l'interesse verso la lingua italiana, insegnata e appresa oggi in 33 istituti universitari cinesi (Maugeri & Scolaro, 2015).

Benché negli ultimi decenni vi sia stata un'apertura verso un approccio glotto-didattico di stampo comunicativo, l'insegnamento delle lingue straniere e, quindi, anche dell'italiano LS in Cina è tuttora prevalentemente centrato sugli aspetti formali e i metodi più diffusi rimangono quello grammaticale-traduttivo e quello audiovisivo (Consalvo, 2012). In questo contesto, lo sviluppo delle abilità comunicative, soprattutto orali, è tralasciato o affidato esclusivamente ai docenti madrelingua, a vantaggio delle abilità di lettura, scrittura e comprensione, esercitate attraverso un approccio allo studio prevalentemente mnemonico (Hu, 2002).

Il docente assume nelle classi di lingua un ruolo dominante e direttivo, ha il pieno controllo dell'interazione e gode della stima e del rispetto dei suoi studenti, i quali non sono incoraggiati a partecipare attivamente alle lezioni ma vivono una

---

<sup>1</sup> Università degli studi Napoli "L'Orientale".

costante preoccupazione di “perdere la faccia”: «*They are afraid of being ashamed, they are afraid of asking foolish questions [...] They pay too much attention to other's opinion of them [...]*» (Jin & Cortazzi, 1996: 195).

Ne consegue che, per i numerosi studenti cinesi che raggiungono le Università italiane, alle difficoltà di carattere puramente linguistico nell'approccio alla lingua italiana, già osservate in numerose precedenti ricerche (Banfi, 2003; Della Putta, 2008; Bonvino & Rastelli, 2011; De Meo & Vitale, 2021), si affiancano anche difficoltà di natura più genericamente culturale, legate alle differenti convenzioni comunicative, agli atteggiamenti, alle routine e alle diverse aspettative nel contesto della classe, che possono erroneamente essere date per scontate sia dai docenti sia dagli studenti, con il risultato di un processo di insegnamento/apprendimento poco efficace e scarsamente motivante (Mariani, 2006).

In Italia e in generale all'estero, gli studenti cinesi ampliano la propria “cultura di apprendimento” (Mariani, 2007), approcciando nuove metodologie di insegnamento e anche di valutazione delle competenze.

Alla luce di queste considerazioni, nel presente studio si vuole verificare la possibilità di una formazione di docenti cinesi esperti di italiano LS sul tema della valutazione della competenza scritta dei propri studenti universitari, attraverso l'applicazione e il confronto di una scala di valutazione “tradizionale”, più attenta agli aspetti formali, e una scala di valutazione “globale”, focalizzata sull'adeguatezza funzionale del testo. Lo studio ha inteso verificare anche se la variabile legata a un percorso formativo condotto in parte in Italia dagli studenti possa incidere in qualche modo sui giudizi ottenuti.

## 2. *Lo studio*

### 2.1 Studenti e corpus

Sono stati coinvolti nella ricerca 156 studenti cinesi di italiano come lingua straniera dell'Università degli Studi Internazionali di Xi'an e dell'Università di Lingue Straniere di Tianjin, suddivisi nei seguenti gruppi:

- 54 studenti iscritti al secondo anno (II-cin), con un percorso di formazione linguistica di un anno e tre mesi, condotto esclusivamente in Cina;
- 53 studenti iscritti al terzo anno (III-ita), con due anni di studio in Cina e tre mesi di soggiorno in Italia;
- 49 studenti iscritti al quarto anno, di cui 13 con un percorso di studio della lingua italiana di tre anni e tre mesi solo in Cina (IV-cin) e 36 con un percorso della stessa durata ma con un intero anno di studio (il terzo) svolto in Italia (IV-ita).

A ciascuno studente, nell'arco di tre settimane e nel contesto del regolare corso di lingua italiana in Cina, sono stati somministrati tre diversi task di produzione scritta: un task descrittivo, in cui era richiesto di descrivere una fotografia raffigurante diversi soggetti in un contesto scolastico; un task narrativo, che prevedeva la

trasposizione nella scrittura di una breve storia illustrata; un task argomentativo, che proponeva di enumerare i pro e i contro di una dieta vegetariana.

I compiti di produzione scritta sono stati svolti in classe, in modalità manoscritta e senza l'ausilio di dizionari o altri strumenti multimediali.

In totale sono stati, quindi, raccolti 468 testi: 156 descrizioni, 156 narrazioni e 156 argomentazioni.

## 2.2 Docenti-valutatori e scale di valutazione

Hanno partecipato alla ricerca sei docenti cinesi esperti di italiano LS, con oltre 10 anni di esperienza professionale, i quali si sono occupati sia della somministrazione dei task nei diversi gruppi classe, sia della successiva fase di valutazione delle produzioni scritte.

Prima dell'avvio della valutazione, i sei docenti coinvolti nello studio hanno seguito un laboratorio di formazione della durata di 6 ore (3 incontri di 2 ore), mirato alla presentazione e all'analisi di due diverse scale di valutazione, una a loro più familiare, attenta agli aspetti formali del testo scritto, e una più innovativa, centrata sull'adeguatezza funzionale del testo.

La prima scala, che sarà qui definita della Valutazione Standard (V-ST), è stata elaborata a partire dalle scale attualmente utilizzate nelle prove di produzione scritta erogate dagli Enti certificatori per l'italiano L2 e già utilizzata in studi precedenti (De Meo, Maffia & Vitale, 2019). Si riportano di seguito i diversi parametri della scala V-ST, a ciascuno dei quali era possibile attribuire una valutazione da 1 a 10 punti<sup>2</sup>:

- *Appropriatezza Testuale (AT)*: aderenza della *performance* del soggetto rispetto all'input fornito, coerenza dei concetti sviluppati e adeguatezza dei meccanismi coesivi usati, appropriatezza della lingua utilizzata rispetto al contesto pragmatico di uso della stessa;
- *Correttezza Morfosintattica (CM)*: rispetto delle regole che riguardano le forme e gli usi degli elementi di una lingua (accordi, concordanze, modi e tempi verbali, posizione dei costituenti frasali, presenza/assenza di preposizioni ecc.);
- *Appropriatezza Lessicale (AL)*: uso di un repertorio lessicale che permette all'apprendente di gestire la propria produzione adeguandosi all'argomento del testo e al registro richiesto dalla situazione comunicativa, efficacia e precisione terminologica;
- *Ortografia e Punteggiatura (OP)*: rispetto delle convenzioni ortografiche dell'italiano standard, compreso il criterio di impaginazione come uso funzionale dello spazio pagina.

<sup>2</sup> Per garantire la comparabilità tra le valutazioni ottenute con le due scale, a ogni criterio è stato attribuito lo stesso valore in termini di punteggio sebbene solitamente esso andrebbe calibrato per rapporto al livello di competenza.

La seconda scala proposta ai docenti cinesi è stata quella dell’Adeguatezza Funzionale (V-AF), elaborata da Vedder e colleghi (Kuiken & Vedder, 2011; 2014; 2016; Vedder, 2016) e applicata di recente in diverse ricerche sulla scrittura in italiano L2 (Del Bono, 2019; Orrù, 2019; De Meo, Maffia & Vitale, 2019). Costruita a partire dalle massime conversazionali di Grice e applicabile sia a produzioni orali sia scritte, è una scala che tenta di rilevare il grado di efficacia comunicativa e di appropriatezza pragmatica di un testo: l’adeguatezza funzionale è considerata la “felice” esecuzione del compito ed è misurata in maniera indipendente dagli aspetti formali di correttezza morfosintattica. Tale scala è composta dalle seguenti quattro dimensioni, a ciascuna delle quali, come nella V-ST, era possibile attribuire un giudizio da 0 a 10 punti:

- *Contenuto (CO)*: adeguatezza del numero e del tipo di unità informative espresse nel testo, loro pertinenza e rilevanza, indipendentemente dai requisiti del task;
- *Requisiti del Task (RT)*: rispondenza ai requisiti specifici del task da svolgere, relativi al genere testuale, all’atto linguistico, alla macro-struttura e al registro del testo;
- *Comprensibilità (COMPR)*: misura dello sforzo richiesto al ricevente per capire lo scopo e le idee espresse nel testo;
- *Coerenza e Coesione (CC)*: presenza/assenza di anafore, connettivi, salti logici e argomenti non collegati, ripetizioni ecc.

A conclusione del percorso di formazione, a ciascuno dei docenti è stato chiesto di procedere alla valutazione dei testi scritti del corpus precedentemente raccolto, utilizzando entrambe le scale ma applicandole nella valutazione di testi diversi. Al fine di ridurre al minimo le variabili soggettive e garantire la validità e l’affidabilità del processo valutativo, sono state svolte sessioni di standardizzazione dei risultati in fase iniziale, intermedia e finale.

Durante gli incontri sono stati effettuati dei controlli incrociati sui punteggi dati da diversi docenti ad alcuni testi scelti come campione. Si è riflettuto in gruppo sulle scale di valutazione e sulle procedure di attribuzione dei punteggi e, qualora i risultati tra i valutatori differissero, si è cercato di raggiungere un accordo di valutazione.

### 3. Risultati

In tab. 1 sono riportati i valori medi delle valutazioni attribuite dai docenti cinesi con le due diverse scale a tutte le produzioni scritte dei diversi gruppi di studenti.

Al fine di valutare la significatività delle differenze osservate tra i giudizi ottenuti con le due scale di valutazione nei diversi gruppi di studenti, è stato applicato un *test T* per campioni appaiati con il software R. La soglia di significatività è stata fissata a 0.05.

Tabella 1 - *Valori medi delle valutazioni attribuite con le due scale a tutte le produzioni, per gruppo di studenti*

	V-ST	V-AF
II-cin	6.5	7.4
III-ita	6.5	7.2
IV-cin	6.6	7.4
IV-ita	6.9	7.4

È immediatamente evidente come l'applicazione della scala V-AF produca giudizi più positivi rispetto alla V-ST in ciascuno dei gruppi (II-cin e III-ita:  $p < 0.001$ ; IV-cin:  $p = 0.006$ ; IV-ita:  $p = 0.004$ ), sebbene in tutti i casi il punteggio medio ottenuto superi il valore di sufficienza (6 pt). Mentre si riscontrano, inoltre, risultati in media omogenei con la V-AF, gli studenti iscritti al IV anno, in particolare del gruppo IV-ita, sono quelli che ricevono le migliori valutazioni con la V-ST. Questo dato sembra indicare che la maggiore durata del percorso formativo e l'immersione e lo studio in contesto L2 influenzano positivamente la correttezza formale nella scrittura in lingua seconda. Nel caso, invece, del gruppo III-ita, i tre mesi trascorsi in Italia sono forse un periodo troppo breve perché influisca positivamente e in misura tangibile sulle competenze linguistiche, anche dato il fisiologico periodo di adattamento al nuovo contesto.

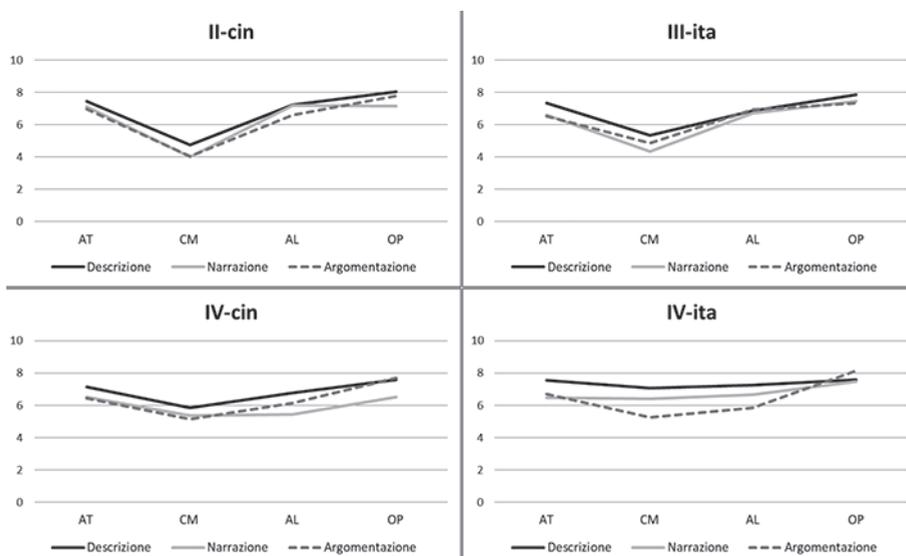
I valori medi suddivisi per task, riportati in tab. 2, rivelano come il compito descrittivo sia quello che registra in media i punteggi più alti, sia con la V-ST (da 6.9 a 7.4) sia con la V-AF (da 7.6 a 7.9). I risultati ottenuti con gli altri due task, invece, risultano meno costanti, muovendosi nel range di 1 punto con la V-ST e di 0.7 punti con la V-AF.

Tabella 2 - *Valori medi delle valutazioni attribuite con le due scale, per task e gruppo di studenti*

	V-ST			V-AF		
	descr.	narr.	argom.	descr.	narr.	argom.
II-cin	6.9	6.4	6.4	7.6	7.3	7.2
III-ita	6.9	6.3	6.4	7.6	7.0	6.9
IV-cin	6.9	6.0	7.0	7.9	6.8	7.5
IV-ita	7.4	6.8	6.5	7.9	7.1	7.3

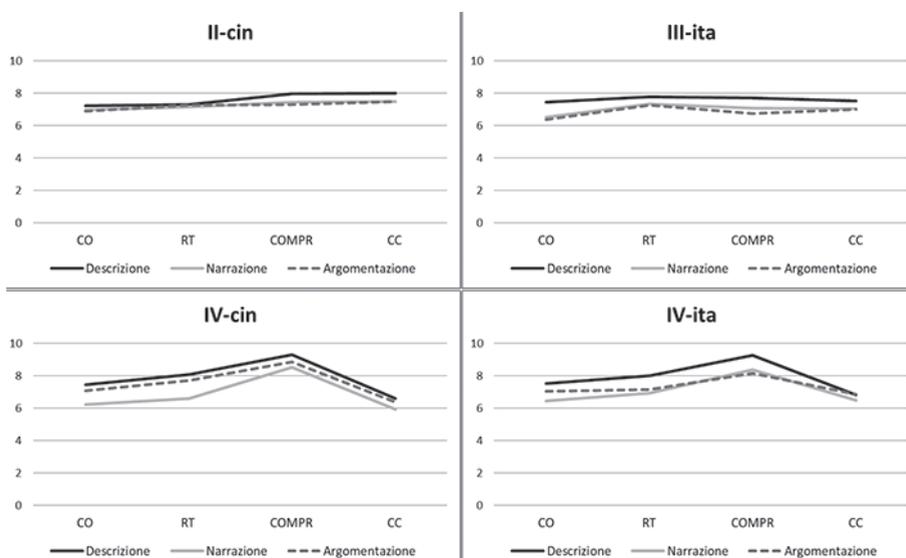
Se si osservano i giudizi medi ottenuti con la scala V-ST per ciascun parametro, nei tre task e nei quattro gruppi di studenti, riportati nella fig. 1, si nota che il parametro che influisce più negativamente sui valori medi è quello relativo alla Correttezza Morfosintattica, che raggiunge la soglia dei 6 punti esclusivamente nelle descrizioni e nelle narrazioni del gruppo IV-ita. I punteggi più alti sono spesso ottenuti in tutti i gruppi con il parametro dell'ortografia e punteggiatura.

Figura 1 - Valori medi delle valutazioni attribuite con la scala V-ST, per parametro, task e gruppo di studenti



Nella fig. 2 sono riportati, invece, i giudizi attribuiti a ciascuna dimensione della V-AF, nei tre task e in tutti i gruppi di studenti

Figura 2 - Valori medi delle valutazioni attribuite con la scala V-AF, per parametro, task e gruppo di studenti



Se nelle produzioni del II e del III anno non si evidenziano sostanziali differenze né tra i quattro parametri né tra i tre task, nelle valutazioni dei due gruppi del IV anno

i dati sono più variati: la Comprensibilità raggiunge e supera un punteggio medio pari a 9, mentre le dimensioni del Contenuto e della Coerenza e Coesione ottengono i giudizi più bassi, avvicinandosi alla soglia del 6, in particolare nelle narrazioni.

#### *4. Discussione e conclusioni*

Nella presente ricerca si è voluto sperimentare con un gruppo di docenti cinesi di italiano LS l'uso di due scale di valutazione per produzioni scritte di diversa natura: la scala V-ST, in cui gli aspetti formali hanno un peso notevole sul giudizio totale, e la scala V-AF, con cui si misura l'adeguatezza funzionale del testo, prescindendo dalla correttezza morfosintattica e ortografica. Le due scale sono state applicate alla valutazione di un corpus di 468 testi, equamente divisi in descrittivi, narrativi e argomentativi, prodotti da studenti cinesi di italiano LS, iscritti in due Università cinesi e con percorsi di formazione linguistica differenziati per durata e contesto di apprendimento.

I risultati ottenuti hanno evidenziato come con la scala V-AF i docenti abbiano attribuito ai diversi testi in media giudizi più alti, per tutti i gruppi di studenti. A incidere negativamente sulle valutazioni con la scala V-ST è stato, in particolare, il parametro relativo alla Correttezza Morfosintattica, valutata in media con un punteggio sotto la sufficienza nei gruppi II-cin, III-ita e IV-cin. Il gruppo IV-ita, costituito da studenti con un anno di formazione linguistica e di studio svolto in Italia, è stato, invece, valutato in maniera migliore in relazione agli aspetti grammaticali dei testi prodotti. Il parametro Ortografia e Punteggiatura, in linea con l'attenzione posta alle convenzioni della scrittura nel contesto formativo cinese, ha ottenuto in media i punteggi migliori.

I giudizi attribuiti a ciascuna dimensione della scala V-AF sono stati in media omogenei, in particolare nei gruppi II-cin e III-ita, dato che conferma l'indipendenza di tale scala dalla correttezza formale dei testi e, quindi, l'efficacia delle valutazioni stesse. La maggiore durata del percorso di apprendimento linguistico, nei gruppi IV-cin e IV-ita, hanno una ricaduta positiva, secondo i docenti valutatori, sulla dimensione della Comprensibilità, cioè sulla quantità di impegno richiesta al lettore per la comprensione del testo. La variabile legata al soggiorno in Italia non sembra influire in maniera significativa sui risultati ottenuti con la scala V-AF.

Sia con la scala V-ST, sia con la scala V-AF, il task descrittivo, linguisticamente più semplice rispetto alla narrazione e all'argomentazione, è quello che ha ottenuto risultati migliori per tutti i gruppi di studenti.

In conclusione, i docenti cinesi hanno riconosciuto e apprezzato le potenzialità di una scala di valutazione non ancorata rigidamente alla dimensione della correttezza formale, che ha permesso di valorizzare maggiormente le competenze comunicative degli allievi, pur nello sforzo complessivamente maggiore richiesto per portare a conclusione il processo di valutazione.

## Bibliografia

- BANFI E. (a cura di) (2003), *L'italiano L2 di cinesi. Percorsi acquisizionali*, Franco Angeli, Milano.
- BONVINO E. & RASTELLI S. (a cura di) (2011), *La didattica dell'italiano a studenti cinesi e il progetto Marco Polo*, Atti del XV seminario AICLU Roma, 19 febbraio 2010, University Press, Pavia.
- CONSALVO G. (2012), L'influenza delle differenze culturali nell'apprendimento dell'italiano L2 da parte di sinofoni: il metodo di studio e l'insegnamento della lingua straniera in Cina, in *Italiano LinguaDue* 1: 31-45.
- CORTAZZI M. & JIN L. (1996), Cultures of learning: Language classrooms in China, in COLEMAN H. (a cura di) *Society and the Language Classroom*, Cambridge University Press, Cambridge: 169-206.
- DE MEO A. & VITALE M. (2020), *Aspetti prosodici dell'acquisizione dell'italiano da parte di sinofoni*, Aracne, Roma.
- DE MEO A., MAFFIA M. & VITALE G. (2019), La competenza scritta in italiano L2 di apprendenti vulnerabili. Due scale di valutazione a confronto, in *Educazione Linguistica. Linguistic Education ELLE* 8 (3): 637-653.
- DEL BONO F. (2019), Aspetti pragmatici nella valutazione di testi scritti: uno studio sull'adeguatezza funzionale in italiano L2, in NUZZO E. & VEDDER I. (a cura di), *Lingua in contesto: la prospettiva pragmatica*, Studi AitLA 9, Officinaventuno, Milano: 231-244.
- DELLA PUTTA P. (2008), Insegnare l'italiano ai sinofoni: contributi acquisizioni, tipologici e glottodidattici, in *Studi di glottodidattica* 2: 52-67.
- HU G. (2002), Potential Cultural Resistance to Pedagogical Imports: The Case of Communicative Language Teaching in China, in *Language Culture and Curriculum* 15(2): 93-105.
- KUIKEN F. & VEDDER I. (2011), Task complexity and linguistic performance in L2 writing and speaking: the effect of mode, in ROBINSON P. (a cura di), *Second language task complexity: researching the cognition hypothesis of language learning and performance (Task-based language teaching, 2)*, John Benjamins, Amsterdam: 91-104.
- KUIKEN F. & VEDDER I. (2014), Raters' decisions, rating procedures and rating scales, in *Language testing* 31(3): 279-284.
- KUIKEN F. & VEDDER I. (2017), Functional adequacy in L2 writing: Towards a new rating scale, in *Language Testing* 34(3): 321-336.
- MARIANI L. (2006), Le diversità culturali negli stili cognitive e comunicativi. Dal conflitto al compromesso alla sinergia, in *Lingua e Nuova Didattica* 5: 90-99.
- MARIANI L. (2007), Learning style across cultures, in *Perspectives. A journal of TESOL-Italy* 2: 343-348.
- MAUGERI G. & SCOLARO S. (2015), Segnali di qualità nell'insegnamento dell'italiano in Cina: un caso esemplare, in *In. It* 31-32: 17-24.
- ONETO C. (1998), L'insegnamento dell'italiano in Cina, in *Mondo Cinese* 97 (1): 55-72.
- ORRÙ P. (2019), Misurare l'adeguatezza funzionale in testi scritti di apprendenti di italiano L2, in *Italiano LinguaDue* 11 (1): 45-58.
- VEDDER I. (2016), Il ruolo dell'adeguatezza funzionale nelle produzioni scritte in lingua seconda: proposta per una scala di valutazione, in VEDDER I. & SANTORO E. (a cura di), *Pragmatica e interculturalità in italiano lingua seconda*, Franco Cesati Editore, Firenze: 79-92.