

L'IMPATTO E LA VALIDITÀ DEI PROCESSI DI VALUTAZIONE DELL'ITALIANO L2. RIFLESSIONI SUGLI APPRENDENTI SINOFONI IN ACCESSO AL SISTEMA DELLA FORMAZIONE SUPERIORE IN ITALIA

*Sabrina Machetti, Paola Masillo*¹

1. INTRODUZIONE

La valutazione della competenza linguistico-comunicativa in italiano L2 di studenti internazionali in accesso al sistema della formazione superiore in Italia rappresenta ad oggi un tema poco affrontato nell'ambito delle scienze linguistiche applicate. Tranne rare eccezioni (Grego Bolli *et al.*, 2017; 2017; Machetti, 2020), l'attenzione continua ad essere dedicata alle dinamiche che caratterizzano i percorsi di apprendimento e di insegnamento linguistico e all'impatto che tali percorsi hanno nel più generale percorso formativo degli studenti internazionali (Bagna, Machetti, 2015; Bagna *et al.*, 2017; Rastelli, Bagna, 2020). La valutazione della L2, quando menzionata, viene generalmente vista ed analizzata nella sua funzione meramente accessoria e/o strumentale, nella maggior parte dei casi con l'obiettivo di evidenziare le criticità dei risultati che questa produce (Rugge, 2019).

Procedendo dall'assunto secondo cui la valutazione, in quanto processo, riveste un ruolo centrale nei più ampi processi di apprendimento / insegnamento linguistico ed ha un impatto che segue il loro sviluppo e ne può condizionare, in positivo e in negativo, gli esiti (Council of Europe, 2001; 2020), questo contributo analizza l'impatto che l'uso di un particolare strumento di valutazione dell'italiano L2 nella fase di accesso degli studenti internazionali alle università italiane e all'Alta Formazione Artistica e Musicale (AFAM) ha per gli stessi studenti e, più in generale, per il loro percorso formativo in Italia. Il contributo fa riferimento al caso degli studenti di origine cinese che, prima dell'ingresso nel sistema della formazione superiore, accedono al programma MP/T - Marco Polo e Turandot (Bagna *et al.*, 2017)² e tratta il caso degli esami CILS – Certificazione dell'Italiano come Lingua Straniera, Università per Stranieri di Siena – di livello B1 e B2 destinati a questa particolare tipologia di pubblico³.

L'analisi dei dati relativi all'impatto si lega al più generale processo di validazione degli esami e alla necessità di fornire per essi, ciclicamente, evidenze di validità ed affidabilità

¹ Università per Stranieri di Siena.

Il contributo è frutto della ricerca e della riflessione di entrambe le autrici, che insieme hanno redatto la Bibliografia. Sono di Sabrina Machetti i §§ 1, 2, 4.3, 5; di Paola Masillo i §§ 3, 3.1, 3.2, 4.1, 4.2.

² Il Programma MP/T ha origine nei primi anni 2000 (2004-2006), promosso dalla Conferenza dei Rettori delle Università Italiane con l'intento di sostenere e stimolare l'accesso e la frequenza da parte di studenti di origine cinese nelle università italiane.

³ La Certificazione CILS è un titolo di studio ufficialmente riconosciuto che misura il grado di competenza linguistico-comunicativa in italiano L2. Il sistema di certificazione si pone in linea con il modello teorico proposto dal *Quadro Comune Europeo di Riferimento per le Lingue* (Council of Europe, 2001) che descrive lo sviluppo della competenza linguistico-comunicativa su una scala di sei livelli. Sull'argomento, si rimanda a Machetti, Vedovelli, in stampa. La Certificazione CILS propone dal 2016 esami di livello B1 e B2 rivolti agli studenti MP/T, somministrati nella sessione del mese di agosto.

(Chaloub-Deville, O'Sullivan, 2020). In relazione a ciò, il contributo illustra la funzione di tali dati, accompagnati da dati quantitativi CILS relativi al biennio 2017-2019, ai fini della revisione del costruito dell'esame CILS destinato a studenti del programma MP/T, già in corso dal 2015 (Bagna, Machetti, 2015).

2. STATO DELL'ARTE

La ricerca dedicata al ruolo che la valutazione dell'italiano L2 ha nel percorso formativo dello studente MP/T conta ad oggi, e come si diceva poco sopra, un numero limitato di studi che ne evidenziano le diverse criticità. In più di un caso, i risultati di tali studi sono stati estrapolati dal contesto in cui questi nascono e si sono andati ad inserire nel più generale contesto degli studenti internazionali in accesso e/o presenti nel nostro sistema universitario (Machetti, 2020). Come è facile immaginare, tale trasferimento non ha avuto altro effetto se non quello di leggere in chiave ancor più negativa un quadro già di per sé abbondantemente critico (Rugge, 2019), in cui lo studente sinofono è più di altri esposto a fenomeni di abbandono pre o post-iscrizione; mostra una media dei voti ampiamente al di sotto della media degli studenti internazionali; evidenzia ritardi nel percorso universitario e, quando lo porta a termine, nei tempi di conseguimento del titolo; è lo studente che più di altri – scenario, questo, poco o per niente analizzato in relazione al contesto della formazione italiana – è esposto a depressione, malattie psicosomatiche, ecc.

Guardando allo scenario internazionale, gli studi riguardanti la presenza in ambito accademico di studenti di origine cinese risultano prevalentemente legati al contesto anglofono e a quello delle università coreane (Qian, Cumming, 2017; Zou, Xu, 2017). Tali studi presentano alcune analogie con quanto registrato per l'Italia, ma appaiono attenti anche ad elementi che, a nostro avviso, possono aiutarci meglio ad interpretare il contesto del nostro studio.

Il primo elemento che accomuna le ricerche condotte su studenti di inglese L2 e coreano L2 a quelle condotte sugli studenti del programma MP/T che apprendono l'italiano è che la L2 rappresenta un elemento di fondamentale criticità, ma non l'unico elemento che mette in crisi l'accesso e il percorso universitario degli studenti di origine cinese. Alle difficoltà contribuirebbe infatti

- la presenza di barriere più generalmente comunicative e culturali, aumentata dall'atteggiamento dei docenti e dei pari;
- la debole o del tutto assente *background knowledge*, intesa come consapevolezza di che cosa il mondo accademico richiede a ciascuno studente, assenza che spesso si accompagna ad un livello di ansia particolarmente elevato;
- la scarsa attenzione da parte dei docenti e di altri addetti alla formazione accademica alla *academic self-efficacy* e al tema della *effort regulation*. Le ricerche condotte in questa direzione evidenziano infatti come la media dei voti degli studenti sinofoni sia solo in minima parte condizionata dalla *language proficiency* degli stessi studenti in coreano L2 o in inglese L2⁴.

In generale, gli studi citati conducono a relativizzare il peso della *language proficiency* nel percorso accademico degli studenti sinofoni in contesto internazionale e ci paiono

⁴ Le medesime ricerche prendono in esame gli esiti della stessa popolazione di studenti negli esami IELTS e TOEFL ed evidenziano misure di correlazione modeste o discordanti con la media dei voti riportata dagli studenti in ambito universitario.

perfettamente in linea con quanto, a più riprese, proposto per l'italiano di sinofoni (Banfi, 2003; Della Putta, 2008; Rastelli, 2011)⁵.

La riflessione su questo aspetto impone il riferimento al concetto di *language proficiency* e di *academic language proficiency*. In valutazione, il concetto di *language proficiency* rappresenta la base per l'interpretazione dei dati e delle misure che un test di lingua fornisce (Messick, 1989). Si tratta dunque di un concetto centrale della valutazione, ma di certo anche del concetto più complesso da definire (Chapelle *et al.*, 2008). Di tutte le definizioni a nostra disposizione, non ne esiste una che possa dirsi migliore di un'altra, ma forse una più o meno ampia di un'altra, sulla quale il consenso non è comunque mai scontato (Jin, 2022). Come noto, gli esami CILS (Barni *et al.*, 2009; Machetti, Vedovelli, in stampa) fondano il loro costruito su un modello di *language proficiency* risultante dall'integrazione tra conoscenze linguistiche e strategie linguistico-comunicative necessarie per svolgere un qualsivoglia compito comunicativo in qualsivoglia dominio della comunicazione, aderendo in ciò al modello proposto dal *CEFR* (Council of Europe, 2001; 2020). Tale modello, attento alle caratteristiche del contesto della comunicazione, sembrerebbe rendere poco utile la proposta di test che lavorano su un modello di *academic language proficiency*.

Come già proposto in altra sede (Machetti, 2020), i diversi tentativi, volti a definire che cosa sia la competenza linguistico-comunicativa in italiano necessaria ad uno studente straniero per gestire in maniera autonoma la comunicazione in contesto accademico entro i sotto-contesti che lo costituiscono, sembrano fallire nel non iscriverne tale competenza lungo un *continuum* di conoscenze e competenze linguistiche, extra-linguistiche e più generalmente semiotiche. Un *continuum* che si estende e distende entro uno spazio linguistico che diviene esso stesso «modello della competenza individuale e degli usi collettivi, non più la prima intesa in senso 'verticale' come il progressivo avvicinamento ai livelli linguistici considerati più alti e avanzati entro un'unica norma linguistica funzione dei valori del gruppo dominante, ma come la capacità di scegliere gli usi e perciò le norme di riferimento in rapporto alle condizioni contestuali in cui avviene l'evento linguistico» (Machetti, Vedovelli, 2012: 188). Da tali tentativi deriva anche l'equivoco secondo cui la lingua italiana, in contesto accademico, assolva per gli studenti internazionali una mera funzione strumentale, dunque consista semplicemente in una chiave di accesso a determinati contenuti e a determinati contesti della comunicazione; come deriva la persistente frammentarietà di una riflessione sull'effettiva fruibilità, sempre in contesto accademico, di questi stessi contenuti e sulle modalità della loro messa a disposizione per l'intera comunità studentesca, italiana o straniera che sia.

La ricerca svolta su altre lingue e in altri contesti fornisce in proposito indicazioni interessanti ma non prive di problematicità, in particolare nella messa a confronto di test di lingua generalisti – come le certificazioni linguistiche – e test mirati o per scopi speciali, orientati cioè verso lo studente e i bisogni linguistici che questi matura in relazione al contesto accademico o ad altri specifici contesti, in cui dovrà agire e muoversi linguisticamente (Knoch, Elder, 2016; Elder, 2017). Più precisamente, nella comunità scientifica internazionale non esiste univocità di giudizio in relazione alla preferenza, basata su argomenti di maggiore validità ed affidabilità, verso l'uso di test mirati, ad esempio test di lingua per lo studio, rispetto a test di competenza linguistica generalisti. In aggiunta a ciò, nel contesto accademico italiano, i test standardizzati di competenza linguistico-comunicativa sviluppati per l'italiano L2, le cosiddette certificazioni, hanno nel tempo fornito buone evidenze in relazione alla validità di contenuto e costruito, che nei

⁵ In particolare, la proposta di Rastelli (2011) relativa alla distinzione tra grammatica e processing e al fattore distanza tipologica tra le lingue appare fondante per la necessità di un approccio innovativo della didattica dell'italiano rivolta a questa particolare tipologia di studenti, sia in relazione ai tempi che ai contenuti.

livelli B contemplano quanto previsto dalla comunicazione nel dominio educativo e nei contesti della formazione accademica. Più in generale, le certificazioni hanno permesso e permettono di fornire per i livelli della non completa autonomia comunicativa come per i livelli della piena autonomia una fotografia ampia e comparabile dello stato della competenza linguistico-comunicativa in possesso dello studente in un dato momento del proprio sviluppo (Barni, Machetti, 2022). Tale fotografia e le informazioni che da essa derivano possono essere usate per i più diversi scopi, che vanno anche al di là di quelli più legati alla spendibilità sociale che ad ogni certificazione pertiene (Machetti, Vedovelli, in stampa) e che hanno un impatto diretto sui processi di apprendimento e insegnamento linguistico: diagnostici, di incentivo alla motivazione, di auto-correzione ecc.

Circa il tema dell'impatto della valutazione della L2, la ricerca condotta in questa direzione mira a comprendere gli effetti e le conseguenze che derivano dall'uso della valutazione linguistica nei contesti educativi e nella società. In particolare, e come argomentato da Saville (2000), le prove fornite in sede di valutazione certificatoria possono avere un impatto, positivo o negativo, sul processo formativo e sulla società in generale, impatto che di certo influenza il futuro degli individui che al test si sottopongono e la loro possibilità di successo anche in contesto accademico. Diverse ricerche rilevano inoltre come gli esami, tra cui è possibile collocare anche gli esami di certificazione linguistica, influenzino gli atteggiamenti, il comportamento e la motivazione di insegnanti, apprendenti e genitori (Alderson, Wall, 1993).

Le evidenze necessarie ai fini dell'impatto e della validità del test riguardano, infine, anche la percezione del livello di complessità delle prove e le aspettative nutrite nei confronti dell'esperienza certificatoria da parte di chi al test si sottopone. Tale prospettiva di analisi nasce dalla necessità di riflettere sul requisito più importante di un buon test, secondo quanto sostenuto da Pilliner (1973: 4), ovvero che il test sia vantaggioso dal punto di vista educativo. Tale piano di analisi concorre, come si diceva poco sopra, a mettere in atto il più ampio e complesso processo, continuo e costante, di validazione degli esami di certificazione, in cui per una misurazione del grado di validità dei punteggi diventa utile studiare anche l'impatto che tali prove possono avere su chi ad esse si sottopone (Messick, 1989).

3. METODO

I dati discussi nel contributo sono stati raccolti attraverso la somministrazione dell'esame CILS UNO-B1 e DUE-B2 per studenti MP/T nella sessione Agosto 2017⁶, somministrazione accompagnata da quella di un questionario di valutazione post-esame. Di entrambi gli strumenti si illustreranno le principali caratteristiche nei paragrafi che seguono.

I dati ricavati dalla somministrazione degli esami e da quella del questionario post-esame vengono analizzati ed impiegati ai fini della misurazione dell'impatto dell'esame nel percorso formativo degli studenti e ai fini della validazione degli esami dei livelli interessati.

La validazione segue il *validity argument* proposto da Mislevy *et al.* (2002), in cui la base per l'interpretazione dei punteggi è rappresentata

- a) dal modello di competenza e dal relativo costrutto, che rendono conto della performance linguistica in un ampio ventaglio di contesti della comunicazione;

⁶ La sessione è stata scelta per significatività del numero dei candidati e loro rappresentatività. Si tratta infatti di studenti che hanno svolto da 9 a 11 mesi di formazione in italiano L2 presso l'Università per Stranieri di Siena, e che provengono da 7 diverse aree della Cina.

- b) dai processi e dalle strategie necessarie a definire e adattare la performance ai differenti domini e contesti della comunicazione.

In tale modello, costruito, processi e strategie rappresentano la base per interpretare una serie di dati (*Grounds*), dati che per supportare la dichiarazione finale (*Claim*, ad es. “lo studente ha una competenza adeguata per esprimersi in forma scritta e parlata nei contesti della comunicazione accademica”) – che tiene conto di un principio generalmente valido (*Warrant*) e di informazioni che derivano da basi empiriche di diverso tipo (ad es., l’esperienza del docente che ha lavorato in quella classe con quel dato studente, il giudizio dei valutatori ecc.) – devono anche rispondere in maniera efficace ad eventuali obiezioni (*Rebuttal*, ad es. “il compito a cui è esposto lo studente è troppo elevato per il suo livello di competenza”; “lo studente non ha familiarità con l’argomento” ecc.).

3.1. *Il questionario*

Il questionario di valutazione post-esame è composto da 10 domande ed è stato elaborato in lingua italiana. Le domande sono suddivise in 4 sezioni, come di seguito riportato:

- Sezione 1: Biografia (sesso, età)
- Sezione 2: Formazione linguistica (apprendimento dell’italiano L2)
- Sezione 3: Esame CILS
- Sezione 4: Impatto dell’esame (*face validity*)

Le domande sono in prevalenza a risposta chiusa, con la possibilità per il candidato o la candidata di indicare con una crocetta la risposta tra le opzioni possibili proposte rispetto all’esperienza certificatoria appena conclusa (*È stato difficile l’esame CILS?; Quale parte dell’esame è stata più difficile?; Quale parte dell’esame è stata più facile?; L’esame CILS che hai appena svolto risponde alle tue aspettative?*); le domande riservate alla parte anagrafica e al percorso pregresso di apprendimento della lingua italiana sono a risposta aperta con la possibilità di indicare il numero di anni oppure il numero di mesi di studio della lingua italiana, sia nel proprio Paese di origine sia in Italia (*Età; Per quanto tempo hai studiato la lingua italiana nel tuo Paese?; Da quanto tempo studi la lingua italiana in Italia?*). L’ultima domanda aperta è dedicata alla possibilità per il candidato o la candidata di esprimere i propri commenti, illustrando eventuali aspetti dell’esame che non si aspettava di trovare nel test, relativamente alle abilità testate, alla tipologia di prova utilizzata, o agli argomenti trattati, ecc. Tali aspetti concorrono a definire quanto in letteratura trova espressione nella cosiddetta *face validity*, ovvero la validità apparente o di facciata con cui l’attenzione si concentra sul punto di vista del non esperto, di chi cioè ha appena vissuto l’esperienza diretta della prova, il candidato, ed è chiamato ad esprimere la propria impressione sullo strumento di valutazione con cui si è cimentato (Green, 2013).

Ai fini di una compilazione valida, si è stabilito di formulare le domande in una lingua chiara, con frasi brevi e semplici per entrambi i livelli in esame, al fine di garantire a tutti i candidati, inclusi quelli con un livello di competenza linguistica più basso, la possibilità di comprendere le richieste. In aggiunta, la gestione della somministrazione del questionario post-esame è stata svolta in collaborazione con un tutor di madrelingua cinese, qualora fosse emersa la necessità di intervenire per fornire chiarimenti o spiegazioni sulle domande.

3.2. L'esame CILS

L'esame CILS misura il raggiungimento di un dato livello di competenza linguistico-comunicativa in italiano L2, fondandosi su un costrutto che comprende le abilità linguistiche di comprensione orale (*Ascolto*), comprensione scritta (*Lettura*), produzione e interazione scritta (*Produzione scritta*), produzione e interazione orale (*Produzione orale*), cui si aggiunge una quinta componente che misura la capacità d'uso delle forme e delle strutture della lingua oggetto di valutazione in contesto, ovvero mediante test a completamento o *cloze* a partire da un *input* testuale (*Analisi delle Strutture di comunicazione*). La verifica e la valutazione della competenza in italiano L2, come si evince dalle componenti oggetto della misurazione, chiama in gioco non solo competenze linguistiche, ma include le competenze pragmatiche e sociolinguistiche che insieme concorrono a definire il profilo di competenza in esame (Council of Europe, 2001; 2020).

I livelli oggetto del nostro studio coincidono con la fascia intermedia che descrive i due profili di competenza riconducibili all'*independent user*, ovvero il livello B1 (*Threshold*) e il livello B2 (*Vantage*).

Riprendendo quanto illustrato nelle *Linee Guida CILS* (Barni *et al.*, 2009: 29), in riferimento alla descrizione globale del profilo di competenza, l'esame di livello B1 misura

il livello che attesta la competenza in riferimento al profilo dell'apprendente autonomo del QCER. Verifica le capacità comunicative necessarie per usare la lingua italiana con autonomia e in modo adeguato nelle situazioni più frequenti della vita quotidiana in Italia. L'apprendente con questo livello di competenza è in grado di comunicare in italiano nelle situazioni di tutti i giorni in forma sia scritta sia orale, di comprendere i punti essenziali di un discorso articolato chiaramente in lingua standard, di leggere i testi scritti che incontra più frequentemente nella vita quotidiana. La produzione orale e scritta è comunicativamente efficace, anche se contiene errori.

L'esame B2 si riferisce invece al (ivi: 32)

livello che attesta la piena autonomia della competenza comunicativa in italiano come lingua straniera. Il candidato è in grado di comunicare efficacemente durante un soggiorno in Italia per motivi di studio e di gestire i contatti con la lingua e la cultura italiana anche per motivi di lavoro. L'interazione con i parlanti nativi si sviluppa senza eccessiva fatica e tensione. Il candidato con questo livello di competenza è in grado di comprendere le idee fondamentali di testi anche complessi che si riferiscono sia ad argomenti legati alla realtà della vita quotidiana sia a concetti più astratti. La produzione orale e scritta è comunicativamente efficace, anche se contiene alcuni errori. È il livello minimo di competenza per l'accesso al sistema universitario italiano, per realizzare un ciclo di studi entro un progetto di mobilità breve di studenti [...].

4. DISCUSSIONE DEI RISULTATI

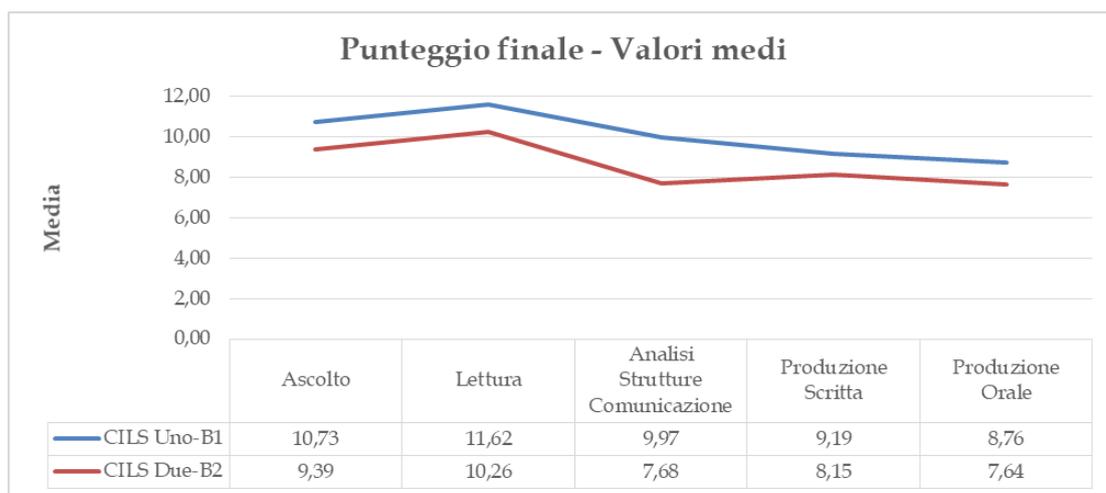
4.1. I risultati del test

Nella sessione Agosto 2017, l'esame CILS UNO-B1 è stato sostenuto da 412 candidati; l'esame CILS DUE-B2 da 101 candidati.

Partendo dall'esito dell'esame, ovvero il superamento o meno e dunque l'ottenimento della certificazione per il livello cui il candidato o la candidata si è sottoposto/a, il 40,3% ha superato le prove di cui l'esame si compone, distribuendosi tra i due livelli come segue: 33% B1 (n. 174) e 6,4% B2 (n. 33).

L'intervallo dei punteggi sulla scala di valutazione per gli esami oggetto del nostro studio è di 100 punti; per ogni singola abilità la scala di valutazione va da 0 a 20 punti. La soglia di sufficienza per ogni abilità coincide con 11 punti (Barni *et al.*, 2009: 18). Il punteggio finale medio ottenuto è di 50,27 punti per il livello B1 e 43,12 punti per il livello B2. Come riportato in dettaglio nel Grafico 1, la media dei punteggi più alta si registra per entrambi i livelli in corrispondenza dell'abilità di Comprensione della lettura (11,62 B1; 10,26 B2). Allo stesso modo, l'abilità che per entrambi i livelli evidenzia la media dei punteggi più bassa è la Produzione orale (8,76 B1; 7,64 B2). Il dato, come vedremo, risulta parzialmente coerente con le percezioni dei candidati raccolte nel questionario con la domanda 7 (*Quale parte dell'esame è stata più difficile?*) e la domanda 8 (*Quale parte dell'esame è stata più facile?*). La componente più complessa si conferma essere la Produzione orale (Bonvino, Rastelli, 2010), mentre la componente meno complessa, mantenendosi sul piano delle abilità scritte, risulta essere quella che il questionario rivela come seconda in base all'ordine di preferenza dei candidati, ovvero la Comprensione della lettura (vs. Produzione scritta, cfr. Grafico 5).

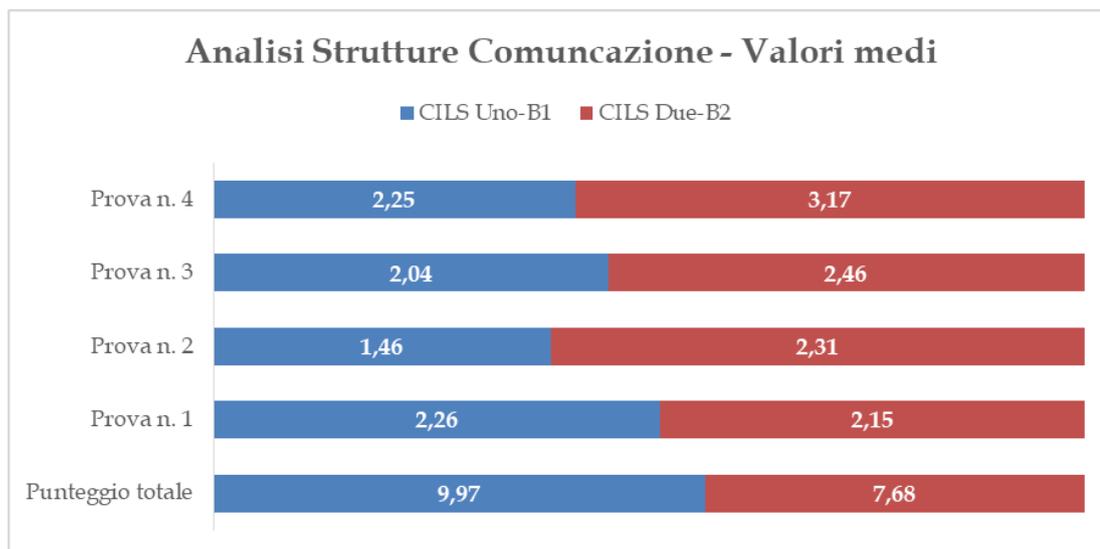
Grafico 1. *Punteggio finale. Media*



Da una lettura dei punteggi grezzi ottenuti per le singole prove della sezione Analisi di Struttura della Comunicazione⁷ (Grafico 2) si ottiene un dato interessante. Per i candidati del livello B1 la prova nella quale incontrano maggiori difficoltà, ottenendo un punteggio più basso (1,46 su 6 punti), è la prova numero 2 dedicata alla capacità di gestire il sistema verbale atteso per il livello in esame: indicativo presente, passato prossimo, imperfetto; infinito presente; imperativo; condizionale presente (Barni *et al.*, 2009: 29). Per i candidati del livello B2, invece, la prova con il punteggio più basso è la numero 1 (2,15 su 6 punti), ovvero la prova dedicata alla gestione del sistema aggettivale e pronominale (pronomi e aggettivi indefiniti; pronomi e aggettivi possessivi; pronomi e aggettivi dimostrativi; pronomi e aggettivi riflessivi).

⁷ Per entrambe le prove evidenziate nella nostra analisi sono assegnati 0,3 punti per ogni risposta esatta e 0 punti per ogni risposta sbagliata o omessa. In entrambi i casi il punteggio massimo ottenibile è 6 punti (Barni *et al.*, 2009).

Grafico 2. *Analisi delle Strutture di Comunicazione. Valori medi*



4.2. I risultati del questionario

Il questionario è stato somministrato a 34 candidati, di cui 23 candidate, 11 candidati, con un'età media di 21 anni⁸.

Per quanto concerne il periodo di formazione linguistica svolto nel proprio Paese d'origine (Grafico 3), prima dell'ingresso in Italia, questo va da un minimo di un mese a un massimo di 12 mesi (1 anno).

Grafico 3. *Sezione 2: Formazione linguistica (apprendimento dell'italiano L2). Domanda n. 3*



Secondo quanto previsto dalla normativa di riferimento⁹ (*par. 1.4. Sede prescelta per il corso di lingua italiana*),

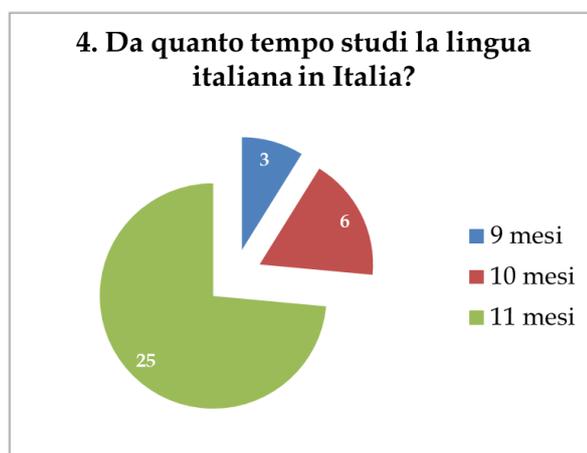
⁸ Il numero degli informanti è piuttosto esiguo se confrontato con quello di coloro che hanno svolto l'esame, ma i dati raccolti vengono trattati ipotizzandone una rappresentatività per una popolazione più ampia.

⁹ Candidati cinesi aderenti al progetto "MARCO POLO" anno accademico 2023/2024: <https://www.studiare-in-italia.it/studenti/stranieri/moduli/2022/MarcoPolo23.pdf>.

gli studenti si possono iscrivere, oltre ai corsi organizzati dall'Università per Stranieri di Perugia, Università per stranieri di Siena, Terza Università di Roma, dalla Società "Dante Alighieri" nonché dall'Università per Stranieri di Reggio Calabria, ad eventuali corsi di lingua presso l'Università prescelta, e quest'ultima, nella propria autonomia organizzativa, ha facoltà di demandare lo svolgimento di tale corso ad un ente esterno o centro linguistico privato di riconosciuta qualità, con il quale abbia stipulato un'apposita convenzione.

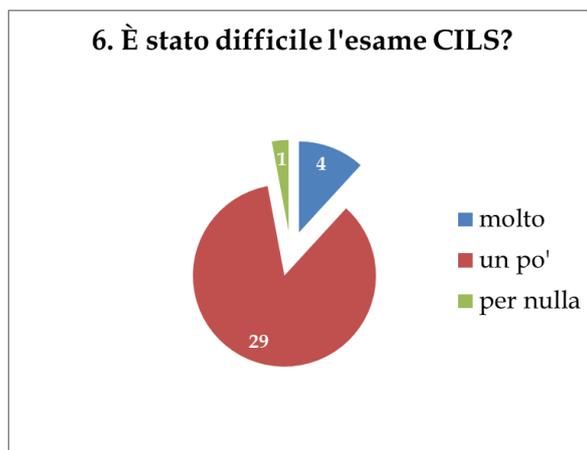
Al percorso di formazione linguistica avviato nel proprio Paese d'origine, si abbina tendenzialmente un periodo di apprendimento linguistico in Italia che, come riportato nel Grafico 4, si estende da un minimo di 9 mesi a un massimo di 11.

Grafico 4. Sezione 2: Formazione linguistica (apprendimento dell'italiano L2). Domanda n. 4



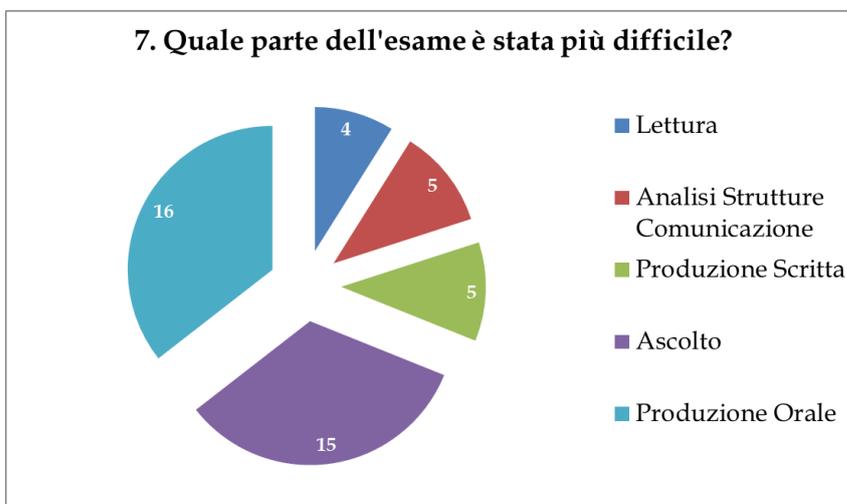
Passando alla sezione 3 dedicata ad indagare come gli studenti abbiano percepito l'esame di certificazione cui si sono sottoposti, il campione da noi analizzato concorda in modo quasi unanime su un livello medio di difficoltà delle prove (Grafico 5). Su una *Likert scale* a tre gradi (da 0 a 2) quasi la totalità degli informanti (n. 29 su 34) ha indicato come risposta *un po'* (1); una piccola porzione (n. 4 su 34) ha indicato come risposta *molto* (2), a segnalare un maggior grado di difficoltà incontrato nello svolgimento delle prove; solo un informante (su 34) ha indicato come risposta *per nulla* (0), a significare presumibilmente una conferma delle aspettative sul livello di complessità dell'esame cui si è sottoposto/a.

Grafico 5. Sezione 3: Esame CILS. Domanda n. 6



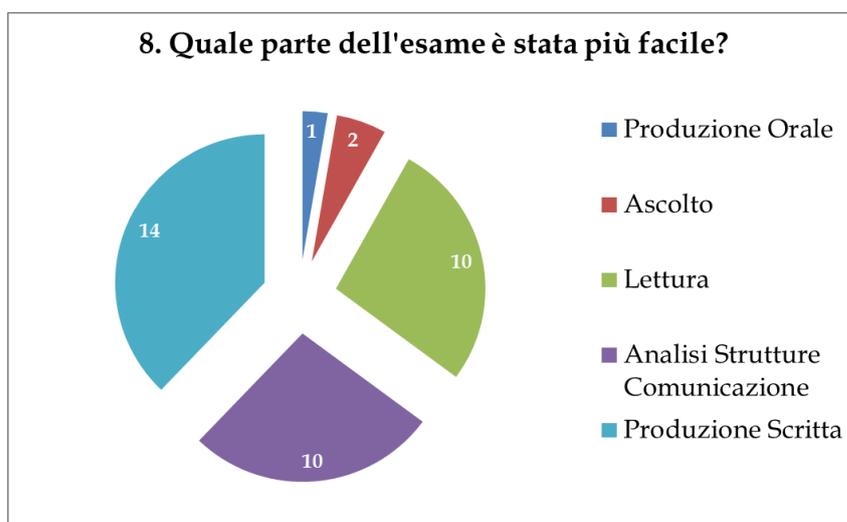
Nelle domande n. 7 (*Quale parte dell'esame è stata più difficile?*) e n. 8 (*Quale parte dell'esame è stata più facile?*) i dati raccolti confermano alcune tendenze già mostrate in altri lavori sul pubblico in questione. In particolare, e come si accennava poco sopra, la prova che pare abbia creato maggiore difficoltà (Grafico 6) concerne l'abilità di Produzione orale (n. 16 su 34), cui segue al secondo posto l'abilità di Ascolto (n. 15 su 34). Il canale dell'oralità, come riflettuto anche nella raccolta di studi curata da Bonvino e Rastelli (2011), comporterebbe maggiori difficoltà per il pubblico degli apprendenti sinofoni, in particolare per una variazione della struttura prosodica della lingua madre, che influirebbe sulla capacità degli apprendenti di orientarsi nella ricezione orale (Costamagna, 2011).

Grafico 6. Sezione 3: Esame CILS. Domanda n. 7



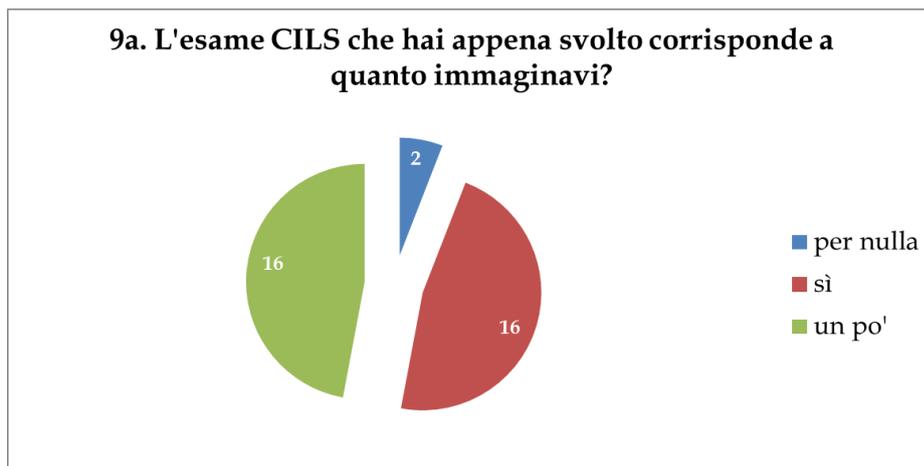
La parte dell'esame che invece è stata considerata dagli studenti come la meno complessa (Grafico 7) è quella concernente l'abilità di Produzione scritta (14 su 34), cui seguono la parte di Analisi delle Strutture di Comunicazione (10 su 34) e la Lettura (10 su 34).

Grafico 7. Sezione 3: Esame CILS. Domanda n. 8



L'ultima domanda del questionario ha sollecitato nei candidati alcune riflessioni di particolare interesse ai fini della ricerca discussa in questo contributo (Grafico 8). In dettaglio, gli informanti sono equamente divisi su due risposte possibili, da un lato una conferma assoluta di quanto atteso (*sì*), dall'altro solo una parziale conferma (*un po'*). Solo due informanti (su 34) dichiarano di essere in qualche modo delusi o sorpresi rispetto al materiale d'esame proposto (*per nulla*).

Grafico 8. Sezione 4: Impatto dell'esame. Domanda n. 9a



Per quanto concerne l'ultima domanda (9b), in cui i candidati hanno potuto scrivere liberamente le proprie riflessioni e osservazioni sulla prova svolta, riportiamo in Tabella 1 una selezione delle più significative, cui abbiamo affiancato tra parentesi l'indicazione dell'aspetto dell'esame oggetto del commento. In particolare, si segnalano alcune perplessità inerenti al tema comunicativo proposto; alcune strutture morfosintattiche e forme lessicali presumibilmente rintracciabili negli *input* testuali proposti o nel *task* per le produzioni scritte e/o orali; e, infine, le condizioni di somministrazione del test di Produzione orale, con particolare riferimento al ruolo del somministratore (Masillo, Machetti, in stampa).

Tabella 1. Sezione 4: Impatto dell'esame. Domanda n. 9b

Sezione 4: Impatto dell'esame. Domanda n. 9b
<i>le immagini dell'orale pensavo che fossino più facile</i> [immagini]
<i>è difficile. Ci sono molte parole quale non ho capisco ma c'è un problema per non studiarci bene</i> [lessico]
<i>mi sento l'ascolto è un po' veloce per me</i> [ascolto]
<i>la produzione scritta ha parlato dei cibi</i> [prod. scritta, tema]
<i>il tema della produzione scritta è nuovo tipo per me</i> [prod. scritta, tema]
<i>la prima volta non ho potuto passare la produzione scritta</i> [prod. scritta]
<i>grammatica è molto difficile</i> [morfosint.]
<i>la produzione orale è difficile perché la grammatica dell'argomento non è facile per me</i> [prod. orale, morfosint.]
<i>il standard dell'esame non è uguale tra gli esaminatori</i> [prod. orale]

4.3. La validazione degli esami

I dati raccolti tramite la somministrazione degli esami CILS B1 e B2 e del questionario post-test sono stati usati ai fini della validazione dell'esame. In relazione a ciò, è stata svolta anche un'analisi dei due esami in termini di abilità, contesti e usi della lingua, tipologie e generi testuali, strutture fonologiche, morfosintattiche, lessicali, così come esplicitati nelle *Linee Guida CILS* (Barni *et al.*, 2009)

L'argomento di validità descritto al § 3 ha evidenziato buone evidenze per le fasi di *Grounds* e *Claim*, mentre diverse sono state le criticità in relazione al *Warrant*, che le evidenze ricavate dall'esame non sono riuscite a supportare, in particolare le evidenze ricavate dalle dichiarazioni degli studenti. Solo per proporre qualche esempio, alcuni errori di morfosintassi rintracciati nelle produzioni scritte ed orali indicherebbero il possesso da parte dei candidati di un livello di competenza in italiano L2 inferiore a quello previsto per i livelli B1 e B2 (Ni, non pubbl.) e sarebbero tali da condurre alla messa in crisi della validità di tali esami.

Tenendo conto di tale criticità, la scelta è stata quella di procedere ad un'integrazione dell'argomento di validità attraverso il modello già adottato per la revisione dell'esame TOEFL (*Test of English as a Foreign Language*, Chapelle *et al.*, 2008). Per gli esami CILS B1 e B2 destinati agli studenti MP/T è stato dunque costruito un argomento di validità funzionante su due diverse basi. La prima è costituita dall'insieme delle caratteristiche della performance dello studente; la seconda dalle caratteristiche dei compiti affidati allo studente per elicitarne la sua performance in italiano L2. Le due basi rappresentano, nell'argomento, i punti di partenza per un'inferenza più ampia che costituisce, a sua volta, la base del *Claim* finale, relativo alla prestazione dello studente in relazione ai domini personale, pubblico, educativo.

L'adozione del nuovo argomento di validità non è risultata esente da criticità. Pur non addentrandosi nella descrizione dettagliata dei diversi passaggi previsti dall'argomento né, tantomeno, nel dettaglio delle singole inferenze (descrizione del dominio, valutazione, generalizzazione, spiegazione, estrapolazione, utilizzazione) discutiamo di seguito che cosa l'uso di tale argomento ha portato a rilevare in relazione alle tre fondamentali, diverse tipologie di validità: la *content-related validity*, la *construct-related validity*, la *criterion-related validity* (Chaloub Deville, O'Sullivan, 2020).

Se degli esami CILS B1 e B2 destinati a studenti MP/T analizziamo la *content-related validity*, le evidenze che ricaviamo dall'analisi del formato dei due esami sono adeguate. Entrambi gli esami presentano misure statisticamente significative relative alla diversità del formato dei *task*, al numero degli *item*, che evidenziano a loro volta una copertura adeguata dei domini della comunicazione su cui lavora l'esame (personale, pubblico, educativo). Inoltre, il tempo previsto per l'esecuzione dei *task* risulta adeguato, ed adeguati risultano gli *input* come le istruzioni delle diverse prove.

Sempre in relazione alla *content-related validity*, anche il grado di autenticità dei due esami risulta adeguato, sia in relazione alle tematiche in esso presenti (legate alla vita quotidiana in Italia, a temi collegati allo studio in Italia ecc.), sia alle tipologie che ai generi testuali (Machetti, Vedovelli, in stampa).

Guardando alla *construct-related validity*, gli esami CILS B1 e B2 destinati a studenti MP/T coprono le abilità di base legate allo sviluppo della competenza linguistico-comunicativa in italiano L2, riflettendo in maniera adeguata il sottostante modello di competenza. Il costruito di tali esami appare inoltre adeguatamente operazionalizzato dal loro formato e le specificazioni riflettono in modo adeguato il costruito.

Le criticità emergono nel raccogliere evidenze rispetto alla *criterion-related validity*, in particolare alla comparabilità degli esami CILS B1 e B2 destinati a studenti MP/T con gli esami CILS UNO-B1 e UNO-B2 standard. A tale criticità si lega, come è facile

immaginare, quella della collocazione degli esami destinati agli studenti MP/T entro l'offerta certificatoria CILS.

Prendendo in considerazione un numero di dati più consistente, relativo agli esami svolti dagli studenti MP/T negli anni 2017-2019 e agli esami CILS UNO-B1 e DUE-B2 svolti da altre tipologie di candidati nello stesso arco temporale – candidati tra cui prevalente è la presenza di studenti internazionali interessati a svolgere un percorso formativo in Italia – la criticità non si attenua e i risultati sono quelli che troviamo riassunti nei sottostanti Grafici 9 e 10.

Grafico 9. *Esiti esami CILS UNO-B1 / B1 MP/T (2017-2019)*

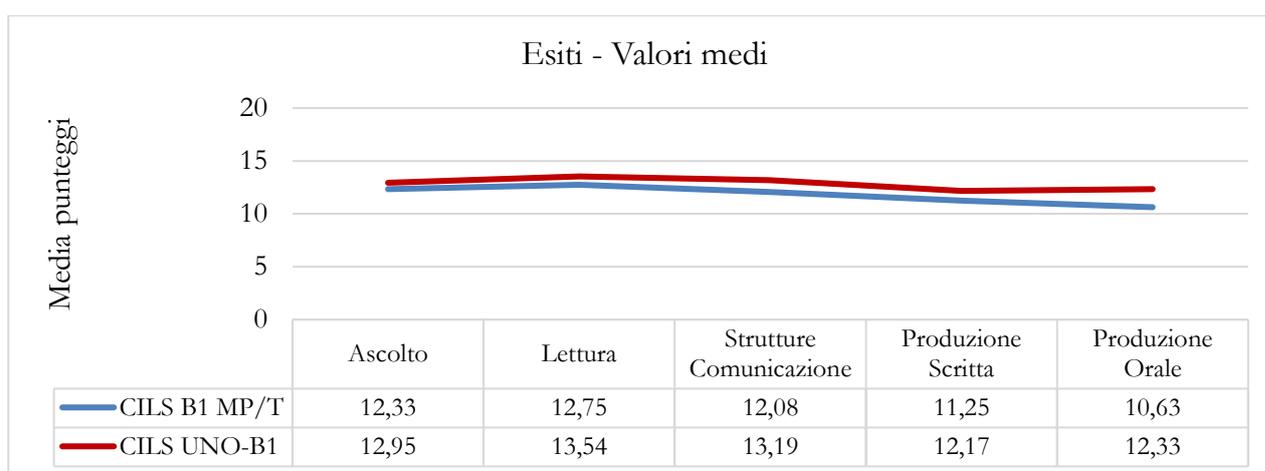
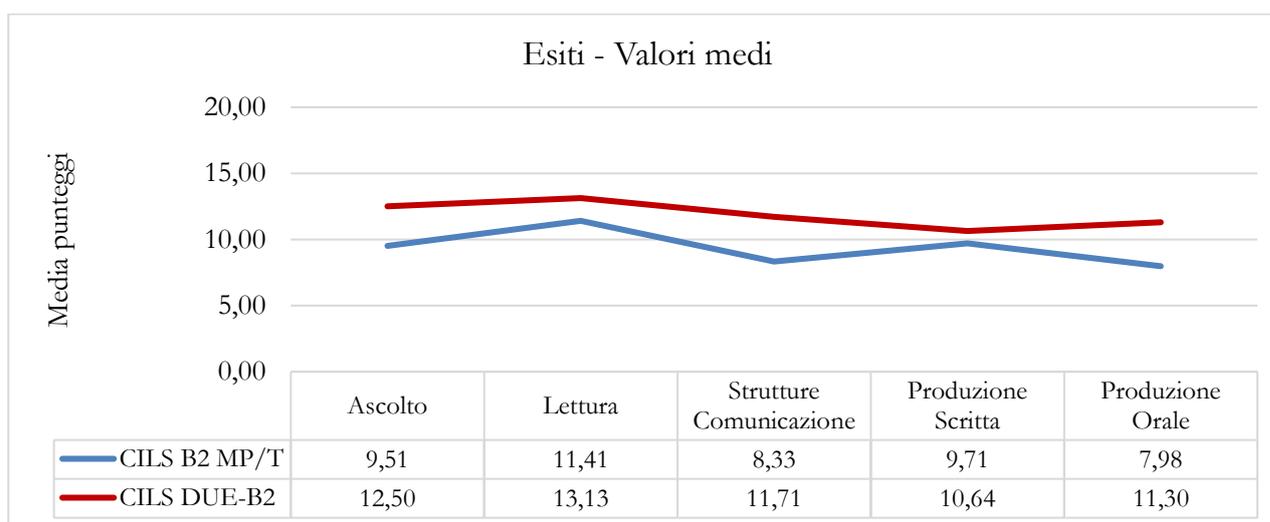


Grafico 10. *Esiti esami CILS UNO-B2 / B2 MP/T (2017-2019)*



Se per il livello B1 non si evidenziano particolari criticità, per gli esami di livello B2 la comparabilità tra l'esame destinato a studenti MP/T e l'esame CILS DUE-B2 standard appare decisamente critica, anche solo analizzando i valori medi. Pur nella consapevolezza della necessità di una discussione più ampia, il dato rappresenta la fonte di quello che l'argomento di validità definisce in termini di *Rebuttal*, ossia elemento che giustifica ed anzi

richiede la messa in discussione della validità dell'esame B2 rivolto agli studenti MP/T. I dati raccolti relativi agli esami MP/T, anche sulla base delle percezioni degli studenti, sono tali da far sorgere dubbi sull'adeguatezza dei *task* e degli argomenti che l'esame propone. Un ampliamento del costrutto dell'esame B2 destinato a tali studenti, capace di tener conto dei contenuti e degli obiettivi del percorso formativo universitario (Bonvino, Rastelli, 2011; Guanfang Zha, Fan, 2016), potrebbe forse portare ad attenuare tale criticità. Tale considerazione trova una sua ragion d'essere anche alla luce di quanto previsto dalla normativa vigente in riferimento ai contenuti della formazione linguistica, ovvero «il corso di lingua [...] dovrà altresì prevedere parte del tempo dedicato alla cultura italiana e/o ai linguaggi settoriali»¹⁰.

5. CONSIDERAZIONI FINALI E PROSPETTIVE FUTURE

Nel condurre uno studio d'impatto, l'obiettivo è comprendere meglio l'interazione tra i contesti macro e micro all'interno del contesto di valutazione in cui i test sono utilizzati e determinare quali elementi facilitino o ostacolino i risultati attesi (Saville, 2000). Uno degli auspici per un impatto positivo implica la comunicazione e la collaborazione tra tutte le parti interessate, i cosiddetti *stakeholders*, ovvero gli attori che a vario titolo prendono parte e svolgono un ruolo del processo di valutazione. Tale condivisione non può che avere un fine ultimo condiviso e condivisibile, ovvero la capacità di cambiare per migliorare per quanto possibile i risultati in termini formativi o mitigare le conseguenze negative connesse agli esami.

I dati analizzati in questo contributo, pur limitandosi a fornire una fotografia della competenza linguistico-comunicativa dei candidati al momento della prestazione in sede d'esame, ci offrono alcuni spunti di riflessione che riguardano non solo la validità degli esami CILS B1 e B2 destinati a studenti MP/T, ma si aprono anche a considerazioni che integrano la valutazione con le dinamiche dei processi di apprendimento ed insegnamento linguistico e ne fanno processi inscindibili.

In relazione alle questioni dell'apprendimento dell'italiano L2, in entrambi i livelli d'esame presi in considerazione in questo contributo i candidati hanno registrato in media punteggi più bassi per le due abilità produttive (*Produzione orale* e *Produzione scritta*) rispetto alle abilità ricettive (*Ascolto* e *Lettura*). Tale dato appare in linea con la discrepanza tra il raggiungimento degli obiettivi delle abilità ricettive rispetto a quelli delle abilità produttive rilevata da ricerche dedicate all'apprendimento dell'italiano L2 da parte di studenti sinofoni (Bettoni, 2001; De Meo, Pettorino, 2011; Luzi, 2011; Costamagna, 2011). Allo stesso modo, i dati discussi nel nostro contributo confermano una difficoltà nel raggiungimento di una riflessione metalinguistica conforme ai profili di competenza della fascia intermedia B1/B2 (Della Putta, 2008; Bagna 2011; Rastelli 2011).

A parziale spiegazione di tale scenario, va detto che dall'avvio del programma gli studenti MP/T hanno evidenziato alcune tendenze rintracciabili non solo nei bisogni specifici degli apprendenti, ma anche nella sfera motivazionale e negli stili di apprendimento (Bonvino, Rastelli, 2010). La lingua d'origine degli studenti e la lingua oggetto di apprendimento e studio presentano evidenti differenze in termini tipologici e strutturali; l'incontro che ne deriva tra due sistemi linguistici così distanti non può che comportare delle ricadute sul percorso acquisizionale (Della Putta, 2008) e, conseguentemente, nei processi valutativi. Come riflette Costamagna (2011), la distanza tipologica della lingua madre, in questi casi specifici, compromette la possibilità di disporre

¹⁰ Candidati cinesi aderenti al progetto "MARCO POLO" anno accademico 2023/2024: <https://www.studiare-in-italia.it/studentistranieri/moduli/2022/MarcoPolo23.pdf>.

di materiale trasferibile e condiziona i processi di riconoscimento o segmentazione dell'*input*, con un impatto rilevabile nei processi di apprendimento e di valutazione linguistica.

In tutte le ricerche fin qui citate, il tema della valutazione linguistica non risulta comunque né centrale né esplicito, ma partendo dal presente contributo tale centralità può essere sfruttata per verificare la fondatezza di alcune ipotesi che nel tempo sono state formulate sulla valutazione linguistica dell'italiano L2 di studenti internazionali (Machetti, 2020). Ne isoliamo alcune, con l'auspicio che negli anni a venire si possano sviluppare attorno a ciascuna di esse nuove ricerche capaci di creare un impatto positivo nel percorso formativo degli studenti internazionali nel nostro Paese:

- l'eterogeneità degli strumenti di valutazione non aiuta la raccolta di dati validi ed affidabili e dunque non ne autorizza la generalizzazione;
- la valutazione condotta con strumenti troppo eterogenei non permette di raccogliere dati omogenei e comparabili;
- i processi e gli strumenti di valutazione linguistica che non rispecchiano il percorso di formazione linguistica degli studenti non forniscono informazioni utili sullo sviluppo della competenza di tali studenti e non possono essere usati con finalità diagnostiche e/o predittive;
- gli strumenti di valutazione certificatori misurano una competenza linguistico-comunicativa di tipo generalista e danno informazioni solo indirette sulla competenza linguistico-comunicativa effettivamente necessaria agli studenti a muoversi in contesto accademico.

In aggiunta a ciò, ci preme sottolineare come gli *Standards for Educational and Psychological Testing* (1999), il Codice di etica ILTA¹¹, il Codice di buone pratiche EALTA¹², i Codici ALTE¹³, gli Standard di qualità dell'associazione CLIQ¹⁴ siano concordi nell'affermare che la valutazione linguistica serve per prendere decisioni importanti, decisive per la vita educativa, per la formazione professionale, ma anche per la vita individuale di ogni studente. Per questo, tutti i codici affermano che l'implementazione e la validazione dei test linguistici debbano ricadere tra le competenze dei professionisti della valutazione, non essere affidata a pratiche e/o giudizi estemporanei, ma a pratiche professionali condivise.

RIFERIMENTI BIBLIOGRAFICI

- Alderson J. C., Wall D. (1993), "Does Washback Exist?", in *Applied Linguistics*, 14, pp. 115-129.
- Bagna C. (2011), "L'esperienza dei progetti Marco Polo e Turandot all'Università per Stranieri di Siena", in Bonvino E., Rastelli S. (a cura di), *La didattica dell'italiano a studenti cinesi e il progetto Marco Polo*, Pavia University Press, Pavia, pp. 93-102.
- Bagna C., Chiapedi N., Salvati L., Scibetta A., Visigalli M. (a cura di) (2017), *La lingua italiana per studenti internazionali: aspetti linguistici e didattici*, Ol3, Perugia.
- Bagna C., Machetti S. (2015), "International mobility and the Impact on LT: the case of Chinese students", in Mader J., Urkun Z. (eds), *Diversity, Plurilingualism and their*

¹¹ <https://www.iltaonline.com/page/CodeofEthics>.

¹² <https://www.ealta.eu.org/guidelines.htm>.

¹³ <https://www.alte.org/Resources-ALTE>.

¹⁴ <https://www.associazionecliq.it/chi-siamo>.

- impact on language Testing and Assessment*, IATEFL Testing, Evaluation and Assessment SIG and University for Foreigners of Siena, Siena 22-23 November 2013, IATEFL (TEA SIG), pp. 9-12.
- Banfi E. (a cura di) (2003), *Italiano/L2 di cinesi*, FrancoAngeli, Milano.
- Barni M., Bandini A., Sprugnoli L., Lucarelli S., Scaglioso A. M., Strambi B., Fusi C., Arruffoli A. M. (2009), *Linee guida CILS*, Guerra Edizioni, Perugia.
- Barni M., Machetti S. (2022), “Le certificazioni di italiano L2”, in Diadori P. (a cura di), *Insegnare italiano L2*, Le Monnier, Firenze, pp. 122-136.
- Bettoni C. (2001), *Imparare un'altra lingua. Lezioni di linguistica applicata*, Laterza, Roma-Bari.
- Bonvino E., Rastelli S. (a cura di) (2011) *La didattica dell'italiano a studenti cinesi e il progetto Marco Polo*, Pavia University Press, Pavia.
- Chaloub-Deville M., O'Sullivan B. (2020), *Validity. Theoretical Development and Integrated Arguments*, Equinox Publishing Ltd, Sheffield-Bristol.
- Chapelle C. A., Enright M. K., Jamieson J. M. (eds.) (2008), *Building a Validity Argument for the Test of English as a Foreign Language*, Taylor & Francis, Routledge.
- Council of Europe (2001), *The Common European Framework of Reference for Languages: Learning, Teaching, Assessment (CEFR)*, University Cambridge Press, Cambridge.
- Council of Europe (2020), *Common European Framework of Reference for Languages: learning, teaching, assessment. Companion volume*, Strasbourg: <https://rm.coe.int/common-european-framework-of-reference-for-languages-learning-teaching/16809ea0d4>. Trad. it. a cura di Barsi M., Lugarini E., *Quadro comune europeo di riferimento per le lingue: apprendimento, insegnamento, valutazione. Volume complementare*, in *Italiano LinguaDue*, 12, 2, 2020: <https://riviste.unimi.it/index.php/promoitals/article/view/15120>.
- Costamagna L. (2011), “L'apprendimento della fonologia dell'italiano da parte di studenti sinofoni: criticità e strategie”, in Bonvino E., Rastelli S. (a cura di), *La didattica dell'italiano a studenti cinesi e il progetto Marco Polo*, Pavia University Press, Pavia, pp. 49-66.
- De Meo A., Pettorino M. (2011), “L'acquisizione della competenza prosodica in italiano L2 da parte di studenti sinofoni”, in Bonvino E., Rastelli S. (a cura di), *La didattica dell'italiano a studenti cinesi e il progetto Marco Polo*, Pavia University Press, Pavia, pp. 67-78.
- Della Putta P. (2008), “Insegnare l'italiano ai sinofoni: contributi acquisizionali, tipologici e glottodidattici”, in *Studi di Glottodidattica*, 2, pp. 52-67.
- Elder C. (2017), “Language Testing in Higher Education”, in Shohamy E., Or I. G., May S. (eds.), *Encyclopedia of Language Education. Language Testing and Assessment*, Taylor & Francis, New York, pp. 271-286.
- Elder C., Knoch U. (2016), “Post-entry english language assessments at university: how diagnostic are they?”, in Aryadoust V., Fox J. (eds.), *Trends in language assessment research and practice: the view from the Middle East and the Pacific Rim*, Cambridge Scholars Publishing, Cambridge, pp. 209-230.
- Green A. (2013), “Washback in language assessment”, in *International Journal of English Studies*, 13, 2, pp. 39-51
- Grego Bolli G., Machetti S., Rini D. (2017), “Researching Italian language assessment in China and Italy: focusing on CELI and CILS examinations”, in Bagna C., Chiapedi N., Salvati L., Scibetta A., Visigalli M. (a cura di), pp. 28-35.
- Grego Bolli G., Machetti S., Bandini A., Rini D. (2017), “The Evaluation of Chinese Students of Italian L2 in China and in Italy: Practices at the Universities for Foreigners of Perugia and Siena”, in ALTE (ed.), *Learning and Assessment: Making the Connections*, pp. 118-123.

- Guanfang Zha C., Fan J., (2016), *The construct of EFL Academic Writing Ability as Operationalized in the Context of Chinese Colleges*, Working-in-progress, LTRC 2016.
- Jin Y. (2022²), "Test specifications", in Fulcher G., Harding L. (eds.), *The Routledge Handbook of Language Testing*, Routledge, Abington-New York, pp. 271-288.
- Luzi E. (2011), "Il certificato di italiano L2 per sinofoni base.IT: ideazione, struttura e primi risultati", in Bonvino E., Rastelli S. (a cura di), *La didattica dell'italiano a studenti cinesi e il progetto Marco Polo*, Pavia University Press, Pavia, pp. 103-122.
- Machetti S. (2020), "La valutazione linguistica. Una riflessione a cinque anni dalla redazione del Manifesto", in Rastelli S., Bagna C. (a cura di), *Manifesto per l'insegnamento della lingua italiana agli studenti internazionali. Otto commenti*, Pacini Editore, Pisa, pp. 99-110.
- Machetti S., Vedovelli M. (a cura di) (in stampa), *Manuale della certificazione dell'italiano L2. Nuova edizione*, Carocci, Roma.
- Machetti S., Vedovelli M. (2012), "Per una storia linguistica dell'emigrazione italiana nel mondo: lo spazio linguistico globale", in Bancheri S. (ed.), *Italian Outside of Italy. The Situation in Canada, USA and The English-Speaking World*, The Frank Iacobucci Center, Toronto, pp. 173-200.
- Masillo P., Machetti S. (in stampa), "Situazioni comunicative asimmetriche. Il test di produzione orale negli esami di certificazione della competenza in L2", in Caruso V., Maffia M. (a cura di), *Vecchie e nuove forme di comunicazione diseguale: canali, strutture e modelli*, Collana Studi AItLA, n. 17, Officinaventuno, Milano, pp. 99-112.
- Messick S. (1989), "Validity and washback in language testing", in *Language Testing*, 13, 4, pp. 241-256.
- Milanovic M., Saville N. (1996), *Considering the impact of Cambridge EFL Examinations*, UCLES Internal Report, University of Cambridge Local Examination Syndicate, Cambridge.
- Mislevy R. J., Steinberg L. S., Almond R. G. (2002), "Design and analysis in task-based language assessment", in *Language Testing*, 19, pp. 477-496.
- Ni Y. (non pubbl.), *Dalla competenza testuale alla didattica della scrittura "accademica" in italiano L2: uno studio pilota sulle strategie di ripresa anaforica nei testi scritti dagli apprendenti sinofoni in contesti LS e L2*, tesi di Dottorato in Linguistica Storica, Linguistica Educativa e Italianistica. L'italiano, le altre Lingue e Culture, Curriculum Linguistica e Didattica della lingua italiana a stranieri, Università per Stranieri di Siena, XXXI Ciclo.
- Pilliner A. (1973), *Assessment - Principles and Practice with Special Reference to Education in Pakistan*, Unpublished ms, The British Council.
- Qian D. D., Cumming A. (2017), "Researching English Language Assessment in China: Focusing on High-Stakes Testing", in *Language Assessment Quarterly*, 14, 2, pp. 97-100.
- Rastelli S. (2011), "L'italiano L2 per i cinesi è più difficile? La distinzione tra grammatica e processing nella glottodidattica sperimentale", in Bonvino E., Rastelli S. (a cura di), *La didattica dell'italiano a studenti cinesi e il progetto Marco Polo*, Pavia University Press, Pavia, pp. 79-92.
- Rastelli S., Bagna C. (a cura di) (2020), *Manifesto per l'insegnamento della lingua italiana agli studenti internazionali. Otto commenti*, Pacini Editore, Pisa.
- Rugge F. (a cura di) (2019), *L'internazionalizzazione della formazione superiore in Italia. Le Università. Aggiornamenti e approfondimenti*, Fondazione CRUI, Roma.
- Saville N. (2000), "Investigating the impact of international language examinations", in *Research Notes*, 2, University of Cambridge, pp. 4-7.
- Zou S., Xu Q. (2017), "A Washback Study of the Test for English majors for Grade Eight (TEM8) in China – From the Perspective of University Program Administrators", in *Language Assessment Quarterly*, 14, 2, pp. 140-159.

