

PAOLA MASILLO¹, SABRINA MACHETTI¹

Situazioni comunicative asimmetriche. Il test di produzione orale negli esami di certificazione della competenza in L2

Abstract

In most language examinations, the communicative situation in which the speaking test is embedded is clearly asymmetrical. This asymmetry consists in the fact that the interviewer and the candidate access in different ways the process of meaning negotiation involved in the communicative situation. Likewise, they have different access, therefore, to the possibilities for success arising from that situation, precisely because of their respective roles within the examination situation.

This asymmetry is strongly affected by the presence of elements that are only partially predictable, but which play a central role in terms of the validity, reliability and fairness of language examinations.

The paper discusses data from a research study analysing the behaviour of the interviewer within the CILS Speaking test, offered at the University for Foreigners of Siena. The paper focuses on the oral interaction test of the CILS UNO-B1, CILS B1 Cittadinanza and CILS DUE-B2 exams. The data collected are not only of some interest for the purposes of the validation process, but also in view of a discussion on the variables that impact the communicative asymmetry and negatively affect the candidate's performance.

1. Oggetto e obiettivi della ricerca²

Il test di produzione orale utilizzato nella maggior parte degli esami che certificano la competenza in L2, diffusi e riconosciuti a livello internazionale (Field 2011), si struttura generalmente attraverso prove di tipo monologico (parlato monodirezionale faccia a faccia) e dialogico (parlato bidirezionale faccia a faccia, con presa di parola libera, semi-libera, non libera), condotte da uno o più somministratori alla presenza di uno o più candidati (Luoma 2004: 35-39). La situazione comunicativa che caratterizza tale test è chiaramente asimmetrica, poiché il somministratore e il candidato, in ragione dei rispettivi ruoli all'interno della situazione d'esame, accedono in modi diversi al processo di negoziazione e mediazione dei significati e dei sensi caratterizzanti tale situazione. L'asimmetria è fortemente influenzata dalla presenza di elementi solo parzialmente prevedibili, ma che giocano un ruolo centrale in termi-

¹ Università per Stranieri di Siena.

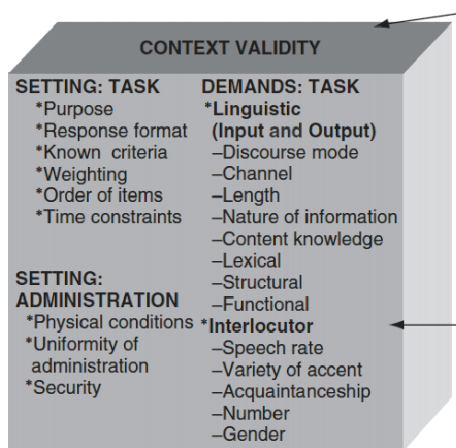
² Il contributo è frutto della ricerca condivisa delle autrici. Sono di Paola Masillo i parr. 3, 4, 4.1. 4.2, 5; di Sabrina Machetti i parr. 1, 2.

ni di validità, affidabilità, equità dell'esame nel suo complesso (Young & Milanovic 1992; O'Sullivan 2002, 2008; Brown 2003; Weir 2005; Nakatsuhara 2009).

Il contributo discute i dati provenienti da uno studio di validazione degli esami CILS – Certificazione di Italiano come Lingua Straniera – dell'Università per Stranieri di Siena (Machetti & Purpura 2018), concentrandosi sull'analisi del comportamento del somministratore all'interno della prova d'interazione proposta nel test di produzione orale degli esami CILS UNO-B1, CILS B1 Cittadinanza e CILS DUE-B2 (Barni & Machetti 2019). Come noto, si tratta di esami e di livelli la cui "spendibilità" (Vedovelli 2005: 15) è andata aumentando negli anni a noi più vicini, considerata la loro rilevanza per il contesto della formazione accademica, per il contesto lavorativo, per la richiesta della cittadinanza italiana ecc.

Applicando il *test validation framework* elaborato da Weir (2005) – che pone l'attenzione sui processi cognitivi che caratterizzano i meccanismi comunicativi verbali che ciascun test di lingua dovrebbe essere in grado di replicare nelle diverse fasi e con i diversi elementi di cui si compone – è possibile sostenere come in un test di produzione orale numerosi siano gli aspetti che entrano in gioco. Come rilevano Galaczi e French (2011: 112) in riferimento al modello di Weir (2005), alcuni degli aspetti di contesto – dalle caratteristiche del task alle modalità di somministrazione (fig. 1) – sono in grado di condizionare, appunto in termini di *context validity*, la performance del candidato e, di conseguenza, le evidenze che il test si prefigge di raccogliere.

Figura 1 - *Aspetti della validità di contesto per la produzione orale (Weir 2005: 45)*



Focalizzando l'attenzione sul momento della somministrazione del test di produzione orale degli esami CILS UNO-B1, CILS B1 Cittadinanza e CILS DUE-B2, la nostra ricerca si focalizza sulla figura del somministratore, prendendo appunto le mosse dall'assunto secondo cui la performance di un candidato può essere condizionata dal comportamento di tale figura, e risultare a causa di ciò qualitativamente differente (Alderson *et al.* 1995; French 2003; O'Sullivan *et al.* 2002). La performance orale può dunque dipendere, se non essenzialmente, almeno in parte da ciò

che il somministratore fa (Lazaraton 1996; McNamara & Lumley 1997; O'Sullivan 2000; Brown 2003; Field 2011).

L'analisi che qui proponiamo mira ad isolare alcuni fattori (*sources of variances*, Weir 2005) – riconducibili al somministratore – che possono condizionare la performance orale del candidato e che sono rilevabili dai valutatori in fase di valutazione³:

- a. facilitazioni (*interviewer's accommodation*: velocità del parlato, semplificazione lessicale, riformulazione delle domande, ...);
- b. limitazioni al parlato del candidato (*interviewer's questions*: modalità di proposta delle domande, domande sì/no, affermazioni che necessitano di una semplice conferma da parte del candidato, ...);
- c. profilo (formazione del somministratore, livello di preparazione, livello di competenza linguistica in italiano se di origine straniera, ...).

2. Il quadro teorico della ricerca

Negli esami che certificano la competenza in L2, la scelta di utilizzare prove di tipo monologico e prove di tipo dialogico risiede nel fatto che queste risultano, almeno potenzialmente, rappresentative di quanto accade nei contesti comunicativi quotidiani, in cui la lingua parlata viene usata con modalità e scopi diversi. Le prove di produzione ed interazione orale, rappresentative del profilo di competenza oggetto di misurazione, sono infatti finalizzate ad elicitarne un campione di lingua parlata, simulare una situazione comunicativa reale, o il più possibile autentica (*situational authenticity*, Bachman & Palmer 1996), e raccogliere evidenze utili al fine di elaborare inferenze e proporre un giudizio sull'abilità di produzione orale del candidato che ad esse si sottopone (Taylor 2011; Nakatsuhara *et al.* 2021). Durante tali prove, il candidato deve sapersi esprimere in maniera adeguata a quanto richiesto dagli input delle prove stesse, in linea con gli standard previsti per il livello d'esame cui si sottopone; il somministratore deve invece saper rendere la performance del candidato quanto più possibile adeguata alle caratteristiche della comunicazione parlata, pur vincolata dal tempo, dagli argomenti, dal fine delle prove, ovvero l'accertamento della competenza in L2. Il somministratore, infatti, non ha solo il compito di comunicare al candidato le modalità di svolgimento delle prove di produzione orale, ma deve anche regolarne i tempi ai fini della corretta esecuzione dell'esame in termini di uniformità di somministrazione e interagire con il candidato (Field 2011). In un certo qual modo, egli condivide la responsabilità del successo dell'esame.

Come si diceva poco sopra, la situazione comunicativa che caratterizza le prove di produzione ed interazione orale è di tipo asimmetrico, anche solo per il fatto di chiamare in causa, all'interno del contesto d'esame, un somministratore nativo (o, se non nativo, con una competenza linguistica almeno di livello C1) e uno o più candidati,

³ Negli esami CILS, il somministratore è responsabile dello svolgimento del test di produzione orale, ed ha l'esclusiva funzione di somministrare le prove di cui il test si compone e di condurne le diverse fasi. La performance del candidato viene registrata dal somministratore su supporto digitale, caricata su apposita piattaforma e valutata in una seconda fase da valutatori esperti (Badini *et al.* 2022).

non nativo/i nella lingua oggetto di verifica. Young e Milanovic (1992) hanno definito l'asimmetria lungo tre direttrici: dominanza, dipendenza e raggiungimento dell'obiettivo (*dominance, contingency and goal-orientation*) e hanno specificato che questa si lega al fatto che il somministratore e il candidato accedono in maniera diversa, proprio in ragione dei rispettivi ruoli all'interno della situazione d'esame, alla gestione e negoziazione dell'evento comunicativo e alle conseguenti possibilità di successo (*unbalanced power distribution*, Ross & Berwick 1992; Young & Milanovic 1992).

Secondo alcuni studi condotti sui test di competenza in inglese L2 (Brown 2012), la naturale condizione di asimmetria della situazione comunicativa che caratterizza il test di produzione orale risulta fortemente dipendente dalla presenza di elementi non solo parzialmente prevedibili ma anche difficilmente controllabili (O'Sullivan & Green 2011: 57). Tali elementi (Bachman 1990; Alderson *et al.* 1995; Brown 1995; Bachman & Palmer 1996; O'Sullivan 2000; Luoma 2004) si legano al momento della somministrazione della prova e, dunque, al contesto, anche fisico, d'esame (ad es. l'acustica dell'aula dove l'esame si svolge, la registrazione su supporto digitale che tendenzialmente inibisce il candidato, ...); riguardano le caratteristiche del candidato (ad es. il suo livello di ansia, la personalità, il livello di preparazione, ...). Infine, e come si sottolineava poco sopra, coinvolgono anche il somministratore (ad es. la formazione di cui è in possesso, la velocità del parlato, l'accento, ...).

Se le caratteristiche del candidato, di natura fisica, psicologica ed individuale (Brown 1995), rappresentano elementi in incertezza in grado di incidere non solo sulla sua performance ma anche sulla misurazione della stessa e determinare l'errore in fase di valutazione (Bachman 1990), lo stesso può dirsi per le caratteristiche del somministratore. Chi somministra l'esame, per i fattori di cui si diceva al par. 1 ed in virtù del ruolo che gli compete nella situazione d'esame, può incidere infatti sulla non prevedibilità degli esiti dell'interazione comunicativa e, in alcuni casi, compromettere l'affidabilità della misurazione (Brown 2003).

3. Metodologia e strumenti della ricerca

Il test di produzione orale CILS si compone di due prove: una prima prova consistente in un'interazione tra somministratore e candidato e una seconda prova consistente in un parlato faccia a faccia di tipo monologico. Le prove sono finalizzate ad elicitarne la capacità del candidato di esprimersi oralmente e di interagire in italiano L2 a un dato livello di competenza linguistica. In base al profilo di competenza, la prova dialogica consiste in uno scambio comunicativo tra somministratore e candidato che si può svolgere in forma guidata (domanda/risposta) o forma libera (*role-play*, intervista, discussione); la prova monologica può vertere su una descrizione, una narrazione o l'espressione di un'opinione (Barni *et al.* 2009). Nel corso dell'intero test di produzione orale CILS, il somministratore assume, come si diceva, il ruolo di interlocutore, in alcune fasi anche di intervistatore, con la funzione di dare avvio allo scambio comunicativo e cercare di portarlo avanti in linea con quanto richiesto dal task scelto dal candidato.

I dati discussi di seguito provengono da un campione di interazioni orali raccolto nelle seguenti sessioni d'esame: giugno 2021 per gli esami CILS UNO-B1 e CILS-DUE B2; febbraio e dicembre 2021 per gli esami CILS B1 Cittadinanza.

Ai fini della discussione, di tali interazioni sono state selezionate alcune evidenze, rilevate in fase di monitoraggio dei valutatori sulla base di un'analisi qualitativa e quantitativa delle note da questi riportate nel corso dell'attività di valutazione⁴.

Tali evidenze, una volta rilevate e selezionate, sono state analizzate al fine di rintracciare possibili linee di tendenza, riferibili agli effetti che il somministratore può in qualche modo contribuire a determinare nella performance dei candidati. Infine, tali linee di tendenza sono state classificate in base alle seguenti quattro variabili, riprese da O'Sullivan (2002), cui è stata aggiunta una quinta, per comprendere evidenze di altro genere:

1. tempi della prova, ovvero la durata effettiva dello svolgimento dell'esame;
2. toni/comportamento del somministratore, ovvero il grado di interazione e/o di supporto del somministratore nello svolgimento della prova dialogica;
3. input e modalità con cui il somministratore ne fa proposta al candidato, ovvero l'argomento selezionato dal candidato/dalla candidata e le strategie del somministratore nel proporre lo stesso;
4. qualità dell'interazione, ovvero il rispetto dei turni di parola, la velocità del parlato, l'accento ecc.;
5. altro, ovvero un'aula rumorosa, la qualità della registrazione audio ecc.

4. *Discussione dei dati*

Dedichiamo i due seguenti paragrafi alla discussione dei dati del campione, condotta alla luce delle variabili sopra elencate.

4.1 Esito degli esami

Il campione di candidati che ha sostenuto l'esame CILS nei livelli oggetto del nostro studio comprende: n. 2252 candidati per l'esame UNO-B1, n. 2345 per l'esame DUE-B2 e n. 3089 per l'esame B1-Cittadinanza.

Per quanto concerne il punteggio finale dell'esame, proponiamo un confronto tra le percentuali di superamento ottenute nei livelli CILS UNO-B1 e DUE-B2 (tab. 1).

Tabella 1 - *Percentuale di superamento candidati UNO-B1 / DUE-B2*

	<i>N. candidati sufficienti</i>	
	Prova 1	Prova 2
<i>Livello B1</i>	48% (1083)	45,4% (1023)
<i>Livello B2</i>	46,5% (1093)	47,8% (1122)

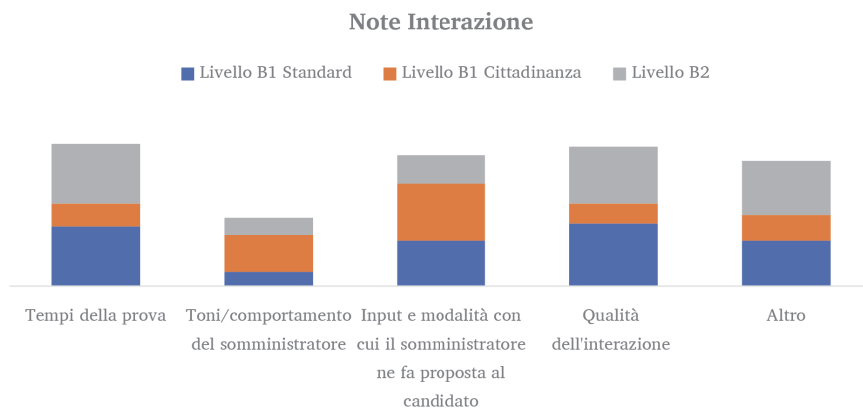
⁴ Sulla fase di monitoraggio dei valutatori condotta dal Centro CILS, si veda Badini *et al.* (2022).

Un dato interessante rilevato per la prova numero 1, ovvero la prova d'interazione orale, concerne la percentuale di maggiore successo per i candidati del livello B1 (48%) rispetto ai candidati del livello B2 (46,5%). La prova per il livello B1 prevede, come indicato nelle *Linee guida CILS* (Barni *et al.* 2009) e in base a quanto previsto dal *Quadro Comune Europeo di Riferimento per le Lingue* (Consiglio d'Europa 2001), semplici scambi routinari di domande e risposte legate ad argomenti familiari inerenti prevalentemente il dominio personale e pubblico. Le domande del somministratore, da un lato, sono relativamente prevedibili; dall'altro, le risposte dei candidati sono tendenzialmente semplici e possono consistere in una o due frasi (Masillo & Strambi 2020). Al livello B2, lo sforzo cognitivo per il candidato tende ad aumentare, e il contenuto dell'input può influenzare la prestazione linguistica: le domande poste dal somministratore raggiungono in molti casi un grado maggiore di complessità, vertendo su argomenti generali e di attualità (Bachman & Cohen, 1999). Su tali argomenti il candidato è chiamato ad esprimere un'opinione, talvolta fornendo risposte articolate e non facilmente prevedibili. Ne deriva, nel complesso, un impegno evidentemente maggiore sia da parte del candidato che del somministratore, rispetto a una più semplice partecipazione a uno scambio routinario di domande e risposte (Field 2011).

Focalizzando la nostra attenzione sulla tipologia di prova, ovvero sulla possibile influenza riconducibile a un'interazione (prova 1, dialogo) o a un'esposizione (prova 2, monologo), notiamo come una discreta variabilità si registra in particolare per il livello B1 (48% P1 *vs.* 45,4% P2). Il dato trova conferma nella letteratura del settore: riprendendo infatti la riflessione di Bachman sull'influenza del *test method* (1991: 687), una delle principali evidenze emerse dalla ricerca sui test di lingua rivela che il rendimento in una prova è influenzato non solo dall'abilità che stiamo cercando di misurare, ma anche dal metodo che utilizziamo per misurarla.

4.2 Analisi delle note dei valutatori

Per quanto concerne le analisi svolte sulle note proposte dai valutatori nel corso della sessione di valutazione, riportiamo in figura 2 una rappresentazione grafica della quantità di occorrenze, distinte per livello d'esame, delle cinque variabili qui prese in considerazione.

Figura 2 - *Classificazione note valutatori*

La distribuzione delle variabili risulta tendenzialmente condizionata dal livello d'esame; in particolare, si nota una maggiore rappresentatività di alcune variabili, ovvero la variabile 2 (Toni/comportamento del somministratore) e la variabile 3 (Input e modalità con cui il somministratore ne fa proposta al candidato), nelle performance dell'esame CILS B1 Cittadinanza.

Procedendo con ordine, per quanto concerne la variabile 1, ovvero la gestione dei tempi per lo svolgimento della prova, riportiamo di seguito (tab. 2) alcune delle note di valutazione rilevate e prese in esame.

Tabella 2 - *Variabile 1 – Lista occorrenze*

Livello B1
Prova breve
Prova troppo lunga
Livello B1 Cittadinanza
I tempi non sono rispettati (matr. *** presentazione personale 11 minuti, prova n. 1 3 minuti)
Presentazione personale lunga (matr. *** 12 minuti; *** 9:31 minuti), prova n. 1 breve (matr. *** 4 minuti)
Il somministratore non rispetta i tempi
La presentazione personale è breve (matr. *** 5:50 minuti; matr. *** 4 minuti). La prova n. 1 è breve (matr. *** 2:45 minuti; matr. *** 2 minuti; matr. *** 4 minuti)
I tempi non sono rispettati, le prove sono generalmente brevi rispetto a quanto previsto (durata meno di 10 min)
Livello B2
Numerosi errori morfosintattici e di lessico, probabilmente dovuti alla eccessiva lunghezza della prova
Prova breve
Prova lunga
[...]

Da un'analisi delle prove si evince che uno dei fattori di maggiore incidenza sembrerebbe risiedere nel livello di competenza del candidato. Ciò conduce talvolta il somministratore ad intervenire – perché più o meno consapevolmente condizio-

nato dal fatto di essere anche un insegnante – stimando un grado di inadeguatezza nella performance del candidato e interrompendo così la somministrazione della prova. In tale comportamento, non è da escludere un livello di preparazione del somministratore non pienamente adeguato al compito che è chiamato a svolgere. Ciò in molti casi determina l'interruzione della prova, senza rispettare il tempo a disposizione del candidato: il somministratore è carente di idee o argomenti da proporre al candidato, ed è dunque incapace di portare avanti e condurre a buon fine l'interazione. Come noto, il fattore relativo alla preparazione e dunque alla formazione del somministratore (Field 2011) rappresenta una delle potenziali fonti di errore che trasversalmente incide su più variabili della somministrazione e comporta un generale impatto nella valutazione (Weir 2005). In particolare, tale osservazione rimanda alle note dei valutatori riconducibili non solo alla variabile 1 (Tempi della prova), ma anche alle variabili 2 (Toni/comportamento del somministratore) e 3 (Input e modalità con cui il somministratore ne fa proposta al candidato).

Analizzando la variabile relativa ai toni/comportamento del somministratore (*features of examiner talk*, Brown & Lumley 1997; Brown 2003), si riporta di seguito una lista di occorrenze delle caratteristiche rintracciate e riconducibili alla suddetta variabile (tab. 3). Si tratta, come affermato in precedenza, di caratteristiche che possono in un modo o nell'altro condizionare la performance del candidato, semplificando o rendendo più complesso il task proposto e le modalità con cui dovrebbe o potrebbe essere svolto (Bachman & Cohen 1999).

Tabella 3 - *Variabile 2 – Lista occorrenze*

Livello B1 Cittadinanza
Il somministratore pone domande in modo confuso, incalzante.
Il somministratore fa domande non pertinenti né previste dalla lista delle domande proposte per la presentazione personale.
Il somministratore interviene a volte con considerazioni/comments di natura politica o con valore di giudizio.
Il somministratore assume un tono a volte annoiato/insofferente.
Il somministratore si lascia andare a commenti e considerazioni non pertinenti al contesto.
Il somministratore esprime giudizi sul livello di competenza.
Il somministratore assume un registro non adeguato al contesto d'esame, presumibilmente perché è stato insegnante dei candidati.
Il somministratore corregge continuamente i candidati nelle scelte sbagliate di natura morfosintattica oppure nelle scelte lessicali.
[...]

Il dato rilevato mediante le note di valutazione trova una possibile ipotesi interpretativa nelle specificità del formato d'esame e, più in dettaglio, nella modalità di somministrazione della prova orale. La prima parte della prova è dedicata a una presentazione personale e ha come obiettivo principale misurare la capacità del candidato di presentarsi e parlare di sé, fornendo informazioni personali sulla base di una serie di domande poste dal somministratore. Quest'ultimo è chiamato ad attenersi alla lista proposta, evitando di introdurre nuovi elementi nella conversazione guidata

che potrebbero compromettere la validità della prova e l'affidabilità dei punteggi (Masillo 2021). Ciononostante, tale comportamento, riconducibile alla variabile 2, occorre con particolare frequenza in questo tipo d'esame.

Per quanto concerne le modalità di svolgimento del task (Brown 2003), ovvero la variabile 3, si riporta di seguito (tab. 4) una lista di occorrenze delle caratteristiche rintracciate e riconducibili alla modalità di proposta dell'input al candidato e di restituzione del feedback da parte del somministratore. Anche in questo caso, si tratta di caratteristiche che possono in qualche modo avere un impatto sulla performance del candidato, influenzando l'esecuzione del compito e le condizioni con cui dovrebbe o potrebbe essere svolto.

Tabella 4 - *Variabile 3 – Lista occorrenze*

Livello B1 Cittadinanza
Il somministratore pone domande di natura enciclopedica.
La somministrazione pare sia condotta da due somministratori, a quello principale subentra un altro che pone ulteriori domande, quasi a sostegno del somministratore principale quando non sa cosa chiedere.
Il somministratore lascia il candidato presentarsi da solo, senza intervenire ponendo le domande, non rispetta la traccia con la lista delle domande.
Il somministratore pone domande non pertinenti all'input scelto.
Da come il somministratore introduce l'argomento della prova, si ha l'impressione che l'input sia stato scelto dal somministratore.
Il somministratore somministra più di un input oppure va fuori tema presumibilmente per carenza di domande.
A volte interviene un secondo somministratore, a supporto del primo, con altre domande oppure commentando quanto detto dal candidato.
[...]

Infine, per quanto concerne la variabile 4 (Qualità dell'interazione), le evidenze rilevate denotano in alcuni casi una tendenza comune per alcuni profili di candidati, ovvero la memorizzazione. I candidati, di frequente dopo non aver superato un esame e dopo essersi iscritti a corsi preparatori, propongono durante la prova un modello di testo che, per contenuto e strutture, appare frutto di un processo di memorizzazione (O'Sullivan & Green 2011). L'esperienza specifica dell'esame da parte del candidato, o la preparazione attraverso un corso, possono infatti favorire una certa familiarità con gli argomenti e con i compiti proposti, in particolare al livello B1 (O'Sullivan & Green 2011; Masillo & Strambi 2020). Il ruolo del somministratore, non solo in termini di esperienza ma anche di formazione per quanto chiamato a svolgere, diventa cruciale per poter risolvere interazioni che perdono di spontaneità e autenticità per le ragioni sopra discusse, e dunque ai fini dell'affidabilità della somministrazione della prova e dei punteggi (Fulcher & Davidson 2012).

Tabella 5 - Variabile 4 – Lista occorrenze

Livello B1
Il candidato non parla
Il somministratore parla più del candidato
La candidata non parla, nonostante l'aiuto del somministratore
Il parlato del candidato è poco spontaneo
Livello B2
Il candidato non parla
La prova viene somministrata come monologo anziché come un dialogo
Il candidato non riesce a portare avanti il discorso
Parlato poco spontaneo
[...]

5. Considerazioni conclusive e prospettive future

I dati discussi in questo contributo non solo risultano utili ai fini del continuo processo di validazione degli esami CILS, ma consentono di sviluppare una riflessione e alcune considerazioni su quanto le variabili prese in esame possano influire sull'asimmetria comunicativa tipica della situazione d'esame ed incidere sulla performance del candidato e sull'esito finale dell'intero esame.

In primo luogo, i dati analizzati permettono di ipotizzare che se la prova di produzione orale è frutto di una costruzione condivisa ed è caratterizzata da un meccanismo di reciprocità a cui gli interlocutori, seppur a livelli diversi, contribuiscono in maniera attiva, l'asimmetria permane, ma non produce effetti significativamente negativi sull'intero processo di valutazione, sia nella prospettiva del candidato che in quella del somministratore.

Nel test di produzione orale proposto negli esami CILS, inoltre, la scelta di proporre una varietà di formati, con una conseguente varietà di task, permette di offrire una rappresentazione più ampia del costrutto e di proporre al candidato differenti canali d'interazione in cui dar prova della propria competenza (*range of interactional formats*, Saville & Hargreaves 1999), riducendo così gli inconvenienti derivabili dai fattori discussi in questo contributo. È noto come ogni formato d'esame presenti vantaggi e svantaggi, ed è altrettanto noto come il grado di adeguatezza del formato vari in relazione agli scopi per cui è utilizzato (Bygate *et al.* 2001). Ne consegue perciò che una combinazione di più formati può garantire una rappresentazione più ampia del costrutto (Alderson *et al.* 1995; Taylor 1999).

La riflessione sui limiti dell'asimmetria nel contesto d'esame non può che portarci a considerare un formato alternativo, ampliamento diffuso nell'ambito delle certificazioni linguistiche, ovvero l'interazione tra pari o in gruppo (*paired formats e group formats*, O'Sullivan 1995, 2008; Fulcher 2003; Luoma 2004). La maggiore simmetria nell'interazione (Taylor 2000), offerta dal formato in coppia con un altro candidato o in gruppo tra pari, così come la maggiore attinenza con le attività in

classe (Chalhoub-Deville 2001) sono nella maggior parte dei casi in grado di determinare una reazione positiva da parte dei candidati. Nonostante ciò, la selezione casuale o intenzionale del partner, così come l'influenza reciproca tra i candidati (*interlocutor effect*), non possono arginare uno dei punti critici relativo al ruolo del somministratore / esaminatore, ovvero la determinazione di punteggi individuali a partire da un'interazione condivisa (Taylor & Galaczi 2011).

Per quanto concerne poi gli effetti del somministratore, l'auspicio è di assicurare un loro monitoraggio costante e un controllo continuo mediante un sempre più alto grado di uniformità tra le sessioni d'esame (Lazaraton 1996; O'Sullivan 2002; Field 2011). Come sostiene Weir (2005: 80), «the examiner/interlocutor's contribution to the interaction must be standardized as far as possible». Per far sì che ciò sia possibile, è necessario garantire un continuo percorso di *examiner training*, a cui abbinare un *examiner monitoring system* con sessioni periodiche che consentano di aggiornare gli attori coinvolti nella somministrazione. Il fine rimane quello di mitigare gli effetti del comportamento del somministratore sulla performance del candidato. Rientra in tale questione la riflessione sullo stato emotivo del candidato e sulle condizioni di possibilità del somministratore di gestire le emozioni transitorie che il candidato può provare in contesto d'esame (ansia, preoccupazione, ...). Un'adeguata formazione dei somministratori sui modi per affrontare queste emozioni transitorie può aiutare i candidati a gestire ciò che accade nel corso dell'esame e a raggiungere la migliore performance di cui i candidati stessi sono capaci (O'Sullivan & Green 2011). In particolare, l'adozione di un *examiner script / frame* (Weir & Milanovic 2003) può garantire una maggiore equità e consistenza tra le somministrazioni e, dunque, ridurre ma non eliminare, il margine di variabilità ed incertezza che ad ogni processo di valutazione si lega.

Come rilevato altrove (Machetti 2023), la verifica, la misurazione e la valutazione della L2 rappresentano nel loro insieme un processo dinamico e non-discreto, una mediazione tra le procedure usate, la lingua, le modalità con cui tali procedure sono gestite e percepite da chi in tale processo è coinvolto. È all'interno di questo quadro di riferimento che la valutazione linguistica, un processo costantemente in precario e variabile equilibrio tra esplicitezza, incertezza, variabilità, vaghezza, richiede a chi ne fa a diverso titolo uso innanzitutto la conoscenza approfondita di un simile scenario, conoscenza che assume la veste di un'assunzione di responsabilità, indispensabile anche solo per tentarne la gestione.

Bibliografia

- Alderson, J. Charles & Clapham, Caroline & Wall, Dianne. 1995. *Language Test Construction and Evaluation*. Cambridge: Cambridge University Press.
- Bachman, Lyle. 1990. *Fundamental Considerations in Language Testing*. Oxford: Oxford University Press.
- Bachman, Lyle. 1991. What does language testing have to offer? *TESOL Quarterly* 25 (4). 671-704.

- Bachman, Lyle & Cohen, Andrew (eds) 1999. *Interfaces between Second Language Acquisition and Language Testing Research*. Cambridge: Cambridge University Press.
- Bachman, Lyle & Palmer, Adrian. 1996. *Language testing in practice*. Oxford: Oxford University Press.
- Badini Alessia & Lucarelli, Silvia & Strambi, Beatrice. 2022. Standard di qualità per la valutazione linguistica della L2: il ruolo del valutatore. In Diadori, Pierangela & Machetti, Sabrina (a cura di), *Italiano per stranieri. Lo stato dell'arte sulla verifica e la valutazione delle competenze linguistiche e didattiche*, 133-146. Firenze: Franco Cesati Editore.
- Barni, Monica & Bandini, Anna & Strambi, Beatrice & Sprugnoli, Laura & Lucarelli, Silvia & Scaglioso, Anna Maria, & Arruffoli, Anna Maria. 2009. *Linee guida CILS, Certificazione di Italiano come Lingua Straniera*. Perugia: Guerra Edizioni.
- Barni, Monica & Machetti, Sabrina. 2019. Le certificazioni di italiano L2. In Diadori, Pierangela (a cura di), *Insegnare italiano L2*, 122-133. Firenze: Le Monnier.
- Bygate, Martin & Skehan, Peter & Swain, Merrill (eds). 2001. *Researching Pedagogic Tasks: Second Language Learning, Teaching and Assessment*. London: Pearson.
- Brown, Annie. 1995. The effect of rater variables in the development of an occupation-specific language performance test. *Language Testing* 12(1). 1-15.
- Brown, Annie. 2003. Interviewer variation and the co-construction of speaking proficiency. *Language Testing* 20(1). 1-25.
- Brown, Annie. 2012. Interlocutor and rater training. In Fulcher, Glenn & Davidson, Fred (eds.), *The Routledge Handbook of Language Testing*, 413-425. Milton Park: Taylor & Francis.
- Brown, Annie & Lumley, Tom. 1997. Interviewer variability in specific-purpose language performance tests. In Ari, Huhta & Viljo, Kohonen & Liisa, Kurki-Suonio & Sari, Luoma (eds), *Current Developments and Alternatives in Language Assessment*, 137-150. Jyväskylä: Centre for Applied Language Studies, University of Jyväskylä.
- Chalhoub-Deville, Micheline. 2001. Task-based assessments: characteristics and validity evidence. In Martin, Bygate & Peter, Skehan & Merrill, Swain (eds), *Researching pedagogical tasks*, 210-228. London: Longman.
- Galaczi, Evelina & ffrench, Angela. 2011. Context validity. In Taylor, Lynda (ed), *Examining speaking. Research and practice in assessing second language speaking*: 112-170. Cambridge: Cambridge University Press.
- ffrench, Angela. 2003. The change process at the paper level, Paper 5 Speaking. In Cyril, Weir & Milanovic, Michael, *Continuity and Innovation: Revising the Cambridge Proficiency in English Examination 1913-2022*, 367-471. Studies in Language Testing 15. Cambridge: Cambridge University Press.
- Field, John. 2011. Cognitive validity. In Taylor, Lynda (ed), *Examining speaking. Research and practice in assessing second language speaking*, 65-111. Cambridge: Cambridge University Press.
- Fulcher, Glenn. 2003. *Testing Second Language Speaking*. Harlow: Longman-Pearson Education.
- Fulcher, Glenn & Davidson, Fred (eds.) 2012. *The Routledge Handbook of Language Testing*. London: Routledge.

- Lazaraton, Anne. 1996. Interlocutor support in oral proficiency interviews: the case of CASE. *Language Testing* 13(2). 151-172.
- Luoma, Sari. 2004. *Assessing Speaking*. Cambridge: Cambridge University Press.
- Machetti, Sabrina. 2023. *Essere vaghi per essere precisi. La vaghezza e la valutazione della L2*. Pisa: Pacini Editore.
- Machetti, Sabrina & Purpura, James E. 2018. Examining Claims Underlying the B1 Level CILS Exam: A Certification Test of Italian as a Foreign Language, paper presented at the *Assessing World Languages Conference 2*. Macau, October 2018.
- Masillo, Paola. 2021. Lingua e cittadinanza italiana: uno studio sulla validità della valutazione linguistica per la cittadinanza. *SILTA* L(1). 249-268.
- Masillo, Paola & Strambi, Beatrice (2020). La valutazione delle produzioni scritte in contesto certificatorio: uno studio di validazione, in Nuzzo, Elena & Santoro, Elisabetta & Vedder, Ineke (a cura di), *Valutazione e misurazione delle produzioni orali e scritte in italiano lingua seconda*, 93-105. Firenze: Franco Cesati Editore.
- McNamara, Tim & Lumley, Tom. 1997. The effect of interlocutor and assessment mode variable in overseas assessments of speaking in occupational settings. *Language Testing* 14 (2). 140-159.
- Nakatsuhara, Fumiyo. 2009. *Conversational Styles in Group Oral Tests: How is the Conversation Co-constructed?* Unpublished phd thesis. University of Essex.
- Nakatsuhara, Fumiyo & Khabbzbashi, Nahal & Inoue, Chihiro. 2021. *Assessing speaking*. London: Routledge.
- O'Sullivan, Barry. 1995. Oral Language Testing: Does the age of the interlocutor make a difference? Unpublished phd thesis. University of Reading.
- O'Sullivan, Barry. 2000. Exploring gender and oral proficiency interview performance. *System* 28(3). 373-386.
- O'Sullivan, Barry. 2002. Learner acquaintanceship and oral proficiency test pair-task performance. *Language Testing* 19(3). 277-295.
- O'Sullivan, Barry. 2008. *Modelling Performance in Tests of Spoken Language*. Frankfurt: Peter Lang.
- O'Sullivan, Barry & Weir, Cyril & Saville, Nick. 2002. Using observation checklists to validate speaking-test tasks. *Language Testing* 19(1). 33-56.
- O'Sullivan, Barry & Green, Anthony. 2011. Test taker characteristics. In Taylor, Lynda (ed), *Examining speaking. Research and practice in assessing second language speaking*, 36-64. Cambridge: Cambridge University Press.
- Ross, Steven & Bermick, Richard. 1992. The discourse of accommodation in oral proficiency interviews. *Studies in Second Language Acquisition* 14(2). 159-176.
- Saville, Nick & Hargreaves, Peter. 1999. Assessing speaking in the revised FCE. *ELT Journal* 53(1). 42-51.
- Taylor, Lynda. 1999. *Study of qualitative differences between CPE individual and paired speaking tests*. Internal UCLES EFL report.
- Taylor, Lynda. 2000. Investigating the paired speaking test format. *Research Notes* 2. Cambridge ESOL. 9-12.

- Taylor, Lynda (ed.). 2011. *Examining Speaking. Research and practice in assessing second language speaking*. Cambridge: Cambridge University Press.
- Taylor, Lynda & Galaczi, Evelina. 2011. Scoring validity, in Taylor, Lynda (ed), *Examining speaking. Research and practice in assessing second language speaking*, 171-233. Cambridge: Cambridge University Press.
- Vedovelli, Massimo (a cura di). 2005. *Manuale della certificazione dell'italiano L2*. Roma: Carocci.
- Weir, Cyril. 2005. *Language Testing and Validation: An Evidence-Based Approach*. Basingstoke: Palgrave Macmillan.
- Weir, Cyril & Milanovic, Michael (eds). 2003, *Continuity and Innovation: Revising the Cambridge Proficiency in English Examinations 1913-2002*. Cambridge: Cambridge University Press.
- Young, Richard & Milanovic, Michael. 1992. Discourse variation in oral proficiency interviews. *Studies in Second Language Acquisition* 14(4). 403-424.