

ANALISI DI PROVE DI ASCOLTO IN ESAMI DI CERTIFICAZIONE ITALIANI (LIVELLO B2)

Italia weist bezüglich Sprachzertifizierung eine Besonderheit auf: Gleich vier Zertifikatsprüfungen sind vom italienischen Bildungsministerium anerkannt: CELI, CILS, PLIDA und .IT. Diese Vielfalt bietet Gelegenheit, die Prüfungen zu vergleichen und die einzelnen Items auf ihre Validität hin zu überprüfen.

161 Maturanden aus Schweizer und Südtiroler Gymnasien haben die vier Hörverständnisübungen auf dem Niveau B2 durchgeführt und zudem die Prüfungen bezüglich ihrer Attraktivität bewertet. Die folgende, eigens für Babylonia durchgeführte Untersuchung legt die Resultate der vier Prüfungen dar und zeigt die Schwachstellen der jeweiligen Prüfungen auf. Die Lektüre dieses Beitrags empfiehlt sich besonders für jeden, der selbst Multiple-Choice-Prüfungen durchführt und um angemessene und unmissverständliche Fragestellungen und Distraktoren ringt.

Die Prüfungen sind auf der Babylonia-Homepage abrufbar:
babylonia.ch > Nummer 2/2016 > Artikel Torresan

● Paolo Torresan | Venezia

Paolo Torresan Ha svolto attività di insegnamento presso l'Università dello Stato di Rio de Janeiro (*Visiting Professor*), il Santa Monica College (*Fulbright Visiting Scholar*) e l'Università di Catania (sede di Ragusa). Ha svolto attività di ricerca presso l'Università Ca' Foscari di Venezia e la Lancaster University.



1. Il sistema plurale delle certificazioni italiane

I *proficiency test* nell'ambito dell'italiano come lingua straniera riconosciuti dal Ministero degli Affari Esteri italiano sono rappresentati dalle 4 certificazioni ufficiali:

- > CELI, Università per Stranieri di Perugia
- > CILS, Università per Stranieri di Siena
- > PLIDA, Società Dante Alighieri
- > .IT, Università degli Studi Roma Tre

Altre certificazioni, come gli esami AIL dell'Accademia di Firenze¹, o gli esami CLIDA² del Centro Linguistico Italiano Dante Alighieri (con sedi a Firenze e Roma), pur essendo implementate in alcuni contesti (es. in Svizzera nel primo caso; in vari paesi, sotto il patrocinio dell'Unione Latina, nel secondo), ricadono fuori dal riconoscimento ministeriale.

Dinanzi a questa pluralità di proposte, domande ricorrenti sono:

- > Perché quattro certificazioni e non una, per una lingua, qual è l'italiano, che non riscuote un interesse internazionale al pari dell'inglese?
- > Esistono dei confronti tra una certificazione e l'altra?

Alla prima domanda non siamo in grado di rispondere, visto che comporta uno sguardo di tipo storico e sociologico esterno alle nostre competenze. Alla seconda rispondiamo che non esistono studi comparativi in termini di risultati, quanto descrizioni, alcune delle quali molto raffinate (Novello, 2014). C'è da riconoscere, ad ogni modo, che una comparazione in termini di risultati tra le 4 certificazioni non è praticabile in sé e di per sé. Si dovrebbe supporre che uno stesso campione si sottoponga alle varie prove; i tempi di gestione di una impresa del genere sono enormi (almeno un giorno per certificazione); inoltre esiste un 'effetto-accumulo', tale per cui, passando dalle prove della prima certificazione alle prove della quarta, il candidato acquisirebbe maggiore competenza (o perlomeno guadagnerebbe in *expertise* nello svolgimento di esercizi simili) e tale fattore inciderebbe a favore della certificazione somministrata per ultima.

In luogo di una comparazione 'in blocco', esiste però la possibilità di un'analisi parallela di singole prove omologhe. Tale indagine ci risulta non essere mai sta-

PIÙ ARTICOLI SU QUESTO TEMA:
WWW.BABYLONIA.CH >
ARCHIVIO TEMATICO > SCHEDA 6

ta eseguita nell'ambito dell'italiano ed è l'oggetto del nostro saggio.

2. Item Analysis parallelo: presentazione della ricerca

In seno a tre corsi di formazione, svoltisi in tempi assai ravvicinati (dicembre 2015-gennaio 2016), tra la provincia di Bolzano (corso promosso dall'Ispettorato delle scuole di lingua tedesca) e la Svizzera (corsi tenuti rispettivamente presso l'Università di Friburgo e l'Alta Scuola Pedagogica di Zurigo), abbiamo raccolto, grazie alla disponibilità di colleghi, le risposte di un campione omogeneo (studenti maturandi) a prove omolo-

ghe (ascolto, livello B2, quesiti a scelta multipla)³, tratte dalle 4 certificazioni di italiano, liberamente consultabili *online*⁴. Il lettore può scaricare le prove dai siti riportati in bibliografia. Ai candidati è stato inoltre somministrato un questionario attraverso il quale, oltre a segnare le preferenze per le singole prove attraverso una scala Likert, ciascuno era libero di esprimere i propri giudizi in merito alla sperimentazione (§ 2.4). Nella tab. 1 rappresentiamo le caratteristiche delle prove.

Tab. 1. Caratteristiche generali delle prove

	CELI	CILS	PLIDA	.IT
Sessione	estiva 2007	estiva 2012	non dich.	non dich.
Minutaggio soundfile, senza istruzioni	9'12"-12'46"	10'56"-14'09"	28'03"-31'42"	1'05"-3'45"
Durata	3'34"	2'13"	3'45"	2'50"
Numero di ascolti	2	2	2	2
Tipologia	intervista	intervista	monologo	intervista
Dominio	sport	un'offerta formativa	storia dell'urbanistica	un'attività commerciale
Numero item	5	7	5	5
Introduzione al tema	✓			
Item-esempio	✓			✓
Layout chiaro ⁵	✓	✓		
Testo autentico ⁶	✓		✓	✓
Citazione della fonte			✓	✓
Ortoepia naturale	✓	✓	✓	✓
Testo interessante	✓		✓	✓
Testo chiaro	✓		✓	✓
Punteggio solo positivo		✓		non dich.
Trascrizione		✓		
Soluzioni	✓	✓	✓	✓

1 Gli esami AIL, nell'ambito della valutazione delle abilità di comprensione, constano soprattutto di prove di *vero/falso*, senza peraltro provvedere ad alcuna misura che consenta di arginare l'arbitrarietà che il *vero/falso* comporta. Si tratta, quindi, di un esame che presenta margini di criticità sin dalla scelta dei *format*.

2 Gli esami CLIDA, nell'ambito della valutazione delle abilità di comprensione, presentano soprattutto *domande aperte*, con numerose ingenuità nella confezione degli *item*, come l'*interdipendenza*, tale per cui si può rispondere a una domanda leggendo una successiva (es. prova CLIDA P3 2014: domanda 1 – "Dove è stato il Campionato Mondiale di Ciclismo 2013?"; Domanda 4 – "Quanti «fan» sono stati presenti lungo le strade tra Firenze e Fiesole?") o domande a cui è facile rispondere grazie a preconcose (es. prova CLIDA P3 2011: domanda 3 – "Dov'è nato il melodramma?").

3 Una leggera differenza è data dal fatto che la prova CILS consta di 7 *item* anziché 5, come le altre; ma ciò non crea difficoltà in termini di comparazione.

4 La prova tratta dal PLIDA corrisponde al vecchio formato; il centro certificatore sta rivedendo l'impianto della certificazione; per ora sono uscite le specifiche del nuovo formato del B1 (Cardillo, Vecchio & Menzinger, 2015), nel corso del 2016 usciranno quelle relative al B2.

5 Si tratta di un giudizio soggettivo. Nel caso del PLIDA, la distribuzione della prova su due fogli (fronte/retro) rende poco agevole, secondo noi, l'esecuzione. Nel caso di .IT l'*item*-esempio è stato inteso, da una buona parte del campione, come facente parte della prova.

6 Anche questo è un nostro giudizio: il testo pare essere studiato a tavolino; inoltre la concettosità gioca a sfavore dell'interesse, da parte di chi ascolta, e della chiarezza del messaggio (in riferimento ai parametri sottostanti nella tabella 1), come del resto emerge dal *feedback* del campione.

Nella tab. 2 riportiamo il tempo a disposizione per leggere i quesiti secondo le indicazioni fornite nei file.

	prima del primo ascolto	tra il primo e il secondo ascolto	dopo il secondo ascolto	tot.
CELI	20" (60" per tutte e tre le prove di ascolto)	25"	35"	1'20"
CILS	-	60"	120"	3'
PLIDA	60"	-	60"	2'
.IT	indicazioni non presenti nell'audio			

La relativa facilità della prova .IT è evidente se si considera la percentuale degli studenti che hanno raggiunto l'eccellenza, e cioè che hanno risposto correttamente a 5 item su 5: 45,3%, cifra ben superiore rispetto a quelle relative a quanti hanno raggiunto l'eccellenza nelle altre prove (tab. 3)⁸. Se anzi, si somma la percentuale che ha raggiunto l'eccellenza (5 su 5) con quella che ha raggiunto il livello pienamente soddisfacente (4 su 5), colpisce ancor di più la cifra che si ottiene: 74,5 %; ciò significa quasi uno studente su 3 ha superato la prova con risultati più che soddisfacenti.

La relativa difficoltà della prova CELI, come dicevamo, è immediatamente inferibile considerato il profilo della curva che presenta una distribuzione pressoché normale (con un leggerissimo disallineamento a destra). Si tratta di un profilo che non ci si attende con un campione rappresentativo come il nostro, cioè con studenti che si stanno preparando sistematicamente ad affrontare una prova di questo livello. Si evidenziano due sottogruppi all'interno del campione, i quali ruotano attorno alla sufficienza: il primo appena al di sopra, totalizzando 3 item su 5 (31%), il secondo appena al di sotto, totalizzando 2 item su 5 (28%).

Il profilo della prova CILS è simile a quello della prova .IT, salvo presentare una percentuale molto bassa in riferimento all'eccellenza (solo 13 candidati hanno risposto a tutti gli item: 8,1%). Ciò si spiega con la forte sovracalibrazione di un item (# 3), come avremo modo di vedere nel paragrafo successivo.

Nel complesso pare che alla prova PLIDA corrisponda il profilo più armonico e atteso.

2.2. La calibrazione degli item

Nella tab. 4 rappresentiamo le percentuali di studenti che hanno risposto correttamente ai singoli item (*Facility Value*, d'ora in poi FV).

A quali valori attenerci, sotto i quali o oltre i quali si danno, rispettivamente, casi di sotto- e di sovracalibrazione? Nella letteratura i pareri sono discordanti, con oscillazioni significative (per un prospetto: Green, 2013). A nostro giudizio, un ente certificatore, nella più flessibile delle ipotesi, si può orientare attorno all'intervallo 35-90, e cioè ritenere che con un $FV \leq 35$ vi sia *evidente* sovracalibrazione (vale a dire che se solo un candidato su 3 indovina la risposta, l'item è *evidentemente* difficile per il livello *target*), così come, al contrario, con



Tab. 2. Tempi di lettura dichiarati nei soundfile prima, tra e dopo gli ascolti

7 Va precisato, in ogni caso, che l'analisi condotta, poiché limitata a una singola prova e soprattutto dato il numero ristretto di item, non è rappresentativa, non consente cioè di formulare un giudizio sul modo con cui le singole certificazioni valutano la comprensione orale a livello B2. Ciò che emerge ha, piuttosto, un valore circoscrivibile ai soli dati a disposizione. L'impianto metodologico dell'indagine può valere comunque da riferimento per analisi più estese.

8 Compresa la CILS, per la quale è necessario realizzare una proporzione.

Fig. 1. Frequenza dei punteggi



Le *research question* che ci siamo posti sono state:

- > I risultati del campione variano *significativamente* da prova a prova?
- > Esistono casi evidenti di sotto- e/o sovracalibrazione?
- > Esistono casi di cattiva scrittura degli item?
- > Qual è il *feedback* dei candidati?

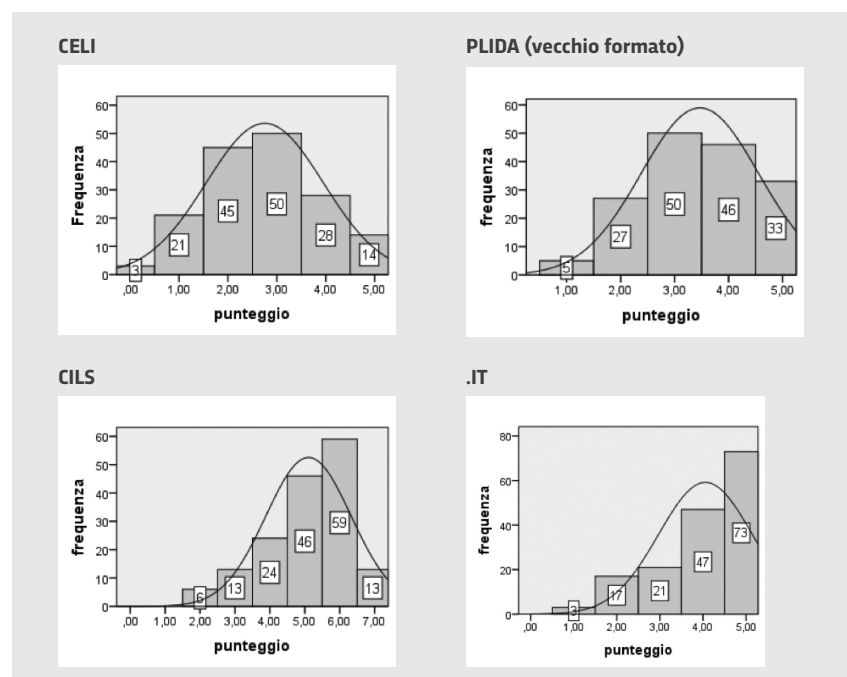
Consideriamo ciascuna nei paragrafi a seguire⁷.

2.1. Variazione tra le prove

Da un'analisi degli istogrammi alla fig. 1. si evince che i risultati variano significativamente da prova a prova.

Agli estremi abbiamo due prove:

- > .IT, la più facile, con un profilo fortemente disallineato a destra
- > CELI, con un profilo compatibile con una distribuzione normale.



un $FV \geq 90$ vi è *evidente* sottocalibrazione (se oltre 9 studenti su 10 indovinan la chiave, l'item è *evidentemente* facile in riferimento al livello *target*). In termini di ricerca, sempre nell'ottica di un contesto certificatorio, suggeriamo valori più ristretti: 40-85; e cioè con item con

- > $FV \leq 40$ abbiamo una sovracalibrazione che merita di essere indagata
- > $FV \geq 85$ abbiamo una sottocalibrazione che, al pari, costringe a un'analisi accurata.

Sempre alla tab. 4, i valori che non rispettano i parametri da noi definiti sono evidenziati così:

- > fondino grigio scuro, item sovracalibrati ($FV \leq 40$)
- > fondino grigio chiaro, item sottocalibrati ($FV \geq 85$)

Alla domanda, dunque, se si diano *item* sotto- o sovracalibrati, la risposta è affermativa. Abbiamo, in effetti:

- > due *item* sovracalibrati nella prova CELI
- > due *item* orientati in maniera opposta nella prova CILS, uno sotto- e uno sovracalibrato
- > un *item* sottocalibrato nella prova PLIDA
- > tre *item* sottocalibrati nella prova .IT

In alcuni casi, come avremo modo di appurare, i casi di sotto- e sovradimensionamento si spiegano per via di problemi legati alla confezione degli item stessi.

2.3. La qualità degli item

La tab. 5 riprende la tab. 4: riportiamo ancora una volta gli FV dei singoli item. In aggiunta, rappresentiamo con un doppio asterisco ed evidenziati in grassetto i casi di cattiva fattura degli item. Come anticipato, in alcuni frangenti, l'errato dimensionamento è attribuibile a una scrittura poco sorvegliata dei quesiti.

In sintesi, più casi richiederebbero una riformulazione dei quesiti e un conseguente riadattamento delle prove:

- > CELI, tre casi: *item* ## 1,2,4
- > CILS, cinque casi: *item* ## 1,2,3,6,7
- > PLIDA, un caso: *item* # 2
- > .IT, tre casi: *item* ## 1,2,3

Ancora una volta, la prova PLIDA pare essere la meno approssimativa, mentre la prova CILS comporta il numero maggiore di criticità. A seguire rappresentiamo, caso per caso, la natura di tali problemi, servendoci dell'analisi dei distrattori.

	0	1	2	3	4	5	6	7
CELI	1,9	13,0	28,0	31,0	17,4	8,7		
CILS	0,0	0,0	3,7	8,1	14,9	28,6	36,6	8,1
PLIDA	0,0	3,1	16,8	31,0	28,6	20,5		
.IT	0,0	1,9	10,6	13,0	29,2	45,3		

Tab. 3. Percentuali dei punteggi raggiunti dal campione
n=161

	1	2	3	4	5	6	7
CELI	55,9	27,3	72,7	37,9	82,6		
CILS	75,8	76,4	37,9	71,4	83,9	93,2	72
PLIDA	88,2	54,7	60,9	73,9	70,2		
.IT	90,1	88,8	71,4	67,2	87,6		

Tab. 4. Facility Value
n=161

	1	2	3	4	5	6	7
CELI	55,9**	27,3**	72,7	37,9**	82,6		
CILS	75,8**	76,4**	37,9**	71,4	83,9	93,2**	72**
PLIDA	88,2	54,7**	60,9	73,9	70,2		
.IT	90,1**	88,8**	71,4**	67,2	87,6		

Tab. 5. Item di cattiva fattura
n=161

CELI

Item #1. Problema: distrattore parzialmente vero

Il distrattore [b] riceve un numero considerevole di consensi (31,1%) ed è, a nostro giudizio, parzialmente vero. Il protagonista infatti è ambiguo nella formulazione del suo stato d'animo: dichiara che, alla fine della maratona, si è sentito "benissimo", ma subito dopo smentisce ("a parte gli scherzi"), e lascia presagire che il dimagrimento è stato un fattore di rilievo (i chili persi sono stati una "bella quantità"), per poi riaffermare, comunque, la sua "soddisfazione" per l'impresa. Questo contorsionismo logico avrebbe dovuto mettere in guardia il certificatore, che avrebbe potuto orientarsi verso un altro testo o, tutt'al più, scegliere un altro punto del testo verso cui orientare l'attenzione dell'ascoltatore, evitando peraltro di far riferimento ad atteggiamenti (inferibili, stando al Quadro, a livello C, ma di fatto suscettibili di interpretazioni che variano da cultura a cultura, e spesso da persona a persona). Riportiamo la trascrizione dell'audio, con evidenziati i passaggi cui facciamo riferimento:

Ci si sente benissimo, cioè, quando vinci ti passa tutto. No, a parte gli scherzi, è stata dura, ho perso sei chili sul campo, che, su di me che sono abbastanza magro, è una bella quantità di muscoli che ho perso. Eh, ci si sente molto bene, è una soddisfazione incredibile, e quindi sia il caldo che il freddo, che la fame, che la sete, passano poi in secondo piano, però indubbiamente sono problemi con i quali hai a che fare quando fai questo genere di corse.

A riprova di quanto detto, si consideri il fatto che si danno ben 5 risposte doppie (alcune delle quali realizzate dagli studenti più competenti).

Item #2. Problema: doppia o tripla chiave.

I distrattori [b] e [c] ricevono un numero alto di consensi, rispettivamente 22,4% e 32,9%.

A ben vedere, sia l'uno che l'altro sono inferenze possibili, sulla base di quanto detto dal testo.

Per quanto riguarda il distrattore [b], in particolare (secondo il quale il maratoneta "corre per distrarsi dal lavoro"), va riconosciuto che il paragone, presente nell'audio, tra l'intervistato e i personaggi del romanzo di Oscar Wilde, dott. Jackill e Mr Hyde (paragone che, peraltro, dà per scontata una conoscenza letteraria), lascia presagire che, se è vero che l'intervistato "svolge una tranquilla attività lavorativa" (alla pari di un dottor Jackill), come recita la chiave e a riflesso dell'"attiva, sedentaria e pacata attività di farmacista" documentata nell'audio, è pur vero che l'"esplosione della passione", di cui si parla ancora nell'audio, può essere colta come una reazione al "sedentario" e "pacato" mondo del farmacista: rientriamo ancora nella logica degli atteggiamenti. Il distrattore [c], secondo il quale lo sportivo "pratica molti sport estremi" è altrettanto plausibile: nel testo si parla di "corse estreme" in generale, inoltre si fa accenno ad attraversamenti del deserto "con delle macchine" e "a dorso di cammello".

Item #4. Problema: distrattore parzialmente vero

Il distrattore [d] raccoglie il 53,4% dei consensi, superando di molto quelli riscossi dalla chiave [a], 37,9%. Anche in questo caso, l'informazione contenuta in [d], il fatto cioè che il protagonista progetti "di dividere in due tappe la corsa in Alaska", può essere colta sulla base di un'inferenza che il marcatore temporale

"per il momento", contenuto nell'audio, lascia supporre. È pur vero che nel testo il marcatore è adiacente al sintagma riferito all'aver "trovato una mezza compagnia per fare il tratto corto", tuttavia nulla vieta di pensare (non abbiamo smentite nel testo) che in un 'altro momento' il protagonista decida di percorrere il tratto lungo in compagnia di altri e potendo disporre di una lunga pausa lavorativa ("ci vuole un mese per farla").

*Un'idea ce l'ho, di fare, di andare a correre la mitica ****, una gara che si corre in inverno, in Alaska, la stessa gara che fanno con le slitte trainate dai cani. [...] Ovviamente, a piedi. Si può fare a piedi, in bicicletta o con gli sci. Per il momento, forse ho trovato una mezza compagnia per fare il tratto corto. Perché di questa **** ci sono due tappe. Una da soli 600 km, e poi c'è quella lunga da 1800. Solo che lì ci vuole un mese di tempo per farla. Ci vuole una compagnia adatta perché senno potrebbe essere pericoloso.*

Da rilevare peraltro come, anche in questo caso, due studenti competenti abbiano segnato una doppia risposta, chiave [a] e distrattore [d].

CILS

Item #1. Problemi: informazione nell'incipit, margini di ambiguità

Il FV di quest'item lascia presagire una buona discriminazione: 75,8. Tuttavia, siamo convinti che siano presenti insidie, in assenza delle quali, con ogni probabilità, l'item risulterebbe sottocalibrato. In effetti, la chiave [a] – il fatto che nel brano si parli di "nuove possibilità di lavoro" – rimanda a un'informazione ridondante: che si parli di lavoro viene ripetuto due volte nel testo (vedi grassetto). Ciò dovrebbe garantire appunto una notevole facilità (anche per via di un fenomeno di riconoscimento) nell'esecuzione dell'item.

Questa sera su Radio Montecarlo parliamo di lavoro o meglio di alternative al lavoro tradizionale. Parliamo del turismo sostenibile, cioè strutture e servizi turistici realizzati in modo tale da non distruggere il panorama, il mare e l'ambiente, perché preservare l'ambiente è etico. Sono sempre più i viaggiatori attenti alla qualità della vacanza e alla salvaguardia del paesaggio e quindi potrebbe anche essere un'idea per trovare una nuova occupazione.

In realtà, a contrastare la risposta immediata concorrono diversi elementi:

1. La prima occorrenza riferita al lavoro ("Questa sera su Radio Montecarlo parliamo di lavoro o meglio di alternative al lavoro tradizionale") è posta all'inizio del brano: l'ascoltatore non ha familiarità con item di questo tipo.
2. La seconda occorrenza è sintatticamente complessa: non è immediatamente chiaro, a chi ascolta (e forse nemmeno a chi legge la trascrizione), a chi si riferisca il sintagma "potrebbe anche essere un'idea per trovare una nuova occupazione" nell'enunciato finale dell'estratto.
3. Vi è un'incoerenza nel testo: è vero che l'intervistatrice sostiene che nel testo si parla di "lavoro", però nell'intervista si parla, piuttosto, di un'offerta formativa, dal momento che l'intero discorso del prof. Malatesta ruota attorno alla descrizione del master.
4. Il distrattore [b] (secondo il quale l'intervistatrice sosterebbe che nell'intervista parlerà dell'"ultima moda in tema di viaggi"), benché riceva un numero ridotto di consensi (16,1%), non è del tutto falsificato dal testo, dal momento che l'inciso "sono sempre più i viaggiatori attenti alla qualità della vacanza e alla salvaguardia del paesaggio", nel contesto del disorientamento di cui abbiamo detto sopra, può lasciar presagire che il nocciolo della discussione sia appunto il turismo.

Item #2. Problema: dipendenza tra gli item

All'item #2 si può rispondere attraverso la lettura dello *stem*⁹ degli item ##3,4,5.

Item #3. Problema: cripticità del testo di riferimento, imprecisione della chiave

Il testo di riferimento è caratterizzato da sintassi complessissima, con difficoltà di interpretazione, all'ascolto, anche per un nativo.

Il master affronta i temi della comunicazione turistica con particolare riferimento al turismo sostenibile, al fine di fornire elementi e basi teoriche sul turismo sostenibile e le politiche ambientali, nonché sulla progettazione legata alla comunicazione multimediale radiofonica, televisiva, editoriale per la sua promozione.

Questo passaggio rende evidente che l'audio è il recitato di un testo scritto

da parte di due attori. Il problema è la dubbia efficacia comunicativa: un ascoltatore, sintonizzato su questa fantomatica trasmissione, con grande difficoltà si sentirà motivato a iscriversi al master di cui parla l'intervistato. Il disorientamento, da parte del campione, è evidenziato dalla dispersione dei consensi tra chiave [a] e distrattori [c] e [d] (quest'ultimo, peraltro, lascia le porte aperte a possibili inferenze), con i distrattori che raggiungono ciascuno la soglia del 30%.

Oltretutto, la chiave [a] non è precisa. Recita: "Il master in Comunicazione del turismo sostenibile dell'Università di Milano Bicocca cerca di coniugare principi teorici solidi con strategie di diffusione e promozione". Riprendendo il testo citato sopra, il master piuttosto fornisce "elementi e basi teoriche" su entrambi i livelli:

- > "turismo sostenibile e politiche ambientali"
- > "progettazione legata alla comunicazione multimediale radiofonica"

Non vi è una "coniugazione" di teoria e "strategie": si forniscono, invece, principi ("elementi e basi teoriche") in riferimento sia a un ambito ("turismo sostenibile e politiche ambientali") che (= "nonché") all'altro ("la diffusione e la promozione"). Entrambi i sintagmi preposizionali "sul turismo sostenibile", "sulla progettazione legata alla comunicazione multimediale" fanno riferimento alla subordinata finale ("al fine di fornire..."), la quale, a sua volta, dipende dalla reggente con cui si apre la macroproposizione ("il master affronta...").

Paradossalmente, quindi, la complessità dell'enunciato ha tratto in inganno il certificatore stesso nell'atto di redigere la chiave.

Item #6. Problema: risalto della chiave

La chiave risalta sulle altre opzioni, dal momento che è ben più lunga e ben più elaborata. I distrattori, oltretutto, sono inoperativi.

Item #7. Problema: inadeguata distribuzione degli item

La risposta all'item #7 è 'incastrata' nella risposta all'item #6. Ciò comporta un'evidente difficoltà, da parte dell'ascoltatore, di individuare dove si trovi la risposta all'ultimo quesito: molti formandi, nei corsi da noi tenuti, giunti alla fine dell'ascolto, erano ancora in attesa dell'infor-

⁹ Per *stem* è da intendersi la frase da completare o la domanda a cui si riferiscono le varie opzioni.

mazione corrispondente. Tale criticità si riflette, peraltro, nel numero consistente di risposte mancanti da parte del campione (n=6).

PLIDA

Item #2. Problema: interpretabilità degli atteggiamenti

Il distrattore [a] attira un numero consistente di consensi (37,3%). Il distrattore si riferisce a un presunto atteggiamento "indifferente" espresso dalle magistrature nei confronti della decisione del Duca Cosimo I di smantellare le vecchie sedi per rimpiazzarle con gli Uffizi. La chiave [a] pare non determinare alcun dubbio: vi fu una reazione di "fastidio", anziché di "indifferenza", dato che, recita la voce narrante:

La costruzione [...] è un gesto di razionalizzazione [...] della vita cittadina e della burocrazia, che non fu accolto molto benevolmente dalle varie magistrature.

Il problema è che la comprensione degli atteggiamenti, come già espresso, pur se prevista dal Quadro a livello avanzato (C), presenta margini di interpretabilità. L'"indifferenza", come il "fastidio", sono sfumature che possono coprire, entrambe, la gamma delle reazioni indesiderate. In particolare, l'"indifferenza" può manifestare tanto una reazione neutra (come da etimologia), quanto una reazione negativa

(nel linguaggio comune, cfr. "reagì con indifferenza al mio saluto"), quanto, in rari casi, una reazione positiva (l'indifferenza alle avversità della vita, la "Divina Indifferenza" di Montale). Si tratta di un termine a cui, dunque, per via di polisemia, può corrispondere una ricca paletta di colori per rappresentare gli stati d'animo. Chiedere agli studenti di disambiguare tra varie opzioni, di cui una è semanticamente ambigua, è ingiustificato.

.IT

Item #1. Problema: informazione nell'incipit, non progressione

La risposta dell'item #1 è all'attacco del testo, nelle prime parole formulate dall'intervistatrice; troppo prematuramente, a nostro avviso. Da segnalare, in aggiunta, la non progressione dell'item rispetto all'esempio [0]. La risposta (data dal certificatore) all'esempio [0] si trova infatti nell'enunciato successivo a quella riferita all'item #1.

Item #2. Problema: dipendenza tra gli item

La risposta all'item #2 è desumibile dall'item-esempio [0]. Di conseguenza, l'apprendente può rispondere correttamente, anche senza ascoltare il brano di riferimento.

Item #3. Problema: nessuna chiave

La chiave segnata dal certificatore [c] non è corretta. Allo stem "Spesso, gli abiti delle bambole provengono da", la chiave [c] completa "clienti della bottega". In realtà, il testo afferma che i clienti forniscono "merletti" e più in generale "tessuti", mentre la confezione degli abiti spetta alla titolare del negozio.

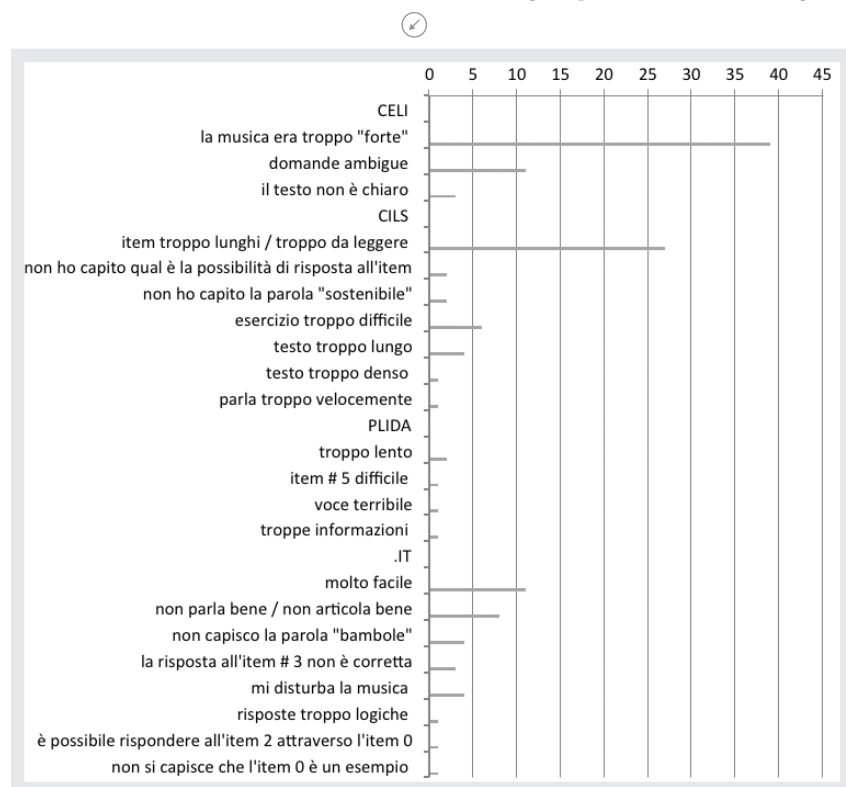
2.4. Il feedback del campione

Mentre i dati ottenuti tramite la scala Likert circa il grado di preferenza non risultano significativi, data la dispersione delle risposte, i commenti critici alle prove (fig. 2) ci illuminano su aspetti sfuggiti all'indagine, oltre a ribadire questioni già trattate durante l'analisi degli item. In particolare, quali punti finora non trattati, rileviamo:

Per la prova CELI

> Il disagio provocato dalla musica di sottofondo, che, oltre ad essere presente negli stacchi si sovrappone spesso al parlato; per alcuni, la musica catalizza l'attenzione ("se sento la musica non mi

Fig. 2. Osservazioni critiche di una parte del campione



concentro più sulle parole”)

Per la prova CILS

- › La generale complessità del testo (denso, lungo, intricato, con velocità di eloquio eccessiva)
- › La lunghezza degli item e la complessità, in particolare, di alcuni. Ciò costituisce un fattore di invalidità, data la presenza di una variabile esterna rispetto al costrutto (construct-irrelevant variance): la lettura, cioè, rispetto all’ascolto
- › L’incomprensione della parola-chiave “sostenibile”

Per la prova PLIDA:

- › La lentezza del monologo
- › La qualità della voce
- › La densità delle informazioni

Per la prova .IT:

- › La difficoltà a capire il parlato di una delle donne intervistate
- › L’incomprensione della parola-chiave “bambole”
- › La predittività delle risposte (“risposte troppo logiche”), anch’esso fattore di invalidità, in quanto a variabile esterna
- › Difetti di layout (non è chiaro che l’item 0 costituisca un esempio).

Bibliografia

Cardillo, G., Menzinger, C. & Vecchio, P. (ed.) (2015). *Livello B1 PLIDA. Quaderno delle specifiche*. Roma: Società Dante Alighieri. <http://plida.it/certificazione-plida/documenti.html> (consultazione: 18.01.2016)

Green, R. (2013). *Statistical analyses for language teachers*. Basingstoke: Palgrave MacMillan.

Novello, A. (2014). *La valutazione delle lingue straniere e seconde nella scuola. Dalla teoria alla pratica*. Venezia: Edizioni Ca' Foscari.

Sitografia

Prova ascolto B2 CELI

<http://www.cvcl.it/categorie/categoria-104> (consultazione: 18.01.2016)

Prova ascolto B2 CILS

http://cils.unistras.it/89/198/Prove_Liv._B2.htm (consultazione: 18.01.2016)

Prova ascolto B2 PLIDA (vecchio formato)

<http://plida.it/certificazione-plida/che-cos-e-la-certificazione-plida/esempi-di-prove-d-esame/esempi-sessioni-esame-plida.html> (consultazione: 18.01.2016)

Prova ascolto B2 .IT

<http://www.certificazioneitaliano.uniroma3.it/B2-intlT.aspx> (consultazione: 18.01.2016)