

Bergamo, 1975

PARTE I: I FONDAMENTI DEL *LANGUAGE TESTING*

CAP. 1 PRINCIPI METODOLOGICI

1.1. Verifica dell'apprendimento linguistico

La verifica del profitto dell'allievo è esigenza primaria di ogni insegnante che collochi la propria opera educativa in una corretta visione pedagogica, focalizzando la propria attenzione sull'allievo per conoscerne i processi di apprendimento e per accertare quanto, di ciò che l'insegnante ha cercato di trasmettere, sia stato effettivamente acquisito.

Nel campo delle lingue straniere, è particolarmente importante controllare di volta in volta che l'allievo possieda determinate abilità linguistiche: essendo una lingua, nell'accezione saussuriana, un sistema "où tout se tient", non può instaurarsi un reale apprendimento se non sviluppando adeguatamente ed in modo funzionale tutti gli elementi che entrano in gioco nel sistema.

Sul piano didattico ciò significa che non si debbono far affrontare nuove difficoltà prima di aver accertato la sicura acquisizione, da parte di ciascun allievo, di tutto ciò che si presuppone già noto. Pertanto, al di là di ciò che all'insegnante viene richiesto dalla scuola in quanto struttura burocratizzata (voti, classificazioni, pagelle, ecc.), la misurazione del profitto si configura come una necessità insopprimibile, un problema dalla cui risoluzione può dipendere, in misura notevole, il buon andamento di un corso di lingue straniere moderne.

La ricerca di nuovi strumenti, di soluzioni sempre più adeguate, ha dato origine agli studi sul "Language Testing" (LT), in cui rifluiscono i recenti sviluppi di varie discipline, tra cui la linguistica generale ed applicata, la psicolinguistica, la docimologia, la didattica delle lingue straniere e la statistica metodologica.

Premesso che il termine verrà precisato man mano che ne verranno presi in esame i singoli aspetti, definiamo il LT come *il complesso dei principi metodologici e delle tecniche che presiedono alla realizzazione, somministrazione e utilizzazione delle prove oggettive di profitto nell'apprendimento delle lingue straniere.*

La comprensione dei principi metodologici è di preminente importanza rispetto alla conoscenza di modelli e formule, in quanto permette di sviluppare su fondamenti scientifici la preparazione di materiale adatto alla situazione didattica in cui si opera. Proprio alla carenza di scientificità vanno ascritti gli aspetti negativi che il LT ha assunto nelle prime fasi del suo sviluppo.

In concomitanza col metodo tradizionale grammatica-traduzione, si ebbe l'identificazione tra *esercizio* e *test*: le versioni in lingua straniera venivano usate sia in fase di apprendimento (fissazione delle strutture e del lessico), sia in fase di controllo. Sul piano psicologico e su quello didattico, si tratta di un grave errore; il momento dell'acquisizione non può e non deve confondersi con quello della verifica; un esercizio è sostanzialmente diverso (per struttura e per funzione) da un test, anche quando entrambi si basano su materiale apparentemente simile.

Successivamente, sono state prese a modello le tecniche impiegate in psicomatria per la costruzione dei test mentali; in questo settore, infatti, si era registrato un notevole sviluppo delle prove oggettive e standardizzate ed un moltiplicarsi di pubblicazioni su tale argomento. Negli Stati Uniti, la vastissima diffusione nelle scuole dei test d'intelligenza, attitudinali e caratterologici è stata di per sé un veicolo di divulgazione ed ha fatto sì che gli insegnanti, insoddisfatti delle prove tradizionali, vi si ispirassero largamente.

Inizialmente, però, questo accostamento ai test è stato caratterizzato da un alto grado di superficialità e di approssimazione. Delle tecniche di valutazione è stato colto soprattutto l'aspetto esteriore, meccanico, e si sono create delle prove che assomigliavano sempre più ai quiz enigmistici e sempre meno ai test mentali. Mancava, alla base, un'approfondita analisi dell'oggetto da verificare, la L2¹, condotta secondo i criteri elaborati dalla scienza linguistica; nessuno strumento di misurazione è idoneo se non viene costruito tenendo conto delle caratteristiche di ciò che deve essere misurato.

Mancava altresì il riferimento costante ad un preciso itinerario didattico. Per verificare cosa è stato appreso occorre prima sapere come si è instaurato l'apprendimento, perché gli aspetti qualitativi sono preminenti rispetto a quelli quantitativi, e non ha senso sottoporre un allievo ad una verifica condotta in modi difformi rispetto al tipo di preparazione acquisita.

Date queste premesse, prima di procedere alla presentazione delle tecniche per la verifica delle abilità linguistiche, delineeremo, sia pure per sommi capi, il quadro generale nel quale si colloca il LT.

1.2. Testing e valutazione

Un esame, per quanto sommario, del tema della valutazione non può certo esaurirsi in poche righe, in quanto coinvolge direttamente tutta la problematica relativa ai fini e agli esiti dell'educazione e, quindi, anche alla natura stessa della scuola. D'altra parte, anche se il nostro discorso sul LT si muove in un orizzonte assai più ristretto, e ad un livello essenzialmente tecnico, non possiamo ignorare le controversie sulla "selezione meritocratica" che hanno agitato la scuola italiana in questi ultimi anni. Siamo infatti consapevoli delle implicazioni e degli effetti di un uso delle prove di controllo che, prescindendo da una chiara coscienza delle finalità educative, non avvertisse i pericoli di una presunta "oggettività" non sufficientemente qualificata. Proporre

¹ Qui e in seguito l'espressione "lingua due" (abbreviata L2) indica genericamente la lingua straniera oggetto di apprendimento da parte di uno studente; la situazione-tipo a cui si fa costante riferimento è quella dello studente italiano che apprende una lingua estera in una scuola italiana. Pertanto — salvo esplicite indicazioni — non terremo normalmente conto di tutte le possibili varianti relative alle interferenze dei dialetti locali, assumendo come L1 la lingua italiana "standard". Sul significato di L1 e L2 v. FREDDI, *Metodologia e didattica delle lingue straniere*, Bergamo Minerva Italica, 1970, pp. 51-55.

semplicemente tecniche nuove e più sofisticate in sostituzione delle prove di vecchio tipo è antiscientifico nella misura in cui non si tiene conto dei presupposti pedagogici e metodologici: è questa la ragion d'essere della docimologia come scienza².

Per quanto riguarda gli aspetti sociopolitici, ci limiteremo ad osservare come la valutazione scolastica sia servita prevalentemente ad instaurare e tramandare stratificazioni di classe, selezionando sulla base di un modello precostituito ed emarginando gli studenti i cui caratteri individuali sono difformi da tale modello. Sotto il profilo psicologico, ad esempio, sono stati privilegiati in misura talora eccessiva gli aspetti verbali e numerici dell'intelligenza a discapito degli aspetti psicomotori o "spaziali" che presiedono alle abilità manuali e pratiche.

Su questa diagnosi registriamo un'ampia concordanza di vedute; sono invece assai discordi i pareri sui possibili rimedi, e cogliere un orientamento preciso non è agevole, anche perché in vari casi è difficile distinguere tra le proposte e i progetti fondati su una seria impostazione sociopedagogica e quelli propugnati strumentalmente da vari gruppi politici, magari per fini diversi da quelli dichiarati³.

Possiamo tuttavia ritenere che certe tensioni tra allievi e insegnanti non si sarebbero acuite così clamorosamente se da entrambe le parti, ma soprattutto da parte degli insegnanti, fosse stata chiarita, sviluppata e tenuta sempre presente la distinzione fondamentale tra *misurazione del profitto* e *valutazione dell'allievo*, così come essa è stata elaborata dalla pedagogia contemporanea.

La misurazione del profitto consiste nella verifica oggettiva del valore di un risultato scolastico, esclusivamente sulla base della esattezza nozionale e prescindendo da ogni considerazione sull'alunno che ha fornito quel risultato. Il processo di valutazione è invece un atto di natura complessa, nel quale i dati ottenuti mediante la misurazione del profitto sono posti in relazione alla personalità degli studenti, alla loro condizione umana sotto il profilo biologico, psichico, ambientale, socioeconomico, ecc.

In ogni insegnante, perciò, accanto alla dimensione che fa di lui un tecnico, un esperto della propria disciplina, deve coesistere la dimensione dell'educatore, ossia di colui che, inserito nel gruppo di lavoro costituito dal consiglio di classe, mira ad una valida formazione globale della persona umana.

Parallelamente ad una valutazione sommativa in cui, sotto forma di sentenza, si "tirano le somme" di un processo di apprendimento, deve svilupparsi sempre più una valutazione *formativa* o *programmatica*, allo scopo di accertare le capacità esistenti non in vista di una sanzione, bensì per stabilire, per ogni alunno, un piano d'azione didattica che possa condurlo al livello più alto di istruzione e di educazione da lui raggiungibile. Infine, possiamo identificare una valutazione *orientativa*, intesa a stabilire in quali direzioni la personalità dell'allievo potrà meglio svilupparsi.

I vari aspetti del problema della valutazione, ai quali abbiamo accennato brevemente, ci offrono una visione ampia ed articolata su tutto questo settore e ci consentono quindi di cogliere le dimensioni e i confini del testing. I test di profitto sono strumenti di misurazione e non di valutazione, né sono gli unici di cui dispone l'insegnante: non sarebbero buoni educatori coloro i

² Sul termine "docimologia" e sulle origini della stessa si veda: H. PIÉRON, *Esami e Docimologia*, Roma, Armando, 1965, parte I.

³ Le alterne vicende relative agli "esami di gruppo" ed ai "voti formalizzati" nelle scuole ed università italiane dal 1968 ad oggi sono indice della crisi della valutazione scolastica, della difficoltà di realizzare soluzioni soddisfacenti e della molteplicità di motivazioni addotte a sostegno delle varie proposte: motivazioni che spaziano dal campo ideologico a quello psicologico, dal campo strettamente politico a quello produttivistico, ecc.

quali, vittime di un tecnicismo esasperato, attribuissero valore esclusivo alle prove oggettive, trascurando gli altri mezzi di conoscenza e verifica.

1.3. Testing e analisi linguistica

Non è casuale, a nostro avviso, che Robert Lado, il linguista a cui si deve una delle opere più importanti sul LT, sia anche l'esponente più qualificato della linguistica contrastiva. Infatti la verifica dell'apprendimento linguistico presuppone l'esatta comprensione del meccanismo delle interferenze della LI sulla L2, a tutti i livelli, così come, sul piano didattico, "i materiali più efficaci sono quelli basati su una descrizione scientifica della lingua da apprendere, accuratamente confrontata con una descrizione parallela della lingua materna dell'allievo"⁴

A livello fonologico, il confronto avviene attraverso una triplice indagine su ciascun fonema della L2:

- 1) La lingua materna ha un fonema foneticamente simile?
- 2) Le varianti del fonema sono simili in entrambe le lingue?
- 3) I fonemi e le loro varianti sono distribuiti in modo simile?⁵

Assumendo come L1 l'italiano e L2 l'inglese, rispondendo alla prima domanda troviamo che i fonemi inglesi / ð θ ʒ h / non corrispondono a suoni italiani simili. Ulteriori controlli (sulla base dei punti 2 e 3) rivelano, tra l'altro, che:

— il suono [ŋ] è una variante del fonema / n / in italiano, mentre funziona come fonema autonomo in inglese;

— in inglese, ma non in italiano, / s / e / z / sono capaci di costituire coppie di opposizione minimale (ice-eyes);

— esiste in inglese, ma non in italiano, una variante velare del fonema / l /, ricorrente in posizione post-vocalica.

L'indagine, che ovviamente deve essere molto più accurata ed approfondita di quanto non possa apparire dalla sommaria esemplificazione che ne abbiamo dato, abbraccerà non solo i fonemi segmentali (vocali, consonanti, dittonghi) ma anche i valori soprasegmentali (accento, intonazione); inoltre, poiché le differenze nella distribuzione dei fonemi sono la causa di un gran numero di problemi di apprendimento, dovranno essere esaminate contrastivamente tutte le sequenze fonemiche, ad esempio i nessi consonantici in posizione iniziale, mediana e finale.⁶

A livello grammaticale, il Lado⁷ postula la necessità di accertare se ad una struttura della L2 corrisponda, nella LI, una struttura segnalata allo stesso modo (ossia con gli stessi mezzi formali), avente lo stesso significato, e analogamente distribuita nel sistema.

⁴ C.C. PRIES, *Teaching and Learning English as a Foreign Language*, Ann Arbor, University of Michigan Press, 1945, p. 9. Sul problema dell'interferenza si vedano: U. WEINREICH, *Languages in Contact*, The Hague, Mouton, 1970; G. FREDDI, *Metodologia e didattica delle lingue straniere*, cit. cap. 9° pp. 27-254, e il contributo di E. SONINO, *Linguistica contrastiva e insegnamento delle lingue*, in "Lingue e civiltà", Brescia, C.L.A.Di.L., a. II N. 1 pp. 3-5.

⁵ R. LADO, *Linguistics across cultures*, Ann Arbor, University of Michigan Press, 1957, p. 13.

⁶ Per *nesso consonantico* si intende quindi l'incontro di due o più consonanti o nella stessa sillaba (come nella prima sillaba di *strano*) o tra le vocali di due sillabe consecutive come in *inglese*).

⁷ R. LADO, *Linguistics across cultures*, cit., p. 66.

Se volessimo confrontare la formazione del plurale in italiano e in francese, dovremmo - anzitutto descrivere il fenomeno in entrambe le lingue:

a) ITALIANO

Forma: Il plurale si ottiene mediante sostituzione dei morfemi inflessionali vocalici, secondo il seguente schema (semplificato):

Singolare	Plurale
mano, buono, libro	mani, buoni, libri
casa, buona, penna	case, buone, penne
nave, agile, ponte	navi, agili, ponti
poeta, idioma, idiota	poeti, idiomi, idioti

I sostantivi monosillabici e quelli terminanti in consonante sono invariabili al plurale (re, tram, sport, ecc.).

Un certo numero di vocaboli (per lo più sostantivi) hanno forme autonome (“irregolari”) per il plurale.

Significato: più di uno (nei sostantivi, aggettivi, ecc.); nei pronomi personali il plurale indica l'aggiunta di altri referenti a quello indicato dalla corrispondente persona singolare. In altre parole, mentre “libri” significa “libro + libro + ...”, “noi” non equivale a “io + io + ...”, bensì a “io + qualcun altro”.

Distribuzione: Formano il plurale i sostantivi, gli articoli determinativi, gli aggettivi qualificativi, possessivi, dimostrativi, molti indefiniti, i pronomi, ecc.

b) FRANCESE

Forma: Occorre in primo luogo distinguere il messaggio scritto da quello orale, come appare dal seguente esempio, in cui il segno + indica la presenza di tratti distintivi tra singolare e plurale, mentre il segno — indica l'assenza di tratti:⁸

La	porte	était	fermée	
Les	portes	étaient	fermées	
+	+	+	+	(nello scritto)
+	—	—	—	(nell'orale)

Mentre sul piano grafico ogni parola è stata modificata nella formazione del plurale, sul piano orale la sola variazione è nell'articolo.

Pertanto, nel descrivere la forma del plurale francese, diremo che mentre nello scritto essa è caratterizzata dalla desinenza -s, presente nella maggioranza dei casi, nel codice orale esiste un'alta percentuale di forme invariate.

Anche in francese esistono forme irregolari (nella grafia e/o nella pronuncia), e forme invariabili anche nella grafia. Da notare il gruppo di vocaboli terminanti in -al e -ail formanti il plurale con la desinenza -aux (sul piano orale: singolare / -al, -aj /, plurale /-o/).

⁸ G. FREDDI, *Metodologia e didattica delle lingue straniere*, cit., p. 75.

In quanto alla distribuzione e al significato del plurale francese, in linea di massima valgono le osservazioni fatte per l'italiano.

Sulla base di questa analisi, per quanto sommaria, si può presumere che le sole interferenze capaci di creare problemi di apprendimento si collochino a livello formale.

Se l'indagine si sposta su altre lingue, i risultati possono essere notevolmente diversi: nei rapporti tra italiano e inglese, ad esempio, esistono anche problemi di distribuzione del plurale, essendo invariabili, in inglese, l'articolo, gli aggettivi qualificativi e la maggior parte degli altri aggettivi e pronomi (con l'eccezione di other pronome e dei dimostrativi).

Problemi a livello di significato si evidenziano nei rapporti con quelle lingue il cui sistema comprende, oltre al singolare ed al plurale, altri numeri, quali il duale. La dissimmetria può essere così schematizzata:

N. oggetti	1	2	3 4 5 6 ...
L1	Singol.	Plurale	
L2	Singol.	Duale	Plurale

Lo schema offre la visualizzazione di un principio molto importante: nei rapporti tra due lingue, i problemi di apprendimento e di valutazione cambiano a seconda di quale delle due sia la L1 e quale la L2.

Nell'esempio, il passaggio da un sistema più semplice ad uno più complesso prefigura l'attendibilità di certe difficoltà; nel passaggio inverso, i problemi sarebbero minori, e comunque di tipo diverso.

Da ciò la necessità di elaborare materiali sulla base di una analisi contrastiva, tenendo presente la direzione nella quale opera l'apprendimento; in caso contrario, il lavoro potrebbe risolversi in un mero confronto di sistemi, interessante in sé ma di scarsa utilità per l'insegnante di lingue.⁹

È stato merito del Lado l'aver mostrato come i concetti di forma, distribuzione e significato siano operanti non solo a livello morfo-sintattico, ma anche nel confronto di due sistemi lessicali e di due culture,¹⁰ nonché l'aver indicato procedure di analisi assai feconde di risultati.

Il modello linguistico a cui il Lado si ricollega si colloca entro i confini della linguistica tassonomica, le cui origini risalgono all'opera di Bloomfield e di Sapir.¹¹ Lo strutturalismo americano si è distinto per la capacità di raccogliere materiale linguistico estremamente importante;

⁹ A titolo esemplificativo si citano, per quanto riguarda i rapporti tra le lingue inglese ed italiana, i volumi: F.B. AGARD e R. J. Di PIETRO, *The Sounds of English and Italian*, Chicago, University of Chicago Press, 1965; F.B. AGARD e R. J. Di Pietro, *The Grammatical Structures of English and Italian*, Chicago, University of Chicago Press, 1966. I rapporti tra le due lingue sono esaminati assumendo come L1 l'inglese e pertanto i problemi di apprendimento evidenziati non sono quelli che interessano direttamente gli insegnanti italiani. Recenti programmi, come il progetto ITALS del CLADiL, intendono condurre ricerche sul piano contrastivo assumendo l'italiano come L1.

¹⁰ R. LADO, *op. cit.*, capp. 4 e 6.

¹¹ E. SAPIR, *Language*, New York, Harcourt, Brace & Co., 1921; L. BLOOMFIELD, *Language*, New York, Henry Holt & Co., 1933. Sulle origini e gli sviluppi della linguistica tassonomica americana si veda l'introduzione (a cura di G. Freddi) a: R. LADO, *Per una didattica scientifica delle lingue*, Bergamo, Minerva Italica, 1974 (versione italiana di *Language Teaching. A Scientific Approach*, McGraw-Hill, New York, 1964).

ma per ciò che si riferisce agli sviluppi dei criteri di analisi linguistica, pur riconoscendo il valore degli studi di Hockett, Gleason, Fries (di cui il Lado fu allievo), e degli altri maggiori strutturalisti americani, concordiamo con coloro che giudicano la linguistica funzionale, e in particolare l'opera di Jakobson e Martinet, più rispondente all'esigenza di una visione organica del problema globale del linguaggio.¹²

Attraverso il recupero delle dicotomie saussuriane *langue/parole* e *signifiant/signifié*, l'orizzonte linguistico si è ampliato oltre i limiti della linguistica descrittiva, pur senza abbandonarne i canoni di rigorosa aderenza ai criteri fondamentali dell'analisi strutturale. Con Jakobson, salgono alla ribalta della linguistica i problemi della genesi del linguaggio (attraverso gli studi sull'afasia); l'atomo-fonema è scisso nelle sue particelle, i tratti distintivi; anche la letteratura trova nuove forme di interpretazione e di analisi. Al Martinet si deve una sintesi estremamente utile e chiara delle articolazioni del linguaggio, nonché l'enunciazione di una sintassi funzionale ricca di spunti e di indicazioni valide per una corretta impostazione della didattica delle lingue.

1.4. Testing e didattica delle lingue straniere

Un'esposizione sintetica dei fondamenti della didattica delle lingue straniere non è agevole, sia per la vastità dell'oggetto, sia per le profonde implicazioni sul piano culturale e sul piano operativo. Preferiamo perciò far riferimento ad altri volumi già pubblicati in questa collana¹³ e limitarci a presentare alcuni principi essenziali, più strettamente collegati ai problemi del LT.

- a) La lingua è strumento di comunicazione verbale; il codice scritto è una rappresentazione secondaria e approssimativa del codice orale.
- b) La didattica delle lingue straniere si basa su un'analisi condotta secondo i principi della linguistica moderna.
- c) Per grammatica di una lingua si intende essenzialmente il sistema di strutture rilevate mediante una descrizione organica e funzionale, non una serie di regole a carattere prescrittivo.
- d) Una lingua è intimamente connessa, in vari modi, con la civiltà del popolo che la parla; il termine *civiltà* è qui inteso in senso antropologico.
- e) Apprendere una lingua straniera significa essenzialmente formare un nuovo sistema di abiti fonologici, morfosintattici e lessicali.
- f) Le quattro abilità fondamentali sono: il saper comprendere, il saper parlare, il saper leggere, e il saper scrivere la lingua straniera; esse vanno affrontate e sviluppate seguendo rigorosamente l'ordine indicato.
- g) A causa dell'interferenza del codice scritto, si postula la necessità di una fase iniziale di condizionamento in cui si impieghi esclusivamente la lingua orale (*oral approach*).
- h) Non è possibile apprendere una lingua straniera se non attraverso l'impiego della lingua stessa; il discorso sulla lingua (ossia la "metalingua"), specialmente se tenuto nella lingua materna dello studente, non conduce ad un effettivo apprendimento linguistico.

¹² G. C. LEPSCHY, *La linguistica strutturale*, Torino, Einaudi, 1966 offre un panorama delle tendenze e degli sviluppi della linguistica contemporanea.

¹³ Oltre al già citato *Metodologia e didattica delle lingue straniere* di G. Freddi, ricordiamo: AA. VV., *Strutturalismo e didattica delle lingue*, Bergamo, Minerva Italica, 1967; *La civiltà nell'insegnamento delle lingue*, idem, 1968; R. TITONE, *Introduzione alla metodologia della ricerca nell'insegnamento linguistico*, idem, 1974; G. FREDDI, *Gli adulti e le lingue*, idem, 1974.

i) La lingua insegnata sarà lingua autentica, sempre presentata in situazioni tipiche della civiltà straniera; se possibile, ci si avvale dei parlanti nativi, o direttamente o mediante sussidi (registrazioni su nastro, disco, ecc., programmi radiofonici e televisivi).

l) Il programma di insegnamento è basato su un'accurata graduazione del lessico e delle strutture morfosintattiche; attraverso l'analisi contrastiva di due lingue è possibile evidenziare le dissimmetrie e prevedere le difficoltà di apprendimento.

m) Un corso o programma si articola in una serie di unità didattiche, ciascuna delle quali racchiude, in un contesto di civiltà, materiale per 5-8 ore di lezione.

n) Le fasi dell'unità didattica sono:

A) la presentazione di motivi di civiltà straniera;

B) l'accostamento globale al brano di lingua-civiltà;

C) l'induzione delle strutture morfosintattiche e del lessico;

D) la fissazione delle strutture mediante opportuni esercizi;

E) la sistemazione grammaticale;

F) la verifica dell'apprendimento.

1.5. Oggettività dei test

L'aver definito i test come prove oggettive di profitto impone che venga precisato che cosa si intende per "oggettività". Sul piano generale, essa è l'esclusione di ogni soggettività, ossia di qualsiasi intervento arbitrario dell'esaminatore, tale da rendere incostanti le prove e aleatori i risultati. Sul piano tecnico specifico, l'oggettività di un test è il prodotto di due fattori primari, la validità e l'affidabilità.¹⁴

La validità di un test è la sua attitudine a verificare l'oggetto. Un test di grammatica è valido in quanto lo si impieghi per controllare l'apprendimento della grammatica; non può essere valido se impiegato per verificare la pronuncia o la padronanza del lessico.

In base ai principi *d* e *i* (vedi paragrafo precedente) non si possono considerare validi test che non impieghino la lingua naturale di conversazione, sufficientemente situazionata, in un contesto di civiltà. Il principio *e* implica che un test di lingua deve verificare il possesso di abilità specifiche, e non deve essere rivolto a misurare altri fattori (p. es. l'intelligenza e la memoria) in misura tale da annullare il contenuto linguistico. Il principio *h* esclude che un test valido possa servirsi della lingua madre dell'alunno o della terminologia grammaticale tradizionale.

Per affidabilità di un test si intende la sua capacità di fornire dati attendibili, ossia costanti e indistorti. Tale capacità è connessa all'univocità dei quesiti ed alla quantificabilità del test. È univoco il quesito che ammette una ed una sola soluzione esatta, chiaramente predeterminata, con esclusione di varianti, risposte parziali o incomplete, soluzioni incerte che richiedano l'intervento

¹⁴ L'uso del termine "oggettivo" per i test di profitto è criticato da alcuni autori, in quanto l'oggettività è limitata al momento della correzione e della determinazione del punteggio, mentre la definizione dell'oggetto della prova e la formulazione delle domande risentono delle scelte soggettive dell'esaminatore. L'obiezione, che pure — come vedremo — non è priva di fondamento, non ci sembra però tale da indurci ad abbandonare la terminologia di uso corrente. Cfr. M. GATTULLO, *Il controllo delle risposte a caso nei test di profitto*, in "Scuola e Città", XXIII, n. 6, Giugno 1972, nota 1, p. 267. Lo stesso autore aveva usato il termine "prove oggettive" nella sua opera *Didattica e Docimologia*, Roma, Armando, 1971.

soggettivo dell'esaminatore. È quantificabile il test da cui si ricava direttamente e oggettivamente il punteggio conseguito da ciascun esaminato e/o la graduatoria di merito in un gruppo di esaminati.

L'affidabilità di un test non è espressa da valori assoluti ma da un indice di affidabilità compreso tra 0 (affidabilità nulla) e 1 (affidabilità massima); l'indice è in relazione anche al contenuto del test, in quanto, ad esempio, le prove sulla morfosintassi sono di solito più facilmente oggettivabili di quelle sulla pronuncia.¹⁵

I test possono infine essere valutati sulla base della loro praticità: il tempo che richiedono per essere approntati, somministrati e corretti, la facilità di impiego nelle normali situazioni scolastiche, la necessità di sussidi o attrezzature, ecc. Queste considerazioni sono più importanti di quanto non sembrino a prima vista, sia per l'esigenza di economizzare tempo nelle prove di controllo, in modo da non sottrarlo all'attività didattica in senso stretto, sia perché non è sempre facile far coesistere accuratezza e praticità, ed è inutile costruire test tanto sofisticati ed elaborati da risultare poi inadoperabili.

1.6. Testing e finalità educative

Non è compito di chi si occupa di testing stabilire quali siano le mete generali e specifiche di un programma di lingue straniere. L'enunciazione dei fini e degli obiettivi di un insegnamento-apprendimento deve preesistere alla formulazione delle prove attraverso cui si verifica il raggiungimento di tali fini. Non si può finalizzare un'attività didattica al mero superamento di una prova d'esame; il solo itinerario corretto è quello inverso, che adegua le prove di controllo ai contenuti ed agli scopi dell'attività didattica.¹⁶

Questo principio, la cui importanza è fondamentale, è stato disatteso in moltissimi casi. Su vasta scala, questa è tuttora la situazione per quanto riguarda gli esami di maturità: temendo l'arrivo di commissari particolarmente severi e formalisti, lo svolgimento dei programmi, soprattutto dell'ultimo anno, fa riferimento costante alle esigenze dell'esame. Non è paradossale affermare che troppe volte si fa scuola sulla base del peggior esaminatore ipotizzabile. Si tratta di un macroscopico nonsenso, dato che anche le disposizioni vigenti lasciano chiaramente intendere come sia la commissione esaminatrice a doversi adeguare alle scelte metodologiche e programmatiche degli insegnanti, e non viceversa. Resta poi da chiedersi dove possano annidarsi questi commissari-spauracchio, visto che ogni professore giura di essere una persona sensibile, aperta al dialogo, disponibile e "moderna". Al di là di ogni facile ironia, si ha l'impressione che il formalismo conformista giochi un ruolo spesso decisivo nell'impostazione delle scelte di fondo.

Nel settore specifico delle lingue, un'azione analoga viene esercitata dagli esami per il conseguimento dei titoli di studio rilasciati da università o enti stranieri. Poiché alcuni di questi diplomi godono di un indiscusso prestigio, numerose scuole di lingue orientano il loro programma verso il superamento di quegli esami e una enorme mole di materiale didattico (corsi, testi, sussidi, ecc.) viene approntata esclusivamente a tale scopo. Il fatto che le prove siano solitamente elaborate da équipes di esperti ad alto livello, se per certi aspetti garantisce la buona qualità della preparazione richiesta, finisce però per instaurare un circolo vizioso: non si ha un rinnovamento didattico senza un aggiornamento delle prove d'esame, e le prove d'esame non possono essere modificate in tempi brevi senza danneggiare coloro che hanno già impostato in un certo modo la propria preparazione. Come osservava M. West già nel 1952, "gli esaminatori sono creature

¹⁵ L'indice di affidabilità si determina normalmente mediante il calcolo della correlazione statistica (v. par. 10.5.).

¹⁶ Tra le opere che esaminano le finalità educative e gli obiettivi didattici dell'insegnamento linguistico, segnaliamo R. TITONE, *Le lingue estere*, Zurigo, PAS-Verlag, 1966, pp. 17-28.

conservatrici e tendono... ad esaminare ora così come essi stessi sono stati esaminati in passato”.¹⁷ L'effetto complessivo può quindi essere paralizzante, ostacolando ulteriormente l'evoluzione della didattica, come se già non bastassero tutte le altre remore burocratiche e psicologiche.

Il LT ha invece una duplice funzione ausiliaria nella definizione delle mete e degli obiettivi. Anzitutto, l'impostazione di prove di controllo valide esige che gli obiettivi siano formulati in termini comportamentistici. Per fare ciò occorre collocarsi nell'ottica dell'allievo esaminato e descrivere, in termini inequivocabili, quello che egli dovrà fare per dimostrare di avere imparato un certo argomento o acquisito determinate capacità; bisogna inoltre indicare il modo in cui si svolgerà la prova di controllo e quale sia la prestazione minima ritenuta sufficiente. Mancando un criterio che determini i requisiti minimi per cui una prova si ritiene superata, ogni esaminatore lo fisserà più o meno arbitrariamente secondo il proprio metro di valutazione; non di rado questo varia, nello stesso insegnante, da una somministrazione all'altra della prova, sotto l'influenza di fattori estranei alla prova stessa.

A titolo esemplificativo, consideriamo il caso di un insegnante che voglia verificare se i suoi studenti “sanno una certa poesia”. In concreto, che cosa chiederà? Conoscere una poesia in lingua straniera, a seconda delle circostanze e del livello della classe, può voler dire saperla leggere senza errori di pronuncia, saperla tradurre nella lingua materna, saperla recitare a memoria, saperla commentare criticamente, o altro ancora. Sono tutte attività che rientrano nel “saper parlare” e riferite allo stesso oggetto, ma diversissime tra loro per il tipo di prestazioni richieste, per il grado di difficoltà e complessità, per il modo in cui si potrà misurare il profitto. L'uso di termini astratti e generici può dare luogo a una carenza di informazioni che a sua volta può causare incomprensioni e conflitti, come quelli che emergono quando lo studente si trova a dover effettuare delle prove diverse da quelle che si aspettava di sostenere e per le quali si era preparato.

Non riteniamo che l'arbitrarietà e l'indeterminatezza, proprie di tutte le relazioni interpersonali, possano essere totalmente eliminate dal rapporto insegnante-alunno, anche solo in ordine alla misurazione del profitto; crediamo però che esse debbano interferire nella misura minore possibile, e che sia quindi auspicabile un maggior impegno verso una più dettagliata comunicazione degli obiettivi e dei criteri di valutazione.¹⁸

La seconda funzione ausiliaria del testing si esplica nelle sperimentazioni didattiche. Perché il termine sperimentazione non sia un'etichetta abusiva data ad iniziative scolastiche più o meno valide, è necessario che gli insegnanti seguano una rigorosa metodologia sperimentale e che possiedano strumenti precisi di misurazione del profitto: questi permetteranno di confrontare i risultati degli allievi con i quali è stata effettuata la sperimentazione (gruppo sperimentale) con i risultati degli altri allievi (gruppo di controllo). Solo attraverso accurate verifiche si possono rilevare oggettivamente gli effetti positivi o negativi di una sperimentazione; in caso contrario, ogni giudizio sul successo dell'esperimento è lasciato alla valutazione soggettiva di chi l'ha condotto.¹⁹

¹⁷ M. WEST, *Examinations in a Foreign Language*, in “E.L.T. Selections I”, a cura di W. R. Lee, Londra, Oxford University Press, 1967, p. 193.

¹⁸ v. G. PORCELLI, *Obiettivi comportamentistici nello studio delle lingue*, in “Lingua e Civiltà” V, n. 2, Maggio 1972, pp. 15-19.

¹⁹ Questa impostazione del problema della sperimentazione trova conferma, a livello legislativo, nel Decreto Delegato n. 419 (D.P.R. 31 Maggio 1974); si veda in particolare il secondo comma dell'art. 3.

1.7. Prognosi e attitudini

Pur non rientrando nella tematica del LT in senso stretto, i test di attitudine linguistica meritano un cenno a parte, in relazione all'interesse che essi suscitano nell'insegnante di lingue straniere. In effetti le capacità finali C, rilevabili al termine di un processo di apprendimento, equivalgono alla somma delle attitudini A e dell'addestramento specifico o "esperienza" E secondo la formula $A + E = C$. Pertanto un giudizio su un'esperienza educativa deve tener conto, oltre che dei risultati finali, delle attitudini dei discenti.

Lo studio delle attitudini linguistiche, inoltre, ci rivela quali processi mentali sono coinvolti nell'apprendimento linguistico, orientando così la nostra azione didattica; esso serve anche a costruire i test di attitudine attraverso cui si può fare la prognosi del successo e delle difficoltà degli alunni, permettendo una programmazione delle esperienze didattiche, la scelta dei mezzi più efficaci e la distribuzione del carico di lavoro in relazione alle capacità di apprendimento degli allievi.

Lo psicolinguista americano J. B. Carroll, in una serie di indagini e ricerche negli U.S.A., ha evidenziato gli elementi attitudinali specifici dell'apprendimento delle lingue straniere. Tra i più rilevanti, ricordiamo:

- la capacità di intuire certe leggi morfofonematiche presenti, sotto forme diverse, in tutte le lingue; sono tipiche, ad esempio, quelle relative alla formazione dei numeri (tre-tredici-trenta-trecento, quattro-quattordici-quaranta-quattrocento, ecc.);
- la percezione auditiva e visiva;
- la capacità di associare i suoni con i simboli grafici corrispondenti;
- la sensibilità nel distinguere le strutture logico-grammaticali (per esempio, la capacità di riconoscere, in qualunque tipo di frase, gli elementi che hanno funzione di soggetto, complemento, ecc.).

A questi fattori specifici si associano altri elementi più generalizzati, quali la memoria immediata e a lungo termine, i fattori verbali e psicomotori dell'intelligenza, la capacità di comprendere le istruzioni e di organizzarsi, ecc.

Tenendo conto delle varie componenti, il Carroll ha elaborato il *Modern Language Aptitude Test*, una prova attitudinale che ha rivelato un alto indice di correlazione coi risultati scolastici degli studenti che ad essa erano stati sottoposti. I buoni risultati prognostici del test testimoniano non solo della validità di tale strumento ma anche che il modello attitudinale elaborato dal Carroll, trovando conferma nelle sue applicazioni scolastiche, può essere assunto come precisa indicazione sulle modalità dell'apprendimento linguistico.²⁰

²⁰ J. B. CARROLL e S. M. SAPON, *Modern Language Aptitude Test*, New York, The Psychological Corporation, 1959 (versione italiana a cura di Mario Ferencich, Firenze, Organizzazioni Speciali, 1964).

CAP. 2 LE PROVE TRADIZIONALI

2.1. Caratteri delle prove tradizionali

Col termine *prove tradizionali* designiamo quelle prove di controllo del profitto che tutti gli insegnanti ben conoscono, anche perché su di esse si basano gli esami scolastici di lingue straniere: la traduzione, il riassunto, il questionario, la composizione, il dettato, l'interrogazione/colloquio.

Una caratteristica comune a tutte queste prove è la *complessità*; ognuna di esse coinvolge infatti numerose variabili e richiede il superamento di difficoltà di vario tipo. La possibilità di errori di natura diversa rende molto problematico e soggettivo il giudizio.

Che la soggettività dei mezzi tradizionali di controllo non sia un'ipotesi, bensì una realtà, è stato dimostrato da numerose esperienze, in Italia e all'estero.²¹ Sottoponendo un elaborato a gruppi di docenti, e chiedendo a ciascuno di dare un voto in decimi, si sono rilevate fortissime variazioni (in alcuni gruppi, addirittura dal 4 al 9, o dal 3 all'8). Ciò si è verificato non solo con temi, riassunti, traduzioni o prove analoghe, ma anche con problemi di matematica, che a prima vista potrebbero sembrare prove oggettive, ma in realtà rientrano nel campo delle prove operanti su molteplici variabili. Si è potuto inoltre rilevare che la stessa dispersione, cioè il ventaglio di voti, si aveva quando l'elaborato veniva valutato da docenti della materia e quando esso veniva invece valutato da insegnanti di materie diverse; non vi sarebbe, in altre parole, differenza tra gli esperti e i non competenti.

Per limitare la soggettività nella valutazione delle prove tradizionali sono state proposte alcune procedure. Anzitutto, l'insegnante deve analizzare le prove stesse per accertare quante e quali variabili entrino in gioco. Fatto ciò, si prospettano due soluzioni. La prima prevede che si stabilisca una gerarchia di valori tra i vari elementi della prova; ad esempio, per una composizione in lingua straniera, la graduatoria potrebbe contemplare, in ordine di importanza; la aderenza al tema proposto; la scelta appropriata del lessico; la correttezza morfosintattica; l'ortografia, ecc., con tutte le possibili varianti che tengano conto degli obiettivi che l'insegnante si prefigge. Naturalmente gli allievi debbono essere preventivamente informati della graduatoria di valori (e quindi del "peso specifico" degli eventuali errori) non solo per potersi concentrare su ciò che è più importante, ma anche perché il voto globale dato dall'insegnante risulti giustificato dalle premesse.

La seconda soluzione consiste invece nell'attribuire tanti voti quanti sono gli elementi, le variabili o le parti in cui si articola la prova (nel nostro esempio, daremmo un voto per l'aderenza al tema, un altro per la grammatica, un terzo per il lessico, un quarto per l'ortografia, ecc.). Questa proposta di soluzione ci fa chiaramente intendere come l'ideale che si vuole raggiungere sia una serie di dati oggettivi sui singoli fattori. Ma se questo è il fine, il solo mezzo adeguato è una batteria di test oggettivi predisposti opportunamente, mentre nessuna delle due procedure suggerite risolve alla radice il problema della soggettività del giudizio dell'insegnante.

Un difetto comune a molte prove tradizionali — per esempio certi "compitini di grammatica" — è la mancanza di una visione funzionale della lingua: infatti è inutile verificare l'acquisizione di questo o quel dettaglio, quando ciò che conta è la capacità di adoperare la lingua. L'errore risiede non tanto nel controllare l'acquisizione di sistemi parziali, quanto piuttosto nell'assumere elementi di dettaglio come indice di abilità assai più complesse. Troppo spesso, coi sistemi tradizionali, si è confusa la capacità di svolgere un certo tipo di prova con il possesso delle abilità linguistiche: tipicamente, il saper tradurre identificato con la conoscenza della lingua straniera.

²¹ Per l'Italia, ricordiamo le esperienze condotte a più riprese presso il Centro Europeo dell'Educazione (Frascati, Villa Falconieri), alle quali si riferisce l'esemplificazione da noi riportata. Per le esperienze straniere, si veda H. PIÉRON, *Esami e docimologia*, cit., pp. 14-22.

2.2. La traduzione

Robert Lado²² ha messo in luce i limiti della traduzione come test:

“ 1. Gli studenti meglio preparati non traducono mentre usano la lingua straniera.

2. Ci sono vari modi di tradurre e di giudicare una traduzione: dal punto di vista artistico, dell'accuratezza delle informazioni, della correttezza grammaticale, dell'appropriatezza lessicale. Una traduzione può essere valutata da questi (e altri) punti di vista. Se lo studente è costretto a tradurre a scopo di controllo lessicale o grammaticale, può risentirne negativamente la sua capacità di accostarsi alla letteratura.

3. La votazione delle traduzioni tende a non essere affidabile a causa dei vari modi di tradurre e delle varianti possibili, accettate o no da chi valuta.

4. La traduzione è un'abilità speciale, diversa dal capire, parlare, leggere e scrivere.

5. La traduzione è un test lento. A meno che abbia avuto un addestramento speciale, lo studente bravo impiega più tempo a tradurre una lettera che a comporla. Nel tempo occorrente per tradurre un brano può, usando altre tecniche, affrontare una quantità maggiore di materiale.

6. La traduzione è lenta da correggere, perché l'esaminatore deve considerare ogni soluzione per vedere se sia accettabile.

7. L'uso della traduzione nei test incoraggia l'abuso della traduzione in aula”.

Alle osservazioni del Lado possiamo aggiungere:

a) Non è facile reperire del materiale da tradurre che sia ben graduato per difficoltà sia sul piano lessicale che su quello morfosintattico; i branetti creati “ad hoc”, imbottiti di problemi grammaticali, oltre a non avere alcun valore letterario, sono spesso penosamente insulsi.

b) Si tende ad iniziare la traduzione mediante frasi staccate, svincolate da ogni contesto o situazione, e perciò talora ambigue o interpretabili in vari modi.

c) L'impiego della traduzione incoraggia l'uso indiscriminato dei dizionari bilingui, con la conseguente mistificazione sulla presunta “equivalenza” dei vocaboli nelle due lingue — errore inevitabile ove non si tenga conto dei contesti di civiltà.

L'uso della traduzione come test è valido là ove occorra verificare specificamente la capacità di tradurre (per esempio, nei corsi per traduttori); negli altri casi, esso non appare sufficientemente giustificato, in quanto gli svantaggi superano abbondantemente i vantaggi: il Lado afferma che “forse le sole cose favorevoli da dire sulla traduzione come test è che i temi di traduzione sono facili da assegnare e non occupano molto spazio”.²³ A favore della traduzione stanno i vecchi metodi di insegnamento, le tendenze conservatrici, i programmi ministeriali per le scuole medie superiori (v. gli esami di Maturità Scientifica), e i piani di studio universitari. Sappiamo come queste forze esercitino un'azione frenante nei confronti dei metodi più avanzati; di fatto, in molte situazioni, la traduzione è tuttora considerata non come un momento terminale, un'abilità che si acquisisce *in seguito alla* conoscenza della lingua, bensì come un indispensabile strumento di apprendimento della lingua straniera.

Un cenno a parte meritano le “retroversioni”, ossia quelle traduzioni dall'italiano che hanno come punto d'arrivo dialoghi o brani in lingua straniera già noti agli studenti: nella misura in cui sono esercizi di *rievocazione* di materiale già appreso, le retroversioni non soggiacciono alle critiche

²² R. LADO, *Language Teaching*, cit., pp. 157-158.

²³ *ibidem*.

mosse alle traduzioni vere e proprie. Vanno però usate con cautela, e solo se si ha la certezza che vi sia stato l'apprendimento di *tutto* il brano richiesto; pertanto le retroversioni come test sono o superflue (in quanto presuppongono una verifica già effettuata) o pericolose (in quanto rischiano di essere vere e proprie traduzioni laddove gli studenti non avessero sufficientemente memorizzato il materiale in lingua straniera).

Le versioni dalla lingua straniera sono a volte usate per verificare la comprensione di brani o poesie; lo stesso scopo, però, può essere raggiunto mediante prove oggettive, e senza ricorrere alla lingua madre (v. cap. 6). Se invece le si impiega come vere e proprie traduzioni, esse coinvolgono tutta la problematica linguistica e di civiltà che interessa le versioni in lingua straniera.

2.3. La composizione, il riassunto, il questionario aperto

Premesso che per questionario *aperto* intendiamo ogni tipo di questionario in cui le domande ammettano più risposte, diverse tra loro nel contenuto e/o nella forma, raggruppiamo queste tre prove, malgrado le differenze che intercorrono tra di esse, in quanto, come test, hanno caratteristiche comuni, e anche nella prassi didattica spesso si considerano i riassunti ed i questionari aperti come “composizioni guidate”, e li si impiegano come avviamento al comporre.

Queste prove richiedono all'esaminato di:

- formulare mentalmente i concetti più adeguati a svolgere il tema, riassumere il brano, o rispondere alle domande, e organizzarli opportunamente in sequenza logica;
- redigere frasi corrette in lingua straniera, idonee ad esprimere i concetti.

A sua volta, la fase di redazione implica:

- la scelta di elementi lessicali appropriati;
- il rispetto delle strutture morfosintattiche;
- il rispetto dell'ortografia e della punteggiatura;
- a livello avanzato, il rispetto di norme stilistiche e retoriche: chiarezza, efficacia, persuasività, ecc.

Tutte queste abilità possono essere efficientemente misurate per mezzo di test oggettivi, eliminando il ricorso a composizioni, riassunti o questionari che, in quanto prove complesse, non sono in grado di fornire misurazioni indistorte del profitto. La principale obiezione contro i test è la loro presunta inidoneità a verificare la capacità di scrivere brani di prosa. E mentre per la traduzione si possono opporre critiche di fondo, la capacità di comporre, riassumere, redigere lettere o resoconti, rispondere per iscritto a questionari è di solito considerata parte integrante del “saper scrivere”. In effetti, nel criticare queste prove *come test* non si intende affermare che esse non debbano rientrare negli obiettivi di un corso di lingue straniere. Si vuole invece sfatare l'opinione corrente secondo cui solo prove di composizione possono verificare la capacità di comporre. Poiché esiste un'alta correlazione tra i dati ottenuti mediante opportuni test e quelli ricavati attraverso prove di composizione, è possibile fare a meno di queste ultime *a scopi puramente valutativi*. Anche in questo caso, gli esami (di Stato e non) esercitano un certo tipo di influenza, concorrendo a far identificare esercizi e test e a confondere gli obiettivi con i mezzi di controllo.

2.4. Il dettato

Il dettato non è una prova oggettiva anche se, da un certo punto di vista, potrebbe sembrarlo: infatti, nella correzione, gli errori sono sicuramente rilevabili, in quanto si tratta semplicemente di osservare se il testo dettato sia stato riprodotto fedelmente oppure no. Ma due importanti fattori ne invalidano l'oggettività. Anzitutto, anche il dettato è una prova complessa, in cui entrano in gioco la

comprensione auditiva, l'ortografia e l'interpunzione. A loro volta i problemi ortografici si suddividono a seconda che interessino il lessico (ortografia *lessicale*) ovvero la morfologia (accenti, apostrofi, morfemi inflessionali — suffissi e desinenze —: ortografia *grammaticale*). L'esecuzione di un dettato può quindi comportare errori di natura diversa, la cui gravità è stabilita soggettivamente da chi li corregge.

Un secondo fattore di distorsione è inerente al modo in cui le prove di dettato vengono effettuate. Nella prassi scolastica sono in uso tecniche diverse; c'è chi detta subito e rilegge, e chi invece legge una volta anche prima di dettare; alcuni dettano parola per parola, altri per segmenti significativi o intere frasi; il ritmo di dettatura può essere lento o veloce, ecc. Queste variabili fanno sì che ben difficilmente un brano venga dettato allo stesso modo da due persone diverse, o addirittura che la stessa persona detti due volte (a gruppi diversi) senza variazioni.

Inoltre, la ricezione di chi è seduto vicino a colui che detta è diversa da quella di chi è seduto in fondo all'aula, e gli psicologi hanno rilevato differenze di prestazioni anche in relazione alla diversa illuminazione e rumorosità di fondo dell'ambiente. Pertanto, se si vogliono comparare due dettature dello stesso brano, occorre che il dettato sia registrato su disco o nastro e diffuso mediante un sistema di cuffie, riducendo al minimo i disturbi ambientali. L'uniformità di somministrazione è perciò ottenibile solo in particolari situazioni (ad esempio il laboratorio linguistico), ma queste condizioni non sono facilmente realizzabili, né prive di svantaggi. L'audizione da un mezzo di registrazione impedisce la “lettura delle labbra” e i registratori o i giradischi di solito non riproducono l'intera gamma delle frequenze sonore, per cui va perduta una parte del messaggio acustico.

Sul dettato come esercizio i pareri sono discordi: alcuni autori lo ritengono superfluo: i più lo ritengono invece utile, soprattutto nella fase di passaggio dal codice orale allo scritto.²⁴ Riteniamo che le divergenze di opinione possano essere in relazione sia alle caratteristiche della lingua straniera considerata, sia al fatto che alcuni studiosi si sono interessati al problema della “lingua straniera”, ossia di una lingua estera insegnata fuori dai paesi in cui tale lingua si parla (p. es. il francese in Italia) mentre altri si sono occupati prevalentemente della “seconda lingua”, ossia dell'insegnamento di una lingua, a studenti stranieri, nel paese in cui tale lingua è parlata (p. es. l'inglese in Gran Bretagna o Stati Uniti).

Sulla scarsa produttività del dettato come test esiste, al contrario, una notevole concordanza di vedute. Premesso che le abilità interessate dal dettato possono essere valutate più efficacemente con altre tecniche, analizzando a fondo il dettato si può notare come esso in realtà non misuri certi fattori che in apparenza vi sono coinvolti. Ad esempio, esso non è di per sé un test di discriminazione auditiva. Se si detta la parola inglese *ship* e questa viene scritta correttamente, noi possiamo dedurre che lo studente discrimina tra il suono /i/ di *ship* e il suono /i:/ di *sheep* solo a condizione che:

- a) entrambi i vocaboli siano stati appresi dallo studente;
- b) il contesto sia tale da non privilegiare un vocabolo rispetto all'altro.

Se la frase dettata è *I can see a ship entering Dover Harbour* il contesto è tale per cui le probabilità di *sheep* sono pressoché nulle: lo studente che commettesse questo errore rivelerebbe non tanto problemi di discriminazione auditiva, quanto piuttosto problemi di apprendimento del lessico.

²⁴ Sull'uso del dettato nel passaggio dall'orale allo scritto v. G. FREDDI, *Metodologia e didattica delle lingue straniere*, cit., pag. 297. Sul dettato come test, vedi anche: R. LADO, *Language Testing*, Longmans, Londra, 1961, p. 34; R. TITONE, *Le lingue estere*, cit., p. 446; N. BROOKS, *L'apprendimento delle lingue straniere (teoria e pratica)*, Firenze, La Nuova Italia, 1968, pp. 280-281.

Rispetto alle altre prove tradizionali, comunque, il dettato ha il vantaggio della rapidità di somministrazione e della facilità di correzione, qualità che ne fanno una prova abbastanza utile, sempreché coloro che se ne servono siano consapevoli dei suoi limiti e preparino del materiale idoneo.

2.5. Le prove orali tradizionali

L'interrogazione tradizionale assumeva vari aspetti a seconda delle circostanze, ma in genere comprendeva una o più delle seguenti attività:

- traduzione di frasi nella lingua straniera, oralmente o alla lavagna;
- lettura di brani;
- retroversione di brani, con eventuale commento grammaticale e lessicale;
- dialogo dello studente con l'insegnante (conversazione libera o risposte a domande predeterminate).

Ovviamente, la traduzione di frasi alla lavagna non è affatto un'attività orale, anche se era comunemente adoperata dagli insegnanti che impiegavano il metodo grammatica/traduzione; del resto, a tutti è noto come, per mettere voti nella colonna ORALE del registro si usasse ricorrere anche a dettati, compitini di grammatica o di vocaboli, e simili. Tutti questi espedienti servivano a dare classificazioni collettive, risparmiando il tempo richiesto dalle verifiche individuali, ma ciò non toglie che tanti, troppi “voti di orale” fossero dai falsi ideologici, in quanto poco o nulla avevano a che fare con il saper comprendere e parlare.

L'abilità del saper comprendere, in particolare, era la più trascurata, e a ciò si devono in larga misura gli insuccessi, all'estero, di coloro che apprendevano la lingua coi metodi tradizionali. Infatti anche nei colloqui si richiedeva spesso che lo studente rispondesse a domande prestabilite ed a lui già note, esigendo solo un minimo sforzo di comprensione.

Il colloquio è il test meno facilmente valutabile, ossia la prova in cui si riscontrano i più bassi indici di affidabilità. Le cause vanno ricercate nella complessità della prova (un colloquio abbraccia problemi di pronuncia, intonazione, ritmo, accento, fluidità, appropriatezza lessicale, correttezza grammaticale, e altri) e *nell'effetto di alone* che, nel colloquio, è avvertibile in misura maggiore che nelle altre prove soggettive. Esso viene definito come la “distorsione di giudizio dovuta alla tendenza di valutare una caratteristica particolare di un individuo, sulla base dell'impressione globale che il valutatore si è fatta nei riguardi del valutato o, viceversa, la tendenza ad allargare il giudizio di una caratteristica su tutta la personalità del valutato”.²⁵

In altre parole, ciò significa che chi valuta è esposto a due tendenze che possono falsare il suo giudizio. La prima è quella di estendere un giudizio generale a tutti i particolari: dopo aver giudicato un allievo positivamente o negativamente nel suo complesso, incontriamo una certa difficoltà ad accettare una prestazione di segno opposto, ossia il bel voto dell'allievo scadente o il brutto compito dello studente bravo. Eppure, anche prescindendo da cause accidentali (indisposizioni, stati emotivi o ansiosi), ed esclusa la possibilità di inquinamento della prova (copiature, suggerimenti), un risultato divergente dagli altri è perfettamente plausibile qualora la prova si riferisca ad un'abilità particolare, di tipo specialistico, che presupponga doti, attitudini e preparazione difformi, in qualche misura, da quelle richieste per le altre prove. È indubbio che esista una certa correlazione fra i diversi aspetti di una disciplina, ma ciò non deve far escludere la possibilità che uno studente riesca in certe prove assai meglio che in altre.

²⁵ E. SPALTRO, *La psicologia del lavoro*, Etas-Kompass, Milano, 1967, p. 207.

La seconda tendenza è quella di generalizzare un aspetto particolare. È il caso di chi, accertate in un allievo buone capacità in una prova di lettura, viene indotto a supporre buone capacità in tutte le abilità linguistiche, comprese quelle che non hanno alcun rapporto diretto col saper leggere. Accade di frequente, nelle riunioni dei consigli di classe, che un insegnante, avendo giudicato in un certo modo un allievo, si meravigli del fatto che altri colleghi abbiano espresso, per le loro materie, giudizi molto diversi dal suo.

All'effetto di alone si debbono anche le interferenze sulla misurazione del profitto di elementi che col profitto stesso non hanno nulla a che vedere: la condotta dello studente, il suo aspetto fisico più o meno gradevole, le sue doti di simpatia e comunicatività, la sua capacità di assumere un comportamento gradito all'insegnante, ecc. Dell'esistenza di questi condizionamenti siamo tutti abbastanza consapevoli, se non altro perché abbiamo dovuto tenerne conto nella nostra vita di studenti; non sempre, però, ci rendiamo conto di quanto penetranti, sottili e subdole possano essere certe interferenze, che non si manifestano in modo sensibile ma incidono più sui nostri riflessi inconsci che sulle nostre valutazioni razionali, portandoci a giudizi errati mentre noi, in buona fede, riteniamo di essere obiettivi, sereni ed equilibrati.

2.6. Il voto

Poiché nel corso di queste pagine facciamo più volte riferimento ai voti scolastici, riteniamo opportuno inserire a questo punto alcune osservazioni in proposito.

La legislazione vigente prevede che ogni insegnante formuli nel corso di un certo periodo (trimestre o quadrimestre) un giudizio sul profitto di ciascun allievo. Al momento dello scrutinio i giudizi vengono tramutati in voti da 1 a 10. La prassi è tuttavia ben diversa, in quanto i voti in decimi vengono usati in riferimento a ogni singola prova. Ciò porta alla istituzionalizzazione della confusione tra misurazione del profitto e valutazione dell'allievo.

Come strumento di misurazione, i voti in decimi sono del tutto inadeguati; lo dimostra l'uso dei "mezzi voti", dei "più", dei "meno", dei "meno meno", e così via, senza che peraltro vi sia accordo sul valore preciso di questi artifici. Il nostro stesso ordinamento scolastico dimostra di essere consapevole dell'insufficienza dei voti in decimi, prevedendo voti in trentesimi per gli esami universitari, in sessantesimi per gli esami di maturità, in centesimi o centodecimi per gli esami di laurea; in quanto ai punteggi delle abilitazioni e dei concorsi, è prevista una scala di 75 o 100 punti ampliata mediante l'uso di due cifre decimali, consentendo così 7.500 o 10.000 punteggi diversi.

Se per la misurazione del profitto la scala decimale è troppo ristretta, per la valutazione complessiva essa è sovrabbondante. La opinione più diffusa, al riguardo, esclude che il risultato di un processo valutativo possa essere adeguatamente espresso da un numero o dalla media di una serie di numeri. In effetti gli aspetti qualitativi sono di gran lunga più importanti di quelli quantitativi, per cui un "profilo" o una scheda di valutazione in cui compaiano, oltre al profitto, tutte le notizie relative alla personalità dello studente, sono strumenti assai più idonei della pagella tradizionale.

Qualora a fini pratici, ovvero in ossequio alle leggi e alle tradizioni, si debba esprimere una valutazione globale mediante una scala (cercando di quantificare degli aspetti qualitativi), le possibilità si riducono a cinque: lo studente si colloca o nella media dei valori considerati, o al disopra della media, o al disotto di essa; a loro volta, gli scostamenti dalla media possono essere più o meno ampi. La seguente tabella riporta, per ciascuna delle cinque possibilità, gli aggettivi e i voti decimali normalmente adoperati per esprimere i relativi giudizi.

<i>Valutazione</i>	<i>Aggettivi</i>	VOTO
Molto al disotto della media	pessimo, negativo	4
Al disotto della media	insufficiente mediocre	5
Nell'ambito della media	sufficiente medio	6
Al disopra della media	discreto, distinto, buono	7

Molto al disopra della media	ottimo, eccellente	8
------------------------------	--------------------	---

I voti compresi tra il 4 e l'8 sono infatti quelli che normalmente compaiono nelle pagelle scolastiche, mentre i voti inferiori o superiori (1, 2, 3, 9, 10), sono in genere usati più per singole prove che per giudizi complessivi, a meno che non concorrano circostanze eccezionali. In quanto alla frequenza dei voti, essi sono tanto più numerosi quanto più si avvicinano al valore centrale: il 6 è di gran lunga il voto più frequente, seguito dal 5 e dal 7 e, a distanza, dal 4 e dall'8. Questo schema, tuttavia, soggiace a numerose alterazioni dovute in parte ad usi ormai invalsi nella scuola (ove spesso

si tende a considerare il voto non tanto come un giudizio quanto piuttosto come un premio o una punizione) e in parte alla disposizione di legge che fa coincidere la sufficienza col voto 6. Poiché

i concetti di sufficienza, idoneità e promovibilità non coincidono necessariamente con quello di appartenenza alla media, a fine anno si mettono in moto vari meccanismi, tra cui i “voti di consiglio”, che hanno l'effetto pratico di tramutare in 6 certi voti negativi e si giustificano per la necessità di conciliare istanze diverse. Vi è quindi difformità tra le votazioni dei trimestri intermedi (non vincolate alla promozione) e quella finale, in cui sono di solito presenti più voti sufficienti, indipendentemente dagli effettivi progressi degli studenti.

Per superare tutti questi equivoci riteniamo opportuno che la misurazione del profitto venga effettuata sulla base di scale assai più ampie di quella in decimi: tra tutte, quella in centesimi è senz'altro la più facilmente adoperabile, e scale diverse si possono ricondurre ad essa in modo abbastanza semplice. Al tempo stesso, occorre abolire i voti come strumento di valutazione, sostituendoli con profili valutativi sufficientemente articolati. Ciò richiederà un cambiamento di abitudini non solo da parte degli insegnanti ma anche degli allievi e delle loro famiglie, e occorrerà una paziente e penetrante opera di chiarimento per evitare di giungere all'incomunicabilità tra scuola e famiglia. D'altronde, solo superando gli schemi attuali si potrà impostare su base scientifica il problema della selezione nella scuola e della “meritocrazia”.

Nel frattempo ogni insegnante dovrà contemperare le diverse esigenze nel modo più proficuo; se usa tecniche rinnovate di misurazione e valutazione è necessario che ne renda edotti allievi e famiglie, perché il linguaggio dei voti (come ogni linguaggio) ha essenzialmente il fine di comunicare; qualora si parlassero linguaggi diversi, ciò sarebbe ancor peggio del sistema vigente che, pur con i suoi grossi difetti, ha il pregio di essere noto a tutti.

CAP. 3 I VARI TIPI DI TEST

3.1. Ipotesi di classificazione

I test possono essere classificati secondo diversi parametri, in relazione alle caratteristiche che vengono considerate di volta in volta preminenti. Ecco anzitutto un quadro di riferimento riguardante i vari tipi di test esaminati in questo capitolo.

A) Secondo l'ampiezza del campo d'indagine:

- a) *Test fattoriali*
- b) *Test di abilità*
- c) *Test globali di lingua*
- d) *Test di lingua-civiltà.*

B) Secondo i livelli linguistici:

- a) *Test fonologici*
- b) *Test morfosintattici*
- c) *Test lessicali.*

C) Secondo il momento di applicazione:

- a) *Test intermedi*
- b) *Test di livello*
- c) *Test di padronanza.*

D) Secondo il contenuto:

- a) *Test specifici*
- b) *Test aspecifici.*

E) Secondo il grado di sistematicità ed elaborazione:

- a) *Test episodici*
- b) *Test sperimentali*
- c) *Test standardizzati.*

L'ipotesi di classificazione qui proposta si differenzia in certa misura dagli schemi suggeriti da altri autori.²⁶ Riteniamo che le nostre scelte siano giustificate da tre motivi fondamentali. Anzitutto vi è discordanza tra i vari autori sui termini adottati, soprattutto a proposito di vocaboli piuttosto generici, quali *profitto*, *idoneità*, *progresso*, *capacità* e simili. Ciò è in parte dovuto alla difficoltà di tradurre adeguatamente certi termini stranieri, come gli inglesi *achievement*, *attainment*, *proficiency*. Su un punto importante vi è comunque un certo accordo, e cioè nel considerare i test di profitto in contrapposizione ai test psicologici, psicometrici, attitudinali, ecc. Poiché, tuttavia, tutti i test di cui ci occuperemo rientrano in questa categoria, l'impiego del termine "test di profitto" risulta troppo generico per essere di alcuna utilità pratica ai nostri fini.

Una seconda giustificazione si ricollega al fatto che noi ci interessiamo esclusivamente di *language testing*, ossia di un settore specifico e ristretto con particolari presupposti metodologici e precise caratteristiche di impiego, mentre altre classificazioni sono rivolte a coprire l'intera gamma delle possibilità di uso dei test nella scuola. Anche la terminologia assume valori ben precisi: un termine di per sé generico, quale "abilità", possiede nel campo della didattica delle lingue un significato ben

²⁶ Si vedano R. TITONE, *Le lingue estere*, cit., p. 444 segg. e N. BROOKS, *L'apprendimento delle lingue straniere (teoria e pratica)*, cit., p. 211 segg.

determinato (vedi paragrafo 1.4, punto *f*), il quale rifluisce nella definizione dei “test di abilità” così come essa è formulata nel paragrafo seguente.

Infine, ci è parsa primaria l'esigenza di mettere in rilievo i diversi ordini di fattori che entrano in gioco nel testing delle lingue straniere; lo schema che ne è risultato evidenzia i cinque parametri fondamentali con le relative sotto-classificazioni.

3.2. Il campo d'indagine

Un test²⁷ può vertere su un campo più o meno ristretto di elementi di lingua-civiltà.

I test *fattoriali* misurano un solo fattore od elemento, ossia prendono in esame un singolo problema di apprendimento della lingua-civiltà. Per esempio, un test sugli aspetti morfologici di un tempo verbale è un test fattoriale. È intuibile come un test complesso sia costituito da più test fattoriali coordinati.

I test *di abilità* misurano il grado di apprendimento nell'ambito di una abilità linguistica (capire, parlare, leggere, scrivere), rilevando e interrelando i diversi fattori che operano nell'ambito di tale abilità. Un test di comprensione della lettura è un esempio di test di abilità relativo al saper leggere.

I test *globali di lingua* misurano il grado di apprendimento linguistico complessivo da parte di un allievo ed operano su tutte le abilità linguistiche nel loro complesso.

I test *di lingua-civiltà*, oltre al grado di apprendimento della lingua, mirano a rilevare il livello di comprensione interculturale sulla base di un approfondimento degli elementi della civiltà straniera.

I test globali di lingua e i test di lingua-civiltà sono difficili e complessi da realizzare; vedremo nei successivi capitoli i metodi, le tecniche e le procedure di costruzione dei test. Qui vogliamo rilevare come i test contenuti in certi corsi e volumi e presentati come prove globali spesso siano invece test parziali, soprattutto in quanto non forniscono materiale per la verifica delle abilità orali (capire e parlare).

3.3. I livelli linguistici

Poiché ogni lingua viene analizzata su diversi piani o livelli, avremo test caratterizzati dal vertere primariamente su problemi a livello fonologico, morfosintattico o lessicale.

I test *fonologici* misurano la padronanza del sistema fonologico della lingua straniera. Essi riguardano quindi i fonemi segmentali (vocali, dittonghi, consonanti) ed i valori soprasegmentali (ritmo, accento, intonazione, giuntura). Tipici delle abilità orali, essi verificano sia l'aspetto passivo o ricettivo (percezione e discriminazione di suoni, comprensione della lingua parlata) sia quello attivo o produttivo (il parlare nei suoi vari aspetti).

I test *morfosintattici* rilevano la padronanza delle strutture morfosintattiche e misurano l'acquisizione del sistema grammaticale nelle sue varie estrinsecazioni. Rifacendoci alla lezione della linguistica strutturale²⁸ riteniamo infondata la distinzione classica tra morfologia e sintassi; il concetto di lingua come sistema conduce al superamento di confini artificiosi, mal definiti e non rispondenti ai criteri di analisi linguistica.

²⁷ Si noti che qui e in seguito, per comodità e brevità — conformandoci all'uso corrente — usiamo il termine “test” anche per indicare le *batterie di test* (serie di prove costituenti un unico blocco). Ove occorra designare le singole parti di un test complesso, il termine più frequentemente usato, specialmente nell'area anglosassone, è “subtest”.

²⁸ Per un esame di alcuni rapporti tra la linguistica strutturale e i vari aspetti della didattica delle lingue straniere si rimanda ai contributi di G. FREDDI e N. PERINI in *Strutturalismo e didattica delle lingue*, cit., 1967.

Anche per i test *lessicali* è essenziale la distinzione tra lessico passivo (riferito alle abilità del capire e del leggere) e lessico attivo (impiegato nel parlare e nello scrivere). Nella lingua materna e nelle lingue straniere il lessico passivo è di norma assai più ampio di quello attivo.

Malgrado alcuni autori, specialmente nell'area linguistica francese, abbiano talora considerato i test *ortografici* come categoria a sé stante, noi, accogliendo l'impostazione di R. Lado,²⁹ non riteniamo giustificata questa distinzione in quanto i problemi di ortografia lessicale rientrano nel livello lessicale e i problemi di ortografia grammaticale nel livello morfosintattico. Non esistono test puramente ortografici, a meno di giungere a un tal grado di artificiosità da contrastare con le premesse sulla funzionalità del linguaggio. Il negare che i test ortografici costituiscano una classe autonoma non significa tuttavia escluderli dall'impiego sistematico, che è anzi necessario nella misura in cui il codice scritto presenta difficoltà di apprendimento.

Abbiamo invece escluso dal nostro discorso, ritenendoli di scarso interesse per la maggioranza degli insegnanti italiani di lingue straniere, i problemi relativi all'uso di alfabeti diversi da quello latino (greco, cirillico, arabo, ecc.) ovvero di sistemi ideografici.³⁰

3.4. Impiego dei test

Nel corso di un programma di lingua-civiltà straniera i test vengono impiegati in momenti e per scopi diversi.

Se il nostro insegnamento si basa sull'unità didattica come modello operativo,³¹ è indispensabile che non si proceda all'unità didattica successiva prima di aver accertato che gli allievi abbiano acquisito in modo sicuro e funzionale tutti i contenuti dell'unità precedente. Per questo si afferma che il testing si colloca, per sua natura, nella fase conclusiva di ciascuna unità didattica.

Questa collocazione naturale ha la sua giustificazione non solo nel concetto di unità didattica, ma anche in relazione all'esigenza di condurre controlli sistematici. Non presentano invece carattere di sistematicità, anche se costituiscono una fonte d'informazione preziosa per l'insegnante, tutti quei dati che è possibile raccogliere durante ogni lezione, specialmente se questa è condotta in modo che gli allievi partecipino attivamente. Le rilevazioni sul profitto condotte durante la fissazione dei dialoghi o delle strutture — o, in genere, in qualunque momento diverso da quello specifico del testing — se tradotte in giudizi o espresse sotto forma di voti, rischiano di riportare alla confusione tra esercizio e test, già denunciata come ascientifica. Inoltre può essere demotivante per gli allievi, nel momento in cui stanno compiendo degli sforzi per apprendere, vedere il loro insegnante trasformarsi da colui che li aiuta ad imparare in colui che li giudica e sanziona. Per questi motivi riteniamo opportuno che si elabori una serie completa di test *intermedi*, uno per ciascuna unità didattica. I test *di livello* misurano il grado di apprendimento complessivo ad un certo punto dello sviluppo del programma; servono a stabilire eventuali fasce di livello, oppure a collocare nuovi allievi nel gruppo o nella classe cui sono idonei.

Per test *di padronanza* intendiamo quelli che misurano il grado di autonomia linguistica raggiunto al termine di un programma completo: è il caso dei test finali di un corso di lingua-civiltà.

²⁹ In *Language Testing*, cit., il Lado non considera l'ortografia in modo autonomo, bensì come uno degli elementi interessati dalla verifica del "saper scrivere".

³⁰ Per il confronto di due sistemi grafici si veda R. LADO, *Linguistics across Cultures*, cit., cap. 5°.

³¹ Al concetto di unità didattica abbiamo già accennato (par. 1.4); per un esame più approfondito si vedano le due "conversazioni" ad esso dedicate in: G. FREDDI, *Fondamenti e metodi della didattica delle lingue*, Ed. Cafoscarina, Venezia, 1972, pp. 111-133.

3.5. Il contenuto dei test

Per quanto riguarda i contenuti, sia linguistici che di civiltà, esistono due possibilità.

Un test può fare preciso ed esplicito riferimento ad un particolare corso, testo o materiale di apprendimento di una lingua straniera, ed esaminare l'allievo sulla base e nei limiti di esso. L'impiego di questi test *specifici* è rigorosamente riservato agli allievi che usano quel materiale, mentre essi sono inutilizzabili in situazioni diverse.

Esistono tuttavia anche test *aspecifici*, che prescindono totalmente da qualsiasi corso, testo o materiale e sono pertanto di uso generalizzato. Essi si basano di norma su indagini e ricerche attraverso cui si stabilisce il "corpus" di materia che si presuppone noto a studenti di un certo tipo di scuola o a un dato livello; tali indagini sono condotte o mediante questionari rivolti a gruppi sufficientemente numerosi di esperti oppure indirettamente, per esempio comparando un numero adeguato di testi per ricavarne gli elementi comuni. Sono aspecifici, di solito, i test di profitto e le schede di valutazione poste in commercio, salvo quelli che accompagnano esplicitamente un dato corso.

Il vantaggio principale dei test aspecifici è quello di permettere il confronto tra gruppi diversi di studenti, la cui preparazione è stata effettuata con materiali e metodi disomogenei; indirettamente, quindi, si può giudicare della validità del materiale didattico, ma perché il giudizio sia probante occorre che il test sia neutrale, ossia non privilegi indebitamente alcuni gruppi o testi a danno di altri. Inoltre, non vi deve essere stato un addestramento particolare al contenuto del test, e tutti i gruppi debbono essere egualmente preparati alla tecnica di esecuzione del test.

Le difficoltà di un test si riferiscono sia al contenuto dei quesiti sia al tipo di tecnica usata per la risposta. Questo secondo ordine di difficoltà deve essere rimosso preventivamente, addestrando gli allievi a tecniche di risposta a cui non sono abituati, per evitare che al momento della somministrazione della prova si sovrappongano fattori emotivi quali l'insicurezza ed il disagio, tali da influire negativamente sul punteggio. Alcuni test in commercio³² sono accompagnati da fogli per le esercitazioni preliminari, attraverso i quali gli allievi apprendono la tecnica di esatta esecuzione del test vero e proprio. Non si deve confondere l'addestramento alla tecnica di risposta con l'addestramento al contenuto del test, ossia con l'anticipazione delle difficoltà inserite nei quesiti e delle relative soluzioni; questo addestramento snaturerebbe il test e toglierebbe ogni validità ai punteggi conseguiti. Un test aspecifico, in quanto tale, comprende normalmente una grande quantità di problemi ed è normale che ciascun esaminato risponda solo ad una parte dei quesiti, tralasciando quelli dei quali non è in grado di individuare la soluzione (per ignoranza del lessico, delle strutture, degli elementi di civiltà, ecc.). Un test di questo genere può quindi evidenziare le lacune individuali o dei gruppi e segnalare la necessità di integrazioni e sviluppi nel programma didattico; è chiaro che le lacune si colmano mediante apposite esercitazioni, non contrabbandando le risposte a qualche quesito.

Vorremmo infine sottolineare che non esiste un conflitto tra test specifici ed aspecifici, nel senso che gli uni siano da preferire agli altri: entrambi trovano la loro naturale collocazione nell'attività scolastica. Mentre i test specifici verificano che gli allievi apprendano il materiale didattico con cui la classe lavora, e sono quindi di uso regolare e frequente, i test aspecifici sono usati come test di livello o di padronanza ogni volta che occorre uscire dai limiti del testo adottato e controllare l'apprendimento sul piano generale.

³² V., tra gli altri, il *Test obiettivo di profitto della lingua Inglese*, di FERENCICH, COMBATTI, LUZZATTI e PIRONA, ed. O.S., Firenze.

3.6. La sistematicità dei test

Uno dei caratteri tipici delle prove tradizionali è la mancanza di sistematicità, nel senso che si tratta di prove somministrate “una tantum”, senza che poi i risultati vengano elaborati statisticamente né confrontati con quelli ottenuti somministrando la stessa prova ad altri gruppi di allievi. Le prove di questo tipo rientrano nella classe dei test *episodici*; essi si contrappongono ai test *sperimentali* che hanno lo scopo di permettere le rilevazioni statistiche, attraverso cui si giunge all'elaborazione dei test *standardizzati*, ossia delle prove oggettive la cui validità è stata verificata. Poiché ai processi di rilevazione, analisi ed elaborazione statistica sono dedicati i capitoli 9 e 10, ci limitiamo qui alla pura e semplice classificazione, rimandando ad essi per la discussione dettagliata dei principi e delle tecniche statistiche.

Concludiamo accennando brevemente ad alcune definizioni di tipi particolari di test che noi non abbiamo incluso nella nostra classificazione generale perché riguardanti aspetti secondari o specifiche tecniche di somministrazione. La distinzione tra test *individuali e collettivi* (o *di gruppo*) è di per sé chiara e non ha bisogno di ulteriori precisazioni. In quanto alle tecniche di somministrazione, si parla di test *verbali* e di test *carta-matita*, a seconda che le risposte vengano fornite oralmente dagli esaminati ovvero mediante l'uso di appositi fogli per le risposte, su cui gli esaminati debbono scrivere, disegnare, apporre segni convenzionali o contrassegnare la risposta esatta scegliendola tra le alternative. In relazione a certe prove sono in uso termini particolari, che preferiamo illustrare nei capitoli successivi man mano che prenderemo in esame le singole tecniche di testing.

PARTE II LA VERIFICA DELLE ABILITÀ INTEGRATE

CAP. 4 II SAPER COMPRENDERE

4.1. La comunicazione verbale

La prima delle quattro abilità fondamentali, il saper comprendere, si riferisce all'aspetto *passivo* della comunicazione *orale*. La ricezione di messaggi parlati può aver luogo in circostanze e sotto forme diverse, che occorre esaminare preliminarmente.

Nel *colloquio*, cioè nella normale conversazione, il messaggio verbale è accompagnato da svariate informazioni complementari che giungono al ricevente attraverso gli altri sensi: la vista, l'olfatto, il tatto, più raramente il gusto.³³ La presenza di tale corredo di informazioni permette che il messaggio linguistico sia ridotto all'essenziale; di rado si incontrano periodi complessi; solitamente le battute sono di poche parole, strutturate in frasi brevi con molte forme ellittiche e qualche anacoluto. Il messaggio linguistico è completo in tutti i suoi aspetti sia segmentali (vocali, consonanti, dittonghi) che soprasedimentali e acustici (accenti, ritmo, intonazione, giunture, volume, timbro, ecc.).

Dalle *comunicazioni telefoniche* sono assenti tutte le informazioni non-verbali e vi sono alcune differenze — di portata peraltro limitata — rispetto ai colloqui nella qualità acustica della ricezione, interessanti soprattutto il timbro di voce. Nel messaggio linguistico troviamo meno ellissi ed anacoluti, più periodi composti o complessi, ma la forma è ancora essenzialmente quella del dialogo.

Con le *audizioni radio* entriamo nel campo della trasmissione unidirezionale del messaggio — manca cioè (a parte il caso dei radioamatori) lo scambio di ruolo tra parlante ed ascoltatore che si verifica nel colloquio. Il messaggio linguistico è assai curato dal punto di vista formale ed espositivo, ed è spesso ridondante. Non potendo verificare se l'ascoltatore abbia capito o no, si ricorre a ripetizioni, chiarimenti, apposizioni, forme antonomastiche e simili, in misura tale da far ritenere pressoché sicura la comprensione del messaggio.

Il linguaggio della *televisione* (in parte analogo, in parte diverso da quello del *cinema* e del *teatro*) presenta caratteristiche peculiari intrinseche e connesse alla totalità del messaggio audio-visivo.³⁴

L'audizione di *canti* si differenzia dagli altri tipi di audizione di messaggi verbali perché nel canto vengono alterati i valori soprasedimentali; le alterazioni del ritmo, dell'accento e dell'intonazione si riflettono anche sulla qualità dei singoli fonemi segmentali, e ciò spesso rende più difficile la comprensione di un canto in lingua straniera.

Qualunque sia la forma di comunicazione, l'audizione è disturbata da interferenze di vario tipo: rumori ambientali, disturbi e limiti del mezzo tecnico, ecc. L'intensità delle interferenze varia a seconda delle circostanze; quando si superano i livelli normali, si hanno distorsioni più o meno intense nella ricezione del messaggio, fino a giungere all'impossibilità di comunicare.

Per quanto riguarda il LT, non possiamo prendere in considerazione le variabili extralinguistiche; a noi interessa il messaggio verbale nelle sue manifestazioni ordinarie, cioè la lingua dei dialoghi e delle conversazioni percepita direttamente o riprodotta senza particolari interferenze.

³³ In effetti, con l'udito si percepisce solo l'11% dei messaggi informativi, contro l'83% percepito mediante la vista. Cfr. A. D'ELIA, *I sussidi audiovisivi nella Scuola Media*, Roma, Ed. di "Audiovisivi", 1967, pp. 19-20.

³⁴ Per gli aspetti del linguaggio televisivo direttamente connessi alla didattica delle lingue si veda l'introduzione al volume: G. FREDDI, *Fondamenti e metodi della didattica delle lingue*, cit.

4.2. Fattori della comprensione auditiva

Nel fenomeno della comprensione di messaggi verbali interagiscono vari fattori, a diversi livelli:

- riconoscimento e discriminazione di elementi soprasegmentali: accenti verbali e frasali,³⁵ schemi intonativi, ecc.;
- riconoscimento e discriminazione di fonemi segmentali;
- comprensione delle strutture morfosintattiche;
- comprensione del lessico.

Ciascuno di questi fattori è verificabile per mezzo di prove oggettive, anche senza far ricorso alla lettura o comunque all'impiego di frasi scritte. Ciò è importante per consentire il testing degli allievi nella fase di accostamento orale alla L2, quando ancora non è stato effettuato il passaggio al codice scritto. Sono indubbi i vantaggi di un esame precoce del profitto, sin dai primi momenti. Il testing nella fase di *oral approach* si rivela di particolare utilità per superare difficoltà psicologiche insorgenti in certe categorie di principianti.

Alcune nostre esperienze in corsi serali di lingue per adulti³⁶ hanno messo in luce come varie persone, specialmente se non più giovani o già abituate a imparare le lingue coi sistemi tradizionali, si trovino a disagio nella fase orale e abbiano l'impressione di non stare imparando nulla. L'uso di test appropriati ha permesso di oggettivare l'apprendimento orale e di dimostrare a quegli studenti che, contrariamente alle loro sensazioni, si erano instaurati degli apprendimenti linguistici in misura tale da rendere quantificabile il profitto. Il più delle volte questo è bastato perché vi fosse un atteggiamento più favorevole nei riguardi del metodo audio-orale, contribuendo a migliorare ulteriormente l'andamento delle classi. La sperimentazione, che si riferiva a classi di inglese, francese e tedesco, ha dato risultati positivi per tutte e tre le lingue; non sono state rilevate differenze significative che potessero essere attribuite alle differenze tra le lingue in oggetto.

4.3. Procedure di somministrazione

Le tecniche qui di seguito suggerite possono realizzarsi abbastanza semplicemente in classe, mediante lettura degli stimoli da parte dell'insegnante; se si vuole tuttavia che il test sia oggettivo e abbia un sufficiente grado di affidabilità, occorre che gli stimoli siano registrati su nastro, per assicurare uniformità nelle successive somministrazioni della prova. Per le risposte, vengono distribuiti fogli appositi — riprodotti a stampa o ciclostile — sui quali, oltre agli spazi per le indicazioni abituali (nome, classe, data, ecc.) sono predisposte, a fianco del numero che contrassegna ciascun quesito, le alternative tra cui l'esaminato deve scegliere.

La correzione avviene mediante il sistema della griglia trasparente, anch'essa molto facile da realizzare: basta prendere un foglio-allievo e segnare su di esso le risposte esatte, poi sovrapporre un foglio trasparente o traslucido (acetato di cellulosa, carta oleata, plastica trasparente o simili) e tracciare con un pennarello o una matita vetrografica un cerchio attorno a ciascun segno corrispondente alle risposte esatte. La griglia di correzione, sovrapposta ai vari fogli-allievo, permetterà di individuare immediatamente le risposte esatte (segni nei cerchi) e quelle errate (segni fuori dai cerchi).

³⁵ Col termine *accenti verbali* denotiamo quelli operanti all'interno delle singole parole, mentre gli *accenti frasali* operano nell'ambito di enunciati di maggiore estensione (intere proposizioni, segmenti di periodo, frasi, ecc.) In entrambi i casi si hanno accenti *primari*, *secondari* e *deboli*.

³⁶ Ci riferiamo in particolare alle indagini effettuate presso la Civica Scuola Serale Sperimentale di Lingue, gestita dal Comune di Milano in Via Orseolo, 1 nell'anno scolastico 1970-71 e presso le classi sperimentali funzionanti nelle altre Civiche Scuole Serali di Lingue.

Ogni prova o parte di essa è preceduta dalle necessarie istruzioni, che saranno nella lingua materna o nella lingua straniera a seconda dei criteri adottati, nonché della preparazione della classe in rapporto alla difficoltà delle istruzioni stesse. Queste devono, in ogni caso, risultare ben chiare agli esaminati, altrimenti noi non verificheremo il possesso delle abilità su cui verte il test, bensì la capacità dell'esaminato di comprendere le istruzioni. Se la prova è registrata su nastro magnetico, anche le istruzioni devono essere impartite mediante il registratore. Così facendo:

- si garantiscono pari condizioni di lavoro a tutti i gruppi;
- ci si cautele contro la possibile omissione di istruzioni essenziali;
- si abitua il gruppo all'ascolto dal mezzo tecnico immediatamente prima di dar corso alla prova vera e propria.

In questo capitolo e nei successivi, fino all'ottavo compreso, riportiamo solo esemplificazioni di singoli quesiti, rimandando alla parte finale il discorso sulla costruzione e sperimentazione di un intero test.

Per ciascun tipo di test daremo le istruzioni, uno o più esempi di quesiti riferiti alla lingua francese (Fr.) o inglese (Ing.), e le eventuali indicazioni riguardanti il foglio-allievo. Le istruzioni qui riportate sono solo indicative per il lettore e ridotte all'essenziale: quelle da impartire di volta in volta agli esaminandi non possono essere qui ipotizzate, in quanto non dipendono solo dalla natura dei quesiti ma anche da altri fattori, quali la precedente esperienza in prove analoghe, l'età, il livello scolastico, ecc.

4.4. Test esclusivamente orali

4.4.1. Riconoscimento e discriminazione di suoni

Esempio 1 (Ing.). Viene fatta ascoltare una serie di vocaboli, contenenti il suono / i / o il suono /i: /; l'allievo deve segnare A per indicare che il vocabolo contiene il suono / i / e B per indicare il suono / i: /.

Stimoli: 1.	<i>lip</i>	Il foglio-allievo reca:	1.	A	B
	2.	<i>leap</i>	2.	A	B
	3.	<i>reach</i>	3.	A	B
	4.	<i>rich</i>	4.	A	B
	ecc.		ecc.		

Si noti che questo test, come molti dei successivi, si basa sulle coppie di opposizione minimale,³⁷⁵ isolando un problema alla volta.

Esempio 2 (Fr.). Stessa tecnica dell'esempio 1; suoni / ð / e /o/.

- Stimoli: 1. *beau*
2. *bon*
3. *son*
4. *sot*

³⁷ Le coppie di opposizione minimale (*paires minimales* o *minimal pairs*) sono coppie di vocaboli che si differenziano per un solo suono vocalico o consonantico: in italiano, *pace/pece* e *pece/pesce* sono esempi di opposizioni minime. Sui *minimal pairs* si basano gli esercizi sulla discriminazione e la fissazione dei fonemi.

ecc.

Esempio 3 (Ing.). Ogni stimolo consiste in una coppia di parole. Si deve segnare U per indicare che sono uguali, D per indicare che sono diverse. Suoni / e / ed / ae /.

Stimoli: 1. <i>beg, beg</i>	Il foglio-allievo reca: 1. U D
2. <i>set, sat</i>	2. U D
3. <i>pan, pen</i>	3. U D
ecc.	ecc.

Nel caso di istruzioni impartite nella lingua straniera, varieranno anche le lettere impiegate sul foglio per le risposte: S e D (*same/different*) in inglese, M e D (*même/différent*) in francese, e così via.

Esempio 4 (Fr.). Stessa tecnica dell'esempio 3. Suoni / u / e / y /.

Stimoli: 1. *lu, loup*
2. *pour, pour*
3. *sous, su*
ecc.

Questa tecnica può essere impiegata con intere frasi anziché con vocaboli isolati.

Esempio 5 (Ing.). Come gli esempi precedenti. Suoni /a:/ e /ou/.

Stimoli: 1. *Throw that ball. Throw that bowl.*
2. *That's a big hall. That's a big hall.*
3. *They sow it in the field. They saw it in the field.* ecc.

È indispensabile che i vocaboli e le frasi di ciascuna coppia vengano pronunciati con identica intonazione.

Esempio 6 (Ing.). Ogni stimolo comprende tre parole o tre frasi, in opposizione minimale. Occorre segnare 1 e 2 sul foglio-allievo per indicare che la prima e la seconda sono uguali; 2 e 3 per la seconda e la terza; 1 e 3 per la prima e la terza; 1, 2 e 3 se sono tutte uguali; 0 se sono tutte diverse.

Stimoli: 1. <i>fat, sat, fat</i>	Il foglio-allievo reca: 1. 1 2 3 0
2. <i>sin, thin, fin</i>	2. 1 2 3 0
3. <i>thick, sick, sick</i>	3. 1 2 3 0
4. <i>through, through, true</i>	4. 1 2 3 0
ecc.	ecc.

Come si vede, è possibile realizzare quesiti abbastanza difficili anche con stimoli semplici, aumentando il numero delle alternative e complicando la tecnica di risposta.

Esempio 7 (Fr.). Tecnica dell'esempio 6, con frasi intere invece di parole isolate.

Stimoli: 1. *Voilà des hommes. Voilà dix hommes. Voilà deux hommes.*
2. *Je suis Parisien. Je suis Parisienne. Je suis Parisienne.*
3. *Je finis mon devoir. J'ai fini mon devoir. Je finis mon devoir.* ecc.

4.4.2. Intonazione e accento

Esempio 8 (Fr.). Intonazione. Tecnica dell'esempio 6.

Stimoli: 1. *Il vient. Il vient? Il vient.*

2. *J'ai lui dit pourquoi. J'ai lui dit pourquoi. J'ai lui dit: pourquoi?*

3. *Un cognac? Un cognac! Un cognac! ecc.*

Esempio 9 (Ing.). Intonazione. Tecnica dell'esempio 3.

Stimoli: 1. *I beg your pardon. I beg your pardon?*

2. *John's here? John's here.*

3. *How are you? How are you? ecc.*

Esempio 10 (Ing.). Accento. Tecnica dell'esempio 3.

Stimoli: 1. *còmment, còmment*

2. *recòrd, rècord*

3. *insight, incite ecc.*

Questi ultimi esempi dimostrano come le varie tecniche (scelta binaria, scelta quintupla, "uguale/diverso") siano in buona misura intercambiabili, permettendo di formulare quesiti adeguati al livello degli esaminati, indipendentemente dal contenuto del quesito. Si noti che queste tecniche possono venire impiegate anche con studenti a livello avanzato: il secondo quesito dell'esempio 9 richiede che gli studenti riconoscano come, in una frase a struttura affermativa, l'intonazione ascendente sia usata per esprimere stupore o sorpresa.

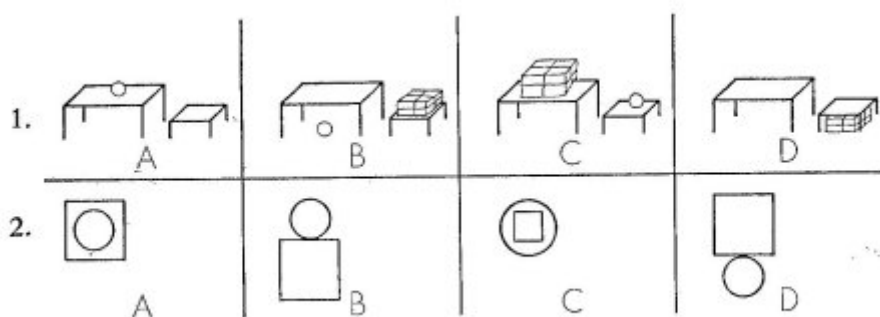
4.4.3. Comprensione auditiva

Esempio 11 (Ing.). Gli stimoli sono intere frasi. A ciascuna di esse, sul foglio-allievo, corrispondono diverse figure, una sola delle quali rappresenta adeguatamente il significato della frase. L'allievo deve segnare la lettera corrispondente alla figura prescelta.

Stimoli: 1. *The box is on the small table.*

2. *There is a circle in a square.*

Foglio allievo:



Quesiti di questo tipo si prestano facilmente ad essere formulati ponendo scarsa attenzione al contenuto linguistico, ovvero sovrapponendo le difficoltà; occorre quindi analizzarli attentamente, anche per determinarne la difficoltà che essi presumibilmente presentano all'esaminato. Il primo quesito è piuttosto complesso: richiede che si conosca il significato di *small* e di *on*, e inoltre che non si confonda *box* con *ball*. Solo a queste condizioni si può giungere senza errori alla risposta esatta. Il secondo quesito è più semplice: *circle* e *square* non sono facilmente confondibili (sul piano fonologico), e la difficoltà maggiore risiede nel capire esattamente *in* invece di *on*.

La difficoltà del quesito (e questo vale per tutti i test a scelta multipla) è anche in relazione alla capacità delle risposte errate di apparire esatte. Se in un quesito in francese entra in gioco la parola *bateau*, figure rappresentanti un *baton* o un *pataud* possono servire a controllare l'esatta percezione del vocabolo, mentre figure rappresentanti un *avion*, una *vulture*, o un *train* sono del tutto inefficaci a tale scopo, anche se possiedono una contiguità di significato con la parola-chiave dello stimolo.

4.5. Test implicanti l'uso della lingua straniera scritta

4.5.1. Posizione dell'accento tonico

Esempio 12 (Ing.). Individuare la sillaba su cui cade l'accento primario della parola-stimolo e contrassegnare la lettera corrispondente.

1. un be liev a ble
A B C D E
2. mis - un - der - stand - ing
A B C D E
3. ne - ces - sa - ry
A B C D

Esempio 13 (Ing.). Individuare la sillaba su cui cade l'accento frasale primario e contrassegnare la lettera corrispondente.

1. Get out of it!
A B C D
2. What are you up to?
A B C D E
3. Please te - le - phone me.
A B C D E

Dei quesiti nn. 12 e 13 esistono due versioni. La prima prevede che le parole o frasi vengano ascoltate dall'esaminato, e verifica la capacità di riconoscere la collocazione degli accenti primari (verbali e frasali); la seconda versione è invece basata solo sullo stimolo scritto e richiede che l'esaminato pronunci mentalmente le parole o frasi, e individui la posizione degli accenti. In questa seconda ipotesi usciamo però dall'ambito della vera e propria sfera auditiva, in quanto sono interessati anche aspetti produttivi (ossia il saper parlare), e ciò comporta ulteriori problemi per l'esaminato: nel caso dell'es. 13 si richiede che sappia determinare la posizione normale dell'accento frasale primario, ossia la sillaba su cui cade l'accento più forte quando la frase è pronunciata senza forme enfatiche, nel modo più naturale e consueto. La maggior difficoltà della seconda versione è in parte compensata dalla possibilità, per l'esaminato, di lavorare seguendo un ritmo personale anziché il ritmo con cui gli giungono gli stimoli auditivi.

4.5.2. Intonazione

Esempio 14 (Fr.). Ascoltare le frasi-stimolo e individuare, tra le alternative date, quella che meglio esprime l'atteggiamento globale del parlante.

Stimolo: 1. *Donnez-moi vos cahiers!*

- Foglio-allievo: 1. A Domanda
B Ordine
C Affermazione
D Incertezza

Esempio 15 (Ing.). Come il precedente.

Stimolo: 1. *John said so? Well, I never...*

- Foglio-allievo: 1. A Domanda
B Esclamazione
C Stupore, meraviglia
D Affermazione

Prove di questo tipo presuppongono che agli studenti siano state illustrate le intonazioni operanti in una lingua, in relazione non solo alle strutture grammaticali ma anche agli atteggiamenti e sentimenti che il parlante può esprimere alterando l'intonazione normale (dubbio, sorpresa, irritazione, meraviglia, ecc.). Sono quindi adatte a studenti a livello avanzato, ma devono essere preparate con molta cura per evitare stimoli ambigui, riconducibili a più significati. Si osservi come il significato globale non sia dato dalla sola intonazione, bensì dall'accoppiamento degli elementi fonologici con quelli lessicali e morfosintattici; queste prove introducono quindi il discorso sulla verifica globale della comprensione auditiva.

4.5.3. Test di comprensione auditiva

Esempio 16 (Fr.). Dopo aver ascoltato ciascuna frase-stimolo segnare, tra le alternative proposte sul foglio-allievo, quella che per significato si ricollega allo stimolo.

Stimolo: 1. *Yvette est la femme de Pierre.*

- Foglio-allievo: 1. A Pierre est le fiancé d'Yvette.
B Pierre n'est pas marié.
C Pierre est le mari d'Yvette.
D Yvette et Pierre viennent de se marier.

Esempio 17 (Ing.). *Stessa tecnica.*

Stimolo: 1. *Are you feeling better, today?*

- Foglio-allievo: 1. A Much better, thanks.
B Hullo! Here you are, after your flu!
C Have an aspirin, then.
D Most people don't care about feelings.

Esempio 18 (Ing.). Dopo aver ascoltato molto attentamente il brano, rispondere alle domande sul foglio-allievo, scegliendo tra le alternative proposte.

Stimolo: *Here's Mr. Brown. Today has been a very hard day for him at the office. There was a lot of traffic this morning, the bus went slowly and he arrived late. Then he found that his secretary was ill and had remained at home. As there were a lot of important things to do, he had to stay there after time.* (61 parole)

Foglio-allievo:

1. How has Mr. Brown's day been?
 - A Not so bad.
 - B Just as usual.
 - C Rather tiring.
 - D Very hard.

2. Why was Mr. Brown late this morning?
 - A He missed the bus.
 - B The train was late.
 - C There was a lot of traffic.
 - D He felt ill.

3. What about Mr. Brown's secretary?
 - A She was late, too.
 - B She stayed at home.
 - C She was on time.
 - D We don't know.

4. What is Mr. Brown's job?
 - A Manager.
 - B Director.
 - C Clerk.
 - D We don't know.

5. What did Mr. Brown have to do in the evening?
 - A He stayed after time.
 - B He worked faster than usual.
 - C He did only the most important things.
 - D He had to go slowly.

Oltre a quanto già osservato a proposito dell'esempio n. 11, nella redazione di quesiti di questo tipo occorre prestare attenzione ad altre esigenze. Anzitutto, il brano non deve essere né troppo breve né troppo lungo; in genere, una lunghezza compresa tra i 40 e gli 80 vocaboli appare la più idonea a presentare abbastanza informazioni per permettere di formulare da 4 a 7 domande. Non è produttivo presentare brani molto lunghi che però si prestano a poche domande; d'altro canto la concentrazione di molte notizie (e relative domande) in poche parole aumenta la produttività ma anche la difficoltà del quesito, per cui sarà sempre necessario ricercare un buon equilibrio tra le varie componenti (stimolo, domande, alternative). Nel quesito debbono essere inserite e verificate quelle difficoltà che l'insegnante ha affrontato nel suo programma. Nell'esempio 18, tra l'altro, si verifica la comprensione delle espressioni di tempo *late*, *on time*, *after time*.

L'esempio seguente illustra un altro tipico errore di redazione dei quesiti.

Esempio 19 (Fr.). Tecnica dell'esempio 16.

Stimolo: 1. *Il y a des canaris qui gazouillent dans cette cage.*

Foglio-allievo: 1. A Tous les oiseaux qui gazouillent vivent dans des cages.

B Pas de cages sont grandes assez pour ce canari.

C Il y a une cage avec des canaris.

D C'est le canari le plus menteur de la cage.

Notiamo subito che tutte le alternative, tranne la C, sono frasi palesemente false, assurde o quantomeno del tutto improbabili; l'esaminato non ha quindi bisogno di comprendere lo stimolo per giungere alla risposta esatta: gli basta saper leggere. Questo quesito è perciò da scartare o da revisionare.

CAP. 5 IL SAPER PARLARE

5.1. La produzione di messaggi verbali

Nell'ambito del “saper parlare” rientrano tutti gli aspetti *attivi* della comunicazione verbale. Delle modalità della comunicazione — e dei fattori che vi sono coinvolti — abbiamo già discusso nel capitolo precedente, mettendo in risalto gli aspetti passivi. Per ciò che attiene alla *produzione* dei messaggi verbali, notiamo anzitutto che essa risulta dall'interazione di numerosi elementi, tra cui i principali sono:

- a livello fonologico, le abitudini articolatorie, ossia la capacità di riprodurre esattamente intonazioni, accenti, suoni;
- a livello morfosintattico, la padronanza delle strutture sul piano orale, in termini di *scelta* delle strutture appropriate e di *impiego* efficiente;
- a livello lessicale, la padronanza di un adeguato lessico *attivo* (*scelta* dei vocaboli richiesti e corretto *impiego* nei periodi);
- a livello di civiltà, la capacità di partecipare attivamente al dialogo nelle situazioni tipiche della civiltà straniera.

La verifica del saper parlare è, tra gli aspetti del testing, quello in cui i progressi sono stati più lenti, le difficoltà maggiori, e i risultati meno soddisfacenti. Ancora nel 1964 R. Lado scriveva che è “un settore nel quale si registra un fallimento quasi generale”.³⁸ Oggi, malgrado non sia stata trovata una soluzione adeguata a molti problemi, si è tuttavia giunti ad una definizione delle tecniche più efficienti per impostare un'indagine accurata e sufficientemente oggettiva del saper parlare.

Si è infatti constatato che i sistemi tradizionali non davano sufficiente affidamento soprattutto in quanto essi tendevano genericamente a “far parlare” l'esaminato, per mezzo di stimoli vari; questi mancavano però di una precisa strutturazione, tale da far sì che l'esaminato fornisse una prestazione già stabilita in precedenza. La reazione dell'esaminato veniva considerata e valutata globalmente, con l'instaurarsi di tutti gli elementi perturbanti da noi già analizzati in rapporto all'effetto di alone.³⁹

Le tecniche proposte per una verifica efficace del saper parlare si basano su tre presupposti fondamentali:

- a) la preparazione di materiale adeguato che serva di stimolo al parlante;
- b) la precisazione preventiva degli elementi considerati;
- c) la valutazione delle risposte sulla base del solo elemento su cui era imperniato lo stimolo, tralasciando sia gli altri singoli fattori, sia ogni tentativo di valutazione globale.

5.2. Misurazione del saper parlare

Nel valutare la produzione orale non è possibile eliminare totalmente la soggettività di giudizio dell'esaminatore, tuttavia si può ridurla notevolmente mediante opportuni accorgimenti.

L'uso di un registratore consente all'esaminatore di riascoltare più volte le prestazioni dell'esaminato: oltre ad avere modo di accertare esattamente forma e contenuto delle frasi pronunciate, il valutatore evita di emettere un giudizio “a caldo”, secondo l'impressione del momento. Anche il confronto tra diverse prestazioni viene notevolmente agevolato.

³⁸ R. LADO, Per una didattica scientifica delle lingue, cit., p. 199.

³⁹ Cfr. par. 2.5.

In secondo luogo, bisogna evitare di ricorrere alle scale tradizionali. Numerose esperienze⁴⁰ hanno dimostrato che i migliori risultati, dal punto di vista dell'affidabilità dei punteggi, si ottengono lavorando su una scala di tre punti:

10 = prestazione corretta;

5 = prestazione imprecisa ma intelligibile;

0 = prestazione del tutto scorretta, inintelligibile.

In questa scala, 10, 5 e 0 non sono assolutamente da porre in relazione con i voti in decimi, ed è significativo il fatto che altri autori abbiano proposto scale diverse, come 5, 3 e 0 oppure 2, 1 e 0. Qualunque siano i valori prescelti, essi hanno solo lo scopo di quantificare il giudizio sulle prestazioni, in modo tale che si possano confrontare tra di loro successive prove dello stesso studente o prove di studenti diversi. Questi confronti si effettuano, di solito, dopo aver calcolato il punteggio in centesimi corrispondente ad ogni serie di rilevazioni, secondo la seguente procedura:

1) Annotare i singoli giudizi: *esempio* 0 5 5 10 5 10

2) Calcolare la somma dei punti: 35

3) Calcolare il punteggio massimo possibile in rapporto al numero dei giudizi: 60 (per 6 giudizi)

4) Dividere la somma dei punti per il punteggio massimo e moltiplicare per 100: $35 / 60 \times 100 = 58$

Si tenga presente che i punteggi in centesimi non vanno confusi con le percentuali, e che punteggi ottenuti con scale diverse non sono sempre confrontabili tra di loro. Chiariremo quest'ultima affermazione con un esempio, calcolando con due scale diverse il punteggio in centesimi di uno studente che, in cinque rilevazioni, abbia sempre ottenuto il giudizio intermedio.

Scala 10/5/0.

Punteggi: 55555; somma: 25; punteggio massimo: 50; punteggio in centesimi: $25/50 \times 100 = 50$.

Scala 5/3/0.

Punteggi: 33333; somma: 15; punteggio massimo: 25; punteggio in centesimi: $15/25 \times 100 = 60$.

5.3. Produzione di suoni e di parole

La tecnica più semplice per verificare la buona pronuncia di un suono consiste nel procurarsi un numero adeguato di parole o frasi in cui esso è contenuto e chiedere allo studente di ripeterle.

I suoni consonantici dovranno essere controllati in posizione iniziale, mediana, finale e nei vari nessi; le vocali e i dittonghi debbono comparire sia in sillaba aperta che in sillaba chiusa.

Esempio 20 (Fr.). Ripetere le frasi-stimolo.

Stimoli: 1. *Luc a vu une lagune.* (suono /y/)

2. *Ton oncle est bon, donc!* (suono /õ/) ecc.

Il giudizio va formulato solo sul suono-chiave di ogni frase. Il limite più notevole di questa tecnica sta nel fatto che non si ha un'autentica produzione orale quanto piuttosto una riproduzione di suoni appena uditi. Gli esempi che seguono illustrano alcune tecniche che permettono di evitare la ripetizione.

⁴⁰ V. l'ultimo capitolo del volume di E. M. STACK, *The Language Laboratory and Modern Language Teaching*, New York, Oxford University Press, 1960. In Italia, la tecnica è stata sperimentata dal C.L.A.Di.L. a Monza con gruppi di studenti a vari livelli sotto il controllo di uno psicolinguista, il Prof. Umberto Dell'Acqua.

Esempio 21 (Ing.). Gli stimoli sono parole contenenti la vocale /a:/ come in *cart*; le risposte sono le parole corrispondenti, contenenti però la vocale /ʌ/ come in *cut*.

Stimoli: 1. *heart*

2. *pari*

3. *carp* ecc.

Nelle istruzioni agli allievi è opportuno inserire uno o più esempi: ciò vale per la maggior parte delle tecniche presentate in questo capitolo.

Esempio 22 (Fr.). Gli stimoli sono vocaboli femminili, con la consonante nasale; le risposte sono i corrispondenti maschili, con la vocale nasale.

Stimoli: 1. *parisienne*

2. *bonne*

3. *patronne* ecc.

Esempio 23 (Ing.). Nomi singolari da volgere al plurale.

Stimoli: 1. *boy*

2. *match*

3. *book* ecc.

Alternando opportunamente i vocaboli, questo test verifica l'uso corretto delle desinenze /-s// -z// -iz/.

Esempio 24 (Ing.). Verbi al presente da volgere al passato.

Stimoli: 1. *play*

2. *start*

3. *stop* ecc.

Alternando opportunamente i vocaboli, questo test verifica l'uso corretto delle desinenze /-t// -d// -id/.

I test nn. 22, 23 e 24 possono ovviamente essere impiegati nel senso opposto, ossia per il passaggio dal maschile al femminile, dal plurale al singolare, dal passato al presente. In essi non debbono invece essere introdotte le forme irregolari, perché ciò impedirebbe di concentrare l'attenzione sui problemi di pronuncia; la verifica di tali forme va effettuata separatamente, con quesiti opportunamente predisposti, e preferibilmente mediante strutture complete e non singoli vocaboli.

Per stimolare la produzione di vocaboli si può far ricorso a figure, disegni o altro materiale visivo, a due condizioni:

— l'immagine deve rappresentare chiaramente il vocabolo desiderato, senza possibilità di confusione (ciò è spesso difficile, e può diventare del tutto irrealizzabile per certi termini, come quelli astratti);

— il materiale visivo deve essere strutturato in modo da permettere la verifica di un problema di pronuncia alla volta.

5.4. Produzione di frasi

Se è difficile che le immagini rappresentino inequivocabilmente un vocabolo, è ancor più difficile che esse presentino situazioni tali da condurre esclusivamente a una frase prestabilita. Mentre il disegno o la fotografia di un albero possono ottenere che l'allievo produca *arbre o tree*, è pressoché impossibile suscitare frasi come *Voilà un arbre o There's a tree over there*, escludendo nel contempo ogni altra frase.

Specialmente in Francia⁴¹ il problema dei rapporti tra immagine e parola è stato analizzato in maniera approfondita, e sono state elaborate delle soluzioni molto interessanti. Si è tentato di giungere alla definizione dell'immagine, che è per sua natura polisemica, soprattutto ricorrendo a disegni integrati da segni convenzionali chiarificatori: X per le frasi negative, ? per le domande, calendari per le espressioni di tempo, ecc. I risultati, che hanno condotto alla produzione di copioso materiale didattico, non ci sembrano però soddisfacenti in quanto a possibilità di impiego nel testing, ove le esigenze di chiarezza e univocità si collocano ad un livello assai più rigoroso.

Per le frasi che esprimono azioni, un tentativo di soluzione può essere ricercato nell'immagine in movimento (film, videonastri, ecc.); tuttavia l'impiego di questi mezzi è tutt'altro che agevole, nelle normali situazioni scolastiche, e poi nemmeno il movimento riesce a fornire esattamente ed esclusivamente le informazioni necessarie. Supponiamo che la sequenza filmata mostri una donna nell'atto di aprire una finestra: quale sarà la frase implicata dall'azione, poniamo, in inglese? Forse *The woman is opening the window* oppure *She's opening that window* ovvero *That lady is opening a window*, o ancora altrimenti. Come rappresentare *Helen is opening this window* senza, ovviamente, ricorrere a scritte di alcun genere dato che non stiamo verificando il saper leggere? Lo studente che reagisse all'apertura della finestra con *She's probably hot* come lo giudicheremmo?

Accantoniamo perciò l'approccio visivo per prendere in considerazione gli altri metodi che ci permettono di far produrre frasi in lingua straniera.

La tecnica più diffusa si basa sulle domande a risposta obbligata.

Esempio 25 (Fr.). Rispondere alle domande con un'intera frase.

Stimoli: 1. *Quelle est la capitale de la France?*

2. *Dans quel département se trouve Bordeaux?* ecc.

La ragione per cui si richiede "un'intera frase" risiede nel fatto che in questo tipo di prova noi non stiamo controllando il possesso di nozioni, bensì la capacità di parlare, e quindi dobbiamo evitare di ricevere risposte troppo brevi, di una sola parola (*Paris, Gironde*); nel caso vi sia il fondato timore che il contenuto nozionistico della risposta possa presentare difficoltà per l'esaminato, completeremo la domanda con la parola-chiave della risposta:

3. *Quel est le fleuve qui traverse Nantes? (Loire)*

Le domande meno produttive sono quelle che prevedono una risposta *sì o no*, mentre quelle che iniziano con un pronome o avverbio interrogativo (chi, dove, quando, perché, ecc.: in inglese, *wh-questions*) possono molto facilmente ammettere più di un modo di formulare la risposta. Per i nostri scopi, le domande comprendenti un'alternativa sono spesso assai utili.

Esempio 26 (Ing.). Rispondere con un'intera frase.

Stimoli: 1. *Is it hot or cold in summer?*

2. *Does Tom's father work in a factory or on a farm?* ecc.

Per ottenere risposte negative, formuleremo delle domande terminanti con "... o no?".

Esempio 27 (Ing.). Rispondere negativamente con un'intera frase.

Stimoli: 1. *Is Cardiff in Scotland or not?*

2. *Do they work on Sundays or not?* ecc.

⁴¹ Si veda il copioso materiale prodotto per iniziativa di organismi quali il C.R.E.D.I.F. e il B.E.L.C., e gli studi di vari Autori, tra cui il Capelle.

Il grosso limite delle tecniche basate sulle domande è che esse conducono necessariamente alla formulazione di risposte, ossia di frasi dichiarative, affermative o negative che siano. Volendo far produrre delle frasi interrogative, si può impostare la prova su stimoli del tipo “chiedimi se...”.

Esempio 28 (Ing.). Formulare le domande richieste.

Stimoli: 1. *Ask me if Ted is here* (per ottenere: *Is Ted here?*).

2. *Ask me if they work hard* (per ottenere: *Do they work hard?*) *ecc.*

Un altro sistema, pure abbastanza diffuso, è quello che implica la trasformazione delle frasi-stimolo secondo le tecniche degli esercizi strutturali. Questo sistema ha il vantaggio di permettere la produzione di qualsiasi struttura e di sfruttare, in parte, lo stesso materiale degli esercizi.

Esempio 29 (Fr.). Trasformare le frasi secondo l'esempio. Es.: *J'ai pris une boîte / Voilà la boîte que j'ai prise.*

Stimoli: 1. *J'ai offert une cigarette.*

2. *J'ai écrit une lettre.* *ecc.*

5.5. Testing e laboratorio linguistico

L'accento agli esercizi strutturali ci porta ad alcune considerazioni sul testing nel laboratorio linguistico. Il materiale registrato si presta infatti al controllo delle capacità dello studente per quanto riguarda *immediatezza, appropriatezza, intonazione e struttura.*

Il criterio dell'*immediatezza* è conseguente alla funzione degli esercizi di fissazione delle strutture e all'esigenza di formare delle *abitudini* linguistiche, ossia di far sì che l'allievo reagisca prontamente, senza tradurre, a stimoli in lingua straniera, come avviamento alla capacità di dialogare.

Per *appropriatezza* si intende la corrispondenza tra stimolo e risposta. Uno studente potrebbe produrre frasi corrette dal punto di vista fonologico e morfosintattico, ma diverse da quelle richieste: in questo caso diremo che manca di appropriatezza.

Il termine *intonazione* sta ad indicare non solo la verifica degli schemi intonatori, ma della pronuncia delle frasi nel loro complesso. La scelta di questo termine intende sottolineare l'importanza degli aspetti prosodici (o soprasegmentali) della lingua nella comunicazione verbale, e la necessità che l'insegnante sviluppi e controlli (non soltanto in laboratorio) questi aspetti del messaggio, tradizionalmente trascurati.

Il controllo della *struttura* consiste nella verifica della correttezza grammaticale delle frasi prodotte, indipendentemente dalla loro appropriatezza.

Per la rilevazione sistematica delle prestazioni in laboratorio è consigliabile predisporre un congruo numero di apposite schede. Su ognuna di esse vi sarà spazio per l'indicazione della classe, della data, e della struttura su cui verte il test. Ad ogni allievo corrisponde un riquadro in cui annotare, oltre al nome, i punteggi delle singole rilevazioni, la somma e il punteggio in centesimi. Ciascun riquadro è collocato nel foglio in modo da riprodurre la pianta del laboratorio, e reca il numero della cabina occupata dallo studente.

Chi fa del testing in laboratorio deve tenere presenti due avvertenze. La prima, di ordine generale, si riferisce ai rischi di confusione tra esercizio e test; avendo già sviluppato questo concetto, qui ci basta ribadire la necessità di separare il momento dell'apprendimento da quello del controllo.

La seconda avvertenza si riallaccia alla prima, in quanto in laboratorio è estremamente importante cogliere il momento esatto per effettuare rilevazioni oggettive. Quando lo studente è in fase di apprendimento, in ogni esercizio strutturale si attua una progressione dalle prime battute alle ultime: nel procedere dell'esercizio, infatti, lo studente supera gradualmente varie difficoltà, cosicché verso la fine di esso le sue prestazioni, globalmente considerate, sono migliori che all'inizio. La verifica

operata quando lo studente è ancora nella fase di fissazione delle strutture risente negativamente della disomogeneità delle prestazioni nei diversi momenti, pertanto non è possibile comparare un controllo effettuato su uno studente alle prime battute con un controllo su un altro studente alle ultime battute dello stesso esercizio. Se invece la verifica viene effettuata soltanto in una fase finale, dopo che il materiale è stato sufficientemente impiegato dagli studenti, la disomogeneità dei controlli è eliminata o quantomeno ridotta entro limiti sufficientemente ristretti.⁴²

5.6. Valutazione della conversazione

Il saper parlare non può in alcun modo essere limitato alla produzione di suoni o frasi secondo schemi predeterminati; già dalla prima unità didattica è necessario giungere ad esercizi di tipo creativo e all'espressione autonoma sia pure, è ovvio, nel ristretto ambito del materiale appreso. A questo punto non è più possibile effettuare prove oggettive nel senso tecnico del termine (v. par. 1.5.).

Ciò non significa, tuttavia, che l'insegnante sia abbandonato a se stesso, soprattutto se avrà avuto cura di seguire un rigoroso itinerario didattico che procede dalla ripetizione e memorizzazione, per giungere gradualmente alle risposte obbligate, ai questionari aperti, alla conversazione guidata e, infine, all'espressione libera e alla drammatizzazione. Le osservazioni e valutazioni condotte lungo tutto l'arco di questo processo di apprendimento permettono non solo l'analisi dei singoli problemi, la correzione degli errori e il superamento delle difficoltà, ma anche l'accertamento dell'abilità del parlare in tutte le sue componenti e manifestazioni. È uno dei tanti casi in cui insegnamento e valutazione si sostengono e si illuminano a vicenda.

Questo è il settore in cui il language testing potrà svilupparsi maggiormente nei prossimi anni, attraverso opportune sperimentazioni; noi speriamo di aver sufficientemente illustrato le premesse metodologiche e le linee direttrici lungo le quali potrà procedere la ricerca, e di aver fornito i parametri essenziali per la corretta impostazione del problema.

⁴² La maggior parte dei laboratori linguistici permette di registrare in qualsiasi momento le prestazioni degli allievi su uno dei registratori a disposizione dell'insegnante. Ciò consente di effettuare il *monitoring* normalmente, e di impiegare in un secondo momento, con tutta tranquillità, il nastro coi test dei singoli studenti, per una valutazione accurata.

CAP. 6 II SAPER LEGGERE

6.1. Aspetti generali

Nel “saper leggere” rientrano, oltre alla lettura in senso stretto, tutti i problemi inerenti alla comprensione dei messaggi grafici. Escludiamo però dal nostro discorso, ritenendola di scarso interesse per la stragrande maggioranza degli insegnanti italiani di lingue straniere, la problematica relativa all'uso di alfabeti diversi da quello latino, come il greco, il cirillico, l'arabo, ecc. Ovviamente, codici grafici diversi richiedono, anteriormente alla lettura, apposite esercitazioni per la produzione ed il riconoscimento dei nuovi grafemi.⁴³

Assai più importante ci sembra ricordare come il codice scritto sia una rappresentazione secondaria e approssimativa del codice orale. È secondaria nel senso che una lingua è essenzialmente orale, costituita da simboli vocali, a cui non corrispondono necessariamente simboli scritti.⁴⁴ È approssimativa perché un testo scritto non può riprodurre tutti gli aspetti del messaggio orale, essendo soprattutto carente nei riguardi dei valori prosodici (ritmo, accento, intonazione), degli aspetti acustici (volume e timbro di voce), e di quelli collaterali (cadenze dialettali e simili). Si pensi a quanto inadeguata sia la punteggiatura di uso corrente a rappresentare l'infinita, gamma di pause, schemi intonatori ed effetti retorici possibili in qualsiasi lingua.

Col termine *lettura*, salvo espressa indicazione, non intendiamo il leggere a voce alta: la lettura ad alta voce è, nella vita quotidiana, un'attività assai meno frequente della lettura silenziosa. Al di fuori della scuola, la ritroviamo nelle cerimonie ufficiali (civili e religiose), nelle trasmissioni radiotelevisive, e in pochi altri casi. La lettura a voce alta è in realtà la somma delle abilità del saper parlare e del saper leggere.

Insegnando le lingue straniere con i metodi tradizionali, la lettura a voce alta aveva un ruolo primario, sia come esercizio che come test. Il presupposto era l'esistenza delle “regole di pronuncia”, ossia di quel complesso di norme che permetterebbero, teoricamente, di desumer il codice orale da quello scritto. Su tale presupposto sono state costruite lingue artificiali, come l'Esperanto, presentate come “perfettamente regolari”. Esiste invero un alto grado di regolarità per quanto riguarda i fonemi segmentali (consonanti, vocali e dittonghi) ma questa regolarità vien meno a livello soprasegmentale, nella misura in cui ciascun parlante di una lingua artificiale trasferisce in essa i ritmi e le intonazioni della propria lingua materna.⁴⁵

⁴³ Chi fosse interessato ad approfondire il problema può utilmente leggere il capitolo “How to compare two writing systems” in: R. LADO, *Linguistics Across Cultures*, cit, pp. 93-109.

2

3

4

5

6

7

8

⁴⁴ Un esempio moderno è costituito dalla lingua somala, la cui forma scritta è stata introdotta il 21 ottobre 1972 mediante un decreto governativo. Anteriormente a tale data, la lingua era esclusivamente parlata, e nelle scuole si insegnava a leggere e scrivere l'arabo, l'italiano o l'inglese.

⁴⁵ Le lingue artificiali ci aiutano a comprendere quanto labili siano le “regole di pronuncia” per le lingue vive, comprese quelle ritenute più “regolari”, come il tedesco, l'italiano o lo spagnolo. Sui rapporti tra codice orale e codice scritto v. G. FREDDI, *Fondamenti e metodi della didattica delle lingue*, cit., pp. 66-67.

I metodi attuali, che prevedono l'acquisizione del codice orale prima del passaggio allo scritto, hanno fatto sì che la lettura assumesse un ruolo diverso da quello tradizionale. Non si parte più da un testo scritto per ricavarne il corrispondente orale, ma ci si basa sull'acquisizione di materiale sul piano orale, e sulla sua fissazione. Quando dello stesso materiale viene presentata la forma scritta, lo studente non compie un lavoro di ricomposizione bensì di riconoscimento. Questa procedura, se ben condotta, permette di evitare le interferenze della grafia sulla pronuncia, e di realizzare la lettura a voce alta in condizioni simili a quelle esistenti per la lingua materna, in cui si giunge a leggere quando si è già padroni della lingua parlata.

6.2. Lessico e strutture

Essendo di particolare rilievo ai fini della comprensione della lingua scritta, richiamiamo una distinzione importante, che ritroviamo nella linguistica strutturale, tra vocaboli a preminente funzione lessicale (nomi, aggettivi qualificativi, verbi principali, avverbi, ecc.) e vocaboli a preminente funzione morfosintattica (articoli, preposizioni, congiunzioni, verbi ausiliari, pronomi, ecc.). Dal punto di vista della comprensione di un testo scritto, la conoscenza degli elementi lessicali è più importante della conoscenza degli elementi grammaticali. Se da un testo togliamo i primi, lasciando inalterato tutto il resto — punteggiatura compresa — il brano non è comprensibile (i trattini rappresentano le parole omesse):

“Il -, - la -, e -, com'era -, lo - al -, - una - che non s' -, e che non avrebbe - -. Due - -, l'uno - all'altro, al -, per - così, delle due -: un di costoro, a - sul - -, con una - - al di -, e l'altro - - sul - della -; il -, in -, - al -, con le - - sul -”.

Ecco cosa otteniamo col procedimento opposto:

“curato voltata stradetta dirizzato solito sguardo tabernacolo vide cosa aspettava voluto vedere uomini stavano dirimpetto confluyente dir viottole cavalcioni muricciolo basso gamba spenzolata fuori piede posato terreno strada compagno piedi appoggiato muro braccia incrociate petto”.

Dalla seconda versione il lettore avrà probabilmente afferrato, in buona misura, il senso del paragrafo, e forse avrà riconosciuto il famoso brano, tratto dal primo capitolo dei *Promessi Sposi*, in cui si descrive l'incontro di Don Abbondio con i bravi.

Analizzando ulteriormente, osserviamo che la prima versione comprende 45 parole e 20 segni d'interpunzione, mentre la seconda comprende solo 35 vocaboli; inoltre, sono state smembrate alcune espressioni idiomatiche (“per dir così”, “in piedi”). Ciò malgrado, il brano composto dagli elementi prevalentemente lessicali conserva un discreto grado di comprensibilità.

La distinzione tra i due tipi di vocaboli, che potrebbe essere ulteriormente qualificata e precisata, per quanto illuminante non è tuttavia sufficiente a dare indicazioni precise sulle prevedibili difficoltà di un test lessicale. È indispensabile far riferimento agli studi sul lessico di base (BASIC English, Français Fondamental, Grund-deutsch, etc.) ed alle liste di frequenza dei vocaboli.⁴⁶

D.P. Harris⁴⁷ presenta un tipico esempio di quesito mal formulato perché le alternative non sono allo stesso livello di difficoltà:

very evil

⁴⁶ C. K. OGDEN, *Basic English*, Londra, Kegan Paul, Trench, Trubner & Co., 1930; M. WEST, *A general service list of English words*, Londra, Longman, 1953; R. DOTRENS, D. MASSARENTI, *Vocabulaire fondamental du Français*, Neuchâtel, Delachaux et Niestlé, s.d.; G. GOUGENHEIM, P. RIVENC, R. MICHEA, A. SAUVAGEOT, *L'élaboration du Français Fondamental (1er degré)*, Parigi, Didier, 1964.

⁴⁷ D. P. HARRIS, *Testing English as a Second Language*, New York, McGraw-Hill, p. 55.

- A. *nefarious*
- B. *delicious*
- C. *prosperous*
- D. *courageous*.

Poiché *nefarious*, la risposta esatta, è un vocabolo assai meno frequente degli altri, è prevedibile che lo studente vi giunga per eliminazione delle altre alternative. Si impone quindi una revisione del quesito, o portando tutte e quattro le alternative allo stesso grado di difficoltà (il che renderebbe il quesito adatto a studenti di livello assai progredito), ovvero cambiando la parola-chiave e sostituendo *nefarious* con un vocabolo della stessa frequenza degli altri.

Se un test specifico è basato su materiale accuratamente graduato, il pericolo di dislivelli così rilevanti è alquanto ridotto; sussiste tuttavia la necessità di verificare, in ciascun quesito, la frequenza di tutti i vocaboli impiegati.

Altre tecniche di controllo del lessico conducono a risultati incerti e poco probanti perché non tengono conto di alcune premesse linguistiche. Una di queste tecniche, abbastanza diffusa, consiste nell'assegnare un vocabolo e richiedere agli studenti di comporre una frase comprendente quel vocabolo.

Supponiamo che la parola in questione sia *courageux* e che un allievo produca la frase *Henri est un garçon très courageux*. Come giudicarla? Da un lato, essa è corretta sotto ogni punto di vista: ortografico, morfosintattico e lessicale. Dall'altro, non ci rivela affatto se lo studente conosca il significato esatto di *courageux*, dato che quel contesto può accogliere qualsiasi aggettivo qualificativo attribuibile a persona. Il quesito ha fallito il suo scopo, lasciando dubbioso l'esaminatore, e questa tecnica si dimostra incapace di condurre ad una verifica oggettiva della padronanza del lessico.

6.3. Test lessicali

Esempio 28 (Ing.). Scegliere l'alternativa migliore.

beach

novel

A. *shore*

A. *a short story*

B. *harbour*

B. *any fiction*

C. *seaside*

C. *a narrative in prose*

D. *pier*

D. *a long poem*

Essendo rari i casi di perfetta sinonimia, i quesiti di questo tipo presentano un certo grado di ambiguità, che è difficile eliminare totalmente; per questo motivo, anziché richiedere “il sinonimo” o “la definizione esatta”, è più corretto chiedere quale vocabolo (o quale definizione) si avvicini maggiormente, per significato, alla parola-stimolo. La principale fonte di ambiguità risiede nella mancanza di contestualizzazione del vocabolo, il che rende impossibile la verifica di tutte quelle parole — e sono la maggioranza — che assumono significati diversi a seconda del contesto in cui

sono inserite.⁴⁸ Una precisazione parziale è possibile attraverso la scelta delle alternative. Nell'esempio 26, sono proprio le definizioni e le scelte presentate a farci intendere che le parole-stimolo sono due sostantivi (*novel* potrebbe essere aggettivo e *beach* verbo).

Per contestualizzare un vocabolo, di solito è sufficiente inserirlo in una frase adeguata:

Esempio 29 (Fr.). Scegliere l'alternativa migliore per il vocabolo in corsivo.

Il est *étonné* par la réponse qu'il a reçue.

A. très surpris B. très fâché C. très satisfait D. très amusé.

In molti casi, anziché alle forme sinonimiche si può ricorrere al completamento di frasi:

Esempio 30 (Fr.). Completare scegliendo l'espressione adatta.

Ma maison a deux étages: et le premier étage.

A. le vestibule B. le rez-de-chaussée C. la mansarde D. le parterre.

Un test lessicale che viene spesso usato si basa sull'accoppiamento di vocaboli sulla base dell'attinenza o della contiguità di significato, ovvero sulla capacità di combinarsi in composti:

Esempio 31 (Ing.) Accoppiare ogni parola della colonna di destra con ogni parola della colonna di sinistra, scrivendola a fianco di questa.

1. railway pie
2. brick plane
3. apple sitter
4. baby wall
5. office station
6. jet building

Esempio 32 (Ing.). Combinare i vocaboli per ottenere nomi composti, con la stessa tecnica dell'esempio precedente.

1. foot. board
2. grand. paper
3. cup seller
4. bed ball
5. book. father
6. news. room

Questo tipo di test non ci sembra molto convincente per almeno due motivi. Anzitutto, manca la contestualizzazione dei vocaboli, il che è fonte di ambiguità. Secondariamente, quando l'esaminato ha ricostituito cinque coppie, può ricavare la sesta anche senza conoscere il significato dei vocaboli. L'esaminatore non ha alcun modo di verificare se vi sia stato o no un lavoro di ricostruzione per eliminazioni ed incastri né quale sia la sesta coppia eventualmente formata per esclusione. Il vantaggio, rispetto agli altri sistemi di verifica del lessico, risiede nella possibilità di operare su un

⁴⁸ L'esigenza di contestualizzare i vocaboli e di presentarli intuitivamente in strutture autosufficienti anche in vocaboli adatti ad essere impiegati nella scuola media ha condotto alla realizzazione (nell'ambito delle attività del C.L.A.Di.L.) di due "vocabolari attivi": *My English Word-Book* (a cura di P. BERNARDI e G. FREDDI) e *Mon Dictionnaire Français* (a cura di G. FREDDI), ed. Minerva Italica, Bergamo.

notevole numero di vocaboli contemporaneamente (dodici per quesito); con un numero ridotto di prove si possono controllare tutti i vocaboli di un'unità didattica, sempreché siano accoppiabili.

6.3.1. Prove lessicali integrate

Un tipo particolare di test lessicale è quello che gli Americani chiamano *cloze test*. La tecnica su cui si basa è molto semplice: da un brano (noto agli allievi) si tolgono delle parole ad intervalli regolari, sostituendole con degli spazi bianchi, e si chiede agli studenti di ricostruire il testo scrivendo negli spazi le parole omesse. Questa prova, oltre a richiedere capacità di comprensione del testo, memoria e padronanza del lessico, interessa anche l'abilità del saper scrivere, sia pure strumentalmente e in misura marginale.

I vantaggi che essa presenta sono molteplici. Anzitutto, è particolarmente flessibile, potendosi costruire numerose forme parallele ed equivalenti dello stesso test. Stabilito, ad esempio, di omettere una parola ogni dieci, una forma mancherà della 10a 20a, 30a... parola, un'altra della 5a, 15a, 25a..., e così via. La difficoltà del test è graduabile semplicemente aumentando o diminuendo la frequenza dei vocaboli omessi: un testo in cui manca una parola su cinque è notevolmente più difficile da ricostruire di un testo da cui è stata tolta solo una parola ogni quindici.

La prova è oggettiva in quanto ammette una sola soluzione esatta, ed il punteggio si ricava direttamente dal numero dei vocaboli correttamente ricollocati (l'esattezza riguarda l'individuazione del vocabolo, prescindendo da considerazioni accessorie — ad esempio, sulla correttezza ortografica).

Per queste caratteristiche, la prova si rivela particolarmente utile nelle classi di studenti non principianti: permette di riutilizzare più volte lo stesso brano, sfruttando a fondo il materiale didattico e verificando contemporaneamente i progressi di acquisizione; consente, attraverso l'impiego delle forme parallele, la differenziazione delle singole prove, premessa indispensabile per l'individualizzazione dell'insegnamento.⁴⁹

6.4. Comprensione dei testi scritti

La tecnica per la verifica della comprensione del contenuto di brani scritti è analoga a quella adottata per i brani orali e descritta al paragrafo 4.5.3., Es. 18. I passi possono essere più lunghi (da 100 a 250 parole, a seconda della difficoltà e del livello di preparazione degli esaminandi); contenendo una quantità maggiore di informazioni, permettono di formulare più quesiti. Anche in questo caso, occorre ricercare l'equilibrio ottimale tra lunghezza del brano e numero dei quesiti. Non sarebbe produttivo, infatti, produrre brani molto lunghi solo per due o tre domande, e non sarebbe facile concentrare molti quesiti su poche notizie senza generare confusione o incorrere in ripetizioni. Poiché l'esaminato ha davanti a sé il brano mentre risponde alle domande, se si vuole evitare che la scelta dell'alternativa esatta si riduca al semplice riscontro della espressione contenuta nel testo, bisogna formulare le domande e le alternative in modo tale da impegnare effettivamente le capacità di comprensione dei significati.

Esempio 33 (Ing.). Leggere attentamente il brano e rispondere alle domande finali a scelta multipla.

History of the Washington Cherry Trees

About a month after W. H. Taft became President of the United States on March 4, 1909, Mrs. Taft expressed a desire to have Japanese cherry trees planted in Potomac Park. Dr. Jokichi Taka-mine, the Japanese chemist famous as the discoverer of adrenalin, was visiting

⁴⁹ J. Donaid Bowen ha illustrato ampiamente l'impiego di questi materiali, inseriti in una tecnica di insegnamento di rilevante interesse, in alcuni recenti contributi. Tra gli altri, segnaliamo "Lecturettes" for Mature Learners, in "English Teaching Forum" XII, 1, Gennaio-Marzo 1974, pp. 8-15.

Washington at that time and learned of Mrs. Taft's interest in the trees. He played an important part in having the city of Tokyo make the gift of cherry trees.

The cherry trees destined for Washington were raised at the Okitsu Imperiai Horticultural Experiment Station at Shizoukan, Japan, under the supervision of Professor Inokichi Kuwana, director of the Yokohama Imperiai Plant Quarantine Station, assisted by Dr. Manabu Miyoshi, the noted authority on Japanese cherry trees. The scions for these trees were taken in December 1910 from the famous collection of cherry trees on the bank of the Arakawa River in the Adachi Ward, a suburb of Tokyo. These grafted trees were lifted for shipment in December 1911.

The 3,000 cherry trees of 12 varieties were shipped from Yokohama on board the S. S. *Awa Maru* on February 14, 1912. After the trees arrived in Seattle, they were transferred to insulated freight cars and sent the 2,500 miles across the continent to Washington, arriving on March 25.⁵⁰

1. *Who wanted Japanese cherry trees to be planted in Washington?*

- A. President Taft
- B. Mrs. Taft
- C. Dr. Takamine
- D. Dr. Miyoshi

2. *Why was Dr. Takamine in Washington?* A. He was touring the U.S.A.

- B. He was attending a scientific meeting
- C. He was a guest of President Taft's
- D. The passage doesn't say

3. *Where were the trees grown?*

- A. At Shizoukan
- B. In Tokyo
- C. In Yokohama
- D. Along the Arakawa River

4. *Who donated the cherry trees?*

- A. Dr. Takamine
- B. The city of Tokyo
- C. The Okitsu Horticultural Station
- D. Professor Kuwana

5. *How long after President Taft's election did the cherry trees reach Washington?*

- A. About a month
- B. About a year
- C. Two and a half years
- D. About three years

⁵⁰ Da *Cherry-Blossom Time in Washington*, in "English Teaching Forum", VIII, 2, Marzo-Aprile 1970, p. 31.

6. *What is a scion?*

- A. A branch
- B. A root
- C. A shoot
- D. A blossom

7. *As used in the passage, authority means a person:*

- A. belonging to a government office
- B. with much knowledge and experience
- C. having power of command
- D. making final decisions

8. *As used in the passage, what does to play a part mean?*

- A. To be active in order to obtain something
- B. To perform a rôle, like an actor in a play
- C. To let people believe something false
- D. To behave correctly, according to rules

Il brano, composto di 206 parole, permetterebbe di formulare altri quesiti, sia pure ad un più alto livello di difficoltà; ci sembrano tuttavia sufficienti gli otto quesiti proposti, sia per una verifica generale del contenuto del brano, sia per un controllo della comprensione di termini poco frequenti o usati in modo particolare.

6.5. Comprensione delle strutture

Oltre che sul lessico e sul significato globale di brani, i test di comprensione possono focalizzarsi sugli aspetti più tipicamente strutturali della lingua.

La linguistica ci insegna quanto sia difficile operare una distinzione rigida tra lessico e strutture morfosintattiche laddove ci si riferisca non ai singoli aspetti del linguaggio ma ai rapporti tra significante e significato. È la struttura che determina il valore e la funzione di una parola (ricordiamo quanto abbiamo detto sopra a proposito di *novel* e *beach*) ed è la struttura che evidenzia i referenti e la loro qualità di soggetto, predicato, oggetto, complemento, ecc. Nella frase “Gigi diventò il direttore dell'azienda” abbiamo lo stesso referente, ossia Gigi e il direttore sono la stessa persona; nella frase “Gigi invidiava il direttore dell'azienda” vi sono invece due referenti diversi.

N. Chomsky ha sviluppato questi temi, postulando una distinzione tra “struttura superficiale” e “struttura profonda”, cioè tra una struttura così come si manifesta sensibilmente nel linguaggio e le strutture di base che l'hanno generata.⁵¹

Le frasi ambigue, secondo la teoria chomskiana, sono quelle in cui ad una struttura superficiale possono corrispondere strutture profonde diverse. La frase “Anna andò ad abitare dalla zia perché era ammalata” si presta ad una duplice interpretazione, a seconda che, nella struttura profonda, si abbia “Anna era ammalata” oppure “la zia era ammalata”.

L'ambiguità può essere sfruttata a fini umoristici, come nella frase “Prima di dare il latte al bambino, fallo bollire”, a cui il solito spiritoso risponde: “Ma devo proprio bollirlo, povero bimbo?”

⁵¹ Per un approccio alla teoria chomskyana si veda N. CHOMSKY, *Filosofia del linguaggio*, Torino, Boringhieri, 1969.

equivocando su un possibile “bolli il bambino” nella struttura profonda della prima frase. Anche lo slogan, attribuito all'imbonitore di uno zoo viaggiante: “Più persone entrano, più bestie si vedono!” trae l'effetto umoristico dalla possibilità che i due sostantivi vengano attribuiti allo stesso referente anziché a referenti diversi.

Si notino, infine, le differenze tra coppie di frasi la cui struttura appare simile, ma che differiscono nella struttura profonda:

“Sono disposto a partire” (io voglio / io parto)

“Sono costretto a partire” (X vuole / io parto)

Queste premesse postulano la necessità di una verifica rivolta specificamente al controllo del valore delle strutture morfosintattiche attraverso quesiti appositamente predisposti, secondo i seguenti esempi:

Esempio 34 (Ing.). Scegliere, tra le alternative, quella che meglio descrive il senso globale della frase-stimolo.

Dick wants Lou to wash the car

- A. Dick must wash the car
- B. Lou must wash the car
- C. Dick and Lou must wash the car together

Questo quesito verifica la discriminazione tra le strutture / *want to do something* e *I want someone to do something*.

Esempio 35 (Fr.). Stessa tecnica del precedente.

Jean m'a dit quoi faire

- A. Je dois faire quelque chose
- B. Jean doit faire quelque chose
- C. Nous devons faire quelque chose

Confronta con la frase-stimolo: *Jean m'a dit: “Quoi faire?”*

Esempio 36 (Ing.). Come i precedenti.

Those shoes want mending

- A. Someone wants to mend those shoes
- B. Those shoes can be used for mending something else
- C. It is necessary to mend those shoes

CAP. 7 II SAPER SCRIVERE

7.1. Scrittura e ortografia

La quarta abilità fondamentale, il saper scrivere, non ha bisogno di essere analizzata dettagliatamente perché è quella che ha sempre ricevuto le maggiori attenzioni, fino ad eclissare le altre. Questo atteggiamento è responsabile anche del predominio di certe prove tradizionali, di cui ci siamo già occupati nel secondo capitolo, e che risentono, a loro volta, del fatto che si prende coscienza della lingua materna solo quando si impara a scrivere.

Non tutte le tecniche di verifica proposte in questo capitolo richiedono all'esaminato di scrivere nella lingua straniera; alcune di esse si basano infatti sulla lettura dei quesiti. Le abbiamo raggruppate non tanto per comodità di discorso, quanto piuttosto perché hanno in comune l'oggetto, che è l'aspetto attivo e produttivo della lingua scritta.

Il controllo dell'ortografia è bene che sia effettuato in tempi e modi diversi dal controllo della padronanza delle strutture e del lessico. Al paragrafo 2.4. abbiamo operato una distinzione tra i problemi dell'ortografia lessicale e quelli dell'ortografia grammaticale; a ciascuno di essi si collegano tipi diversi di quesiti specifici.

Per la verifica dell'ortografia lessicale, occorre per prima cosa isolare un singolo problema, dopo di che si predispongono delle frasi idonee a contestualizzare i vocaboli in esame. Nel seguente esempio, l'oggetto del quesito è la verifica delle grafie *augh* e *ough* corrispondenti al suono / ɔ: / in Inglese.

Esempio 37 (Ing.). Completare le seguenti frasi.

Professor M. Davis t..... t Latin here last year.

I th t she was Irish, but I was wrong.

I c t the first train in the morning.

He br t back the book I had lent him.

Julius Caesar f . . . t against the Britons.

Where have you b . . . t that gadget? At Selfridge's?

Si possono anche realizzare quesiti a scelta multipla:

Esempio 38 (Fr.). Scegliere la grafia corretta.

Les philantropes aiment tout le g e humain.

(enr) (èn) (endr) (eni)

A B C D

Le docteur reussit a r ... mer la femme évanouie.

(éani) (eani) (ani) (eà)

A B C D

Nella verifica dell'ortografia grammaticale, i problemi variano in misura notevole a seconda della lingua straniera considerata. Nella lingua inglese, i problemi ortografici relativi alle desinenze inflessionali sono relativamente pochi e riguardano soprattutto: la *y* finale; la *e* “muta” finale; il raddoppiamento della consonante finale; la desinenza *-es* invece di *-s*. Tali problemi si instaurano in un numero circoscritto di casi: il plurale dei nomi; la 3a persona singolare del presente, il passato, il

participio passato e la forma *-ing* dei verbi; il comparativo e superlativo degli aggettivi; la formazione di derivati mediante l'aggiunta di suffissi come *-ly*, *-ness*, *-dom*, *-able*, ecc.

Per il controllo, sono utili i quesiti a scelta multipla.

Esempio 39 (Ing.)

He is (studing) (studying) (studieing) the lesson.

A B C

She is (lying) (lieing) (liing) ill in her bed.

A B C

Ben diversa è la situazione nella lingua francese, in relazione al complesso meccanismo di concordanze del genere e del numero e alla coniugazione dei verbi. Come si può facilmente verificare, un test basato su quesiti a scelta multipla risulterebbe assai poco valido e produttivo. F. Manouvrier suggerisce una tecnica interessante, basata sull'impiego di frasi in cui alcune parole sono scritte in tutte maiuscole. Alcune di queste parole sono esatte, altre invece mancano di accenti, desinenze, ecc. All'esaminato si chiede di scrivere le correzioni necessarie, ovvero di apporre un segno convenzionale (per esempio, una croce) sopra le parole ritenute esatte.⁵²

Esempio 40 (Fr.)

Pour ALLE AU Pantheon, PRENE CE autobus.

Ils AIME les voitures FRANCAIS .

LE train est DEJA ARRIVE .

ecc.

Il test è oggettivo in quanto nella correzione si adotta il criterio “o tutto o niente”, ossia si assegna un punto per ciascuna soluzione esatta e nessun punto per le soluzioni omesse, parziali, errate o comunque imperfette, di qualunque tipo sia l'errore commesso.

Per una valutazione sufficientemente accurata, occorre predisporre una batteria di quesiti tale da permettere la verifica estensiva di tutti i problemi ortografici presi in considerazione.

7.2. Verifica della produzione di strutture

Per il controllo oggettivo della padronanza delle strutture morfosintattiche sono state elaborate numerose tecniche, anche assai diverse tra loro. In questo settore il LT ha registrato gli sviluppi più notevoli, con dovizia di sperimentazioni e di risultati.

Una tecnica frequentemente adottata richiede allo studente di eseguire per iscritto alcuni esercizi di trasformazione, analoghi a quelli usati per la fissazione delle strutture o per la verifica del saper parlare (vedi paragrafo 5.4.). Nell'applicazione pratica, tuttavia, questo test perde di oggettività nella misura in cui l'esaminatore non si limita a controllare la correttezza delle strutture, ma permette che il giudizio venga influenzato da rilievi non pertinenti (ad esempio, sull'ortografia o sul lessico). Sappiamo infatti che la considerazione contemporanea di più elementi non correlati impedisce una precisa quantificazione del profitto, pregiudicando l'affidabilità del test. Si consiglia perciò di verificare solo la struttura, assegnando un punto per ogni frase corretta dal punto di vista morfosintattico, e nessun punto per le frasi errate dallo stesso punto di vista. Eventualmente, se il test lo consente, si possono dare valutazioni *separate* anche in relazione all'ortografia, al lessico, ecc., ma ciò deve essere evitato nella maggioranza dei casi.

⁵² 1 F. MANOUVRIER, *Tests d'orthographe grammaticale*, in “Revue Belge de Psychologie et de Pédagogie”, n. 119, 1967.

Il Lado ⁵³ ha proposto un impiego della traduzione come test che ci sembra assai illuminante ai fini della comprensione della natura delle tecniche oggettive.

Esempio 41 (Ing.). Tradurre:

Vorrei sapere quando verrà Dan.

I'd like to know when.....

Così facendo si isola un singolo problema specifico (nell'esempio, l'ordine soggetto-verbo nella frase interrogativa indiretta) e si accettano *tutte* le soluzioni che dimostrino che lo studente conosce la costruzione esatta, anche se coesistono errori ortografici, lessicali o addirittura morfosintattici, diversi però da quello che noi verifichiamo. Sulla base di questo principio, le soluzioni valide non saranno solo *Dan comes, Dan is coming o Dan will come*, ma anche *Dan come, Dan came o Dan will go*; soluzioni errate sono invece tutte quelle che presentano un ordine diverso (verbo-soggetto, ausiliare-soggetto verbo, ecc.): *comes Dan, will Dan come*.

Allo stesso risultato si giunge peraltro, in modo più rapido, diretto, e meno soggetto a dubbi di valutazione, attraverso un quesito a scelta multipla:

Esempio 42 (Ing.).

I'd like to know when

(will Dan come) (will come Dan) (Dan will come)

A

B

C

Le domande a scelta multipla sono spesso presentate in modo leggermente diverso:

Esempio 43 (Fr.).

Il faut que vous avec nous.

A. venez

C. viendrez

B. veniez

D. viendriez

Per la verifica dei costrutti sintattici, si può ricorrere anche alle frasi rimescolate:

Esempio 44 (Ing.).

When . . . ?

A. parents

B. are

C. your

D. coming

Indicare l'ordine di successione dei vocaboli.

Esistono due modi per segnare la risposta. Il primo consiste nello scrivere la frase nello spazio dopo "When", ovvero la sequenza esatta BCAD; il secondo, controllabile mediante griglia trasparente, richiede che il foglio allievo sia così predisposto:

1a parola: A B C D 2a parola: A B C D 3a parola: A B C D 4a parola: A B C D

Indipendentemente dal modo di segnare le risposte, questo tipo di test è criticato⁵⁴ perché è difficile trovare quesiti che presentino il livello di difficoltà desiderato; i problemi tendono ad essere troppo facili e ad assomigliare ai puzzle. Con frasi più complesse si rischia facilmente di ottenere

⁵³ R. LADO, *Language Testing*, cit., pag. 266.

⁵⁴ D. P. HARRIS, *Testing English as a Foreign Language*, cit., p. 28.

quesiti con più soluzioni, specialmente in lingue come l'Inglese in cui molti sostantivi e verbi hanno forme comuni.⁵⁵

Per superare questo inconveniente, è stato proposto di associare le frasi rimescolate alla scelta multipla, così:

Esempio 45 (Fr.). Scegliere la sequenza esatta.

Rugby - France - est - populaire - le - très - en.

A. 7 5 2 3 6 4 1 B. 5 1 3 6 4 7 2

⁵⁵ Un esempio classico in Inglese: "*Time flies*". "*I can't: they go by too fast*". La risposta rivela che **la** battuta iniziale ("Il tempo vola") è stata reinterpretata assumendo *time* come verbo e *flies* come sostantivo ("Cronometra le mosche").

C. 1 3 6 4 7 5 2 D. 5 1 6 4 3 7 2

Si tratta, come appare evidente, di una tecnica macchinosa, complessa e di aspetto sgradevole, e quindi di impiego limitato; la verifica dei costrutti sintattici può essere condotta con le altre tecniche, tutte molto più efficienti delle frasi rimescolate. Queste possono essere utili, specialmente a livello elementare, come esercizio, anche perché ai ragazzi piace misurarsi con quiz di questo tipo (che infatti, più che abilità linguistica, richiedono intelligenza verbale);

la loro utilità come test è invece assai ridotta. Ne abbiamo parlato diffusamente sia perché molti testi presentano prove di questo tipo, sia perché è importante conoscere il maggior numero possibile di soluzioni proposte nel settore del testing, sia, soprattutto, perché attraverso la critica delle tecniche si possono definire e puntualizzare i requisiti dei test validi.

Il seguente esempio presenta dei quesiti analoghi a quello dell'Esempio 43, ma realizzati mediante il completamento anziché la scelta multipla:

Esempio 46 (Fr.). Completare con la voce verbale richiesta:

1. Je ce film hier soir. (voir)
2. Quand maman arriva, Thierry déjà son devoir. (finir) ecc.

Nell'ambito delle singole frasi è relativamente facile costruire quesiti appropriati; è invece meno facile costruire test per verificare la capacità di scrivere paragrafi, brani o lettere. I test oggettivi, per loro natura, non possono andare al di là di certi limiti nel valutare le capacità creative nella lingua straniera; la situazione è analoga a quella che abbiamo già delineata a proposito del saper parlare. Esistono tuttavia varie prove che danno indicazioni utili sulla capacità dell'esaminato di coordinare frasi e periodi, e di scegliere elementi lessicali e morfosintattici in relazione a contesti più ampi.

7.3. Verifica della coordinazione di strutture

È relativamente semplice verificare il possesso della capacità di produrre periodi articolati e complessi dato che, in misura notevole, ciò equivale a saper usare i connettivi: congiunzioni coordinanti e subordinanti, pronomi relativi, correlativi, ecc.

Esempio 47 (Ing.). Unire ogni coppia di frasi usando i connettivi indicati.

1. Mr. Brown is an engineer.
He lives next door.
Who
2. He played in the garden.
He didn't write the exercise.
Instead of
- 3, She didn't feel well.
She went to the cinema all the same.
Although

Analogo procedimento può servire ad inserire frasi in periodi complessi, rendendo sempre più difficili i quesiti, ma permettendo di giungere alla composizione di strutture altamente articolate, ricche di coordinazioni, subordinazioni, frasi parentetiche, e così via.

Un altro approccio al problema è stato realizzato mediante la verifica della punteggiatura, in quanto questa evidenza gli elementi costitutivi del periodo e le loro connessioni.

Esempio 48 (Fr.). Scegliere, tra le alternative offerte, l'interpunzione adeguata. Il simbolo *N* indica nessuna punteggiatura.

1. Avant de fermer la porte ₁ il aperçut deux hommes ₂ qui étaient dans la chambre ₃

1. A , 2. A : 3. A ?

1. B : 2. B N 3. B .

1. C ; 2. C , 3. C ,

2. L'homme ₄ qui était a gauche de le fenêtre ₅ dit ₆ bonjour ₇ M. Bertrand ₈

4. A , 5. A , 6. A , 7. A , 8. A !

4. B N 5. B N 6. B : “ 7. B ”, 8. B .”

4. C ; 5. C ; 6. C ; 7. C N 8. C ?”

Quesiti di questo tipo richiedono particolare attenzione per evitare ambiguità di interpretazione e frasi sibilline.⁵⁶ Una seconda fonte di ambiguità può essere costituita da usi regionali diversi. Ecco come la stessa frase viene punteggiata, di solito, in Gran Bretagna e negli stati Uniti;

'I'm back at 6.45', Mr Brown said.

*“I'm back at 6:45,” Mr. Brown said.*⁵⁷

Per verificare la capacità dello studente di organizzare sequenzialmente del materiale sono stati sviluppati quesiti del tipo seguente:

Esempio 49 (Ing.). Disporre le frasi nel loro ordine logico.

A. In 1954 he was appointed as a minister of the Baptist Community of Montgomery, Ala.

B. Therefore he organized a boycott of public transportation that lasted over one year.

C. In those days racial prejudice was even stronger than today, and Negroes were segregated on buses, too.

D. Martin Luther King, Jr. was the most famous leader who fought for the rights of American black people.

1a frase: A B C D

2a frase: A B C D

3a frase: A B C D

⁵⁶ Come appare dall'esempio riportato alla nota precedente, non tutte le frasi ambigue sono famose come “Ibis redibis non morieris in bello”.

⁵⁷ Occorre sempre verificare se il quesito presenti diverse soluzioni a seconda delle varianti regionali; in caso affermativo, esso potrà essere impiegato solo se si abbia la certezza che tutti gli studenti conoscono una sola variante. Nella maggior parte dei casi, tuttavia, e specialmente con studenti non principianti, è meglio evitare di correre rischi, modificando opportunamente i quesiti. Il problema assume una particolare rilevanza nei riguardi delle varietà britannica ed americana della lingua inglese, che presentano differenze a vari livelli: fonologico (/gra:s/ vs. /græs/), ortografico (plough - plow; travelied - traveled), morfosintattico (he hasn't - he doesn't have; help me to do this - help me do this), lessicale (lift - elevator), ecc.

4a frase: A B C D

Questa tecnica è solo apparentemente analoga a quella delle frasi rimescolate. Per risolvere il quesito, lo studente deve dimostrare di riconoscere quegli elementi che segnalano la connessione tra le frasi. Solo nella frase D il soggetto è espresso da un nome, mentre nelle altre abbiamo il pronome *he*; essa perciò deve essere la prima. Nella frase A troviamo una data, e nella frase C c'è la espressione *in those days*, che a quella data si ricollega, quindi A precede C. Gli eventi descritti nelle frasi C e B sono legati dal rapporto di causa ed effetto, evidenziato da *therefore*, perciò B segue C e la sequenza completa non può essere che DACB.

7.4. Ipotesi di verifica degli elementi stilistici

La difficoltà primaria da superare per la realizzazione di test oggettivi riguardanti lo stile è la divergenza di opinioni su cosa debba considerarsi “buono stile”, e addirittura su che cosa debba intendersi per “stile” *tout court*. Nessuna prova può dirsi oggettiva se la risposta esatta non può essere individuata in modo inequivocabile e inconfutabile. Quesiti che chiedessero genericamente di scegliere, tra le varie formulazioni di un concetto, la “frase migliore”, resterebbero sul piano dell'opinabile e dell'aleatorio. In un raggio d'azione più ristretto, tuttavia, sembra possibile preparare materiale utile.

Se, per ipotesi, agli studenti è stato insegnato che un periodo composto da due frasi coordinate è meglio costruito rispetto a due frasi principali giustapposte, e che un periodo con frasi subordinate è meglio articolato rispetto a un periodo composto di frasi coordinate, quegli studenti sono in possesso del criterio di giudizio necessario per rispondere al seguente quesito.

Esempio 50 (Fr.). Quale periodo è meglio articolato?

- A. Pierre étudia sa leçon. Plus tard, il écrit une lettre.
- B. Après d'avoir étudié sa leçon, Pierre écrit une lettre.
- C. Pierre d'abord étudia sa leçon, puis il écrit une lettre.

Date le premesse, appare chiaro che la risposta richiesta è la **B**, in quanto il periodo complesso è ritenuto meglio articolato del periodo composto **C**, e questo a sua volta è migliore delle frasi principali giustapposte **A**.

Nel caso di microlingue specialistiche,⁵⁸ questa tecnica è invece suscettibile di interessanti sviluppi; l'esempio seguente presenta una sua possibile applicazione alla corrispondenza commerciale inglese.

Esempio 51 (Ing.). Lo stile commerciale deve essere cortese, chiaro e conciso; quale delle tre frasi rispetta tali esigenze?

- A. We regret to say that in case we do not receive your remittance by return we shall have to pass the matter to our solicitors.
- B. Send us the money by return, or else we will take legal steps and get you into trouble.
- C. Any further failure in your remitting the outstanding balance will compel us to consider entrusting our solicitors with the whole matter, to which unpleasant action may ensue.

⁵⁸ Per *microlingue specialistiche* intendiamo (in contrapposizione a *lingua standard o macrolingua*) i linguaggi propri di settori particolari di attività (scienza, tecnica, commercio, sport, ecc.). Tali linguaggi si differenziano rispetto alla macrolingua non solo sul piano lessicale, ossia nell'uso di vocaboli o espressioni particolari, ma anche su quello delle strutture morfosintattiche.

Ogni studente sufficientemente addestrato riconosce facilmente A come frase migliore, dato che la frase B non è affatto cortese e la frase C, prolissa, involuta e di tono arcaico, manca di chiarezza e di concisione.

Riassumendo, possiamo concludere affermando che gli indubitabili limiti delle prove oggettive non impediscono di affrontare tematiche che sembrerebbero esulare dalla possibilità di una verifica accurata, purché venga svolto un adeguato lavoro preliminare di chiarificazione degli obiettivi e di determinazione dei parametri di giudizio.

CAP. 8 LA COMPrensIONE INTERCULTURALE

8.1. Confronto di civiltà

Come già accennato al paragrafo 1.3, i criteri di analisi contrastiva elaborati da R. Lado⁵⁹ offrono le basi per la realizzazione di test oggettivi intesi a verificare l'esatta comprensione del significato dei comportamenti tipici della cultura straniera.⁶⁰

Sinteticamente, i presupposti sono i seguenti:

1. L'analisi dei fatti di civiltà si effettua in modo analogo all'analisi dei fatti linguistici, secondo tre parametri essenziali: forma, distribuzione e significato.
2. L'interferenza si manifesta sotto forma di giudizi negativi sui comportamenti degli appartenenti alla C2, a causa della mancata comprensione del loro esatto significato.

Per chiarire il primo punto, occorre anzitutto spiegare che cosa si intenda per “fatti di civiltà”; abbiamo volutamente usato un termine generico, che abbraccia i “modi di vivere”, il “patrimonio culturale”, il “sistema dei valori e degli ideali”, e così via. Nella accezione antropologica, le *civiltà* (o *culture*) sono sistemi strutturati di modelli di comportamento. Ogni gruppo sociale analizza i fatti, i singoli elementi del comportamento individuale, e li classifica secondo determinati schemi. Malgrado non esistano due manifestazioni identiche dello stesso atto, certe attività vengono percepite come “uguali” ad altre. Passare una settimana in una tenda da campeggio, o un mese in un albergo di lusso, possono essere considerati due aspetti dello stesso comportamento, “andare in villeggiatura”. Pertanto la *forma* di un modello di comportamento può avere estrinsecazioni oggettivamente diverse, non sempre facilmente definibili e classificabili; tuttavia gli appartenenti ad un gruppo culturale nel cui sistema è operante un certo modello di comportamento sono in grado di riconoscerlo sotto qualsiasi forma. Pur essendovi infiniti modi di andare in villeggiatura, ognuno di noi è in grado di stabilire se questo o quel vicino di casa sia o non sia andato in villeggiatura, semplicemente per avere osservato (o sentito descrivere) le loro azioni.

Tutte le manifestazioni dei modelli di comportamento hanno una tipica *distribuzione* nel tempo, nello spazio e in rapporto ad altri comportamenti significativi. La villeggiatura si effettua in alcuni caratteristici periodi dell'anno, in determinate località e in relazione agli impegni di lavoro o di studio. Altri fatti hanno una distribuzione nel tempo basata su cicli giornalieri, settimanali o mensili, anziché annuali.

Ad ogni modello di comportamento viene attribuito un *significato* primario, spesso accompagnato da una serie di significati secondari. La villeggiatura è primariamente “un periodo di vacanza trascorso in una località diversa da quella in cui si risiede abitualmente”; secondariamente, le forme in cui essa si attua ci danno indicazioni sullo *status* socio-economico, sui gusti, sullo stato di salute, ecc. dei villeggianti.

Il secondo punto, relativo all'interferenza della CI sulla C2, è essenziale per la comprensione dei meccanismi attraverso cui si instaura l'incomprensione interculturale. Osserva il Lado:

⁵⁹ R. LADO, *Linguistics across Cultures*, cit., pp. 110-123. Sui criteri generali di analisi contrastiva vedi anche U. WEINREICH, *Languages in Contact*, L'Aia, Mouton, 1970, e G. NICKEL, (ed.), *Papers in contrastive linguistics*, Cambridge, Cambridge University Press, 1971.

⁶⁰ Qui ci limitiamo a prendere in considerazione i rapporti tra la cultura propria del parlante nativo (indicata con CI) e la cultura straniera (C2). Per un esame più approfondito dei rapporti tra due lingue e culture si veda di G. FREDDI, *Bilinguismo e biculturalismo* in “Quaderni per la promozione del bilinguismo”. Brescia, CLADiL, I, n. 1, 1973.

“Quando i visitatori stranieri, provenienti da zone in cui il caffè viene servito nero e molto forte, assaggiano il caffè americano, non dicono che è diverso; dicono che il caffè americano è cattivo. Similmente, quando gli americani vanno all'estero in paesi in cui il caffè è nero e forte, assaggiano il caffè e non dicono che è diverso; anch'essi dicono che è cattivo”.⁶¹

Come il sistema linguistico, così anche il sistema culturale è costituito da una serie di abiti, per lo più inconsci, e comunque profondamente radicati nella personalità di ogni individuo. **Da** ciò la difficoltà di descrivere oggettivamente i propri comportamenti e quelli del gruppo sociale cui si appartiene, e l'istintivo rifiuto di comportamenti diversi, appartenenti ad altri sistemi e che pertanto rischiano di sconvolgere l'equilibrio esistente e costringono a ricercare nuovi equilibri.

Il possesso del sistema della CI si realizza attraverso un lunghissimo processo di apprendimento, nel quale non solo si acquisiscono i modelli di comportamento, ma si attribuiscono loro anche particolari valori. Tagliare col coltello un pezzo di cibo e portarlo alla bocca con la forchetta, oltre a essere un modo pratico per ingerire le vivande, è anche e soprattutto il comportamento che distingue la persona “perbene”, “bene educata”, “che sa stare a tavola”, “che non fa brutta figura in società”. Lo stesso non si può dire di chi ficca in bocca le dita o il coltello, e nemmeno di chi taglia tutto il cibo a pezzetti, posa il coltello, quindi passa la forchetta nella mano destra e mangia, di seguito, tutti i pezzetti già tagliati: è “un comportamento da maleducati”, secondo il galateo. Poiché così si insegnano le “buone maniere”, chi apprende a comportarsi in un certo modo è indotto a pensare che esistano modelli universali di comportamento, e quindi a respingere forme diverse; per ovvi motivi, di solito non si insegna che tagliare un pezzo per volta è buona educazione *in Italia*, mentre tagliare tutti i pezzetti prima è buona educazione *in altri Paesi* (come gli Stati Uniti).

La presunzione di possedere schemi e modelli aventi valore assoluto e universale si associa spesso alla presenza di preconcetti, false informazioni e stereotipi sulla C2. L'ignoranza della realtà e la generalizzazione di qualche aspetto particolare hanno creato una serie di miti che si tramandano di generazione in generazione: l'avarizia degli scozzesi, dei genovesi e degli ebrei; il militarismo tedesco; la flemma britannica; il fatalismo degli arabi; la musicomania e l'ingordigia di spaghetti degli italiani, la megalomania americana, e così via. Spinti all'estremo, questi preconcetti sono alla radice di fenomeni gravissimi e irrazionali, quali il razzismo, e dei pregiudizi contro le minoranze etniche e religiose. A livello più superficiale, essi danno luogo ad atteggiamenti di ripulsa, disagio, sospetto, scherno, o ad altre manifestazioni che evidenziano sintomi di mancata o errata comprensione del sistema di modelli e di valori operante nella C2.

Un programma completo di lingua-civiltà straniera non deve perciò limitarsi a presentare alcuni aspetti tipici della “way of life” e del folklore del paese straniero: esso fallirebbe il suo scopo se non mirasse a far comprendere nel loro giusto valore i modelli di comportamento inseriti nel sistema della C2.

8.2. Il relativismo culturale

Un'analisi contrastiva di due civiltà è un compito formidabile per la vastità dell'oggetto, per la complessità dell'indagine, per il pericolo di affermazioni soggettive ed arbitrarie e per la difficoltà di individuare esattamente le interferenze. Tornando agli esempi citati, nel caso del caffè e delle buone maniere a tavola si tratta di interferenze derivanti da forme diverse. In altre situazioni, si hanno forme uguali e significati diversi: fischiare durante uno spettacolo è segno di disapprovazione nei paesi latini, mentre è segno di gradimento negli U.S.A.

Significati diversi possono dar luogo a distribuzioni diverse: la stretta di mano in Italia è una forma di saluto, mentre in Inghilterra ricorre solo in occasioni speciali (nelle presentazioni, ovvero per congratulazioni o condoglianze); di conseguenza, essa è assai più frequente da noi che da loro. I fraintendimenti sorgono da questa differenza di distribuzione: un italiano, notando che l'inglese lo

⁶¹ R. LADO, *Linguistics across Cultures*, cit., p. 119.

saluta senza porgere la mano, crede di cogliere un atteggiamento di freddezza, distacco o superbia; a sua volta l'inglese, di fronte alle ripetute offerte di stretta di mano, è indotto a cogliere una forma di eccessiva invadenza o uno stato di ipereccitabilità emotiva, dato che nella sua cultura (presentazioni a parte) lo stringersi la mano è associato a situazioni ricche di connotazioni emotive.

Mancano, a tutt'oggi, adeguate analisi contrastive a livello culturale, ed è prevedibile che esse non saranno disponibili se non fra molto tempo. Ciò tuttavia non deve scoraggiare gli insegnanti di lingue: un confronto organico ed esauriente tra due culture è opera altamente specialistica, e come tale esula dai normali obiettivi di un corso di lingua-civiltà straniera. A livello di scuola preuniversitaria può bastare l'analisi degli elementi di civiltà presentati dal materiale didattico impiegato, ovvero accostati attraverso le esperienze personali dell'insegnante e degli studenti. È sufficiente esaminare criticamente un numero ristretto di situazioni per:

- evidenziare le dissimmetrie fra i vari sistemi di valori culturali;
- studiare l'insorgere delle interferenze e delle distorsioni di giudizio;
- avviare alla nozione di *relativismo culturale*,⁶² una delle più valide conquiste nel campo delle scienze umane, secondo cui ogni cultura rappresenta un modo diverso ed originale di organizzare la vita comunitaria di un gruppo.

Gli studenti che avranno capito ciò saranno sempre meno facilmente vittime dei pregiudizi e degli stereotipi, e non saranno indotti a cercare di stabilire impossibili graduatorie tra culture più o meno “civili”, evolute, organizzate, buone, belle, “umane”, e così via.

8.3. L'approccio linguistico

Lo studio di una lingua permette di cogliere direttamente molti degli aspetti della civiltà che in quella lingua si esprime. Parallelamente, solo attraverso la conoscenza di elementi culturali si possono apprezzare e interpretare correttamente certe espressioni della lingua straniera.

Ad esempio, in italiano, francese e spagnolo abbiamo coppie di parole come *gamba/zampa*, *jambe/patte*, *pierna/pata*; in inglese, invece, un solo vocabolo, *leg*, è usato per denotare sia gli arti inferiori dell'uomo, sia gli arti degli animali. Questa dissimmetria, secondo il Lado,⁶³ è indice di un modo diverso di considerare gli animali nella cultura spagnola e in quella anglosassone: a sua volta, la difformità di vedute fa sì che gli anglosassoni non comprendano esattamente il significato della “*corrida*”. Mentre nella cultura spagnola la *corrida* è simbolo della lotta tra la forza bruta animale e l'intelligenza umana (con l'inevitabile trionfo di quest'ultima), in culture diverse, e in particolare in quelle che tendono ad umanizzare gli animali, l'impressione globale che se ne ricava è di una lotta impari, di un macello elaborato e sofisticato, spettacolare ma sostanzialmente crudele.

Espressioni come *grandeur*, *force de frappe*, *gauchiste* e *franglais* riportano alla mente, in modo immediato, alcuni aspetti della Quinta Repubblica francese; a loro volta, esse non sono comprensibili senza il riferimento agli ideali, alle aspirazioni ed ai problemi del periodo. A volte, la comprensione di frasi idiomatiche o di altre espressioni richiede nozioni assai approfondite ed estese sulla C2. Così, per capire pienamente il significato del nome di un noto complesso musicale, i “Rolling Stones”, occorre conoscere il proverbio inglese *A rolling stone gathers no moss* (“Pietra che rotola non raccoglie muschio”), che si riferisce a coloro i quali, passando da una attività all'altra senza soffermarsi su qualcosa di costruttivo, non sono capaci di mettere a frutto le proprie esperienze e di costruirsi una posizione solida. Il proverbio evoca l'immagine dei terreni franosi e

⁶² Su questa tematica si veda, oltre al contributo già citato, di G. FREDDI, *Fondamenti e metodi della didattica delle lingue*, cit., pp. 89-95.

⁶³ R. LADO, *Linguistics across Cultures*, cit., pp. 115-117.

instabili, sui cui non cresce alcuna vegetazione. “Rolling stone”, quindi, non è semplicemente una “pietra rotolante”, ma anche un giovane alla ricerca del proprio inserimento nella società.

Ancor più complesso è il caso della frase *A land flowing with milk and money* (“Una terra in cui scorrono il visone e il danaro”) usata a volte come definizione degli U.S.A. La frase è una parafrasi dell'espressione biblica *A land flowing with milk and honey* (“Una terra in cui scorrono il latte e il miele”; Esodo, 3,8) con cui si descrive la Terra Promessa. Dal punto di vista linguistico, la parafrasi si basa su due cambi di consonante, per cui *milk and honey* diventa *mink and money*. A livello culturale, il discorso si allarga fino a considerare il ruolo esercitato dalla Bibbia nella civiltà anglosassone, e come il testo sacro sia ben conosciuto da vasti strati della popolazione.⁶⁴ Per quanto riguarda specificamente gli Stati Uniti, ci si può ricollegare all'influenza dei Puritani sugli sviluppi del sentimento religioso oltre oceano. Sulla scorta di queste considerazioni si giunge a percepire la connotazione umoristica della frase; il suo effetto ironico deriva, strumentalmente, dal gioco di parole e, concettualmente, dal dislivello tra l'ideale biblico e la realtà consumistica. Come si vede, una frase che alla prima traduzione poteva apparire povera di significato, analizzata adeguatamente è ricca di spunti per affrontare temi di notevole rilievo culturale.

8.4. Tecniche di controllo della comprensione interculturale

Anche se poco noti in Italia (motivo per cui ci siamo soffermati su di essi piuttosto a lungo) i principi di analisi contrastiva a livello interculturale appaiono abbastanza ben definiti. Non si può invece dire altrettanto delle tecniche proposte per la verifica della comprensione della C2.

Il Catford⁶⁵ propone un test basato sulle allusioni:

Esempio 52 (Ing.). Rispondere alle domande.

1. *There's one born every minute.* Question: One what?
2. *A land flowing with milk and money.* Question: What phrase is alluded to?

La risposta alla prima domanda è *sucker*: si allude al detto attribuito a Phineas T. Barnum (impresario dei circhi omonimi) *There's a sucker born every minute* (“Ogni minuto nasce un gonzo”); della seconda frase ci siamo già occupati nel paragrafo precedente.

Si tratta di un tipo di test con un indice di difficoltà presumibilmente assai alto e con quesiti ai margini dell'oggettività, soprattutto per il modo in cui devono essere formulate le risposte. Riteniamo che questa tecnica debba essere ulteriormente perfezionata per non rischiare di produrre prove ambigue.

Il Lado⁶⁶ preferisce orientarsi verso i quesiti a scelta multipla e presenta alcuni esempi basati sui possibili fraintendimenti dovuti a interferenze tra la cultura spagnola e quella degli Stati Uniti. Ne riportiamo due (in versione italiana) riferiti alla civiltà spagnola.

Esempio 53 (Spagn.)

Si sta svolgendo un incontro di calcio in Spagna. C'è un'azione molto rapida davanti a una delle porte. La maggior parte degli spettatori si alza in piedi e FISCHIA FORTE. Essi stanno:

- 1) dimostrando entusiasmo per l'azione di gioco.

⁶⁴ Può essere sintomatico il fatto che un umorista di successo, G. MIKES abbia intitolato *Milk and Honey* il suo libro su Israele (Londra, Ed. Deutsch).

⁶⁵ J. C. CATFORD, *Some Aspects of Linguistics in Language Testing*, in “English Teaching Forum”, v. VI, n. 5, 1968, p. 19.

⁶⁶ R. LADO, *Language Testing*, cit., pp. 286-8. L'intera parte IV di questo volume è dedicata ai problemi della verifica della comprensione interculturale e dei valori culturali.

- 2) dimostrando delusione e disapprovazione.
- 3) dimostrando sollievo per uno scampato pericolo.
- 4) chiedendo che l'azione sia ripetuta.
- 5) segnalando la fine dell'incontro.

Esempio 54 (Spagn.)

Era nuovamente primavera a Madrid, il 21 marzo, ed un sole allegro filtrava attraverso la finestra. Il grande orologio della sala da pranzo battè la mezz'ora. “Le due e trenta”, disse il babbo, “e il cibo, dall'aspetto, sembra buono”.

L'ora del pasto:

- 1) è l'ora normale di ogni giorno.
- 2) è l'ora delle occasioni speciali.
- 3) non è spagnola.

La risposta esatta al quesito dell'esempio n. 53 è la seconda. Si noti come questo quesito non presenti alcuna difficoltà per il lettore italiano mentre, per i motivi già visti, uno studente americano potrebbe essere facilmente indotto a scegliere la risposta n. 1; le alternative 3, 4, e 5 ipotizzano una scarsa conoscenza del gioco, ma sono plausibili nel contesto. Il quesito dell'esempio n. 54 potrebbe invece essere valido anche per gli studenti italiani, specialmente in quelle regioni in cui le 14.30 non sono l'ora abituale della seconda colazione; la risposta esatta infatti è la n. 1.

La funzione essenziale degli esempi riportati è di chiarire come la verifica debba avere per oggetto non aspetti qualsiasi della C2, bensì solo quelli che subiscono interferenze da parte della CI a causa delle dissimmetrie tra i due sistemi.

Sia il Catford che il Lado, a conclusione dei passi citati, riconoscono che il problema della verifica della comprensione interculturale deve essere ulteriormente esplorato; lo stato attuale degli studi lascia scorgere ampie possibilità di sviluppo attraverso una ricerca su vasta scala e sperimentazioni sistematiche condotte da esperti e insegnanti.

PARTE III ASPETTI TECNICI E STATISTICI

CAP. 9 LA REDAZIONE DI PROVE OGGETTIVE

9.1. Formulazione dei quesiti

Più volte, nel corso dei precedenti capitoli, abbiamo accennato alla necessità di organizzare i quesiti nell'ambito di test opportunamente strutturati. Vediamo ora quale possa essere un itinerario-tipo che conduca alla redazione ed alla sperimentazione efficace di un test.

Le principali fasi della costruzione di un test sono:

- 1) La definizione esatta del tipo di prova che si intende realizzare: oggetto, scopi, metodi, tecnica, livello, ecc.
- 2) L'inventario del lessico, delle strutture e degli elementi di civiltà che si presuppongono noti agli allievi e sui quali deve essere condotta la verifica.
- 3) La redazione di un numero adeguato di quesiti ed il loro raggruppamento in subtest corrispondenti alle varie sezioni della prova (pronuncia, morfosintassi, ecc.).
- 4) La somministrazione del test ad un congruo numero di allievi per la raccolta di dati sulla validità dei quesiti.
- 5) L'elaborazione statistica dei dati ottenuti.
- 6) La revisione dei quesiti poco validi (nel qual caso occorre un secondo *pre-testing*) o la loro eliminazione.
- 7) La somministrazione della forma finale e l'elaborazione dei punteggi.

La prima fase è fondamentale perché dall'accuratezza delle premesse dipenderà la validità di tutto il lavoro successivo. Prima di procedere a qualunque altra operazione, il redattore di un test deve chiarire a se stesso gli obiettivi che si prefigge ed il modo in cui tende a raggiungerli. Si consiglia di indicare i punti essenziali per iscritto, anche se schematicamente, ad esempio così:

“Inglese, II media. Passato dei verbi regolari ed irregolari: difficoltà ortografiche. Quesiti a scelta tripla (almeno 25). Lessico e strutture: fino al cap. XV del testo”.

In questo caso, non si pongono problemi particolari per la definizione del lessico e delle strutture interessati, in quanto si tratta di un *test specifico*, nel senso indicato al par. 3.5. Sarà sufficiente controllare che effettivamente i quesiti operino solo nell'ambito prestabilito.

È piuttosto facile incorrere nell'errore di non rispettare i limiti del patrimonio linguistico dell'allievo; tale errore lo si riscontra, solitamente, non nel punto focale del test (su cui il redattore concentra la propria attenzione) ma in quelli apparentemente marginali: nelle frasi che servono a contestualizzare il quesito, o nelle alternative errate. Naturalmente la presenza di elementi estranei *in qualunque punto del quesito* è sufficiente per bloccare l'esaminato o per condurlo fuori strada. In questa fase di redazione la massima cura dovrà essere prestata per evitare gli altri errori di redazione dei quesiti, di cui ci siamo già ampiamente occupati, e che pertanto ci limitiamo a richiamare:

- formulazione di quesiti non funzionali al tipo di verifica desiderato;
- formulazione di quesiti poco chiari, ambigui, o che ammettono più risposte;
- ingiustificata concentrazione di difficoltà di tipo diverso in un solo quesito, ovvero insufficiente definizione del contesto.

Lavorando con molta attenzione e criticando rigorosamente ciascun quesito è possibile evitare sin dall'inizio gli errori di formulazione; è tuttavia consigliabile predisporre più quesiti di quanti ne

occorreranno nella redazione definitiva del test, in modo da poter eliminare quelli che, al controllo, risulteranno meno validi.

Se il test si compone di quesiti diversi tra loro per l'oggetto della verifica e/o per la tecnica impiegata (quesiti sulla pronuncia, sul lessico, sulla morfosintassi; domande a scelta multipla o completamento, ecc.) è buona norma raggruppare i diversi tipi di quesiti in altrettanti *subtest*, che costituiranno le varie parti in cui si articolerà il test. Ogni *subtest* inizia fornendo agli allievi le istruzioni per lo svolgimento della prova.

Salvo che particolari motivi di ordine metodologico inerenti, ad esempio, al contenuto del test, non lo sconsiglino, nell'ambito di ogni *subtest* (o dell'intero test, se la sua omogeneità è tale da non richiedere suddivisioni) i quesiti saranno riordinati a seconda della loro difficoltà, iniziando con i più facili. In un primo momento l'insegnante valuterà le difficoltà soggettivamente, sulla base della propria esperienza; in seguito al pre-testing, vi sarà un riordinamento che tiene conto dell'indice di facilità.⁶⁷

In questa fase di assemblaggio del test si farà attenzione a che le risposte dei quesiti a scelta multipla si susseguano in modo assolutamente casuale e imprevedibile, come BCAABCACC... onde evitare schemi ciclici (ABCABCABC...; AACCBBAACCB...; ABCBA-ABCBA...) sui quali potrebbe basarsi qualche allievo. È anche opportuno che le alternative compaiano con frequenza pressoché uguale: su 25 quesiti a scelta tripla, la distribuzione delle alternative sarà 8+8+9, 7+9+9 o anche 7+8+10; altre distribuzioni sarebbero eccessivamente sbilanciate, con la netta preponderanza di una delle tre alternative rispetto alle altre. Tutto ciò ha ovviamente lo scopo di ridurre quanto più possibile l'eventualità che l'allievo individui la risposta esatta lavorando a caso o basandosi su fattori estranei al contenuto linguistico della prova.

9.2. Scelte multiple e risposte date a caso

È altamente improbabile indovinare tutte le risposte grazie ad una combinazione fortunata, anche con quesiti a due sole alternative come quelli del tipo *uguale/diverso*, *vero/falso*, *giusto/sbagliato*, ecc.

In un test comprendente 50 di tali quesiti, si ha una sola probabilità su oltre un milione di miliardi.⁶⁸ Agli appassionati del Totocalcio o di concorsi analoghi è ben nota la difficoltà di azzeccare tutte le risposte a 12 o 13 quesiti a scelta tripla, quali sono appunto i pronostici dati col sistema 1-X-2.

Se la probabilità di fornire il 100% di risposte esatte lavorando a caso è pressoché trascurabile, è invece rilevante la possibilità di indovinare un numero tale di risposte da influire sensibilmente sul punteggio complessivo. Questa probabilità (espressa in frazioni o in percentuali) è facilmente determinabile: 1/2 (50%) per i quesiti a scelta binaria; 1/3 (33,3%) per quelli a scelta tripla; 1/4 (25%) per quelli a scelta quadrupla; 1/5 (20%) per quelli a scelta quintupla, e così via: il denominatore della frazione corrisponde al numero delle alternative.

Nel calcolo dei punteggi in test basati sulle scelte multiple si usa ricorrere ad una formula correttiva per scoraggiare la tendenza ad indovinare, penalizzando le risposte errate. Poiché invece nessuna penalizzazione è prevista per le risposte omesse, ai fini del punteggio globale all'allievo conviene non rispondere a quei quesiti dei quali non conosce la risposta con sufficiente sicurezza. Ciò, ovviamente, richiede che si informino gli allievi dell'esistenza di una penalizzazione ogni volta che si intende calcolare il punteggio usando la formula correttiva.

⁶⁷ Vedi oltre, par. 9.4.

⁶⁸ Esattamente 1.125.899.906.842.624. Vedi M. GATTULLO, *Il controllo delle risposte a caso nei test di profitto*, cit., p. 268.

Tale formula si basa sulla cosiddetta *ipotesi nulla*. Lo studente che desse 25 risposte esatte a 100 quesiti a scelta quadrupla fornirebbe esattamente la stessa prestazione che avrebbe presumibilmente raggiunto lavorando a caso. Nell'ipotesi nulla, tale prestazione coincide con il punteggio *zero*. A tale punteggio si giunge osservando che, per ogni quesito, le risposte errate possibili sono tante quante le alternative meno una (dato che una delle alternative corrisponde alla risposta giusta). Pertanto il punteggio *probabile* si calcola con la formula

$$P = E - S / (n - 1)$$

in cui P è il punteggio che tiene conto della probabilità di dare risposte esatte lavorando a caso; E è il numero delle risposte esatte, da cui va sottratto il numero delle risposte sbagliate S dopo averlo diviso per il numero *n* delle alternative diminuito di 1.

Applicando la formula all'esempio precedente (25 risposte esatte e 75 sbagliate in un test a scelta quadrupla) si ottiene:

$$P = 25 - 75 / (4 - 1) = 25 - 75/3 = 25 - 25 = 0$$

L'uso della formula presenta limiti precisi e non è esente da critiche,⁶⁹ tuttavia essa può essere utilmente adottata nella maggior parte delle normali situazioni di impiego dei test. Si noti che con l'applicazione della formula noi riusciamo a distinguere tra loro lo studente che dà (per tornare all'esempio) 25 risposte esatte e 75 errate, dallo studente che dà solo le 25 risposte esatte, omettendo di rispondere a tutte le altre. In tal caso la formula dà:

$$P = 25 - 0/(4-1) = 25 - 0 = 25$$

ossia un punteggio positivo. Impostando la questione in questo modo si presuppone che questo secondo allievo abbia fatto una scelta precisa tra i quesiti, e non abbia risposto a caso. In tal senso sarebbe giusto che anche nel punteggio la sua prestazione venga giudicata diversamente dall'altra, in senso positivo. D'altronde si potrebbe osservare che l'omissione di una risposta è, di fatto, un'ammissione di ignoranza, di incapacità a risolvere il quesito; non vi sarebbe, in altre parole, una notevole differenza qualitativa tra una risposta sbagliata ed una omessa. Evidentemente entrambe le tesi sono plausibili, anche perché non è possibile, di solito, conoscere la genesi degli errori (ignoranza, confusione, distrazione, tentativo di indovinare, o altro) né l'esatta motivazione dell'astensione (ignoranza della risposta, insicurezza, mancanza di tempo, ecc.). Le stesse risposte esatte potrebbero essere casuali non solo in quanto frutto di tentativi fortunati, ma anche in quanto conseguenza di scelte effettuate in funzione di un ragionamento sbagliato. Sarà compito dell'insegnante giudicare nelle singole situazioni il valore da attribuire ai punteggi ed agli effetti della formula correttiva, sulla base del tipo di prova effettuata, dell'atteggiamento degli allievi nello svolgimento della prova, del tipo di preparazione prossima e remota e di ogni altro dato disponibile.

Nessuna prova, oggettiva o tradizionale che sia, può infatti garantire l'assenza di "inquinamenti", quali appunto le risposte esatte casuali; anche dei processi mentali attraverso cui l'allievo giunge a formulare le sue risposte o a svolgere le prove non è solitamente possibile avere la chiara visione, e bisogna da un lato prendere atto dei dati oggettivi relativi alle varie prestazioni, dall'altro elaborare tutta una serie di prove che possano permettere all'insegnante una attenta verifica anche del *modo* in cui l'allievo lavora.

9.3. Il pre-testing

La quarta fase (somministrazione del test sperimentale per la raccolta di dati) è anche detta "fase di *pre-testing*". Attraverso di essa si tende a controllare la validità dei singoli quesiti ed a determinare

⁶⁹ Si veda M. GATTULLO, *Didattica e Docimologia*, cit., pp. 273-276; D. ADKINS WOOD, *Test Construction*. Columbus, Ohio, Merrill Books, 1961, pp. 39-41; D. P. HARRIS, *Testing English as a Second Language*, cit., pp. 128-129; R. LADO, *Language Testing*, cit., pp. 366-367.

il grado di efficienza globale della prova. Di conseguenza, i risultati dei singoli allievi debbono essere impiegati *prima* per verificare l'attendibilità del test e *poi* (solo nel caso che tale verifica dia esito positivo) per giudicare del grado di preparazione degli allievi stessi.

Il campione⁷⁰ di studenti ai quali viene somministrato il test sperimentale dovrebbe essere il più vasto e rappresentativo possibile; ciò tuttavia non è facilmente realizzabile nella maggioranza delle situazioni in cui si trovano ad operare gli insegnanti, per cui ci si potrà adattare a lavorare sui risultati ottenuti in due o tre classi parallele o (al limite) anche in una sola classe. È evidente che anche nel nostro caso, come per tutte le sperimentazioni, la collaborazione di più insegnanti nell'ambito della stessa scuola consentirebbe di raggiungere i migliori risultati.

Agli allievi vengono distribuiti i fogli (o fascicoli) d'esame, contenenti i quesiti e le istruzioni per la corretta esecuzione della prova, e i fogli per le risposte. Questa procedura di somministrazione è la stessa alla quale abbiamo fatto cenno al par. 4.3. Il foglio-allievo per il pre-testing recherà appositi spazi nei quali annotare i dati ricavati dalla correzione: risposte esatte, sbagliate ed omesse, nonché altre eventuali notizie di cui si volesse tener conto — per esempio, il tempo impiegato a terminare la prova.⁷¹

La massima cura dovrà essere prestata affinché sul foglio-allievo siano chiaramente e facilmente determinabili gli spazi nei quali scrivere le risposte (quesiti con completamento) oppure le lettere corrispondenti alle alternative (quesiti a scelta multipla). Per ridurre al minimo l'eventualità di errori materiali nel segnare la risposta ai quesiti il foglio-allievo riporterà tutte le indicazioni contenute nel fascicolo d'esame e che servono a identificare i quesiti: ad esempio, i titoli dei subtest, il numero progressivo che contrassegna i quesiti, eventuali parole-chiave. È opportuno che la numerazione sia unica anche se la prova si articola su più parti; in caso contrario potrebbe darsi che qualcuno dia la risposta al quesito n. 3 del *secondo subtest* segnalandola nello spazio per la risposta al n. 3 del *terzo* (o primo, quarto, quinto, ecc.) *subtest*. Ciò si evita molto facilmente, contrassegnando con un dato numero *un solo quesito* in tutto il test.

A questo punto tralasciamo altri dettagli tecnici relativi alla redazione e presentazione del test, ritenendo sufficienti le indicazioni fornite perché ognuno possa preparare materiali di testing ben costruiti anche dal punto di vista dell'accuratezza della realizzazione. Tale accuratezza è peraltro essenziale se non si vuole che la prova possa risultare inficiata da carenze o inesattezze estranee al contenuto del test ma che influiscono ugualmente sul punteggio.

Allo scopo di raccogliere dati su tutti i quesiti in fase sperimentale si possono invitare gli allievi a dare una risposta a tutte le domande, scegliendo l'alternativa ritenuta migliore o, comunque, tentando in ogni caso di risolvere il quesito. Ovviamente questa procedura non è sempre applicabile, specialmente in quelle prove in cui l'esattezza nozionale è di primaria importanza e quindi il chiedere una risposta anche a coloro che non conoscono certi dati si risolverebbe di fatto in una serie di tentativi a caso; ove possibile, però, è preferibile analizzare i quesiti senza l'interferenza delle omissioni. Nel caso che agli allievi si chieda di dare comunque una risposta non dovrà poi essere applicata la formula che penalizza le risposte sbagliate.

⁷⁰ In statistica, il termine *popolazione* designa l'intero oggetto delle rilevazioni; il termine *campione* indica una parte della popolazione, opportunamente selezionata, su cui viene effettuata un'indagine da cui inferire dati relativi a tutta la popolazione. Se si verifica con un test l'attitudine linguistica in una classe, gli studenti di quella classe costituiscono la popolazione se l'indagine verte solo su di loro; costituiscono invece un campione se l'indagine riguarda, ad esempio, tutti gli studenti di quella scuola.

⁷¹ Per le notevoli complicazioni metodologiche che il calcolo dei tempi comporta, prescindiamo in questa sede dai cosiddetti *speed tests* (test che verificano la velocità di esecuzione delle prove).

Rilevato, per ciascun esaminato, il numero delle risposte esatte, sbagliate ed omesse, si calcolano i singoli punteggi grezzi: questi coincideranno col numero delle risposte esatte se non si usa la formula correttiva, oppure con i punteggi probabili se la formula viene applicata. Sulla base dei punteggi grezzi i fogli per le risposte (detti anche *protocolli*) vengono riordinati dal migliore al peggiore. Nel caso che sia stato rilevato per ognuno il tempo impiegato per terminare la prova, a parità di punteggio precederanno i test eseguiti in minor tempo.

9.4. Indici di facilità e di discriminazione

Dal punto di vista statistico, la facilità di un quesito altro non è che il rapporto tra il numero di risposte esatte ed il numero di allievi che hanno affrontato il quesito. Pertanto *Vindice di facilità* (IF) è espresso dalla formula:

$$IF = E \times 100 / N$$

in cui E rappresenta il numero delle risposte esatte ed N il totale dei protocolli (risposte esatte + sbagliate + omesse). Si esclude quindi ogni giudizio soggettivo sulla presumibile difficoltà della prova: è facile il quesito a cui risponde esattamente il 75% o più degli allievi, è di media difficoltà il quesito con un IF compreso tra 25 e 75, è difficile il quesito a cui rispondono esattamente meno di 25 allievi su cento.

Per *discriminazione* di un quesito si intende la sua capacità selettiva, ossia di condurre a differenziazioni tra le prestazioni dei vari allievi. Se ad un quesito rispondono tutti esattamente (IF = 100) ovvero nessuno risponde esattamente (IF = 0), tale quesito presenta una capacità di discriminazione nulla, in quanto non è in grado di differenziare gli studenti migliori dai peggiori. È pure nulla la discriminazione in quei quesiti ai quali tanto i migliori che i peggiori danno la stessa percentuale di risposte esatte. Se il quesito è ambiguo o male formulato vi può essere addirittura una discriminazione negativa, ossia i peggiori (seguendo un ragionamento sbagliato) danno più risposte “esatte” dei migliori.

Tra i metodi più usati per il calcolo dell'*indice di discriminazione* ricordiamo quello di Flanagan.⁷² Si debbono dividere i protocolli riordinati in tre gruppi: il primo comprende i migliori (27%), il secondo quelli centrali (46%), e il terzo i peggiori (27%). Per ogni quesito si calcola la percentuale delle risposte esatte del primo gruppo e la percentuale di risposte esatte del terzo gruppo (ossia, su un ipotetico campione di 100 studenti, le prestazioni — in percentuale — dei 27 migliori e dei 27 peggiori). Tali percentuali, riportate su un'apposita tabella, consentono di ricavare l'ID.⁷³ Non disponendo della tabella, è ugualmente possibile calcolare l'indice di discriminazione mediante la seguente formula:

$$ID = (M - P) / N$$

in cui M indica il punteggio (in dati assoluti, non percentuali) del gruppo degli studenti migliori; P il punteggio (in dati assoluti) del gruppo dei peggiori; N il numero degli studenti che compongono il gruppo M (oppure il gruppo P, che ovviamente deve essere composto da un ugual numero di soggetti).

Si supponga, ad esempio, che la risposta esatta sia stata data da 17 studenti su 20 nel gruppo dei migliori e da 6 su 20 nel gruppo dei peggiori:

$$ID = (17 - 6) / 20 = + 0,55$$

⁷² J. C. FLANAGAN, *Units, Scores and Norms*, in E. F. LINDQUIST, *Educational Measurement*, Washington, American Council on Education, 1951, pp. 695-763.

⁷³ La tabella è pubblicata nei manuali sul testing e nei repertori statistici; v. M. GATTULLO, *Didattica e Docimologia*, cit., p. 548.

Questi indici variano da + 1,00 (discriminazione massima: tutti i soggetti M rispondono esattamente e tutti i P danno risposte errate o si astengono) a — 1,00 (discriminazione rovesciata massima: il caso opposto del precedente). Allo zero corrisponde la discriminazione nulla (M = P); normalmente si ritiene accettabile una selettività minima pari a + 0,20. Rispetto alla formula ora esaminata, l'indice di Flanagan tiene conto non solo della selettività, ma della discriminazione in rapporto anche alla difficoltà del quesito: è infatti intuitivo come i quesiti molto facili o molto difficili non possano agevolmente esprimere una discriminazione alta.

Quando i protocolli da analizzare sono pochi, si preferiscono altri criteri per la verifica della validità dei quesiti a scelta multipla. Ecco una procedura semplice ma efficace:

1. dividere i protocolli in due gruppi (50% dei migliori e 50% dei peggiori punteggi); se i protocolli sono in numero dispari si accantona quello centrale;
2. per ogni quesito, indicare quante volte è stata prescelta ciascuna alternativa dai migliori e quante volte dai peggiori;
3. contrassegnare con un segno convenzionale l'alternativa corrispondente alla risposta esatta;
4. sommare le risposte esatte dei migliori e dei peggiori, per il calcolo dell'IF;
5. analizzare l'andamento delle risposte esatte ed errate in rapporto ai due gruppi.

Esamineremo una serie di casi possibili, supponendo di disporre di trenta protocolli relativi ad un test a scelta tripla.

Quesito n. X	A	B	C°
15 migliori	3	2	10
15 peggiori	7	6	2
TOTALI	10	8	12

Il contrassegno ° indica che la risposta esatta è la C. Rispetto ad essa, i migliori hanno risposto assai meglio dei peggiori; il contrario avviene nei confronti delle altre due alternative. Il quesito appare pertanto valido in quanto a selettività. Poiché il totale delle risposte esatte è 12, avremo

$$IF = 12 / 30 \times 100 = 40$$

Quesito n. Y	A°	B	C
15 migliori	6	2	7
15 peggiori	4	9	2
TOTALI	10	11	9

Analizzando i risultati, si rileva un andamento corretto nei riguardi dell'alternativa A (risposta esatta) e B (errata). L'alternativa C presenta invece l'andamento tipico delle risposte esatte (i migliori l'hanno scelta più dei peggiori); anzi, in tal senso appare più valida dell'alternativa A. Con ogni probabilità il quesito è formulato in maniera ambigua, oppure ammette due soluzioni: esso deve essere scartato ovvero riveduto e corretto. In questo caso, ovviamente, non ha senso calcolare l'IF.

Quesito n. Z	A	B°	C
15 migliori	4	11	0
15 peggiori	8	7	0
TOTALI	12	18	0

L'analisi delle risposte ci rivela un quesito che solo apparentemente è a scelta tripla, mentre in realtà solo due sono le alternative effettivamente operanti. Il fatto che nessuno (nemmeno tra i peggiori) abbia scelto C come risposta, indica che l'alternativa offerta è palesemente errata, inverosimile o assurda. Anche in questo caso il quesito deve essere eliminato o migliorato.

Come appare dagli esempi riportati, anche disponendo di un numero esiguo di elaborati è possibile accorgersi della presenza di quesiti mal formulati o comunque di anomalie nella redazione. Ciò deve indurre gli insegnanti che elaborano i propri test a verificare ogni volta la validità del proprio operato. Sarebbe estremamente grave abbandonare le prove tradizionali — avendone riconosciuto limiti, carenze e pericoli — per ricorrere a test altrettanto poco attendibili, e magari con l'illusione di possedere strumenti più efficaci di misurazione del profitto. Solo attraverso la sperimentazione dei test si potrà disporre di *forme finali* utilizzabili con sicurezza ogni volta che gli allievi raggiungono certi punti nello sviluppo del programma.

9.5. Affidabilità dei test

Il concetto di affidabilità (in parte già illustrato al par. 1.5.), si riferisce alla capacità di un test di fornire dati costanti e coerenti nelle sue successive somministrazioni. Accenniamo brevemente ai tre metodi più frequentemente usati per il controllo dell'affidabilità di un test.

Il primo metodo, detto del *test-ritest*, prevede la somministrazione del test ad uno stesso gruppo di allievi in due momenti successivi, con un congruo intervallo tra le due prove. È applicabile in quei casi in cui — per il contenuto del test o per le condizioni nelle quali opera il gruppo di allievi — non sussiste il pericolo che intervengano elementi perturbatori tali da alterare sensibilmente i risultati della seconda somministrazione. I risultati delle due prove vengono raffrontati mediante l'uso delle formule per il calcolo degli indici di correlazione (vedi par. 10.5).

Il secondo metodo comporta la redazione di due *forme equivalenti* dello stesso test, somministrate contemporaneamente, e la correlazione dei risultati ottenuti. La maggiore difficoltà risiede nel garantire l'effettiva equivalenza delle forme parallele e situazioni di somministrazione del tutto equiparabili.

Per ovviare, almeno in parte, a tali difficoltà, si fa talora ricorso al metodo dei *semitest*: non potendo redigere due forme equivalenti di un test, si somministrano, come se fossero due test paralleli, le due metà del test opportunamente diviso (ad esempio, i quesiti pari e i quesiti dispari). Anche qui il confronto è operato mediante il calcolo della correlazione statistica tra le due serie di risultati.

Questi metodi, ovviamente, non si escludono a vicenda: essi, anzi, possono integrarsi l'un l'altro permettendo di misurare l'equivalenza e la stabilità combinate, mediante la somministrazione di forme equivalenti in momenti successivi. A sua volta, una buona correlazione positiva può confermare l'effettiva *coerenza interna* del test, sulla base della quale era stata condotta la revisione dei quesiti.

Si tenga tuttavia presente che, pur operando con il massimo rigore metodologico, nessun metodo è in grado di garantire la sicura oggettività dei test, anche se sono possibili alti livelli di sofisticazione nel controllo dell'affidabilità; a volte, si rischia di entrare in circoli viziosi. Ad esempio, la validità dei quesiti viene verificata sulla base dei punteggi offerti dal test di cui i quesiti stessi fanno parte. In altre parole, il quesito Y, ritenuto poco valido sulla base delle risposte degli studenti migliori e peggiori, ha esso stesso contribuito a determinare quali fossero i “migliori” e i “peggiori”. Gli esperti di testing concordano in genere ad attribuire un valore più teorico che reale a simili obiezioni, riconoscendo peraltro la loro fondatezza: ne abbiamo riferito affinché non si sopravvalutino le capacità dei test di fornire dati attendibili e indistorti. Ogni insegnante, poi, nella sua situazione didattica specifica, sarà in grado di interpretare correttamente i risultati dei test e di inquadrarli criticamente nell'ambito delle rilevazioni effettuate sul profitto della classe.

CAP. 10 ELABORAZIONE DEI PUNTEGGI

10.1. Le medie

Il lavoro di analisi statistica che il compilatore dei test deve svolgere non termina con la redazione della stesura definitiva del test. Occorre effettuare tutta una serie di rilevazioni che permettano di interpretare correttamente i punteggi ricavati dalla somministrazione del test ai gruppi di allievi. Per giudicare la difficoltà globale del test, l'ampiezza del "ventaglio" di punteggi, il grado di concordanza dei risultati del test con altre serie di risultati, il valore da attribuire ai singoli punteggi, è necessario calcolare una serie di indici e parametri. Il primo e più semplice di essi è la *media aritmetica*, che tutti ben conosciamo anche perché nella prassi scolastica a fine trimestre o quadrimestre si "fanno le medie" per determinare il voto.⁷⁴

La media aritmetica (M) si ottiene dividendo la sommatoria (Σ) dei punteggi (x) per il numero (N) dei punteggi stessi:

$$M = \Sigma x / N$$

Il calcolo del punteggio medio conseguito dagli allievi in un test ci offre un primo termine di raffronto: anzitutto ci permette di osservare quali allievi abbiano fornito una prestazione superiore alla media e quali invece una prestazione inferiore; in secondo luogo possiamo assumere la media aritmetica come indice della difficoltà globale di un test. Una prova in cui il punteggio medio (espresso su una scala di cento punti) è pari a 58 è, globalmente considerata, assai più difficile di un test la cui media aritmetica dei punteggi (sempre riportata a 100) è pari a 90. Di conseguenza, *i punteggi grezzi hanno tutti un valore relativo*. L'allievo che consegue un punteggio di 78 nel primo test si colloca ben al di sopra della media, mentre lo stesso punteggio nel secondo test potrebbe equivalere ad una prestazione scadente.

La media aritmetica possiede un'importante proprietà: se si sommano algebricamente gli *scostamenti* (ossia le differenze) di ciascun punteggio dalla media, il totale è uguale a zero. Ricorriamo ad un semplice esempio per chiarire il concetto.

Si abbiano i punteggi: 2, 4, 7, 11; la media aritmetica è:

$$M = (2+4+7+11) / 4 = 24 / 4 = 6$$

Gli scostamenti saranno: $2-6 = -4$; $4-6 = -2$; $7-6 = 1$; $11-6 = 5$, e la loro somma algebrica:

$$-4-2+1+5=0.$$

In una serie di punteggi (ordinati dal migliore al peggiore) la *mediana* (ME) corrisponde al punteggio centrale. Esempio: si abbiano i punteggi 82, 47, 41, 38, 29, 28, 24. La mediana è 38, essendo questo il punteggio centrale della scala, preceduto e seguito da altri tre punteggi. Qualora i punteggi fossero di numero pari (non esiste in tal caso un punteggio centrale), si assume convenzionalmente come mediana la media dei due punteggi centrali; ad esempio, con i punteggi 40, 38, 35, 31, 30, 22, la mediana è

$$(35 + 31) / 2 = 33$$

⁷⁴ L'accento alla prassi delle medie dei voti per stabilire le votazioni finali è stato qui fatto solo per richiamare il concetto di "media aritmetica", e non certo come implicita accettazione di tale procedura.

Calcolando la media aritmetica per i due esempi precedenti, si può verificare che media e mediana corrispondono spesso a valori diversi;⁷⁵ dell'utilità e della funzione della mediana ci occuperemo in seguito.

La *frequenza* di un dato statistico corrisponde al numero delle volte in cui tale dato compare in una certa rilevazione. Se in un test 5 allievi hanno conseguito il punteggio 36, si dirà che 5 è la frequenza di tale punteggio. Negli esempi che abbiamo preso in considerazione finora (per il calcolo di medie e mediane) ciascun punteggio appariva una sola volta, ossia la sua frequenza (f) era uguale a 1. Se i singoli punteggi hanno frequenze diverse da 1, per il calcolo della media aritmetica occorre anzitutto moltiplicare ogni punteggio per la rispettiva frequenza, indi sommare i prodotti e dividere il totale per il numero globale dei punteggi. La formula data in precedenza per il calcolo della media aritmetica deve essere così modificata:

$$M = \Sigma fx / N$$

Questa media si chiama *ponderata* perché ciascun dato viene moltiplicato per il suo “peso” (*pondus*) prima di essere immesso nella formula.

La *moda* è il dato che presenta la massima frequenza; nei test, è il punteggio conseguito dal maggior numero di allievi.

La media, la mediana e la moda sono *misure di tendenza centrale*: per loro natura tendono infatti a collocarsi al centro di una graduatoria o di una serie di punteggi.

Per rilevare la moda e calcolare la media ponderata conviene di solito ordinare i dati sotto forma di tabella. Nella prima colonna si riportano i punteggi conseguiti in un test, nella seconda le frequenze di ciascun punteggio e nella terza i prodotti dei punteggi per le rispettive frequenze. La seguente tabella riporta i dati ricavati dalla somministrazione di un test a cento allievi:

x	f	fx
24	1	24
23	4	92
22	6	132
21	11	231
20	20	400
19	18	342
18	15	270
17	9	153
16	7	112
15	4	60

⁷⁵ Nel primo caso: M = 41,2, ME = 38; nel secondo: M = 32,6, ME = 33.

La posizione della mediana (rispetto alla media) indica la tendenza generale dei punteggi in un test. Come si può desumere dal primo esempio, una mediana sensibilmente inferiore alla media segnala che i punteggi inferiori alla media sono più numerosi di quelli superiori alla media (5 e 2 rispettivamente); nel secondo esempio, il valore della mediana assai prossimo a quello della media indica una distribuzione equilibrata (3 punteggi superiori alla media e 3 inferiori).

14	2	28
13	1	13
12	1	12
11	1	11
N = 100		Σ = 1880

Inserendo i dati nella formula, otteniamo $M = 1880 / 100 = 18,8$; la moda si ricava direttamente dalla colonna delle frequenze: poiché il punteggio 20 è stato conseguito dal maggior numero di allievi, esso costituisce la moda per quel test. Il calcolo della mediana, in situazioni come questa, presenta qualche difficoltà; per i nostri scopi potrà tuttavia bastare un valore approssimato, desunto dalla distribuzione delle frequenze. Sommando progressivamente le frequenze iniziando dall'alto, riscontriamo che il cinquantesimo e il cinquantunesimo posto in graduatoria sono entrambi occupati da allievi che hanno conseguito il punteggio 19; questo punteggio può quindi essere assunto come valore mediano.⁷⁶

Un'ultima misura di tendenza centrale da prendere in esame (ci servirà in seguito) è la *media quadratica*, che è la radice quadrata della media aritmetica dei quadrati dei dati:

$$M_2 = \sqrt{\frac{\sum x^2}{N}}$$

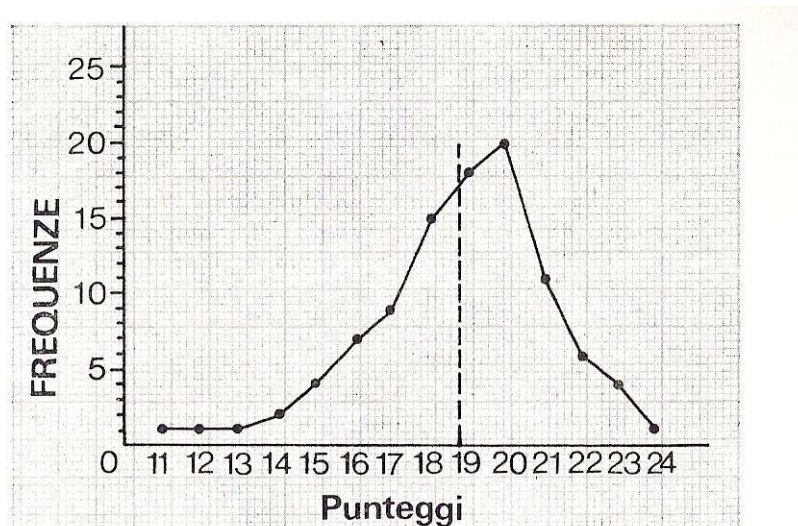
In pratica, si calcolano i quadrati dei dati e si sommano; il totale viene diviso per il numero dei dati, indi si estrae la radice quadrata del quoziente ottenuto. Il vantaggio principale della media quadratica consiste nell'eliminazione dei segni negativi; infatti qualsiasi numero, positivo o negativo, diventa positivo quando viene elevato al quadrato.

10.2. Distribuzione delle frequenze

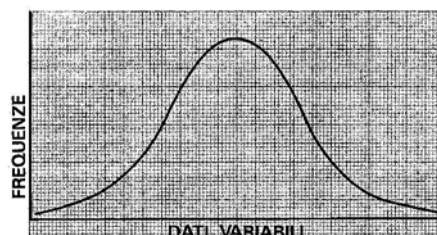
Possiamo rappresentare graficamente l'andamento dei dati riportati alla tabella precedente, ottenendo il *poligono delle frequenze*: l'asse orizzontale del grafico (*ascissa*) reca i punteggi; l'asse verticale (*ordinata*) le frequenze.

La linea tratteggiata verticale, in corrispondenza del punteggio 18,8, rappresenta la media aritmetica. Il grafico mette in evidenza come la maggioranza degli studenti abbia conseguito un punteggio superiore alla media; ne è conferma il fatto che il valore della mediana è più alto del valore della media.

⁷⁶ In realtà il calcolo esatto della mediana è un poco più complicato; qui, come altrove nei capitoli nono e decimo, abbiamo in parte semplificato il discorso, non sembrandoci utile presentare troppe formule.



Indagini statistiche a vasto raggio hanno dimostrato che se la popolazione considerata è sufficientemente numerosa⁷⁷ le frequenze relative a dati antropometrici o psicometrici si distribuiscono secondo una tipica curva a campana che, dal nome del matematico che per primo ne mise in evidenza le proprietà, prende il nome di *curva di Gauss*. I dati antropometrici sono la statura, il peso, la circonferenza toracica, ecc.; i dati psicometrici sono, ad esempio, i quozienti d'intelligenza, ma anche i punteggi dei test di profitto tendono a distribuirsi secondo lo stesso modello matematico.



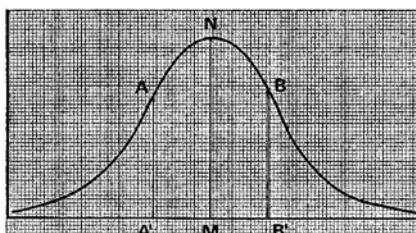
Rispetto ai poligoni delle frequenze che possiamo ricavare da un numero limitato di dati, la curva gaussiana presenta alcune notevoli e caratteristiche differenze. Anzitutto, media, mediana e moda coincidono: il dato medio è anche il più frequente, e divide la popolazione in due parti uguali. Se, ad esempio, la statura media dei giovani alla visita di leva, in un certo anno, è cm. 170, il gruppo dei giovani alti cm. 170 è più numeroso di ogni altro gruppo, ed i giovani la cui statura supera tale dato equivalgono per numero a quelli la cui statura è inferiore alla media.

Inoltre, la curva è simmetrica, il che significa che le frequenze diminuiscono in pari misura — quando ci allontaniamo dal dato medio — sia nella direzione dei valori bassi che in quella dei valori alti. Così, ad esempio, il gruppo dei giovani alti cm. 200 (cm. 30 più della media) tende ad essere altrettanto numeroso del gruppo dei giovani alti cm. 140 (cm. 30 meno della media).

Osservando l'andamento della curva, da sinistra verso destra, notiamo che essa all'inizio è concava verso l'alto, poi è convessa per tutta la parte centrale, indi è nuovamente concava verso l'alto. I punti in cui la curva cambia direzione (*punti di flesso*) hanno notevole importanza per la determinazione della *variabilità*. Con questo termine indichiamo la possibilità che hanno i valori di disporsi in una gamma più o meno estesa. Nell'esempio da noi considerato, i punteggi variavano da 11 a 24; una variabilità maggiore avrebbe dato luogo ad un poligono più largo, una variabilità minore ad un poligono più stretto. Ritornando alla curva normale, se riportiamo sull'ascissa i punti di flesso A e B

⁷⁷ È noto che le leggi della statistica (così come quelle sul calcolo della probabilità) valgono per i “grandi numeri”, si applicano cioè a rilevazioni numericamente estese.

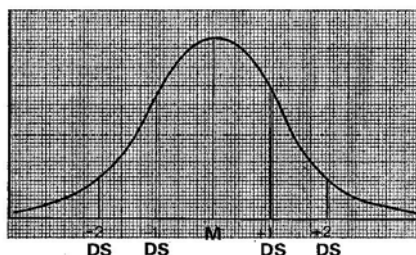
e rileviamo a quali dati corrispondono (rispettivamente A' e B'), disponiamo di un dato preciso sulla variabilità dei dati.



Il dato, chiamato *deviazione standard*, corrisponde all'intervallo tra il valore di M e quello di A oppure B'. Ad una deviazione maggiore corrisponde una curva più larga, ad una deviazione minore una curva più stretta; esse rappresentano rispettivamente una dispersione verso i valori estremi e una concentrazione verso il valore medio. La deviazione standard può essere calcolata, per ogni serie di dati, abbastanza agevolmente per mezzo della formula apposita (v. paragrafo 10.4).

Prima però di passare al calcolo del suo valore occorre approfondire il significato della determinazione dei punti-chiave in una distribuzione normale delle frequenze. È stato calcolato che nell'area compresa tra AA' e BB' si concentra il 68% circa⁷⁸ della popolazione; convenzionalmente si attribuisce a questo nucleo centrale la denominazione di *norma*. Al di fuori della norma si colloca quindi il 16% della popolazione con i dati più bassi (nel caso della statura, i nani o comunque le persone notevolmente basse) ed il 16% della popolazione con i dati più alti (i giganti o le persone notevolmente alte).

Nella figura seguente la deviazione standard è stata riportata esternamente una seconda volta in entrambe le direzioni:



Sulla scala dei dati si sono quindi calcolati cinque punti importanti: $-2 DS$, ossia due volte la deviazione standard in senso negativo; $-1 DS$, una deviazione standard in senso negativo; M, il valore della media aritmetica; $+1 DS$, una deviazione standard in senso positivo; $+2 DS$, il doppio della deviazione standard in senso positivo. Nell'area compresa tra le ascisse corrispondenti a $2 DS$ (in altre parole, tra le due linee verticali esterne) si colloca il 96% circa della popolazione; all'esterno troviamo quindi il 2% dei dati più bassi ed il 2% dei dati più alti.

10.3. Raffronto di dati e punteggi

Ritorniamo all'esempio della statura e supponiamo di aver calcolato che la deviazione standard è pari a cm. 10. Ne deduciamo che:

- a) il 68% dei giovani di leva ha una statura compresa tra cm. 160 ($170-10$) e cm. 180 ($170+10$);
- b) il 96% di essi ha una statura compresa tra cm. 150 ($170-2 \times 10$) e cm. 190 ($170+2 \times 10$);

⁷⁸ I dati nel testo sono approssimati all'unità; le percentuali approssimate a meno di un centesimo sono rispettivamente: 68,26; 15,87; 95,44; 2,28.

c) su un campione rappresentativo di 100 soggetti, un giovane alto cm. 180 supera per statura 84 soggetti, ed è superato da 16;

d) raffrontato allo stesso campione, un giovane alto cm. 150 supera per statura 2 soggetti, ed è superato da 98; analogamente si procede per gli altri punti considerati.

L'indice che riporta i singoli dati su una scala teorica di cento soggetti, tenuto conto della distribuzione normale delle frequenze, si chiama *centile*. Per le ragioni sopra considerate, al valore della media aritmetica corrisponde il centile 50; a +1 DS il centile 84; a +2 DS il centile 98; a —1 DS il centile 16; a —2 DS il centile 2.

Quale sarà (sempre riferito all'esempio precedente) il centile di un giovane alto cm. 175? Occorre anzitutto calcolare la differenza tra il dato e la media, indi dividere per la deviazione standard:

questo calcolo ci fornisce il *punto Z* relativo al dato considerato.

$$Z = (x - M) / DS = (175 - 170) / 10 = 0,5$$

Un'apposita tabella⁷⁹ permette di ricavare il centile corrispondente al punto **Z**; nell'esempio considerato, troviamo il valore 69,15, che significa che la statura di cm. 175 fa sì che quel giovane superi altri 69 soggetti e sia superato da 31.

Il calcolo dei centili ci permette quindi di ottenere un parametro per la corretta interpretazione di qualsiasi dato o punteggio. All'inizio del capitolo avevamo accennato al valore relativo dei punteggi; evidentemente, se anziché raffrontarli alla sola media aritmetica possiamo valutarli alla luce di altri indici, usciamo dal vago e dal generico. Il confronto tra dati assoluti non tiene conto della variabilità: un punteggio che supera la media, ad esempio, di 20 punti (su una scala di 100) può essere ottimo o appena discreto a seconda che la deviazione standard sia rispettivamente pari a 8 o a 40.

Il centile è anche un dato di raffronto tra i punteggi ottenuti dallo stesso allievo in test diversi; non è altrimenti possibile confrontare il punteggio 85 ottenuto nel test **A** con il punteggio 73 ottenuto nel test **B**, anche sapendo che la media aritmetica è pari a 80 nel primo e a 70 nel secondo.

10.4. Calcolo della deviazione standard

Data una serie di dati o punteggi, la deviazione standard equivale alla media quadratica degli scostamenti dalla media aritmetica, ed è espressa dalla formula:

$$DS = \sqrt{\frac{\sum (x - M)^2}{N}}$$

Analiticamente, la procedura è la seguente:

- 1) calcolare la media aritmetica (M);
- 2) sottrarre la media aritmetica da ciascun punteggio, per ottenere gli scostamenti (x — M);
- 3) elevare al quadrato gli scostamenti (x — M)²;

⁷⁹ Reperibile nei compendi di tabelle statistiche. Vedi (con qualche modificazione) M. GATTULLO, *Esami e Docimologia*, cit., p. 549.

- 4) sommare i quadrati degli scostamenti: $\Sigma (x-M)^2$;
- 5) dividere la sommatoria per il numero dei punteggi N;
- 6) estrarre la radice quadrata.

Questa procedura si applica quando tutti i punteggi ricorrono una sola volta (ossia la frequenza è uguale a 1); in caso contrario, prima di sommare i quadrati degli scostamenti, occorrerà moltiplicare ciascuno di essi per la rispettiva frequenza.

Ancora una volta, il modo migliore di procedere consiste nel predisporre una tabella riportando tutti i dati occorrenti: completeremo quindi la tabella iniziale con altre tre colonne calcolando gli scostamenti dalla media, i quadrati degli scostamenti, ed i prodotti dei quadrati per le rispettive frequenze:

x	f	fx	x—M	(x—M) ²	f(x—M) ²
24	1	24	+5,2	27,04	27,04
23	4	92	+4,2	17,64	70,56
22	6	132	+3,2	10,24	61,44
21	11	231	+2,2	4,84	53,24
20	20	400	+1,2	1,44	28,80
19	18	342	+0,2	0,04	0,72
18	15	270	—0,8	0,64	9,60
17	9	153	—1,8	3,24	29,16
16	7	112	—2,8	7,84	54,88
15	4	60	—3,8	14,44	57,76
14	2	28	—4,8	23,04	46,08
13	1	13	—5,8	33,64	33,64
12	1	12	—6,8	46,24	46,24
11	1	11	—7,8	60,84	60,84
N = 100		Σ = 1880		Σ = 580,00	

Poiché la tabella, a prima vista, potrebbe apparire più complicata di quanto essa in effetti non sia, sarà bene chiarire che:

- a. le prime tre colonne sono identiche a quelle della tabella precedente;
- b. la quarta colonna reca i risultati ottenuti sottraendo 18,8 (che è il valore della media aritmetica) a ciascun punteggio: $23-18,8 = 4,2$; $17-18,8 = -1,8$, e così via;
- c. la quinta colonna è costituita dai quadrati dei numeri della quarta colonna: il calcolo è estremamente facile e rapido disponendo di comunissime tavole numeriche;
- d. nell'ultima colonna sono riportati i prodotti ottenuti moltiplicando i quadrati (quinta colonna) per le rispettive frequenze (seconda colonna). La sommatoria di questi prodotti va inserita nella formula per il calcolo della deviazione standard:

$$DS = \sqrt{\frac{580}{100}} = \sqrt{5,80} = 2,40832$$

Il risultato (anch'esso desunto dalle tavole numeriche, senza bisogno di effettuare l'operazione di estrazione della radice quadrata) si può arrotondare a 2,41.

Questo dato, unitamente alla media aritmetica, ci serve per valutare il valore di ciascun punteggio e, eventualmente, per calcolare il centile. Sottraendo la deviazione standard alla media, ad esempio, troviamo il punteggio che corrisponde a -1 DS, e che a sua volta corrisponde al centile 16. Nel nostro caso avremo $18,8 - 2,41 = 16,39$: secondo le percentuali calcolate sulla base della curva teorica, al disotto di tale punteggio si dovrebbe trovare circa il 16% dei soggetti. In effetti, se sommiamo il numero di allievi che hanno un punteggio inferiore a 16,39, troviamo che essi sono sedici (su un totale di 100).

Allo stesso modo, possiamo calcolare $+1$ DS: $18,8 + 2,41 = 21,21$. Al punteggio 21 dovrebbe quindi corrispondere il centile 84: effettivamente possiamo osservare che gli allievi che hanno conseguito quel punteggio occupano nella graduatoria (a pari merito) i posti dal 12° al 22°, e assumendo come valore centrale il 17° posto in graduatoria, si vede come vi sia una notevole vicinanza tra i dati teorici e quelli reali, tenuto anche conto della parte decimale.

Non dobbiamo tuttavia attenderci una corrispondenza perfetta tra la distribuzione delle frequenze nel campione e quella normale espressa dalla curva di Gauss: tale corrispondenza si ha solo con un campione sufficientemente ampio, omogeneo ed equilibrato, mentre nel nostro caso già il poligono delle frequenze denunciava una distribuzione asimmetrica rispetto al valore medio. Ciò spiega come mai, ad esempio, riscontriamo 73 punteggi su 100, invece dei 68 prevedibili, compresi nell'area centrale tra -1 DS e $+1$ DS ossia tra 17 e 21.

Desiderando calcolare per i punteggi di un test centili attendibili, che possano servire da punto di riferimento per chiunque riutilizzerà il test standardizzato, bisogna effettuare la *taratura* su un notevole numero di soggetti, così da ottenere un poligono delle frequenze che si avvicini abbastanza alla curva teorica. Nelle normali situazioni scolastiche ciò si realizza difficilmente, e quindi i dati raccolti mediante l'elaborazione dei punteggi vanno opportunamente interpretati.

Il centile, pur rappresentando una misurazione del profitto molto più accurata rispetto al punteggio grezzo fornito da un test, non può di per sé tradursi in un voto convenzionale. Anzitutto occorrerebbe aver precisato ove collocare il livello di sufficienza: se questo viene fatto coincidere con il limite inferiore della norma, al centile 16 può già corrispondere il voto 6. È quindi totalmente assurdo pensare che si possa far corrispondere al centile 80 il voto 8, o al centile 30 il voto 3. Così facendo si dimostrerebbe di non aver capito nulla sulla distribuzione normale delle frequenze; inoltre si avrebbe una preponderanza di voti negativi, dato che al valore medio (centile 50) verrebbe assegnato un 5. Portando il discorso ad un livello più generale, occorre richiamare la distinzione basilare tra valutazione dell'allievo e misurazione del profitto, e mettere in guardia contro ogni tentativo di "meccanizzare" i giudizi.

L'uso corretto dei centili (così come di tutti gli altri dati statistici) consiste nell'impiegarli per verificare i progressi di ogni allievo, o evidenziare le lacune di preparazione per intervenire con una strategia didattica che permetta di condurre lo studente a superarle. In altre parole, la maggior precisione delle misurazioni deve tradursi in un migliore impegno didattico, e non già in un alibi per un atteggiamento selettivo che pretenderebbe di trovare una giustificazione nei risultati "oggettivi" di un test.

10.5. La correlazione statistica

Tutti gli strumenti di misurazione del profitto e di valutazione, siano essi test oggettivi o prove tradizionali, conducono all'attribuzione di classificazioni (voti, punteggi, ecc.) e/o alla formazione di graduatorie. Può essere quindi necessario rispondere a domande come le seguenti:

— In quale misura i risultati della prova A sono confermati da quelli della prova B, ovvero sono in contrasto con essi?

— I punteggi ricavati dalla somministrazione di un test sono, in complesso, compatibili con gli altri dati in possesso dell'insegnante?

La risposta a queste domande è di primaria importanza: se si dovesse accertare che i dati forniti da un test non hanno alcuna conferma negli altri elementi di giudizio a disposizione dell'insegnante, ovvero sono contraddetti da prove analoghe sicuramente affidabili, ciò significherebbe che il test è privo di validità. In tal caso non bisogna tenere conto dei risultati di quel test, che deve essere revisionato radicalmente.

Per risolvere questi problemi la statistica ha elaborato una serie di formule per il calcolo dell'*indice di correlazione* (r), che esprime il grado di corrispondenza esistente tra due serie di dati. L'indice è espresso da un numero compreso tra +1,00 e -1,00. Al valore di +1,00 corrisponde una correlazione diretta perfetta, come quella che intercorre tra il peso e il peso specifico di una serie di oggetti (a parità di volume); a -1,00 equivale una correlazione perfetta inversa, come quella tra le velocità ed i tempi necessari per percorrere un dato tragitto; lo zero indica l'assenza di correlazione tra le due serie di dati prese in esame.

Il concetto di *correlazione* non deve essere confuso con quello di *causalità*. Se si confrontano i dati annuali relativi all'andamento delle morti per annegamento con quelli sull'andamento annuo delle vendite dei gelati, si riscontra un'alta correlazione positiva, senza che peraltro uno dei due fenomeni sia la causa (o l'effetto) dell'altro: una terza variabile, il caldo, è la verosimile causa di entrambi.

Il metodo più semplice per il calcolo di r si basa sulle differenze di posizione nelle due graduatorie (*rank difference method* di Spearman). Dopo aver riordinato gli studenti dal migliore al peggiore in base ai punteggi (separatamente per le due graduatorie), si attribuiscono a ciascuno le posizioni in classifica: al migliore 1, al secondo 2, e così via. Nel caso di studenti a pari merito, si fa la media delle posizioni occupate: se il terzo ed il quarto hanno lo stesso punteggio, attribuiremo ad entrambi 3,5; se il sesto, il settimo e l'ottavo sono a pari merito, avranno tutti la posizione 7. Fatto ciò con entrambe le graduatorie, si confrontano le posizioni, studente per studente, e si calcolano le differenze. Se l'allievo A è ottavo nella prima graduatoria e decimo nella seconda, la differenza (D) è 2. Non occorre indicare se sia +2 o -2, perché l'operazione successiva consiste nell'elevare le differenze al quadrato, e con ciò (come abbiamo già visto) si eliminano i segni negativi.

La sommatoria dei quadrati delle differenze di posizione viene inserita nella seguente formula:

$$r = 1 - \frac{6 \cdot \sum D^2}{N(N^2 - 1)}$$

in cui D indica le differenze di posizione in graduatoria e N il numero dei soggetti.

Illustriamo la procedura con un esempio, basato sui risultati conseguiti da un gruppo di sette allievi in due test. Accanto a ciascuno indichiamo: il punteggio nel primo test (x), il punteggio nel secondo test (y), la posizione in graduatoria nel primo test (g_x) e nel secondo (g_y), la differenza di posizione in graduatoria (D) ed il suo quadrato (D^2).

Allievo	x	y	g_x	g_y	D	D^2
A	35	90	1	1	0	0

B	26	75	2,5	3,5	1	1
C	26	73	2,5	6	3,5	12,25
D	15	82	4	2	2	4
E	12	75	5	3,5	1,5	2,25
F	8	74	6	5	1	1
G	5	65	7	7	0	0

La sommatoria dei dati dell'ultima colonna ($\Sigma D^2 = 20,50$), assieme al numero degli allievi ($N=7$), può ora essere immessa nella formula:

$$r = 1 - (6 \times 20,50) / 7(7^2 - 1) = 1 - 123 / 7(49 - 1) = 1 - 123 / 7 \times 48 = 1 - 123 / 336 = 1 - 0,36 = 0,64.$$

Poiché normalmente si assume 0,40 come indice minimo accettabile di correlazione positiva, possiamo ritenere la correlazione tra i le due graduatorie discreta ma non altissima. Vi sono infatti alcuni I casi notevoli di dissimmetria: ad esempio, l'allievo C, secondo a i pari merito nel primo test, è penultimo nel secondo; l'allievo D, secondo nel secondo test, è solo quarto nel primo. In alcuni casi la formula di Spearman non basta a darci un indice esatto della correlazione tra due serie di punteggi. Nell'esempio seguente:

Allievo	x	y
A	9	8
B	4	7
C	3	6
D	1	5

la correlazione tra le *graduatorie* è perfetta, e quindi dalla formula ricaviamo $r = 1,00$. Osservando i *punteggi* si rilevano invece delle dissimmetrie: i valori della prima serie sono distribuiti irregolarmente su un arco molto esteso (da 1 a 9); quelli della seconda serie si collocano ad intervalli regolari e in un arco molto più ristretto. Calcolando il coefficiente di correlazione sulla base dei punteggi (*product-moment correlation* di Pearson) si ottiene $r = 0,945$, un indice molto alto ma che evidenzia come la correlazione non sia perfetta.

La formula per il calcolo di questo indice è:

$$r = \frac{N \Sigma xy - (\Sigma x) (\Sigma y)}{\sqrt{[N \Sigma x^2 - (\Sigma x)^2] [N \Sigma y^2 - (\Sigma y)^2]}}$$

In questa formula, oltre a N (numero degli allievi) vanno immessi i seguenti dati:

- la sommatoria dei punteggi del primo test: Σx ;
- la sommatoria dei punteggi del secondo test: Σy ;
- la sommatoria dei quadrati dei punteggi del primo test: Σx^2 ;
- la sommatoria dei quadrati dei punteggi del secondo test: Σy^2 ;
- la sommatoria dei prodotti ottenuti moltiplicando ciascun punteggio x per il corrispondente punteggio y: Σxy ;
- il quadrato della sommatoria delle x: $(\Sigma x)^2$;
- il quadrato della sommatoria delle y: $(\Sigma y)^2$.

Anche in questo caso il ricorso ad una tabella si rivela utile e serve a chiarire che anche questo calcolo, sebbene più laborioso dell'altro, non è però più difficile. Utilizzeremo ancora i dati impiegati per il calcolo dell'indice di Spearman:

Allievo	x	y	X ²	y ²	xy
A	35	90	1225	8100	3150
B	26	75	676	5625	1950
C	26	73	676	5329	1898
D	15	82	225	6724	1230
E	12	75	144	5625	900
F	8	74	64	5476	592
G	5	64	25	4096	320
Sommatorie	127	533	3035	40975	10040

Inserendo le sommatorie nella formula otteniamo:

$$r = \frac{7 \cdot 10040 - 127 \cdot 533}{\sqrt{(7 \cdot 3035 - 127^2)(7 \cdot 40975 - 533^2)}}$$

da cui, svolgendo i calcoli, si ricava $r = 0,69$.

La differenza (peraltro contenuta nei limiti di un 5%, valore relativamente modesto) tra i risultati forniti dalle due formule è da porre in relazione alla diversità dell'oggetto rilevato.

Si noti che con questa seconda formula si possono confrontare direttamente tra di loro due serie di dati, prescindendo dalla scala su cui essi si collocano: ciò permette, tra l'altro, di raffrontare i punteggi di un test (in centesimi o in qualsiasi altra scala) con i tradizionali voti in decimi. La possibilità di scegliere tra due formule diverse deve essere sfruttata a seconda degli obiettivi che ci poniamo nell'effettuare l'indagine statistica.⁸⁰

10.6. Il testing nella prassi scolastica

Non possiamo ignorare che l'impiego dei test nella normale *routine* di insegnamento ha suscitato e suscita varie perplessità ed obiezioni. Ad alcune di esse, pensiamo di avere dato una risposta trattando dei singoli problemi; altre, che spesso ci vengono presentate dagli insegnanti, meritano di essere trattate specificamente. Ad esse dedichiamo queste pagine conclusive.

1) La preparazione dei test richiede un notevole dispendio di tempo

È vero. Molto di questo tempo viene però recuperato attraverso una più rapida somministrazione e correzione delle prove, e ciò costituisce anche un vantaggio dal punto di vista didattico nella misura

⁸⁰ Se gli allievi sono pochi, i dati abbastanza omogenei, e non vi è necessità di rilevazioni accurate, la correlazione può essere valutata mediante un semplice confronto dei punteggi o delle graduatorie; soprattutto sulla base di queste ultime l'insegnante può osservare quali allievi abbiano compiuto progressi rispetto ai compagni e quali invece abbiano perduto posizioni in graduatoria. Si tratta di una prima indicazione nel caso occorra stabilire fasce di livello ovvero dividere in gruppi la classe. Nessuna valutazione "a occhio" può essere invece attendibile se i dati sono parecchi, disomogenei e di difficile interpretazione, ovvero se le circostanze richiedono il calcolo di indici precisi.

in cui si riesce a comunicare sollecitamente agli allievi il risultato delle prove sostenute e la correzione degli errori. Le più comuni prove tradizionali richiedono tempi relativamente lunghi per essere corrette e accade spesso che gli elaborati vengano riconsegnati agli allievi quando questi hanno dimenticato la genesi dei propri errori e quindi sono interessati essenzialmente al voto. Si tenga poi presente che il tempo impiegato per redigere un test è inversamente proporzionale all'esperienza acquisita in questa attività: le prime volte si procede lentamente e con difficoltà, ma in seguito l'abitudine permette di preparare prove abbastanza complesse anche senza molto tempo disponibile.

2) *L'elaborazione statistica dei dati e dei punteggi è difficile, richiede parecchio tempo e la disponibilità di apposite apparecchiature*

È solo parzialmente vero. In quanto alla difficoltà e al dispendio di tempo, vale ciò che abbiamo appena osservato a proposito della redazione dei quesiti: dopo le prime volte, l'applicazione delle formule diventa agevole e relativamente rapida. Si consideri poi che non occorre percorrere ogni volta l'intero itinerario di rilevazioni ed elaborazioni statistiche; ciò è indispensabile solo per la preparazione di test standardizzati. Spesso può bastare il calcolo della media aritmetica e della deviazione standard. Abbiamo preso in considerazione alcune formule (tralasciandone peraltro numerose altre) soprattutto per illustrare l'ampia gamma di informazioni che si ricavano da un'analisi statistica e per puntualizzare alcuni concetti essenziali, quali quelli di *validità* e *affidabilità* dei test, *distribuzione delle frequenze* e *correlazione*. Per i calcoli occorrono le tavole numeriche dei quadrati e delle radici quadrate,⁸¹ disponendo di una calcolatrice, o anche solo di un'addizionatrice tascabile, si può accelerare notevolmente l'esecuzione dei conteggi.

3) *La costruzione dei test implica la disponibilità di mezzi ed attrezzature particolari*

Non è vero. Un ciclostile e (per i test orali) un magnetofono sono, in linea di massima, tutto quello che occorre; non crediamo che, nella situazione attuale delle scuole italiane, sia eccezionale il disporre.

4) *La redazione di test è compito di gruppi di esperti, non di singoli insegnanti*

È vero qualora i test debbano servire ad un impiego generalizzato su vasta scala: è il caso dei test di profitto posti in commercio. Altrimenti ciascun insegnante non solo può, ma deve prepararsi i "suoi" test, che vanno bene nelle "sue" classi, esattamente come predispone le prove tradizionali. Naturalmente la collaborazione di altri insegnanti può essere preziosa e contribuire ad evitare errori di impostazione del test, redazione dei quesiti e interpretazione dei risultati.⁸²

5) *L'impiego dei test richiede all'insegnante conoscenze specifiche che, in genere, l'Università non fornisce agli studenti di lingue*

È vero. Tranne alcune eccezioni, le Facoltà di Lingue Straniere non hanno ancora dato adeguato spazio alla linguistica applicata, alla psicolinguistica, ed alla metodologia e didattica delle lingue straniere. L'obiezione quindi si riferisce non al LT come tale, bensì alla preparazione professionale complessiva dell'insegnante di lingue straniere. Coloro che invece possiedono sufficienti nozioni linguistiche e metodologiche non incontrano particolari difficoltà ad affrontare la problematica del testing.

6) *Le prove finali previste dagli ordinamenti scolastici non possono essere sostituite dai test*

⁸¹ Reperibili nella maggior parte dei testi di matematica per le scuole, o acquistabili separatamente (numerossime edizioni).

⁸² Lo svolgimento di un'organica attività di testing può rientrare nei progetti di sperimentazione previsti dai Decreti Delegati; in tal caso la collaborazione tra insegnanti si realizzerebbe usando opportunamente una parte delle venti ore mensili previste dallo stato giuridico.

È vero. Del resto non abbiamo mai affermato che i test debbano soppiantare *in toto* le prove tradizionali: essi servono invece ad assicurare che tali prove vengano affrontate dagli allievi in condizioni ottimali. Far eseguire un riassunto o una composizione dopo aver accertato la padronanza delle strutture linguistiche è assai più produttivo che servirsi delle stesse prove *anche* per verificare tale padronanza.

7) *L'uso dei test provoca un'eccessiva meccanicità nella valutazione ed induce ad accentuare atteggiamenti selettivi*

Non è vero: ciò potrebbe accadere solo qualora si confondesse la verifica delle abilità (che è un aspetto della misurazione del profitto) con la valutazione dell'allievo: abbiamo già operato una netta distinzione tra i due concetti e indicato quale sia la corretta impostazione del problema sotto il profilo pedagogico.⁸³ In quanto all'atteggiamento selettivo, non crediamo che esso possa essere posto in relazione con certi strumenti di verifica piuttosto che con altri, bensì riteniamo che esso debba essere collegato ad una visione globale della funzione della scuola nella società. Sappiamo bene che qualcuno, in particolari situazioni, cerca di avallare le proprie tesi sciorinando risultati di vari test: tenta cioè di far assurgere al rango di valutazioni i dati offerti da un mezzo di misurazione. A questo proposito occorre riaffermare la natura *strumentale* dei test: come ogni arnese funziona a dovere se si trova in mani esperte, così i test di profitto possono essere mezzi più efficaci — nei confronti delle prove tradizionali — solo se gli insegnanti che ne fanno uso sanno servirsene nell'ambito di una illuminata visione pedagogica, metodologica e didattica.

⁸³ V. paragrafo 1.2.