

ORIENTAMENTI ATTUALI NELLA VERIFICA DELL'APPRENDIMENTO DELLE LINGUE STRANIERE

di Gianfranco Porcelli (Centro di linguistica dell'Università Cattolica, sede di Brescia)

[in L. SORIANI CUCCHI (a cura di), *Insegnare la lingua: verifica e valutazione*, Milano, Bruno Mondadori, 1982, pp. 72-106]

Premessa

Da sempre il problema dell'accertamento del profitto e della valutazione dell'allievo ha interessato gli insegnanti di lingue straniere e, in particolare, coloro tra essi che si sono posti in atteggiamento critico nei confronti degli approcci tradizionali alla questione. L'attenzione verso il *language testing* (di seguito useremo l'abbreviazione LT) si è sviluppata di pari passo con la sensibilità alle istanze di rinnovamento metodologico. Il fiorire delle iniziative di ricerca e sperimentazione e il ripensamento in atto sugli obiettivi dell'insegnamento delle lingue estere (in termini di bisogni dei discenti e in un'ottica curricolare) hanno esaltato il ruolo e l'importanza del momento valutativo ai diversi livelli. L'interesse della scuola e dell'università italiana per questa materia (sia a livello strategico sia per ciò che si riferisce alle tecniche) è testimoniato dal moltiplicarsi di convegni, seminari e pubblicazioni sull'argomento.¹ Per questi motivi cercheremo qui di delineare gli orientamenti attuali, così come emergono da recenti studi e dalle ricerche in corso.

1. Sviluppo storico del language testing

Il volume *Language Testing* di Robert Lado, considerato da molti la pietra miliare che ha segnato l'inizio del LT moderno, fu pubblicato nel 1961. Sembrerebbe quindi poco opportuno parlare di "storia del testing" trattando di una disciplina affrontata in maniera scientificamente apprezzabile da soli vent'anni. Peraltro, si è effettivamente registrata un'evoluzione secondo linee e fasi di sviluppo chiaramente individuabili, e l'analisi di questo divenire permette di cogliere con maggior precisione dimensioni e valore degli indirizzi metodologici e delle tecniche di accertamento che oggi vengono proposti.

1.1. Le prove tradizionali

Con questo termine (in mancanza di uno migliore) si designano comunemente gli strumenti di controllo del profitto impiegati di solito prima che il rinnovamento dei metodi mettesse in crisi l'approccio grammaticale-traduttivo. Le prove tradizionali comprendono la traduzione (dalla e nella lingua straniera), la composizione e il riassunto, per quanto riguarda la lingua scritta, e le "interrogazioni" orali. Il dettato, che pure era spesso presente ai vari livelli (non escluso l'esame di abilitazione all'insegnamento), richiederà di essere trattato a parte.

Nella critica alle prove tradizionali è bene tenere distinte le motivazioni di ordine strategico/metodologico da quelle riferite agli aspetti più propriamente tecnici e docimologici. Non si possono accomunare impunemente traduzioni e temi, riassunti e interrogazioni, se non a prezzo di gravi confusioni.

¹ Citiamo, per restare in Lombardia, il Seminario nazionale promosso dall'università di Bergamo su *Testing English for Academic Purposes in Italian Social Science Faculties*, e il convegno internazionale di Milano sull'analisi dei bisogni linguistici, entrambi a pochi giorni di distanza da questo convegno regionale.

1.1.1. “L'uso della traduzione come test di controllo ne incoraggia l'abuso durante il corso”.² Questa frase del Lado sintetizza bene una serie di istanze metodologiche:

- a) l'esigenza di una corretta saldatura tra il modo in cui si insegna e quello in cui si accerta (v. *infra* il paragrafo sulla “validità di impostazione”);
- b) la necessità di tener conto dei riflessi che la verifica finale tende ad avere sul comportamento di insegnante e allievi già nella fase di apprendimento. Nei casi peggiori, questo rifluire all'indietro (il termine inglese *backwash* è particolarmente significativo) condiziona negativamente l'andamento di tutto il corso;
- c) l'importanza di analizzare la prova in termini di rispondenza con l'uso che se ne fa nella realtà extrascolastica. Che parentela esiste (se esiste) tra le traduzioni letterarie o scientifiche e le versioni di frasi o branetti *ad hoc*?

Se si accettano le critiche all'approccio grammaticale-traduttivo, e se si riconosce che il tradurre è una quinta abilità, diversa dal comprendere, parlare, leggere e scrivere, occorre concludere che l'uso della traduzione come test non è compatibile con gli attuali indirizzi metodologici nella didattica delle lingue, e che ogni tentativo di recupero delle versioni e retroversioni - sia pure con adattamenti e alterazioni più o meno apprezzabili - si scontra con ostacoli che, allo stato attuale delle nostre conoscenze in materia, appaiono insormontabili.

Un'affermazione così drastica può sollevare almeno due obiezioni. La prima si riferisce al problema dell'accertamento di questa quinta abilità nei corsi specificamente destinati ai traduttori. Evidentemente qui la traduzione ha pieno diritto di cittadinanza e il problema si sposta sugli aspetti tecnici: saper tradurre a qual fine? Quali materiali? In quali condizioni? Con quali sussidi? In quanto tempo? Sulla base delle risposte si potrà tentare di definire i criteri di giudizio e i parametri riferiti alla correttezza ortografica, morfosintattica, lessicale, allo “stile” ecc. Ritourneremo su questi aspetti, almeno indirettamente, quando affronteremo il problema delle prove “oggettive”. Qui ci basterà sottolineare che la nostra presa di posizione sulla traduzione non va riferita a situazioni speciali e, in certa misura, marginali, bensì ai normali corsi di lingua straniera, ad esempio nella scuola media e nel biennio superiore.

La seconda obiezione può riferirsi all'opportunità di ricorrere a retroversioni o a forme particolari di “traduzione guidata” in alcune fasi di apprendimento della lingua, ad esempio nel momento in cui un'unità didattica acquisita sul piano orale viene “recuperata” per quanto riguarda il codice scritto. Ebbene, anche ammettendo l'eventuale opportunità di tali forme di *esercizio* resta esclusa la validità della traduzione come *test*. Occorre sempre tener distinti due momenti sostanzialmente diversi: quello in cui l'allievo acquisisce le abilità (mediante *esercizi* di fissazione o di reimpiego della lingua) e quello in cui viene controllata l'avvenuta acquisizione.

A questo proposito va osservato che per quanto riguarda il testing il metodo grammaticale-traduttivo era viziato da un equivoco fondamentale: gli “esercizi” non erano esercizi, ma prove di controllo. Le frasi da tradurre non erano presentate e ordinate in modo da facilitare l'allievo e consentirgli di addestrarsi a produrre un congruo numero di strutture corrette in L2. Anzi, l'alternarsi di applicazione di “regole” e di “eccezioni” era tale da indurre facilmente in errore proprio chi più avrebbe avuto bisogno di essere aiutato ad apprendere. È paradossale come, in un momento in cui il LT in quanto tale non era ancora presente sulla scena didattica, si siano usate in modo così massiccio le prove di controllo, sia pure in forma surrettizia.

² R. LADO, *Language Teaching: a scientific approach*, McGraw-Hill, New York, 1964 (trad. ital. *Per una didattica scientifica delle lingue*, Minerva Italica, Bergamo, 1974, p. 196). Useremo qui il termine “scientifico” nel senso proposto dal Lado, ossia “sistematico e basato su un'accurata analisi contrastiva”.

1.1.2. Le obiezioni fondamentali, di ordine metodologico, che sono state sollevate nei confronti della traduzione, non possono essere trasferite alle altre prove tradizionali, scritte ed orali. Il parere pressoché unanime degli esperti è che la composizione in L2, ad esempio, è perfettamente coerente con gli indirizzi metodologici attuali e non presenta indesiderati effetti negativi di *backwash*; al contrario, tende a favorire l'uso costante della L2 anche negli esercizi intesi a favorire l'acquisizione delle strutture morfosintattiche e del lessico. Qualche riserva si può formulare per ciò che riguarda l'utilità extrascolastica del saper comporre, che appare limitata alla redazione di lettere d'affari. Questo limite è tuttavia abbondantemente compensato dai vantaggi a cui si è già accennato.

Nei confronti di tali prove, peraltro, si possono opporre obiezioni di carattere docimologico. È esperienza comune - suffragata, peraltro, da indagini attendibili³ - che l'interpretazione soggettiva del valutatore incide in misura ragguardevole sul giudizio. Ciò significa non solo che valutatori diversi sono normalmente in disaccordo (più o meno accentuato) sul giudizio da dare su un certo elaborato, ma anche che un valutatore incontra difficoltà ad essere coerente con se stesso nell'attribuire i punteggi ad una serie di elaborati (di solito, un "pacco" di compiti in classe). Una prova indiretta di ciò la si è avuta quando dei docenti sono stati chiamati a rigiudicare, a distanza di tempo, un elaborato: le differenze di voto - in alcuni casi, pari a due decimi - possono essere assunte come indice di questa incostanza di giudizio. Discrepanze ancora maggiori si riscontrano nelle prove orali, nelle quali l'"effetto di alone" è avvertibile in maniera più immediata.

La compresenza di fattori diversi (dalla "forma" al "contenuto" e, nell'ambito della "forma", dall'ortografia alla "eleganza" dello stile) reca con sé la difficoltà di isolare i singoli elementi, attribuire a ciascuno un "peso specifico" e giungere ad un giudizio che sia una rappresentazione equa ed attendibile delle prestazioni, non viziata da oscillazioni o interferenze indesiderate.

L'analisi sempre più attenta di queste prove, in termini sia di *fattori* o elementi da tenere sotto controllo, sia di *operazioni* o abilità coinvolte (v., più oltre, la sezione dedicata all'analisi degli errori), sia infine di *parametri* di giudizio, può condurre ad un minor grado di arbitrarietà e indeterminatezza, senza che peraltro sia eliminabile la soggettività del valutatore. È questo il motivo di fondo che ha indotto molti a ricercare una soluzione nei *test*.

A nostro avviso, non bisogna creare una falsa dicotomia, un'antitesi artificiosa tra prove tradizionali e test. In una corretta impostazione di un corso di lingua straniera, i due approcci alla verifica dell'apprendimento non sono in contrapposizione, ma debbono integrarsi, secondo un itinerario che passa attraverso i test (man mano sempre meno elementari e più articolati), per giungere alle prove più complesse che sole sono in grado di documentare certi aspetti *produttivi e creativi* (in senso tecnico) dell'apprendimento. E questo un tema di basilare importanza, su cui avremo modo di ritornare, ma che sin d'ora era importante evidenziare.

1.2. Test o quiz?

I primi test entrati nelle scuole - ci riferiamo soprattutto alle esperienze anglosassoni e scandinave - non erano test di profitto, bensì strumenti diagnostici usati da psicologi e (in minor misura) sociologi: test di intelligenza, indagini sulla personalità e sulle motivazioni, sociogrammi, questionari caratterologici, sociologici ecc. In molti casi si trattava (e si tratta tuttora) di strumenti che rispondono ad una serie di requisiti tra loro interrelati. I principali tra questi sono:

³ Ci riferiamo in particolare alle indagini condotte da De Castro e altri. Cfr. G. PORCELLI, *Il language testing: problemi e tecniche*, Minerva Italiana, Bergamo, 1975, capp. 1 e 2.

- 1) il test si compone di quesiti⁴ a uscita unica, che ammettono un'unica risposta esatta chiaramente predeterminata; anche i criteri per discriminare tra una risposta accettabile e una non accettabile sono definiti accuratamente e preventivamente. Ovvero,
- 2) i quesiti ammettono più risposte, ciascuna delle quali è però riconducibile ad un punteggio prestabilito;
- 3) il calcolo dei punteggi parziali e complessivi avviene secondo procedure prefissate;
- 4) il test è di norma costruito sulla base dei risultati di una fase di pre-testing; dopo una somministrazione preventiva ad un campione adeguato, si effettua l'analisi dei singoli quesiti, dei subtest e del test nel suo complesso per individuare eventuali disfunzioni;
- 5) redatto il test nella forma definitiva, esso viene sottoposto a *taratura* per calcolare, su un campione rappresentativo, una serie di dati che permettono di interpretare i punteggi grezzi (medie, deviazione standard, centili ecc.).

I primi tre punti tendono a garantire che il punteggio sia calcolato in modo da non risentire delle valutazioni e interpretazioni soggettive di chi corregge; il quarto si riferisce ad una procedura di controllo interno tendente a garantire che lo strumento funzioni a dovere; il quinto, infine, riguarda la ricerca di chiavi di lettura e interpretazione corretta dei risultati.

I due ultimi punti hanno trovato scarsa applicazione nei comuni test di profitto, non perché se ne disconoscesse l'importanza, ma per le obiettive difficoltà sia di disporre di campioni adeguati, sia di svolgere i calcoli richiesti. L'attenzione degli insegnanti si è invece focalizzata sulle tecniche di redazione dei quesiti, e in particolare su due di esse: la scelta multipla e il completamento. Così facendo, si è privilegiato l'aspetto tecnico-formale dei test rispetto alla ricerca sui metodi e contenuti. In altre parole, tra le due componenti primarie dell'oggettività, *la fedeltà* (o *affidabilità*, come oggi si preferisce chiamarla; in inglese, *reliability*) è stata oggetto di attenzioni molto maggiori rispetto alla *validità*.

Ciò ha fatto sì che la fase sistematica del LT fosse preceduta da una fase parascientifica, o dell'approssimazione. Dobbiamo ai guasti operati in questa fase le confusioni tra test e quiz: infatti anche le prove enigmistiche e i programmi radiotelevisivi di quiz si basano sul principio dell'unica risposta esatta.

Nessun test è valido se non risponde allo scopo a cui è destinato; nessuna forma di testing e, in generale, di controllo opera *in vacuo*, in una sorta di universo astratto e a sé stante: al contrario, la verifica nasce e si sviluppa come risposta ad esigenze concrete ed individuate. Il quiz può essere gratuito, senza alcun fine se non quello ludico, con contenuti casuali, occasionali, talora anche paradossali o totalmente assurdi, frutto di una fantasia che si sbriglia lungo linee autonome e non necessariamente nel rispetto di una realtà oggettiva: il test, assolutamente no.

Un indice della scarsa comprensione di questa fondamentale distinzione lo si ritrova in quei quesiti che contengono forme linguistiche ipotetiche, inesistenti, prive di un sufficiente contesto, o la cui probabilità di ricorrere in una normale situazione comunicativa è scarsa o nulla. In altre parole, quesiti che si basano non sulla lingua viva, ma su artifici di vario genere. Ecco un esempio:

John (plays) (plais) (playes) (plaies) tennis.

Le tre forme errate o *distrattori* sono inesistenti, e l'enunciato è ridotto al minimo, senza nemmeno un attualizzatore temporale (ad esempio, un avverbio di frequenza) che giustifichi l'uso del *present simple*. Si potrebbe argomentare che le forme usate come distrattori di fatto esistono, sia pure come errori frequenti commessi non solo da studenti, ma anche, a volte, da nativi. Tuttavia:

⁴ Usiamo questo termine per indicare i *test items* di qualsiasi tipo, non necessariamente sotto forma di domanda, così come “risposta” indica qualsiasi *performance* dell'allievo oggetto di osservazione.

a) questi errori si riscontrano nella produzione scritta, mentre il quesito opera nell'ambito del saper leggere;

b) il quesito presuppone l'esplicitazione formale di una regola di ortografia; ciò potrebbe non corrispondere al modo in cui di fatto in classe si è affrontato il problema dello *spelling*;

c) soprattutto, la presentazione di forme errate può nuocere allo sviluppo dell'apprendimento, per il rischio che esse possano fissarsi nella mente del discente invece della forma corretta.

Se, come avviene normalmente, un test non è un episodio isolato e collocato nel momento terminale di un corso, bensì una tappa in un processo di apprendimento, è importante che esso contribuisca a precisare e chiarire (pur non essendo questa la sua funzione specifica), e non certo a confondere le idee.

1.2.1. In questo quadro, è opportuno definire più puntualmente il concetto di *validità* dei test. Globalmente intesa, essa è la capacità di un test di verificare ciò che intende verificare e nient'altro. In altre parole (e a titolo esemplificativo) un “test di comprensione orale” riesce davvero ad accertare l'abilità di comprendere? E che cosa veramente controllano certi test sul lessico della L2? Non è forse vero che alcuni si basano in misura eccessiva sulla memoria visiva dell'esaminato?

L'esemplificazione rischia di banalizzare il problema, che è invece estremamente complesso. Ricerche recenti (in parte tuttora in corso) stanno indagando se i “test di intelligenza” non siano in realtà test di lingua: problema delicatissimo, con profonde e ramificate implicazioni psicopedagogiche.⁵ E se negli anni Settanta molti studiosi, negli U.S.A. e altrove, hanno assunto un atteggiamento più cauto nei confronti di certe indagini sulla personalità condotte mediante test, ciò fu dovuto al fatto che dei test veniva messa in discussione non l'affidabilità, bensì la validità. La crisi più acuta si ebbe quando furono evidenziati i limiti del behaviorismo skinneriano, e modelli di tipo cognitivista apparvero più idonei a rendere conto dei processi e delle modalità che presiedono all'acquisizione del linguaggio, allo sviluppo dell'intelligenza ecc. Come vedremo in seguito, queste critiche interessano solo indirettamente i test scolastici; in ogni caso, la validità di questi deve essere accertata autonomamente, prescindendo da considerazioni riferite ad altri campi di applicazione. L'insegnante di lingue che rifiuta i test “perché ormai le grandi aziende li hanno abbandonati” mostra di confondere i problemi della selezione del personale con quelli del controllo del profitto scolastico.⁶

Per precisare meglio il concetto di validità, con riferimento al LT, è opportuno richiamare i tipi di validità che sono stati individuati.⁷ Per *validità facciale* (dall'inglese *face validity*) si intende l'appropriatezza del test rispetto ai suoi fini, sotto almeno tre aspetti:

a) in rapporto alla situazione didattica: un test di profitto finale (*achievement o proficiency test*)⁸ non può essere correttamente impiegato come test prognostico o di accertamento iniziale (*assessment test*), o viceversa;

b) nei confronti della popolazione scolastica a cui il test è somministrato: un test redatto per studenti di scuola media con ogni probabilità non risulterà valido se usato in una classe elementare o a livello universitario;

⁵ Si veda J. W. OLLER, Jr., *Language Tests at School*, Longman, Londra, 1979, Appendice (p. 423 segg.).

⁶ Per inciso, noteremo che tale abbandono non è affatto generalizzato, anzi vi sono molte ricerche per lo sviluppo di nuovi test, soprattutto in campo linguistico.

⁷ Cfr. J. B. HEATON, *Writing English Language Tests*, Longman, Londra, 1975, pp. 153-155; R. M. VALETTE, *Modern Language Testing*, Harcourt Brace Jovanovich, New York, 1977, p. 46; J. P. B. ALLEN & DAVIES (a cura di), *Testing and Experimental Methods*, The Edinburgh Course in Applied Linguistics, Volume 4, Oxford University Press, 1977, pp. 18-22, 58; J. W. OLLER, *op. cit.*, pp. 50-69.

⁸ I due termini non sono sinonimi. V. oltre, al par. 3.4.

c) riguardo all'oggetto della verifica: un test sulla padronanza del lessico non deve coinvolgere altri elementi (ad esempio, l'ortografia, la pronuncia o la morfosintassi) né può validamente essere adottato per verificare abilità o contenuti diversi.⁹

La *validità dei contenuti* (*content validity*) postula un'attenta analisi della L2 in rapporto agli obiettivi didattici. Non si richiede solo l'assenza di elementi estranei o spuri: occorre piuttosto garantire la presenza di tutti i quesiti idonei a verificare il conseguimento degli obiettivi. Molti test sono carenti in questo senso: privilegiano contenuti sui quali è agevole costruire quesiti a scelta multipla o con altre tecniche elementari, e al tempo stesso trascurano elementi essenziali, ma difficili da verificare mediante tecniche univoche.

La *validità di impostazione* (*construct validity*) richiama l'esigenza che un test operi la verifica in coerenza con una data teoria dell'apprendimento linguistico: se l'insegnamento della L2 è condotto secondo un approccio comunicativo o funzionale-semantico, il controllo del profitto non può validamente operarsi con strumenti e procedure che si ricollegano a modelli diversi (ad esempio, di matrice behaviorista).

Infine, si riconosce una *validità statistica* (o *empirica*), che si realizza quando i risultati di un test trovano conferma in altri dati omogenei e affidabili di cui il valutatore disponga. Si tratta di un dato eminentemente tecnico, espresso normalmente mediante indici di correlazione.

1.3 L'apporto di Robert Lado

Abbiamo già fatto cenno all'importanza del contributo del Lado nella evoluzione del LT. La sua opera si inquadra nelle attività di ricerca e di sperimentazione didattica condotte presso l'Università del Michigan, a Ann Arbor, durante e dopo la seconda guerra mondiale sotto la guida di C. C. Fries.

Sulla scia dell'Army Specialised Training Program, che durante la guerra aveva segnato un'autentica svolta nell'impostazione e conduzione dei corsi di lingue estere, furono svolti studi importanti di linguistica generale e applicata (oltre a Fries e Lado, occorre ricordare almeno E. A. Nida e K. L. Pike) e realizzati corsi e strumenti didattici di vario tipo, noti come "Michigan material".

Lo stesso Lado è autore di *Linguistics across Cultures*, un volume pubblicato nel 1957 in cui si analizzano i metodi e i campi dell'analisi contrastiva a livello fonologico, grammaticale, lessicale, grafemico e interculturale. Dall'analisi strutturale di due lingue a confronto emergono le dissimmetrie esistenti, riferibili a parametri che il Lado chiama *forma*, *distribuzione* e *significato*. Nell'apprendere la lingua materna, ognuno impara non solo a cogliere i contrasti significativi (diremo meglio, le *opposizioni pertinenti*) che in essa operano, ma anche a ignorare e trascurare i tratti non pertinenti, rispetto ai quali diventa insensibile. Sono quelli che il Fries, nella premessa al volume, chiama i "blind spots", i punti ciechi presenti a tutti i livelli: si può essere sordi ad un'opposizione fonologica non presente nella pronuncia della L1, così come si può interpretare erroneamente un tratto culturale o un comportamento sociale rispetto a cui la civiltà straniera ha una diversa attribuzione di significato. Poiché modelli e strutture della lingua-civiltà materna interferiscono nell'apprendimento di una seconda lingua e civiltà,¹⁰ l'individuazione dei "punti ciechi" equivale all'individuazione dei problemi di apprendimento.

L'approccio di Lado al problema del testing prende le mosse proprio da questa interpretazione della dinamica delle lingue a confronto. "Secondo questa teoria, verificare il superamento dei problemi di

⁹ Peraltro, un test che isoli i problemi lessicali potrebbe *ipso facto* non adeguarsi agli altri requisiti di validità.

¹⁰ Rimandiamo a: U. WEINREICH, *Lingue in contatto*, Boringhieri, Torino, 1974; G. NICKEL (a cura di), *Papers in contrastive linguistics*, Cambridge University Press, 1971; E. ZUANELLI SONINO, *Italiano/Tedesco/Inglese: analisi contrastiva a livello fonico*, Mmerva Italica, Bergamo, 1975, cap. 1.

apprendimento equivale a verificare la padronanza della lingua. I problemi sono quelle unità o strutture linguistiche che non trovano corrispondenza nella lingua nativa o che corrispondono a unità e strutture aventi distribuzione o significato diversi". (Lado, 1961, p. 24).

Questa saldatura tra un serio lavoro di analisi linguistica e la realizzazione di test di profitto costituisce l'apporto più valido offerto dal Lado allo sviluppo del LT. Si tratta di un salto di qualità di particolare rilievo in ordine alla validità dei contenuti e della coerenza di impostazione. Con il superamento dello strutturalismo bloomfieldiano, sul versante linguistico, e dei modelli comportamentisti in psicologia, buona parte delle posizioni del Lado sono state sottoposte a revisione critica. Resta però sempre attuale il rigore dell'impostazione scientifica del problema del testing, sia nella ricerca di una teoria coerente ed esaustiva, sia nella lucida analisi delle singole tecniche.

1.4. Le critiche a R. Lado

Un primo test di inglese per ispanofoni (a livello di principianti) consisteva di soli tredici quesiti, tanti quanti erano i "problems of learning" riconosciuti presenti a quel livello. Ben presto lo stesso Lado superò posizioni così estremiste, riconoscendo la difficoltà di attenersi strettamente all'impostazione teorica da lui stesso elaborata: "Il compito di preparare test distinti per tutti i livelli di preparazione linguistica è tanto vasto che non potremo mai sperare di avere simili test salvo che per un numero limitato di lingue" (Davis, 1968, p. 13).

Al di là dei problemi concreti di realizzazione, vi sono obiezioni che riguardano alcuni punti-chiave della teoria del Lado, e che sono il risultato degli sviluppi successivi in linguistica, glottodidattica e psicolinguistica.

Un vizio di fondo - del resto comune a gran parte della produzione scientifica del tempo, soprattutto negli Stati Uniti - risiede nella identificazione, di fatto, della glottodidattica con la linguistica applicata. Senza dubbio, lo strutturalismo ha fornito apporti decisivi nella comprensione del linguaggio come sistema e, quindi, per l'elaborazione di nuovi metodi di insegnamento; tuttavia, in una corretta visione pedagogica, l'analisi dell'oggetto da insegnare è solo uno dei termini dell'equazione. Gli altri termini, la cui definizione è altrettanto importante, si riferiscono al binomio insegnante-allievo e all'ambiente socioscolastico.

Si pensi, ad esempio, al ruolo della motivazione nell'apprendimento linguistico, così come si configura in quello che R. Titone chiama il livello *egodinamico*.¹¹ Anche per quanto riguarda il testing sono necessari ampliamenti dell'orizzonte rispetto agli ambiti nei quali normalmente si conducono ricerche e sperimentazioni. In parte ciò si sta oggi realizzando nell'ottica di una programmazione curricolare, nella quale la *valutazione* (o *verifica* o *accertamento*, a seconda degli Autori) assume un ruolo nodale, collegando gli esiti dell'intervento didattico con gli *obiettivi*, i quali, a loro volta, tengono conto dei *bisogni* dei discenti.¹²

¹¹ R. TITONE, "The Holistic Approach to Second Language Education", in J. E. ALATIS, H. B. ALTMAN, P. M. ALATIS (a cura di), *The Second Language Classroom: Directions for the 1980's*, Oxford University Press, 1981, pp. 67-78 (ripreso, con integrazioni, da "A psycholinguistic definition of the "glossodynamic model" of language behaviour and language learning", *Rassegna Italiana di Linguistica Applicata*, vol. V, 1973, n. 1, p. 19).

¹² Con riferimento specifico alla didattica delle lingue, si vedano: G. PORCELLI, "Obiettivi comportamentistici nello studio delle lingue", in *Lingua e Civiltà*, a. V, n. 2, maggio 1972; G. ZANI, "Programmazione curricolare e didattica delle lingue", in *Lingua e Civiltà*, a. IX, n. 1, febbraio 1981, pp. 7-15; M. GUIDI, L. MARIANI, A. MAZZA Ferraris, "Gli obiettivi comportamentali nella programmazione didattica", in *Lingua e nuova didattica*, a. X, n. 3, luglio 1981, pp. 34-38.

In quest'ottica, appare anche possibile superare l'aspetto potenzialmente conflittuale del momento valutativo (che può tradursi nella cosiddetta "ansia da test") e rendere operante il concetto di valutazione formativa. Non possiamo qui affrontare questi temi, che pure fanno da cornice al discorso più tecnico sul testing; li abbiamo richiamati solo per illustrare come, anche per ciò che attiene al LT, la glottodidattica trovi i suoi fondamenti non solo nella linguistica, ma anche nella psicologia e nelle scienze dell'educazione.

Proprio dalla psicolinguistica provengono molte voci critiche nei confronti del Lado.¹³ In effetti, l'analisi dell'interferenza sia in *Linguistics across Cultures*, sia in *Language Testing* appare piuttosto elementare e viziata da meccanicismo. In particolare, gli studi sull'analisi degli errori hanno fornito importanti puntualizzazioni. Di ampio rilievo è la distinzione tra l'errore sistematico, che rivela lacune nella padronanza del sistema della L2, e lo sbaglio occasionale, legato a problemi di esecuzione (in inglese si usano rispettivamente i termini *error* e *mistake*). Un test che prevede un solo quesito per ciascun problema rinuncia in partenza ad ogni possibile distinzione tra i due tipi di errore. Perché ciò avvenga occorrono più verifiche dello stesso problema, possibilmente mediante quesiti congegnati in modo da consentire riscontri e controlli incrociati. Ciò si realizza regolarmente nei test *diagnostici*, l'obiettivo dei quali è l'accertamento non delle singole prestazioni bensì - in certa misura - della "competenza", sia pure circoscritta agli ambiti linguistici esaminati.¹⁴

Un'altra distinzione rilevante è quella che colloca l'errore *interlinguistico*, da un lato, e l'errore *intralinguistico* dall'altro. Con questi termini taluni indicano, rispettivamente, l'errore riconducibile a dissimmetria L1-L2 e l'errore derivante da ipercorrettismo o da errata generalizzazione nel sistema della L1. Si consideri, ad esempio, il seguente quesito:

"Volgere il verbo tra parentesi alla forma richiesta:

What... (happen) there after I had left?"

Se pensiamo all'italiano "Che cosa accadde là dopo che me n'ero andato?" osserviamo un'esatta simmetria tra *accadde* e *happened*, che è la risposta esatta. Poiché oltre la metà degli esaminati diede risposte errate quali *did happen* o *did it happen*, ci troviamo dinnanzi non a un caso di interferenza della lingua italiana, bensì o a un'indebita generalizzazione dell'uso dell'ausiliare, o a una falsa analogia con frasi come *When* (o *How*, *Where*, *Why*, *How often* ecc.) *did it happen?*

Sviluppando questa tematica, Selinker (1974) elabora il concetto di *interlingua*: l'acquisizione di una L2 avviene per una serie di successive approssimazioni, ciascuna delle quali costituisce un sistema linguistico completo. In altre parole, il discente che ha acquisito una conoscenza molto limitata di una lingua straniera, tende comunque a comportarsi come se quel bagaglio ristretto fosse *tutta* la L2; ciò spiega le generalizzazioni, gli ipercorrettismi, le associazioni incrociate, in breve, tutti i tipi di errore intralinguistico.

Il discente-tipo implicitamente postulato dal Lado sarebbe perciò una persona della quale vengono ignorati i costituenti affettivi della personalità, e per la quale l'apprendimento di una seconda lingua si traduce in un mero superamento delle dissimmetrie. Per comprendere questo grado di meccanicismo (che a noi, sulla base degli sviluppi successivi, può sembrare ingiustificabile) occorre ricordare entro quali orizzonti si collocava la linguistica strutturale americana del tempo. Il seguente brano di C. C. Fries può essere illuminante: "La condanna vigorosa dell'uso del significato, come è stato impiegato negli studi linguistici della generazione che ci ha preceduto, tende a far sì che il solo menzionare la parola *significato* sia anatema presso alcuni linguisti attuali. Affermare che uno studio linguistico odierno è "basato sul significato" equivale a porre sotto accusa l'integrità

¹³ Cfr. in A. DAVIES (a cura di), 1968, pp. 12-13 e cap. 4.

¹⁴ Anche se l'obiettivo è la *competence* (in senso chomskyano), molti studiosi dubitano che i test possano essere altro che accertamenti della *performance*.

scientifica dell'autore di quello studio. Il ripudio globale ed acritico di qualsiasi rilevanza del “significato” nell'attività dei linguisti sembra essere diventato l'atteggiamento convenzionale di alcuni di coloro che sono reputati tra i maggiori linguisti” (Fries, 1952, p. 293). Le critiche successive hanno ampiamente messo in luce come questo ripudio del significato fosse stato, in realtà, realizzato solo parzialmente, con numerose incoerenze, incertezze e sbandamenti. Ciò tuttavia non ha particolare rilievo nell'economia del nostro discorso: se la glottodidattica era identificata alla linguistica applicata, che a sua volta mutua indirizzi e strumenti dalla linguistica generale, il fatto che quest'ultima fosse orientata a prescindere dalla semantica (e anzi, per molti aspetti, a contrapporsi ad essa) è di per sé estremamente significativo.

Il Lado autore di *Language Testing* opera in un periodo in cui la linguistica è giunta ad elaborazioni sufficientemente ampie, a “certezze” abbastanza solidamente acquisite, solo nel campo della fonologia e della morfosintassi. È quindi del tutto coerente che egli si sia mosso in tali direzioni, anche se ciò ha condotto ad una “parcellizzazione” della verifica dell'apprendimento, con quesiti spesso privi di contestualizzazione, o in cui il significato assume un ruolo accessorio e incidentale. Eccone un tipico esempio:

“Lo stesso problema [il contrasto tra /u:/ e /u/] testato con un quesito a parole multiple e che coinvolge altri contrasti vocalici. Istruzioni: 'Confrontare i suoni rappresentati dalle lettere in neretto. Scrivere i numeri dei suoni che sono uguali al primo'. Quesito:

1	2	3	4	5	6	7	8	9
WOOD	FOOD	LOAD	WALK	UNDERSTOOD	CUT	SHOULD	GOOD	RUDE

Risposta: 1 5 7 8.

Lo stesso quesito con le parole inserite in un contesto:

1	2	3	4	5
THE	WOOD	FOR THE	FOOD	WAS CUT AS IT SHOULD. WE WALKED WITH A
6	7	8	9	
GOOD	LOAD	AS UNDERSTOOD	AND WERE NOT	RUDE.

Risposta: 4 6 8”. (Lado 1961, p. 98).

Notiamo anzitutto che la frase del “contesto” significa: *La legna per il cibo fu tagliata come si doveva. Camminammo con un buon carico, come inteso, e non fummo maleducati.* Sembra legittimo domandarsi se davvero la frase significhi qualcosa e quindi fornisca un verosimile contesto alle parole. La questione più importante è tuttavia un'altra, e cioè se il quesito sia idoneo ad accertare la capacità di discriminare i suoni /u:/ e /u/ tra di loro e da altri simili, attraverso la lettura silenziosa da parte dell'esaminato. Anche ammettendo che le validazioni ipotizzate dal Lado per queste “tecniche di produzione parziale” (Lado, 1961, p. 96) siano risultati positivi, qual è l'effettiva incidenza della discriminazione fonologica sull'abilità di comprendere e di parlare? In *Language Testing* non troviamo alcuna risposta precisa; nel capitolo dedicato al “controllo globale della lingua” si ipotizza un test in 12 parti, la prima delle quali è “Consonanti e vocali. Venti quesiti”.¹⁵ A questo test carta-matita potrebbe affiancarsi un test di comprensione orale.

Certo, non mancano gli accenni ad un superamento dei limiti che il Lado si autoimpone per operare con rigore metodologico. In particolare sono significativi i capitoli sulla verifica della comprensione interculturale e dei “valori superiori” mediante test. Il richiamo del binomio lingua/cultura è in sé un germe per il superamento dello strutturalismo; un superamento che, in questa direzione, si attuerà solo con lo sviluppo dell'analisi sociolinguistica.

¹⁵ Cfr. LADO, 1961, p. 272 (su un totale di 200 quesiti).

2. I fondamenti delle attuali tendenze nel “language testing”

In questo capitolo passeremo in rassegna (sia pure per ampie sintesi) gli apporti che le recenti acquisizioni della linguistica, sociolinguistica e metodologia didattica hanno offerto al language testing. In una fase, come la attuale, caratterizzata dalla ricerca di nuovi indirizzi, di risposte più valide ed articolate, di strumenti di verifica che assolvano a funzioni estremamente delicate, l'interrogarsi sugli orientamenti, i fini e i metodi del LT appare molto più importante del semplice possesso di un bagaglio di tecniche di controllo.

2.1. La grammatica generativo-trasformatzionale

L'approccio chomskyano è caratterizzato dalla postulazione di un parlante-ascoltatore ideale, capace di riconoscere e di produrre “frasi ben costruite” e di esprimere giudizi di grammaticalità/non grammaticalità degli enunciati. Sono intenzionalmente esclusi i tratti paralinguistici, le “false partenze”, le interiezioni ecc., interpretati o come elementi marginali della *performance* o (come nel caso delle false partenze, interruzioni, anacoluti, esitazioni, balbettii ecc.) come deviazioni dalle frasi ben costruite. Ai fini dell'analisi generativo-trasformatzionale (in senso stretto) è irrilevante il fatto che anche le “mezze frasi”, le esclamazioni, le stesse sgrammaticature intenzionali, possano essere prodotte deliberatamente dal parlante per comunicare (dare ordini, esprimere stati d'animo, o altro ancora) e venire comprese perfettamente dall'ascoltatore.

È stato osservato¹⁶ che, dati i fini propri perseguiti dalla linguistica generativo-trasformatzionale, non si è avuta da essa una filiazione diretta di uno specifico metodo di insegnamento della L2; né, sia ben chiaro, ciò rientrava negli obiettivi di Chomsky e degli altri linguisti. A parte qualche contributo marginale e indiretto (ad esempio, la distinzione tra frasi *nucleari* e *trasformazioni*), l'apporto offerto alla linguistica applicata (e, tramite questa, alla glottodidattica) si colloca su un piano più ampio ed elevato. La linguistica affronta finalmente in modo organico il problema della sintassi e, riconoscendo una dicotomia tra le strutture di superficie e le strutture profonde, studia i processi che conducono da queste a quelle, con ciò recuperando a pieno titolo la componente semantica che lo strutturalismo aveva messo al bando. E da qui prendono le mosse, come naturale sviluppo o come reazione critica, gli indirizzi odierni della linguistica.

2.2. La linguistica del testo

A grandi linee, l'evoluzione degli studi linguistici si muove lungo un itinerario che procede dal fonema¹⁷ al testo, dai segmenti elementari della catena parlata ad unità sempre più ampie, fino a superare i confini della singola frase allo scopo di cogliere gli elementi di *coesione* e *coerenza* testuale. Entrambi gli aspetti assumono notevole importanza anche per quanto si riferisce all'apprendimento di una seconda lingua. Sul versante recettivo, è indispensabile saper cogliere non solo il significato “di superficie” dei singoli enunciati, ma anche le connessioni tra di essi - sequenze temporali, periodi ipotetici, relazioni causa/effetto ecc. - le implicazioni e presupposizioni e, in breve, tutto ciò che il parlante intende comunicare (non esclusi i propri atteggiamenti: ironia, disgusto, irritazione, adulazione...).

L'attenzione al testo e alla *ridondanza* ha condotto ad un profondo riorientamento della didattica delle lingue straniere. Mentre nell'ambito dello strutturalismo si mirava ad un rigoroso controllo di tutte le strutture e di tutto il lessico a cui lo studente era esposto, in questa ottica appare più produttivo mettere lo studente a contatto con testi autentici, non “addomesticati”, affinché sviluppi quelle strategie che gli consentono di cogliere le informazioni necessarie anche in presenza di

¹⁶ Cfr., tra gli altri, G. FREDDI, *Metodologia e didattica delle lingue straniere*, Minerva Italica, Bergamo, 1970, p. 187.

¹⁷ Tra le più remote formulazioni del concetto è notevole quella di D. JONES, *The Phoneme: its Nature and Use*, Heffer, Cambridge, 1898.

parole o strutture nuove o di disturbi vari (quali, ad esempio, i rumori d'ambiente nell'ascolto, o gli errori e le mende tipografiche nella lettura). Come ricorda Van Ek, “la capacità di... non lasciarsi disorientare da elementi non conosciuti... è forse uno dei fattori più potenti nell'apprendimento delle lingue straniere” (Van Ek, 1976, p. 22).

La tematica delle tecniche di lettura è stata di recente oggetto di particolari attenzioni, soprattutto in rapporto alle esigenze di coloro che si accostano a settori specialistici e desiderano acquisire la capacità di comprendere la *microlingua* del settore che loro interessa professionalmente.¹⁸ Si pone ora in rilievo non la *linearità* del testo scritto, da leggere in sequenza dalla prima parola all'ultima, bensì l'uso di tecniche di *scanning* e *skimming*: con questi due termini inglesi, il cui uso si va diffondendo anche da noi, ci si riferisce all'enucleazione di una singola informazione da un testo, e alla lettura “in diagonale” per cogliere il concetto-base sviluppato in ciascun paragrafo.

Sul versante della produzione, l'obiettivo è la capacità di sviluppare testi (scritti e orali) *coerenti* sia al loro interno, sia soprattutto con il contesto situazionale, e *coesi*, ossia legati tra loro mediante tutti i necessari nessi cataforici e anaforici, nelle loro diverse realizzazioni: pronominalizzazione, ripetizioni, parallelismo, sinonimia, antonimia ecc. Di conseguenza, non basta più produrre “frasi ben costruite”, ma occorre mirare a produrre “testi efficaci”, cioè capaci di realizzare l'intenzione comunicativa del parlante. Come ben si intuisce, e come vedremo diffusamente più oltre, questo riorientamento negli obiettivi comporta un preciso mutamento di rotta anche nel testing.

2.3. La competenza comunicativa

Mentre la linguistica generale allargava il proprio campo d'indagine fino ad abbracciare i rapporti testo <--> senso, la sociolinguistica registrava importanti sviluppi sia in termini epistemologici, di autodefinizione come disciplina, sia di ricchezza e produttività delle analisi “sul campo”. Attraverso il recupero delle intuizioni di Firth e Malinowski¹⁹ e l'elaborazione della tematica delle *funzioni comunicative* (soprattutto sulla scorta del modello di Jakobson, 1966), il problema di “chi parla quale lingua a chi e quando” (Fishman, 1972, pp. 15-32) è stato esplorato in diverse direzioni. Tra le nozioni più importanti elaborate in sociolinguistica, quella che più direttamente interessa la glottodidattica è probabilmente quella di *registro*. Saper comunicare in L2 significa anche saper scegliere, tra le diverse opzioni, la più idonea sulla base dei livelli di formalità/non formalità, deferenza /non deferenza, familiarità/non familiarità ecc., connessi ai ruoli dei parlanti: da pari a pari (con ulteriori differenze tra adulto-adulto e bambino-bambino), da superiore a inferiore e viceversa, da adulto a bambino e viceversa, da uomo a donna e viceversa (soprattutto in alcune lingue e culture), e così via. Le locuzioni inglesi *Thank you very much* e *Thanks a lot* sono esponenti della stessa microfunzione “ringraziare qualcuno”, ma in un diverso registro, formale il primo, colloquiale/informale il secondo. Poiché in un testo si ritrovano normalmente enunciati appartenenti allo stesso registro, è possibile una verifica controllando (per restare nell'esempio) che *Thank you very much* sia associato a *Goodbye* e *Thanks a lot* a *Byebye*, e non viceversa.

¹⁸ Cfr., G. CORTESE (a cura di). *La lettura nelle lingue straniere: aspetti teorici e pratici*. Franco Angeli, Milano, 1980; Gruppo Lend Torino (a cura di). *Testi di orientamento e materiale bibliografico* (documentazione preliminare al Convegno Lend sull'Educazione alla lettura, Martinafranca, 22-24/X/1980); il tema della lingua “per scopi speciali (o accademici)” è ora diffusamente trattato: tra i primi *readers* citiamo S. HOLDEN (a cura di), *English for Specific Purposes*, Modern English Publications, 1977.

¹⁹ J. F. FIRTH, *Papers in Linguistics 1934-1951*, Oxford University Press, Londra, 1957; B. MALINOWSKY, “The Problems of Meaning in Primitive Languages”, in C. K. OGDEN, I. A. RICHARDS (a cura di), *The Meaning of Meaning*, Kegan Paul, Londra, 1923 (trad. ital. *Il significato del significato*. Il Saggiatore, Milano, 1966); cfr., G. FREDDI, *Didattica delle lingue moderne*. Minerva Italica, Bergamo, 1979, cap. 1.

Un'ulteriore, importante verifica riguarda la conoscenza delle restrizioni: il “saper tacere” o il “sapere che cosa non dire” accanto al “sapere che cosa dire, quando e come dirlo”. Insegniamo ai nostri allievi a dire “*Quelle âge avez-vous?*” confidando che essi non andranno, per fare sfoggio del loro francese, a porre la domanda alla prima signora francese che incontrano. Se ciò avvenisse, essi dimostrerebbero di possedere una *competenza linguistica* non associata ad un'adeguata *competenza comunicativa*. La corretta interpretazione degli enunciati non può prescindere dalla comprensione della loro funzione. A questo proposito si cita spesso l'esempio costituito dalla frase: *Hai dei gettoni del telefono?* Le risposte più probabili sono *Sì, tieni!* oppure *No, mi dispiace* (con le possibili varianti o aggiunte quali ... *ma li vendono all'edicola qui di fronte*). Nessuna di queste risposte è compatibile, di norma, con la domanda *Hai dei fratelli?*, domanda apparentemente analoga alla precedente; ma mentre la prima, con un oggetto inanimato, è esponente di “chiedere per avere”, la seconda, con un oggetto riferito ad esseri umani, è esponente del “chiedere per sapere”. L'eventuale incapacità di cogliere la differenza pregiudica il corretto progredire della conversazione.

La conseguenza è che l'obiettivo della verifica non può limitarsi all'accertamento della capacità di produrre enunciati formalmente corretti in sé, prescindendo dal contesto situazionale (*competenza linguistica*), né tantomeno della capacità di formulare “regole” che descrivano il sistema della L2 (*competenza metalinguistica*); l'oggetto del controllo deve essere la capacità di impiegare la L1 nell'interazione con parlanti nativi e/o testi autentici. In altre parole, dell'allievo importa conoscere “che cosa sa fare con la L2”. Ovviamente, questo obiettivo può essere raggiunto solo per gradi; per questo motivo sarà utile esaminare gli orientamenti didattici oggi prevalenti.

2.4. L'approccio funzionale-semantic

Wilkins (1976) opera una distinzione tra strategie glottodidattiche di tipo *sintetico* e di tipo *analitico*. Le prime comportano un insegnamento graduale dei singoli elementi del linguaggio, e l'acquisizione è un processo di accumulazione delle parti fino a ricomporre, nelle fasi finali, il sistema complessivo della L2.

Negli approcci analitici manca invece la preselezione del materiale linguistico a cui il discente viene esposto. Il suo compito è di approssimarsi quanto più possibile al comportamento linguistico richiesto dalle operazioni che viene chiamato a compiere. A tal fine, potrà essere necessario focalizzare di volta in volta l'attenzione su alcune forme linguistiche, ma il corso è essenzialmente basato sullo sviluppo delle funzioni comunicative, che a loro volta si estrinsecano in categorie semantico-grammaticali. Questo procedere dalla lingua globale ai suoi componenti lessicali e morfosintattici giustifica la definizione di “approccio analitico”; le considerazioni sulla struttura della L2 (in termini di complessità o di dissimmetria dalla L1) sono secondarie, ai fini della scelta del materiale linguistico da presentare agli studenti, -rispetto alla individuazione di obiettivi linguistici in termini operazionali. In altre parole, se sviluppare la capacità di “chiedere per avere” è ritenuto di primaria importanza, allo studente verrà presentato il materiale linguistico appropriato, prescindendo dall'eventuale difficoltà costituita dal condizionale (*vorrei..., mi darebbe...? ecc.*).

In realtà, come osserva lo stesso Wilkins, non esistono corsi che realizzino *in toto* un approccio esclusivamente analitico o sintetico, ma si ha un *continuum* lungo il quale ogni corso si colloca verso l'uno o l'altro polo. Tuttavia, poiché “anche quando abbiamo descritto il significato grammaticale (e lessicale) di una frase, non abbiamo spiegato il modo in cui è usata come enunciato” (Wilkins, 1976, p. 10), non si possono non esprimere delle riserve nei riguardi degli approcci sintetici. L'orientamento oggi prevalente è verso corsi che prendono le mosse dalle funzioni comunicative più frequenti e più strettamente collegate ai bisogni degli studenti. Per ciascuna di esse viene favorita l'acquisizione degli esponenti linguistici più accessibili, salvo poi ampliarla in seguito con strutture sempre più complesse, in un procedere a spirale che accresce e diversifica le competenze dei discenti.

2.5. I livelli-soglia

“Poiché l'inglese è la lingua internazionale dell'aviazione, la conoscenza di una decina di parole ed espressioni permette di girare il mondo: *departures, arrivals, check-in, passport control, ticket, boarding pass, duty-free shop, fasten seat belts, no smoking* e, purtroppo con sempre maggior frequenza, *delayed* e *cancelled*”. L'affermazione è chiaramente paradossale e la realtà non è così semplice, soprattutto se qualcosa non fila liscio. Tuttavia, ipotizzata una situazione ben precisa, è possibile indicare con sufficiente accuratezza i bisogni linguistici delle persone che vi sono coinvolte. Gli studi promossi dal Consiglio d'Europa per definire il livello-soglia²⁰ muovono da un presupposto analogo: dall'osservazione del comportamento adulto in termini di interazioni sociali essenziali, si possono desumere i livelli minimi di padronanza linguistica per svolgere le funzioni comunicative indispensabili: presentarsi, chiedere e dare informazioni, comprendere avvisi e istruzioni ecc.

La definizione di livello-soglia e i risultati delle ricerche (sia pure in certa misura divergenti per il francese e l'inglese) hanno favorito la produzione di corsi destinati ad adulti basati su un approccio comunicativo. Al tempo stesso sono stati avviati altri studi e ricerche per rispondere ad esigenze diverse. Ad esempio, una grande impresa industriale del nord Europa ha commissionato alla English Language Teaching Development Unit (ELTDU) della Oxford University Press uno studio sugli obiettivi dei corsi di inglese collegandoli con i compiti che i dipendenti della società sono chiamati a svolgere in inglese. Ne è risultata una scala dei livelli di padronanza (*stages of attainment scale*) in cui al livello A corrisponde la capacità di svolgere operazioni elementari, quali “capire ed eseguire semplicissime istruzioni”, mentre al livello H corrisponde una padronanza della lingua assimilabile a quella di un parlante nativo.²¹ I livelli A e B, tipici dello studente principiante e pre-intermedio, sono al di sotto del livello-soglia; in altre parole, ciò che egli sa fare con la lingua straniera non è sufficiente per svolgere alcuno dei compiti che gli si possono effettivamente richiedere sul posto di lavoro. Ad esempio, per ciò che si riferisce alla capacità di comprendere e parlare nei rapporti con clienti e fornitori, come minimo è necessario ricevere o trasmettere ordinazioni, in termini di quantità, prezzi, date di consegna delle merci ecc. Una persona al livello C è in grado di fare questo solo limitatamente alla normale routine, ma nel caso di richieste inattese o di informazioni complesse deve richiedere l'intervento di qualcun altro.

Tenendo conto che i livelli e gli obiettivi devono essere esplicitati separatamente per le diverse abilità (comprendere/parlare, leggere, scrivere), per le diverse situazioni operative (rapporti col pubblico, collaborazione coi colleghi di lavoro ecc.), e in rapporto ad altri parametri che potrebbero risultare importanti nei casi specifici, si rischia facilmente di moltiplicare le definizioni dei livelli ed i criteri di giudizio. Il pericolo è però teorico più che reale, in quanto l'esigenza stessa di redigere test per verificare l'appartenenza ai vari livelli conduce naturalmente ad accorpare le “mansioni linguistiche” in unità organiche e coerenti. Un sano empirismo è spesso la miglior risposta alla tentazione di sottilizzare al di là delle effettive esigenze di precisione della verifica.

3. Il ruolo del testing

Anche se nel precedente capitolo abbiamo cercato di tener distinte le varie istanze, da uno sguardo d'insieme emerge la circolarità nei rapporti tra i diversi momenti: dai bisogni agli obiettivi, all'intervento didattico, alla verifica, alla ridefinizione degli obiettivi, e così via, ciclicamente, fino al soddisfacimento dei bisogni (o, come accade più spesso, all'abbandono dello studio della lingua

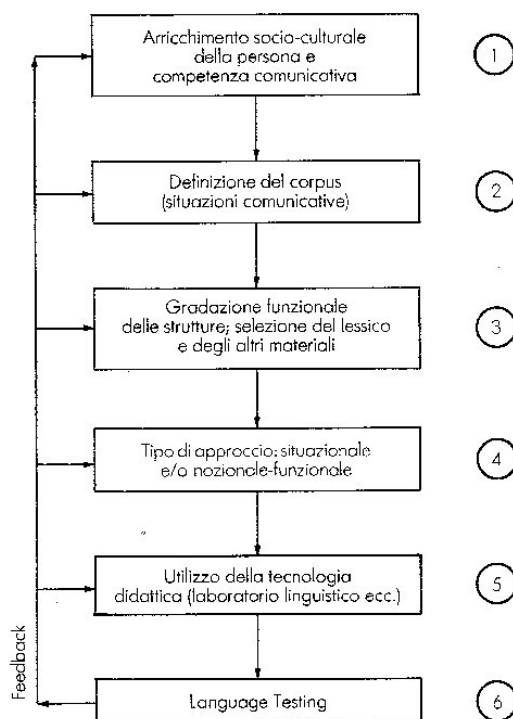
²⁰ Per una visione d'insieme della tematica del “livello soglia”: *Un système européen d'unités capitalisables pour l'apprentissage des langues vivantes par les adultes. Articles extraits de la Revue “Education et Culture”*, n. 28, Consiglio d'Europa, Strasburgo, 1978.

²¹ Cfr. C. YATES, “What objectives?”, in S. HOLDEN (a cura di), *English for Specific Purpose*, cit., p. 47.

prima che esso abbia prodotto i frutti desiderati). Cercheremo ora di sviluppare questo tema, ponendo l'accento sul ruolo che il testing è chiamato a svolgere nelle varie fasi.

3.1. Aspetti della programmazione curricolare

Nell'articolo citato (v. nota 12) lo Zani rielabora il diagramma di Pellerey adattandolo allo specifico della didattica delle lingue.



Nel diagramma va notato l'andamento delle frecce relative all'effetto retroattivo (*feedback*): gli esiti della verifica si riversano su ciascuna delle fasi del ciclo, in termini di ridefinizione degli obiettivi, dei contenuti, delle strategie e sequenze di apprendimento, e quindi della scelta dei metodi e degli strumenti.

La collocazione del LT nell'ultimo riquadro non deve perciò trarre in inganno: i ruoli e le funzioni del testing sono molteplici e non limitati all'accertamento finale. Per comodità di esposizione possiamo distinguere le verifiche in *iniziali*, *intermedie* e *finali*.

3.2. Il controllo della situazione iniziale

Vi sono tre aspetti distinti delle verifiche iniziali:

- a) l'analisi della situazione di partenza nell'ambito di una ricerca sperimentale. L'esito di una sperimentazione deve essere valutato sulla scorta di dati oggettivi riferiti sia alla situazione iniziale sia a quella finale, omogenei tra gruppo sperimentale e gruppo di controllo;²²
- b) l'accertamento delle acquisizioni pregresse. Iniziando un corso con studenti non principianti è necessario conoscere con precisione quanto è stato appreso e ritenuto, quali sono le lacune e le dimenticanze, quali settori devono essere recuperati e sviluppati. Sotto questo profilo, si parla spesso di *funzione diagnostica* dei test;

²² Cfr., R. TITONE, *Introduzione alla metodologia della ricerca nell'insegnamento linguistico/Methodology of research in language teaching*, ediz. bilingue, Minerva Italica, Bergamo, 1974.

c) un uso frequente dei test iniziali riguarda la suddivisione degli iscritti in gruppi omogenei per livelli di padronanza. Il ricorso alla verifica iniziale sistematica è utile sia che si intenda formare gruppi in cui sono equamente distribuiti i livelli di padronanza, sia che si vogliano raggruppare assieme gli appartenenti allo stesso livello per operare interventi didattici differenziati.

La verifica di tipo a) rientra nelle attività di pianificazione di una sperimentazione. La ricerca del termine *a quo* è indispensabile per formulare qualsiasi giudizio sul termine *ad quem*. In assenza di dati affidabili, la valutazione sul valore dell'intervento sperimentale si limiterebbe ad un giudizio soggettivo che, soprattutto se espresso dallo sperimentatore anziché da un osservatore indipendente, risulta viziato dal legame affettivo che lega ricercatore e oggetto della ricerca.²³

La verifica di tipo b) interessa qualsiasi insegnante che inizi a lavorare con una classe di non principianti. L'accertamento sistematico è molto importante se la classe è “nuova” per l'insegnante; ma anche nel caso di continuità didattica con l'anno precedente è utile accertare le variazioni intervenute durante le vacanze, sia in senso positivo (in seguito a viaggi all'estero, incontri con stranieri, o corsi integrativi), sia in senso negativo (dimenticanza dovuta a scarso esercizio). Disponendo di dati precisi, il “ripasso iniziale” generico e differenziato lascia il posto ad un lavoro di “recupero” tarato sugli effettivi bisogni presenti nella classe.

La verifica di tipo c) riguarda le situazioni in cui i gruppi non sono precostituiti sulla base di criteri diversi da quello della padronanza della L2. Vi sono valide ragioni sia a favore della costituzione di gruppi uniformi, sia contrarie. La scelta si baserà sulle mete del corso, sulla disponibilità di tempo, sussidi e materiali ecc. In ogni caso il controllo accurato della situazione iniziale assume un ruolo fondamentale.

Abbiamo accennato al valore diagnostico della verifica iniziale. In quanto alla funzione *prognostica*, diremo soltanto che essa è di pertinenza non tanto dei test di L2 in quanto tali, quanto piuttosto dei test di *attitudine linguistica*. Questi ultimi, che normalmente prescindono dalla conoscenza di una specifica lingua, sono gli unici che possono essere somministrati ai principianti assoluti.²⁴

La trattazione adeguata del problema dell'attitudine, con profonde implicazioni psicolinguistiche, richiede ben altro spazio rispetto a quello che qui possiamo dedicarvi, e perciò rimandiamo il lettore alle opere citate nella nota precedente.

3.3. Le verifiche intermedie

I metodi recenti elaborati in glottodidattica presentano, come caratteristica comune, un'articolazione in *unità didattiche* (UD).²⁵

I modelli di UD man mano proposti e realizzati si differenziano in varia misura, a seconda dei presupposti teorici (strategie), dei destinatari dei materiali (adulti o bambini, principianti, intermedi o avanzati ecc.), della quantità e tipo dei supporti audiovisivi disponibili, e altro ancora. Per l'aspetto che ci riguarda, tuttavia, l'orientamento è concorde e può essere così formulato:

²³ Sull'uso corretto di termini quali *sperimentazione* e *ricerca* si veda anche R. TITONE, *Insegnare oggi le lingue seconde: breviario di glottodidattica*, SEI, Torino, 1979, p. 227 segg.

²⁴ Un'esposizione critica del problema dell'attitudine linguistica (con ampi riferimenti alle ricerche di J. B. CARROL) si trova in R. TITONE, *Psicolinguistica Applicata*, Armando, Roma, 1971, e nelle altre opere già citate dello stesso Autore; segnaliamo inoltre la versione italiana (a cura di M. FERENCICH) di J. B. CARROLL, S. M. SAPON, *Modern Language Aptitude Test*, The Psychological Corp., New York, 1959 (ediz. ital. Organizzazioni Speciali, Firenze, 1964).

²⁵ Cfr., R. TITONE, *Insegnare oggi...*, cit., cap. 5; G. FREDDI, opere citate e *Fondamenti e metodi della didattica delle lingue*. Ed. Cafoscarina, Venezia, 1972, conversazioni VIII e IX.

Se una UD costituisce un nucleo organico di materia da apprendere, e se un corso si articola su un numero adeguato di unità in sequenza appropriata, è di primaria importanza che si verifichi l'acquisizione (nel senso più ampio del termine) dell'UD precedente prima di affrontare la successiva.

È questo il ruolo del testing intermedio, la cui collocazione naturale è nella fase finale di ogni unità. Se questa esigenza è spesso disattesa, ciò si deve o al persistere del “compito in classe” mensile o alla difficoltà, per l'insegnante, di reperire test idonei (o di crearne autonomamente). Siamo convinti che quanto più i docenti saranno sensibili al problema del testing, tanto più indurranno gli editori a corredare i materiali didattici con schede di valutazione. Anche se ciò avverrà, tuttavia, gli insegnanti non potranno esimersi dall'impadronirsi dei principi e delle tecniche di testing, sia per giudicare dell'applicabilità *nelle loro classi* dei materiali offerti, sia per produrre i testi integrativi idonei ad operare le verifiche specificatamente richieste dalla propria attività didattica.²⁶

Le verifiche intermedie non vanno confuse con il continuo riflusso di informazioni che l'insegnante riceve sull'andamento della classe ogni volta che gli studenti partecipano attivamente alla lezione; non solo, per i motivi già esposti, è importante tener distinto il momento dell'esercizio da quello del controllo, ma è anche necessario ricorrere a momenti formali e sistematici di verifica per minimizzare l'effetto di alone.²⁷

3.4. Le verifiche finali

La prima domanda da porsi è: quanto finali sono le verifiche *finali*? Si pensi alle prove di accertamento al termine della classe prima media: da un lato non sono altro che una verifica intermedia nell'ambito di un corso triennale sostanzialmente unitario; dall'altro lato concorrono, assieme ai giudizi sull'andamento nelle altre materie, a decidere sulla promozione dello studente. Ci si trova di fronte all'intersecarsi di istanze e modalità diverse, le une interne al corso di lingua straniera, le altre esterne ad esso ma interne all'istituzione in cui il corso si esplica. Una prima risposta va quindi ricercata osservando l'utilizzazione che si intende fare dei risultati della verifica.

Una seconda domanda riguarda l'oggetto della verifica: che cosa è il profitto in lingua straniera? Sinora non abbiamo usato il termine *profitto* a causa della possibile ambiguità di interpretazione. Bisogna infatti distinguere tra l'acquisizione delle abilità e dei contenuti relativi a un dato corso, materiale o testo, e il raggiungimento di un dato livello di padronanza della lingua, verificato prescindendo dall'itinerario didattico percorso per giungervi. In inglese si parla rispettivamente di *achievement* e di *proficiency* che potremmo rendere in italiano (in mancanza di, termini migliori) con *realizzazione* e *padronanza*.²⁸

²⁶ Si obietterà che questa esigenza è generale, non limitata alle verifiche intermedie. Ne abbiamo parlato qui per sottolineare come i test debbano essere presenti nella routine quotidiana. Per le verifiche iniziali e finali, e nelle sperimentazioni, è anche ipotizzabile il ricorso ad esaminatori esterni ed esperti di valutazione; ciò non può (e non deve) avvenire nella valutazione formativa che accompagna passo passo l'attività di insegnamento.

²⁷ Per *effetto di alone* si intende 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” in E. SPALTRO, *La psicologia del lavoro*, Etas-Kompass, Milano, 1967, p. 207.

²⁸ In A. DAVIES (a cura di), 1968, a p. 6 si afferma che “the terms *achievement*, *attainment*, *proficiency*, *diagnosis* and *aptitude* are often confused”. A. AMATO (pp. cit., pp. 14-15) traduce *achievement* con “istruzione”, *proficiency* con “profitto” e *attainment* con “realizzazione”. Sui termini *aptitude* “attitudine” e *diagnosis* “diagnosi” non sembra esservi controversia.

A chiarimento dei termini, il Davies (1968, p. 7) propone il seguente schema nel quale X rappresenta il test e le frecce indicano la direzione temporale (il passato a sinistra, il futuro a destra) verso cui si rivolge il test:

Realizzazione: <———— X
 Padronanza: <- - - - - X ———>Y
 Attitudine: (X)————>X
 Diagnosi: <———— X ———>

I test di realizzazione si occupano delle acquisizioni pregresse in relazione ad un dato programma di studio. Si tratta di test *specifici* in senso tecnico, ossia riferiti a particolari materiali didattici.²⁹

I test di padronanza sono solo indirettamente interessati al passato, sono per la loro natura *aspecifici* e spesso sono orientati verso il futuro o verso l'esterno, in quanto verificano la capacità di usare la lingua per impieghi particolari (simboleggiati con Y nello schema).

Un test di attitudine è proiettato verso il futuro e si basa sulle abilità linguistiche attuali rilevabili in L1 o comunque, normalmente, in lingua diversa da quella da apprendere. Questa difformità è rappresentata dalle parentesi.

I test diagnostici sono tipicamente bidirezionali, analizzando il passato per impostare il futuro. Essi comunque costituiscono categoria a sé: possiamo parlare di realizzazione, di padronanza e di attitudine da parte degli studenti, ma di diagnosi da parte dell'insegnante. Ciò che caratterizza i test diagnostici è l'uso che se ne fa, non l'oggetto della verifica.

Un terzo problema si riferisce all'ampiezza e al livello della verifica nei test di padronanza. Quanto deve estendersi il campo di indagine? Nella tassonomia citata alla nota precedente si riconoscono *test fattoriali*, rivolti al controllo di un solo elemento del sistema linguistico della L2; *test di abilità*, aventi per oggetto ognuna delle quattro abilità linguistiche di base; *test globali di lingua*, intesi a verificare la competenza linguistica generale attraverso la interrelazione delle abilità integrate; *test di lingua-civiltà* che mirano a rilevare il livello di comprensione interculturale (oltre alla competenza linguistica globale).

I test fattoriali trovano la loro collocazione e impiego come subtest, soprattutto nei test di realizzazione, mentre i test di padronanza si configurano essenzialmente come test sulle abilità integrate, e la “nuova frontiera” nel testing non va più nella direzione della comprensione interculturale (o, almeno, non direttamente), ma piuttosto verso la verifica della competenza comunicativa.

3.5. Il controllo della situazione didattica

Nei tre paragrafi precedenti abbiamo esaminato il ruolo dei test nel controllo dell'apprendimento della L2 da parte degli studenti, ponendo l'accento sul successo o insuccesso nel conseguimento dei livelli di realizzazione e/o di padronanza desiderati. Focalizzando l'attenzione sulle attitudini, motivazioni e capacità dei discenti si rischia facilmente di dimenticare che altre due variabili sono contemporaneamente sotto controllo:

- a) l'efficacia dell'intervento didattico nei suoi diversi aspetti e componenti: competenza linguistica e didattica degli insegnanti, validità dei materiali e sussidi, disponibilità di tempo adeguato e utilizzazione ottimale delle “risorse” ecc.;
- b) l'idoneità degli strumenti di accertamento (a meno che non si disponga già di test validati).

²⁹ La dicotomia *test specifici/test aspecifici* è discussa in G. PORCELLI, 1975, pp. 35-36.

Supponiamo, per il momento, che la verifica sia stata effettuata rispettando tutte le garanzie di oggettività, affidabilità e validità. I risultati ottenuti si prestano ad una duplice lettura:

individuale, riferita alle acquisizioni di ciascuno studente;

collettiva, da collegare alla situazione del gruppo-classe.

Per interpretare i dati relativi all'intero gruppo-classe occorre tabulare adeguatamente i punteggi analitici ottenuti nelle somministrazioni del test; così facendo si possono calcolare:

- 1) la media e la mediana³⁰ per ciascun subtest; questi dati indicano in quali parti del test il gruppo, nel suo complesso, è riuscito a far meglio e in quali peggio. L'analisi dei quesiti³¹ conduce ad individuare, per ciascuno di essi, gli indici di facilità e di discriminazione; in rapporto a questi parametri è possibile giudicare più correttamente il “peso specifico” degli eventuali errori, che a loro volta saranno ricondotti ai problemi di apprendimento che evidenziano;
- 2) la deviazione standard, un indice statistico che rivela il “ventaglio” dei livelli presenti nella classe: quanto più è alto, tanto maggiore è la disparità delle forze in campo e la differenza di prestazioni tra i migliori e i peggiori;
- 3) le correlazioni con altri parametri riferiti all'intero gruppo-classe.

I dati grezzi così ottenuti debbono essere “filtrati” attraverso opportune griglie che ne consentano la corretta interpretazione. Non è agevole stabilire con sicurezza a quale fattore si debba attribuire l'eventuale insuccesso di un programma, a meno che l'indagine non sia condotta nell'ambito di una precisa procedura sperimentale.³² A questo proposito occorre rammentare che un test (anche se particolarmente sofisticato, sperimentato e prestigioso) non è che uno strumento di *misurazione* e di accertamento; l'utilizzazione corretta dei dati ricavati (ai fini della *valutazione* degli allievi e della situazione didattica) è demandata al giudizio esperto e meditato dei docenti. Si può tentare un'analogia con i test clinici: un termometro che funzioni a dovere fornirà dati sulla temperatura corporea molto più precisi e affidabili delle impressioni soggettive del paziente, o dell'idea approssimativa che si può avere posando una mano sulla fronte. Tuttavia il termometro non ci dirà mai nulla sulle cause della febbre e sul quadro clinico in cui questo sintomo si colloca: la diagnosi resta compito del medico.

4. I test pragmatici

Oller definisce “test pragmatico... qualsiasi procedura o compito per cui il discente affronta sequenze di elementi, in una lingua, conformi alle normali restrizioni contestuali di quella lingua, e che richieda al discente di correlare le sequenze di elementi linguistici con il contesto extralinguistico mediante strategie pragmatiche” (Oller, 1979, p. 38). I due requisiti essenziali sono:

- l'impiego di lingua *autentica* e *naturale*, coerente e coesa, rispettosa di tutte le norme d'uso, registro ecc.;
- il costante riferimento al *valore comunicativo* del dialogo o brano in questione (registro, valore illocutivo - atteggiamento e intenzione del parlante - richiami al contesto situazionale, implicazioni, presupposizioni ecc.).

³⁰ Porcelli, 1975, cap. X.

³¹ Porcelli, 1975, cap. IX.

³² Cfr. R. TITONE, *Introduzione alla metodologia della ricerca nell'insegnamento linguistico*, cit. Un esempio di griglia si trova anche in G. PORCELLI, “Un approccio all'analisi dei libri di testo”. *Lingue e Civiltà*, a. IX, n. 2, aprile 1981, pp. 3-6.

Un test pragmatico è quindi sostanzialmente diverso da un test integrato: questo cerca di ricomporre in unità i frammenti ricavati attraverso test fattoriali e quesiti a uscita unica, mentre in un approccio pragmatico si fa riferimento all'analisi del discorso in testi adeguati per contenuto e dimensioni.

Il dettato e la “cloze procedure” sono, secondo l'analisi di Oller, i due tipi più caratteristici di test pragmatici. In entrambi i casi, chi è sottoposto alla prova deve mettere in opera strategie adeguate per la comprensione del testo (orale e scritto, rispettivamente). Abbiamo già fatto cenno ai concetti di ridondanza, coerenza e coesione; essi portano a concepire la capacità di comprensione e di lettura come abilità di formulare previsioni corrette sul contenuto e sullo sviluppo del testo e di procedere alla “riduzione di incertezza” attraverso il rigetto delle opzioni impossibili o improbabili e la scelta delle alternative più verosimili. Tale scelta viene operata tenendo conto sia delle restrizioni operanti nel codice della L2 (a livello fonologico, grafemico, morfosintattico e lessicale), sia del contesto (linguistico ed extralinguistico), delle connessioni logiche, delle informazioni e conoscenze possedute sull'argomento. Se una persona mi sta parlando della Polonia, anche se mi sembra di aver sentito “Dublino” (o anche se il mio interlocutore ha effettivamente detto “Dublino” a causa di un lapsus), rigetterò questa ipotesi in favore di quella, assai più probabile, secondo cui egli ha detto (o intendeva dire) “Lublino”. La mia eventuale ignoranza del dato geografico mi impedirebbe la comprensione corretta dell'enunciato.

4.1. Il dettato

Se si usano testi tratti da brani di lingua autentica, e se questi vengono presentati in sequenze “abbastanza lunghe per mettere alla prova la memoria a breve termine degli studenti, il dettato tradizionale risponde ai requisiti di naturalezza dei test pragmatici. Anzitutto, tale compito richiede l'elaborazione di sequenze di materiale nella lingua con restrizioni temporali e, in secondo luogo, il compito di suddividere la catena parlata e trascrivere ciò che si è udito richiede la comprensione del significato del materiale, ossia il porre in relazione il contesto linguistico (che, in un certo senso, è dato) con il contesto extralinguistico (che deve essere inferito)” (Oller, 1979, p. 39). Mentre Lado e gli altri strutturalisti consideravano il dettato essenzialmente come test di ortografia - e, per di più, poco significativo - oggi si pone l'accento sugli aspetti attivi e creativi del comprendere; il fatto che occorrono ulteriori ricerche per accertare quali operazioni compiano esattamente coloro che scrivono sotto dettatura non sembra infirmare la validità complessiva della prova. Anzi, il coinvolgimento simultaneo di più operazioni mentali e procedimenti è ora considerato un chiaro vantaggio rispetto ai test miranti ad accertare un elemento alla volta.

Oller sostiene che nell'elaborazione dei punteggi relativi a un dettato non si dovrebbe tenere conto degli errori puramente ortografici, quali quelli che a volte si riscontrano negli scritti di nativi. Tra gli esempi presentati troviamo *yung* per *young*, *directer* per *director*, e *abillity*, *abilitie*, *abilaty* per *ability*. Sono invece considerati “errori fonologici”, rispettivamente, *lung*, *dractor* e *a meladic* (Oller, 1979, p. 276 segg.).

Pur riconoscendo la difficoltà che a volte si incontra nel discriminare tra i due tipi di errore, Oller è favorevole a “depurare” i punteggi dei dettati dagli errori di mera ortografia; la sua argomentazione si basa principalmente sul fatto che la capacità di scrivere in modo ortograficamente corretto presenta, nelle ricerche effettuate sull'argomento, un bassissimo indice di correlazione con le altre abilità linguistiche. In altre parole, il livello di padronanza dello *spelling* ci dice nulla o ben poco sui livelli di padronanza del lessico e della morfosintassi, o sulle capacità di comprendere e leggere.

Questo atteggiamento è probabilmente influenzato, in certa misura, dalla tipica difficoltà, per i nativi, di acquisire la grafia corretta in inglese; per un italiano o uno spagnolo (o anche, sia pure con modalità diverse, per un francese o un tedesco) i termini del problema si configurano in modo da condurre, di solito, a conclusioni opposte. È significativo che tra gli “errori ortografici” Oller includa *fathers* per *father's*. Così facendo, mostra di non distinguere affatto tra *ortografia lessicale* (non regolata da norme identificabili) e *ortografia grammaticale*, codificata in morfologia. Altro è

scrivere *sieze* per *seize* (o *seige* per *siege*), e altro è scrivere *longer* invece di *long*, o *studys* per *studies*. Le desinenze sono di notevole importanza in tutte le lingue: perché siano riprodotte correttamente occorre cogliere tutte le relazioni sintagmatiche (concordanze di genere, numero, persona; tempo dei verbi; declinazione di sostantivi, aggettivi e determinanti ecc.). Queste relazioni sintagmatiche costituiscono il tessuto connettivo del discorso; soprattutto nei casi di omofonia, la corretta riproduzione ortografica attesta l'esatta comprensione dell'enunciato. Si pensi alla lingua francese: *ami/amis/amie/amies*; *vien/viens/vient/viennent*; *aimer/aimez* ecc. Il solo criterio fonologico a cui si attiene Oller non ci pare sufficiente a render conto degli errori ortografici; una simile impostazione contraddice la considerazione che egli ha del dettato come test pragmatico. Se la comprensione del significato deve essere il criterio fondamentale, né *fathers* né *father* ci dicono se lo studente ha capito la presenza di un caso possessivo in *father's*. Considerare il primo come errore ortografico (in virtù dell'omofonia?) ed il secondo come errore fonologico costituisce un'arbitraria semplificazione del problema.

In termini operativi, è difficile offrire indicazioni che soddisfino alle varie istanze, le quali sono in certa misura contraddittorie. Se da un lato la distinzione tra i vari tipi di errore può offrire un quadro più preciso sulle prestazioni dei singoli studenti, essa presenta varie difficoltà:

- non è sempre possibile accertare se si tratti di errore sistematico o occasionale (cfr. par. 1.4 e nota 14);
- la definizione di una tipologia degli errori non è, di norma, compito agevole; in essa rifluiscono principi attinti dalla linguistica generale ed applicata, dalla psico- e socio-linguistica, dalla glottodidattica;
- definita una tipologia, occorre classificare correttamente gli errori; in questa fase si presentano molti casi incerti e l'attribuzione è in certa misura soggettiva (*studys* potrebbe essere classificato come errore di ortografia o di morfologia a seconda del contesto, ma anche a seconda del valore che ciascun esaminatore attribuisce alle “regole ortografiche”);
- altrettanto soggettiva è l'attribuzione di un “peso specifico” agli errori di ciascun tipo; l'ortografia lessicale può avere “peso zero” (come suggerisce Oller), ma può anche essere ritenuta importante sulla base del lavoro svolto con un certo gruppo di studenti, o in relazione agli obiettivi del corso o ad altri criteri esterni.

Oltre al rischio della soggettività - che può peraltro essere contenuto entro limiti accettabili attraverso un'accurata e puntuale definizione dei criteri e parametri a cui si è fatto cenno - questa procedura presenta lo svantaggio di richiedere tempi più lunghi per la correzione degli elaborati. Molto spesso - se la prova di dettato non ha uno scopo diagnostico, se fa parte di una prova finale ai fini di una valutazione sommativa, e se i protocolli da correggere sono numerosi - gli esaminatori adottano una procedura più grezza, ma più spedita: il semplice conteggio delle parole non riprodotte correttamente. Parafrasando Getrude Stein si direbbe che ci si affida al principio per cui *word is a word is a word*; chi scrive *phragment* invece di *fragment* viene penalizzato allo stesso modo di chi non ha compreso la parola e ha lasciato lo spazio bianco ovvero ha scritto tutt'altro vocabolo.

Questo approccio è giustificabile presupponendo un'equa distribuzione dei tipi di errore e una progressione parallela del numero e nella gravità degli errori. In altre parole, se si riscontra che chi fa pochi errori commette in genere errori lievi (o semplici sbagli occasionali), mentre gli errori gravi, le lacune e le omissioni si accompagnano a numerosi errori di vario tipo, ne deriva che anche il dato grezzo (il semplice conto degli errori) è un valido indicatore del livello di padronanza degli esaminati. Per accertare che così avvenga in realtà, occorre valutare un congruo numero di dettati adottando entrambe le procedure e calcolare la correlazione esistente tra le due serie di risultati. Se questa fosse molto alta, potremmo inferire che la procedura più semplice è sufficientemente esplicativa e pertanto il ricorso alla procedura più complessa può risultare superfluo in molti casi.

L'accertamento della correlazione, a nostro avviso, deve essere effettuato per ciascun livello (principianti, intermedi, progrediti), tipo e grado di scuola (media inferiore, liceo, istituto tecnico o professionale, facoltà universitaria): è infatti possibile che popolazioni scolastiche diverse presentino variazioni anche sensibili sotto questo profilo; ulteriori indagini dovrebbero poi esplorare eventuali differenziazioni collegabili a situazioni regionali o locali.

Indipendentemente dai risultati di queste ricerche - da cui è lecito attendersi una migliore comprensione dei fattori che interagiscono in una prova di dettato - va sottolineato che l'uso regolare del dettato come prova di controllo contribuisce a dare il giusto rilievo all'abilità del saper comprendere la lingua straniera, un'abilità che tradizionalmente è la più trascurata tra le quattro fondamentali. L'impiego del dettato non sarà di per sé sufficiente, perché non può sostituire gli altri aspetti della comprensione così come essa si configura nelle normali situazioni quotidiane (dialogo, ascolto di radio, TV, registrazioni, comunicati all'altoparlante ecc.); tuttavia è certamente utile, purché siano rispettati i requisiti di naturalezza già esaminati. Ad esempio, dettare una parola alla volta è gravemente scorretto, innaturale, distorce ritmo, intonazione e giuntura, e costituisce una facilitazione solo apparente: ad un'analisi attenta, si rivela un'attività non solo inutile, ma controproducente, in quanto abitua gli studenti ad un ascolto non realistico e preclude un valido accostamento ad un ascolto in situazioni assimilabili a quelle della comunicazione autentica.

4.2. La “cloze procedure”

Si attribuisce a Wilson L. Taylor (Aller, 1979, p. 341) l'introduzione del termine *cloze* (omofono del verbo inglese *close* “chiudere”), in una serie di studi sulla leggibilità dei testi e sulla misura della comprensione; il termine richiama il concetto di “chiusura” così come è stato elaborato nell'ambito della psicologia della forma (*Gestalttheorie*). Nel percepire, cogliamo prima l'insieme (la “forma”, appunto) e poi i dettagli; questi ultimi, anzi, acquistano significato solo nell'ambito dell'unità maggiore. Se qualcosa ostacola la nostra percezione, riusciamo tuttavia a ricostruire mentalmente l'oggetto intero (che può essere un'immagine visiva incompleta, un oggetto parzialmente visibile, un messaggio disturbato dal rumore ecc.) completando gli elementi percepiti. Per far ciò ci serviamo della nostra esperienza: più lo stimolo (visivo, uditivo, olfattivo, tattile ecc.) ci è familiare, tanto più semplice è ricostruire nella sua interezza ciò di cui abbiamo potuto cogliere solo qualche dettaglio.

Un testo linguistico a cui sono state tolte delle parole costituisce un problema analogo: la conoscenza del codice e dell'argomento consente di reintegrarlo. Si consideri la frase seguente: “Il tema _____ fondo [nel romanzo postumo di Silone] è quello capitale dello scrittore, la natura della _____.”³³ Per riempire il primo spazio con la preposizione *di* ci basta la conoscenza della lingua italiana e non occorre che spingiamo lo sguardo più in là del contesto immediato (i vocaboli *tema* e *fondo*). Per individuare la seconda parola omessa, al contrario, dobbiamo non solo capire il significato globale della frase, ma anche conoscere il ruolo che la tematica della *fede* (questa è, infatti, la parola mancante) svolge nelle opere di Silone. Se la frase non fosse stata enucleata dall'articolo, ma fosse stata presentata nell'ambito di un testo più esteso, la soluzione esatta si sarebbe potuta inferire attraverso l'analisi testuale. La frase successiva è: “Come credere, è possibile credere, in che cosa e dove dobbiamo fare convergere le nostre aspirazioni religiose?” Qui il verbo *credere* riprende ed esplica l'*aver fede*. Se il testo da ricomporre è sufficientemente lungo, la ricostruzione si basa sugli elementi di ridondanza, coesione e coerenza testuale necessariamente presenti, e solo in misura ridotta sulla conoscenza dell'argomento trattato.

Sarà bene sottolineare la differenza tra i *cloze test* e i test di grammatica o di lessico basati sulla tecnica del completamento. Se da un brano vengono tolte sistematicamente, ad esempio, le preposizioni, la ricostruzione avviene essenzialmente sulla base della conoscenza delle norme di uso relative ai diversi complementi, ovvero della conoscenza degli usi idiomatici; la comprensione

³³ C. Bo, “Severina o la libertà”. *Corriere della sera*, 13.10.1981, p. 3.

del significato non può certo essere assente, tuttavia essa assume un ruolo per molti aspetti ancillare nel processo di risoluzione. Questo fa sì che un test di completamento così strutturato non presenti il requisito della naturalezza e pertanto non possa essere assimilato ai test pragmatici.

Paradossalmente, la tecnica che più facilmente porta a realizzare un *cloze test* vario e valido è quella più semplice e meccanica, che richiede di eliminare sistematicamente una parola ogni *n* vocaboli del testo. In genere si lascia intatta la prima frase (nei testi più lunghi, due o tre frasi, o l'intero paragrafo introduttivo, se è breve), per fornire sufficienti dati per la comprensione. Il valore di *n* è di norma compreso tra 5 e 10, ed è scelto in rapporto alla lunghezza e complessità del testo ed al livello di difficoltà desiderato. Lasciare meno di quattro parole tra uno spazio e il successivo rende estremamente complessa la prova e può facilmente condurre ad interpretazioni alternative plausibili, ma sostanzialmente difformi dal testo originale. Lasciarne più di nove porta all'utilizzazione di brani molto lunghi per verificare pochi elementi, e quindi si abbassa la produttività del test. Con studenti a livello intermedio o avanzato, si rivelano in genere utili brani di circa 360-380 vocaboli. Togliendo una parola ogni sette si hanno 50 items, che rappresentano una campionatura sufficientemente estesa e, al tempo stesso, un numero molto pratico per l'elaborazione dei dati. Rispetto alla meccanicità delle eliminazioni, il solo intervento di chi costruisce il test dovrebbe limitarsi ad evitare che la stessa parola sia tolta troppe volte; in altre parole, anziché eliminare per la terza volta un certo articolo (o verbo ausiliare, o congiunzione ecc.) si opterà sistematicamente per la parola successiva. Di norma, così facendo si constaterà che tra le parole da reintegrare vi è la stessa percentuale di vocaboli a prevalente funzione lessicale³⁴ esistente nel testo completo. Solo se si dovesse riscontrare una notevole sproporzione (ad esempio, un eccesso di *function words*) varrebbe la pena di provare a realizzare una forma alternativa; ad esempio, invece di togliere la 10a, 15a, 20a... parola, si toglieranno la 11a, 16a, 21a ecc. Con la stessa tecnica è anche possibile costruire "forme parallele" dello stesso test.

Per quanto riguarda la correzione, sono possibili due tecniche. La prima richiede che venga individuata esattamente la parola esistente nel testo originale, rigettando qualsiasi alternativa. Con la seconda procedura occorre invece esaminare le soluzioni alternative per vedere quali siano accettabili. Verranno ritenuti validi non solo i sinonimi dei vocaboli esatti, ma anche tutte le parole in qualche modo equivalenti e comunque compatibili col testo. Questa procedura richiede la definizione precisa dei criteri di accettabilità - ed è quindi soggetta, in certa misura, alle decisioni del correttore - ma è la sola applicabile quando il test ha funzione prognostica o diagnostica, ed è importante l'analisi degli errori. Se invece la finalità del test è quella di ottenere un punteggio da assumere come dato-indice, la prima procedura presenta il vantaggio della rapidità ed affidabilità; l'indice di correlazione tra i punteggi (o tra le graduatorie) che si ottengono applicando le due procedure a campioni adeguati è un indice molto alto e quindi non c'è molto da guadagnare in precisione e significatività dei risultati seguendo la strada più lunga e difficile.³⁵

³⁴ Le cosiddette "parole piene": sostantivi, aggettivi, verbi principali, avverbi di modo, tempo e luogo ecc.; il termine inglese *function words* designa invece le "parole vuote": articoli, pronomi, verbi ausiliari, congiunzioni ecc.

³⁵ Cfr., J. W. OLLER Jr., *op. cit.*, p. 367 segg.; B. J. CARROLL, *Testing Communicative Performance*, Pergamon Press, Oxford, 1980, p. 96 segg. In un nostro studio (non pubblicato) su 48 protocolli, abbiamo ottenuto un indice di correlazione pari a 0,967 usando la formula di Spearman (*rank-difference method*) e pari a 0,978 con la formula di Pearson (*product-moment correlation*). Chiaramente questi indici molto alti si devono al fatto che le risposte esatte costituiscono un sotto-insieme delle risposte accettabili.

Non tutti i ricercatori concordano sulla validità dei *cloze* test. Osserva B. J. Carroll: “Questo [approccio] è ancora basato essenzialmente sulla competenza linguistica.³⁶ La procedura non rappresenta un'autentica comunicazione interattiva ed è perciò soltanto un indice indiretto dell'efficienza potenziale nell'affrontare la comunicazione quotidiana” (Carroll, 1980, p. 9-10). In effetti, il compito di ricostituire un testo è in certa misura artificioso, anche se - come già osservato - ricalca parzialmente le strategie di lettura. Inoltre, mentre la capacità di leggere presuppone una conoscenza “passiva” del lessico, la capacità di affrontare prove ricostruttive richiede anche la padronanza di un adeguato vocabolario “attivo”. Ciò rappresenta un vantaggio se si preferiscono le prove “integrate”, che coinvolgono più abilità; al tempo stesso, tuttavia, non vi può essere una superficiale assimilazione dei *cloze* test con le prove mirate a verificare esclusivamente la comprensione lessica.

Pur accogliendo queste riserve soprattutto come monito ad approfondire le indagini sulla natura e le implicazioni della procedura *cloze*, si può concludere affermando che allo stato attuale delle conoscenze in materia essa appare una delle più promettenti e valide tra le tecniche di testing praticabili. In particolare:

- a) la semplicità di costruzione, somministrazione e correzione è favorevole alla diffusione di questa procedura;
- b) i risultati dei *cloze* test si correlano positivamente con quelli dei dettati e dei test strutturali;
- c) assumendo come criterio il giudizio globale sulla competenza comunicativa in L2, il dettato e la procedura *cloze* (nell'ordine) presentano i più alti indici di correlazione (Carroll, 1980, p. 97);
- d) secondo Titone, “un contributo positivo allo sviluppo della comprensione lessica... consiste nell'organizzare piccoli gruppi strutturati di discussione, dove i testi vengono dibattuti dopo averli letti in silenzio. Lo scopo di tali attività è di incoraggiare una lettura dei testi in forma ricostruttiva (*cloze reading*) in un'atmosfera non competitiva. In tale attività ricostruttiva svolta in gruppo, certe parole vengono cancellate da un brano e il compito dei lettori è quello di studiare il testo e arrivare a una visione consensuale della parola più adatta a riempire gli spazi vuoti” (Titone, 1981, p. 46). Quanto più si diffonderanno tecniche analoghe - con gli adattamenti opportuni e necessari - anche nell'insegnamento della lettura in lingua straniera tanto maggiore sarà la validità di impostazione dei *cloze* test.

4.2.1. Varianti della procedura “cloze”

Oller (Oller, 1979, pp. 44-50) accenna ad un certo numero di test di tipo pragmatico, molti dei quali rappresentano varianti della procedura *cloze* o combinazioni tra questa ed altre modalità di verifica.

Tra essi segnaliamo:

la procedura cloze orale: il brano è presentato in forma orale, con pause al posto degli spazi bianchi; l'esaminato deve fornire (oralmente o per iscritto) le parole omesse;

la combinazione di dettato e cloze: l'esaminato ha davanti a sé il brano da completare e contemporaneamente ascolta la lettura integrale del brano, il che facilita notevolmente il compito di ricostruzione;

procedura cloze facilitata: vengono lasciate le lettere iniziali delle parole omesse, oppure le parole stesse sono ricostruibili in base a un criterio noto agli esaminati, così da restringere il campo delle scelte possibili.

³⁶ Abbiamo reso con “competenza linguistica” il termine *usage*, che si contrappone a *use* “competenza comunicativa”. Per un'ampia e più puntuale illustrazione dei termini v. H. G. WIDDOWSON, *Teaching Language as Communication*, Oxford University Press, 1978, cap. 1.

Non occorre procedere ulteriormente nell'esemplificazione per illustrare la flessibilità della procedura e la sua adattabilità ai vari livelli di apprendimento e alle diverse situazioni didattiche.

4.2.2. Varianti del dettato ed altri test pragmatici

Oltre alla combinazione di dettato e *cloze*, ricordata sopra, vi sono altre possibili varianti del dettato:³⁷

- *il dettato parziale*: alcune parti del brano da dettare sono presentate anche per iscritto all'esaminato, mentre altre (paragrafi, frasi, sintagmi o parole isolate) dovranno essere ricostituite;
- *il dettato-composizione o "dicto-comp"*: un brano viene letto una o più volte agli esaminati, i quali devono poi riprodurlo il più fedelmente possibile usando i medesimi vocaboli, le medesime costruzioni e colmando le eventuali lacune della memoria con parole proprie;³⁸
- *il dettato di numeri*, che elimina il problema ortografico (gli allievi devono scrivere solo delle cifre) ed esplora una delle aree semantiche verso la quale si riscontrano forti resistenze psicologiche, data l'enorme difficoltà per chiunque di fare calcoli in una lingua non materna, anche se ben conosciuta.³⁹

Altre varietà di dettato ("focalizzato", "di frasi" ecc.)⁴⁰ non vengono qui considerate in quanto si tratta di attività utili come esercizi per l'apprendimento delle lingue estere, ma difficilmente utilizzabili come prove di controllo.

Sulla base della definizione "aperta" che egli dà di *test pragmatico*, Oller tende ad includere in questa categoria molte di quelle che avevamo chiamato "prove tradizionali": il rispondere a questionari, il colloquio orale, la composizione scritta, la narrazione orale, e perfino la traduzione (o "almeno alcune varianti di essa" [Oller, 1979, p. 50]). Egli riconosce tuttavia la difficoltà di individuare le forme e i modi di un impiego strutturato in maniera confacente alle esigenze del testing. A proposito della traduzione sostiene che occorrono ulteriori ricerche: un'affermazione con la quale è difficile non essere d'accordo.

5. La verifica del "saper comunicare"

Altri orientamenti attuali prendono le mosse non tanto da un'analisi linguistica - sia pure orientata in senso pragmatico ed estesa all'analisi testuale - quanto piuttosto dall'analisi dei bisogni dei discenti in ordine alla competenza comunicativa.

5.1. Le micro-abilità

Uno dei presupposti per il controllo sistematico della capacità di comunicare è l'individuazione precisa e dettagliata dei fattori che sono implicati nello svolgimento di un "atto comunicativo" qualsiasi. Si supponga di voler determinare se uno studente straniero conosca la lingua 2 abbastanza bene per potersi servire della biblioteca universitaria allo scopo di consultare testi scientifici riguardanti la materia in cui intende specializzarsi. Per far ciò dovrà "saper leggere" e "saper consultare" (ad esempio, uno schedario, od altra forma di catalogo). A loro volta, queste abilità linguistiche si articolano in una serie di *micro-abilità*, ed è solo facendo riferimento a queste ultime che possiamo uscire dal vago e dall'indeterminato, decidere quali siano essenziali nella situazione ipotizzata e quali no, e accertare che siano state acquisite. Fra le tassonomie delle abilità, una delle

³⁷ Cfr., T.M. SPEER, "Putting variety into dictation", *English Teaching Forum*, a. XVIII, n. 3, luglio 1980, p. 28.

³⁸ P. M. RILEY, "The dicto-comp", *English Teaching Forum*, a. X, n. 1, gennaio-febbraio 1972, p. 21.

³⁹ R.V. WHITE, *Teaching Written English*, Allen & Unwin, Londra, 1980, pp. 99-100.

⁴⁰ Cfr., T.M. SPEER, *op. cit.*

più dettagliate ed esaurienti è quella proposta da J. Munby (1978). Tra i molti punti che si ricollegano al saper leggere, ne riportiamo due a titolo esemplificativo:

30. Capire relazioni tra parti di un testo attraverso i seguenti elementi di coesione lessicale:

- 30.1 ripetizione
- 30.2 sinonimia
- 30.3 iponimia
- 30.4 antitesi
- 30.5 apposizione
- 30.6 insiemi lessicali / “collocazioni”
- 30.7 forme sostitutive/nomi generali

32. Capire relazioni tra parti di un testo attraverso i seguenti elementi di coesione grammaticale:

- 32.1 referenza (anaforica e cataforica)
- 32.2 comparazione
- 32.3 sostituzione
- 32.4 ellissi
- 32.5 indicatori di tempo e di luogo
- 32.6 connettivi logici.

È appena il caso di osservare come questi due punti si riferiscano alla *coesione* testuale, la quale a sua volta è una soltanto delle componenti che intervengono nella comprensione di un testo (non necessariamente scritto). Se a qualcuno l'analisi apparisse eccessivamente frammentata, non dimentichi che:

I. Solo conoscendo esattamente fattori, meccanismi ed elementi operativi è possibile cercare di comprendere le modalità della comunicazione linguistica, come premessa essenziale a qualsiasi processo valutativo.

II. Una certa difficoltà (almeno iniziale) ad avvalersi di tassonomie di questo tipo è da collegare alla scarsa dimestichezza con queste problematiche anche in rapporto alla L1; l'uso di classificazioni puntuali può, al contrario, costituire una base comune per la determinazione degli obiettivi e dei criteri di valutazione sia nella lingua materna che in quella straniera, in un'ottica interdisciplinare.

L'uso - anzi, spesso, l'abuso - che è stato fatto di termini quali “bisogni, obiettivi, curricula” in anni recenti deve indurre ad una rigorosa selezione degli strumenti operativi affinché le istanze di programmazione didattica non restino parole vuote. Una buona tassonomia delle microabilità, integrata in un'organica visione del LT in un'ottica comunicativa, costituisce uno degli strumenti indispensabili per uscire dal vago e dall'indeterminato.

5.2. I livelli di competenza comunicativa

Accanto all'analisi dei bisogni e a quella, intimamente connessa, delle microabilità, si rende necessaria la definizione dei livelli di competenza comunicativa. In altre parole, occorre individuare una serie di parametri ai quali rapportare le prestazioni dell'esaminato, in termini non di “come sa” la lingua straniera, bensì di “che cosa è in grado di fare” servendosi di essa. Brendan J. Carroll (1980) ha sviluppato una “scala” di nove punti, di cui presentiamo qui le specificazioni in ordine alla valutazione globale.

9. *Utente esperto*. Comunica con autorevolezza, accuratezza e stile. Completamente a proprio agio nell'inglese⁴¹ idiomatrico e specialistico.

8. *Ottimo utente*. Presentazione dell'argomento chiara e logica con discreto stile e sensibilità agli indicatori di atteggiamento. Si approssima spesso ad una competenza bilingue.

7. *Buon utente*. Sa affrontare la maggior parte delle situazioni in un ambiente di lingua inglese. Lapsus occasionali e limitazioni nel linguaggio non impediscono la comunicazione.

6. *Utente competente*. Sebbene sappia far fronte alla maggior parte delle situazioni in cui può trovarsi, è un po' carente nella fluidità e precisione di eloquio e presenterà occasionali fraintendimenti o errori significativi.

5. *Utente modesto*. Sebbene riesca in genere a comunicare, usa spesso un linguaggio scorretto o improprio.

4. *Utente marginale*. Mancandogli stile, fluidità e precisione, non è facile comunicare con lui e vi sono fraintendimenti dovuti all'accento straniero e all'improprietà nell'uso della lingua. In genere riesce a cavarsela senza gravi inconvenienti.

3. *Utente estremamente limitato*. Non possiede una conoscenza operativa della lingua ai fini dell'uso quotidiano, ma è meglio di un principiante assoluto. Né le abilità ricettive né quelle produttive consentono una comunicazione continuativa.

2. *Utente intermittente*. Prestazioni ben al di sotto del livello di una conoscenza operativa della lingua dell'uso quotidiano. La comunicazione si realizza solo sporadicamente.

1/0. *Non-utente*. Può addirittura non essere in grado di riconoscere con sicurezza quale lingua si stia usando.

Le altre griglie elaborate da Carroll e presentate nel volume citato riguardano la valutazione del colloquio e del saper scrivere per fini "accademici" (tesi, relazioni, sintesi, appunti ecc.). Il valore principale di simili strumenti è quello di fornire una cornice onnicomprensiva nell'ambito della quale collocare qualsiasi utente della L2, dal principiante assoluto al bilingue assimilabile a un parlante nativo. All'interno di ciascuna "scala" debbono poi essere individuati i livelli *proredito*, *intermedio* e *minimo*, sempre in termini di competenza comunicativa. A tal fine vengono indicati dieci criteri: *dimensioni* (del patrimonio linguistico attivo), *complessità*, *varietà*, *velocità*, *flessibilità*, *accuratezza*, *proprietà*, *autonomia*, *ripetizione* ed *esitazioni*.

Ne deriva un sistema "a due ranghi"; all'interno di ciascun livello si ha un'assegnazione di punteggi (secondo modalità che per brevità non possiamo illustrare qui), i quali però vengono successivamente interpretati alla luce della loro significatività in termini comunicativi.

5.3. Nuovi criteri per nuovi test

Alla luce delle considerazioni sviluppate in questa quinta sezione, appare chiaro come alcuni requisiti che apparivano fondamentali in un testing di matrice strutturalista risultino ora marginali, mentre ne sono emersi altri la cui importanza è stata generalmente sottovalutata.

In primo luogo, si è passati da test riferiti alla norma a test riferiti al criterio. Nei primi, la prestazione di un individuo era interpretata (mediante il calcolo della media, della deviazione

⁴¹ Riteniamo che queste specificazioni possano essere applicate ad altre lingue senza necessità di modificazioni.

standard, dei *centili* ecc.)⁴² in rapporto alle prestazioni degli altri appartenenti allo stesso gruppo, o a quelle di un campione idoneo. Oggi si ritiene più utile comparare le prestazioni del singolo con obiettivi comportamentali prestabiliti. Solo in questo modo, ad esempio, è possibile superare l'ambiguità del concetto di “sufficienza”, che pure assume tanto rilievo nella prassi scolastica. Che significa affermare “Gigi Rossi è sufficiente”? Che le sue prestazioni sono normali per uno studente della sua età e classe, oppure che conosce la lingua (nativa o straniera, poco importa) in modo adeguato allo svolgimento di certi compiti comunicativi, chiaramente predeterminati? Si noti che nel linguaggio comune nessuno dice “Mille lire sono sufficienti” se non con riferimento preciso al costo di un certo bene o servizio, ossia a un *criterio*. La seconda istanza oggi sottolineata si riferisce all'*autenticità* dei materiali e delle prove di accertamento. Ciò significa accertare la competenza comunicativa degli esaminati sulla base di compiti realistici, che comportino una vera interazione, realizzata mediante il codice orale e/o il codice scritto e/o tutti gli altri elementi non-verbali della comunicazione, appropriati alle circostanze specifiche. Naturalmente vi sono delle difficoltà ad organizzare prove di questo tipo - in particolare per coloro che non dispongono di materiali didattici autenticamente interattivi. Tuttavia, o si accerta davvero la competenza comunicativa, o ci si accontenta di verificare la pseudo-comunicazione (in termini, ad esempio, di correttezza grammaticale, di lessico non contestualizzato, o di compiti che non hanno alcun riscontro nella realtà quotidiana extrascolastica).

Sempre seguendo il Carroll (1980), osserviamo come un buon test sia il risultato di una presenza equilibrata di quattro fattori:

Pertinenza

Accettabilità

Comparabilità

Economicità

(da cui l'acronimo “P.A.C.E.”)

La *pertinenza* si riferisce alla connessione tra il comportamento soggetto a verifica e il soddisfacimento dei bisogni comunicativi.

L'*accettabilità* è la capacità di un test di apparire valido agli esaminati, senza suscitare reazioni ansiose o atteggiamenti ostili.

Per *comparabilità* si intende la possibilità di ottenere risultati confrontabili con quelli ottenuti da gruppi diversi e/o in momenti diversi.

La *economicità* o *economia* di un test implica il conseguimento di un rapporto ottimale tra la quantità di informazioni ricavabili sugli esaminati e il costo in termini di tempo, fatica e materiali impiegati.

Nessuno dei quattro fattori è, di per sé, né interamente nuovo né in contrasto con quelli esaminati in precedenza; la *pertinenza* è, in effetti, la controparte della *validità* (con una sottolineatura del ruolo della competenza comunicativa), e la *comparabilità* è, in buona sostanza, la versione aggiornata dell'*affidabilità*. Tuttavia il quadro complessivo è significativamente innovatore, e oggi disponiamo di un'impostazione sicuramente più completa ed avanzata del problema del testing. In quanto alle

⁴² Una succinta ma chiara distinzione tra *norm-referenced* e *criterion-referenced tests* si trova in R. M. VALETTE, *op. cit.*, p. 11. Per un'analisi più ampia rinviarno a E. INGRAM, “Basic Concepts in Testing”, in J. P. B. ALLEN & A. DAVIES (a cura di), *Testing and Experimental Methods*, cit., pp. 26-32. Indicazioni sul significato del calcolo dei centili (e degli indici analoghi) con riferimento al LT vengono date in G. PORCELLI, *Il Language testing*, cit., cap. 10; R. LADO, *Language Testing*, cit., cap. 22; D. P. HARRIS, *Testing English as a Second Language*, McGraw-Hill, New York, 1969, p. 135 segg.

realizzazioni concrete, siamo ancora in una fase di pre-sperimentazione delle nuove tecniche; i test “funzionali” già in circolazione costituiscono una tappa intermedia tra quelli di impostazione strutturalista e la “nuova frontiera”. Uno dei problemi più ardui da superare non si riferisce tanto alla preparazione di materiali idonei a suscitare una comunicazione “autentica”, quanto piuttosto alla definizione di procedure abbastanza semplici per consentire un normale impiego scolastico, ma al tempo stesso sufficientemente sofisticate per poter utilizzare le griglie, le scale, i criteri e i parametri di valutazione più consoni alla determinazione della competenza comunicativa.

6. Osservazioni conclusive

In questo *excursus* ci siamo soffermati più sui principi e problemi che sulle tecniche di testing. Questa scelta - a prescindere da ogni considerazione sullo spazio necessario per presentare campioni di test di sufficiente ampiezza - si giustifica per la preminenza che una chiara visione del ruolo e della funzione del testing deve avere rispetto al mero uso di tecniche diverse dalle convenzionali. Non basta avere molte frecce al proprio arco (anche se particolarmente penetranti) se manca un bersaglio preciso verso cui indirizzarle. Analogamente, la padronanza di un certo numero di tecniche non è di per sé garanzia di un miglior accertamento del livello di competenza raggiunto dagli allievi.

Un quesito che viene posto frequentemente dagli insegnanti riguarda il rapporto tra i test e le altre prove di controllo (in particolare quelle previste dai vigenti programmi d'esame). I test debbono costituire *un'alternativa o un complemento*? Il quesito è mal formulato e richiede alcune puntualizzazioni.

Sotto il profilo docimologico, non c'è dubbio che uno strumento affidabile e validato debba essere preferito, ogni volta che ciò sia possibile, a mezzi più aleatori e meno probanti. Le procedure di elaborazione e interpretazione dei punteggi possono assicurare una utilizzazione critica, intelligente e non meccanicistica dei dati numerici grezzi.

Sul piano pedagogico, la varietà degli strumenti di verifica è un indubbio vantaggio. “Quando ho cominciato ad usare i test, mi sono resa conto che riuscivo a cogliere certi aspetti della preparazione della classe che prima mi sfuggivano”. Questa dichiarazione di una collega, in occasione di un seminario sul testing, è emblematica: finché adoperiamo un solo tipo di modalità di accertamento ricaviamo un solo tipo di dato. Al tempo stesso privilegiamo chi si trova a suo agio lavorando in un certo modo e danneggiamo gli altri, che potrebbero esprimersi meglio se le prove mettessero in gioco fattori diversi. Un sano eclettismo non va però confuso con l'incoerenza e il pressappochismo; il Guénou (1964) ammoniva che non dobbiamo essere come coloro che per cinque minuti adorano Gesù Bambino, e per altri cinque minuti il bue e l'asinello! Senza ripeterci, richiamiamo però con vigore l'istanza della coerenza e validità di impostazione, sul piano metodologico in senso lato.

Per ciò che attiene agli aspetti più propriamente didattici, occorre distinguere tra i test capaci di costituire prove di accertamento comprensive ed autonome, e quelli idonei soltanto a verificare alcuni aspetti settoriali della L2. Abbiamo già fatto cenno alla utilizzazione di questi ultimi come prove *preliminari* ad altre più complesse. Riteniamo che, anche alla luce degli orientamenti attuali, non vi sia nulla di intrinsecamente errato nel ricorrere a test focalizzati su aspetti specifici e nell'ambito della competenza linguistica; a patto, però, che si abbia precisa coscienza del loro valore limitato e strumentale, e non si assumano i risultati come indicativi di competenze a livello superiore. Né, d'altra parte, ci si può illudere che esistano test, per quanto ampi, complessi e ben articolati, in grado di risolvere ed esaurire il compito dell'accertamento globale del livello di padronanza linguistica dei nostri allievi. I test, come abbiamo cercato di mostrare, sono utili in tutte le fasi e gradi dell'insegnamento: è solo una questione di scelta oculata dei mezzi rispetto ai fini. Ma nessun test, in quanto tale, potrà mai sollevare l'insegnante dal compito di esprimere un giudizio globale qualitativo su ciascun allievo.

A qualcuno, a cui il compito di giudicare è particolarmente ingrato, ciò dispiacerà; ma davvero vorrebbe personalizzare e disumanizzare il rapporto educativo al punto di affidare ad altri un compito così delicato? Una *oggettività* assoluta non esiste: i test “oggettivi” sono tali soltanto nella determinazione dei punteggi, mentre è soggettiva la scelta dei contenuti e delle tecniche. Se per certi aspetti questo costituisce un limite spiacevole, in ultima analisi ci garantisce che un insegnante non potrà mai essere sostituito nemmeno dal più sofisticato dei computer.

Bibliografia

- Davies A. (a cura di), 1968, *Language Testing Symposium*, Oxford University Press, London. Volume parzialmente tradotto in Amato A. (a cura di) *Il testing nella didattica linguistica*, Bulzoni, Roma, 1974.
- Fishman J. A., 1972, *Who speaks what language to whom and when*, in Pride J. B. - Holmes J. (a cura di), *Sociolinguistics*, Penguin, Harmondsworth pp. 15-32.
- Fries C. C., 1952, *The Structure of English*, Longman, London.
- Guénot J., 1964, *Clefs pour “les langues vivantes”*, Seghers, Paris.
- Jakobson R., 1966, *Saggi di linguistica generale*, Feltrinelli, Milano.
- Lado R., 1961, *Language Testing: the construction and use of foreign language tests*, Longman, London.
- Munby J., 1978, *Communicative Syllabus Design*, Cambridge University Press, London.
- Oller J. W., 1979, *Language tests at school*, Longman, London.
- Porcelli G., 1975, *Il “language testing”*. Minerva Italica, Bergamo.
- Selinker L., 1974, *Interlanguage*, in J. C. Richards (a cura di). *Error Analysis*, Longman, London, pp. 31-54.
- Titone R., 1981, *Una tendenza emergente: la ricerca sulla lettura in lingua straniera in prospettiva psicolinguistica*, in “Rassegna Italiana di Linguistica Applicata”, Bulzoni, Roma, a. XIII, n. 1-2.
- Van Ek J. A., 1976, *The Threshold Level for Modern Language Learning in Schools*, Longman, London.
- Wilkins D. A., 1976, *Notional Syllabuses*, Oxford University Press, London.