

La progettazione di prove di produzione e interazione orale

in ambito di certificazione linguistica per scopi accademici

L. Bavieri, P. Polselli, A. Duguid

1. Introduzione

2. La produzione orale monologica

2.1 *Valutare le prove monologiche: problematiche e indicazioni della ricerca*

2.2 *Tipologia di test e ambito d'uso linguistico di riferimento*

2.3 *Tipologia di formati e di task*

3 Le scelte effettuate

3.1 *Tipi di task scelti e motivazioni*

3.2 *La scelta dell'argomento*

3.3 *Tempi*

3.4 *L'uso della tecnologia informatica*

4. La produzione orale interattiva

4.1 *Valutare le prove interattive: problematiche e indicazioni della ricerca*

4.2 *Tipologia di test*

4.3 *Tipologia di formati e di task*

5. Le scelte effettuate

5.1 *Tempi*

5.2 *L'uso della tecnologia informatica*

6. Bibliografia

1. Introduzione*

Il tratto distintivo della comunicazione orale rispetto alla codificazione scritta consiste in prima istanza nella sua modalità di trasmissione fisica e semiotica, cioè nell'utilizzo del canale fonico-uditivo, non solo dal punto di vista della produzione fonica ma anche relativamente a tutti gli aspetti comunicativi legati all'ideazione e alla progettualità nell'uso che si fa dei suoni. (De Mauro *et al.*, 1993). Ma, per una definizione più esaustiva del parlato, occorre aggiungere, accanto alla canalizzazione che gli è propria, altri parametri che Bazzanella (1994:12) denomina, in un prospettiva pragmatica, "tratti situazionali", ossia il contesto extralinguistico comune e la compresenza di parlante ed interlocutore/i.

È indubbio che, quando si pensa alla comunicazione orale, questa sia immediatamente associata al parlato dialogico e, infatti, la conversazione, cioè la situazione di libero scambio comunicativo tra due o più parlanti, è sicuramente la situazione comunicativa predominante. In quanto tale, essa può essere considerata il modello primario alla base di qualsiasi uso linguistico, la forma di comunicazione più naturale e spontanea utilizzata dai parlanti (De Mauro *et al., cit.*). Tuttavia la comunicazione orale non è univoca e le sue varietà sono legate a situazioni comunicative specifiche: dal contesto interazionale dipendono le modulazioni del discorso in generi, ovvero le variazioni funzionali del registro in relazione al tipo di rapporto che esiste tra parlante e interlocutore o tra parlante e ascoltatore (Bernardelli, Pellerey, 1999). Per arrivare a definire le diverse tipologie della comunicazione orale, si descriveranno dunque le varie situazioni comunicative che derivano dall'incrocio di diversi tipi di condizioni e di relazioni tra interlocutori: si fa qui riferimento alla modellizzazione definita nel *Lessico di frequenza dell'italiano parlato* che, rispetto agli altri corpora di parlato precedentemente pubblicati, inseriva per la prima volta testi monologici. Partendo dal genere conversazione, il LIP delinea una scala rappresentativa di una gamma di usi, tutti consolidati, da un massimo ad un minimo grado di naturalezza e distingue le diverse situazioni di dialogo e di parola rivolta a un pubblico incrociando 4 criteri generali (Bernardelli, Pellerey, *cit.*): scambio reciproco vs non-reciproco, dialogo in presenza vs a distanza, dialogo con presa di parola libera vs non libera, scambio tra pari grado vs non pari grado.

Si ottengono in questo modo 5 modelli base di situazione comunicativa (De Mauro *et al., cit.*):

1. scambio bidirezionale faccia a faccia con presa di parola libera (conversazione in tutte le sue possibili forme). Il corpus prende in considerazione le conversazioni in casa e le conversazioni sul luogo di lavoro, tra pari grado, tra non pari grado, con estranei, in uffici pubblici, banche, ecc.;
2. scambio bidirezionale non faccia a faccia con presa di parola libera (conversazioni telefoniche). Vengono inserite in questa tipologia le normali conversazioni telefoniche, quelle registrate alla radio, i messaggi registrati nelle segreterie;
3. scambio bidirezionale faccia a faccia con presa di parola non libera (dibattiti, interviste, interrogazioni ecc.). Si tratta in questo caso di assemblee legislative, dibattiti culturali, assemblee studentesche e sindacali, incontri di lavoratori, interrogazioni nella scuola elementare e secondaria, esami universitari, interrogatori processuali, interviste alla radio e alla televisione;
4. scambio unidirezionale in presenza di destinatario/i (lezioni, conferenze, omelie, comizi ecc.). Sono inserite nel corpus lezioni di scuola elementare, secondaria e universitarie,

* La stesura di questo contributo risale a settembre 2003. All'epoca, si inseriva nell'ambito del progetto C.E.R.C.L.U. (Certificazione dei Centri Linguistici Universitari) promosso da A.I.C.L.U. (Associazione Italiana Centri Linguistici Universitari). Si rimanda, per informazioni e dettagli, al sito: <http://www.aiclu.it/cerclu/index.htm>
Seppur concepito insieme, il contributo è opera di Luisa Bavieri per le sezioni 1, 2.3, 3, 3.1, di Paola Polselli per le sezioni 2, 2.1, 2.2, 3.2, 3.3, 3.4 e di Alison Duguid per le sezioni 4, 4.1, 4.2, 4.3, 5, 5.1, 5.2

relazioni a congressi o convegni politici, sindacali e scientifici, comizi politici, omelie, conferenze non specialistiche, arringhe giudiziarie;

5. scambio unidirezionale o bidirezionale a distanza (trasmissioni radiofoniche o televisive, che vengono messe all'ultimo posto della scala in quanto i comportamenti comunicativi che le caratterizzano come genere sono perlopiù costruiti e pianificati).

Per mettere a fuoco le differenze che intercorrono tra i diversi generi di discorso orale, è necessario ricorrere ancora una volta alla distinzione tra lingua parlata e lingua scritta, ponendo in questo caso come parametro di riferimento il rapporto tra lessico e grammatica: diversamente da quanto avviene nel testo scritto, dove il contenuto è espresso più dal vocabolario che dalla grammatica, il testo parlato è caratterizzato da una bassa densità delle parole lessicali e da un'alta densità delle parole grammaticali (Rossini Favretti, 2002). Nel LIP, dove viene sottolineata l'influenza che le diverse situazioni enunciative esercitano sulle strutture linguistiche, la distanza tra i modelli descritti è misurata sulla base dei fenomeni grammaticali citati più frequentemente come caratteristici dell'italiano parlato, ossia la riduzione del numero e della varietà dei connettivi e la semplificazione dei paradigmi verbali (De Mauro *et al.*, *cit.*).

Il processo di riduzione ha dimensioni diverse nei 5 gruppi di testi: mentre nei gruppi 1 e 2 si trovano connettivi ad alta frequenza caratterizzati da polifunzionalità sintattica e testuale, i testi dei gruppi 3, 4 e 5 presentano un'ampia gamma di connettivi sia per numero che per varietà; tale fenomeno è giustificato dal ritmo veloce e dai tempi brevi di progettazione dello scambio comunicativo delle conversazioni faccia a faccia e telefoniche, che favoriscono l'utilizzo di parole altamente disponibili con ampia flessibilità funzionale. La presenza di introduttori di frase e segnali di articolazione è segnalata invece in tutti i gruppi e, data l'alta frequenza e la mancanza di corrispettivi negli usi scritti, essa viene considerata un tratto tipico del parlato nel suo complesso, sia dialogico che monologico (*ibid.*).

Per quanto riguarda invece il fenomeno della contrazione dei paradigmi verbali, nella fattispecie l'espansione dell'indicativo presente a spese del futuro semplice e del congiuntivo presente, i gruppi 1 e 2 rappresentano il polo di maggior riduzione, laddove il polo di minor riduzione è rappresentato dai gruppi 4 e 5.

Considerato dunque che i gruppi 1, 2 e 3 sono costituiti da testi bidirezionali e i gruppi 4 e 5 da testi prevalentemente unidirezionali e che i fenomeni linguistici analizzati presentano le stesse caratteristiche ai due poli della scala di naturalezza, gli autori giungono alla conclusione che è possibile individuare nella presenza o assenza di turni di presa di parola il vero punto discriminante tra i vari generi di parlato. Ciò significa che i testi con la stessa struttura comunicativa presentano, indipendentemente da altri fattori, analogie e somiglianze, ossia che è più probabile rilevare lo stesso comportamento linguistico in un dialogo piuttosto che in un dialogo e in un monologo (*ibid.*).

Anche i criteri formulati dal *Quadro comune europeo di riferimento* distinguono le attività comunicative in bidirezionali (interazione) e unidirezionali (produzione) e attribuiscono all'interazione una posizione predominante rispetto alla produzione poiché, se è fondamentale saper produrre testi monologici in quanto "svolgono una funzione importante in molti settori accademici e professionali e ad essi è attribuito un grande valore sociale", "imparare a interagire comporta qualcosa di più che imparare a comprendere e a produrre enunciati. Nell'uso e nell'apprendimento di una lingua generalmente si attribuisce grande importanza all'interazione in considerazione del fatto che essa riveste un ruolo centrale nella comunicazione" (Consiglio d'Europa, 2002:18).

Nelle attività di produzione orale l'emittente produce un testo orale che è ricevuto da un pubblico di uno o più ascoltatori (gli esempi citati sono: diffusione pubblica di informazioni ed istruzioni, discorsi in incontri pubblici, lezioni universitarie, sermoni, intrattenimento, commenti sportivi, presentazioni di vendita, ecc). Esse possono richiedere al parlante di (*ibid.*,72):

- leggere un testo scritto ad alta voce,

- parlare con l'aiuto di appunti scritti o di stimoli visivi (diagrammi, immagini, tabelle, ecc),
- recitare una parte,
- parlare spontaneamente,
- cantare.

Nelle attività di interazione orale il soggetto agisce alternativamente come emittente e ricevente con uno o più interlocutori al fine di costruire congiuntamente, attraverso la negoziazione del significato, il discorso conversazionale. Oltre all'impiego costante di strategie di ricezione e produzione, durante l'interazione vengono adottate strategie discorsive e di cooperazione che gestiscono il discorso rispetto all'alternanza dei turni, alla ripresa e sintesi di quanto detto in precedenza, alla scelta della modalità d'approccio a un problema e alla valutazione delle soluzioni, alla mediazione in caso di conflittualità.

Gli esempi di attività interattive riportati nel QdR sono (*ibid.*,92):

- transazioni,
- conversazione casuale,
- discussione informale,
- discussione formale,
- dibattito,
- intervista,
- negoziazione,
- progettazione comune,
- cooperazione finalizzata al raggiungimento di un obiettivo pratico.

Nel capitolo in cui vengono affrontate le questioni legate alla verifica della competenza linguistico-comunicativa dell'apprendente, il QdR fa presente che un test, per essere valido, deve necessariamente fare riferimento a campioni di comportamento linguistico riconducibili a diversi tipi di discorso (*ibid.*): viene dunque considerata condizione primaria e imprescindibile l'inclusione di compiti la cui esecuzione prevedano la produzione sia di natura bidirezionale che unidirezionale.

Nelle sezioni che seguono le tipologie di comunicazione orale verranno analizzate in modo più dettagliato e considerate nell'ambito di un progetto di certificazione linguistica che includa la verifica della competenza di produzione/interazione orale.

2. La produzione orale monologica

In questa sezione, si cercherà di definire in maggiore dettaglio le caratteristiche delle produzioni orali unidirezionali o monologiche. Come abbiamo visto nell'introduzione, la comunicazione orale accoglie al suo interno una pluralità di forme espressive che si realizza diversamente in relazione a contesti linguistici e situazionali specifici. Per lungo tempo negati o circoscritti dalla letteratura critica a caratteri di evanescenza e di intrinseca imprecisione, i tratti di dinamicità e flessibilità del parlato sono oggi concordemente riconosciuti (Nencioni, 1983; Bazzanella, *cit.*, Simone, 2000). In modo analogo, il discredito che ha interessato inizialmente le *modalità* comunicative orali, caratterizzandole limitatamente al canale espressivo adottato e avallando l'idea di una corrispondenza diretta tra parlato e registri espressivi bassi (contrapposta a quella tradizionale tra forme espressive scritte e registri alti), ha lasciato il posto a un fervore di studi che interessa trasversalmente molti campi disciplinari, secondo prospettive diverse (dalla linguistica teorica e applicata alla pragmatica, alla semiotica, alla psicolinguistica, alla sociologia).

Oggi, dunque, non è più possibile trattare il parlato con sufficienza e, anche in virtù dei recenti rivolgimenti tecnologici, di fenomeni di ibridazione e di "stilizzazione del parlato" (Dardano, 1994:177), numerose ricerche hanno evidenziato come alcuni tratti lessicali e grammaticali siano,

più che espressione intrinseca del mezzo comunicativo, realizzazioni concrete e socioculturalmente definite di modalità comunicative diverse. Come evidenzia Della Casa (1994:10), infatti, “possono esservi testi orali molto controllati e formali, e testi scritti realizzati in modo rilassato e informale [...] In questi casi, si parla assumendo le modalità linguistico-cognitive della scrittura, o si scrive utilizzando quelle del parlato”. Tale dato trova conferma anche nel summenzionato LIP, in cui emergono i due elementi di variazione già accennati precedentemente:

- il parlato è, per sua natura, un fenomeno poliedrico da considerare in prospettive microscopiche e non deterministiche;
- nelle diverse modalità espressive, è rinvenibile una chiara biforcazione dialogico-monologica (Sornicola, 1994).

La natura generalmente dialogica o plurilinguistica dei testi orali può, quindi, variare a seconda della tipologia testuale fino a esprimere tratti di affinità con i testi scritti, la cui natura è, appunto, eminentemente monologica. Per quanto sia ridimensionata la portata della variazione diamesica in rapporto a variazioni di tipo diastratico e diafasico, queste affinità, tuttavia, non devono indurre a considerare le modalità comunicative come indipendenti rispetto al mezzo espressivo. Nell'introduzione, si è detto che la conversazione rappresenta, tra gli usi della lingua orale, la forma orale più “naturale” o “classica” (Bachtin, 1988) della comunicazione verbale: l'alternanza di turni, infatti, garantisce condizioni d'uso ideale allorché consente una ripartizione equilibrata sia dei compiti di programmazione e produzione sia di quelli di ricezione ed elaborazione del messaggio. Di contro, le forme di parlato monologico o prevalentemente monologico si collocano sul versante opposto del continuum di tipi di parlato che le diverse tipologie testuali possono generare. Il parlato monologico, come sottolinea Voghera, esprime “una tensione tra condizione discorsiva primaria dialogica, verso la quale tende qualsiasi discorso parlato, e strutture testuali, che devono, in un certo senso, adattarsi a sostenere situazioni comunicative meno naturali” (Voghera, 2001:76-77). In altri termini, queste forme si allontanano dalla condizione di co-costruzione della comunicazione dialogica senza per questo perderne la fondamentale aspirazione interattiva allorché, appunto, “i testi parlati non costituiti in turni di presa di parola presentano caratteristiche che sono attribuibili al sostrato dialogico della comunicazione orale” quali, ad esempio, le forme di ripetizione, i processi di riformulazione o i segnali discorsivi con funzione pragmatica.

Il parlato monologico si caratterizza, dunque, in primo luogo in quanto abilità primaria in cui entrano in gioco processi cognitivi peculiari rispetto all'abilità integrata rappresentata dal parlato dialogico e, in secondo luogo, per una sua specifica articolazione a livello formale. Si tratta di un tipo di parlato unidirezionale con presenza, anche a distanza, di un destinatario. È una comunicazione che non contempla uno scambio di ruoli e che è caratterizzata da una maggiore coerenza tematica, da un più elevato grado di pianificazione, nonché da una certa regolarità strutturale in termini di coesione lessicale, grammaticale e fonologica (Lavinio, 1990).

Se questi tratti del parlato monologico possono definirsi per contrasto con quello dialogico, ulteriori caratteristiche possono essere precisate a seconda che si contempli il caso di un parlato monologico spontaneo, semispontaneo o pianificato, che si consideri la costruzione di un testo che si affida esclusivamente all'oralità o la resa orale di un testo scritto. In proposito, Lavinio sottolinea come anche l'interazione con la scrittura possa a sua volta modularsi e contemplare il ricorso a supporti scritti (tenuti presenti durante la produzione orale) che incidono in modo tangibile sulla spontaneità e sul livello di pianificazione del monologo: schemi discorsivi precostituiti, testi scritti a fini di estensione della memoria, tracce da sviluppare oralmente, scalette, appunti¹.

Anche per il testo orale unidirezionale, quindi, così come per la produzione orale interattiva, le ricerche hanno messo in luce aspetti peculiari sul piano della testualità e della grammatica, nonché caratteri estrinseci e intrinseci che concorrono a definirlo in modo affatto univoco in un continuum di usi. Quest'ultimo varia in relazione al contesto in cui si svolge la comunicazione (caratteri socio antropologici), alla struttura dell'evento comunicativo (la compresenza o l'assenza in loco del

¹ Per il rapporto tra orale e scrittura nel contesto pedagogico e per la commistione di codici e dei passaggi da un modello semiotico ad un altro, *cf.* anche Nonnon (2002).

destinatario), alle modalità, libere o regolate, della comunicazione. Tutto ciò risulta evidente qualora si consideri che, nel parlato monologico, rientrano a pieno titolo sia forme di oralità con un alto livello di progettazione, strutturazione e controllo (quali, ad esempio, il parlato del docente, l'intervento del conferenziere o la relazione accademica) sia forme di oralità minimamente mediate nella pianificazione testuale oppure caratterizzate da un progetto discorsivo e linguistico simultaneo all'atto espressivo stesso (si pensi, ad esempio, ai messaggi in segreteria telefonica, ai turni lunghi di un dibattito non preordinato nella forma e nei contenuti ecc.).

Nel parlato funzionale pianificato (Sobrero, 1988), ad esempio, laddove sono assenti molte delle variabili che costellano l'interazione, si richiedono, per una realizzazione efficace, una buona capacità di sfruttamento del tempo disponibile, il controllo di pause e di altri fenomeni di esitazione (quali false partenze, riempitivi, ripetizioni, ecc.), la consapevolezza degli aspetti paralinguistici della comunicazione orale (velocità e volume della dizione, chiarezza dell'articolazione, gestualità, ecc.), la gestione delle informazioni di cornice, il decentramento sul destinatario, la capacità di sfruttare adeguatamente eventuali supporti di pianificazione scritta nel rispetto dei caratteri dell'oralità (Colombo, 2001).

Nelle conversazioni telefoniche, viste in rapporto alle interazioni in presenza, Bercelli e Pallotti (2002) rilevano una riduzione della complessità comunicativa imposta dal mezzo telefonico, una tangibile diminuzione delle routine stereotipate e la notevole creatività che caratterizza le telefonate tra intimi, veri "capolavori di ingegneria conversazionale" (Bercelli e Pallotti, *cit.*:82).

A partire dal loro studio, è possibile evidenziare, per contrasto, alcuni elementi peculiari della sottocategoria monologica rappresentata dai messaggi in segreteria telefonica. Rifacendosi a Schegloff, infatti, gli autori identificano cinque fasi che, in molti casi, caratterizzano l'apertura dell'interazione telefonica: l'apertura del canale comunicativo, l'identificazione, i saluti, gli interessamenti e l'entrata nel merito della conversazione. Se l'aspetto più evidente prodotto dall'intermediazione meccanica del mezzo è nella riduzione della tipica evanescenza e non permanenza del parlato (*cf.* Simone, *cit.*), altre peculiarità distintive dei messaggi lasciati nella segreteria telefonica sono riconducibili alla non contiguità spaziale e temporale, cioè al fatto che non soltanto i parlanti non condividono lo stesso contesto fisico, ma si verifica anche quella distanza tra tempo di codifica e tempo di ricezione del messaggio tipica dei testi scritti. Mancano, perciò, o sono ridotti e meno usuali, gli elementi che contraddistinguono il contesto d'enunciazione comune e la compresenza temporale tra parlante e interlocutore (Bazzanella, *cit.*), vale a dire i mezzi non linguistici, l'importanza rivestita dalla funzione fatica, la presenza di un possibile *feedback* e la possibilità di riferirsi a conoscenze condivise (enciclopedia).

Naturalmente, tali cambiamenti sono a loro volta definiti dalle variabili relative al genere di rapporto interpersonale tra gli interlocutori e al grado di formalità della comunicazione. Le comunicazioni unidirezionali tra intimi, ad esempio, pur presentando un uso ridotto dei deittici, ne potranno comunque registrare la presenza, anche se l'interazione non è in simultanea. Modifiche sostanziali si avranno, conseguentemente, anche nella presenza e nell'ordine delle diverse fasi di apertura su elencate. Nei casi di primo contatto tra parlante e ricevente si procederà, in genere, direttamente ai saluti, all'identificazione e all'entrata nel merito della comunicazione e si produrranno ulteriori variazioni a seconda che si tratti di comunicazioni formali, professionali o informali².

Ulteriori tratti distintivi del parlato monologico riguardano la morfologia, la sintassi e la struttura fonologica dell'enunciazione ed emergono dall'analisi dei dati del LIP. Si tratta di elementi che interessano specificatamente il gruppo 2 (*scambi bidirezionali non faccia a faccia con presa di parola libera*), in cui sono compresi anche i messaggi in segreteria telefonica, il gruppo 3 (*scambi bidirezionali faccia a faccia con presa di parola non libera*), comprendenti situazioni enunciative in cui ricorrono turni lunghi di parola quali, in particolare, gli esami universitari, e il gruppo 4 (*scambi*

² In relazione al contesto pedagogico, si veda anche la distinzione semplificata proposta in A. De Benedetti, F. Gatti, *Routine e rituali nella comunicazione*, Paravia Scriptorium, 1999, p. 89.

unidirezionali in presenza di destinatario/i), che rappresenta il punto massimo di realizzazione di una situazione enunciativa monologica.

Relativamente alle comunicazioni telefoniche (*gruppo 2*), Mancini e Voghera (1994) segnalano l'uso prevalente di parole brevi, il numero circoscritto di forme e lemmi diversi, un'alta velocità dell'eloquio, seconda soltanto al parlato radiotelevisivo. A ciò si aggiunge il fatto che il numero di connettivi utilizzati risulta essere il più basso tra le diverse categorie, così come ridotto appare l'uso del futuro semplice e irrisorio il ricorso al congiuntivo presente (Voghera, 1993). Rispetto agli scambi bidirezionali faccia a faccia con presa di parola non libera (*gruppo 3*), a una fluenza ridotta, (Mancini e Voghera, *cit.*), corrispondono un'ampia varietà di connettivi usati e, nel quadro della generale semplificazione verbale nel parlato, una ricorrenza rispettivamente media e massima del futuro semplice e del congiuntivo presente (Voghera, 1993). In relazione agli scambi unidirezionali in presenza di destinatario/i (*gruppo 4*), si evidenziano poi la maggiore lunghezza media di parola, il maggior numero di forme e lemmi distinti e una velocità dell'eloquio non particolarmente ridotta, ma comunque inferiore al valore medio dell'intero corpus del LIP (Mancini e Voghera, *cit.*). In questo gruppo, inoltre, si realizza la massima articolazione nell'uso dei connettivi nonché delle forme verbali del futuro semplice e del congiuntivo presente (Voghera, 1993). Nell'insieme, infine, "maggiore è la varietà del lessico usato, maggiore è la variazione nella lunghezza delle parole. Le conversazioni telefoniche e le conferenze rappresentano rispettivamente i due poli di minore e maggiore varietà lessicale" (Mancini e Voghera, *cit.*:242-243).

L'analisi complessiva dei dati del corpus del LIP rende conto, dunque, delle specificità degli usi rispetto alla norma linguistica e, in particolare, della "massima rilassatezza del parlato" nelle conversazioni telefoniche (De Mauro, 1993:157) nonché della forte riduzione di alcune forme di integrazione tipiche dell'oralità nel cosiddetto parlato di qualità. In tal senso, è opportuno segnalare che, se è oramai superata la netta opposizione binaria che vedeva contrapposto il canale scritto a quello orale, altrettanto può dirsi per la bipartizione tra testi dialogici e testi monologici, vista la condivisa consapevolezza circa l'esistenza di testi intermedi il cui statuto testuale non è pienamente riconducibile all'uno o all'altro estremo di questa polarizzazione espressiva (Voghera, 1992:193).

Un ulteriore fenomeno di commistione si produce, inoltre, nel parlato "secondario": si tratta di quelle forme orali tecnologicamente mediate o fissate su supporto tecnico (da tempo, integrate in vari *corpora*) in cui risultano sfumati i vincoli di contingenza e di estrema volatilità della comunicazione orale e in cui gli elementi extralinguistici, caratteristici del parlato, cessano di rappresentare il naturale corredo della comunicazione orale (Simone, *cit.*; Rossini Favretti, 2002).

Nel parlato unidirezionale riconosciamo, insomma, una varietà e una plasticità di forme che possono esprimere una maggiore o minore stabilità e normatività in rapporto al grado di standardizzazione da esse raggiunto nel relativo ambito socioculturale. Una classificazione esaustiva di tutte le forme possibili è, naturalmente, un'impresa improba allorché ogni minima differenza nelle variabili della comunicazione (quali la circostanza, l'ambiente, la posizione sociale e il background individuale del parlante e del ricevente, ecc.) produce conseguenti mutazioni la cui significatività può essere tale da invalidare gli intenti tassonomici. Questa considerazione, tuttavia, non implica che non si possa giungere a una ripartizione d'insieme, com'è quella proposta da Bachtin (*cit.*), tra generi sostanzialmente stabili e prescrittivi e generi liberi e creativi di comunicazione orale³. Secondo Bachtin, infatti, i generi si richiamano a pratiche sociali che li hanno costituiti, sviluppandosi e assumendo caratteri costanti, riconducibili ad un livello di rappresentazione astratta. Tra generi e pratiche sociali si viene dunque a stabilire un rapporto dinamico che si evolve con il modificarsi dei due elementi interagenti, in base al loro essere costantemente definiti e situati a livello spaziale e temporale. Quest'ultimo punto risulta particolarmente rilevante nella definizione del contesto accademico, considerato nel punto 2.2.

³ Si veda anche Sabatini (1999).

2.1 Valutare le prove monologiche: problematiche e indicazioni della ricerca

La certificazione della competenza comunicativa nelle abilità produttive ha tradizionalmente posto una lunga serie di interrogativi a cui non sempre la ricerca può fornire risposte univoche⁴. Ciò è particolarmente vero per la valutazione della produzione e dell'interazione orale. Diversi esperti hanno contestato la validità dei costrutti delle prove esistenti (Fulcher, 1987; Raffaldini, 1988; Ambroso, 1995; D'Addio Colosimo, 1995); alcuni hanno criticato il concetto di autenticità applicato a contesti d'esame o considerato problemi di trasferibilità dei dati risultanti da ricerche in ambito pedagogico (Spolsky in Alderson e Banerjee 2002; Spence-Brown, 2001; Salaberry, in McNamara *et al.* 2002); altri hanno riaffermato l'importanza di valutare comunque il parlato soppesandone i pro e i contro (Garcia-Debanc 1999) o confrontando l'esito e l'atteggiamento dei candidati in relazione ad esami scritti e orali (Zeidner e Bensoussan, 1988; Hall, 1993); altri ancora sono giunti ad affermare l'opportunità di sospendere tale valutazione in attesa che maggiori certezze siano fornite da ricerche teoriche ed empiriche opportunamente validate (Matthews, 1990).

In generale, nel dibattito tuttora aperto sulla certificazione dell'abilità orale, le considerazioni che sconsigliano una valutazione del parlato rimandano alla difficoltà di definire criteri di valutazione validi ed affidabili che possano contemplarne la complessità costitutiva e i caratteri di co-costruzione presenti a più livelli, la globalità dell'investimento individuale che vi si esprime e il carattere socialmente connotato che esso veicola. Young (2002), ad esempio, testimonia della complessità fenomenologica e operativa che sottostà a una valutazione accurata dell'interazione conversazionale per la quale occorrerebbe indicizzare i diversi livelli descrittivi (la struttura discorsiva, il contesto situazionale, linguistico ed extralinguistico) che compongono il processo di interpretazione dell'interazione orale. Rifacendosi poi esplicitamente alla ricerca di Anderson e Ciliberti (1999) illustrata nel punto 2.2, Young segnala come, nella discussione teorica sul tema, spesso si trascurino contesti espressivi specifici come, ad esempio, la situazione degli esami orali universitari. Come vedremo, questi contesti situazionali rispondono a criteri costitutivi distanti da quelli che intervengono in una situazione dialogica di tipo più generico, sia per quanto riguarda la lunghezza dei turni di parola sia per la referenzialità qualitativa e quantitativa al contesto e al contesto discorsivo.

Nel rendere conto dei diversi aspetti teorici relativi alla produzione orale, Bygate (1998), infine, richiama l'attenzione sui vari meccanismi di elaborazione e produzione che distinguono un parlante nativo da un non nativo e sui diversi processi che possono intervenire a livello lessicale, di pause, di strategie comunicative di compensazione e di equilibrio tra le dimensioni di selezione e combinazione di significati e forme linguistiche.

Relativamente ai principali snodi specifici nella definizione delle prove orali, molti aspetti hanno interessato la ricerca. Ai fini del nostro lavoro, è utile evidenziare che sono stati privilegiati "punti caldi" quali le componenti dei costrutti e le caratteristiche dei task (tipologie e difficoltà di compiti, istruzioni e input; impegno cognitivo, scelta degli argomenti e presupposti socioculturali; tempi di esecuzione e pianificazione, ecc.) (Upshur e Turner, 1999), e se ne è discussa l'incidenza in relazione alla performance attesa (Bachman, 2002) e alla capacità delle prove di riflettere l'abilità orale che si intende valutare (Alderson e Banerjee, *cit.*). In proposito, la letteratura critica degli ultimi decenni si presenta vasta e diversificata per approcci e grado di focalizzazione, come del resto è intuibile viste l'ampia gamma di interazioni che ogni singolo elemento intrattiene con gli altri e le implicazioni che ne conseguono.

Focalizzare l'attenzione sul "fattore tempo", ad esempio, presuppone un discorso articolato su vari livelli. In termini generali, si dovrà segnalare come il problema dei tempi abbia spesso giocato a sfavore della valutazione della lingua parlata, nella tradizionale convinzione che valutare adeguatamente la produzione orale implichi un investimento di gran lunga superiore a quello richiesto da altre abilità (*cfr.* Hall, *cit.*). Richterich (1992), in particolare, ricorda come, in una

⁴ Questa problematica interessa anche aspetti generali della docimologia, *cfr.* Alderson e Banerjee, 2001, 2002; McNamara *et al.* 2002; Bachman, 2002; Ellis, 2003.

certificazione della competenza orale, il fattore tempo incida e caratterizzi svariati aspetti di un esame. Le scelte d'insieme che si devono compiere interessano, infatti, la definizione della durata totale della valutazione orale nel suo insieme, di quella parziale delle singole prove, il numero di prove previsto e la ripartizione dei tempi tra le stesse, la relazione tra l'assenza o la presenza di preparazione e le modalità di preparazione previste (pianificazione libera, guidata, preliminare, in itinere, ecc.). Relativamente ai singoli task, poi, gli elementi di rilievo considerati riguardano il rapporto tra tempi di somministrazione (cioè di introduzione, istruzione, fruizione di stimoli e di eventuali materiali di complemento) e tempo effettivo di produzione/interazione concesso al candidato nonché le modalità di definizione degli aspetti temporali di ogni specifica prova (definizione teorica, sperimentale o arbitraria)⁵.

L'indagine dei ricercatori ha interessato anche i caratteri di autenticità in rapporto alla vita reale (cfr. punti 2.2 e 3), l'impatto delle caratteristiche del candidato sulla prestazione richiesta e sul compito assegnato, la tipologia e la validità delle scale di misurazione utilizzate, le modalità di somministrazione, il ruolo e il peso assunti da interlocutori e misuratori nella valutazione. Nell'insieme, gli studi dimostrano quanto sottovalutare le problematiche connesse alla definizione di tutti questi aspetti possa incidere sull'affidabilità della valutazione orale, inficiandone la validità.

Queste oggettive difficoltà, tuttavia, non possono far eludere le implicazioni cognitive e sociali connesse ai diversi usi linguistici in cui l'abilità orale si articola, né permettono di ridimensionare lo statuto che essa assume in ambito formativo nel suo essere trasversale rispetto a qualsiasi disciplina. Decidere di ignorare il ruolo assolto dalla competenza comunicativa orale e interazionale (competenza socioculturale, identità sociale e autopromozione degli individui) non è, dunque, di per sé, meno problematico. Di conseguenza, l'orientamento teorico e operativo che sottostà alla scelta di procedere alla valutazione dell'abilità di produzione e interazione orale si realizza oggi, da una parte, adottando politiche di trasparenza circa i presupposti teorici, gli obiettivi perseguiti, le scelte effettuate, le procedure adottate e i criteri di valutazione definiti (Monnerie, 1992) e, dall'altra, perseguendo sia forme diverse di validazione sia approcci più articolati alla valutazione del parlato attraverso la misurazione di campionature quanto più possibile rappresentative degli usi linguistici pertinenti (Coste e Moore, 1992; Hall, *cit.*; D'Addio Colosimo, *cit.*; Bachman e Palmer, *cit.*).

Presupporre un approccio permanente nel senso di una validazione in itinere costante appare, a nostro avviso, particolarmente importante sia perché il quadro complessivo delle variabili e della loro interazione reciproca si presenta particolarmente complesso (Bachman, 2002), sia perché la ricerca in questo campo è in perdurante sviluppo, richiedendo quindi una continua revisione critica dei presupposti stessi della certificazione dell'abilità di produzione orale (McNamara, 2001). Un'altra considerazione fondamentale al riguardo, inoltre, interessa anche i tratti socioculturali degli usi discorsivi selezionati per la campionatura del parlato, per i quali pare opportuno perseguire un'analisi critica sistematica su più ampi campioni di prove, che vada a considerare l'impatto sulla prestazione dei candidati sia dei tratti costitutivi presupposti nei generi discorsivi scelti sia degli specifici contenuti proposti.

Per quanto sopra, infine, è ormai diffusa la consapevolezza circa l'importanza di definire criteri comuni di condotta etica⁶ e di vagliare attentamente l'impatto delle pratiche di valutazione sulla società nel suo insieme (Hamp-Lyons in Alderson e Banerjee, 2002; McNamara, 2001), non limitando la ricerca ai fenomeni di ricaduta sull'insegnamento (*washback*). Anche su questo punto, apparentemente così immediato, il dibattito in corso mostra come il raggiungimento di regole condivise e una loro concretizzazione a livello operativo possano risultare tutt'altro che agevoli da perseguire (Boyd e Davies, 2002). Ad ogni modo, dato che un esame di certificazione può avere un'incidenza notevole sui candidati, è cruciale che, nelle varie fasi di documentazione, ideazione,

⁵ Un'altra possibilità, non espressa dall'autore, è che tale definizione sia frutto di un compromesso tra criteri di priorità diversi, quali possono essere, ad esempio, quelli dei committenti e quelli degli ideatori delle prove.

⁶ Per una panoramica dei documenti ufficiali pubblicati in proposito, si consulti la pagina di risorse <http://www.surrey.ac.uk/ELI/links2.html>

creazione, sperimentazione e validazione puntuale e in itinere delle prove, si ponga la massima attenzione affinché le diverse parti costitutive risultino coerenti e, per quanto sempre discutibili e potenziabili nei loro singoli aspetti, dotate di un proprio rigore nei presupposti teorici e nella realizzazione operativa.

In tal senso, Bachman e Palmer (*cit.*) esordiscono nella loro trattazione mettendo in guardia da fraintesi presupposti di sistematizzazione assoluta e da errate aspettative totalizzanti circa la portata conoscitiva e operativa che il campo della valutazione è oggi in grado di esprimere. Un elemento rilevante dell'opera dei due autori è il riconoscimento esplicito del carattere non oggettivo o neutro del metodo di valutazione (caratteristiche ambientali, istruzioni, input, risposta prevista e rapporto tra risposta attesa e input) nonché del peso che l'adozione di un determinato metodo di valutazione, combinato alle caratteristiche dei singoli candidati (abilità linguistiche, conoscenze, convinzioni e atteggiamenti affettivi), può avere sulle performance che si vogliono elicitarle. Nella proposta di Bachman e Palmer, sei "qualità critiche" di un test linguistico (validità di costrutto, affidabilità, impatto, interattività, autenticità e praticabilità) concorrono a costituire il criterio cardine di *test usefulness* che deve essere definito in relazione all'ambito d'uso linguistico di riferimento (*Target Language Use domain*). In questo fondamentale tentativo di sistematizzare gli aspetti teorici e metodologici che contribuiscono alla creazione di un test e dei relativi compiti comunicativi, le suddette sei componenti fondamentali non sono proposte secondo una rigida scala gerarchica o di priorità operativa. In fase di progettazione, quindi, la definizione del loro reciproco rapporto rappresenta un momento cruciale di scelta per gli ideatori di un test.

Nell'intento di mettere a punto delle prove che presentassero condizioni ottimali di valutazione e misurazione delle specifiche componenti dell'abilità di produzione orale monologica (conoscenze linguistiche e strategie metacognitive), si è quindi partiti dai su elencati presupposti teorici e si è proceduto, in primo luogo, a un'analisi dei compiti e delle situazioni d'uso linguistico pertinenti nel contesto universitario di riferimento (*cf.* punto 2.2) e, in secondo luogo, a un'indagine esplorativa delle certificazioni linguistiche internazionali più diffuse e accreditate.

Relativamente alle tipologie di prove orali monologiche più praticate ai fini della certificazione di questa abilità primaria, si è considerata l'ampiezza dell'articolazione delle prove all'interno della campionatura complessiva delle forme di parlato monologico e dialogico. Da questa ricognizione, sono risultati ricorrenti i task orali unidirezionali di descrizione e narrazione basati su testi iconici (fotografie o disegni), seguiti da compiti di esposizione monologica libera o su traccia con funzioni prevalentemente esplicative, informative o argomentative, a seconda dei livelli di competenza comunicativa presupposti. Si è poi riscontrato come, in genere, le prove di produzione orale monologica si distinguano da quelle di interazione soprattutto in termini di tipologie o generi discorsivi sottesi ai diversi task e quindi in termini di variabilità della dimensione testuale e della prestazione attesa rispetto alle diverse tipologie possibili, ai diversi supporti forniti per la loro realizzazione, all'intenzione comunicativa espressa dal parlante e al fuoco prospettico della sua comunicazione, all'incidenza del fattore tempo sull'assetto del microtesto e sulla struttura sintattica, al rapporto tra fattore prosodico e struttura sintattica nonché alla scelta degli argomenti.

Differenziare le procedure di somministrazione non è, invece, una procedura altrettanto diffusa allorché i compiti di produzione orale unidirezionale sono spesso somministrati congiuntamente alle prove di parlato interattivo dallo stesso interlocutore che, all'occorrenza, svolge anche il ruolo di misuratore in presenza, risultando talvolta affiancato da un secondo somministratore e/o misuratore (*cf.* punto 4.1). In altri casi, le fasi di somministrazione e di misurazione sono tenute distinte e la misurazione avviene in un secondo momento su parlato registrato. Comunque, le condizioni ambientali e procedurali d'esame delle prove monologiche e interattive variano soprattutto in relazione all'eventuale copione seguito nel corso del colloquio dall'interlocutore⁷.

⁷ Tra gli altri, si differenziano da questo modello, l'*Oral Situation Test* illustrato da Raffaldini (1988), in cui la somministrazione avviene con testo scritto e la performance è audioregistrata e il test del lavoro di gruppo del Centre de

Come si è visto nel punto 2, ciò è perfettamente coerente con l'intrinseca vocazione dialogica che qualsiasi testo orale unidirezionale reca in sé e con il fatto che, nel parlato monologico come in quello dialogico, si deve valutare una prestazione che consenta di realizzare delle inferenze non soltanto sulla competenza linguistica del candidato, quanto piuttosto sulla sua competenza comunicativa, valutandone anche il grado di consapevolezza e maestria espressiva in relazione alle norme condivise a livello sociale e culturale che regolano una comunicazione adeguata al contesto, al destinatario e allo scopo dell'atto linguistico. Di contro, questo dato consente di isolare due aspetti che possono risultare in qualche modo problematici in sede di *testing*: uno attiene alla penuria di indicazioni operative specifiche che la ricerca teorica ed empirica offrono sulle peculiarità delle prove orali monologiche e un altro riguarda l'obiettivo di controllare o compensare il genere e il numero di variabili che possono condizionare la prestazione dei candidati nell'ambito di una campionatura di forme di parlato quanto più articolata possibile.

Riguardo la somministrazione del task da parte dell'interlocutore, gli esperti segnalano l'incidenza che possono assumere sia le caratteristiche dell'interlocutore sia le variabili di prestazione che questi può esprimere nel corso dell'esame. La natura stessa della co-costruzione interattiva, infatti, invalida l'idea di un interlocutore neutro o totalmente passivo: anche se silenzioso, questi non può comunque evitare d'influire sulla prestazione del candidato attraverso il proprio atteggiamento benevolo, ostile, indifferente o incoraggiante (Carton, 1992). In proposito, la letteratura registra diversi modi di porsi rispetto al ruolo di interlocutore/somministratore, riconducibili ad atteggiamenti di fondo (Brown e Hill, in Ellis, 2003; Brown e Lumley, in Alderson e Banerjee 2002; Lazaraton, in McNamara *et al.* 2002), al fatto di essere o non essere parlanti nativi della lingua (Wigglesworth, in Ellis, *cit.*), all'eventuale familiarità tra gli interlocutori (Katona, in McNamara *et al. cit.*) e a possibili sensibilità dei candidati rispetto alle aspettative nutrite dall'esaminatore o dall'istituzione (Upshur in Ellis, *cit.*; Lazaraton, in McNamara *et al. cit.*).

Anche rispetto alla somministrazione tecnologicamente mediata, si registrano perplessità, seppur diverse. Nella letteratura sull'argomento, si avanzano dubbi in relazione agli effetti che questa modalità di somministrazione può avere sulla prestazione del candidato soprattutto in termini di filtri affettivi e di inautenticità del *setting*, cioè di assenza di quelle caratteristiche che dovrebbero ricreare condizioni di tipo pragmatico di una comunicazione verosimile. L'analisi contrastiva di due dei maggiori esami in lingua inglese che ricorrono a queste diverse modalità di somministrazione, vale a dire l'*Oral Proficiency Interview* (OPI) dell'American Council on the Teaching of Foreign Languages (ACTFL) e il *Simulated Oral Proficiency Interview* (SOPI) del Center for Applied Linguistics (CAL), ha da una parte evidenziato una correlazione tra i risultati ottenuti nelle prove dirette e semidirette e, dall'altra, messo in luce che la gamma e la tipologia delle funzioni e degli argomenti relativi ai task di elicitazione variano in modo tangibile; nelle prove registrate, risultano, inoltre, inferiori i campioni linguistici relativi a strategie comunicative e tratti discorsivi quali, ad esempio, la densità lessicale (Alderson e Banerjee, 2002).

In termini positivi, possiamo concludere che la somministrazione dei compiti ad opera di un interlocutore garantisce condizioni ottimali di interattività, di informazioni paralinguistiche, di maggiore autenticità del contesto interattivo nonché di eventuale personalizzazione della prova d'esame, mentre la somministrazione tecnologicamente mediata garantisce una maggiore standardizzazione di somministrazione dell'input, condizioni affidabili di uniformità nel contesto d'esecuzione del task, una maggiore flessibilità gestionale e una notevole funzionalità in termini di somministrazione a grandi numeri di candidati.

Ulteriori problemi emergono per ambedue i tipi di somministrazione e riguardano la definizione dei parametri di misurazione e la misurazione del parlato registrato, procedura adottata in prove con o senza interlocutore. In proposito, la ricerca testimonia del vivace dibattito che interessa le modalità empiriche o teoriche di definizione delle scale di misurazione (scale nominali, ordinali o ad intervalli) nonché le problematiche connesse alle variabili rappresentate dalle caratteristiche

individuali dei misuratori (nativi o non nativi, status e background socioculturale, ecc.) (North e Schneider, 1998; Benvenuto, 2001; Upshur e Turner, *cit.*). Su quest'ultimo punto, in relazione alla misurazione del parlato secondario, si segnalano fenomeni di interferenza soggettiva dovuti al fatto che nella performance audioregistrata sono assenti elementi extralinguistici, che la misurazione può avvenire con criteri che attengono più all'espressione scritta che a quella orale (ad esempio, per le trascrizioni scritte spesso utilizzate), andando così ad avvantaggiare o svantaggiare i diversi candidati a seconda del livello di competenza dei candidati, del background e della formazione dei misuratori (Nambiar e Goon, 1993, Young, 2001).

Gli aspetti connessi alle problematiche della somministrazione e della misurazione sono dunque vari e complessi; Turner (1998) ed altri ritengono che gli studi disponibili permettano di evidenziare la presenza di variabili che incidono sulla performance dell'interlocutore e del misuratore, ma non di giungere a una conclusione condivisa. Allo stato attuale della ricerca, l'approccio più rigoroso sembra essere quello che Daniel Coste (1992:114) definisce "una geometria variabile" nella valutazione; se infatti "*language use in context is a chameleon that changes its color from one context to another*" (Young, *cit.*:12), appare opportuno praticare una politica di variazione non soltanto nella scelta dei task, ma anche nelle modalità di somministrazione, di misurazione e di applicazione dei criteri stessi in modo da permettere una migliore indicizzazione degli elementi ritenuti rilevanti nei diversi casi. Tali scelte non possono essere definite in modo arbitrario, ma devono trovare una corrispondenza nei criteri di significatività che conseguono dalle caratteristiche dei task e dei generi testuali prescelti. In altre parole, è necessario che la scelta delle condizioni ambientali di somministrazione e misurazione delle prove e l'intento di compensare le variabili che incidono sulla prestazione dell'esaminando nell'arco delle varie prove si traducano in un ulteriore elemento di caratterizzazione specifica.

Le scelte operate (*cf.* punto 3) nel differenziare le prove di produzione orale unidirezionale da quelle di parlato interattivo, si giustificano, quindi, in virtù sia dei dati della ricerca sulle diverse procedure, sia della valutazione complessiva dell'abilità del parlato che le prove di interazione monologica e dialogica congiuntamente esprimono, sia, infine, del massimo investimento in termini di risorse umane che è cruciale garantire nelle prove interattive (*cf.* punto 4.2). In tal senso, la nostra proposta di valutazione della produzione orale vuole fornire una campionatura adeguata su più fronti: parlato interattivo e unidirezionale, parlato spontaneo e pianificato, turni brevi e lunghi, performance orali in presenza e in assenza di interlocutore, misurazioni in simultanea e su parlato registrato.

2.2 Tipologia del test e ambito d'uso linguistico di riferimento

Riguardo alla valutazione delle abilità linguistiche e del parlato, Bachman e Palmer (*cit.*) rigettano la funzionalità di una trattazione incentrata sulle singole abilità in quanto tali e si richiamano piuttosto al concetto di uso linguistico socialmente e culturalmente connotato (*cf.* punto 2). Valutare la competenza comunicativa significa, quindi, contemplare la conoscenza delle regole sociali, psicologiche e culturali che governano l'uso del parlato nell'interazione sociale. Obiettivo di una prova di produzione orale unidirezionale sarà quindi quello di provocare una performance da cui inferire la competenza comunicativa del candidato in relazione alla dimensione linguistica, strategica, sociopragmatica e culturale.

I concetti chiave di *costrutto*, *task* e *performance*, tuttavia, possano veicolare contenuti e sfumature di significato diversi a seconda che si sposti l'enfasi sull'uno o sull'altro termine o che si voglia sbilanciare in un senso o nell'altro il rapporto che li interessa (McNamara, 1995; Bachman e Palmer, *cit.*; Bachman 2000; Skehan, 2003). Un approccio integrato ed equilibrato tra queste componenti consente di definire chiaramente le specifiche aree linguistiche che si vogliono valutare (*costrutto*), di selezionare compiti significativi su cui prevedere la prestazione attesa e poter esprimere delle generalizzazioni sulla sottostante competenza comunicativa (caratteristiche inerenti

al task) nonché di contemplare le complesse interazioni che si esprimono nell'esecuzione del task (performance, caratteristiche dei candidati, interazioni tra queste ultime e le caratteristiche del task) (Bachman, 2002).

Considerando le "qualità critiche" (*cf.* punto 2.1) indicate da Bachman e Palmer (*cit.*), si sono quindi definiti dei task di produzione orale monologica in grado di provocare una prestazione comunicativa significativa in rapporto all'esigenza che ha dato origine al progetto, allo specifico ambito d'uso linguistico di riferimento (*Target Language Use domain*) e agli eventi comunicativi che vi ricorrono.

Una certificazione linguistica in ambito accademico deve rispondere all'esigenza, già espressa in documenti ed atti ufficiali⁸, secondo cui "l'università, nella sua versione riformata ed europea, ha bisogno di verifiche eque, attendibili, valide, sia interne sia esterne per favorire la comparabilità dei curricula e la mobilità degli studenti... [nel rispetto delle] specificità del contesto universitario" (Zagrebelsky, 2001:282). Per rilevare tali specificità, si è dunque proceduto a un'analisi dei bisogni e a un'indagine accurata del contesto universitario e degli usi linguistici specifici, in modo da definire delle prove:

1. rappresentative dello specifico ambito socioculturale,
2. rilevanti rispetto ai contenuti prescelti (aree di abilità valutate),
3. adeguate in termini di corrispondenza tra task d'esame ed eventi comunicativi accademici.

In generale, il mondo accademico richiede, più di altri contesti esperenziali, la capacità di padroneggiare un contesto discorsivo di tipo monologico o comunque caratterizzato da turni lunghi di parola. In particolare, si sono considerati in maggior dettaglio i tratti ricorrenti in contesti italiani e anglosassoni per verificare le costanti dei diversi generi discorsivi vagliati in relazione alle specifiche pratiche sociali.

Rispetto al mondo accademico italiano, Ciliberti e Anderson (1999) affermano che vi si premiano non soltanto le abilità di ricezione e fruizione dei testi monologici prodotti dal corpo insegnante nell'ambito di lezioni frontali, ma anche e soprattutto la capacità dei discenti di produrre turni lunghi di parlato unidirezionale nell'ambito di "cornici discorsive" proposte dai docenti. Il discorso monologico, inoltre, caratterizza trasversalmente tutte le principali attività universitarie (lezioni, seminari ed esami) e appare contraddistinto dall'aspirazione a un'unitarietà topicale. Il discorso unidirezionale rappresenta, infatti, "il modello implicito di *comunicazione esperta*" e "la capacità di produrre monologhi estesi costituisce [...] per gli studenti un obiettivo da raggiungere se vogliono aspirare a diventare membri della comunità esperta cui appartiene il docente" (Anderson e Ciliberti, 2002:92). Non solo, questa capacità espressiva arriva a svolgere un ruolo chiave nella verifica e nella valutazione delle conoscenze possedute dai discenti tanto da indurre a un'implicita identificazione tra le conoscenze dello studente e "le capacità auto-gestionali dell'esaminato, le sue strategie interattive, o, più semplicemente, la sua capacità di ripetere in modo puntuale una precedente esplicitazione di determinati contenuti fatta dal docente o espressa dai libri di studio" (Ciliberti e Anderson, 1999:27). Relativamente agli studenti stranieri, le risposte fornite dai non nativi nei questionari di valutazione della didattica segnalano le difficoltà dei discenti stranieri in relazione all'uso del parlato monologico nel sistema universitario italiano (Pugliese e Ciliberti, 1999). Un'analisi contrastiva sulla lezione accademica italiana e sul caratteristico ricorso alle digressioni mette in luce le specificità del modello culturale di organizzazione discorsiva dell'insegnamento accademico italiano rispetto a quello inglese, sottolineando come la comprensione di una lezione universitaria così definita richieda lo sviluppo di una competenza metalinguistica e metadiscorsiva (Zorzi, 1994).

Sullo studio della lingua inglese nell'università italiana, Eerdmans e Lipson (1995) lamentano una diffusa difformità nelle procedure di valutazione delle abilità scritte e orali nei corsi di laurea

⁸ Si veda, ad esempio, la Dichiarazione di Bologna del 1999.

linguistico-letterario. Duguid (2000) conferma la prevalenza di un modello discorsivo fondamentalmente monologico nel sistema scolastico e universitario italiano: in aula come in sede di verifica, infatti, risulta premiata la capacità di ripetere i contenuti espressi dall'insegnante. Secondo la studiosa, la situazione è diversa in contesto anglosassone dove si privilegia l'espressione scritta su quella orale, cosa che giustificherebbe anche la diversa enfasi sulle due abilità produttive. Swales (1990) articola la sua trattazione sull'uso della lingua in ambiente universitario anglofono intorno ai concetti cardine di comunità di discorso e di genere discorsivo. Nel riconoscere la qualità esperienziale e formativa che la conversazione veicola e il suo essere fondante e preordinata al concetto stesso di genere, lo studioso ricorda che spesso esistono discrepanze tangibili tra le capacità conversazionali e non-conversazionali di un individuo e richiama l'attenzione sulle capacità verbali implicite nella padronanza di specifici generi testuali. Restringere l'attenzione sul mondo accademico non implica, a suo parere, una delimitazione e una definizione degli specifici eventi comunicativi prive di intrinseche difficoltà laddove "una ragione [di cautela] riguarda il fatto che il nome degli eventi comunicativi [...] tende a risolversi in etichette istituzionali più che in etichette descrittive" in relazione a variabili individuali e a mutamenti occorsi nella concezione stessa delle diverse attività svolte in ambito universitario. In tal senso, "le *lectures* possono non esser più quei monologhi recitativi di un tempo e possono, di contro, invitare attivamente ad alternare momenti di discussione e a svolgere compiti in gruppi ristretti. I *tutorials* possono consistere in interazioni tra studente e programmi su computer o su audioregistratore e non prevedere più un *tutor* nel senso tradizionale del termine" (*ibid.*,55-56).

Come si è visto, il quadro risultante dall'analisi del mondo accademico italiano registra notevoli costanti: in tale ambito, una buona padronanza dei diversi generi testuali monologici, considerati nella gamma delle diverse forme, più o meno regolate in termini di codificazione socioculturale, appare, infatti, un prerequisito ineludibile affinché un discente universitario sia posto in grado di attuare progetti di discorso adeguati ai diversi contesti comunicativi in cui si troverà ad operare, potendovi realizzare, al tempo stesso, quei caratteri stilistici ed espressivi che attengono alle esperienze, alle conoscenze e agli obiettivi personali di ogni singolo individuo. Riguardo al mondo accademico anglosassone, invece, il parlato monologico è presente, ma non preponderante, mentre è possibile che in tale contesto siano maggiori le situazioni di interazione studente-macchina.

Bachman e Palmer (*cit.*) ricordano che non sempre l'ambito d'uso di riferimento è chiaramente definibile per cui una fedele riproduzione di compiti deputati dalla vita reale può risultare inappropriata per ragioni riconducibili all'inadeguatezza dei "compiti reali" per la valutazione dell'abilità specifica, alla loro scarsa praticabilità in termini operativi e al fatto che determinati compiti possono svantaggiare alcuni candidati laddove presuppongano specifiche conoscenze o esperienze acquisite. Per tale ragione, si è deciso di non scegliere la situazione "esame universitario orale" e, nella definizione delle prove, si è cercato di comprendere dei generi testuali stabili e prescrittivi; generi in cui ricorrono routine esemplificative delle norme comunicative che caratterizzano il mondo accademico e prevedano, in relazione ai livelli di competenza B1 e B2, gradi variabili di impegno cognitivo e di individualizzazione del compito. Si evita così il rischio di penalizzare un'utenza composta da candidati provenienti da background anche molto diversi tra di loro.

2.3 Tipologia di formati e di task

In questa sezione si passeranno in rassegna le tecniche di elicitazione di parlato unidirezionale che contemplano esclusivamente tipologie di formati adeguati a far produrre al candidato un campione di comportamento linguistico, ovvero un testo che, secondo la funzione, la finalità e il canale, è riconducibile ad uno specifico genere di discorso parlato. È stata pertanto omessa la descrizione di tecniche che prevedono attività di tipo meccanico - la ripetizione, trasformazione e correzione di frasi, o il loro completamento a partire da uno stimolo audio o scritto - così come non sono presenti attività che sviluppano generi di parlato non utilizzabili a fini di

verifica - come la lettura ad alta voce (Balboni) – o utilizzabili solo nella fase preparatoria di *warm up* – come il testo orale libero.

Il candidato deve dunque produrre un testo orale che è recepito, nella maggior parte a distanza, da uno o più ascoltatori non tenuti a rispondere (CdE, *cit.*). Tenuto conto che le variazioni delle caratteristiche del discorso dipendono dalla relazione tra emittente e ascoltatore, dalla presenza o assenza di quest'ultimo nel luogo e nei tempi in cui viene prodotto, dai suoi scopi, la sua lunghezza e il suo grado di progettazione (Brighetti, Minuz, 2001), si è scelto di raggruppare le tipologie di *task* nelle 4 categorie di attività di produzione utilizzate dal CdE per la costruzione di scale di livello:

- Monologo articolato (descrivere esperienze)
- Monologo articolato (argomentare)
- Annunci pubblici
- Discorsi pubblici

Monologo articolato (descrivere esperienze)

➤ Narrazione

Genere universale, prima forma di esposizione orale, può assumere diversi formati in rapporto al soggetto di trattazione (Zuccherini, 1988):

- *il racconto di sé*, forma di espressione che contiene un ricco potenziale linguistico perché il candidato parla di esperienze personali. In Weir (1990) si riscontra che la produzione orale riguardante informazioni specifiche in ambito personale fornisce valide indicazioni di *language proficiency*;
- *il racconto di altri*, cioè di fatti che riguardano altre persone; in questo caso il candidato deve saper padroneggiare maggiormente i legami logici della narrazione;
- *il racconto di invenzione*, che permette al candidato di costruire liberamente il proprio testo narrativo, ma che richiede anche la capacità di controllare molto bene le strutture della narrazione.

Più confacenti alla verifica appaiono essere i task “guidati”, che per la narrazione utilizzano stimoli orali, scritti o visivi.

- *Narrazione a partire da stimolo orale o scritto*: integrare l'attività con testi letti o ascoltati può far equivalere il *test task* ad un *real life task* (Weir, *cit.*); d'altra parte si corre il rischio che le abilità di comprensione scritta od orale interferiscano negativamente nell'esecuzione del compito. Tra gli esempi di stimolo scritto si segnala l'uso di alcune brevi frasi sotto forma di appunti, appropriato per una verifica in ambito accademico; l'impiego di un *set* di informazioni generali su un racconto: dopo che il candidato ha letto l'insieme di informazioni, gli viene mostrata una parte del testo che dovrà collocare nel racconto stesso; infine, l'utilizzo di note biografiche per la ricostruzione del racconto della vita di un personaggio famoso (Zuccherini, *cit.*).
- *Narrazione a partire da uno stimolo visivo*: un'immagine o una striscia di qualità costituiscono degli ottimi stimoli per la produzione di un discorso esteso, molte sono le opportunità per l'espressione e l'interpretazione personale e il compito richiesto al candidato è chiaro: non richiede lettura o ascolto e pertanto si evita che la contaminazione di diverse abilità incida sull'esecuzione. È una tecnica efficace per far produrre un campione di discorso che permette l'applicazione di un'ampia gamma di criteri di valutazione: è utile, ad esempio, per valutare l'abilità morfosintattica del candidato, come l'uso appropriato di tempi e aspetti verbali, necessari a produrre un racconto efficace (Brighetti, Minuz, *cit.*). Con questo tipo di stimolo, si rischia però che il candidato perda la parte essenziale del significato per ragioni personali o culturali, o che possa mancargli l'opportunità di dimostrare il meglio della sua performance se la qualità delle immagini non è buona, o che, infine,

possa risultare svantaggioso per chi non conosce l'argomento e quindi i termini essenziali, a meno che non venga preventivamente fornito il lessico relativo (Underhill, 1987, Weir, *cit.*). Una variante consiste nell'uso di 2 o 3 immagini simili: il candidato sceglie per la narrazione una delle immagini e un altro candidato oppure l'interlocutore devono indovinare di quale immagine si tratta; oppure il candidato deve rilevare nella narrazione le differenze tra le immagini. La valutazione di questo formato punta soprattutto alla scioltezza e scorrevolezza della lingua generata. Un'altra variante proposta è la sistemazione di immagini in sequenza per la ricostruzione di una storia; un'altra ancora si basa sull'utilizzo di azioni filmate, che è un tipo di stimolo particolarmente gradito agli studenti (Anderson *et al.*, 1984).

- *Narrazione riferita*

a. Da uno stimolo audio: il candidato ascolta un brano audioregistrato e gli viene richiesto di riferirlo o di riassumerlo. Le istruzioni fornite sottolineano che dovrebbe usare il più possibile parole sue piuttosto che quelle del testo e che verrà valutata la qualità e non la quantità delle informazioni prodotte. Tale criteri devono essere contemplati nel sistema di misurazione che dovrebbe tener conto della qualità della parafrasi e della riproduzione dei punti principali del testo. Sebbene i brani registrati debbano essere necessariamente piuttosto brevi – da 6 a 10 frasi – sono comunque troppo lunghi per essere memorizzati completamente, quindi il compito che il candidato deve portare a termine richiede una sequenza complessa di abilità e strategie che includono la comprensione, la memorizzazione, la pianificazione e la decisione di cosa includere nel testo, l'elaborazione e la produzione. Trattandosi dunque di un compito ad abilità integrate, può risultare difficoltoso utilizzarlo nel caso in cui si voglia verificare la sola competenza orale, anche se esso possiede sicuramente tutte le qualità di attività autentica.

b. Da uno stimolo scritto: il candidato legge un brano o una serie di brani brevi che poi riferisce con le sue parole subito dopo. Non gli viene dato un tempo limite per la lettura, ma non gli è permesso di ritornare al testo scritto una volta che ha cominciato il suo racconto. Questa tecnica può essere usata per tutti i livelli: per i livelli più bassi, può essere necessario offrire lessico e controllare la comprensione, così che la riuscita nell'esecuzione del compito dipenda non dal grado di comprensione ma dall'attivazione delle abilità integrate. Per i livelli alti, si può segnalare al candidato (nelle istruzioni, nella presentazione dello stimolo) l'importanza che verrà attribuita, nella valutazione, all'uso dei connettivi interfrastici e all'inserimento appropriato di lessico non presente nel testo originale.

Le differenze tra *a* e *b* riguardano ovviamente le abilità usate: *a.* ascolto - parlato e *b.* lettura - parlato. Nel primo caso i tempi di somministrazione sono predeterminati, il candidato non ha il controllo su di essi e deve eseguire il compito basandosi solo su ciò che è riuscito a recepire in quel preciso momento; nel secondo caso, invece, avendo più controllo sullo stimolo, può recepirlo e rielaborarlo con i suoi tempi e le sue modalità; pertanto il testo scritto può essere linguisticamente più complesso.

➤ *Descrizione fattuale*

Il candidato descrive un oggetto, un sistema o una procedura a lui familiari. La familiarità consente al candidato di produrre un discorso esteso su un determinato argomento e allo stesso tempo gli permette una certa libertà di espressione. Le istruzioni dovrebbero sottolineare che la valutazione riguarda solo i punti considerati rilevanti all'interno dell'argomento scelto. Solitamente vengono forniti un elenco di 5-7 argomenti tra cui effettuare una scelta e alcuni minuti di tempo per la preparazione (esempio: "Come farebbe

una buona tazza di caffè?”). Una variante può consistere nella spiegazione di termini accademici o collegati a una professione.

Monologo articolato (argomentare)

Questo genere di discorso richiede un discreto livello di padronanza linguistica: il Quadro di Riferimento non assegna nessun descrittore ai livelli A1 e A2.

➤ Espressione di opinioni ed attitudini personali

Il candidato è invitato a scegliere un tema di carattere generale come, ad esempio, “Cosa si può fare per risolvere il problema della disoccupazione?”. Oltre ad esprimere la sua posizione, il candidato è invitato a esprimerne le ragioni. Lo scopo non è valutare i contenuti, quanto la capacità di usare efficacemente la lingua per sviluppare un’argomentazione.

➤ Discussione del contesto

Fornito uno stimolo scritto oppure audio o videoregistrato (come una notizia giornalistica), il candidato è invitato a fare speculazioni sul contesto della situazione e a prendere posizione (“Cosa farebbe se fosse...?”).

➤ Assunzione di ruolo

Il discorso, il comizio, la predica o l’arringa possono essere riprodotte attraverso attività di simulazione come il *role making*: al candidato viene descritta una situazione e gli viene chiesto di assumere un ruolo e produrre un discorso appropriato (Zuccherini, *cit.*). Se lo stimolo è costituito solo da una traccia di massima, questa tecnica lascia molta libertà al candidato nella conduzione dell’attività ed è in grado di verificare la competenza argomentativa effettivamente raggiunta. Bisogna comunque prestare molta attenzione a fornire situazioni che non siano ambigue o che possano costituire un pericolo di discriminazione culturale, favorendo ad esempio i candidati che hanno un background simile all’esaminatore e sfavorendo quelli che ne sono culturalmente lontani.

Annunci pubblici

Le forme tradizionali di annuncio pubblico, quali il bando, l’appello, il proclama, che sono normalmente prodotte dall’autorità verso un pubblico di sottoposti (Zuccherini, *cit.*), possono venire impiegate come campione di comportamento linguistico che lo studente produce in momenti particolari della sua vita sociale: gli si può ad esempio richiedere di fare un appello per una raccolta di fondi, di annunciare l’apertura di una manifestazione culturale o sportiva, o di proclamare una notizia di pubblico interesse.

Discorsi rivolti a un pubblico

➤ Conferenza

Tenuta da esperti che intrattengono un pubblico selezionato (colleghi, persone interessate all’argomento), è solitamente il risultato di un lavoro di ricerca e richiede dunque di essere preparata in anticipo (Zuccherini, *cit.*). Per l’esecuzione di questo task, al candidato viene richiesto di scegliere un argomento da un elenco (o, in alternativa, è l’interlocutore che identifica l’argomento più idoneo ponendogli preventivamente una serie di domande generali); gli viene poi fornito un corredo di testi scritti relativi al discorso da sviluppare. Si tratta di un’attività comunicativa autentica sia per scopi accademici che per scopi professionali, ma presenta due aspetti negativi: l’abilità di comprensione scritta può incidere negativamente nella produzione del discorso orale e la conoscenza dell’argomento molto tempo prima dell’esecuzione del test può indurre il candidato ad imparare l’esposizione a memoria. Per quanto riguarda la valutazione, può essere necessario identificare criteri di punteggio specifici per le diverse abilità che essa presuppone, per esempio esposizione di dati, espressione di opinione, capacità di sintesi, ecc.

➤ Esposizione orale o relazione

Gli scopi e i contenuti sono gli stessi della conferenza, ma diminuisce il grado di preparazione e progettazione. Il candidato prepara un'esposizione orale di durata variabile senza leggere o documentarsi in precedenza, ma avvalendosi solo degli stimoli grafici, iconici o sonori forniti al momento dell'esecuzione della prova. L'aspetto negativo rilevato da Hughes (1989), secondo cui le differenze personali e culturali di esecuzione e la molteplicità delle interpretazioni di argomenti di contenuto molto generale possono creare problemi al momento della valutazione, può essere ovviato fornendo al candidato una scaletta predeterminata dei punti del discorso da svolgere. Anche in questo caso siamo in presenza di un compito comunicativo autentico, sebbene la fase di preparazione ravvicinata all'esecuzione del test possa rappresentare per il candidato una fonte notevole di stress.

➤ Dimostrazione

La presentazione di specifici oggetti di studio presenta, in sostanza, le stesse caratteristiche di una relazione breve (Zuccherini, *cit.*).

➤ Lezione e spiegazione

Attività autentica per scopi accademici, richiede un ampio grado di preparazione e progettazione e può essere sviluppata con le stesse modalità del genere conferenza.

Tutti i compiti fin qui descritti possono essere somministrati a distanza, senza la presenza di interlocutore, misuratore o altri candidati. Poiché il test *performance-based* non si presta così facilmente all'autosomministrazione come altre tipologie di *testing* più tradizionali, la canalizzazione tecnologica di *test task* è un fenomeno relativamente recente (Chalhoub-Deville, 2001). Tale canalizzazione prevede l'uso di strumentazioni quali l'audioregistratore, il videoregistratore o il computer, mediante il quale il test può essere distribuito in rete locale (CBT) o via *World Wide Web* (WBT).

La procedura con uso di audioregistratore prevede la presentazione di stimoli registrati su cassetta o minidisk associati a stimoli scritti; la risposta del candidato viene registrata mediante un secondo audioregistratore, archiviata e in seguito valutata. Del tutto simile è la somministrazione video, come nella prova orale *VOCI, Video Oral Communication Instrument* dell'Università di San Diego, con l'unica differenza che i task sono presentati mediante videocassetta (Kenyon, Malabonga, 2001).

L'utilizzo di questo tipo di strumentazione presenta notevoli vantaggi di praticabilità, quali:

- la somministrazione del test a diversi candidati nello stesso tempo;
- la somministrazione contemporanea di prove registrate e prove in presenza;
- la valutazione delle esecuzioni nei luoghi e nei tempi che si ritengono più convenienti, con la possibilità di riascoltare la registrazione tutte le volte che si vuole (*cf.* punto 2.2);
- l'economia del tempo impiegato per l'attribuzione del punteggio, grazie a forme di controllo sulle registrazioni, come la compressione e l'eliminazione delle parti che contengono istruzioni e *prompt*.

L'introduzione del computer aggiunge a questi vantaggi tratti di maggiore efficienza, sicurezza, flessibilità e individualizzazione, permette il rilevamento del percorso seguito dal candidato nell'esecuzione del compito e un feedback immediato, consente infine la creazione di nuovi formati di task grazie all'integrazione dei media (*cf.* Rossini Favretti *et al.*, 2005⁹).

Tra gli svantaggi che la somministrazione a distanza comporta sono da annoverare:

- la proposta frequente di compiti non autentici, in quanto le situazioni in cui il parlante si trova a produrre un discorso in assenza di destinatari sono relativamente scarse nella vita reale (per l'autenticità delle attività in un test, *cf.* punto 3); in caso di parlato unidirezionale, le situazioni autentiche in cui il parlante si rivolge a destinatari che ricevono il messaggio in tempi e luoghi diversi sono riconducibili

⁹ Articolo reperibile in <http://amsacta.cib.unibo.it/>

solo alla comunicazione telefonica non in tempo reale (messaggio in segreteria telefonica) e alle trasmissioni radiofoniche e televisive in differita;

- la perdita di tutti gli aspetti visivi della comunicazione, come le espressioni del viso e la gestualità che, insieme all'interpretazione faccia a faccia dei silenzi e delle pause, possono costituire elementi decisivi per la valutazione del discorso prodotto;
- la mancata opportunità da parte dell'interlocutore di fare aggiustamenti durante l'esecuzione se trova difficile valutarla: in un test registrato si trova di fronte ad un prodotto finito;
- i problemi tecnici che possono influire sulla qualità dell'esecuzione registrata;
- l'assenza del contatto umano con la lingua straniera;
- la difficoltà e l'imbarazzo di molte persone a parlare al microfono, oltre ad un senso di mancanza di controllo della situazione, che possono generare nervosismo e incidere negativamente sull'esecuzione del compito. Si può comunque ovviare a questi inconvenienti di tipo affettivo fornendo le istruzioni nella lingua madre dei candidati o per iscritto, aumentando il controllo dei candidati sui tempi di pianificazione ed esecuzione del compito e assicurandosi preventivamente che i candidati abbiano preso confidenza con la strumentazione loro fornita (Norris, 2001).

Tra le tecniche di somministrazione a distanza, una nota a parte meritano, per le caratteristiche intrinseche che li contraddistinguono, i compiti con uso, simulato o non, del telefono. La comunicazione telefonica non permette di usare le componenti visive della comunicazione faccia a faccia, quindi può essere utilizzata come espediente per rendere un test registrato più autentico; inoltre, lo stesso tipo di comunicazione in lingua straniera richiede abilità particolari perché, a differenza delle telefonate in lingua madre che presentano un'alta frequenza di contatti sociali informali, avviene frequentemente in situazioni personali o professionali che esulano dalla ristretta cerchia di amici e parenti e possiede, dunque, contenuti altamente informativi e scarsamente ridondanti (Underhill, *cit.*).

Ci si può avvalere di questa tecnica richiedendo, ad esempio, al candidato di produrre un messaggio in segreteria telefonica per:

- rispondere a una persona che ha lasciato un messaggio,
- chiedere di chiarire un messaggio confuso e incompleto,
- fare una richiesta orientata a fini pratici,
- fornire informazioni su persone o eventi.

In conclusione, rimane da ribadire che in un test di verifica del parlato dovrebbero essere utilizzate almeno 2 tipologie di formato di discorso unidirezionale, associate ad altrettante tipologie di parlato bidirezionale: il test che ne risulterebbe sarebbe più equilibrato, in termini di formati più o meno controllabili, più autentico, riflettendo diverse modalità di uso della lingua, più onesto, perché le persone riescono meglio in determinati compiti piuttosto che in altri e, infine, più flessibile e adattabile a circostanze e bisogni diversi, nel caso in cui, ad esempio, si abbia la necessità di omettere alcuni task o di inserirli solo come prove di test.

3 Le scelte effettuate

Il nucleo inferenziale attorno al quale si costruisce un test *performance-based* è costituito dalla correttezza delle informazioni che si ricavano dai quesiti di fondo: a che cosa serve? A chi è rivolto? Che cosa devono sapere e saper fare gli studenti? In quali situazioni? Quali dati se ne vogliono ricavare?

Dunque, per la progettazione, è necessario innanzitutto definire e descrivere (Bachman e Palmer, *cit.*):

- gli scopi del test,
- gli ambiti d'uso della lingua,
- le caratteristiche e i bisogni dei candidati,
- i *constructs* sottostanti alle prove da somministrare.

Una volta stabilite le coordinate di base, ai candidati deve essere richiesto di produrre un campione di comportamento linguistico dall'osservazione del quale si possano ricavare informazioni sulle competenze sottostanti utilizzate per eseguire il compito (Klein-Braley, 1991), non essendo possibile esercitare su di esse un controllo diretto (CdE, *cit.*).

Così, dopo aver analizzato le caratteristiche degli ambiti d'uso della lingua in cui vengono attivate le competenze che ci interessa rilevare, bisogna identificare quali compiti inducano a produrre prestazioni significative per il contesto di apprendimento e per le caratteristiche e i bisogni degli esecutori del test. Le proprietà di tali compiti, la loro influenza sulla prestazione del candidato e l'inerente questione del grado di autenticità delle attività del test, costituiscono attualmente ampio oggetto di ricerca (Elder, Iwashita, McNamara, 2002): se alcuni studi si concentrano sugli effetti ottenuti dalla manipolazione delle caratteristiche del compito sulla qualità della prestazione, altri sostengono vigorosamente una necessaria corrispondenza tra *test task* e *target language use task*, annoverandola tra i fattori determinanti l'utilizzabilità di un test (Bachman e Palmer, *cit.*). Comunque, gli stessi fattori di quest'ultima posizione non mancano di mettere in luce come in alcuni casi la percezione di autenticità del *test task* appaia non determinante per la performance (Bachman, 2000): ad esempio, le ricerche svolte da Lewkowicz (in Bachman, 2000) suggeriscono che il modo in cui i candidati percepiscono i task sembra essere determinato, più che dal grado di autenticità, dalla qualità della loro esecuzione, dalla capacità di identificare esattamente ciò viene loro richiesto e dalla familiarità con quel tipo di compito.

Se la prospettiva incentrata sull'esecuzione di compiti è una base di partenza appropriata per rilevare le caratteristiche delle situazioni d'uso della lingua da cui ricavare poi le relative tipologie di performance, l'accertamento per *construct*, che si basa su una definizione precisa delle componenti dell'abilità linguistica oggetto di rilevamento, permette una valutazione generalizzata delle competenze del candidato a partire da performance situate in contesti specifici; è necessario, infatti, un modello che renda esplicito il procedimento di generalizzazione dei risultati ottenuti dall'osservazione del campione linguistico (Mislevy, Steinberg, Almond, 2002).

Il modello qui prescelto si basa sull'apparato concettuale fornito da Bachman e Palmer (*cit.*), che include, oltre ad una serie di procedure per la definizione delle qualità di un test, una sistematizzazione delle caratteristiche dei *test task* e dell'abilità linguistica richiesta nel test. Lo schema proposto per la descrizione dei *tasks* ne raggruppa le caratteristiche in 5 categorie: ambiente, struttura del test, input, risposta prevista all'input, rapporto tra input e risposta prevista (*ibid.*); l'abilità linguistica, intesa come combinazione di conoscenza della lingua e competenza strategica, viene invece scomposta nelle sue componenti grammaticali, testuali, funzionali, sociolinguistiche e metacognitive: la lista di controllo che se ne ottiene è lo strumento adottato per la definizione dei *constructs* (*ibid.*).

3.1 Tipi di task scelti e motivazione

Nella scelta della campionatura dei diversi tipi di discorso, si è tenuto conto dei fattori che possono incidere sulla prestazione dei candidati, tra cui (Bachman, *cit.*):

- le caratteristiche dei candidati stessi, nella fattispecie le conoscenze linguistiche, le competenze strategiche, le conoscenze pregresse sull'argomento e la componente affettiva, che concorrono a produrre interattività, una delle qualità di utilizzabilità del test;

- le caratteristiche delle procedure del test, quali la specificità dell'argomento, la tipologia del materiale di stimolo fornito, i tipi di compito da eseguire, il tempo di preparazione e la quantità di elaborazione richiesta nell'esecuzione del compito.

Sono stati inoltre presi in considerazione i criteri di praticabilità imposti da grandi numeri di candidati, che portano a considerare la registrazione (idealmente digitale) della prova orale monologica un canale privilegiato per la somministrazione.

Tra i compiti che prevedono la produzione di un discorso unidirezionale a distanza, si è scelto di utilizzare la tecnica con uso del telefono per la simulazione di una comunicazione telefonica non in tempo reale, che possiede la qualità di rendere autentico un discorso orale registrato: il messaggio in segreteria telefonica, testo conciso ad alta densità informativa e utilizzabile in lingua straniera sia in ambito personale che professionale. Il messaggio per il livello B1 rientra in un contesto d'uso privato ed è finalizzato alla ricerca di un'abitazione in un paese straniero, mentre per il livello B2 la comunicazione, inserita nel dominio educativo e caratterizzata da un grado più alto di formalità, riguarda la richiesta di partecipazione a un docente universitario ad attività culturali organizzate da un'associazione studentesca. Per entrambi i livelli, i *prompt* selezionati sono testi scritti che associano la brevità e l'immediatezza all'autenticità (annunci di offerte di appartamenti, brochure di presentazione di un'iniziativa studentesca).

Non si poteva inoltre prescindere dall'inserire nella campionatura l'esposizione orale, per la fondamentale funzione che essa ricopre in ambito accademico (CdE, *cit*): parlare a uno o più ascoltatori, anche con l'ausilio di appunti, testi scritti o supporti visivi, è un *real life task* che viene normalmente richiesto a uno studente universitario. Poiché questo tipo di compito richiede, per scopi professionali, un buon livello di padronanza linguistica e l'attivazione di strategie complesse di pianificazione, esecuzione e controllo, è parso appropriato assegnarne l'esecuzione ai candidati di livello B2. I materiali di stimolo, testi scritti iconici e grafici, differiscono nei contenuti in relazione alle 3 macroaree di competenza professionale, umanistica, scientifica e sociale, che il candidato sceglie prima di essere sottoposto alla prova. Per l'esecuzione del compito, i candidati si avvalgono di una scaletta preordinata, fornita nel *prompt*, dei punti del discorso da sviluppare e possono prendere appunti durante la fase di preparazione. L'abbondanza di risorse messe a disposizione del candidato (contestualizzazione del genere, scelta della macroarea, titolo dell'argomento, materiali di stimolo diversificati per formato, scaletta del discorso, possibilità di prendere appunti) contribuisce ad aumentare il grado di autenticità della simulazione registrata e può risolvere, in parte, problemi di filtro affettivo, ovviando all'inconveniente di dover produrre un discorso in pubblico...senza pubblico. Per un'ampia trattazione dell'argomento, si rimanda a Rossini Favretti *et al.* (*cit.*¹⁰).

3.2 La scelta dell'argomento

Le prove di produzione orale qui descritte non rientrano nell'ambito della valutazione a scopi specifici (*special purpose testing*), in quanto non è contemplata la valutazione combinata della competenza comunicativa del candidato e delle sue conoscenze pregresse in uno specifico ambito disciplinare (Alderson e Banerjee, 2001). In tal senso, si sono previste sia prove di parlato monologico relative a contesti comunicativi di tipo quotidiano (B1) sia prove di tipo formale più specificatamente collegate all'area educativa rappresentata dal contesto universitario (B2). Gli argomenti prescelti non si caratterizzano per un contenuto tematico marcatamente specialistico, ma attengono a una generale sfera d'azione comunicativa nell'ambito d'uso linguistico di riferimento o, nel caso della seconda prova di livello B2, a una macroarea disciplinare, comunque gestibile anche da non specialisti del settore.

¹⁰ Articolo reperibile in <http://amsacta.cib.unibo.it/>

Nel loro insieme, le prove di produzione orale monologica si presentano come prove semistrutturate (stimolo chiuso, risposta aperta) che si distinguono in relazione alle intenzioni comunicative, al destinatario e al genere discorsivo presupposti. Le prove *Lasciare un messaggio in segreteria telefonica* richiedono il controllo di formule linguistiche e routine di comunicazione a livello informale (B1) e formale (B2), in rapporto alle regole che governano l'uso delle segreterie telefoniche (concisione, sequenza delle fasi di apertura della comunicazione, ecc.)

La prova di livello B1 richiede una consapevolezza delle proprietà tipiche del parlato informale, mentre quelle di livello B2 presuppongono una padronanza sia del parlato formale a fini sociali (*messaggio in segreteria telefonica per un docente universitario*) sia delle proprietà del parlato di qualità finalizzato alla trasmissione dei contenuti (*esposizione orale in ambito accademico*).

In particolare, in quest'ultima prova di macroarea, il candidato deve saper comunicare secondo un genere orale caratterizzato da turni lunghi di parola, saper gestire la comunicazione nei limiti di tempo prefissati, definendo e organizzando una quantità di informazioni idonea all'esposizione, saper sfruttare l'apporto fornito dai supporti proposti e "transcodificarne" i contenuti in un'esposizione orale adeguata alla situazione enunciativa e ai destinatari della comunicazione indicati, ricorrendo a opportune risorse linguistiche e paralinguistiche. In particolare, i materiali di supporto prevedono un primo testo iconico a carattere connotativo e un secondo testo verbale o misto a carattere denotativo, comprendente elementi lessicali e nozionali di base.

A questo livello, si è quindi tenuto conto del fatto che "tra i parlanti nativi, gli studenti [...] costituiscono un sottogruppo peculiare, che deve imparare nuove varietà di lingua: dei sottocodici, caratterizzati, oltre che dal lessico tecnico, da usi formali della lingua" (Berretta, 1991:101). In tal senso, la preparazione sul tema prevista nel task di macroarea B2 risponde sì all'esigenza di evitare che i contenuti incidano sull'esecuzione e sulla misurazione del task, ma al contempo prevede l'opportunità di valutare la capacità del candidato di gestire le informazioni di base nel quadro di un genere discorsivo ben definito. Di conseguenza, non si valuterà la presenza o l'assenza di conoscenze pregresse sull'argomento proposto quanto piuttosto la capacità di strutturare i contenuti e il lessico fornito in un discorso coerente e coeso a partire dalla scaletta presentata.

3.3 Tempi

Diversi studi hanno analizzato questo importante aspetto in contesti pedagogici o d'esame, in esami di produzione orale monologica, interattiva o combinata. Le indagini hanno riguardato, da prospettive e con modalità diverse, soprattutto le implicazioni a livello di scorrevolezza dell'eloquio, accuratezza e complessità espressiva e cognitiva delle performance elicitate (Alderson e Banerjee, 2002; Ellis, *cit.*). Un altro aspetto di rilievo è rappresentato dagli eventuali meccanismi di *trade-off* che la presenza, la quantità o l'assenza di tempi di pianificazione impongono al candidato. Come teorizzato da Levelt (1989), infatti, il processo di pianificazione si realizza in momenti e su fronti diversi, tra loro interferenti, facendo presupporre che la pianificazione in un'area piuttosto che in un'altra influisca selettivamente sui sistemi di rappresentazione coinvolti nel pensiero e nel linguaggio.

Su questi singoli aspetti e sul loro rapporto con il fattore tempo, si sono anche considerati gli effetti prodotti su prestazioni linguistiche monologiche dall'assenza di pianificazione, dalla pianificazione preliminare e da quella in itinere. Yuan ed Ellis (2003) hanno riscontrato come l'opportunità di una pianificazione in itinere sembri favorire l'accuratezza e la complessità della comunicazione, incidendo negativamente sulla fluenza. La possibilità di una pianificazione preliminare, invece, parrebbe incentivare una maggiore scorrevolezza e complessità linguistica. Su questo punto, però, Wigglesworth (1997) ha rilevato soltanto una ricaduta positiva limitata ai candidati di livello linguistico alto.

Altrettanto controversi appaiono il rapporto tra il tipo di task e gli effetti della pianificazione (Ellis, *cit.*) e l'assunto secondo cui dagli esiti rinvenuti in contesti pedagogici si possano trarre indicazioni trasferibili nell'ambito del *testing* (cfr. Spence-Brown, 2001). Ne è un esempio la

ricerca svolta da Elder *et al.* (2002) in contesto d'esame, i cui dati hanno mancato di confermare quanto precedentemente riscontrato in ambito didattico.

Per le prove orali monologiche *Lasciare un messaggio in segreteria telefonica* (tempi di esecuzione: 2 minuti), si è definito di prevedere un tempo minimo di pianificazione (1 minuto), in considerazione delle maggiori condizioni di stress che generalmente contraddistinguono una comunicazione telefonica in lingua straniera e del fatto che la concisione del messaggio rappresenta un elemento di valutazione essenziale delle prove.

Riguardo alla prova di macroarea *Esposizione orale in ambito accademico* (tempi di esecuzione: 3 minuti), invece, considerazioni relative al genere discorsivo presupposto, all'impegno cognitivo richiesto, all'articolazione dei materiali di supporto nonché ai riscontri giunti da centri di sperimentazione delle prove hanno portato a definire il tempo di preparazione in 3 minuti.

3.4 *L'uso della tecnologia informatica*

Le prove di produzione orale monologica sono state concepite in una duplice versione: quella informatica (in fase progettuale) e quella presentata in questo contributo che si basa sull'uso di materiali cartacei e audioregistratori. Tra l'una e l'altra versione delle medesime prove, alcune coordinate devono essere necessariamente definite in modo diverso, riviste cioè alla luce delle mutate condizioni esterne dell'ambiente d'esame, delle procedure di somministrazione dello stimolo, delle modalità di registrazione delle prove e autonomia di gestione procedurale del task da parte del candidato. Il ricorso all'uso della tecnologia informatica per la somministrazione delle prove orali monologiche presenta innegabili vantaggi così come è passibile di suscitare delle perplessità. Per una trattazione sistematica di questo punto in relazione alle prove monologiche, si rimanda all'articolo "Somministrazione informatizzata di prove di produzione orale: ipotesi di fattibilità in ambito di certificazione linguistica per scopi accademici" di Rossini Favretti *et al.* (*cit.*¹¹).

4. **La produzione orale interattiva**

Molte delle problematiche poste dalla valutazione della produzione orale interattiva sono già state illustrate al punto 2.1. Ai fini della certificazione e in riferimento al contesto d'uso linguistico presentato dal Quadro europeo di riferimento (QdR), si sono definiti i seguenti punti:

- le aree scelte (*domains*) dal committente sono tre: il personale, il pubblico, e l'educativo, di cui la sotto-area specifica in termini di contesto educativo è quello dell'ambiente universitario – non sarebbe ugualmente accessibile per una popolazione di candidati legati per esempio alla scuola elementare;
- le situazioni, invece, sono state definite in base a quelle in cui gli studenti universitari si trovano normalmente a operare, compresi i limiti o le condizioni inerenti a tali contesti, allo scopo di ricercare una validità apparente (*face validity*).
- le motivazioni e gli interessi dei futuri candidati sono stati considerati in relazione alle attività prescelte, ponendo attenzione alla selezione di temi, sottotemi o nozioni presentati per i due livelli definiti per la certificazione (B1 e B2) e ricorrendo alle specificazioni disponibili per il livello Soglia e il livello Vantage (Van der Ek and Trim 1990, 2000).

¹¹ Articolo reperibile <http://amsacta.cib.unibo.it/>

Si è quindi compilata una lista di compiti comunicativi appropriati per l'ambito di riferimento e per i singoli livelli. Una volta individuati i tipi di testo più frequenti, la ricerca si è focalizzata sull'identificazione e sulla definizione di una serie di *tasks* che potessero risultare adeguati all'accertamento. Allo scopo di valutare la produzione orale complessiva, sono stati esaminati i diversi descrittori per i livelli prescelti. Nel QdR vengono forniti quattro esempi relativi alla padronanza monologica, mentre, per l'interazione, sono riportati nove esempi d'attività per l'interazione orale generica, compendiate dai relativi descrittori, e otto esempi specifici. Per la valutazione e la creazione degli *items*, sono state prese in considerazione le competenze illustrate nel QdR, considerando sia le competenze linguistiche (grammaticale, lessicale, semantica, fonologica) sia le competenze sociolinguistiche, pragmatiche e funzionali.

Inoltre, rispondendo alla richiesta del committente di definire un componente per scopi specifici, si sono usate come canovaccio le griglie già adottate in alcune università italiane per il Portfolio Linguistico (Argonizzio, 2001). Fra le indicazioni offerte per la produzione orale, le griglie fanno riferimento alla capacità:

- di interagire a lezione con espressioni semplici, ma corrette sul tema trattato dal docente (livello B1),
- di riassumere per linee generali i contenuti ascoltati o letti (livello B2), e
- di potere fornire in modo adeguato opinioni sull'argomento (livello B2),
- di fornire descrizioni chiare su una discreta gamma di argomenti riguardanti i propri campi di interesse e di spiegare un punto di vista su un argomento di studio, illustrando i vantaggi e gli svantaggi delle varie opzioni (livello B2).

Un'etica del *testing* (cfr. punto 2.1) richiede che il prodotto sia oggetto di un controllo di validità, affidabilità e praticità. A questo scopo, si era inizialmente deciso di utilizzare le domande di controllo proposte da Bachman e Palmer per valutare la *usefulness* (Bachman e Palmer, *cit.*, cfr. cap. 7). Il tempo disponibile per il progetto è giunto a termine prima che si concludesse la validazione della produzione orale. Questo è avvenuto per una serie di motivi intrinseci alla natura della produzione orale, che presenta dei problemi particolari già indicati in precedenza. Integriamo quanto già detto con ulteriori considerazioni tratte dalla ricerca attuale.

Siamo programmati biologicamente per il dialogo e le prime abilità comunicative che si possono osservare nell'interazione madre-bambino sono proprio quelle che riguardano i turni di parola. La natura interattiva della comunicazione orale ricorre in tutte le forme espressive, anche se in modo diverso nelle forme scritte e monologiche (cfr. Hoey 1983 ;Bakhtin; 1981, Voloshinov 1995). Ci sono una varietà di risorse con cui un parlante/scrittore può aggiustare lo status dialogico di un enunciato; in altri termini l'interazione è alla base del linguaggio (Voloshinov 1995: 139) In proposito, si deve tenere presente che la capacità d'interazione in una comunicazione orale interattiva implica diverse competenze, quali:

- la competenza linguistica,
- le competenze relative all'organizzazione del discorso e
- le competenze necessarie alla gestione dell'interazione stessa.

La produzione orale dimostra una sua specificità rispetto alle altre forme di comunicazione in termini di grammatica, di lessico e di *patterns* discorsivi (tratti, questi, documentati dai vari corpora linguistici, si veda, ad esempio, Biber *et al.*, 1999).

In un tipo di comunicazione com'è quella orale interattiva, si richiedono determinate abilità di elaborazione, fra le quali ricordiamo la concettualizzazione, la formulazione, l'articolazione ed il monitoraggio (cfr. Scovel, 1998). Inoltre, il tutto si svolge in tempo reale – cioè troppo velocemente per essere compiuto senza un certo grado di automazione – perché non abbiamo la capacità di controllare questi processi in modo conscio nei pochi attimi che ci sono concessi. Infine, il parlare è un fenomeno tipicamente reciproco, un'interazione fra persone, e questo porta automaticamente ad una certa imprevedibilità. Come si è detto nei punti 2 e 2.1, una serie di fattori collegati alle macro-funzioni del discorso (Halliday 1978) influisce in modo notevole:

- i ruoli dei partecipanti (simmetrici o asimmetrici),
- la familiarità dell'interlocutore,
- l'atto comunicativo di cui si tratta e
- il fatto che l'interazione sia situata in un contesto fisico in cui vengono sfruttati, anche se inconsciamente, segnali fisici che segnalano l'attenzione, l'intenzione di comunicare, l'atteggiamento rispetto a quello che viene detto e così via.

Per questo ultimo motivo l'interazione orale può anche permettersi di essere meno esplicita di quella scritta. Il problema del dover processare tutto in tempo reale rende difficile una pianificazione del discorso e induce a una serie di aggiustamenti, esitazioni, correzioni a metà frase e pause. Tutto ciò incide sul linguaggio che viene prodotto. Tenere presenti tutti questi fattori rende molto complessa la scelta di attività adatte sia per l'apprendimento sia per la valutazione della produzione orale di parlanti non-nativi, soprattutto a livelli di competenza limitata.

Infatti, com'è stato detto nel punto 2.1, soltanto negli ultimi 20 anni circa, la produzione orale in L2 è stata studiata e analizzata nelle sue peculiarità dalla linguistica applicata e dalla glottodidattica. Ne è emerso un corpus di studi che ha permesso di vedere nella produzione orale in L2 un argomento ricco di implicazioni, l'oggetto di una vasta letteratura critica paragonabile a quella riservata alle altre abilità. Ciò avviene per una serie di motivi; in parte, vi sono motivi storici che derivano dall'influenza ancora duratura della tradizione grammaticale-nozionale, che ha marginalizzato a lungo l'insegnamento delle abilità comunicative. Anche a distanza di anni dopo l'avvento dell'era comunicativa, l'osservazione in classe e il racconto di studenti e tirocinanti nella scuola italiana ci indicano la poca attenzione per l'uso della lingua per scopi comunicativi durante l'apprendimento (Duguid 2001:cap.3). Il secondo motivo per il ritardo nella ricerca dedicata all'interazione orale è tecnologico, solo a partire dall'ultimo quarto del secolo scorso è stato possibile studiare a fondo la produzione orale grazie a mezzi di riproduzione audio e video; in precedenza, infatti, la povertà dei mezzi tecnici aveva reso più facile lo studio della lingua scritta. Adesso invece la ricerca ha potuto rivolgervi un'attenzione maggiore.

Tipicamente, dunque, l'interazione orale si caratterizza per il fatto di essere parlata, interattiva, prodotta in tempo reale e in una situazione condivisa dai partecipanti, i quali esprimono un'intenzione comunicativa personale a un pubblico per lo più individuale su un ambito o un'area di contenuto che varia secondo le circostanze. Tutto ciò rappresenta una sfida per chi cerca di riprodurre le condizioni per l'apprendimento o per la valutazione della produzione orale interattiva. Anche se la tradizione audio-orale ha messo in prima linea le abilità orali, tale approccio ha ignorato l'importanza sia del contesto sociale sia dell'intento comunicativo. Con il concetto di competenza comunicativa (Hymes, 1979), lo sviluppo dell'approccio funzionale-nozionale ha portato a elaborare materiali e procedure per l'apprendimento delle abilità comunicative, e a sviluppare una nuova consapevolezza circa l'importanza di fattori come la formalità o l'informalità di una situazione comunicativa, in relazione ai ruoli dei partecipanti. Anche allora però tutto si basava sull'identificazione di questi *speech acts* e vi erano pochissimi studi che si interessassero

approfonditamente all'uso naturale della lingua orale in situazioni interattive o all'apprendimento delle abilità orali. Adesso, però, possiamo contare su una folta letteratura dedicata alla L1 (Yule, 1996; Cook, 1989; Hoey, 1991; Carter e Mc Carthy, 1997) e sono stati individuati i tratti salienti del discorso orale interattivo. Gli studi sull'uso della L2 ci hanno portati a poter individuare i problemi degli apprendenti e le abilità necessarie per essere in grado di comunicare in L2. Attualmente, si sta cercando di esplorare la produzione orale in contesti comunicativi per capire in che modo i compiti possano influenzare la produzione in termini di fluenza, accuratezza e complessità. Altri studi importanti riguardano l'interazione fra parlanti nativi (*native speakers, NS*) e non-nativi (*non-native speakers, NNS*) di cui si indagano i tratti distintivi particolari (Long 1981, 1983). Tutto ciò indica la necessità di dare vita a una metodologia distinta e stilare programmi particolari per l'apprendimento di questa specifica abilità, in modo da permettere al discente di trovare un giusto equilibrio fra accuratezza, fluenza e complessità. Una varietà di tipi di interazione va dunque offerta nei compiti scelti, le condizioni devono essere diverse da quelle fornite per la lingua scritta, mentre la capacità di improvvisazione va allenata e valutata. Le implicazioni per la valutazione della produzione orale interattiva sono evidenti, ma ci sono problemi particolari per quanto riguarda la valutazione della performance in interazione.

4.1 Valutare le prove interattive: problematiche e indicazioni della ricerca

Le considerazioni espresse nella sezione 2.1 ci portano a definire con cura le modalità con cui viene valutata la produzione orale interattiva. Un test di esecuzione deve sfruttare elementi sia analitici che integrativi; l'aspetto analitico rappresenta il rigore ed evita la non-generalizzabilità dei risultati, quello integrativo garantisce una contestualizzazione, così importante in un test di produzione orale, visto che una lingua non è un insieme di pezzi separati e che tali pezzi devono essere testati insieme. Quello che va testato, dunque, dev'essere l'abilità comunicativa e non la conoscenza formale; per questo, servono test d'uso della lingua: è questa la sfida che raccoglie chi vuole creare un test di produzione orale. Morrow, in uno dei primi contributi sulla valutazione scrive: "*the use of language in a communicative situation has a number of features which are not measured in conventional language tests*" (Morrow, 1977). Canale e Swain (1980) hanno fornito un quadro di riferimento per coloro che vogliono creare un test comunicativo, ma è stato Skehan a proporre la soluzione della difficoltà ("*there seems no way in which we can ever, as testers, truly satisfy the conditions of purposive communication and contextual grounding*") con l'uso del *performance testing*, paragonato al processo di isolare un campione d'acqua dal corso di un fiume attraverso l'immersione di un tubo. Questa modalità, in cui il test riproduce più o meno mimeticamente un comportamento del mondo reale, garantisce con maggiori probabilità il rispetto delle condizioni della comunicazione genuina. Il problema che permane riguarda il fatto che un *performance test* è per forza un *test di performance* e non della competenza sottostante. Il problema viene discusso a fondo in Bachman 1990. Comunque sia, l'organizzazione di questa campionatura crea problemi non indifferenti per tutti quelli che cercano di ricreare in un test condizioni di autenticità, imprevedibilità e fattibilità, senza però perdere le qualità di validità e affidabilità che sono fondamentali per una certificazione degna di tale nome.

Una parte importante nella valutazione della produzione orale riguarda la definizione delle condizioni della performance. Come si è visto nel punto 2.1, queste condizioni possono avere un impatto definitivo sui risultati dell'interazione orale: il numero, la familiarità e lo status degli interlocutori, il tempo e la predisposizione alla preparazione. A questi fattori si aggiungono quelli relativi all'input, anch'esso fonte di variabilità come accade per il contributo dell'interlocutore. Anche i fattori che possono cambiare la qualità dell'output sono numerosi in quanto le caratteristiche personali possono intrecciarsi in modo evidente con la capacità di comunicare con gli

altri (cfr. QdR, 5.1). Come si è detto, l'interazione è un'espressione di reciprocità, in cui si realizza l'interazione interpersonale. La sfida risiede nel costruire delle prove che riproducano condizioni fattibili e verosimili rispetto all'area e al contesto prescelti, ma che non compromettano la validità e generalizzabilità delle inferenze che si possono trarre dai risultati. Lazaron (1996) e Bachman (1990) hanno discusso la difficoltà di mantenere costanti le condizioni di elicitazione, soprattutto per quanto riguarda il comportamento dell'interlocutore. Walsh (1999, 2000) fornisce un'utile recensione della ricerca e discute le differenze fra conversazione e colloquio di competenza linguistica, analizzando nel dettaglio le variazioni interazionali, i problemi nell'uso di un protocollo fisso, i problemi relativi all'asimmetria del rapporto interpersonale nei colloqui individuali, nonché i problemi attinenti alla formazione dei misuratori. La soluzione adottata in molti casi è ricorrere all'uso di colloqui, malgrado le differenze riscontrate a livello di contesto situazionale e di output ottenuto rispetto alla conversazione. Questa soluzione, di fatto, viene preferita semplicemente perché risulta più facile somministrare un colloquio che cercare di ricreare le condizioni necessarie per valutare una conversazione autentica.

Il *proficiency test* di lingua inglese più diffuso nel mondo universitario (a parte il TOEFL che solo tardivamente si è adattato alla rivoluzione comunicativa, aggiungendo pezzetti alla sua struttura di base) è stato l'ELTS, adesso IELTS. Questo test è anche quello più studiato e criticato, rispetto al quale la ricerca ha molto da offrirci (cfr. Cambridge Research Notes 1-11, una serie di relazioni brevi in cui si documentano le ricerche in atto e si esaminano in dettaglio i cambiamenti e le motivazioni delle scelte adottate per il test orale). Altre indicazioni e spunti pertinenti vengono dai dati di una ricerca empirica sui risultati dei test di livello B1 in un'università italiana. Questi test sono stati somministrati a un campione di oltre 2000 studenti provenienti da una varietà di background facendo particolare attenzione ai risultati del test orale interattivo di livello B1 (cfr. Duguid, 2001, cap.9). Il test di produzione orale è risultato poco discriminante rispetto alle altre prove: solo il 9% di chi non ha superato il test è stato bocciato per problemi inerenti all'abilità di produzione orale interattiva, e solo l'8% di quelli che non hanno superato il test per un margine minimo di punteggio (i *narrow-fail*) ha avuto problemi con la prova orale. Relativamente all'ascolto, vediamo invece che, tra i bocciati, il 64% aveva avuto problemi con l'ascolto mentre quelli appena insufficienti sono risultati l'86%. È possibile che ciò sia dovuto al fatto che gli studenti italiani sono più abituati ai test orali, la forma di valutazione più frequente nel sistema scolastico, ma questa tende ad essere piuttosto un susseguirsi di monologhi e non presenta le caratteristiche di una vera interazione (cfr. Anderson e Ciliberti, 1999). Tuttavia, l'impressione di tutti gli insegnanti coinvolti nel progetto è stata che il test orale fosse particolarmente blando e che una preparazione ai compiti potesse limitarsi all'allenamento nell'uso di certe enunciazioni formulaiche e ad una conoscenza del lessico. Il fattore più significativo è sembrato essere la preparazione al formato del test. I criteri di valutazione, infine, sono apparsi molto generosi. Ci sembra improbabile che un ente di fama mondiale, molto criticato e osservato, ma con un'equipe di ricercatori, un database enorme e strutture di ricerca invidiabili, il quale investe somme ingenti nella ricerca e nel controllo dei test, possa fare errori grossolani di tale livello. Sembra molto più probabile, invece, che fare un test diretto di produzione orale interattiva nel rispetto dei criteri di validità, affidabilità e fattibilità sia molto difficile, cosa che tende a far ridurre i compiti e i criteri di valutazione delle prove a livelli poco discriminanti.

Per quanto riguarda altri fattori come l'etica o gli effetti di ricaduta sull'insegnamento (*washback*), sono stati adottati i criteri di Bachman and Palmer (*cit.*) per arrivare a un concetto di *usefulness* da applicare alla valutazione della produzione orale. Anche qui comunque la tensione fra validità, affidabilità e fattibilità o praticità ha creato qualche perplessità. Un'analisi in questi termini era prevista nel progetto originale della sperimentazione in discussione ma per motivi logistici non è stata completata. Il test di produzione orale interattiva dunque non è da considerarsi ancora validato nel numero di fasi di preparazione. È particolarmente importante completare:

- l'analisi della performance dei candidati (*cfr.* Fulcher 1996) per quanto riguarda la sintassi, il lessico e la fluenza, e l'organizzazione del discorso; analisi necessaria per assicurare scale di valutazione basate su quello che può succedere e non solo su modelli teorici di competenza;
- la standardizzazione attraverso la formazione dei misuratori è un altro elemento fondamentale; tale formazione includerebbe la familiarizzazione con il formato e le procedure del test, nonché con i criteri di valutazione attraverso l'uso di campioni video di performance a vari livelli, la discussione dei descrittori di valutazione e la pratica di valutazione di un certo numero di esempi di performance. Le specificazioni degli *items* e dei descrittori della valutazione vanno rivisti alla luce della performance dei candidati e delle decisioni prese dai misuratori per scegliere i criteri analitici con cognizione di causa;
- tutto deve essere ritestato con altri compiti, diversi da quelli del prototipo. La validità e l'affidabilità dei nuovi compiti dovrà quindi essere testata;
- i compiti campione vanno testati con NS e NNS per controllare i tempi e la chiarezza delle istruzioni, la loro fattibilità ecc.; andrebbe inoltre organizzato un programma di registrazioni video su una percentuale significativa di test, con conseguente analisi di gruppo;
- ulteriori materiali dovranno essere prodotti per la banca degli *items*.
- Anche i risultati parziali della sperimentazione perdono, dunque, di affidabilità se è stato usato un unico set di *items*.

Per un test fornito da un insegnante per l'uso all'interno della sua classe i criteri di validazione possono essere forse meno rigorosi ma per potere parlare di certificazione è auspicabile che l'ente certificatore cerchi di corrispondere ai principi delineati nei documenti e linee guida ALTE (Association of Language Testers in Europe) e fornire dati e prove che documentino, in particolar modo, la validità ed affidabilità del processo di validazione delle prove e dei singoli *items* utilizzati. Il Language Policy division del Consiglio d'Europa infatti avverte: *“Relating an examination or test to the CEF is a complex endeavour. The existence of such a relation is not a simple observable fact but is an assertion for which the examination provider needs to provide both theoretical and empirical evidence. The procedures by which such evidence is put forward can be summarized by the term ‘validation of the claim’ ”* (North 2001:Introduction) Le procedure per la validazione di compiti interattivi sono particolarmente complesse e richiedono forse più risorse di altre forme di valutazione.

4.2 Tipologia di test

Per una trattazione completa dell'argomento si rimanda anche al punto 2.2. Nel creare il test si è cercato di tenere in mente che la comunicazione è basata sull'interazione, è imprevedibile, ha luogo in un contesto, è il raggiungimento di un obiettivo, che il linguaggio è autentico e basato sul comportamento. Questi criteri, soprattutto quelli dell'imprevedibilità e autenticità, sono quasi impossibili da realizzare nel contesto di una certificazione. Inoltre, l'esigenza da parte dei committenti di fornire un test di linguaggio specifico anche a livello B1 ha introdotto la difficoltà aggiuntiva dell'interazione con la conoscenza di contenuti specifici – problema molto discusso nel mondo del testing (*cfr.* Douglas, 2001, Elder, 2001, Zuengler, 1991). Sono state identificate almeno tre aree problematiche: fino a che punto sia possibile identificare i contesti distinti di linguaggi specifici, l'autenticità, sia interazionale che situazionale, e l'impatto e interazione di fattori non linguistici. Di nuovo, solo la ricerca empirica sul linguaggio prodotto in prove compiute ci può fare procedere in questo campo. Per quanto riguarda la costruzione di item per scopi specifici (in questo caso il committente ha chiesto item per tre macroaree), in seguito a un esame della letteratura e della storia della costruzione di test per scopi simili (IELTS, per esempio, dove un'ampia scelta di

macroaree è stata ridotta a tre e poi ad un'unica area, quella dell'uso accademico), Davies (2001) afferma che vi sono poche prove evidenti che attestino la differenza profonda nell'uso della lingua per scopi specifici; l'uso di vocaboli appropriati è un criterio che vale sia per scopi generali che per scopi specifici e, anche se il lessico è diverso, il criterio rimane lo stesso. Il costrutto è comune, la differenza risiede nella scelta di temi o argomenti, tipi di testo, funzioni e situazioni presentate. Si deve sempre tenere presente che più l'argomento è specifico, meno generalizzabili saranno i risultati. Comunque, perché il test risulti validato, rimane da fare ulteriore lavoro per la costruzione di liste di parole e funzioni basate su un'analisi dei corpora specifici esistenti.

Tutto questo assume un'importanza notevole quando si cerca di arrivare ad un test con validità di contenuto e di costrutto e soprattutto quando il test è mirato ad una certificazione che avrà un impatto sulla carriera accademica di un numero considerevole di studenti. Il costrutto deve riflettere l'attuale accettazione teorica delle abilità in questione, il contenuto deve fornire una campionatura adeguata a permettere una valutazione. I problemi rappresentati da un test di performance sono dunque tanti, ed esiste il rischio reale di fare tanti tipi di errore (errori di misurazione, errori sistematici, errori di carattere culturale) nell'assegnare un punteggio le cui variazioni siano causate da fattori non pertinenti all'abilità sotto esame. Come afferma Messick, non è tanto il test in sé che è più o meno valido, ma l'uso che si fa delle inferenze tratte dai risultati.

3.2 *Tipologia di formati e di task*

Non c'è spazio qui per una rassegna delle tecniche per la produzione del parlato bidirezionale contenente le tipologie di formati adeguati a far produrre al candidato un campione di comportamento linguistico (si rimanda per esempio a manuali come Weir, 1993): rimane da ribadire che in un test di verifica del parlato dovrebbero essere utilizzate almeno 2 tipologie di formato di discorso bidirezionale per garantire un test più equilibrato, in termini di formati più o meno controllabili, e più autentico, che rifletta diverse modalità di uso della lingua.

Fra le varie possibilità si trovano l'interazione studente – studente (tra pari) per scambio di informazioni con *information gap*; la cosiddetta conversazione libera, il colloquio - come si vedrà in seguito, questa tipologia tende a non assomigliare alla conversazione naturale, dati i ruoli dei partecipanti - e il colloquio controllato o *fixed frame* (Walsh, 2002); il *role play* fra candidato e esaminatore, l'*information gap* tra esaminatore e candidato.

Importanti sono le considerazioni già espresse nella sezione 3, molte delle quali riguardano anche la produzione orale interattiva. È chiaro che i compiti scelti per la valutazione delle abilità interattive orali devono in qualche modo rispecchiare il tipo di task dell'uso della lingua in situazioni autentiche. È stato dimostrato (Skehan e Foster, 1997) che diversi tipi di compito tendono a produrre un impatto diverso, in termini della accuratezza, fluenza o complessità della lingua prodotta. La complessità cognitiva del compito tende ad avere un effetto anche sulla complessità linguistica (QdR, *cit.*, cap.7; Bachman e Bannerjee, 2002). Si vede dunque che la scelta del compito ha un ruolo chiave nella qualità della produzione. Anche la possibilità di ripetere un compito può permettere al discente di migliorare la formulazione: anche se questo non è un fattore naturale e autentico, fa comunque parte dell'esperienza del candidato. Attualmente si discute molto sulla scelta dei compiti in rapporto alla complessità, accuratezza e fluenza linguistica dell'output e in rapporto all'input, ma, come afferma Bygate, "*studies into the impact of tasks on students' processing skills are in their infancy*" (Bygate in Carter e Nunan, 2001:19). Inoltre, la scelta dei compiti deve corrispondere ai soliti criteri di validità, affidabilità, praticità. In termini di praticità il testing della interazione orale richiede un notevole dispendio di risorse.

5. Le scelte effettuate

Sono pertinenti anche per la produzione orale interattiva molte delle considerazioni fatte al punto 3. Varie scelte sono state effettuate *pro tempore* e il lavoro richiede ulteriori ricerche. Le decisioni prese riguardano l'utilizzo di una varietà di compiti con un protocollo di amministrazione *fixed frame* per fornire sia una campionatura più ricca che affidabilità maggiore; l'impiego di un formato a due candidati; l'inserimento di una componente di *warm-up* dove i contenuti sono molto familiari e poco discriminanti per superare al meglio il problema del filtro affettivo creato dalla situazione; l'uso del *role play* per creare contesti con una certa autenticità per uno studente universitario.

L'uso di un formato in cui due candidati interagiscono rappresenta il tentativo di ridurre l'impatto dell'asimmetria nella relazione di potere inerente alla situazione, che ha effetti inibenti e tende ad elicitare un'unica varietà: tale formato permette una maggiore affidabilità perché l'esaminatore non deve concentrarsi su due cose contemporaneamente, fornisce una campionatura più varia dell'interazione e riduce la natura asimmetrica del discorso (*cfr.* Ross e Berwick, 1992 e Young e Milanovic, 1992, in cui si dimostra che il discorso fra candidato e esaminatore ha tratti molto particolari). Dai risultati della ricerca si evince che nell'interazione tra pari viene aumentata sia la quantità, in termini di contributi, lunghezza dei turni e numero dei turni dai candidati, che la qualità del discorso, in termini di funzioni comunicative elicitate (Taylor, 2001, Ross e Berwick, 1992, Weir, 1993, Walsh 1999, 2002). Meno l'esaminatore deve occuparsi del ruolo di interlocutore, più affidabile e più valida diventa la valutazione. Il copione per l'esaminatore-interlocutore assicura un input simile per tutti i candidati e facilita la standardizzazione in termini di tempi e formato, importante in una certificazione di uso diffuso. Questo ha ovviamente un impatto su fattibilità e costi e ha rappresentato problemi per il committente che cercava un test più snello e veloce e meno dispendioso di tempo e risorse. Nei commenti sulla sperimentazione la prova viene criticata perché molto lunga e complicata da somministrare. È chiara la tensione fra validità e affidabilità da una parte e fattibilità ed economicità dall'altra.

Per potere valutare invece la capacità di variare il registro e di gestire una situazione di relazione asimmetrica, è stato introdotto un compito nel quale il candidato deve affrontare una situazione di *role play* in cui l'interlocutore è l'esaminatore. Tutti i fattori inibenti o di filtro affettivo (*cfr.* punto 2.3) che si possono trovare in una situazione d'esame si trovano anche nella situazione di uso della lingua e dunque questo non dovrebbe portare ad errori di misurazione. Il *roleplay* è stato concepito simulando situazioni in cui lo studente potrebbe trovarsi realmente: per esempio, in contesti di scambio Erasmus, aiutare un professore o uno studente in visita in Italia che chiede informazioni, oppure affrontare un colloquio di orientamento con un professore in paese anglofono che cerca chiarimenti e informazioni sul piano di studi. Qui subentra anche il fattore della conoscenza dell'argomento nonché l'esperienza culturale, e mette in evidenza tutta la problematica di creare un test specifico per ambienti accademici (affrontata a suo tempo dal test IELTS *cfr.* anche Clapham, 1996.)

5.1 Tempi

La durata di un test di produzione orale interattiva, con l'impiego di uno o due esaminatori (interlocutore e misuratore – fattore che è stato oggetto di commenti positivi nella sperimentazione in termini di validità e affidabilità), ha rappresentato comunque per il committente un grave scoglio in termini di praticità e costi. Da un'analisi della sperimentazione risulta che la parte di produzione orale è stata spesso esclusa dai centri sperimentatori in quanto rappresentava un enorme impiego di risorse e una grande fatica di reclutamento degli studenti candidati, spesso riluttanti a prestare così tanto tempo per scopi non percepiti come utili. Questo dato, e il fatto che il committente ha preferito tagliare parti del test per ridurre i tempi anche a discapito della validità, sottolinea come il tempo sia fattore critico, una risorsa preziosa in sé (Duguid, 2001:cap.4) e fonte di costi ingenti. È stato

tentato di mantenere un minimo di 15 minuti, ma la distribuzione del tempo all'interno del test è anche un fattore condizionato dalla formazione degli esaminatori, fase mancante nel progetto.

5.2 L'uso della tecnologia informatica

Con l'espansione dell'apprendimento a distanza, lo sviluppo della tecnologia informatica, l'uso della rete con un'accessibilità in aumento di software e hardware e un incremento nell'uso della tecnologia nell'apprendimento, l'uso della tecnologia nella valutazione è oggetto di ampia discussione. Non è tanto chiaro comunque se il mezzo rappresenta un aiuto o se ha un impatto e *washback* su metodi e contenuti, né in che modo incida sul costrutto. (Clapham 2000, Alderson e Bannerjee 2001). Nel nostro caso, anche se la prova monologica può essere informatizzata, la prova interattiva è più difficilmente inquadrabile e sono proprio le qualità intrinseche del tipo di test, le parti integrali del costrutto che rendono l'operazione difficile. La registrazione delle prove è fondamentale nella fase di validazione, potrebbe rendere la centralizzazione del processo di valutazione più fattibile ed esiste qualche ricerca che fa pensare che valutare una prova interattiva con una registrazione invece che valutarla in presenza non crei problemi di affidabilità (*cfr.*, per esempio, l'uso di dati audio o video in Brooks, 2003). Alcuni centri che hanno preso parte nella sperimentazione hanno registrato le interazioni e i dati risultanti potrebbero essere utilizzati per ulteriori studi di interesse per il settore.

6. Bibliografia

- Alderson, J.C. and Banerjee, J., 2001, "Language testing and assessment (Part 1)" *Language Teaching*, 34, pp.213-236
- Alderson, J.C. and Banerjee, J., 2002, "Language testing and assessment (Part 2)" *Language Teaching*, 35, pp.79-113
- Ambroso, S., 1995, "Misurare la competenza in italiano L2. Alcune considerazioni e proposte.", *Cuadernos de filología italiana*, 2, pp. 11-24
- Anderson, F., Brown G., Shillcock R., Yule G., 1984, *Teaching Talk*, CUP
- Anderson, L. e Ciliberti, A., 1999, *Le forme della comunicazione accademica*, Franco Angeli
- Anderson, L. e Ciliberti, A., 2002, "Monologicità e di(a)logicità nella comunicazione accademica", in Bazzanella, C. (a cura di), *Sul dialogo. Contesti e forme di interazione verbale*, Guerrini Studio, pp.92-105
- Bachman, L., 1990, *Fundamental Considerations in Language Testing.*, OUP, Oxford
- Bachman, L., 2000, "Modern language testing at the turn of the century: assuring that what we count counts", *Language Testing*, 17 (1), pp.1-42
- Bachman, L., 2002, "Some reflections on task-based language performance assessment", *Language Testing*, 19 (4), pp.453-476
- Bachman, L., Palmer, A., 1996, *Language Testing in Practice*, O.U.P., Oxford
- Bachtin, M., 1979/1988, "Il problema dei generi nel discorso", in Bachtin, M., *L'autore e l'eroe*, Einaudi, pp.245-290
- Balboni, P., "L'oralità e il saper parlare", www.educational.rai.it/corsiformazione/secondalingua, Puntata 3 e 4
- Bazzanella, C., 1994, *Le facce del parlare. Un approccio pragmatico all'italiano parlato*, La Nuova Italia
- Benvenuto, G., 2001, "L'utilizzazione di descrittori, indicatori e standard", in Gattullo, F. (a cura di), *La valutazione degli apprendimenti linguistici*, La Nuova Italia, 2001, pp.165-191
- Berceci, F. e Pallotti, G., 2002, "Conversazioni telefoniche", in Bazzanella, C. (a cura di), *Sul dialogo. Contesti e forme di interazione verbale*, Guerrini Studio, pp.117-192
- Bernardelli, A., Pellerey R., 1999, *Il parlato e lo scritto*, Bompiani
- Berretta, M., 1991, "(De)formazione del lessico tecnico nell'italiano degli studenti universitari", in Lavinio, C. e Sobrero, A.A. (a cura di), *La lingua degli studenti universitari*, Firenze, La Nuova Italia, pp.101-121
- Biber, D., et al, 1999, *Longman Grammar of Spoken and Written English*, Longman, London
- Brooks, L., 2003, "Converting an Observation checklist for use with the IELTS Speaking Test." *UCLES Research Notes 11*, pp 20-21

- Boyd, K and Davies, A., 2002, "Doctors' orders for language testers: the origin and purpose of ethical codes", *Language Testing*, 19 (3), pp.296-322
- Brighetti C., Minuz F., 2001, *Abilità del parlato*, Paravia
- Bygate, M., "Speaking." In Carter,R. and Nunan, D. eds., *The Cambridge Guide to Teaching English to Speakers of Other Languages*, CUP, Cambridge
- Bygate, M., 1998, "Theoretical perspectives on speaking", *Annual Review of Applied Linguistics*, 18, pp.20-42
- Canale, M. and Swain, M., 1980, "Theoretical bases of communicative approaches to second language teaching and testing.", *Applied Linguistics* 1, pp 1-47
- Carter, R. and McCarthy, M., 1997, *Exploring spoken English*, CUP, Cambridge
- Carton, Francis, 1992, "Limites d'une évaluation rigoureuse des capacités communicatives à l'oral : compte rendu d'une expérimentation" in Coste, D. et Moore, D. (éd.), *Autour de l'évaluation de l'oral*, numéro spécial du Bulletin CILA, n. 55, pp. 49-62
- Chalhoub-Deville M., 2001, "Language Testing and Technology: Past and Future", in *Language Learning and Technology*, 5 (2), pp.95-98
- Clapham, C., 2000, "Assessment and testing", *Annual Review of Applied Linguistics* 20, pp.147-161
- Colombo, A., 2001, "Educare al monologo", *Italiano e oltre*, 3, pp.157-159
- Consiglio d'Europa, 2002, *Quadro comune europeo di riferimento per le lingue: apprendimento, insegnamento, valutazione*, La Nuova Italia OXFORD (Traduzione dall'inglese di Franca Quartapelle e Daniela Bertocchi). Council of Europe, 2000, *A Common European Framework of Reference for Language Learning and Teaching*, Council for Cultural Co-Operation, Strasbourg
- Cook, G., 1989, *Discourse*, OUP, Oxford
- Coste, D. et Moore, D., 1992, "Examiner des modalités d'évaluation : vers plus de transparence ? ", in Coste, D. et Moore, D. (éd.), *Autour de l'évaluation de l'oral*, numéro spécial du Bulletin CILA, n. 55, pp.5-12
- Coste, D., 1992, "A partir et autour de l'évaluation de l'oral", in Coste, D. et Moore, D. (éd.), *Autour de l'évaluation de l'oral*, numéro spécial du Bulletin CILA, n. 55, pp.109-116
- D'Addio Colosimo, W., 1995, "Cammin facendo", *Italiano e oltre*, X, pp.57-60
- Dardano, M., 1994, "Testi misti", in De Mauro, T. (a cura di), *Come parlano gli italiani*, Fondazione IBM Italia, La Nuova Italia, p.175-181
- De Benedetti, A. e Gatti, F., *Routine e rituali nella comunicazione*, Paravia Scriptorium, 1999
- De Mauro, T., Mancini F., Vedovelli M., Voghera M., 1993, *Lessico di frequenza dell'italiano parlato*, Etas libri
- Della Casa, M., 1994, *Scrivere testi*, Firenze, La Nuova Italia
- Douglas, D., 2001, "Language for specific purposes assessment criteria: where do they come from?", *Language Testing*, 18/2, pp 133-147
- Duguid, A., 2001, *Anatomy of a context. English language teaching in Italy*, Granville Publishing
- Eerdmans, S. and Walsh, P., 2002, "Co-costruire l'asimmetria: language proficiency interviews.", in Bazzanella (ed), *Sul dialogo: Contesti e forme di interazione verbale*, Guerini, Milano
- Eerdmans, S. and Lipson, M., 1995, *Language teaching in English Degree Courses in Italian Universities*, CLUEB
- Egyud, G. and Glover, P. 2001 "Oral testing in Pairs – a secondary perspective", *English Language Teaching Journal* , 55/1, pp. 70-76
- Elder, C., 2001, "Assessing the language proficiency of teachers: are there any border controls?" *Language Testing* 18/2, pp 149-170
- Elder, C., Iwashita, N., McNamara T., 2002, "Estimating the difficulty of oral proficiency tasks: what does the test taker have to offer?", *Language Testing*, 19 (4), pp. 347-368
- Ellis, R., 2003, *Task-based Language Learning and Teaching*, OUP
- Fulcher, G., 1987, "Tests of oral performance: the need for data-based criteria", *ELT Journal*, Volume 41/4, pp. 287-291
- Fulcher, G., 1996, "Does thick description lead to smart tests? A data-based approach to rating scale construction", *Language Testing* 13, pp 208-238
- Garcia-Debanc, C., 1999, "Évaluer l'oral", *Pratiques*, 103/104, pp.193-121
- Hall, C., 1993, "The direct testing of oral skills in university foreign language teaching", *IRAL*, 31 (1), pp.23-38

- Halliday, MA.K., 1978, *Language as a Social Semiotic: the Social Interpretation of Language and Meaning.*, Edward Arnold, London
- Hoey, M., 1983, *On the Surface of Discourse*, Allen and Unwin, London
- Hoey, M., 1991, "Some properties of Spoken Discourse"., in Bowers,R. and Brumfit, C.J., eds, *Applied Linguistics and English Language Teaching*. Macmillan, The British Council, pp 65-84
- Hughes, A., 1989, *Testing for Language Teachers*, CUP
- Hymes, D., 1979, "On communicative Competence", in Brumfit, C.J. and Johnson, K.,eds, *The Communicative Approach to language Teaching*. OUP, Oxford
- Kenyon D.M., Malabonga V., 2001, "Comparing examinee attitudes toward computer assisted and other oral proficiency assessments" in *Language Learning and Technology*, 5 (2), pp.60-83
- Klein-Braley, C., 1991, "Ask a stupid question...: testing language proficiency in the context of research studies", in K. De Bot, R.B. Ginsberg, C. Kramersch (eds), *Foreign Language Research in Cross-cultural Perspective*, J. Benjamins P. Company, pp. 73-93
- Lazon, A., 1996, "Interlocutor support in oral proficiency interviews: the case of CASE", *Language Testing* 13, pp 151-172
- Levelt, W.J.M., 1989, *Speaking. From Intention to Articulation*, The MIT Press
- Long, M., 1983, "Native speaker/non-native speaker conversation and the negotiation of comprehensible input". *Applied Linguistics*, 4, pp 126-141
- Long, M.H., 1981, "Input, interaction and second language acquisition" in Winitz H. ed *Native language and Foreign language Acquisition*, New York academy of sciences, New York.
- Mancini, F. e Voghera, M., 1994, in De Mauro, T. (a cura di), *Come parlano gli italiani*, Fondazione IBM Italia, La Nuova Italia, pp.217-245
- Matthews, M., 1990, "The measurement of productive skills: doubts concerning the assessment criteria of certain public examinations", *ELT Journal*, Volume 44/2, pp. 117-121
- McNamara, T., 1995, "Modelling performance: opening Pandora's Box", *Applied Linguistics*, 16 (2), 139-159
- McNamara, T., 2001, "Language assessment as social practice: challenges for research", *Language Testing*, 18 (4), pp.333-34
- McNamara, T., Hill, K. and May, L., 2002, "Discourse and assessment", *Annual Review of Applied Linguistics*, 22, pp.221-242
- Messick, S., "Validity", in Linn, R. (ed.), *Educational Measurement*, Macmillan, New York
- Mislevy, R., Steinberg, L., Almond, R., 2002, "Design and analysis in task-based language assessment", in Arnold, *Language Testing*, 19 (4), pp. 477-496
- Monnerie, A., 1992, "La transparence dans une certification officielle décentralisée", in Coste, D. et Moore, D. (éd.), *Autour de l'évaluation de l'oral*, numéro spécial du Bulletin CILA, n. 55, pp.25-34
- Morrow, K., 1977, *Techniques of evaluation for a Notional Syllabus*, London Royal Society of Arts
- Nambiar, M.K. and Goon, C., 1993, "Assessment of oral skills: a comparison of scores obtained through audio recordings to those obtained through face-to-face evaluation", *RELC Journal*, 24 (1), pp.15-31
- Nencioni, G., 1983, *Di scritto e di parlato*, Zanichelli
- Nonnon, E., 2002, "Des interactions entre oral et écrit: notes, canevas, traces écrites et leurs usages dans la pratique orale", *Pratiques*, 115/116, Décembre, pp. 73-92
- Norris, J. M., 2001, "Concerns with computerized adaptive oral proficiency assessment" in *Language Learning and Tecnology*, 5 (2), pp.99-105
- North, B.2001 Relating Language Examinations to the Common European Framework of Reference for Languages: Learning, Teaching, Assessment (CEF):
- Language Policy Division Strasbourg.(DGIV/EDU/LANG (15)
- North, B. and Schneider, G., 1998, "Scaling descriptors for language proficiency scales" *Language Testing*, 15 (2), pp.217-263
- Pugliese, R. e Ciliberti, A., 1999, "Analisi dei questionari studenti" in Ciliberti, A. e Anderson, L., *Le forme della comunicazione accademica*, Franco Angeli, pp.249-262
- Raffaldini, T., 1988, "The use of situation tests as measures of communicative ability", *Studies in Second Language Acquisition*, 10 (2), pp. 197-216
- Richterich, R., 1992, "Question de temps pour l'évaluation de l'oral", in Coste, D. et Moore, D. (éd.), *Autour de l'évaluation de l'oral*, numéro spécial du Bulletin CILA, n.55, pp.13-16
- Ross, S., and Berwick, R., 1992, "The discourse of accommodation in oral proficiency interviews", *Studies in second Language Acquisition*, 14/2, pp 159-176

- Rossini Favretti, R., 2002, *Un'introduzione alla linguistica applicata*, Patron Editore
- Rossini Favretti, R. et al., 2005, "Somministrazione informatizzata di prove di produzione orale: ipotesi di fattibilità in ambito di certificazione linguistica per scopi accademici", in <http://amsacta.cib.unibo.it/>
- Sabatini, F., 1999, " 'Rigidità-esplicitzza' vs 'elasticità-implicitzza': possibili parametri massimi per una tipologia dei testi", in Atti del Convegno interannuale della SLI, Linguistica Testuale Comparativa, G. Skytte e F. Sabatini (a c. di), Copenhagen, Museum Tusulanum Press.
- Scovel, T., 1998, *Psycholinguistics*, Oxford University Press, Oxford
- Simone, R., 2000, *La Terza Fase*, Roma/Bari, Editori Laterza
- Skehan, P. and Forster, P., 1997, "Task type and task processing conditions as influences on foreign language performance", *Language Teaching Research* 1 (3), pp 185-211
- Skehan, P., 2003, "Task-based instruction", *Language Teaching*, 36 (1), pp.1-14
- Sobrero, A.A., 1988, "Avvertenze per l'uso del parlato" in *Italiano e Oltre*, 3, pp.218-222
- Sornicola, R., 1994, "Quattro dimensioni nello studio del parlato", in De Mauro, T. (a cura di), *Come parlano gli italiani*, Fondazione IBM Italia, La Nuova Italia, pp.111-130
- Swales, J., 1990, *Genre Analysis. English in Academic and Research Settings*, Cambridge University Press, pp.21-32
- Taylor, L., 2001, "The speaking test format: recent studies", *UCLES Research Notes*, 6, 15-17
- Taylor, L., 2000, "Investigating the paired speaking test format", *UCLES Research Notes*, 2, 14-15
- Turner, J., 1998, "Assessing speaking", *Annual Review of Applied Linguistics*, 18, pp.192-207.
- Underhill, N., 1987, *Testing Spoken Language*, CUP
- Upshur, J.A., and Turner, C.E., 1999, "Systematic effects in the rating of second-language speaking ability: test method and learner discourse", *Language testing*, 16 (1), pp.82-111
- Van Ek, J. and Trim, J., 2000, *The Vantage level*, CUP, Cambridge
- Van Ek, J., and Trim, J., 1990, CUP, Cambridge
- Voloshinov, V.N. 1995. *Marxism and the Philosophy of Language, Bakhtinian Thought - an Introductory Reader*, S. Dentith, L. Matejka & I.R. Titunik, (trans), London, Routledge.
- Voghera, M., 1992, *Sintassi e intonazione nell'italiano parlato*, Il Mulino
- Voghera, M., 1993, "Le scelte per la costituzione del corpus. Le variabili testuali e pragmatiche", in De Mauro et al., *Lessico di frequenza dell'italiano parlato*, Fondazione IBM Italia, Etas Libri, pp.32-38
- Voghera, M., 2001, "Teorie linguistiche e dati di parlato" in Società Linguistica Italiana, *Dati empirici e teorie linguistiche*, Bulzoni, pp. 75-95
- Walsh, P., 1999 (b), "Can a language interview be used to measure interactional skills?", *Working Papers in TEFL*, vol 2, CALS, University of Reading
- Walsh, P., 1999(a), "Investigating Interactional Variation in Interview-Type Speaking Tests", *Dialogue analysis VII: Working with Dialogue, Selected papers from the 7th IADA Conference, Birmingham 1999*, Max Niemeyer Verlag, Tubingen
- Walsh, P., 2000, "Interviewer reactivity in language Proficiency Interviews", *Quaderni del Dipartimento di Linguistica* 10, Università di Firenze 10, pp 195-212
- Weir, C., 1990, *Communicative Language Testing*, Prentice Hall, London
- Weir, C., 1993, *Understanding and developing Language Tests.*, Prentice Hall, London
- Wigglesworth, G., 1997, "An investigation of planning time and proficiency level on oral test discourse", *Language Testing*, 14 (1), pp.85-106
- Young, R., 2002, "Discourse approaches to oral language assessment", *Annual Review of Applied Linguistics*, 22, pp.243-262
- Young, R.F., 2000, "Interactional competence: challenges for validity" disponibile all'URL <http://www.wisc.edu/english/rfyoung/#RecentPublications>, pp.1-15
- Young, R.F., 2001, "The entextualization of talk" disponibile all'URL <http://www.wisc.edu/english/rfyoung/#RecentPublications>, pp.1-15
- Young, S. and Milanovic, M., 1992, "Discourse variation in oral proficiency interviews", *Studies in Second language Acquisition*, 14/4, pp 403-424
- Yuan, F. and Ellis, R., 2003, "The Effects of Pre-Task Planning and On-Line Planning on Fluency, Complexity and Accuracy in L2 Monologic Oral production", *Applied Linguistics*, 24 (1), pp.1-27
- Yule, G., 1996, *Pragmatics*, OUP, Oxford.

- Zagrebelsky, M. T. Prat “La valutazione dei livelli di ingresso all’università” in Gattullo, F. (a cura di), 2001, *La valutazione degli apprendimenti linguistici*, La Nuova Italia, pp.267-282
- Zeidner, M. and Bensoussan, M., 1988, *Language testing*, 5 (1), pp.100-114
- Zorzi, D., 1994, “Le digressioni nella lezione accademica italiana: un’analisi contrastiva”, in Bruni, S. (a cura di), *La formazione linguistica degli studenti Erasmus in Italia*, Atti del II Seminario Permanente dei Centri Linguistici, Università per Stranieri di Siena, pp.13-25
- Zuccherini, R., 1988, *Manuale del parlare*, La Nuova Italia
- Zuengler, J., and Bent, B., 1991, “Relative knowledge of content domain: an influence on native-non-native conversations”, *Applied Linguistics* 12/4, pp 397-415

Per contatti: *bavieri@cilta.unibo.it*, *polselli@cilta.unibo.it*, *duguid@unisi.it*