

Livelli diversi di competenza nella gestione dell'italiano L2. Ipotesi dall'analisi di un corpus di *Serena Ambroso* e *Elisabetta Bonvino*

I

Introduzione

Se lo studio dell'italiano parlato è un ambito della ricerca linguistica relativamente recente e in piena espansione¹, le ricerche sull'italiano dei non nativi (italiano L2) sono ancora più recenti e il loro ambito è tuttora in via di definizione. I lavori in questo campo² riguardano principalmente lo studio teorico longitudinale di alcuni fenomeni morfosintattici dell'interlingua. La descrizione dell'italiano L2 ha anche un interesse didattico in quanto permette di orientare le pratiche di insegnamento e di valutazione dell'apprendimento. Il nostro contributo si inserisce nel quadro dell'analisi del corpus di italiano parlato da non nativi che stiamo costituendo a Roma Tre a partire dai dati dell'Ufficio della certificazione dell'italiano come L2.

Lo studio dei corpora, è noto, oltre a permettere la descrizione di fenomeni realmente attestati, facilita l'identificazione della relazione fra fenomeni appartenenti a livelli di analisi diversi, in particolare permette di rapportare fenomeni di livello micro-sintattico a quelli macro-sintattici³; consente anche di cogliere la relazione esistente tra lessico utilizzato e grammatica; infine, favorisce l'esame non solo degli aspetti qualitativi ma anche di quelli quantitativi dei fenomeni osservati. Tuttavia, specie per chi lavora su corpora di lingua parlata, emergono diversi problemi da risolvere riguardanti soprattutto il modo di trascrivere, presentare e analizzare i dati. Sono infatti ad oggi ancora aperte le discussioni su quale sia un sistema efficace di trascrizione dei dati che tenga conto della complessità, della non linearità, della simultaneità dei livelli di analisi, e che renda soprattutto possibile l'interrogazione automatica del corpus. Problemi che risultano assai più complessi nel campo della L2, dove si rende necessaria, ad esempio, la codifica di tutta una serie di espressioni *non target*.

In questo articolo intendiamo analizzare alcuni fenomeni che caratterizzano i testi parlati in italiano L2, cercando di presentare una metodologia di analisi a nostro avviso fruttuosa e di mostrare alcuni primi risultati dell'analisi stessa. Obiettivo a lungo termine di questo lavoro è quello di contribuire ad affinare gli strumenti di valutazione della competenza linguistica di non nativi.

2 COCERIT

2.1. Il corpus

I dati confluiti nel corpus di italiano L2 COCERIT (Corpus della certificazione dell'italiano) provengono dall'Ufficio della certificazione dell'italiano come L2, struttura attiva dal 1994, incardinata nel Dipartimento di Linguistica di Roma Tre. A questo riguardo, e soprattutto per maggiore completezza, riteniamo necessario dare alcune informazioni dettagliate sulla attività di certificazione da noi svolta presso tale struttura.

Insieme alle due Università per stranieri di Perugia e di Siena, l'Università di Roma Tre è ente certificatore ufficiale dell'italiano come L2, riconosciuto dai ministeri degli Esteri, dell'Istruzione e dell'Università, del Lavoro e delle politiche sociali. Come tale, l'Ufficio della certificazione progetta, realizza e diffonde all'estero e in Italia prove di certificazione della competenza in italiano, comunque acquisita da parte di parlanti non nativi. Al momento sono operativi due sistemi di prove, adeguati ai descrittori dei livelli di competenza d'uso del Quadro comune europeo di riferimento (QCER): *ele.IT* per il livello B1, e *IT* per il livello C2. È anche in corso di sperimentazione – sarà operativo nel 2009 – un terzo sistema, *int.IT*, che attesterà una competenza in italiano di livello B2. Attualmente, pertanto, il corpus dispone di testi parlanti da stranieri che volontariamente hanno affrontato e superato la prova di produzione del parlato di uno dei due certificati.

Tutte le prove, per tutti i livelli di competenza certificati, propongono autentiche situazioni d'uso dell'italiano e – come già accennato – in linea con tutti i descrittori forniti dal QCER. Tenendo ben presenti quelli relativi ai livelli di competenza di nostro interesse, le prove attualmente operative sono state articolate in maniera diversa.

L'esame di livello B1 prevede 4 prove (ascolto, lettura, usi dell'italiano, comunicazione faccia a faccia) da svolgere in 2 ore, oltre ai 10 minuti necessari per l'intervista (comunicazione faccia a faccia). Poiché è nostro convincimento che chi raggiunge una competenza di questo livello impara a fare in italiano cose più o meno dello stesso grado di complessità, riteniamo "monolitica" una capacità di uso a questo stadio di sviluppo. Pertanto i candidati devono affrontare le prove tutte insieme. Il giudizio valutativo è basato sul rendimento globale e, in analogia con il sistema valutativo italiano, il certificato viene rilasciato solo a chi ha raggiunto il 60% del punteggio globale.

L'esame di livello C2 è invece articolato in 5 diverse prove (ascolto, lettura, produzione scritta, usi e forme dell'italiano, comunicazione faccia a faccia) che possono essere sostenute anche singolarmente. Nel suo complesso, l'esame dura 5 ore e 30 minuti, oltre al tempo per lo svolgimento della pro-

va di produzione del parlato, di solito 15 minuti. Il certificato dà conto dei risultati raggiunti nelle singole prove superate. Questa particolarità nasce dalla considerazione che la conoscenza di una L2 a questo livello, quasi vicina a quella di un parlante nativo, deve essere considerata “modulare” poiché in base all’esperienza del singolo apprendente e a fattori legati alla sua personalità, una competenza di livello avanzato può essere raggiunta non in tutti gli stessi ambiti e contesti d’uso e, soprattutto, non in entrambe le varietà diamesiche⁴. Analogamente all’altro sistema di certificazione, è valutata positivamente ogni prestazione che raggiunge il 60% del punteggio attribuito alla prova in questione⁵.

Aspetti qualitativi

Per meglio illustrare le caratteristiche del nostro corpus di parlato e al fine di metterne in evidenza l’omogeneità, è opportuno entrare anche nei dettagli della prova di comunicazione faccia a faccia del parlato che, essendo differenziata per livelli di competenza, è pertanto organizzata tenendo presenti le competenze descritte nel QCER⁶ per la capacità “Interazione orale generale”:

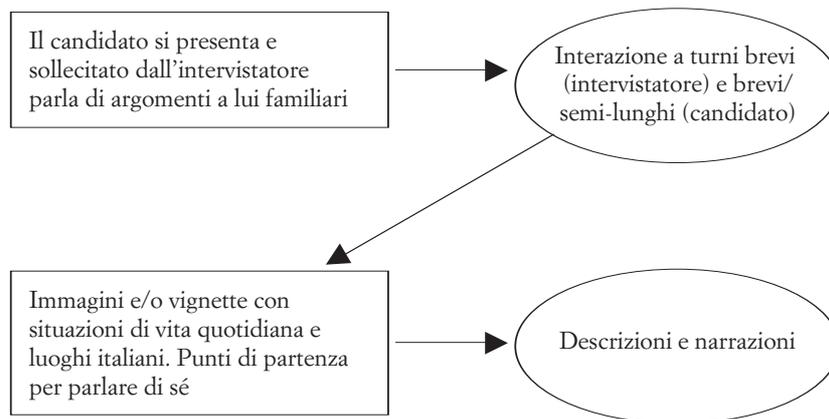
- Livello B1. È in grado di comunicare con discreta sicurezza su argomenti familiari, di routine o no, che lo/la interessino o si riferiscano alla sua situazione. Scambia informazioni, le controlla e le conferma, fa fronte a situazioni meno frequenti e spiega perché qualcosa costituisce un problema. Esprime il proprio pensiero su argomenti più astratti, culturali, quali film, libri, musica ecc. È in grado di utilizzare un’ampia gamma di strumenti linguistici semplici, per far fronte a quasi tutte le situazioni che possono presentarsi nel corso di un viaggio. Interviene, senza bisogno di una precedente preparazione, in una conversazione su questioni familiari, esprime opinioni personali e scambia informazioni su argomenti che tratta abitualmente, di suo interesse personale o riferiti alla vita di tutti i giorni (ad es. famiglia, hobby, lavoro, viaggi e fatti di attualità).
- Livello C2. Ha una buona padronanza di espressioni idiomatiche e colloquiali ed è consapevole delle relative connotazioni. Esprime con precisione sottili sfumature di significato, usando con ragionevole correttezza diversi modificatori del discorso. Aggira le difficoltà ristrutturando il discorso con disinvoltura tale che l’interlocutore quasi non se ne accorge.

In entrambi gli esami, il colloquio viene registrato su supporto audio. Questa procedura, pur con tutti i suoi limiti⁷, è indispensabile per poter valutare le prestazioni in sede centrale e quindi avere dati audio da trasferire in COCERIT. La situazione d’esame prevede la presenza di una persona che interagisce direttamente con il candidato e di un osservatore esterno che assicura la corretta organizzazione e svolgimento dell’intervista. Il candidato

è chiamato a interloquire con l'intervistatore presentandosi e rispondendo a domande finalizzate a stabilire un contesto comunicativo favorevole.

In particolare, per il livello B1 è prevista una breve interazione in cui il candidato è invitato a presentarsi e a parlare di argomenti familiari. Nella seconda fase della prova, in seguito agli stimoli visivi dati (disegni e foto), il candidato sceglie liberamente di illustrare un argomento che rientra nel suo campo di interesse sia nell'ambito privato (famiglia, casa, cucina, feste, arredamento ecc.), sia nell'ambito pubblico (viaggi, mezzi di trasporto, acquisti, ristorante, spettacoli ecc.) che occupazionale (trasporti, lavoro, riunioni, procedure amministrative, pubblicità ecc.)⁸.

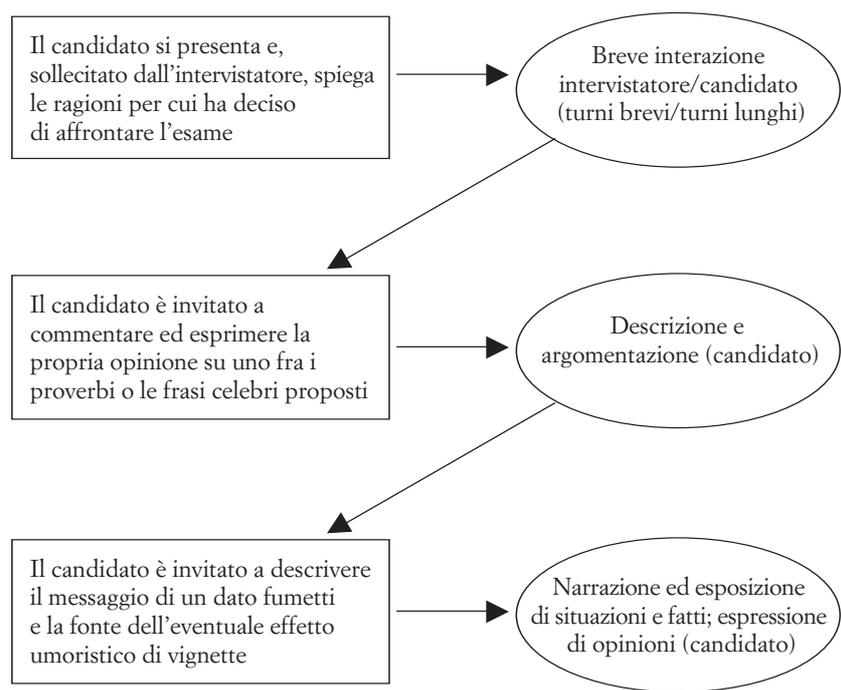
FIGURA 1
Schema della prova "Faccia a Faccia": B1



Per quanto riguarda invece il livello C2, la prova è articolata come segue: dopo una breve interazione in cui il candidato è invitato a presentarsi e a spiegare le ragioni che lo hanno spinto a sottoporsi alla prova, gli viene chiesto di commentare un breve testo scritto (che riporta frasi celebri, proverbi, aforismi, ecc.) esprimendo e anche argomentando la sua opinione in proposito. Al candidato vengono poi proposti vari stimoli visivi (vignette umoristiche su aspetti culturalmente rilevanti nella società attuale, quali il rapporto moglie-marito, genitori-figli, il ruolo del denaro, i vicini di casa, ecc.) fra i quali è chiamato a sceglierne uno, a fornirne una descrizione e a spiegarne il contenuto linguistico o la situazione culturalmente paradossale dell'immagine. Da questa organizzazione si ricavano testi parlati di più varia tipologia e più ampia dimensione rispetto a quelli del livello più basso: si tratta quasi sem-

pre di scambi con turni lunghi che in questo caso, per la struttura argomentale, sono molto vicini al tipo monologico.

FIGURA 2
Schema della prova "Faccia a Faccia": C2



Vista la situazione di esame, non possiamo certo affermare che il corpus in questione presenti esempi di lingua spontanea *stricto sensu*. La situazione è autentica in quanto si tratta di un esame, ma è doppiamente asimmetrica riguardo al rapporto fra gli interagenti, non solo per il ruolo dei partecipanti (esaminatore/esaminato) ma anche per il diverso grado di padronanza della lingua italiana (nativo/non nativo). Quest'ultimo aspetto, però, come vedremo, darà un contributo positivo allo sviluppo del discorso, soprattutto nel caso di candidati di livello B1.

Nonostante ciò, tutti gli scambi registrati possono essere classificati come autentici, in situazione reale e appartenenti alla categoria di "parlato-parlato" in cui i partecipanti condividono il contesto, programmano ed eseguono i propri turni in maniera sincrona, come pure in maniera sincrona

percepiscono e decodificano il parlato. Il tipo di italiano che va a confluire nel nostro corpus, quindi, si può senza esitazione definire: parlato-parlato⁹.

Come emerge da quanto finora detto, tuttavia, la tipologia di discorso del corpus può essere per lo più ricondotta a quella di un'intervista in cui i partecipanti sanno di essere registrati. Per la varietà dei testi e per la modalità della loro raccolta (colloquio registrato), la situazione (esame), il tipo di input proposto (stimoli visivi o linguistici a seconda del livello di competenza dei candidati), i temi trattati e la tipologia testuale prodotta (testi prevalentemente descrittivi o narrativi per il livello B2; testi descrittivi, espositivi, argomentativi e narrativi per il livello C2), il corpus COCERIT può essere considerato molto omogeneo; ed è proprio l'omogeneità dell'insieme dei testi che consente il confronto fra gli stessi. In sintesi, si tratta di conversazioni a due partecipanti che, quando il livello dei candidati lo permette (C2), sono anche di carattere monologico, poiché il candidato produce turni di discorso di una certa lunghezza. Infine, un altro tratto di omogeneità viene conferito a COCERIT dal fatto che nel corpus confluiscono solo prestazioni di candidati che hanno superato la prova. Se per C2 tale tratto è ininfluenza, essendo il livello della prova di certificazione quello più alto di tutti, per il livello B1 invece è rilevante perché la nostra scelta garantisce che il parlato italiano L2 sia almeno – per la maggioranza dei casi – del livello certificato¹⁰.

Aspetti quantitativi

Al momento attuale non tutti i testi parlati dei candidati che hanno ottenuto uno dei certificati di Roma Tre sono confluiti in COCERIT. Nel costruire il corpus prestiamo attenzione a mantenere per quanto possibile lo stesso ordine di grandezza riguardo al numero dei parlanti registrati, mentre non è possibile mantenere per entrambe le categorie lo stesso numero di ore di registrazione data la diversa durata della prova. La TAB. 1 che segue dà conto di alcuni aspetti quantitativi:

TABELLA 1
Corpus COCERIT

Anno di registrazione	2005 – 2006 – 2007
Numero di ore di registrazione	11 h 35'
Numero dei parlanti non nativi divisi per livello	C2 26 B1 27
L1 rappresentate	inglese, tedesco, arabo, bulgaro, spagnolo, francese, giapponese, olandese, ungherese, polacco, portoghese, rumeno, russo, svedese

Per gli aspetti che riguardano i parlanti non nativi, il nostro corpus è davvero molto vario: essi si differenziano per età, origine geografica, livello di scolarizzazione, ambiente sociale di appartenenza, e soprattutto per la madre lingua e la situazione di apprendimento. Questo ultimo aspetto non è preso in considerazione nella nostra ricerca: l'apprendimento può essere avvenuto in contesto di lingua straniera o lingua seconda, in modo spontaneo o guidato. Quanto a noi interessa sono le produzioni in italiano L2 raccolte, però, nel modo sistematico appena illustrato. COCERIT è costituito da testi parlati, registrati prima su supporto magnetico, digitalizzati, e in seguito, trascritti secondo le convenzioni del GARS (*Groupe aixois de recherches en syntaxe*)¹¹.

3

Configurazioni di discorso e griglie

Per l'analisi del corpus abbiamo utilizzato una rappresentazione visiva che tiene conto delle specificità del testo parlato, fornendone una rappresentazione non appiattita sulla dimensione lineare. Nei testi orali, la linearità può essere interrotta a più riprese e l'obiettivo comunicativo può essere raggiunto anche a tappe; infatti, una sequenza spesso è resa per approssimazioni progressive, nelle quali si intravedono le tracce del processo di produzione. Lo strumento che permette di dare conto simultaneamente della dimensione sintagmatica e paradigmatica è la rappresentazione in griglie delle sequenze, un sistema messo a punto da Claire Blanche-Benveniste e collaboratori¹². Si noti comunque che le griglie non costituiscono un'alternativa alla trascrizione o alla annotazione. Appartengono alla fase interpretativa del testo, che – a nostro avviso – non può che essere successiva a quella della trascrizione stessa. Il loro punto di partenza è quindi la trascrizione accurata e affidabile del corpus¹³.

La trasposizione in griglie si basa su due semplici criteri: sull'asse orizzontale, si riporta la sequenza di posizioni – sintagmatiche – che definisce la costruzione; sull'asse verticale si riporta invece la eventuale realizzazione paradigmatica di ogni posizione. I frammenti vengono allineati dove si situerebbero nella costruzione sintagmatica, cioè secondo un criterio di natura prevalentemente sintattica.

L'insieme della costruzione e di questi frammenti costituisce una *configurazione di discorso*, di cui la griglia rende conto. Le griglie forniscono una rappresentazione del testo in cui la “forma” stessa della configurazione ha rilevanza¹⁴. Nell'esempio in TAB. 2, tratto da un corpus di L1, c'è una costruzione matrice (*construction maximale*) “giovedì santo hanno trasmesso una cosa della RAI sul Vaticano”, si tratta di una costruzione costituita da predicato, argomenti ed aggiunti in cui si nota un processo di produzione graduale basato su alcuni di questi fenomeni: anticipazione, ripetizione, rifo-

mulazione, correzione, completamento, arricchimento di struttura. La costruzione matrice può essere preceduta o seguita da frammenti della costruzione stessa. L'insieme della costruzione matrice e dei frammenti che la precedono, la seguono o la riprendono costituisce una configurazione di discorso.

TABELLA 2
Griglia *Radio* corpus Bonvino/ArCoDip – italiano L1

l'altro giovedì giovedì scorso no giovedì santo hanno trasmesso una cosa della RAI sul Vaticano

1	l'altro	giovedì				
2		giovedì	scorso			
3	no					
4		giovedì	santo	hanno trasmesso	una cosa della RAI	sul Vaticano

L'estensione delle configurazioni di discorso può variare, andando dal livello di costituente (o parte di esso) al livello della clausola o a livello di testo. Ad un livello più ampio, si ha una configurazione di discorso⁵ quando una stessa costruzione viene realizzata più di una volta all'interno del discorso, cioè quando nel testo viene mantenuta la stessa forma di linearizzazione con materiale lessicale semanticamente affine⁶, come nell'esempio in TAB. 2, sempre tratto da un corpus di italiano L1.

Come si può notare nelle griglie riportate, le configurazioni possono essere costituite da “disfluenze” (nel caso riportato in TAB. 2) ma anche da strutture perfettamente ben formate e che veicolano un significato (come in TAB. 3).

TABELLA 3
Griglia *Chiacchiere* corpus Bonvino – italiano L1

lui in quel momento mi rifiutava lui quella mia malinconia gli dava fastidio lui la mia malinconia non l'accettava

1	lui	in quel momento		mi	rifiutava
2	lui	quella mia malinconia		gli	dava fastidio
3	lui	la mia malinconia	non	l'	accettava

Trasponendo ampie porzioni di testo in griglie, risulta in maniera evidente che il lavoro di costruzione di una data sequenza (in L1 come in L2) può cominciare molto prima della realizzazione completa della sequenza stessa. Una stessa configurazione può essere ripetuta ad intervalli, come un ritornello, conferendo coesione o comunque una specifica architettura al testo.

Ci si rende conto che le produzioni orali, che in una trascrizione lineare potrebbero sembrare “sgangherate”, hanno un’organizzazione in configurazioni ricorrenti, convenzionalizzate, che veicolano funzioni, per le quali a volte non c’è altro modo di essere espresse nella lingua. In questo senso, riteniamo che le configurazioni di discorso possano essere considerate delle *costruzioni*, nel senso della *Construction Grammar*, cioè delle associazioni convenzionalizzate di una forma e di un significato che possono essere più o meno complesse e più o meno lessicalmente specificate¹⁷.

In Ambroso e Bonvino (in stampa)¹⁸ abbiamo mostrato che sia nella L1 che nella L2 si possono riscontrare delle configurazioni di discorso associate a determinate funzioni, e che pertanto le configurazioni rientrano nella definizione di *costruzione*. L’ipotesi che le configurazioni di discorso possano essere considerate come un tipo di costruzioni implica che si assuma il discorso come livello di analisi, e che di conseguenza si estenda la nozione di costruzione anche a questo livello, come proposto da Östman (2005)¹⁹.

Le griglie sono uno strumento ideale per la rappresentazione del dato linguistico parlato e per l’identificazione delle configurazioni di discorso. Riteniamo pertanto estremamente opportuno estendere l’impiego di questo strumento, la cui efficacia è stata dimostrata in molti studi incentrati su varie L1²⁰, all’analisi della L2 per una serie di motivi.

Innanzitutto perché malgrado lo sviluppo e l’interesse per gli studi sul parlato (in ambiti quali l’insegnamento delle lingue, la valutazione e, talvolta, gli studi sull’interlingua), i riferimenti alla grammatica della lingua *target* sono di solito espressi avendo come riferimento la norma scritta. Non si tiene cioè conto dei fenomeni che riguardano il modo di produzione orale, quali le disfluenze, la ricerca di lessico, e tutti gli aspetti che costituiscono la frammentarietà del parlato. Riteniamo invece questi fenomeni molto importanti per comprendere pienamente il funzionamento della produzione orale in L2. Mantenendo la leggibilità del testo, le griglie rappresentano con grande chiarezza questi fenomeni.

In secondo luogo, come già accennato, codificando fenomeni legati alla produzione orale, rispetto alle trascrizioni lineari, le griglie permettono di evidenziare le configurazioni di discorso, costruzioni che veicolano funzioni non solo non riconosciute dalla grammatica normativa, ma spesso nemmeno individuate. L’attenzione da noi rivolta a tali fenomeni è dunque funzionale a un ripensamento della struttura della grammatica (e il concetto stesso di competenza linguistica), che individua costruzioni al di là delle divisioni abituali in livelli di analisi.

Infine, questo approccio alla lingua parlata da non nativi permette di evidenziare meccanismi di costruzione del significato peculiari dei diversi livelli di competenza d’uso di una L2. Nella nostra ricerca abbiamo notato, ad

esempio, che fenomeni pur presenti nella L₁ dei parlanti vengono trasferiti nella L₂ solo ad un livello avanzato (cfr. 4.1 Livello B₁ e 4.2 La vaghezza).

4 Analisi della L₂

I fenomeni che affronteremo riguardano le dimensioni sintattica, lessicale, pragmatica e testuale del discorso. Si tratta di fenomeni che rientrano in due macro-categorie che abbiamo definito “gestione dell’informazione” e “gestione del lessico”. Si tratta comunque di fenomeni difficilmente isolabili, in quanto interdipendenti, e che riteniamo opportuno considerare nella valutazione della competenza di un non nativo dal momento che sono parte della competenza linguistica di un parlante nativo.

4.1. Gestione dell’informazione

Livello B₁

Un primo fenomeno che, nella sua ovvietà, sembra comunque stabilire un netto spartiacque tra parlanti di livello B₁ e parlanti di livello C₂ è la lunghezza dei turni. I parlanti non nativi di livello B₁ si esprimono attraverso turni poco sviluppati. Infatti un turno supera raramente la costruzione verbale e ancora più raramente la configurazione di discorso. Il candidato a questo livello si limita per lo più a rispondere alle domande dell’interlocutore con strutture nominali o con risposte olofrastiche del tipo sì/no. Tale tendenza è senz’altro legata al contesto di interazione: l’esame.

Questo dato è confermato sia che ci si trovi all’inizio dell’interazione in una fase di presentazione, nella quale la risposta olofrastica e la risposta parziale consistente in un’unica parola o sintagma sono assolutamente naturali, sia più avanti nell’esame, quando le varie consegne sono state ideate per permettere all’apprendente di esprimersi, conformemente al suo livello, con turni il più possibile sviluppati in modo tale da consentire una valutazione della sua capacità di gestione della lingua. Pur con grandi variazioni individuali, circa il 30% dei turni di un candidato di livello B₁ durante l’esame è costituito da risposte olofrastiche o da risposte costituite da un unico sintagma (con punte del 46% e oltre in alcuni casi). Si vedano le griglie in TABB. 4 e 5.

Sempre a livello B₁, i rimanenti turni del gruppo di parlanti non nativi analizzato sono costituiti da un’unica costruzione verbale semplice, realizzata attorno ad un unico verbo di cui il parlante cerca di saturare tutte le posizioni argomentali. Queste ultime realizzazioni, più articolate rispetto a quelle composte da un unico sintagma, costituiscono circa il 40% dei turni del

candidato in un esame B1. Spesso infatti i candidati producono queste strutture sintattiche “canoniche”, non marcate, composte da *soggetto – verbo – oggetto* (si vedano le griglie in TABB. 6, 7 e 8). Il ricorso a queste strutture canoniche potrebbe essere ricondotto all’esigenza del candidato di esprimersi in modo corretto, cioè come gli è stato insegnato. Per non sbagliare in contesto d’esame, il candidato potrebbe quindi preferire attenersi a questo tipo di costruzioni semplici, in cui tutti gli argomenti sono espressi, e che considera ben formate in italiano. Il fatto che anche parlanti che hanno appreso l’italiano in contesto di L2 senza istruzione esplicita adottino questa strategia discorsiva, potrebbe anche farci pensare però, come suggerito dalla *teoria della processabilità*, che questo fenomeno corrisponda ad una tendenza più generale ad apprendere in L2 prima le strutture canoniche e meno marcate²¹.

TABELLA 4
Griglia *Campagna o città* corpus COCERIT – L2 (parlante B1, USA)

Dove vivi? In campagna o in città? No, in città. In città. Sì. Hai sempre vissuto in città? Sì.

1	L1	dove vivi		in campagna
2				in città
3	L2		no	
4				in città
5	L1			in città
6	L2		sì	
7	L1	hai sempre vissuto		in città
	L2		sì	

TABELLA 5
Griglia *Maddalena* corpus COCERIT – L2 (parlante B1, USA)

ti piace la Maddalena? La Maddalena non molto non molto

1	L1	ti piace	la Maddalena?	
2	L2		la Maddalena	non molto
3				non molto

TABELLA 6
Griglia *Matrimonio* corpus COCERIT – L2 (parlante B1, Colombia)

*e sono venuti tutti i tuoi parenti dalla Colombia?
No sono venuti solo la mia sorella de la Germania e un'amica da Berlino*

1	L1		e sono venuti	tutti	i	tuoi parenti	dalla Colombia?
2	L2	no					
3			sono venuti	solo	la	mia sorella	de la Germania
4				e	un'	amica	da Berlino

TABELLA 7

Griglia *Matrimonio* corpus COCERIT – L2 (parlante B₁, Colombia)

*ah che bellina tua sorella è più grande o più piccola di te
è più piccola un anno*

5	L1	ah	che bellina			
6		tua sorella	è	più	grande	
7				o più	piccola	di te
8	L2		è	più	piccola	un anno

TABELLA 8

Griglia *P. B.* corpus COCERIT – L2 (parlante B₁, USA)

*Come ti chiami? Mi chiamo P. B. E da dove vieni? Vengo dagli Stati Uniti. Dal Michigan.
Ok eh quanti anni hai? Io ho 42 anni. E da quale eh da quanto tempo sei in Italia? Sono qui
per quasi sei anni adesso quasi sei anni.*

1	L1	come	ti chiami?		
2	L2	mi	chiamo	P. B.	
3	L1	e da dove	vieni?		
4	L2		vengo	dagli Stati Uniti	
5				dal Michigan	
6	L1	eh quanti anni hai?			
7	L2	io	ho	quarantadue anni	
8	L1	da quanto tempo sei		in Italia?	
9	L2		sono	qui per quasi sei anni	
10				adesso quasi sei anni	

Non troviamo dunque a questo livello architetture testuali più complesse. Anche quando il turno supera la singola costruzione, è in genere costituito da 2 o 3 clausole, tutte molto semplici, collegate per lo più in maniera parattica, in modo che possano essere ricondotte ad un'unica configurazione.

Per esempio, la candidata JLL, di origine latino-americana, nell'interazione con l'esaminatore prende la parola 32 volte, ma solo una volta, nella griglia riportata qui di seguito (TAB. 10), ha un turno relativamente lungo e sintatticamente complesso.

Il parlante non nativo di livello B₁ non è sufficientemente autonomo e risulta completamente dipendente dal suo interlocutore nella gestione dei turni e degli argomenti trattati, come pure nella ricerca del lessico (cfr. 4.2). Questo si riflette anche nella distribuzione dell'informazione. La preoccupazione principale del candidato sembra essere quella di produrre un enunciato ben formato dal punto di vista sintattico e lessicale. La volontà di saturare tutte le posizioni argomentali richieste dal verbo, resa più complessa dalla difficoltà di accesso al lessico, rende l'architettura del testo semplice e la distribuzione dell'informazione poco evidente. Negli enunciati analizzati,

TABELLA 9
Griglia *Viaggi* corpus COCERIT – L2 (parlante B1, USA)

*non conosci quindi la l'Italia settentrionale la parte del nord non la conosci?
sono stato a Venezia a Pisa sono andato anche c'è un altro paese Florence Flor- Flor-*

1	L1	e	non		conosci	quindi	la
2							l'Italia settentrionale
3							la parte del nord
4			non	la	conosci		
5	L2				sono stato		a Venezia
6							a Pisa
7					sono andato	anche	
8					c'è		un altro paese
9							Florence
10							Flor-
							Flor-

TABELLA 10
Griglia *Matrimonio* corpus COCERIT – L2 (parlante B1, Colombia)

*no siamo sposati al comune – in comune – eh però lui ha dovuto andare con le stampelle – perché – la – una settimana prima del matrimonio lui ha caduto *en* un pozzo ebb *se* ha fratturato *el* femore – è per quello lo hanno operato subito – *porque* – lui gli ha detto giovedì devo sposare eh – lui è andato alle stampelle al comune*

1	no						
2			siamo sposati				al comune
3							in comune
4	però	lui	ha dovuto andare	con le stampelle			perché [...]
5							
6	è per quello che non lo hanno operato subito						
7							porque lui gli ha detto [...]
8							
9		lui	è andato	con le stampelle			al comune

le strutture macro sintattiche di *topic* e *comment* sono a volte difficilmente identificabili. Se si trova un'articolazione in *topic* e *comment*, molto spesso si tratta di un *topic* introdotto dall'esaminatore, seguito dalla risposta del candidato che ne costituisce il *comment*, oppure di un *topic* sempre introdotto dall'esaminatore, ripreso dal candidato che vi aggiunge il *comment*.

TABELLA 11
 Griglia *Tv* corpus COCERIT – L2 (parlante B1, Argentina)

La televisione la guarda? No
Poco? Non mi piace

1	L1	la televisione	la guarda?
2	L2		no
3	L1		poco?
4	L2		non mi piace

TABELLA 12
 Griglia *Famiglia* corpus COCERIT – L2 (parlante B1, Russia)

*emb senti e la tua famiglia? È in Italia o nel tuo paese? Mia *familia* tutti sono in Russia a San Pietroburgo.*

1	L1	ehm			
2		senti			
3	L1		la tua famiglia	è	in Italia
4				o	nel tuo paese
5	L2		mia familia		
6			tutti	sono	in Russia
7					a San Pietroburgo

Livello C2

Contrariamente a quanto abbiamo visto per il livello B1, i parlanti non nativi di livello C2 ricorrono a turni molto lunghi. Questo avviene sia nella fase iniziale, come pure nella seconda parte dell'esame in seguito ai vari input volti a stimolare la produzione.

Pur con varianti individuali notevoli, l'intero scambio comporta turni del candidato che superano le due/tre costruzioni (ad esempio nel 66% dei turni prodotti dalla candidata CR e addirittura nell'81% del candidato ESA), mentre le risposte costituite da una singola costruzione, le più frequenti a livello B1, rappresentano circa il 14% delle produzioni dei candidati di livello C2. I parlanti di livello C2 sono in grado di gestire la conversazione, si mostrano autonomi rispetto all'esaminatore malgrado il contesto asimmetrico dell'esame, sanno utilizzare strutture sintattiche complesse e, soprattutto, fanno uso di configurazioni tipiche anche delle produzioni orali dei nativi. Prendiamo ad esempio la TAB. 13. Qui troviamo un'architettura del testo piuttosto elaborata. La candidata, malgrado le frequenti e lunghe interruzioni dell'esaminatore (che non sono state riportate per delle ragioni di spazio e di pertinenza e che sono indicate tramite le parentesi quadre e i punti-

TABELLA 13

Griglia “*noi poveretti*” corpus COCERIT – L2 (parlante C2, Irlanda)

Noi mmm poveretti che cioè se uno se uno è una persona diciamo normale che lavora che cioè non è che ne abbia granché di soldi ehm ci ha sempre delle aspirazioni ci ha sempre dei sogni delle cose che si vuol fare vedere o esperienze che si vuol fare c'è sempre qualcosa in più che si vuol fare...però se uno ha già tutto diventa un po' ehm si annoia...perché se uno va al ristorante ogni sera oh già sì certo cioè non è più una bella uscita quando uno se ne va una volta ogni tanto si ha piacere.

1	noi poveretti												
2	se	uno											
3	se	uno	è	una persona	[diciamo]								
4					normale								
5					che lavora								
6					che								
7					che non è che ne abbia granché di soldi	ci ha	sempre	delle aspirazioni					
8						ci ha	sempre	dei sogni					
9								delle cose	che	si	vuol	fare	
10						o		esperienze				vedere	
11						o		esperienze	che	si	vuol	fare	
12						c'è	sempre	qualcosa in più	che	si	vuol	fare	
13	L1 [...]												
14	L2 però												
15	se	uno	ha già tutto			diventa	un po'						
16						si	annoia						
17	L1 [...]												
18	L2	se	uno	va al ristorante	ogni sera	non	è	più	una bella uscita				
19	L1 o già sì certo												
20	L2 cioè												
21		quando	uno	va	una volta ogni tanto	si	ha	piacere					

ni), mantiene una stessa configurazione, cioè stessa forma di linearizzazione costruita a partire dallo stesso materiale lessicale reiterato (“se uno, se uno”). Il testo presenta caratteristiche tipiche della L1, ripetizioni, ricorso all’asse paradigmatico, ridondanza, costruzioni marcate e poco canoniche. La ripartizione in unità macro-sintattiche è evidente [*se uno ha già tutto*] *topic*, [*si annoia*] *comment*.

Facciamo notare inoltre che architetture testuali di questo tipo mostrano inoltre chiaramente la realtà dell’affermazione di Lewis: «We store much of our mental lexicon in complete, fully-contextualised phrases»²².

Si potrebbe dire che non avendo problemi dal punto di vista dei fenomeni micro-sintattici, i parlanti di livello C2 possono organizzare il testo in strutture macro-sintattiche, a volte anche subordinando la dimensione sintattica alla dimensione macro-sintattica, cioè di gestione dell’informazione. Negli esempi che seguono la struttura dell’informazione è piuttosto evidente: *topic*, *comment* e *appendice* (TAB. 14) o in *topic* e *comment* (TAB. 15).

TABELLA 14
Griglia *Italia* corpus COCERIT – L2 (parlante C2, Spagna)

Che cosa le piace soprattutto dell’Italia? Gli italiani in primo luogo poi mi piace il cibo ovviamente mi piace l’arte come no.

1	L1	che cosa le piace soprattutto dell’Italia?			
2	L1		gli italiani	in primo luogo	
3		poi	mi piace	il cibo	ovviamente
4			mi piace	l’arte	come no

TABELLA 15
Griglia *La gente* corpus COCERIT – L2 (parlante C2, Spagna)

*e come ha trovato gli studenti italiani in spagnolo?
stupendi, stupendi extrasupermotivati, simpatici, non lavoravano – molto a dir la verità però erano veramente piacevolissimi a lezione – sì no la gente adorabile*

1	L1	<i>e come ha trovato gli studenti italiani in spagnolo?</i>			
2	L2				stupendi
3					stupendi
4					extrasuper motivati
5		non	lavorano	molto	
6		a dir la verità			
7		però	erano	veramente	piacevolissimi a lezione
8		sì			
9		no			
10		la gente			adorabile

4.2. Gestione del lessico

Partendo dal presupposto del ruolo fondamentale del lessico nell'apprendimento della lingua, «language consist of grammaticalised lexis, not lexicalised grammar»²³, sulla competenza lessicale e sulla costruzione del significato nel parlato da parte di non nativi, ci limiteremo a due fenomeni che riguardano la gestione della difficoltà di accesso al lessico e il ricorso ad una importante risorsa del linguaggio: la vaghezza. Per entrambi questi fenomeni, per i quali spesso è difficile tracciare una linea di discrimine, vedremo l'importanza del processo di produzione del discorso e vedremo le differenze tra parlanti nativi e non nativi di diversi livelli. Per fare questo partiremo da fenomeni riscontrati in varie L1 (mostreremo ad esempio anche esempi tratti dal francese L1 e dall'italiano L1) per poi passare all'italiano L2.

*Tracce della difficoltà di accesso al lessico
e rilevanza del modo di produzione*

Tra i fenomeni legati al modo di produzione del discorso che mettono in luce l'importanza della competenza lessicale, vi è il processo di ricerca della parola mancante che si realizza, tanto in L1 quanto in L2, nel momento in cui il parlante incontra una qualche difficoltà di accesso al lessico. Questo processo in L1 avviene solitamente con ripetizione di un intero costituente o di parti di esso ripresentando la preposizione o l'articolo, fino ad arrivare alla costituzione di vere e proprie "liste". Prendiamo questo esempio dal francese L1 tratto da Blanche-Benveniste (2008)²⁴. Molto frequentemente vi è l'anticipazione del genere o del numero dell'elemento lessicale mancante (in questo l'articolo *le* anticipa sia il genere che il numero e viene molto prima della "soluzione").

TABELLA 16

Griglia *Buttaf 91,3* corpus GARS francese L1 (citato da Blanche-Benveniste, 2008)

il fallait avoir le le euh ... ah zut j'arrive plus à trouver le nom... le brevet, ...le..., le diplôme là, ...de secouriste: le brevet de secouriste

1	Il fallait avoir	le	euh	
2	ah zut j'arrive plus à trouver le nom			
3		le	brevet	
4		le		
5		le	diplôme	là
6				de secouriste
7		le	brevet	de secouriste

Il lavoro di ricerca del lessico si svolge generalmente sull'asse paradigmatico ed è composto da false partenze, esitazioni e ripetizioni, a volte è anche commentato con chiose del tipo "come si dice?", "non mi viene la parola"²⁵, che attestano una reale difficoltà nel trovare il termine adeguato al contesto percepita dal parlante stesso.

Si tratta comunque di una situazione piuttosto rara in un corpus di parlato L1. Studi sul francese hanno infatti mostrato che raramente i parlanti nativi si bloccano nella ricerca della parola esatta²⁶. Costituiscono un'eccezione le descrizioni più o meno tecniche (cui l'esempio riportato in TAB. 16 appartiene) o il caso di ricerca di idionimi e toponimi che rappresentano gli unici casi in cui si ha effettivamente ricerca lessicale con lista se la parola non viene (come nella TAB. 17).

Altrimenti si accetterebbe di rimanere nel vago o comunque si userebbe una strategia di elusione per evitare la parola che sfugge. La ripetizione che ha come dominio di applicazione il costituente può avere anche altre funzioni che possono essere riunite sotto il titolo di "elaborazione del costituente" e non legate quindi unicamente al processo di produzione, bensì alla costruzione di significato e di cui vedremo altri esempi in italiano L1 nelle TABB. da 24 a 29.

Passando all'italiano L2, a livello B1, il lavoro paradigmatico con ripetizione di materiale lessicale a livello di costituente avviene principalmente per la ricerca di lessico. Abbiamo in altra sede mostrato che il parlante non nativo a questo livello è estremamente legato alla struttura sintagmatica della clausola, come se volesse veder saturati tutti gli argomenti del verbo²⁷. Il parlante si blocca quindi su una posizione sintagmatica e mette in atto una serie di note strategie per farsi capire (ricorso alla L1 o ad altre lingue conosciute, perifrasi ecc.) e completare la costruzione. L'interlocutore nativo è o si sente chiamato in causa e finisce per fornire un'imbecca-

TABELLA 17

Griglia *Vie par 263,3* corpus GARS francese L1 (citato da Blanche-Benveniste, 2008)

Ah, comment est-ce qu'on l'appelle cette accordéoniste ?... Ah, ça me reviendra peut-être, son nom [...] - *Ça y est, mon nom, mon nom, il me revient* : Yvette Corner

1	Ah, comment est-ce qu'on l'appelle cette accordéoniste ?	
2	ah, ça me reviendra peut-être	
3		son nom
4	ça y est	
5		mon nom
6		mon nom
7	il me revient	
8		Yvette Horner

ta risolutiva per la ricerca lessicale. Seguono questo schema le griglie fornite in TAB. 18 e in TAB. 19. Configurazioni dialogiche di questo tipo sono frequentissime a livello B1. In genere il non nativo accetta di buon grado l'imbeccata del nativo.

Nella griglia in TAB. 19 si può riscontrare una configurazione dialogica con un lavoro di co-costruzione del significato tra il parlante nativo e il non nativo, che si estende sull'asse paradigmatico e che utilizza anche l'asse sintagmatico. La parlante giapponese inizia fornendo il verbo *riciclare* che dovrebbe reggere alcuni argomenti della costruzione matrice, tuttavia la particolare pronuncia della parola non permette all'interlocutore il riconoscimento. La parlante fornisce allora gli altri argomenti della costruzione matrice per consentire di identificare il verbo che li regge. Questa strategia di tipo sintagmatico con ricorso al contesto è alquanto efficace e il nativo dimostra di aver capito fornendo l'attesa l'imbeccata, subito adottata dalla nostra parlante.

Il parlante non nativo può anche far appello a codici non verbali. Nell'esempio che segue (TAB. 20), il candidato adotta una strategia iconica per dire che i parlanti della varietà romana di italiano troncano le parole. Inizia dicendo "loro non finiscono la parola", produce poi una sequenza incompleta "la ta-" accompagnata da un fischio e seguita da "metà parola", ad indicare che la metà della parola è stata troncata.

TABELLA 18
Griglia *Mattina* corpus COCERIT – L2 (parlante B1, Irlanda)

*è più efficace eh di alzare presto la mattina eh di perché la *cervesa* la testa cervello la cervello *yeah* è più flessibile fresco fresco {fresco} sì*

1	L2	è	più	eh	eh	la						
		efficace	di	alzare	presto	mattina						
2				eh								
				di								
3												
									la	cervesa		
4									la	testa		
5	L1									cervello		
6	L2								la	cervello		
7										yeah		
8												
											è	più
9	L1											flessibile
10	L2											fresco
11												fresco
12												sì

TABELLA 19
 Griglia *Riciclare* corpus COCERIT – L2 (parlante B1, Giappone)

*ho studiato scienze naturali ho studiato reci- *recicurare* (riciclare) di come la plastica e la carta ah riciclare sì riciclare*

1	L2	ho studiato	scienze naturali	
2	L1	ho studiato	reci-	
3			recicurare	
4			di	
5			come	
6				la plastica e la carta
7	L1		ah	
8			riciclare	
9	L2		sì	
10			riciclare	

TABELLA 20
 Griglia *P. B.* corpus COCERIT – L2 (parlante B1, USA)

e loro non finiscono la parola la tagliano la ta- (fischia) metà parola se tu non conosco conosce la parola la perdi perdi tutto

1	L2	e	loro	non	finiscono	la parola
2					la	
3					ta- (fischia)	
4						metà parola
5	L1				tagliano	

TABELLA 21
 Griglia *Regalo di Natale* corpus COCERIT – L2 (parlante C2, Germania)²⁸

e quando ho visto cos'era sono proprio rimasta scioccata senza parole sono rimasta addirittura era così bello da togliere no anzi

1	L2	e quando ho visto cos'era	sono	proprio	rimasta	scioccata
2						senza parole
3			sono		rimasta	
4	L1	daddirittura	era	così bello	da togliere	
5	L2	no				
6		anzi				

In tutti questi esempi l'intervento del parlante nativo (che è anche l'esaminatore, ma non il valutatore, ricordiamo) è essenziale. La negoziazione e a volte la co-costruzione del significato messa in atto dai due parlanti (nativo e non nativo) sono delle strategie molto comuni per superare le difficoltà legate alla mancanza di lessico. I candidati sembrano proprio sollecitare l'intervento del

TABELLA 22

Griglia “Non lasciarsi la testa” corpus COCERIT – L2 (parlante C2, Germania)

ciòè uno che ha studiato e si è impegnato non ha motivo di preoccuparsi tanto di farsi tanti pensieri cosa succede se non ce la faccio o: quindi avendo studiato avendoci messo un certo impegno può andare anche nell'esame con una certa convinzione

1	ciòè	uno	che	ha studiato				
2			e	si è impegnato		non	ha motivo	di preoccuparsi tanto
3								di farsi tanti pensieri
4								cosa succede se non ce la faccio
5	quindi			avendo studiato				o:
6				avendo messoci	un certo impegno		può	andare anche nell'esame con una certa convinzione

TABELLA 23

Griglia “A caval donato” corpus COCERIT – L2 (parlante C2, Germania)

*un'amica mia mi ha regalato una *malietta* di colore verde che mi dispiace ma proprio è un colore che non mi sta che non è il mio gusto che non che non mi comprerei neanche! E poi la il taglio non era proprio bello bello però tuttavia eh li ho ringraziato tanto e perché ho apprezzato finalmente la il gesto di comprarmi qualcosa di investire tempo ehm per comprarli d'investire anche i soldi per comprarli quindi ho espresso tuttavia il mio ringraziamento per questo regalo*

1		un'amica mia		mi	ha regalato	una *malietta*	di colore verde	che		
2				[mi	dispiace]					
3	ma proprio				è	un colore		che	non	mi sta
4								che	non	è il mio gusto
5								che	non	
6								che	non	mi comprerei neanche
7	e poi	la								
8		il taglio	non		era	proprio	bello			
9							bello			
10	però									
11	tuttavia			li	ho ringraziato	tanto				
	e perché				ho apprezzato	la				
						il gesto				
12										
13							di comprarmi		qualcosa	
14							di investire		tempo	per comprarli
15							di investire	anche	i soldi	per comprarli
16	quindi				ho espresso tuttavia					il mio ringraziamento per questo regalo

nativo che non tarda mai. L'intervento del nativo, anche in questa situazione di esame, dove si potrebbe non attenderselo, è assolutamente naturale, mette il candidato a proprio agio e gli permette di continuare l'interazione. Sottolineiamo che l'obiettivo della prova di produzione orale è di poter valutare la performance dell'esaminato in condizioni più vicine possibili alla situazione di vita reale. Inoltre, come si può vedere dalle griglie appena date, i candidati di livello C2 riescono a modulare il significato del costituente e sottolineano il concetto che stanno esprimendo ricorrendo prevalentemente a strategie diverse: a) la mera ripetizione (“uno che ha studiato”, “avendo studiato”; “investire tempo”, “investire anche soldi”); b) il ricorso a espressioni sinonimiche (“non ha motivo di preoccuparsi”, “di farsi tanti pensieri”; “ho ringraziato tanto”, “ho espresso il mio ringraziamento”); c) l'esplicitazione paradigmatica del rapporto causa-effetto (“che non è di mio gusto”; “che non mi comprerei mai”). Il parlante è inoltre in grado di fare ricorso a circonlocuzioni oppure – come i nativi in contesti non tecnici – si accontenta della vaghezza.

La vaghezza

L'indeterminatezza semantica o vaghezza viene considerata una delle proprietà semiotiche più rilevanti delle lingue naturali. Nel parlato, oltre alla vaghezza intrinseca del significato di alcune parole, cioè la loro proprietà di avere un significato aperto ed estendibile (pensiamo alle «somiglianze di famiglia» di cui parla Wittgenstein), possiamo parlare di vaghezza ogni qual volta nel discorso non sia possibile determinare se un segno si applica o meno ad un referente, quando cioè il parlante non sa e soprattutto non ritiene rilevante trovare il lessema più appropriato a designare un oggetto o un processo. La lingua italiana ha alcuni modi codificati lessicalmente per esprimere vaghezza: possono essere impiegati termini generici, come “cosa”, “affare”, “aggeggio”; esistono specificatori come “sorta di”, “specie di” o forme prefissali come “pseudo” o “quasi”. Vedremo però che esistono altri modi, altrettanto convenzionalizzati, che veicolano lo stesso tipo di rappresentazione semantica.

L'analisi del modo in cui viene espressa la vaghezza nel parlato L1 e ed L2 permetterà ancora una volta di distinguere livelli di competenza, ma al contempo di mostrare come una “forma” del discorso sia associata ad un significato in maniera convenzionale.

Negli enunciati riportati qui di seguito tratti da corpora di L1 (rispettivamente il corpus Bonvino/Arcodip e il LIP²⁹), il costituente evidenziato in grassetto si compone di più sintagmi (per lo più tre) che costituiscono il paradigma di uno stesso elemento grammaticale. Questi sintagmi hanno la stessa funzione sintattica all'interno della costruzione verbale ma soltanto prendendo in considerazione i costituenti in blocco si arriva a cogliere la natura semantica di questo soggetto.

TABELLA 24
Griglia da corpus LIP – L1

lei si era dimenticata eh le mutande reggiseno altre altre cosette di questo genere a casa sua

1	lei	si era dimenticata eh	le mutande	
2			reggiseno	
3			altre	
4			altre cosette di questo genere	a casa sua

TABELLA 25
Griglia da corpus LIP – L1

*sentiva quindi non avevano bisogno di farsi le_ come si chiamano?
le pere le mele non so [ridono]*

1	non	avevano bisogno di farsi	le
2		come si chiamano	
3			le pere
4			le mele
5		non so	

TABELLA 26
Griglia da *Film* corpus Bonvino/ArCoDip – L1³⁰

poi torna sempre 'sta mosca 'sta zanzara quello che sia

1	poi	torna	sempre	'sta mosca
2				'sta zanzara
3				quello che sia

TABELLA 27
Griglia da corpus LIP – L1

gli ha pure detto quanto sei brutta vattene via cose di questo genere

1	gli	ha detto	pure	quanto sei brutta
2				vattene via
3				cose di questo genere

TABELLA 28
Griglia da corpus LIP – L1

sarà mille lire cinquemila lire diecimila lire una cosa del genere la quota di adesione all'associazione

1	sarà	mille lire	
2		cinquemila lire	
3		diecimila lire	
4		una cosa del genere	la quota di adesione all'associazione

La configurazione di discorso è costituita pertanto da una lista generalmente in posizione postverbale composta per lo più da tre elementi: due elementi lessicali che svolgono la stessa funzione sintattica all'interno della costruzione matrice, che sono spesso coiponimi (*le mele/le pere; 'sta mosca/ 'sta zanzara; mutande/ reggiseno*) e che devono intendersi come opzioni possibili di uno stesso paradigma; un terzo elemento che ne attesta l'approssimazione e che può avere la stessa funzione sintattica all'interno della costruzione matrice e quindi far parte dello stesso paradigma (*quello che sia, una cosa del genere*), oppure essere un commento³¹ volto a rendere il tutto più approssimativo (*non so, diciamo*). Abbiamo riscontrato molto frequentemente questo tipo di configurazione di discorso in italiano L1 e riteniamo che sia un modo convenzionalizzato per veicolare la vaghezza.

Se ne osservi un altro esempio nella TAB. 29. In questo caso abbiamo l'approssimazione di "genere" ottenuta con lo stesso procedimento.

Facciamo notare che mentre la vaghezza espressa dagli esempi da TAB. 24 a 28 potrebbe essere resa altrimenti in italiano anche a livello lessicale ricorrendo all'iperonimo (frutto, insetto, biancheria intima) e con la anteposizione di espressioni quali *sorta di, sto cavolo di, circa* ecc. Invece, a nostra conoscenza, l'italiano standard non ha altri mezzi, se non quello riportato in TAB. 29, per esprimere il genere epiceno, se non il ricorso alla parola *creatura* di alcune varietà regionali di italiano.

Abbiamo sin qui inteso dimostrare l'esistenza di una configurazione di discorso specifica che ha la funzione grammaticale di esprimere la vaghezza e che può essere utilizzata come unico mezzo con questa funzione oppure in co-occorrenza con altri strumenti linguistici.

Gli esempi nelle TABB. 30, 31 e 32 possono a vario titolo essere considerati altrettante configurazioni con funzione di espressione della vaghezza prodotte da parlanti non nativi. Nella griglia in TAB. 30, la parlante è bulgara ed ha un livello comunicativo probabilmente superiore al livello B1 accertato dalla prova di certificazione. L'asse paradigmatico è utilizzato nella direzione della modulazione del referente. Nel primo caso, evidenziato in grassetto, *curieri/pulmini piccoli*, la linea di discriminazione tra espressione della vaghezza e strategia di elaborazione lessicale non è facilmente tracciabile. Nel secondo caso evidenziato in grassetto *per diciamo dieci dodici persone*, si trat-

TABELLA 29
Griglia da *Chiacchiere* corpus Bonvino/ArCoDip- L1

è una che ci ha un figlio una figlia non so

1	è	una	che	ci ha	una figlia
2					un figlio
3	non so				

ta di un caso di quantificatori numerali approssimanti assolutamente convenzionalizzato e perfettamente accettabile per la lingua *target*, per giunta preceduto dall'espressione *diciamo*.

Nella TAB. 31, l'idea di vaghezza – ottenuta con lo stesso procedimento – è data dal paradigma *figli/famiglia*. Anche in questo caso non si può affermare con certezza se si tratti di una sorta di ampliamento *vs.* correzione del referente o se la lista è da intendersi in blocco e mira ad ottenere un effetto di approssimazione.

Nella griglia riportata nella TAB. 32, l'espressione della vaghezza è ottenuta con un procedimento molto simile a quello evidenziato per la L1: due elementi lessicali semanticamente affini ma non equivalenti, posti in posizione postverbale³² e seguiti da un commento (*diciamo*) volto a non sottoscrivere la totale corrispondenza fra i termini e l'oggetto che si va a definire.

Gli esempi che abbiamo riportato nelle griglie delle TABB. 30, 31 e 32 appartengono a diversi sottotipi di vaghezza o approssimazione, che non possiamo illustrare in questa sede. Quel che ci preme qui di sottolineare è che il parlante spagnolo di livello avanzato, e come lui i numerosi parlanti esami-

TABELLA 30
Griglia *Pulmini* corpus COCERIT – L2 (parlante B1, Bulgaria)

*questi ragazzi prendono – *currieri* (corriere) pulmini piccoli per diciamo dieci dodici persone no più*

1	questi ragazzi	prendono	*currieri*			
2			pulmini piccoli	per		
3	diciamo					
4					dieci	
5					dodici	persone
6					no	più

TABELLA 31
Griglia *Parenti* corpus COCERIT – L2 (parlante B1, Bulgaria)

*...portano qualche roba quello che mandano i nostri *filii* i nostri *famili**

1	portano	qualche roba			
2		quello	che	mandano	i nostri filii
3					i nostri famili

TABELLA 32
Griglia *Ditta* corpus COCERIT – L2 (parlante C2, Spagna)

4	la	ditta	
5	la	compagnia	
6	diciamo		

nati nel corpus di livello C2, è in grado di utilizzare la vaghezza, nei modi e con gli scopi dei parlanti nativi. A partire dai dati in nostro possesso, quantitativamente importanti, ma spesso troppo difformi per una statistica affidabile, si può invece ipotizzare che i parlanti di livello B1, avendo capacità di gestione del lessico limitata, non sappiano utilizzare con consapevolezza gli strumenti per l'approssimazione, se non in alcuni casi altamente codificati.

5 Conclusioni

Come abbiamo cercato di mostrare, la lingua parlata, anche da parlanti non nativi, ha caratteristiche tipiche per quanto concerne i modi di produzione (ripetizioni, ridondanza, frammentarietà, impiego di paradigmi), che necessitano di essere analizzate e valutate con strumenti adeguati e non basati unicamente sulla grammatica normativa dello scritto. Questi strumenti non sono stati ancora elaborati sufficientemente, e la loro individuazione non è ancora molto diffusa tra gli studiosi. Noi speriamo che uno studio come il nostro (da intendersi come un *work in progress*) possa fornire un contributo in questa direzione nell'ambito degli studi dell'italiano parlato come L2.

Quanto ai risultati, l'analisi dei dati di italiano L2 ci permette di giungere ad alcune – a nostro avviso interessanti – conclusioni.

Considerando il livello B1, abbiamo notato il controllo del livello micro-sintattico, limitato per lo più all'ambito della costruzione verbale: si parte dal verbo e si cerca di saturare in successione “ortodossa” tutte le posizioni argomentali che esso comporta; si riscontrano inoltre una serie di difficoltà di accesso al lessico legate al mancato ricorso alla riformulazione e all'incapacità di usare strategie di elusione. Il ruolo del parlante nativo a questo livello è fondamentale: può suggerire la parola mancante, ma anche contribuire alla costruzione di un testo con un'architettura più ricca.

Al contrario, a livello C2, appaiono strutture sintattiche complesse e turni di parola molto più sviluppati. Queste strutture sono gestite perfettamente dall'apprendente che non ha alcun bisogno dell'aiuto dell'interlocutore nativo. L'organizzazione macrosintattica dei testi è evidente. I dati analizzati ci permettono di affermare che i parlanti non nativi di questo livello ricorrono a strutture non canoniche più spesso rispetto ai parlanti non nativi di livello B1. Dal punto di vista lessicale i parlanti di livello C2 non mostrano particolari problemi e, proprio come i parlanti nativi, ricorrono con naturalezza a circonlocuzioni o a espressioni vaghe per veicolare significato.

La nostra analisi bidimensionale della lingua, che tiene conto sia del processo di produzione che del significato delle costruzioni di discorso, ha delle implicazioni per lo studio della lingua parlata in generale, ma anche per la linguistica acquisizionale, in quanto rimette in discussione la nozione stessa

di competenza di uso della L2. Riteniamo infatti che analizzando produzioni orali in maniera monodimensionale, concentrandosi sulla dimensione sintattica e interpretando qualsiasi deviazione dalla stringa sintagmatica come un fenomeno di natura interlinguistica (ovvero un “errore”), possa portare a delle conclusioni erranee circa la struttura dell'enunciato e i modi di produzione della parola e quindi all'attribuzione di giudizi valutativi impropri.

Il problema principale è che, anche se stiamo procedendo ad un lavoro di censimento a partire dal corpus di alcuni dei fenomeni a nostro avviso più salienti ed interessanti, sia per l'italiano come L1 sia per l'italiano come L2, non disponiamo allo stato attuale di sufficienti individuazioni funzionali in questa ottica, in quanto dal punto di vista della descrizione di fenomeni del parlato le teorie di riferimento non sempre convergono, e inoltre nell'elaborazione di sofisticati strumenti per l'annotazione e soprattutto per l'interrogazione di corpora di lingue non native si è ancora in una fase che possiamo definire sperimentale.

Note

1. Si vedano, tra gli altri, i lavori di R. Sornicola, *Sul parlato*, il Mulino, Bologna 1981; M. Voghera, *Sintassi e intonazione dell'italiano parlato*, il Mulino, Bologna 1992; E. Cresti, *Corpus di italiano parlato*, Accademia della Crusca, Firenze 2000, come pure il sito www.parlaritaliano.it, e i convegni organizzati dal gruppo “La comunicazione parlata”.

2. Si vedano, ad esempio, A. Giacalone Ramat (a cura di), *L'apprendimento spontaneo di una seconda lingua*, il Mulino, Bologna 1986; A. Giacalone Ramat (a cura di), *L'italiano tra le altre lingue*, il Mulino, Bologna 1988; A. Giacalone Ramat (a cura di), *Verso l'italiano*, Carocci, Roma 2004; G. Pallotti, *La seconda lingua*, Bompiani, Milano 1998; G. Bernini, *Come si imparano le parole. Osservazioni sull'acquisizione del lessico in L2*, in “ITALS. Didattica e linguistica dell'italiano come lingua straniera”, I, 2003, pp. 23-47; C. Bettoni, *Imparare un'altra lingua*, Laterza, Roma-Bari 2001; C. Bettoni, B. Di Biase, *Sviluppo obbligato e progresso morfosintattico: un caso di Processabilità in italiano L2*, in “ITALS. Didattica e linguistica dell'italiano come lingua straniera”, III, 2005, pp. 27-48; C. Bettoni, B. Di Biase, *Lessico verbale e questioni di Processabilità in italiano L2*, in M. Barni, D. Troncarelli, C. Bagna (a cura di), *Lessico e apprendimenti. Il ruolo del lessico nella linguistica educativa*, Franco Angeli, Milano 2008; M. Vedovelli, *L'italiano degli stranieri. Storia, attualità e prospettive*, Carocci, Roma 2002.

3. Cfr. E. Bonvino, *Le sujet postverbal. Une étude sur l'italien parlé*, Ophrys, Paris 2005.

4. Questa affermazione è vera anche quando si tratta di una madrelingua: i parlanti nativi gestiscono con diversi gradi di competenza le varie situazioni d'uso. Ad esempio, non tutti i parlanti nativi, anche se acculturati, sanno scrivere bene un testo.

5. La prova di livello B2 è anch'essa articolata in 4 prove: 2 ricettive (ascolto e lettura), e 2 produttive (composizione e comunicazione faccia a faccia). In questo caso, la competenza è considerata come “semi-modulare”: il candidato deve sostenere nella stessa sessione tutte e due prove ricettive; quelle produttive possono essere sostenute in tempi diversi. Per ottenere il certificato, però, il candidato deve superare almeno le prime due prove.

6. Council of Europe, *Common European Framework of Reference*, www.coe.it, 2001; trad. it. *Quadro comune europeo di riferimento*, La Nuova Italia-RCS Scuola, Oxford-Milano 2002, p. 93.

7. In una situazione di esame, la presenza di un microfono può creare maggior stress nel candidato, e pertanto inibirlo nell'uso della L2.

8. Council of Europe, *Common European Framework of Reference*, trad. it. cit., pp. 60-1.

9. Cfr. F. Albano Leoni, *Studiare l'italiano parlato: strumenti, metodi, problemi*, in A. L. Lepschy, A. R. Tamponi (a cura di), *Prospettive sull'italiano come lingua straniera*, Guerra Edizioni, Perugia 2005, pp. 83-93.
10. È possibile che qualche candidato, specie se vive in Italia, abbia un livello superiore al Bi certificato. In questo caso il certificato riporterà il giudizio: Bi-ottimo.
11. Cfr. C. Blanche-Benveniste, C. Jeanjean, *Le français parlé transcription et édition*, Didier, Paris 1987.
12. Riprendiamo qui rapidamente solo gli assunti essenziali dell'analisi in griglie, per un'approfondita esposizione dei principi che ne regolano la elaborazione rimandiamo alla bibliografia sull'argomento: Blanche-Benveniste, Jeanjean, *Le français parlé transcription et édition*, cit.; M. Bilger, *Contribution à l'analyse en grilles*, in "Recherches sur le français parlé", 4, 1982, pp. 195-215; M. Bilger et al., *La transcription de la langue parlée: quelques observations*, in "Recherches sur le français parlé", 14, 1997, pp. 57-86; C. Blanche-Benveniste, *Un modèle d'analyse syntaxique "en grilles" pour les productions orales*, in "Anuario de Psicología", 47, 1990, pp. 11-28; C. Blanche-Benveniste, *Approches de la langue parlée en français*, Ophrys, Paris 1997; C. Blanche-Benveniste et al., *Le français parlé. Études grammaticales*, Éditions du Centre national de la recherche scientifique, Paris 1990.
13. Quando si tratta di dati relativi a lingue non native, le produzioni idiosincratice "non target", vale a dire le produzioni non attese, creano una serie di problemi soprattutto in relazione alla trascrizione (ortografica o fonetica), all'interpretazione (ricostruzione in base al contesto), al loro collegamento alla forma attesa, al loro riconoscimento al momento dell'interrogazione del corpus.
14. Si vedano a questo proposito le argomentazioni che forniscono Francesca Masini e Paola Pietrandrea, *Magari* (in prep.), sulla delimitazione delle unità da analizzare e sulle loro caratteristiche ricorrenti, nell'analisi semantica dell'avverbiale *magari*.
15. Per la definizione di configurazione si veda in particolare C. Blanche-Benveniste, *Des grilles pour le français parlé*, in "Recherches sur le français parlé", 2, 1979, pp. 163-204.
16. Bonvino, *Le sujet postverbal*, cit., p. 79.
17. Cfr. Ch. J. Fillmore, P. Kaye, M. C. O'Connor, *Regularity and Idiomaticity in Grammatical Constructions: The Case of Let Alone*, in "Language", 64, 3, 1988, pp. 501-38; A. Goldberg, *Constructions. A Construction Grammar Approach to Argument Structure*, The University of Chicago Press, Chicago 1995; A. Goldberg, *Constructions at Work*, Oxford University Press, Oxford 2006.
18. S. Ambroso, E. Bonvino, *Configurazioni di discorso. Un'unità di analisi del parlato L2*, in Atti del convegno *Corpora di italiano L2: tecnologie, metodi, spunti teorici*, Università di Pavia, (in stampa).
19. J.-O. Östman, *Construction Discourse, a Prolegomenon*, in J.-O. Östman, M. Fried, *Construction Grammars. Cognitive Grounding and Theoretical Extensions*, Benjamins, Amsterdam-Philadelphia 2005.
20. Per il francese si vedano i lavori di Claire Blanche-Benveniste e del gruppo GARS, citati nelle note 11, 12 e 15; per l'italiano, Bonvino, *Le sujet postverbal*, cit.; P. Pietrandrea, *The Grammatical Nature of some Epistemic-evidential Adverbs and Adverbial Constructions in Spoken Italian*, in "Italian Journal of Linguistics", 2, 2007; P. Pietrandrea, *Certamente e sicuramente. Encoding Dynamic and Discursive Aspects of Commitment in Italian*, in Ph. de Brabanter (ed.), *Commitment*, Benjamins, Amsterdam-Philadelphia, in stampa.
21. Cfr. M. Pienemann (ed.), *Cross-linguistic Aspects of Processability Theory*, Benjamins, Amsterdam-Philadelphia 2005, pp. 199-251.
22. M. Lewis, *Implementing the Lexical Approach: Putting a Theory into Practice*, Language Teaching Publications, Hove 1997, p. 9.
23. M. Lewis, *The Lexical Approach*, Language Teaching Publications, Hove, 1993, p. VI.
24. C. Blanche-Benveniste, *Aspetti lessicali del confronto tra lingue romanze. Esiste un lessico europeo?*, in Barni, Troncarelli, Bagna, *Lessico e apprendimenti. Il ruolo del lessico nella linguistica educativa*, cit., pp. 47-66.

25. In questo caso: *Ab zut j'arrive plus à trouver le nom* – Ah accidenti non riesco più a trovare il nome.

26. C. Blanche-Benveniste, *Aspetti lessicali del confronto tra lingue romanze*, cit.

27. S. Ambroso, E. Bonvino, R. Mastrofini, *Différents niveaux de compétence dans l'italien parlé L2. Phénomènes et hypothèses pour une évaluation appropriée*, in *Atti del Convegno Internazionale "Recherches en acquisition et en didactique des langues étrangères"*, <http://www.grouperca.org/h/colloque2006/actes.html>, 2007.

28. Gli esempi nelle TABB. 21, 22, 30 e 31 sono tratti da E. Fiorenza, *Le strutture del discorso in italiano parlato come lingua seconda. Analisi di un corpus*, tesi di laurea, Università Roma Tre, 2007.

29. T. De Mauro, F. Mancini, M. Vedovelli, M. Voghera, *Lessico di frequenza dell'italiano parlato (LIP)*, Etas Libri, Milano 1993.

30. Per un'analisi sintattico-semantica più dettagliata di questo esempio si veda Bonvino, *Le sujet postverbal*, cit.; Ambroso, Bonvino, *Configurazioni di discorso*, cit.

31. In questo secondo caso, esemplificato dalla griglia in TAB. 29, il commento interrompe la lista ma nello stesso tempo ha la lista nel suo *scope* e ne modifica il significato.

32. Questa griglia costituisce un dettaglio della griglia riportata in TAB. 14.