

AIO



Stefania Ferrari

# **Valutare le competenze orali in italiano L2**

Variazione longitudinale e situazionale  
in apprendenti a livello avanzato





Aracne editrice

[www.aracneeditrice.it](http://www.aracneeditrice.it)  
[info@aracneeditrice.it](mailto:info@aracneeditrice.it)

Copyright © MMXIX  
Gioacchino Onorati editore S.r.l. — unipersonale

[www.gioacchinoonoratieditore.it](http://www.gioacchinoonoratieditore.it)  
[info@gioacchinoonoratieditore.it](mailto:info@gioacchinoonoratieditore.it)

via Vittorio Veneto, 20  
00020 Canterano (RM)  
(06) 45551463

ISBN 978-88-255-2956-2

*I diritti di traduzione, di memorizzazione elettronica,  
di riproduzione e di adattamento anche parziale,  
con qualsiasi mezzo, sono riservati per tutti i Paesi.*

*Non sono assolutamente consentite le fotocopie  
senza il permesso scritto dell'Editore.*

I edizione: dicembre 2019

- 7 *Introduzione*
- II **Capitolo I**  
*Variazione, acquisizione e testing*
- 1.1. Le competenze linguistiche, 15 — 1.1.1. *La competenza*, 16 — 1.1.2. *La proficiency, l'abilità comunicativa e l'abilità d'uso*, 20 — 1.2. L'approccio acquisizionale, 33 — 1.2.1. *La variazione*, 33 — 1.2.2. *La misurazione della variazione*, 63 — 1.3. L'approccio del testing, 93 — 1.3.1. *La variazione*, 94 — 1.3.2. *La misurazione della variazione*, 99 — 1.4. L'interfaccia acquisizione-testing, 114 — 1.4.1. *L'acquisizione e il testing*, 115 — 1.4.2. *I costrutti, la misurazione e le scale di proficiency*, 119 — 1.5. I task, 122 — 1.5.1. *La definizione di task*, 123 — 1.5.2. *I task per insegnamento, SLA e testing*, 128.
- 133 **Capitolo II**  
*Metodologia di ricerca*
- 2.1. Gli obiettivi generali della ricerca, 133 — 2.2. I soggetti, 134 — 2.3. Le rilevazioni, 139 — 2.4. I task, 140 — 2.5. I dati, 142 — 2.6. L'unità e le misure dell'analisi, 143 — 2.6.1. *Le misure CAF*, 143 — 2.6.2. *Le valutazioni QCER*, 170.
- 177 **Capitolo III**  
*Risultati*
- 3.1. La complessità sintattica CAF, 177 — 3.1.1. *La variazione longitudinale*, 178 — 3.1.2. *La variazione situazionale*, 180 — 3.1.3. *La variazione individuale*, 185 — 3.1.4. *Riassunto*, 237 — 3.2. L'accuratezza CAF, 239 — 3.2.1. *La variazione longitudinale*, 239 — 3.2.2. *La variazione situazionale*, 240 — 3.2.3. *La variazione individuale*, 242

— 3.2.4. *Riassunto*, 259 — 3.3. *La fluenza CAF*, 261 — 3.3.1. *La variazione longitudinale*, 262 — 3.3.2. *La variazione situazionale*, 263 — 3.3.3. *La variazione individuale*, 265 — 3.3.4. *Riassunto*, 267 — 3.4. *Complessità, accuratezza e fluenza CAF a confronto*, 270 — 3.4.1. *La variazione longitudinale*, 271 — 3.4.2. *La variazione situazionale*, 272 — 3.4.3. *La variazione individuale*, 275 — 3.4.4. *Riassunto*, 281 — 3.5. *Le valutazioni QCER*, 282 — 3.5.1. *La variazione longitudinale*, 283 — 3.5.2. *La variazione situazionale*, 285 — 3.5.3. *Riassunto*, 287 — 3.6. *Misure CAF e valutazioni QCER a confronto*, 288 — 3.6.1. *La variazione longitudinale*, 289 — 3.6.2. *La variazione situazionale*, 290 — 3.6.3. *Riassunto*, 292.

293 *Conclusioni*

305 *Bibliografia*

# Introduzione

STEFANIA FERRARI\*

La valutazione è un tratto costitutivo dell'insegnamento, come lo sono la programmazione e la realizzazione del piano di lavoro attraverso le attività in classe. Sembra naturale: se si vuole definire la programmazione è necessario avere informazioni sullo stato delle competenze di partenza dell'allievo; per sapere come si sta procedendo con l'attività didattica, e eventualmente apportare dei miglioramenti, è importante avere strumenti di controllo e di monitoraggio in itinere; così come servono dispositivi per conoscere gli esiti alla fine del percorso, quando l'istituzione scolastica richiede di assegnare un valore all'apprendimento di ogni studente. La mia esperienza di insegnante e formatrice mi ha permesso di constatare che non esiste tema più delicato della valutazione nel mondo della scuola: ogni insegnante sembra interpretare diversamente il compito di valutare gli apprendimenti linguistici, e uno stesso studente è facilmente classificato come competente da un collega e di scarso rendimento da un altro. Testimoni di ciò le tante prove e griglie di valutazione incontrate in questi anni o le curiose affermazioni di qualche docente, tra cui la mia preferita è "il mio cinque è un otto". Relativamente all'insegna-

\* Ricercatrice presso il Dipartimento di Studi Umanistici dell'Università del Piemonte Orientale.

mento delle lingue, definire criteri e standard per descrivere i livelli di partenza e registrare i progressi degli apprendenti richiede idee chiare su almeno due aspetti: quali caratteristiche hanno i diversi compiti comunicativi e quali tappe percorre lo sviluppo della competenza linguistica nel tempo. Da qui l'importanza di una stretta collaborazione tra chi si occupa di ricerca acquisizionale e di testing e chi opera direttamente nella scuola.

L'ingresso di studenti immigrati e i nuovi orientamenti curricolari, più orientati alla personalizzazione dei percorsi di apprendimento, nonché i risultati nazionali delle valutazioni del programma OCSE PISA (OECD, 2018) o delle prove INVALSI (Martini, 2017), hanno riacceso negli ultimi tempi la discussione all'interno della scuola sul tema dell'insegnamento e della valutazione. In ambito linguistico, sia per le lingue straniere sia per l'italiano L2, il Quadro Comune Europeo di Riferimento (QCER) e il Portfolio Europeo delle Lingue (PEL) godono di una certa diffusione. Da qui la motivazione per lo studio intrapreso con questo lavoro, il cui obiettivo è appunto contribuire alla descrizione della natura e dello sviluppo delle competenze linguistiche in L2. Nonostante il vivace interesse per l'italiano L2, in letteratura infatti sono pochi gli studi che indagano la variazione longitudinale in apprendenti a livello avanzato inseriti nella scuola secondaria, e a mia conoscenza ancora nessuna ricerca verte sulla variazione situazionale legata a diverse attività comunicative. In questo lavoro ho cercato quindi di affrontare questi temi coniugando le due prospettive dell'acquisizione e del testing, tenendo sempre presenti le implicazioni pratiche sia per la didattica sia per la valutazione delle competenze linguistiche degli studenti immigrati inseriti nella scuola italiana. Lo studio presentato in questo volume si propone da un lato di investigare la variazione situazionale e longitudinale dell'interlingua di apprendenti avanzati im-



pegnati nello svolgimento di task comunicativi monologici e interattivi, e dall'altro verificare sperimentalmente con una prospettiva legata al testing le relazioni che intercorrono tra due tipi di misurazione: quella analitica CAF che rileva la complessità, l'accuratezza e la fluenza della produzione orale e quella soggettiva basata sui descrittori del QCER, maggiormente legata all'ambito del testing.

Il lavoro è diviso in 4 capitoli. Il capitolo 1 fornisce l'ancoraggio teorico dello studio discutendone la letteratura di base. Si inizia definendo la competenza e la proficiency, poi si presenta il tema della variazione e della sua misurazione secondo i due approcci dell'acquisizione e del testing, discutendo anche i punti di contatto tra i due ambiti. Si conclude presentando il task, strumento comune alle due prospettive. La prima parte del capitolo 2 è dedicata alla presentazione dettagliata degli obiettivi della ricerca, delle modalità di raccolta dati, delle informazioni relative ai partecipanti; la seconda parte alla presentazione delle misure impiegate. Il capitolo 3 riporta l'analisi, presentando prima i risultati della misurazione CAF, poi quelli delle valutazioni QCER, e infine discutendo le possibili interazioni tra i due strumenti. Nel capitolo conclusivo vengono sintetizzati i risultati e presentate alcune implicazioni per la ricerca acquisizionale, per il testing e per la didattica.

Ringrazio l'Università degli Studi di Verona, che mi ha dato l'opportunità di svolgere questa ricerca. Un doveroso ringraziamento va inoltre a Camilla Bettoni, Gabriele Pallotti e Ineke Vedder che da tempo motivano e sostengono il mio interesse per la ricerca, a Beatrice Iori che costantemente stimola la mia passione per l'insegnamento e la formazione, a Elena Nuzzo e Giovanna Masiero che hanno condiviso le fatiche del fieldwork (e non solo). Questo studio non avrebbe potuto essere realizzato senza la preziosa disponibilità delle sei stu-

dentesse intervistate, l'efficiente collaborazione dell'I.P.S.I.A. Malpighi di San Giovanni in Persiceto (Bo) e del Multicentro Educativo Sergio Neri di Modena (M.E.MO). Il mio più sincero ringraziamento va anche alle scuole, agli insegnanti e soprattutto agli studenti con cui ho lavorato in questi anni, le loro domande e riflessioni hanno stimolato in me la passione e la curiosità necessarie per realizzare questo lavoro. Infine un ringraziamento particolare va ai miei genitori, a mio marito e ai miei figli, che con la loro amorevole e costante pazienza, rendono possibile il mio lavoro.

## Variazione, acquisizione e testing

La linguistica acquisizionale e il testing si occupano entrambi delle lingue seconde, e dunque appartengono al medesimo ambito di studio, ma costituiscono due campi diversi. Le ricerche acquisizionali si concentrano principalmente sui processi dell'apprendimento e sui fattori che lo favoriscono o lo ostacolano, con l'intento di spiegarli. Le ricerche sul testing invece si concentrano sugli elementi e sulle strategie che compongono l'abilità linguistica, elaborando teorie relative alla performance in grado di spiegarne la variazione durante i test e di controllarne la corrispondenza con l'uso della lingua in contesti reali. La differenza di sguardo che i due campi di ricerca assumono verso la L2 si riflette anche negli approcci metodologici che impiegano. Mentre le ricerche acquisizionali usano tecniche etnografiche o protocolli sperimentali per l'analisi linguistica delle produzioni interlinguistiche, le ricerche sul testing si affidano principalmente a correlazioni statistiche elaborate sui risultati ottenuti nei test.

Queste differenze di sguardo e di metodologia hanno finora tenuto separati i due campi. Eppure già da tempo alcuni lavori (cfr. Bachman, 1988; Bachman, Cohen, 1998a; Hulstijn, Schoonen, 2006; Bartning *et al.*, 2010; Alderson, Huhta, 2011; Alderson *et al.*, 2017 e § 1.4.) riconoscono le potenzialità di uno scambio proficuo tra la linguistica acquisizionale e le ricerche sul testing linguistico. Chi si occupa di testing infat-

ti può ritrovare nella letteratura acquisizionale informazioni importanti per una migliore definizione degli elementi linguistici testabili nei diversi livelli di proficiency. D'altro canto, chi si occupa di ricerca acquisizionale può imparare dalla letteratura relativa al testing l'importanza di una più rigorosa definizione operativa dei costrutti teorici e della verifica di validità e affidabilità degli strumenti di elicitazione e misurazione. Un punto di contatto tra ricerche acquisizionali e testing è sicuramente rappresentato dal tema della variabilità delle produzioni. In questa prospettiva si inserisce il presente lavoro, che intende analizzare la variabilità della L2 integrando appunto l'approccio del testing con quello acquisizionale.

In letteratura è ormai assodato che le produzioni degli apprendenti, così come quelle dei parlanti nativi, siano variabili (cfr. § 1.2.1.). Gli studi acquisizionali distinguono almeno due tipi di variabilità dell'interlingua, dando per scontata la ben nota variabilità individuale. Un primo tipo di variabilità, definita da Corder (1981) come *vertical continuum*, è intrinseca all'interlingua, che come sistema autonomo e transitorio è di per sé variabile. In questo senso l'attenzione di chi la studia viene posta verso l'alternanza tra forme native e forme non native, o tra due o più forme non native per esprimere una determinata nozione che nella lingua target viene espressa con una sola forma. Siccome la variazione riscontrata nell'interlingua è segno che l'apprendente sta acquisendo la forma invariabile della lingua target, l'apprendimento si esprimerà con una riduzione di tale variazione. Un secondo tipo di variabilità, chiamata da Corder (1981) *horizontal continuum*, riguarda gli aspetti della lingua target in cui i parlanti nativi presentano variazione sociolinguistica e situazionale. Il progresso in questo caso è visto come acquisizione dell'intera gamma delle varianti native, usate con la medesima frequenza dei parlanti nativi, nel medesimo rispetto dei fattori linguistici ed extra-linguistici

che le vincolano. Pertanto l'apprendimento si esprimerà con l'aumento di tale variazione, non con la riduzione: se la variabilità orizzontale è parte della competenza del parlante nativo, essa deve essere acquisita dall'apprendente.

Il tema della variabilità è centrale anche nel campo della valutazione. Se le produzioni di un apprendente variano è essenziale determinare quale tipo di performance è rappresentativa delle abilità che si vogliono testare. Come già spiegava Davies (1990: 137) «the fundamental argument/debit in language testing over the last 25 years has been basically about the meaning/realization of language behaviour». Poiché possiamo determinare quello che un apprendente sa esclusivamente da quello che fa (cfr. Ellis, 2001; Chalhoub-Deville, 2001), qualsiasi forma di testing è *performance testing*. Il compito dei ricercatori è dunque quello di stabilire quale tipo di performance testare.

Se allora la variabilità dell'interlingua è centrale in entrambi i campi della linguistica acquisizionale e del testing, diventa importante rilevarla con una metodologia adeguata. Oggi il task sembra essere il mezzo più opportuno (cfr. § 1.5.). Per task, riprendendo la definizione di Skehan (1995), si intende un'attività linguistica in cui l'attenzione del parlante è primariamente volta al significato piuttosto che alla forma di quello che dice, poiché l'esito del compito viene valutato sul risultato raggiunto. In questo modo il tutto — anche se compiuto per scopi di ricerca, di testing o di didattica — ha una qualche relazione con il mondo reale. A partire dagli anni '70 il task è stato proposto in didattica come valida alternativa per l'organizzazione del syllabo. I task pedagogici hanno infatti il vantaggio di fornire agli apprendenti occasioni sia per praticare forme già conosciute sia per apprenderne di nuove, favorendo un uso naturale della lingua e offrendo contemporaneamente specifici momenti di *focus on form*. Per costruire un syllabo efficace è necessario però avere a disposizione un quadro di riferimento che permetta di

concentualizzare la performance richiesta dai vari task, di selezionarli e di graduarli sulla base delle loro caratteristiche, e infine di implementarli per massimizzare le occasioni di *riflessione sulla lingua*. Di conseguenza in ambito acquisizionale la ricerca si è concentrata su una serie di caratteristiche dei vari task, tra cui l'analisi della loro difficoltà in relazione alle competenze e alle caratteristiche degli apprendenti (cfr. I.2.I.2.). Queste ricerche osservano in genere la variazione della lingua prodotta secondo tre dimensioni (cfr. Skehan, 1996a: 22; 1998; 2008):

- a) la complessità sintattica, ossia la lunghezza e il grado di subordinazione degli enunciati;
- b) l'accuratezza, ossia il grado di conformità alla norma della lingua target;
- c) la fluenza, ossia la capacità di evitare eccessive pause o esitazioni.

Secondo Skehan (1998; 2011; 2014; 2015; 2018) e Robinson (2001a; 2011a–b; 2015; 2017), queste tre dimensioni della performance sono indipendenti l'una dall'altra e vengono influenzate dalle caratteristiche del task e dal conseguente tipo di processing che attivano nell'apprendente. Queste tre dimensioni sono centrali anche in questo studio, e ci si riferirà ad esse usando l'acronimo CAF proposto da Housen, Kuiken e Vedder (per es. *Applied Linguistics* 2009, Special Issue; Housen, Kuiken, Vedder, 2012).

Nell'ambito del testing si dibatte l'applicabilità del task quale strumento per aumentare la validità ecologica delle prove e di conseguenza la loro generalizzabilità a contesti di vita reale (cfr. *Language Testing*, 2002; Purpura, 2016; Morrow, 2018). Questo è coerente sia con l'approccio comunicativo alla didattica della L2 sia con l'approccio alla valutazione basato sulla competenza d'azione, così come descritta nel Quadro Comu-

ne Europeo di Riferimento per le lingue, QCER da ora in poi (Consiglio d'Europa, 2001; 2018). Anche in questo caso una maggiore comprensione della variabilità nella performance elicitata tramite i diversi task e del rapporto che intercorre tra abilità e performance in un task pare essenziale, in quanto fornisce una base per selezione i task secondo le difficoltà che pongono ai candidati (McNamara, 1996; Brown *et al.*, 2002; Fulcher, Marquéz Reiter, 2003; Davidson, Fulcher, 2007). In questo ambito la valutazione della performance viene solitamente realizzata a partire da scale di descrittori della proficiency.

Il presente capitolo è dedicato all'approfondimento dei concetti teorici, dei costrutti analitici e delle misure applicative che forniscono la base per questo studio. Pertanto verranno dapprima introdotte le nozioni di competenza e proficiency, e discusse le diverse interpretazioni date a questi termini in letteratura (§ 1.1.). Quindi si affronterà il tema della variabilità delle produzioni interlinguistiche: se la performance è variabile, sarà necessario stabilire cosa vale la pena osservare e con quali strumenti. A questo proposito si considereranno due prospettive: quella acquisizionale (§ 1.2.) e quella relativa al testing linguistico (§ 1.3.), riportando la discussione teorica sul ruolo che la variazione assume in ciascuna e le modalità di valutazione tipiche dei due ambiti, le misure analitiche CAF e le valutazioni olistiche QCER. Si discuteranno poi i punti di contatto tra gli studi acquisizionali e la ricerca sul testing evidenziando i vantaggi di una maggiore interazione tra i due ambiti (§ 1.4.). Infine si presenterà il task (§ 1.5.), quale strumento comune alle due prospettive.

### 1.1. Le competenze linguistiche

Linguistica acquisizionale e ricerca sul testing condividono l'attenzione per le competenze linguistiche: mentre la prima

però si occupa della descrizione dell'interlingua dell'apprendente e del suo sviluppo, la seconda si concentra sulla definizione dell'abilità linguistica. Dal punto di vista acquisizionale è importante capire i meccanismi e i fattori che entrano in gioco nello sviluppo delle competenze linguistiche, mentre dal punto di vista del testing è fondamentale avere una chiara definizione del costrutto dell'abilità linguistica che aiuti a spiegare cos'è la proficiency e come questa si relazioni ai vari test.

Definire in modo teoricamente valido e operativamente utile l'abilità linguistica si è rivelato però un compito estremamente complesso. Gli stessi ricercatori faticano a trovare un accordo e in letteratura l'interpretazione di termini quali competenza, proficiency o performance rimane spesso confusa e il dibattito in proposito è ancora aperto (cfr. Llorca, 2000; Purpura, 2004; Heil, 2014; Turner, Purpura, 2015). I prossimi paragrafi inquadrano i problemi più salienti e riassumono i termini del dibattito. Si conclude introducendo il quadro di riferimento impiegato in questo studio.

### 1.1.1. *La competenza*

Il termine competenza è sicuramente tra i più controversi nel campo della linguistica generale e della linguistica applicata. La confusione è dovuta al fatto che i vari autori usano questo termine riferendosi a nozioni sostanzialmente diverse.

Chomsky interpreta la competenza come una proprietà statica e assoluta dell'individuo. Nel suo testo *Aspects of the theory of syntax* definisce così la nozione di competenza, opponendola a quella di performance:

Linguistic theory is primarily concerned with an ideal speaker-listener, in a completely homogeneous speech community, who knows its language perfectly and is unaffected by such grammatically ir-



relevant conditions as memory limitations, distractions, shifts of attention and interest, and errors (random or characteristic) in applying his knowledge of the language in actual performance. [...] We thus make a fundamental distinction between competence (the speaker–hearer’s knowledge of the language) and performance, the actual use of language in concrete situations. (Chomsky, 1965: 3–4)

La competenza in questo caso è dunque vista come uno stato mentale astratto che esclude l’abilità linguistica «it does not seem to me to be quite accurate to take “knowledge of English” to be a capacity or ability, though it enters into the capacity or ability exercised in language use» (Chomsky, 1975: 23). Quello che interessa al linguista è il prodotto e non il processo, come chiarisce Taylor (1988: 153): l’idea di competenza di Chomsky non ha nulla da dire né rispetto all’abilità a impiegare la rappresentazione mentale nella comunicazione, né rispetto a come i parlanti usano o apprendono tale competenza.

L’idea di competenza intesa come conoscenza di un parlante ideale con un controllo perfetto della lingua in questione è stata presto criticata da chi aveva un approccio sociolinguistico allo studio del linguaggio e vedeva la distinzione effettuata da Chomsky come riduttiva, in quanto confinava tutti i casi di variabilità alla performance, escludendoli così dall’oggetto di indagine. Particolarmente influente in questo senso è il lavoro condotto da Hymes (1971; 1972), che, criticando l’attribuzione di tutti i fattori situazionali e socioculturali alla sola performance, sottolinea «the distorting effect of using the one term “performance” for two distinct things: a theory of performance and a theory of language use» (1971: 11). In altre parole Hymes ritiene che alcuni aspetti di ciò che Chomsky relega alla performance siano sistematici, descrivibili sotto forma di regole, e pertanto da includere in un modello di competenza, che egli definisce appunto “competenza comunicativa”. Lo

stesso Chomsky concorderà con queste osservazioni, aggiungendo alla competenza grammaticale la competenza pragmatica, ossia la capacità di far uso della conoscenza definita come competenza grammaticale (Chomsky, 1980: 59).

Rispetto a Chomsky però, con l'introduzione della nozione di competenza comunicativa Hymes cambia profondamente il concetto di competenza, espandendolo in due direzioni. Innanzitutto va oltre la grammatica includendo «what is feasible, appropriate and actually performed». Poi introduce nella nozione di competenza anche la dimensione sociale del linguaggio, includendo nel quadro l'abilità d'uso: «competence is understood to be dependent on two things: (tacit) knowledge and (ability for) use» (Hymes, 1971: 16; 1972: 282).

Nemmeno l'approccio adottato da Hymes è esente da critiche. Taylor (1988: 156) per esempio sostiene che Hymes, includendo nella definizione di competenza la nozione di abilità, avrebbe trasformato il concetto di competenza a tal punto da avergli fatto perdere un significato preciso:

For Hymes the general term competence covers a number of different elements, ranging from grammatical competence on the one hand to sociolinguistic competence on the other. (It is to be noted that Hymes's "sociolinguistic competence" is much wider than Chomsky's "pragmatic competence".) The term applies then to something that Chomsky would see as biologically based (grammatical competence) and at the same time to something much more socially based (sociolinguistic competence). The former is purely individual, the latter is mainly social. The former concerns form; the latter concerns function. The former characterizes a state; the latter involves processes.

Canale e Swain (1980; 1981) e Canale (1983) riprendendo il lavoro di Hymes sviluppano il concetto di competenza co-

municativa con particolare riferimento alla L2: la competenza non è costituita esclusivamente dalla conoscenza della grammatica della lingua, ma in modo anche più determinante dal modo in cui l'apprendente utilizza tale abilità per comunicare nei diversi contesti. La competenza comunicativa è pertanto costituita da quattro componenti: la competenza grammaticale, la competenza sociolinguistica, la competenza discorsiva e la competenza strategica. La competenza grammaticale è data dalla conoscenza del codice linguistico, include pertanto le regole morfosintattiche, il lessico e la pronuncia. La competenza sociolinguistica riguarda la selezione delle forme linguistiche adatte a un determinato contesto e a un determinato scopo. La competenza discorsiva è l'abilità che permette di scegliere forme e significati utili a organizzare coerentemente un testo o un messaggio. La competenza strategica permette infine di compensare la mancata conoscenza di regole della lingua in questione. Canale e Swain (1980; 1981) usano il termine competenza strategica per indicare le modalità con cui il parlante risolve un problema comunicativo: la competenza strategica è dunque per definizione compensatoria, si attiva quando le altre competenze non ci sono o non sono in quel momento disponibili.

Competenza linguistica e competenza comunicativa indicano dunque due nozioni sostanzialmente diverse, una più restrittiva che esclude l'abilità d'uso e una più ampia che la include. Ellis (1994) ad esempio esplicita la distinzione sottolineando come la competenza sia «a language user's underlying knowledge of language» (p. 697) e la competenza comunicativa «the knowledge that users of a language have internalized to enable them to understand and produce messages in the language» (p. 696). Wiemann e Backlund (1980: 190) affrontano la definizione di competenza comunicativa da un punto di vista didattico, precisando come non possa che esservi un legame molto stretto tra abilità e

competenza in quanto «*proficiency skills* [...] are what is required for the manifestation of communicative competence». D'altronde già Bruner (1973) includeva gli *skill* nella sua definizione di competenza sostenendo che quest'ultima è l'obiettivo dell'intervento didattico e pertanto può essere appresa: "what is learned is competence, not particular *performance*" (Bruner, 1973: 111).

Ricapitolando, possiamo osservare come la confusione che caratterizza la definizione di competenza sia legata all'ambito entro il quale le diverse definizioni sono nate: da un punto di vista strettamente linguistico infatti la competenza è vista come proprietà statica e assoluta, dal punto di vista educativo o psicologico, viene descritta come proprietà relativa e per definizione dinamica. Come riconosce Taylor (1988: 165–166) la confusione è dovuta al fatto che la definizione di competenza in senso chomskiano non sia adatta a far fronte alla variazione e al mutamento tipici dell'apprendimento. Quando come in questo studio l'obiettivo è indagare lo sviluppo linguistico, sia esso in L1 o in L2, diventa essenziale prendere in considerazione i processi ed essere pertanto in grado di distinguere «between what a speaker knows and what he does, between his knowledge and his proficiency in using that knowledge, between what is individual and what is social» (Taylor, 1988: 166). A questo proposito, rifacendoci a Taylor (1988: 166), in questo studio per competenza e performance si intenderanno le nozioni chomskiane e si utilizzerà *proficiency* per indicare la capacità di utilizzare la competenza. I principali modelli utilizzati per descrivere la *proficiency* saranno illustrati nella sezione che segue.

### 1.1.2. *La proficiency, l'abilità comunicativa e l'abilità d'uso*

Il termine *proficiency* è strettamente legato alla letteratura relativa al testing linguistico, e come per la nozione di competenza la sua definizione è stata oggetto di dibattito. Morrow

(1979/81: 18) ad esempio ne propone una definizione intuitiva quanto vaga: «how successful the candidate is likely to be as a user of the language in some general sense». Altrettanto vaga e ingannevolmente semplice è la definizione di Ingram (1985: 220): «what is meant when we say someone is proficient in a language is that that person can do certain things in that language». Altri autori ritengono la proficiency definibile in termini più restrittivi e tecnici come ciò che viene misurato dai test (Vollmer, 1981; Pienemann, 1985).

Stern (1983: 341) offre una definizione più precisa di proficiency descrivendola come «the actual performance of given individual learners or groups of learners». Essa pertanto richiede una serie di diverse conoscenze e abilità (p. 346):

- a) una conoscenza intuitiva delle forme della lingua;
- b) una conoscenza intuitiva dei significati linguistici, cognitivi, affettivi e socioculturali espressi dalle forme linguistiche;
- c) la capacità di utilizzare la lingua con attenzione massima al significato e minima alla forma;
- d) la creatività dell'uso linguistico.

Come evidenzia anche Taylor (1988: 163) Stern include nell'idea di proficiency sia la nozione statica di competenza elaborata da Chomsky, sia la definizione di abilità d'uso introdotta da Hymes. Stern (1983: 357) sottolinea inoltre come la proficiency possa essere indagata secondo due prospettive: una che mira a stabilirne i livelli, in una progressione dal livello base a quello quasi-nativo, l'altra che si concentra invece sulle componenti della proficiency stessa.

Diversi studi, in particolare nell'area della didattica e del testing, si sono occupati della definizione delle componenti della proficiency linguistica. In una fase iniziale, sulla base

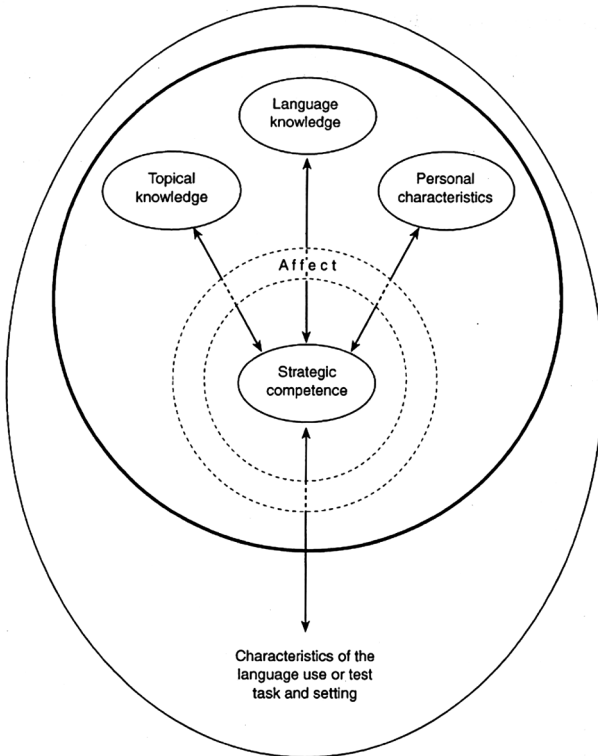
delle ricerche condotte da Oller (1979; 1983) e la conseguente elaborazione della sua *Unitary Trait Hypothesis*, si riteneva che la proficiency fosse un costrutto unitario. Questa proposta subì diverse critiche, in particolare rispetto alle procedure statistiche impiegate. Vollmer e Sang (1983) ad esempio, rianalizzando i dati di Oller, ne contraddicono le conclusioni e sostengono piuttosto l'idea della proficiency come costituita da diverse componenti. Nello stesso periodo Bachman e Palmer (1981), utilizzando una *Multitrait–Multimethod Matrix*, dimostrano come la proficiency nel parlato e nella lettura siano indipendenti l'una dall'altra, e dunque anche loro sostengono l'idea della proficiency come multi–componenziale.

Il modello più influente nell'ambito del testing è certamente quello proposto da Bachman e Palmer (1996; 2010), raffinamento ulteriore del modello proposto in Bachman (1990). Questi autori non utilizzano il termine proficiency, ma preferiscono piuttosto l'espressione *Communicative Language Ability* (CLA da ora in poi). Il loro modello differisce da quelli che lo hanno preceduto per le seguenti caratteristiche: in primo luogo è un modello operativo; in secondo luogo la CLA è solo un aspetto tra altri che possono influire sulla performance, quali il metodo di valutazione o le caratteristiche dell'apprendente. Infine, questo modello assegna un ruolo centrale alla competenza strategica, intesa come autonoma dalla conoscenza linguistica.

Secondo Bachman (1990: 84) la CLA può essere descritta come «consisting of both knowledge, or competence, and the capacity for implementing, or executing that competence in appropriate, contextualized communicative language use». L'uso linguistico coinvolge dunque una serie di interazioni complesse tra le caratteristiche individuali dell'apprendente da un lato e le caratteristiche della situazione comunicativa o di valutazione dall'altro. La CLA viene pertanto inquadrata all'interno di un quadro interazionale dell'uso linguistico (cfr. grafico 1.1.):

[the interactional framework] focuses on the interactions amongs areas of language ability (language knowledge and strategic competence, or metacognitive strategies), topical knowledge, and affective schemata, on the one hand, and how these interact with the characteristics of the language use situation, or test task, on the other. (Bachman, Palmer, 1996: 62)

Il modello proposto in Bachman e Palmer (1996; 2010) è costituito da 4 componenti: la conoscenza linguistica, la competenza strategica, la topical knowledge e le caratteristiche personali.



**Grafico 1.1.** Some components of language use and language test performance, da Bachman, Palmer, 1996: 63.

**Tabella 1.1.** Areas of language knowledge, da Bachman, Palmer, 1996: 68.

organizational knowledge
(how utterances or sentences and texts are organised)
grammatical knowledge
(how individual utterances or sentences are organised)
knowledge of vocabulary
knowledge of syntax
knowledge of phonology/graphology
textual knowledge
(how utterances or sentences are organized to form texts)
knowledge of cohesion
knowledge of rethorical or conversational organization
pragmatic knowledge
(how utterances or sentences and texts are related to the communicative goals of the language users and to the features of the language use setting)
functional knowledge
(how utterances or sentences and texts are related to the communicative goals of the language users)
knowledge of ideational functions
knowledge of manipulative functions
knowledge of heuristic functions
knowledge of imaginative functions
sociolinguistic knowledge
(how utterances or sentences and texts are related to features of the language use setting)
knowledge of dialects/varieties
knowledge of registers
knowledge of natural or idiomatic expressions
knowledge of cultural references and figures of speech

Come mostra la tabella 1.1., la conoscenza linguistica a sua volta è suddivisa in due categorie: conoscenza organizzativa e conoscenza pragmatica. La conoscenza organizzativa coinvolge il controllo formale delle strutture linguistiche necessarie per produrre e comprendere frasi grammaticalmente



accettabili e per organizzare frasi e enunciati in testi scritti e orali; essa comprende pertanto due componenti: la conoscenza grammaticale e la conoscenza testuale. La conoscenza pragmatica guida l'interpretazione del discorso mettendo in relazione frasi ed enunciati con il loro significato, le intenzioni del parlante e le caratteristiche del contesto comunicativo; essa comprende pertanto due componenti: la conoscenza funzionale e la conoscenza sociolinguistica.

La componente competenza strategica è descritta invece come composta da «higher order executive processes that provide a cognitive management function in language use, as well as in other cognitive activities» (Bachman, Palmer, 1996: 70). A differenza dei modelli precedenti, la competenza strategica non è vista come abilità compensatoria, attivata per colmare lacune, ma ha un ruolo centrale: essa ha una funzione di mediazione tra intenzione comunicativa, competenze linguistiche, conoscenza del mondo e contesto situazionale. Tre sono le aree in cui operano le componenti metacognitive della competenza strategica: definizione degli obiettivi, valutazione e pianificazione. La tabella 1.2. riporta le aree d'uso della competenza strategica.

La CLA comprende dunque sia la proficiency linguistica, intesa come capacità d'uso della lingua, sia la proficiency comunicativa, che include la conoscenza del mondo e delle strategie necessarie per applicare la proficiency linguistica alle diverse situazioni. Come nota McNamara (1995: 172) la novità del modello di Bachman (1990) consiste nel riconoscere il ruolo degli aspetti della capacità d'uso delle competenze linguistiche ridefinendo il rapporto tra competenza e performance in modo dinamico.

Il modello di Bachman è stato usato in modo piuttosto diffuso in letteratura, in particolare nel testing. Esso ha il vantaggio, rispetto ai modelli precedenti, di essere molto più

**Tabella 1.2.** Areas of metacognitive strategy use, da Bachman, Palmer, 1996: 71.

goal setting (deciding what one is going to do)
identifying the test tasks choosing one or more tasks from a set of possible tasks (sometimes by default, if only one task is understandable) deciding whether or not to attempt to complete the task(s) selected
assessment (taking stock of what is needed, what one has to work with, and how well one has done)
assessing the characteristics of the test task to determine the desirability and feasibility of successfully completing it and what is needed to complete it assessing our own knowledge (topical, language) components to see if relevant areas of knowledge are available for successfully completing the test task assessing the correctness or appropriateness of the response to the test task
planning (deciding how to use what one has)
selecting elements from the areas of topical knowledge and language knowledge for successfully completing the test task formulating one or more plans for implementing these elements in a response to the test task selecting one plan for initial implementation as a response to the test task

dettagliato nella specificazione delle componenti della competenza linguistica, più preciso nel descrivere il legame che intercorre tra le diverse componenti, più dipendente dalle teorie linguistiche e maggiormente supportato dalla ricerca.

Skehan (1995b; 1998; 2008; 2014) pur riconoscendo i vantaggi del modello, ne evidenzia però un limite sostanziale: il fatto che esso non sia basato su processi e meccanismi psicolinguistici. Questo rende il modello una valida *checklist*, ma non permette di fare considerazioni rispetto alla natura della performance e del suo legame con la competenza (Skehan, 1998: 164). Rifacendosi a modelli psicolinguistici, Skehan

(1995; 1998; 2008; 2014) propone di considerare sia come la competenza linguistica viene rappresentata sia i processi legati al suo uso. In particolare, se la relazione tra competenza e performance è dinamica e mediata dalla competenza strategica, allora è necessario stabilire:

- a) se le operazioni attivate dalla competenza strategica sono legate a una rappresentazione mentale astratta (competenza) o piuttosto a processi in tempo reale che influenzano ciò che viene detto (performance);
- b) se la consapevolezza della competenza strategica, e le possibili conseguenze di tale consapevolezza sul modo di operare di questa competenza, dipendono dalla competenza o dalla performance;
- c) che cosa sono esattamente le diverse operazioni attivate dalla competenza strategica e quale processo le coinvolge.

Per capire come la competenza strategica opera può essere utile vedere come la competenza linguistica è rappresentata psicologicamente. In questo senso il lavoro di Skehan condivide una prospettiva di elaborazione dell'informazione dell'uso linguistico (*information processing view*) con altri autori (cfr. per es. Ellis, 1987a; 1987b; Meisel *et al.*, 1981; Van Patten, 1990). Skehan (1985b; 2008) riprende la distinzione introdotta da Widdowson (1989) tra analizzabilità e accessibilità. La distinzione di Widdowson è legata alla natura dell'immagazzinamento psicologico (*psychological storage*) del materiale linguistico e alla modalità con cui la comunicazione viene eseguita in tempo reale. L'analizzabilità dipende dal modo con cui l'informazione linguistica è immagazzinata ed è pertanto strettamente collegata alla rappresentazione mentale della conoscenza linguistica stessa. Per l'analizzabilità la regola è di primaria importanza, il

sistema pertanto è elegante e compatto e l'unità di riferimento è la parola. L'accessibilità invece ha a che fare con le richieste operative della comunicazione, è legata alla possibilità di accedere alla conoscenza linguistica e interessa l'abilità d'uso, ossia la capacità di mettere in atto la conoscenza. Per l'accessibilità il lessico e le formule sono di primaria importanza, il sistema è facile da usare e soprattutto veloce. Qui l'unità di riferimento è la "frase lessicalizzata". Analizzabilità e accessibilità sono dunque i due processi di un *dual-parallel coding system* basato rispettivamente su regole astratte (*rule-based system*) ed esempi (*exemplar-based system*). Poiché le risorse di elaborazione sono limitate, in una situazione comunicativa che richiede un alto grado di *processing*, l'uso formulaico di unità più lunghe delle parole permette di liberare risorse di attenzione. Il parlante pertanto, cercando di salvaguardare un certo livello di fluenza, tende ad adattare il parlato lungo un continuum di stili di performance. In questo senso la proposta di Skehan (1995) riprende anche l'idea di interazione tra efficacia comunicativa e accuratezza suggerita da Harold Clahsen, Jürgen Meisel e Manfred Pienemann nel loro *Multidimensional Model* (Clahsen, 1980; Meisel *et al.*, 1981; Pienemann, 1987) secondo cui un apprendente si situa in una determinata posizione del continuum variazionale basandosi sulla sua decisione di distribuire le sue limitate risorse di elaborazione. Si identificano così due estremi del continuum, uno rappresentato da apprendenti che preferiscono evitare deviazioni dalla norma della lingua standard, l'altro da apprendenti che concentrano l'attenzione sull'aspetto comunicativo dell'interazione. La proposta di Skehan si rifà anche allo spettro di stili identificato da Tarone (1988) per l'interlingua, dove l'apprendente si muove tra uno stile più accurato e uno più rilassato e colloquiale. Tuttavia Tarone (1988), come Ellis (1987a; 1987b), al contrario di Skehan, vede questo tipo di variazione sistematica come variazione della competenza:

Planning variability is seen as a feature of the learner's competence, not just of his performance. At any stage of development, the learner possesses two or more rules for at least some structures and calls on this differently according to whether the discourse is planned or unplanned. (Ellis, 1987: 14)

Skehan invece interpreta la variazione come parte della performance, qui intesa come esecuzione di un task, quindi la variazione è il riflesso delle priorità di un apprendente rispetto alle sue limitate risorse d'attenzione. Il cambiamento di stile in relazione alle richieste del contesto comunicativo porta così a una ridefinizione del rapporto tra competenza e performance attraverso l'introduzione del costrutto dell'Abilità d'Uso (*Ability for Use*):

Rather than draw upon a generalized and stable underlying competence, the second language performer adjusts to performance conditions by trying to allocate attention in appropriate ways. This enables real-time communication to be more feasible and some degree of fluency achieved. When communicative pressure is not that heavy, when precision is important, or when task demands emphasize form, a syntactic mode assumes greater importance. When pressure is greater and/or when effective communication is paramount, a lexical mode is used more. (Skehan, 1998: 168-69)

L'apprendente dunque attraverso questa competenza di processing è in grado di gestire le diverse richieste della comunicazione rifacendosi a due sistemi di codifica paralleli: la modalità sintattica basata su un sistema di regole (*rule-based system*) e la modalità lessicale basata sugli esempi (*exemplar-based system*). La coesistenza dei due sistemi permette una certa flessibilità nell'uso linguistico. Se la comunicazione può sfruttare sia l'una modalità sia l'altra, o entrambe, allora a volte l'apprendente uti-

lizzerà principalmente la modalità sintattica, e la competenza linguistica sarà il primo fattore di influenza (così come assunto dagli approcci precedenti), a volte invece, quando il tempo reale della comunicazione ha un'influenza significativa, sarà predominante la modalità lessicale, con il risultato che la performance non sarà riflesso diretto della competenza, ma piuttosto di una serie di fattori legati all'elaborazione. Ribaltando così la visione convenzionale che vede la performance basata sulla competenza, ma inevitabilmente influenzata da una serie di fattori che con la competenza non hanno niente a che fare (fatica, complessità del messaggio, stato emotivo, ecc.), Skehan (1995: 106) con questo approccio duale rafforza il ruolo della competenza di Bachman: «Ability for Use, in other words, is what goes beyond Bachman's (1990) assessment, goal-setting, planning and execution, and is what accounts for the balance between analysability and accessibility as the processing dimension of actual communication». Inoltre, introducendo l'Abilità d'Uso come elemento di mediazione tra le competenze del parlante (cfr. l'analizzabilità e il sistema basato sulle regole) e le richieste del contesto (cfr. l'accessibilità e il sistema basato sugli esempi) focalizza l'attenzione sulle capacità di elaborazione. Sulla base di queste considerazioni Skehan (2014) suggerisce come sia più efficace esaminare direttamente le operazioni attivate dai fattori di processing (che egli identifica nell'*Ability for Use*) piuttosto che le diverse componenti della competenza linguistica.

In questa prospettiva il processo di apprendimento è legato essenzialmente a due operazioni, la ristrutturazione e l'automatizzazione (McLaughlin, 1990). La ristrutturazione è il processo che causa la crescita della complessità sintattica attraverso «the evolution of increasingly abstract representations of knowledge» (Foster, Skehan, 1996; Schmidt, 1992), mentre l'automatizzazione risulta dal rafforzamento della memoria per formule e routine. Come sostiene Schmidt (1992) si tratta

di due processi sostanzialmente diversi ma correlati. La ristrutturazione può essere vista infatti come il risultato dell'analisi di formule sempre più complesse e della loro automatizzazione. E a questo proposito McLaughlin (1990) suggerisce come l'automaticità a livelli inferiori libera risorse di elaborazione per la ristrutturazione a livelli superiori. Analogamente Ellis (1996) ipotizza che la memorizzazione delle formule sia alla base non solo dell'automaticità, ma anche della ristrutturazione. Secondo questa visione l'apprendimento parte dall'esposizione a una forma, procede alla crescente esposizione e memorizzazione di più esempi, e arriva alla generalizzazione, mentre l'ulteriore esposizione porta infine alla memorizzazione delle eccezioni. Così gradualmente si sviluppa una rappresentazione sempre più complessa di una forma che include i diversi esempi, sia regolari sia irregolari, mentre l'accesso a tale forma diviene man mano più automatico.

Skehan (1998; 2014) propone di analizzare le produzioni degli apprendenti a partire dalle due dimensioni di processing coinvolte, analizzabilità e accessibilità, e considerando le due modalità, quella sintattica e quella lessicale. Le limitate capacità di memoria richiedono al parlante di scegliere se favorire la forma o il significato. Secondo Skehan, la forma si manifesta nella complessità (se si creano possibilità di *restructuring*, poiché l'apprendente decide di rischiare/sperimentare) o nell'accuratezza (in caso venga data priorità al controllo), mentre il significato è rappresentato dalla fluenza. Sulla base di queste considerazioni Skehan (1998: 270) sostiene che complessità, accuratezza e fluenza possano essere considerati indicatori efficaci nella misurazione della performance in un determinato task.

Ricapitolando, Skehan ridefinisce il rapporto tra competenza e performance introducendo l'Abilità d'Uso, ossia la capacità dell'apprendente di gestire le proprie risorse sulla base delle richieste del contesto. Nel suo modello la variabilità diventa

parte integrante della performance e riflette processi psicolinguistici. In questa prospettiva è dunque possibile analizzare le produzioni degli apprendenti a partire da tre dimensioni: complessità, accuratezza e fluenza. La complessità, ossia il grado di elaborazione della lingua prodotta, viene vista in due modi. In primo luogo, un apprendente può variare nella propensione a usare strutture più complesse e rischiose (Skehan, 2001). In questo caso sono considerate complesse le strutture che si trovano ai limiti del sistema interlinguistico e non ancora completamente automatizzate. In secondo luogo, la complessità si può riferire alla predisposizione dell'apprendente a sperimentare una varietà di strutture. L'accuratezza invece si riferirà a «how well the target language is produced in relation to the rule system of the target language» (Skehan, 1996a: 23). Essa è generata da tre diverse fonti: il grado di accuratezza della rappresentazione stessa, la forza di eventuali rappresentazioni in competizione tra loro, e il grado di automatizzazione della produzione. Durante il processo di apprendimento, infatti, i diversi aspetti della rappresentazione mentale dell'apprendente possono essere più o meno vicini alla lingua target, oltre che differenziarsi per livello di automatizzazione. Gli apprendenti che danno precedenza all'accuratezza esercitano un controllo sugli elementi che hanno internalizzato, adottando un atteggiamento conservativo verso l'uso della L2. La fluenza infine rivela quanto è fluido il parlato di un apprendente, privo di pause o esitazioni inutili. Parte dello sviluppo linguistico, abbiamo visto, è legato chiaramente all'agilità con cui un apprendente può accedere agli elementi linguistici richiesti dall'interazione: la fluenza si ottiene proprio attraverso strategie di processing che permettono all'apprendente di evitare o risolvere problemi velocemente. Viene favorita quando gli apprendenti danno priorità al significato sulla forma così da concludere efficacemente il task.



Rifacendosi a questo approccio in questo studio si analizzano le produzioni dei soggetti coinvolti in termini di complessità, accuratezza e fluenza. In particolare, l'attenzione sarà rivolta alla descrizione della variazione di queste tre dimensioni nel tempo e in diversi task. Poiché il tema della variazione è centrale, le sezioni che seguono sono dedicate alla presentazione della letteratura sia in ambito acquisizionale che nel testing, con attenzione specifica agli strumenti di misurazione adottati.

## **1.2. L'approccio acquisizionale**

Il processo di sviluppo segue percorsi universali comuni a tutti gli apprendenti: l'interlingua di un apprendente in un preciso momento è un sistema coerente e autonomo. Altrettanto riconosciuto il fatto che l'interlingua sia un sistema variabile: nonostante la sistematicità dei percorsi evolutivi, l'interlingua di un apprendente costituisce un sistema instabile, permeabile a nuove forme linguistiche e questa dinamicità si riflette proprio nella variabilità delle produzioni. Descrivere come avviene l'apprendimento della L2 significa dunque spiegare anche il ruolo di tale variabilità in relazione allo sviluppo: la variabilità è "rumore" da regolarizzare o elemento intrinseco del processo di apprendimento, nonché molla per l'acquisizione? Questa sezione è dedicata al tema della variabilità in prospettiva acquisizionale. Dapprima si discuterà degli approcci utilizzati per descrivere la variazione interlinguistica (§ 1.2.1.), poi si affronterà il tema della sua misurazione (§ 1.2.2.).

### *1.2.1. La variazione*

Sistematicità e variabilità possono apparire come caratteristiche contraddittorie. Eppure nell'ambito dello studio delle L1,

sono state individuate modalità per descrivere la variabilità nel parlato dei nativi, interpretabile come risultato dell'applicazione di regole variabili<sup>1</sup> (per es. Labov, 1970; Bayley, 1973) o come grammatiche sovrapposte (per es. Bickerton, 1971), dimostrando appunto come sistematicità e variabilità siano in realtà conciliabili. Rifacendosi a questi studi e ai resoconti relativi ai sistemi linguistici dinamici utilizzati nelle scale implicazionali (Bickerton, 1971; De Camp, 1973), a partire dagli anni '80 sono state indagate tre dimensioni di variabilità in ambito acquisizionale:

- a) il contesto linguistico in cui una certa struttura viene prodotta;
- b) il contesto sociolinguistico con i relativi fattori situazionali, quali il grado di familiarità con l'interlocutore;
- c) il contesto psicolinguistico con i relativi fattori legati alla psicologia della situazione comunicativa, come ad esempio il tempo a disposizione per la pianificazione.

In letteratura si possono così identificare tre distinti approcci allo studio della variabilità e della sistematicità dell'interlingua: l'*Homogeneous Competence Model* di tradizione chomskiana, l'approccio sociolinguistico e quello psicolinguistico. Secondo il modello chomskiano la variazione è vista come aspetto della performance piuttosto che della competenza, e pertanto esula dalla considerazione del ricercatore, il cui obiettivo è regolarizzare, standardizzare e decontestualizzare le proposizioni parlate e avere accesso così a dati invariabili e dunque rappresentativi della competenza linguistica. Poiché esclude di per sé la variazione, questo approccio non sarà qui considerato. L'ap-

1. Una regola variabile viene definita nel seguente modo: la regola x viene utilizzata con probabilità X nel contesto sociale Y dato il contesto linguistico Z.

proccio sociolinguistico analizza la variazione che ricorre nelle produzioni di un parlante come risultato di differenze nel contesto situazionale, occupandosi dunque non solo di ciò che lo studente sa, ma anche della sua abilità nell'utilizzare tali conoscenze nella comunicazione. Infine l'approccio psicolinguistico intende la variazione come risultato dell'abilità dell'apprendente di elaborare le conoscenze in L2 in diversi contesti d'uso.

Nell'ultimo ventennio alcuni ricercatori hanno proposto per lo studio dei processi di acquisizione un ulteriore approccio, detto emergentista (*emergentism*), basato sulla *Dynamic System Theory*, *Chaos/Complexity Theory* e le teorie di *Emergence of Language*. L'apprendimento linguistico viene qui interpretato come un sistema dinamico, ossia un sistema complesso e non lineare in costante mutamento: «each phenomenon is dynamic, complex, nonlinear, chaotic (at times), unpredictable, sensitive to initial conditions, open, self-organizing, feedback sensitive, adaptive, characterized by strange attractors, which are fractal in shape» (Larsen-Freeman, 1997). Se la sociolinguistica e la psicolinguistica giustificano tale instabilità/variabilità come sistematica e governata da regole, l'approccio emergentista interpreta la sistematicità della variazione come il risultato delle capacità di adattamento del sistema e dunque caratteristica intrinseca del sistema stesso.

Nei prossimi paragrafi si presentano i risultati delle ricerche condotte in ambito acquisizionale. Lo scopo di questo capitolo non è individuare il modello migliore, ma piuttosto evidenziare i diversi aspetti messi in luce nei vari approcci e nelle prospettive teoriche adottate in letteratura. Come già suggeriva Zuen-gler (1989: 66) «it is misguided to search for one comprehensive theory, since one theory will most likely be insufficient in explaining the complexity of performance variation». In questo studio pertanto ci si riferirà a diverse prospettive teoriche, cercando di cogliere da ognuna strumenti utili per descrivere la

variazione. Dapprima si presenta l'approccio sociolinguistico (§ 1.2.1.1.), poi quello psicolinguistico (§ 1.2.1.2.), e infine si illustra quello emergentista (§ 1.2.1.3.).

### 1.2.1.1. L'approccio sociolinguistico

I modelli sociolinguistici elaborati in letteratura riguardano principalmente tre paradigmi: quello di William Labov, ripreso da Elaine Tarole e Rod Ellis per la L2, quello dinamico e quello socio-psicologico (per una rassegna dell'ampio dibattito scientifico, cfr. ad es. Eckman, 1994; Dewaele, 2007; Bayley, Lucas, 2007; Adamson, 2008; Bayley, 2007; Tarone, 2012; Chambers, Schilling, 2013). Ciascuno di questi paradigmi ha offerto diversi strumenti per l'analisi delle produzioni di apprendenti, contribuendo in modo diverso alla descrizione della variabilità. Il paradigma laboviano ha permesso di esaminare la variazione stilistica sulla base di due costrutti: lo stile discorsivo e la regola variabile; il paradigma dinamico ha posto le basi per l'esplorazione dei rapporti tra forma e funzione e del loro sviluppo nell'interlingua; e infine il paradigma socio-psicologico della *Speech/Communication Accommodation Theory* ha contribuito allo studio della variazione legata a caratteristiche dell'interlocutore e della situazione di interazione. Nei paragrafi che seguono illustro i tre paradigmi in relazione agli studi acquisizionali.

Il primo paradigma che si considera è quello di Labov. Nato nell'ambito della L1 ha avuto grande influenza negli studi iniziali sulla variazione in ambito acquisizionale. Due sono gli elementi chiave del paradigma: stile discorsivo e regola variabile. Labov (1970; 2001) identifica cinque assiomi legati allo stile discorsivo:

- a) non c'è un unico stile discorsivo;
- b) gli stili possono essere ordinati rispetto al grado di attenzione verso il parlato;

- c) lo stile colloquiale è quello in cui si presta meno attenzione al *monitoring*;
- d) lo stile colloquiale non può essere osservato in contesti formali;
- e) l'unico modo per ottenere "buoni dati" è l'osservazione sistematica.

Le regole variabili registrano invece tre diversi aspetti:

- a) il contesto linguistico;
- b) il contesto sociale in cui una forma linguistica viene utilizzata;
- c) la probabilità di applicazione di una regola.

Tarone (1985; 2007) e Ellis (1985a; 1985b) riprendendo il lavoro condotto da Labov propongono due modelli teorici per l'interpretazione della variabilità dell'interlingua.

Tarone (1983) con il *Continuous Competence Model* sostiene che l'interlingua di un apprendente costituisce un continuum di stili discorsivi che differiscono tra loro per grado di variabilità e coerenza interna. Tarone (1988) presenta 75 studi che mirano a indagare la variabilità dell'interlingua. Buona parte di queste ricerche si riferisce alle differenze dovute al contesto comunicativo, ossia come cambiamenti nelle caratteristiche dei task e nelle procedure di elicitazione possano portare a cambiamenti nel livello di accuratezza d'uso di determinate strutture a tutti i livelli linguistici: fonologico, morfologico, sintattico o lessicale. A livello fonologico, ad esempio, Dickerson (1975) mostra come giapponesi apprendenti di inglese varino in accuratezza la pronuncia di /r/ con il variare del task: lettura di una lista di parole, lettura di un dialogo e conversazione libera, raggiungendo un livello di accuratezza inferiore nella conversazione libera e superiore nella lettura di una lista di parole. A livello

morfologico, Tarone (1985) ad esempio mostra come apprendenti giapponesi e arabi di inglese variano il grado di accuratezza nell'uso del morfema *-s* del plurale con il variare del task (per es., test grammaticale a scelta multipla, intervista e racconto di una storia) ottenendo un maggior livello di accuratezza nel test grammaticale. Opposto il pattern di variazione per l'articolo: in questo caso il maggior livello di accuratezza si riscontra nel task racconto di una storia. A livello sintattico Liu (1991) e Tarone, Liu (1995), in uno studio longitudinale sullo sviluppo interlinguistico di un bambino cinese di 6 anni in tre contesti differenti per grado di familiarità con l'interlocutore, mostrano come lo sviluppo delle forme interrogative sia diverso nei tre contesti situazionali considerati. Ogni nuova forma interrogativa si sviluppa prima nell'interazione con un adulto familiare in contesto di gioco a casa, poi nell'interazione con i compagni in classe, e solo alla fine nell'interazione con l'insegnante.

Nel modello elaborato da Elaine Tarone la scelta dello stile dipende principalmente dalla quantità di attenzione diretta alla forma linguistica rispetto al contenuto: maggior attenzione alla forma porterebbe a una riduzione della variabilità (Tarone, 2000; 2010). L'acquisizione pertanto procede in due modi diversi. Da un lato, le nuove forme interlinguistiche sono inizialmente interiorizzate quando il parlante partecipa a conversazioni spontanee, poi dallo stile colloquiale si propagano negli stili più accurati. Dall'altro, le forme che vengono inizialmente prodotte negli stili più accurati (perché apprese attraverso l'insegnamento) si propagano poi nello stile colloquiale. La variazione viene così strettamente collegata all'acquisizione.

Ellis (1985a; 1985b) propone il *Variable Competence Model*. Questo modello identifica due tipi di variabilità dell'interlingua: la variabilità sistematica, che include quella individuale e del contesto, e la variabilità non sistematica, che include quella libera e quella dovuta alla performance. Nell'un caso l'inter-

lingua è costituita da regole in competizione tra loro e strettamente collegate alla situazione o a fattori contestuali; nell'altro l'alternarsi di regole diverse è invece casuale, e la grammatica dell'apprendente è indipendente da contesto e situazione. L'acquisizione, secondo Ellis, avverrebbe in due tappe: prima si assimilano nuove forme, poi si ricostruisce la relazione forma-funzione. La variazione libera indica dunque che l'apprendente ha individuato una nuova forma, ma non l'ha ancora messa in relazione con la sua funzione specifica. Per Ellis (1999) la variazione libera è dunque la molla che fa scattare l'acquisizione.

Nei due modelli di Elaine Tarone e di Rod Ellis la competenza viene definita come variabile. Tale posizione non è stata esente da critiche. Gregg (1990) ad esempio, fedele alla sua formazione generativista, pur riconoscendo la variabilità della performance come fatto incontrovertibile, pone alcune questioni metodologiche, e mette in dubbio l'utilità stessa di includere la variabilità in una teoria dell'apprendimento della L2. Ai due autori rimprovera una certa vaghezza nella definizione del costrutto competenza variabile:

In principle they claim that linguistic knowledge is itself variable for any given learner, but in practice they avoid any concrete discussion of what that knowledge consists of, leaving the "variable competence" construct up in the air and confining themselves to (often extensive, detailed, and interesting) discussion of output data. (Gregg, 1990: 378)

Secondo Gregg il concetto di competenza variabile rischia infatti di confondere i confini tra conoscenza linguistica e produzione, diventando poco coerente dal punto di vista teorico: se la variazione è parte della performance, è legittimo parlare di una teoria della variazione dell'interlingua all'interno di una teoria della performance, ma non lo è

farlo entro una teoria che mira a descrivere lo sviluppo della competenza.

Il secondo paradigma sociolinguistico è quello dinamico. In questo tipo di ricerche l'attenzione è sullo studio dei rapporti forma–funzione. La variazione interlinguistica non è qui messa in relazione con un'idea di competenza variabile. Preston (1993: 170) ad esempio suggerisce come «there is no inconsistency in believing that variation arises from alternatives (not “variation”) in competence and that the probabilistically determined realisation of those alternatives is due to a variety of contributing factors, not necessarily to a built-in variability in linguistic competence itself». Nemmeno l'idea di variazione libera introdotta da Rod Ellis è condivisa in questo approccio: la variazione non sistematica è considerata solo come apparente, probabilmente frutto di metodi di ricerca o corpora di dati inadeguati (Preston, 1996: 25).

In questo paradigma la variazione nel parlato viene indagata utilizzando le stesse tecniche impiegate in sociolinguistica per studiare la variazione tra parlanti nativi (VARBRUL e analisi multivariata), e mostrando come i vincoli che determinano la variazione siano ordinabili sulla base del peso che hanno nel definire la variazione orizzontale (Bayley, Preston, 1996; 2009; Bayley, Langman, 2004; Preston 2000; 2002; Fasold, Preston, 2006; Young, 1991; 1996). Vediamo alcuni esempi. Wolfram (1985) studiando apprendenti vietnamiti d'inglese mostra come la marca del tempo passato sia sistematicamente influenzata da fattori quali la salienza, definita come differenza tra forma base e forma passata del verbo, e, nel caso dei verbi irregolari, da aspetti del contesto fonologico. Young, (1991; 1996) ha utilizzato l'analisi variazionista (*variable rule analysis*) per studiare la marca del plurale da parte di cinesi adulti apprendenti di inglese. Il suo studio dimostra come l'uso del plurale sia soggetto a una serie di vincoli, tra cui il



livello di proficiency, la ridondanza della marca di plurale, la funzione sintattica del sintagma nominale, le caratteristiche dei segmenti che seguono e precedono. Bayley (1991; 1994), Langman, Bayley (2002) e Howard (2004) utilizzando l'analisi multivariata mostrano come l'utilizzo della marca del plurale sia influenzato da una serie di fattori sia linguistici che sociali: il passato viene impiegato preferibilmente con verbi perfettivi, mentre la forma non flessa con verbi imperfettivi; inoltre tale marca può variare in relazione alla regolarità/irregolarità del verbo in questione, la scelta del passato è influenzata dalla salienza e dal gruppo sociale di appartenenza dell'apprendente. Anche in questi studi la variazione sincronica è vista come precursore del cambiamento, come ad esempio illustra Bayley (1973) nel *Wave Model*: inizialmente una regola ha possibilità di applicazione zero in tutti i contesti, poi nel tempo la probabilità di applicazione cresce contesto per contesto.

A partire dagli anni '90 si sviluppa entro questo paradigma un filone di ricerca che analizza le produzioni di parlanti nativi come controllo in relazione al quale valutare l'acquisizione della medesima variazione sociolinguistica da parte degli apprendenti (cfr. Howard *et al.*, 2013). Così, per esempio, Adamson e Regan (1991) studiano l'acquisizione da parte di immigrati vietnamiti e cambogiani del pattern variabile nativo rispetto all'uso del morfema *-ing* (per es. *working* vs *workin'*) e riscontrano un effetto della variabile genere sulla scelta effettuata dal parlante, nel senso che le donne usano [-*in*] meno frequentemente degli uomini, soprattutto nel parlato controllato. Bayley (1996) invece studia l'acquisizione da parte di cinesi adulti del pattern nativo inglese della riduzione del nesso consonantico e constata come il livello di proficiency e la quantità di contatti con i parlanti nativi correlino con la variazione. Major (2004) studia l'acquisizione di forme fonologiche native da parte di giapponesi e spagnoli apprendenti di inglese, e trova differenze inte-

ressanti in relazione al genere degli apprendenti e allo stile che usano, suggerendo che la variazione preceda quella di stile, nel senso che le donne usano meno forme colloquiali degli uomini. Regan (1995; 1996; 1997; 2004) studia l'acquisizione da parte di alcuni studenti irlandesi del pattern nativo di cancellazione della particella *ne* nella negazione francese, e mostra come, dopo un anno di residenza in Francia, gli apprendenti comincino ad avvicinarsi all'uso nativo (vedi anche Dewaele, 1992; 2007; 2008; Dewaele, Regan, 2002; Donaldson, 2017). Howard *et al.* (2006) mostrano come, dopo un anno di soggiorno in Francia, così come i nativi, gli apprendenti di L2 cancellano /l/ più frequentemente con l'impersonale *il* che con *elle*. O ancora i soggetti femmini e maschili si allineano con il pattern di cancellazione di /l/ con i nativi dello stesso sesso, confermando come gli apprendenti sviluppino la stratificazione di genere prima di quella discorsiva (Major, 2004). Sankoff *et al.* (1997) studiano l'uso dei segnali discorsivi francesi da parte di parlanti non nativi residenti a Montreal, e concludono come l'uso appropriato dei segnali discorsivi sia indice di un buon grado di integrazione con la comunità francofona. In una serie di studi riferiti al francese canadese Mougeon e Rehner (2001), Rehner e Mougeon (1999) e Rehner *et al.* (2003) analizzano 13 variabili distribuite lungo il continuum sociolinguistico formale-informale, ed evidenziano come alcuni apprendenti inseriti in un programma di immersione non usino le forme più colloquiali, ma preferiscano forme intermedie. In studi più recenti poi lo studio della variazione si è dedicato anche all'indagine pragmatica, coinvolgendo non solo aree come la scelta dello stile discorsivo (cfr. Dewaele, 2008), ma anche aspetti più ampi quali gli atti comunicativi, la mitigazione o le implicature (Schneider, Barron 2008; Barron, Schneider 2009; Félix-Brasdefer, Koike, 2012). Le ricerche qui presentate concordano nel dimostrare che la variazione nell'interlingua, così come nella lingua parlata

dai nativi stessi, è sistematica e regolata da diversi vincoli. Di conseguenza apprendere una L2 significa apprendere anche i pattern di variabilità (Rehner, Mougeon, 1999).

Un terzo approccio alla variazione è rappresentato dai modelli di tipo socio-psicologico, come ad esempio la *Speech/Communication Accommodation Theory* di Giles (1977; *Language and Communication*, Special Issue 41, 2015). La variazione d'uso linguistico qui è messa in relazione con lo stato sociale, istituzionale ed etnico del gruppo di appartenenza del parlante. In base al tipo di aggiustamenti che il parlante fa durante lo scambio comunicativo vengono identificate tre dimensioni di variabilità:

- a) la convergenza, nel caso in cui il parlante adatti il suo discorso per renderlo più simile a quello dell'interlocutore;
- b) la divergenza, quando il parlante differenzia il suo eloquio rispetto a quello dell'interlocutore;
- c) la costanza (*speech maintenance*), nel caso in cui il parlante non modifichi il suo stile.

Uno dei punti forti di questo modello riguarda il riconoscimento del ruolo centrale dell'interlocutore nel determinare la variazione della produzione. L'accomodamento è identificabile a tutti i livelli linguistici, e può dunque, in contesti bilingui, riguardare tanto la scelta stessa del codice, quanto la velocità, il tono, la pronuncia, le scelte lessicali o grammaticali, nonché elementi del discorso, quali la lunghezza dei turni o la scelta dell'argomento di conversazione. Usata in linguistica acquisizionale da ricercatori quali Beebe (1981; 1987), Beebe, Zuengler (1983) e Zuengler (1989), la *Speech/Communication Accommodation Theory* ha contribuito sostanzialmente allo studio della variazione linguistica legata al ruolo dell'interlocutore. Beebe (1985) in particolare riconosce l'esistenza

di un'ulteriore dimensione specifica dell'interlingua, che definisce *extravergence*, ossia quella variazione che non rappresenta né convergenza, né divergenza, ma piuttosto è legata all'apprendimento, alla comunicazione e alle strategie di performance o alla variazione sistematica connessa al contesto linguistico tipiche dell'interlingua.

Ricapitolando, i modelli sociolinguistici qui presentati hanno offerto ai ricercatori di SLA una serie di strumenti teorici e metodologici utili per lo studio dell'interlingua. L'idea di sistematicità della variazione e l'identificazione dei fattori che la determinano sono i due elementi di maggior interesse ai fini di questo studio. Non si farà riferimento invece alla metodologia di analisi variazionista. L'approccio sociolinguistico infatti essendo legato alla visione della variazione come fenomeno di gruppo, piuttosto che individuale, può avere limiti quando applicato allo studio della variazione individuale, obiettivo degli studi SLA. Più utile ai nostri fini l'approccio psicolinguistico, illustrato di seguito.

### 1.2.1.2. L'approccio psicolinguistico

Negli studi di stampo sociolinguistico o socio-psicologico, come si è visto in § 1.2.1.1., la variabilità interlinguistica è intesa come variazione sistematica dell'accuratezza fonologica o morfo-sintattica con il mutare di caratteristiche situazionali specifiche quali il tipo di attività svolta, il grado di attenzione rivolto alla forma piuttosto che al significato o il grado di familiarità con l'interlocutore. Vediamo ora gli studi che si sono occupati di modelli di variazione secondo un approccio psicolinguistico analizzando la variazione in termini di processi mentali richiesti nella produzione del parlato.

Alcuni autori, come Meisel *et al.* (1981) e Pienemann (1988; 1992) elaborando un modello multidimensionale spiegano la

variazione dell'interlingua in termini di utilizzo di diverse componenti del sistema cognitivo di produzione linguistica nel realizzare diversi task. Pienemann raffina tale modello proponendo la Teoria della Processabilità (d'ora in poi TP). La TP come approccio permette di fare predizioni sulle traiettorie del percorso di sviluppo nell'acquisizione di qualsiasi L2. La nozione di traiettoria di sviluppo (*developmental trajectory*) implica la presenza sia di una dimensione sistematica e universale (*staged development*) sia di una dimensione di variazione che giustifica le differenze nei percorsi di sviluppo a livello individuale. Secondo questo approccio è possibile individuare diverse traiettorie di sviluppo basate sullo stesso stadio: durante il processo di acquisizione, l'apprendente accumula le regole grammaticali e le loro varianti, percorrendo così una traiettoria di sviluppo individuale, aderendo comunque a una sequenzialità universale. In questo modo la TP definisce uno spazio bidimensionale per lo sviluppo delle ipotesi, entrambe le dimensioni di questo spazio sono regolate dalla gerarchia della processabilità (Pienemann, 2008: 10-1). In questo senso per Pienemann «differences in the percentage of rule application are understood as accuracy differences which do not alter the basic nature of the maturity of the IL system» (1992: 7). Poiché l'acquisizione è definita in termini di emersione, la variabilità è ricondotta entro l'universalità del percorso (Pienemann, 1998). Pienemann suggerisce applicazioni pratiche del suo modello multidimensionale, includendo il testing e suggerendo come la TP possa fornire un'utile base teorica a chi si occupa di valutazione (cfr. Brindley, 1998). La posizione espressa da Pienemann non è esente da critiche. Le potenzialità della teoria sembrano infatti scontrarsi con alcune difficoltà legate alla sua operativizzazione, con particolare riferimento a due aspetti: la definizione operativa del costrutto emersione (cfr. Pallotti, 2007) e la selezione di task efficaci e replicabili. Se una determinata struttura emerge in momenti

diversi con il variare del contesto situazionale (cfr. Liu, 1991; Tarone, Liu, 1995), sarà essenziale prestare attenzione al contesto di elicitazione, sia per quel che riguarda la selezione di task rappresentativi, sia rispetto alla necessità di mantenerli confrontabili nei diversi studi. Poiché la TP riconduce la variabilità individuale entro un percorso universale attraverso il costrutto emersione, questo approccio non sarà qui ulteriormente considerato.

Un ulteriore gruppo di studi indaga la variazione in termini di processi mentali. Qui si inseriscono da un lato le ricerche sui modelli di pianificazione (per es. Crookes, 1989; Ellis, 1987a; 1987b; Yuan, Ellis, 2003; Foster, Skehan, 1996; Mehnert, 1998; Ortega, 1999; Wendel, 1997), dall'altro i modelli di *processing* dell'informazione cognitiva (*cognitive information processing*) dedicati all'analisi dell'uso delle risorse di attenzione durante il completamento di un task, tra i quali in particolare il *Limited Attentional Capacity Model* (per es. Skehan, 1998; 2001; 2003; 2011; 2014; 2015; 2018; Skehan, Foster, 1999; 2001) e la *Cognition Hypothesis* (per es. Robinson, 2001a; 2001b; 2003a; 2005a; 2007b; 2011a; b; 2015; 2017). Vediamo più in dettaglio il tipo di ricerche e risultati ottenuti in questo filone di indagine.

Il modello per la produzione del parlato proposto da Levelt (1989) suggerisce una serie di possibili cause psicolinguistiche di variazione. Nella produzione del parlato, in un primo tempo nel concettualizzatore si effettuano decisioni sulla varietà linguistica da utilizzare sulla base di fattori situazionali e intenzioni comunicative, poi nel formulatore il messaggio pre-verbale viene convertito in discorso pianificato tramite la selezione del lessico e della sua messa in grammatica, e infine nell'articola-tore viene tramutato in onda sonora. A questo sistema di produzione se ne affianca poi uno di comprensione, che durante la produzione fornisce al parlante il *feedback* necessario per mettere a punto eventuali adattamenti di diverso tipo.

Le ricerche condotte in questo ambito si occupano in modo particolare della nozione di attenzione, intesa come meccanismo psicolinguistico attraverso cui elementi sociali, stilistici o cognitivi influiscono sulla performance. Questi studi si rifanno in modo più o meno esplicito all'*information processing theory*, che sostiene come i parlanti abbiano limitate capacità di processing (Miyake, Shah, 1999; Baddeley, 2007) e di conseguenza non siano in grado di controllare tutti gli aspetti di un task (Anderson, 1995; Newell, Simon, 1972). Gli apprendenti, in particolare quelli ai livelli più bassi, possono avere difficoltà a prestare attenzione contemporaneamente a forma e significato, e devono decidere come distribuire le proprie risorse, favorendo un aspetto piuttosto dell'altro (Anderson, 1995; Skehan, 1996a; VanPatten, 1990). Abbiamo già introdotto a proposito in § 1.1.2. i lavori di Skehan, secondo cui la variazione della performance è caratterizzata da tre dimensioni: la complessità, ossia l'elaborazione della lingua prodotta, l'accuratezza, ossia il grado di conformità delle produzioni alla norma linguistica della lingua target, e la fluenza, ossia la capacità di produrre un discorso in tempo reale senza eccessive pause o esitazioni (Skehan, 1996a: 22). Secondo Skehan (1998; 2018), queste tre dimensioni della performance sono rilevanti in quanto influenzate dal tipo di *processing* attivato dall'apprendente. Rifacendosi a due sistemi di codifica paralleli, modalità sintattica (basata su *exemplar-based system*) e modalità lessicale (basata su *rule-based system*), il parlante gestisce le proprie risorse per rispondere alle richieste della situazione comunicativa. In determinate condizioni, l'apprendente deciderà di favorire il significato affidandosi al sistema basato sugli esempi, con un conseguente aumento della fluenza, in altre condizioni cercherà di favorire la forma appoggiandosi maggiormente al sistema basato sulle regole, ottenendo così una maggiore complessità e/o accuratezza (vedi anche Skehan, 2009; cfr. § 1.1.2.).

Un primo gruppo piuttosto cospicuo di studi si è occupato dell'impatto della pianificazione sulla produzione linguistica (cfr. Ellis, 2005; Ortega, 1999; Bygate *et al.*, 2001). Abbiamo visto che poiché hanno risorse limitate i parlanti possono avere difficoltà a prestare attenzione contemporaneamente a forma e significato, quando però hanno la possibilità di pianificare un task sono in grado di compensare le proprie limitate capacità di processing e ottenere una produzione qualitativamente migliore (Skehan, 1996). I risultati di queste ricerche dimostrano tendenzialmente un effetto positivo della pianificazione con un aumento della fluenza nelle produzioni precedute da maggiore pianificazione (per es. Crookes, 1989; Foster, Skehan, 1996; Gass *et al.*, 1999; Tavokoli, Skehan, 2005). In modo analogo un effetto positivo della pianificazione è stato osservato anche rispetto alla complessità (per es. Crookes, 1989; Foster, Skehan, 1996; Mehnert, 1998; Ortega, 1999; Yuan, Ellis, 2003; Skehan, Foster, 2005; Ahmadian, Takavoli, 2010; Ahmadian, 2012; Mehrang, Rahimpur, 2010), mentre risultati meno chiari sono stati ottenuti rispetto all'accuratezza (per es. Ellis, 1987; Crookes, 1989; Foster, Skehan, 1996; 1997; Ortega, 1999; Wigglesworth, 1997; Bygate, 2001; Ahmadian, Takavoli, 2010).

Un punto di interesse comune a queste ricerche sull'impatto della pianificazione riguarda l'attenzione rivolta agli effetti di compensazione (*trade-off*) tra complessità, accuratezza e fluenza, anche se in questo senso i risultati sono in parte contrastanti (per una review si rimanda tra gli altri a Bygate, 2001; Ellis, 2009; Skehan, 2018). Skehan e Foster (1997) ad esempio osservano un effetto compensatorio tra complessità e accuratezza, nel senso che quando cresce la prima la seconda cala. Nello studio di Mehnert (1998) una breve pianificazione sembra favorire l'accuratezza, e una pianificazione più lunga sia l'accuratezza sia la complessità. Infine, nelle ricerche condotte da Yuan e Ellis (2003) si osserva che mentre la



fluenza aumenta con la possibilità di pianificare la produzione linguistica prima del task, l'accuratezza aumenta quando questa possibilità esiste *on-line*, ossia durante la produzione. Se in letteratura è ormai assodato che la pianificazione prima di svolgere il task aiuta gli apprendenti a produrre un linguaggio più fluente e complesso, è meno chiaro se favorisca anche l'accuratezza. Alcuni studi suggeriscono che in questo caso sia più efficace la pianificazione *on-line*. Ahmadian e Takavoli (2010) esplorando gli effetti della combinazione di pianificazione durante il task e ripetizione del task rilevano effetti significativi su tutte e tre le dimensioni CAF, evidenziando come una combinazione di pianificazione e ripetizione possa compensare i possibili trade-off. Come evidenzia Ellis (2009), la maggior parte delle ricerche sulla pianificazione si sono rivolte a livelli intermedi. Gli studi, come ad esempio quelli di Kawauchi (2005) e Tavokoli, Skehan (2005) mostrano come gli effetti del *planning* siano più evidenti a livelli elementari e intermedi, con Mochizuki e Ortega (2008) che confermano come la pianificazione tenda ad avere effetti non significativi sui livelli elementari.

Una seconda area che ha stimolato diverse ricerche riguarda la ripetizione del task (per una review più esaustiva si rimanda a Ellis 2005; Bygate, 2009). Ad esempio, Gass *et al.* (1999) osservano come la ripetizione del task abbia effetti positivi sulla qualità della performance anche se non durevoli nel tempo. Analogamente Bygate (1996; 1999; 2001) in una serie di studi riporta come la ripetizione favorisca la complessità e la fluenza, mentre i risultati per l'accuratezza non sono sempre statisticamente significativi. Lynch e McLean (2001) lavorando con studenti di corsi di medicina riscontrano come nella ripetizione del task "presentazione di un poster" gli apprendenti più avanzati riescano a migliorare le loro presentazioni e cogliere spunti dai contributi degli interlocutori. Ne-

meth e Kormos (2001) usando task argomentativi osservano come la ripetizione porti a una maggior attenzione verso il contenuto e l'uso di forme di supporto. Kim e Tracy-Ventura (2013) investigano gli effetti del tipo di ripetizione, dello stesso task o di un task analogo, mostrando come la ripetizione procedurale promuova complessità e accuratezza, ma non la fluenza. Sample e Michel (2014) mostrano come la ripetizione del task abbia effetti positivi sulla fluenza anche negli apprendenti più giovani e come tre ripetizioni del compito riescano addirittura a portare a un superamento degli effetti di trade-off tra le dimensioni CAF.

Un terzo gruppo di studi si rifà a un approccio cognitivo di elaborazione dell'informazione, dedicandosi all'analisi dell'uso delle risorse di attenzione durante il completamento di un task, della relazione tra caratteristiche del task e caratteristiche della performance, e dell'impatto delle diverse condizioni in cui viene realizzato il task. In questo filone esistono due modelli: il *Limited Attentional Capacity Model* (Skehan, 1998; 2001; 2003; 2018) e la *Cognition Hypothesis/Triadic Componential Framework* (Robinson, 2001a; 2001b; 2003b; 2005a; 2007; 2011). Entrambi descrivono come la performance viene influenzata dalle caratteristiche dei task e dalle condizioni in cui sono svolti, e come le risorse di attenzione hanno un ruolo chiave durante lo svolgimento di un task, mentre raggiungono conclusioni diverse sugli effetti di alcune caratteristiche dei task sulla performance. La tabella 1.3. mette a confronto i due modelli.

Secondo il *Limited Attentional Capacity Model* proposto da Peter Skehan le risorse di attenzione attivate dalle *resource-directing variables* sono limitate, e di conseguenza l'aumento della complessità di un task può ridurre la quantità di attenzione disponibile. Quando i limiti di carico di attenzione vengono raggiunti, gli apprendenti dovranno decidere se dare

**Tabella 1.3.** I modelli cognitivi di elaborazione dell'informazione.

Limited Capacity Model, da Skehan, 1998		
code complexity	cognitive complexity	communicative stress
linguistic complexity and variety	<i>cognitive familiarity</i>	time limits
vocabulary load and variety	familiarity of topic and its predictability	speed of presentation
redundancy and density	familiarity of discourse	number of participants
	genre	length of text used
	familiarity of task	type of response
		opportunities to control interaction
	<i>cognitive complexity</i>	
	information organisation	
	amount of computation	
	clarity and sufficiency of information given	
	information type	
Triadic component framework, da Robinson, 2005		
task complexity	task conditions	task difficulty
(a) <i>resource-directing variables</i>	(a) <i>participation variables</i> , e.g.	<i>ability variables</i> , e.g.
± few elements	open solution	working memory
± here and now	one way/two way	intelligence
± no reasoning demands	convergent/divergent	aptitude
(b) <i>resource-dispersing</i>	(b) <i>participant variables</i> , e.g.	<i>affective variables</i>
± planning time	same/different gender	motivation
± single task	familiar/unfamiliar	anxiety
± prior knowledge	power/solidarity	confidence
Triadic component framework, da Robinson, 2007		
task complexity	task conditions	task difficulty
(a) <i>resource-directing variables making cognitive/conceptual demands</i>	(a) <i>participation variables making interactional demands</i>	(a) <i>ability variables and task relevant resource differentials</i>
± here and now	± open solution	h/l working memory
± few elements	± one way flow	h/l reasoning
± spatial reasoning	± convergent solution	h/l task switching
± causal reasoning	± few participants	h/l aptitude
± intentional reasoning	± few contributions	h/l field independence
± perspective-taking	needed	h/l mind reading
	± negotiation not needed	

(b) <i>resource-dispersing variables making performative/procedural demands</i>	(b) <i>participant variables making interactant demands</i>	(b) <i>affective variables and task relevant state-trait differentials</i>
± planning time	± same proficiency	h/l openness
± prior knowledge	± same gender	h/l control of emotion
± single task	± familiar	h/l task motivation
± task structure	± shared content	h/l processing anxiety
± few steps	knowledge	h/l willingness to communicate
± independency of steps	± equal status and role	h/l self-efficacy
	± shared cultural knowledge	

priorità al significato o alla forma. In altre parole, con l'aumento della complessità cognitiva del task gli apprendenti saranno costretti ad adottare una di queste due strategie: prestare attenzione al significato prima che alla forma, con una conseguente riduzione della complessità e dell'accuratezza della loro produzione a favore della fluenza, oppure prestare attenzione alla forma o alla sperimentazione di nuove strutture, con un aumento di complessità o accuratezza, a scapito della fluenza. Peter Robinson con la sua *Cognition Hypothesis* sostiene una posizione in parte contrastante. Innanzitutto le risorse di attenzione attivate dalle *resource-directing variables* non sono limitate, come sostenuto da Skehan, anzi gli apprendenti possono sfruttare tipi di attenzione che non sono in competizione tra di loro (*non-competing attentional pools*). Di conseguenza, l'aumento della complessità cognitiva del task può portare a una maggiore complessità e accuratezza della produzione. In altre parole, l'aumento della richiesta cognitiva di un task può spingere le risorse di attenzione di un apprendente verso la forma, così che l'input viene processato in modo più profondo ed elaborato. Entrambi i modelli hanno stimolato diverse ricerche che indagano gli effetti delle caratteristiche e condizioni dei task sulla performance, con i primi che si sono rilevati meno influenti rispetto ai secondi

(Skehan, 2016). Vediamo ora i principali risultati emersi. Vista l'ampiezza della letteratura a proposito, si farà riferimento agli studi maggiormente rappresentativi delle tendenze riscontrare.

Peter Skehan e Pauline Foster in una serie di studi dimostrano come alcune caratteristiche dei task influenzino la performance: un task strutturato e familiare favorisce una maggiore fluenza e accuratezza, ma non necessariamente la complessità (Foster, Skehan, 1996; Skehan, Foster, 1997; 1999); un task che richiede una giustificazione come risultato produce lingua più complessa (Skehan, Foster, 1997); i task interattivi portano a produzioni più accurate e complesse, i task monologici favoriscono invece la fluenza (Foster, Skehan, 1996; 1999; Skehan, Foster, 1997; 1999). L'aumento della complessità del task lungo la dimensione *resource-directing* genera una maggior quantità di interazione e negoziazione, quindi più *uptake* e incorporazioni dell'input, favorendo così una produzione più complessa, ma meno fluente per i task più complessi (Robinson, 2000; 2001b). Le ricerche che hanno indagato la dimensione *Here-and-Now/There-and-Then* osservano gli effetti positivi della complessità del task sull'accuratezza (Iwashita *et al.*, 2001; Rahimpour, 1997; 1999; Gilbert, 2007; Ishikawa, 2007). In una serie di studi sulla produzione scritta (Kuiken *et al.*, 2005; Kuiken, Vedder, 2004a; 2004b; 2007a; 2007b; 2007c; 2008) in cui si osservano gli effetti del fattore  $\pm$  elementi l'aumento della complessità del task genera produzioni più accurate e una maggior varietà lessicale, mentre non si riscontrano effetti rispetto alla complessità sintattica. Michel *et al.* (2007) indagando l'effetto dei fattori  $\pm$  elementi e  $\pm$  interattivo osservano come il task dialogico favorisca accuratezza e complessità lessicale, mentre il task monologico porta a un maggior uso di riparazioni e a un aumento di complessità sintattica e lessicale. In generale c'è poco supporto

rispetto alla possibilità che un task complesso generi una sintassi più complessa, maggior supporto rispetto a un aumento dell'accuratezza. Di particolare interesse anche il più recente filone di studi dedicato all'effetto del modo, scritto o orale, sulla performance (Grabowski, 2005, 2007; Granfeldt, 2008, Martínez-Flor, 2006; Kuiken, Vedder, 2011; Kormos, 2014; Vasylets *et al.*, 2017). I risultati in parte contraddittori ottenuti nei vari studi, richiedono alla ricerca approfondimenti ulteriori. Infine, Robinson (2005a) basandosi sullo studio di Niwa (2000) suggerisce come le differenze individuali potrebbero mascherare gli effetti della variazione del task (cfr. Robinson, 2011; 2012). Studi più recenti hanno infatti indagato gli effetti dell'attitudine, della motivazione, dell'ansia e della personalità (Albert, 2011; Robinson, 2007; 2011; 2017; Agnes, 2011; Kim, Tracy-Ventura, 2011; Kormos, Trebits, 2011b; 2012). Nonostante la vivacità delle ricerche in questa direzione è probabilmente prematuro cercare di appoggiare definitivamente uno dei due modelli (cfr. Robinson, 2015; Skehan, 2015). La difficoltà di trovare una soluzione è legata a diversi fattori. Le ricerche ad oggi condotte hanno utilizzato task e condizioni molto diversi, per cui i confronti tra ricerche non sono scontati. Poiché manca ancora in letteratura un accordo su misure standardizzate, i vari studi utilizzano diverse misure per operazionalizzare complessità, accuratezza e fluenza (CAF), rendendo il confronto tra risultati non sempre affidabile. Nonostante questi limiti, poiché gli studi condotti con misure che rilevano queste tre dimensioni della produzione sembrano promettenti e utili, questo studio cerca di offrire un contributo in questo senso: misurare CAF nelle produzioni elicitate per mezzo di task è utile perché può aiutare a definire un modello di proficiency nella seconda lingua con dirette applicazioni in campo didattico (cfr. Ellis, 2003) e nel testing (cfr. Iwashita *et al.*, 2001).

### 1.2.1.3. L'approccio emergentista

L'approccio emergentista riconsidera in modo sostanziale il ruolo della variazione rispetto agli approcci visti in precedenza. Se questi ultimi vedevano la comunicazione come un sistema caratterizzato da una sequenza lineare di eventi, che permette lo scambio di informazioni tra mittente/parlante e ricevente con un meccanismo simile a quello di una macchina per fax, dove i due partner della conversazione si alternano nell'emettere e processare il messaggio, l'approccio emergentista si rifà ai sistemi dinamici e descrive la comunicazione con la metafora della danza. A partire da semplici procedure realizzate in coppie coordinate, come per i passi nella danza, pattern complessi emergono nel tempo, con pattern ancora più complessi che si realizzano quando coppie di ballerini interagiscono con altre coppie (cfr. de Bot *et al.*, 2007: 9). Se l'approccio tradizionale è legato a termini quali segnale e risposta, inviare e ricevere, codificare e decodificare o a comportamenti basati su regole, l'approccio emergentista usa termini quali *engagement* e *disengagement*, sincronia e asincronia, *breakdown* e *repair* nell'interazione per descrivere i processi di comunicazione.

Lo studio dei sistemi dinamici, ossia sistemi complessi e non lineari in costante mutamento, è nato nell'ambito della matematica e delle scienze con la *Dynamic System Theory* (d'ora in poi DST) e la *Chaos/Complexity Theory* (d'ora in poi C/CT). Negli ultimi decenni però ha avuto diverse applicazioni anche nelle discipline umanistiche (cfr. van Geert, 1991; Thelen, Smith, 1994), inclusa la linguistica acquisizionale: con la C/CT proposta da Larsen–Freeman (1997; per una presentazione della letteratura, Larsen–Freeman, Cameron, 2008; Byrne, Callaghan, 2014), la DST illustrata da de Bot *et al.* (2007; cfr. Herdina, Jessner, 2002; Verspoor *et al.*, 2011), e le teorie emergentiste *Emergence of Language* introdotte da El-

lis (1996; 1998) e MacWhinney (1998; 1999; 2001; cfr. anche MacWhinney, O'Grady, 2015).

Sulla base di possibili somiglianze tra sistemi dinamici e apprendimento linguistico, Diane Larsen–Freeman con l'articolo *Chaos/Complexity science and second language acquisition* del 1997 ha proposto l'applicazione dell'approccio emergentista ai processi di apprendimento di una seconda lingua, evidenziando come questo approccio possa essere un utile strumento per descrivere in modo più realistico il processo di apprendimento, includendo aspetti sia cognitivi che sociali. Dal 1997 ad oggi diversi studi sono stati condotti in questa direzione e alcuni autori hanno proposto la C/CT come possibile teoria generale dello sviluppo linguistico (cfr. Larsen–Freeman, 2002; *Applied Linguistics, Special Issue 27*, 2006; *Developmental Review, Special Issue 25*, 2005; *Bilingualism: Language and Cognition, Special Issue 10*, 2007; *Language Learning, Special Issue 59*, 2009 Larsen–Freeman, Cameron, 2008; Verspoor *et al.*, 2011; Larsen–Freeman, 2011). In questa sezione presento le caratteristiche principali dell'approccio emergentista focalizzando l'attenzione in particolare su variazione e andamento dei percorsi di sviluppo.

Un sistema dinamico è caratterizzato da alcune proprietà distintive:

- a) completa interconnessione dei sotto–sistemi, tutte le variabili sono legate l'una all'altra, pertanto il cambiamento di una di esse ha un impatto su tutte le altre. Un esempio il cosiddetto *butterfly effect*, dipendenza dallo stato iniziale; lo sviluppo dei sistemi dinamici è strettamente legato allo stato iniziale del sistema, pertanto piccole differenze iniziali possono avere conseguenze molto evidenti sul lungo periodo. Si generano così traiettorie di sviluppo non facilmente prevedibili. L'intera-



- zione tra variabili porta a cambiamenti che non possono essere anticipati con gli strumenti analitici classici;
- b) emersione di stati di attrazione (*attractor states*) o stati di repulsione (*repeller states*), poiché i sistemi dinamici sono sempre parte di altri sistemi più ampi, i diversi sotto-sistemi si regolano su determinati stati, per definizione temporanei e non fissi, detti appunto stati di attrazione, preferiti ma non forzatamente prevedibili, o stati di repulsione, non preferiti;
  - c) evoluzione costante (*constantly changing*), i sistemi dinamici sono in costante mutamento, grazie all'interazione attraverso una continua riorganizzazione interna: tale condizione rende i sistemi dinamici per definizione variabili.

Riprendendo questo approccio, van Geert (1991) propone un modello di sviluppo della LI basato sulla DST. Perché lo sviluppo avvenga è necessario che una serie di condizioni siano rispettate. In primo luogo deve esserci qualcosa che può crescere, in questo senso van Geert parla di *minimal structural growth condition*; in secondo luogo, dovranno esserci risorse che sostengono il processo di crescita. In questo caso van Geert distingue tra risorse interne, legate all'individuo, quali capacità o tempo per apprendere o elementi quali conoscenze concettuali e motivazione, e risorse esterne, come ad esempio contesto di apprendimento o eventuali interventi didattici. Queste risorse hanno due caratteristiche fondamentali: sono limitate e sono interconnesse, da un lato dovranno essere posizionate su un valore minimo perché avvenga l'apprendimento, dall'altro, quando necessario, il sistema ha a disposizione meccanismi di compensazione che coinvolgono diversi tipi di risorse contemporaneamente. Van Geert (1994) chiama *connected growers* proprio quei sottosistemi che sostengono la reciproca crescita. È il caso

ad esempio della relazione tra sviluppo lessicale e sviluppo della capacità di comprensione: in genere conoscere più parole rende la comprensione del parlato più semplice. Poiché le risorse sono limitate, anche la crescita per definizione è limitata. Van Geert parla in questo caso di *carrying capacity*, riferendosi appunto allo stato di conoscenza che può essere raggiunto da un determinato bambino nello sviluppare strutture tra loro collegate. È il caso ad esempio dell'emersione delle frasi complesse: questa può determinare un rallentamento dello sviluppo lessicale. Se prima infatti tutte le risorse erano impiegate nello sviluppo del lessico, in questa seconda fase le risorse si redistribuiscono tra apprendimento lessicale e sviluppo del sistema grammaticale.

Larsen–Freeman (1997; 2002; 2006; 2007; 2009; 2011a;b), riprendendo il lavoro di van Geert (1991; 1994; 2005) e van Geert, van Dijk (2002), utilizza l'approccio emergentista come utile prospettiva per spiegare due aspetti chiave dell'apprendimento della L2: la possibilità di ottenere diversi *pattern* di sviluppo a livello individuale e l'interpretazione del ruolo della variazione. Catturare il dinamismo dell'interlingua è un tema ampiamente discusso nelle ricerche acquisizionali. Diversi fattori infatti intervengono nel determinare il percorso di sviluppo, alcuni di questi esterni all'individuo, come la lingua da apprendere, la marcatezza della L1, la quantità e tipologia di input, la quantità e il tipo di interazione, la quantità di feedback, l'eventuale insegnamento, altri interni all'individuo, come età, fattori socio-psicologici quali motivazione e attitudine, caratteristiche individuali, stile cognitivo, ecc. Nessuno di questi elementi preso singolarmente giustifica il progresso, mentre la loro interazione ha effetti notevoli in quanto «both individual creativity and social interaction combine to influence the shape of the developing grammar» (Larsen–Freeman, 1997; cfr. meta-analisi di Goldschneider, DeKeyser, 2001). La lingua cresce e si auto-organizza in modo organico dando origine a una

serie di strutture stabili, ma nel complesso poco è fisso e immutabile. Il termine stesso lingua target è ingannevole. Se la lingua è un sistema dinamico che si adatta continuamente al contesto, il target degli apprendenti è in costante movimento; non esiste una fine del percorso di acquisizione poiché non ci sono vincoli che definiscono a priori il punto di arrivo della traiettoria di sviluppo. L'apprendimento poi non procede linearmente, con gli apprendenti che interiorizzano completamente una struttura prima di passare alla successiva, ma piuttosto, così come tipico dei sistemi dinamici, «development is not discrete and stage-like but more like the waxing and waning of patterns; that, from a target-language perspective, certain aspects of the behaviour are progressive, others, regressive; that change can be gradual and it can be sudden» (Larsen-Freeman, 2006: 1; cfr. Thelen, Bates, 2003: 381). Non ci sono stadi discreti in cui l'interlingua non è variabile, nonostante possano esserci periodi in cui alcune strutture sono predominanti: «the dominance of certain structures may arise through a gradual building up process or through a period of fluctuation among competing forms, followed by a shift in the system when a certain critical threshold is crossed, and some wider reorganisation is triggered» (Larsen-Freeman, 2006: 592). La non linearità è data dal fatto che i diversi sotto-sistemi sono interconnessi: in alcuni casi supportano uno lo sviluppo dell'altro, in altri entrano in competizione. Questo accade perché le risorse nei sistemi dinamici sono per definizione limitate, di conseguenza una miglior performance in una dimensione di proficiency può sottrarre risorse alle altre dimensioni.

Secondo l'approccio emergentista, la variazione nella performance in quanto proprietà intrinseca dei sistemi dinamici di un apprendente è riscontrabile sia a livello sincronico che diacronico: non è un errore di misurazione o un rumore da eliminare, quanto piuttosto il risultato della capacità di sviluppo

e di adattamento del sistema: essa è sia origine dello sviluppo che indicatore di un momento specifico del processo di sviluppo stesso. La psicolinguistica, la sociolinguistica, la linguistica dei corpora, i sistemi di analisi probabilistica (*probabilistic analyses*) giustificano tale instabilità come sistematica e governata da regole (per es. Labov, 1970; cfr. § 1.2.1.2.). L'approccio emergentista interpreta la sistematicità della variazione in modo dinamico rafforzando il senso dell'espressione "competenza variabile":

Competence and performance both emerge from the dynamic system that is the frequency-tuned conspiracy of memorized exemplars of construction usage, with competence being the integrated sum of prior usage and performance being its dynamic contextualized activation. (Ellis, Larsen-Freeman, 2006; cfr. Ellis, 2006a; 2006b; 2006c).

I percorsi individuali di sviluppo possono essere molto diversi l'uno dall'altro, anche se visti nell'insieme possono apparire piuttosto simili: gli apprendenti differiscono nelle loro abilità d'uso della lingua non perché hanno o non hanno una conoscenza specifica; piuttosto, le differenze sono il risultato della variazione nei modi e nei tempi in cui il processo è stato realizzato durante l'apprendimento (cfr. Larsen-Freeman, 2006; de Bot *et al.*, 2007). Larsen-Freeman (2006) ad esempio esaminando i dati di cinque cinesi apprendenti di inglese L2 rileva proprio come vi sia un'evidente variazione a livello individuale nei percorsi di sviluppo. L'analisi qualitativa del suo corpus evidenzia infatti come l'apprendimento proceda come in una spirale, con forme complesse che emergono, vengono abbandonate e poi riprese con maggior accuratezza in una fase successiva. Utilizzare un approccio emergentista richiede dunque di concentrarsi non tanto o non solo sulle tendenze

generali a partire da dati aggregati a livello di gruppo, ma anche sull'analisi di diversi casi individuali osservati longitudinalmente, così da ricostruire cosa accade realmente durante il processo di apprendimento. Dati di gruppi e dati individuali risultano così complementari:

Individual data provide us with a window into the complex world of interacting *vincolis* in the learner. Group data, on the other hand, provide us with a window toward society and the distribution of the linguistic input (MacWhinney, 2006: 738).

Dal punto di vista metodologico, un approccio di questo tipo pone diverse sfide alle ricerche future. Come suggeriscono gli stessi Ellis e Larsen–Freeman (2006) uno studio appropriato dei costrutti linguaggio e apprendimento linguistico richiede la sinergia di diversi approcci metodologici:

Each research methodology has its advantages. Ethnography brings together the individual, society, and consciousness in time and place, but it ignores implicit motivations that introspection cannot access (Nisbett, Wilson, 1977). Brain imaging illustrates the dynamic patterns of neural activity involved in mental processing, but it isolates the learner's brain from society and its normal ecology of function. Laboratory experimentation allows the controlled logic of the scientific method, but it sacrifices ecological validity in this goal. And so on. We need to bring these methods together, to strive after the linking relations, to develop the kind of account that Wertsch, a Socioculturalist, called "translation at the crossroads that would make it possible to link, but not reduce, one perspective to another" (Wertsch, 1998) (Ellis, Larsen–Freeman, 2006: 577–78)

Lo studio dell'acquisizione secondo l'approccio emergentista richiede l'impiego di specifici strumenti: sarà necessario

predisporre corpora longitudinali rappresentativi di una molteplicità di condizioni, in modo da osservare i fenomeni di sviluppo in un contesto più ampio (cfr. van Geert, Steenbek, 2005), e selezionare strumenti di misurazione capaci sia di catturare l'intero processo di sviluppo (cfr. van Geert, van Dijk, 2002; Larsen–Freeman, 2006; Wang 2009; MacIntyre, Legatto, 2011) sia di essere sensibili alle variazioni individuali (cfr. Ellis, Larsen–Freeman, 2006).

L'emergentismo rappresenta un nuovo approccio nello studio del linguaggio e certamente offre una serie di metafore utili a descrivere la complessità del processo di apprendimento linguistico, metodi per fare predizioni o interpretare complesse interazioni dialogiche. Nonostante le potenzialità, alcuni ricercatori ritengono che i sostenitori dell'emergentismo siano anche troppo ottimisti (cfr. Horgan, 1995; Stephan, 1999; Cameron, Deignan, 2006; MacWhinney, 2006; Pienemann, 2007; Gregg, 2010) e mettono in luce alcune limitazioni da tenere in considerazione. Pienemann (2007: 43) ad esempio evidenzia come questa prospettiva da un lato non prenda in sufficiente considerazione le ricerche che l'hanno preceduta, dall'altro non definisca in modo chiaro cosa si intende per variazione: manca una definizione nuova e operativa di variazione o più in generale delle dinamiche dello sviluppo di una L2, capace di offrire una valida alternativa agli approcci consolidati usati in letteratura. MacWhinney (2006), pur sostenendo l'utilità dell'emergentismo, invita alla cautela, pone alcune questioni metodologiche e si chiede come l'emergentismo possa raggiungere un livello soddisfacente di applicazione empirica negli studi acquisizionali: è abbastanza semplice elaborare attraenti interpretazioni emergentiste, ma è altrettanto necessario verificare empiricamente che non siano errate. Una volta poi stabilito che l'apprendimento è un sistema dinamico in cui i diversi elementi sono emergenti sarà necessario definire cosa

emerge e soprattutto come. Anche MacWhinney (2006) sottolinea proprio come lo sviluppo futuro di questo approccio è strettamente legato alla creazione di adatti supporti metodologici, legati in particolare alla linguistica dei corpora e al potenziamento dei metodi di analisi computazionale. In ragione di questi limiti, non si adotterà una prospettiva emergentista in questo studio. Si è ritenuto comunque utile presentare questa parte di letteratura sia perché la più recente sul tema della variazione, sia per l'attenzione ai percorsi individuali, sia per l'utilità descrittiva di alcune delle metafore impiegate.

Ricapitolando, la letteratura acquisizionale ha affrontato il tema della variazione dell'interlingua a partire da diversi approcci: quelli di tipo sociolinguistico, che indagano principalmente l'ordine di emersione o la qualità e variazione d'uso di strutture specifiche in relazione a contesto linguistico e situazionale; quelli di tipo psicolinguistico, che tentano di descrivere la proficiency e lo sviluppo attraverso misure quantitative delle dimensioni della performance; e infine il più recente approccio emergentista, che inquadra la variazione entro le teorie dei sistemi complessi. Gli studi condotti hanno dunque contribuito in modo complementare a una maggior comprensione del ruolo e dell'importanza della variazione nell'apprendimento. Questo studio pertanto utilizza l'approccio psicolinguistico come ancoraggio teorico e metodologico principale, mentre l'approccio sociolinguistico motiva l'attenzione alla variazione situazionale e al confronto con il comportamento dei parlanti nativi, e infine l'approccio emergentista giustifica l'attenzione per la variazione individuale in relazione a quella di gruppo.

### 1.2.2. *La misurazione della variazione*

Nella sezione precedente (§ 1.2.1.2.) si è visto che ricerche come quelle di Skehan (1998; 2018) e Robinson (2003; 2011;

2015) dimostrano come l'apprendente possa avere obiettivi diversi nello svolgere un compito linguistico, favorendo pertanto di volta in volta una delle tre dimensioni CAF. L'utilità dell'analisi di queste tre dimensioni è motivata sulla base di modelli psicolinguistici (§ 1.1.2.) e sostenuta dai risultati ottenuti in vari studi (§ 1.2.1.1.). L'interesse per l'analisi CAF è condiviso anche dalle ricerche sugli indici di acquisizione (*developmental index studies*) interessate all'andamento longitudinale di queste tre dimensioni e delle loro relative misure o al legame con il livello di proficiency (Larsen–Freeman, 2009; Wang, 2009). Sin dalla fine degli anni '70 diversi studi hanno tentato di individuare una misura oggettiva dello sviluppo linguistico nella seconda lingua, «a developmental yardstick against which global (i.e. not skill or item specific) second language proficiency could be gauged» (Larsen–Freeman, 1983: 287). Misure di questo tipo sarebbero infatti utili a insegnanti, tester e ricercatori. Dal punto di vista dell'insegnamento permetterebbero di accertare gli effetti di un intervento educativo o di una tecnica didattica; dal punto di vista del testing costituirebbero un utile strumento per definire il profilo di competenza, e per validare prove e scale olistiche. Dal punto di vista della ricerca, renderebbero più affidabile il confronto tra studi diversi e tra lingue diverse.

Negli ultimi decenni la letteratura relativa alle misure quantitative CAF si è dimostrata piuttosto vivace. Ad esempio, con riferimento alla lingua scritta, Polio (1997; 2001) presenta criticamente un buon numero di misure impiegate in diversi studi, Wolfe–Quintero *et al.* (1998) verificano la *sampling reliability* e la *cuncurrent validity* di circa 100 misure CAF impiegate in diversi studi e Ortega (2003) nella sua meta-analisi approfondisce l'utilità di alcune di esse come misure di proficiency, con riferimento alla lingua orale. Freed (2000) offre un'interessante discussione sulle misure di fluenza, Ortega



(1999), Skehan (2003) e Robinson, Ellis (2008) discutono invece della maggiore utilità di misure globali o misure specifiche. Polio (2017), pur riconoscendo il valore ancora attuale del lavoro di Wolfe–Quintero *et al.* (1998) evidenzia come studi più recenti (per es. Manchon, 2012; Sasaki, 2004; Nicolás–Conesa *et al.*, 2014) espandano il focus rispetto al processo di scrittura, rendendo il costrutto più complesso da catturare con misure CAF (vedi anche Pallotti, 2009 rispetto al costrutto complessità). Kormos (2011), Bulté, Housen (2012; 2014), Verspoor *et al.*, (2012) e Connor–Linton, Polio (2014), con particolare riferimento alla scrittura, approfondiscono la relazione tra sviluppo linguistico e misure CAF.

Pubblificazioni di questo tipo offrono indicazioni utili sulle misure più affidabili, ma evidenziano anche alcuni limiti metodologici nelle ricerche condotte (e di conseguenza nelle rispettive analisi cumulative), in particolare in riferimento alle modalità di codifica dell'interlingua, all'uso di definizioni non univoche delle misure impiegate e alla capacità di alcune delle misure impiegate di rilevare il costrutto oggetto di indagine (Pallotti, 2009). Poiché le misure CAF dovrebbero aiutare il ricercatore a osservare con coerenza e affidabilità i propri dati, fornire una base teorica per l'interpretazione dei valori ottenuti e consentire la replica affidabile di uno studio, in letteratura è fortemente sentito il bisogno di migliorare comparabilità, affidabilità e validità delle misure stesse (Housen *et al.*, 2012). In particolare, due sembrano essere gli aspetti problematici su cui è necessario riflettere: da un lato complessità, accuratezza e fluenza sono costrutti complessi, di conseguenza è necessario che le misure impiegate siano in grado di dar conto delle diverse sotto-componenti delle tre dimensioni, dall'altro sono per natura legate allo sviluppo, se quest'ultimo è un sistema dinamico fatto di diversi fenomeni che interagiscono tra di loro in modi non prevedibili, le

misure impiegate dovranno riuscire a riflettere anche questo aspetto (cfr. Pallotti, 2009; 2015; Bulté, Housen, 2012; 2014; Polio, Shea, 2014).

Nonostante le debolezze evidenziate, autori come Wolfe-Quintero *et al.* (1998), Housen *et al.* (2009; 2012; 2019), Polio (2017) sottolineano come proprio le potenzialità di tali misure giustifichino l'interesse per ulteriori indagini in questa direzione e identificano alcuni obiettivi per la ricerca sulle misure CAF: la determinazione della migliore unità di produzione tra clausola o frase complessa, la determinazione della linearità o non-linearità nell'andamento delle misure, il rapporto tra le singole misure CAF e i relativi costrutti, il rapporto tra misure e livelli di competenza, e tra misure, sequenze di acquisizione e sviluppo interlinguistico. Il presente studio cerca di offrire un contributo in questa direzione. Le sezioni che seguono illustrano le unità utilizzate in letteratura per segmentare i testi parlati e per quantificare CAF nelle produzioni.

#### 1.2.2.1. L'unità di produzione

Un'unità di segmentazione del parlato è uno strumento essenziale se l'intento è analizzare aspetti quantitativi della competenza linguistica quali complessità, accuratezza e fluenza, o studiare elementi che caratterizzano il profilo di un apprendente, l'evoluzione nell'apprendimento o la variabilità situazionale dell'interlingua. Se tra gli obiettivi rientrano poi la comparabilità tra studi e tra lingue, o la meta-analisi, allora diventa fondamentale concordare sulla natura dell'unità, sulla sua applicabilità a diversi tipi di produzione nonché accertarsi che sia sufficientemente valida, affidabile e praticabile (Bachman, Palmer, 2010). Per quel che riguarda la prima qualità, si considera valida un'unità quando rileva ciò che intende misurare. In altre parole la validità è intrinseca alla definizione

stessa della misura e del suo costrutto soggiacente. Nel caso dell'unità di produzione del linguaggio orale, come già evidenziava Crookes (1990) in uno dei primi lavori metodologici su questo tema, la validità è legata alla capacità dell'unità di riflettere i processi di pianificazione psicolinguistica. Rispetto all'affidabilità invece, tale qualità è attribuibile a una misura quando questa può essere riapplicata in maniera consistente da uno stesso valutatore in diversi momenti o da diversi valutatori (*inter/intra-rater reliability; inter/intra-rater agreement*). Infine la praticabilità è legata alla possibilità di impiegare l'unità con una certa varietà di dati, senza un eccessivo dispendio di tempo o fatica.

Per l'analisi del parlato la letteratura mette a disposizione una serie di unità di produzione, raggruppabili in quattro macrocategorie: unità semantiche, unità tonali, unità interazionali (turni) e unità sintattiche (per rassegne sull'inglese si vedano Crookes, 1990; Foster et al., 2000; Ellis, Barkhuizen, 2005; per l'italiano L1, Cresti, 2005a; 2005b; 2010; 2012; Moniglia, 2006 e Voghera, 1992a; 1992b; 2006; 2017; per l'italiano L2, Chini, 1998, 2002; 2003). Nelle sezioni che seguono si discuteranno le principali caratteristiche, potenzialità ed eventuali limiti delle unità proposte. La tabella 1.4. offre un quadro complessivo.

Il primo tipo di unità che si considera sono le unità di tipo semantico, basate principalmente su una segmentazione informativa delle produzioni. Tra queste ritroviamo: la *proposition* (Sato, 1988: 375), definita come unità costituita da un argomento principale e una o più predicazioni su questo argomento; la *C-unit* (Brock, 1986: 52; Pica et al., 1989: 72), formata di enunciati, proposizioni, parole, clausole o frasi, grammaticali e non grammaticali, che forniscono un significato pragmatico; l'*Idea unit* con focus semantico, descritta come estratto di informazione percepito come coeso dal parlante/

**Tabella 1.4.** L'unità di produzione.

unità semantiche	
proposizione	unità semantica che consiste di almeno un argomento e uno o più predicati (Sato, 1988: 375)
<i>C-unit</i>	unità formata di enunciati, proposizioni, parole, clausole o frasi, grammaticali e non grammaticali, che forniscono un significato pragmatico (Brock, 1986: 52; Pica <i>et al.</i> , 1989: 72)
<i>Idea unit</i> (con focus semantico)	estratto di informazione percepito come coeso dal parlante/scrivente (Kroll, 1977); unità informativa espressa con una clausola, una frase o una porzione di testo più ampia (Malik, 1990)
unità tonali	
<i>tone unit</i>	configurazione distinta di pitches con un chiaro centro e nucleo. Il nucleo è la sillaba (o in alcuni casi una serie di sillabe) che ha la maggior preminenza nell'unità tonale (Crystal, Davy, 1975: 16)
<i>Idea unit</i> (con focus intonativo)	unità del processing mentale e linguistico, contenuta in un unico profilo intonativo, delimitate da pause, in genere espressa da un'unica clausola formata da un predicato verbale e altri elementi ad esso associati (Chafe, 1980; 1994)
<i>utterance</i>	una stringa di parlato con almeno una delle seguenti caratteristiche: (a) sotto uno stesso profilo intonativo; (b) circondata da pause; (c) contenente un'unica unità semantica, e.g. Crookes (1990: 187)
enunciato	ogni espressione linguistica interpretabile pragmaticamente che esprime un'ilocuzione legata a una condizione semantica di piena significanza dell'espressione in questione e alla sua realizzazione intonata secondo un pattern melodico di valore illocutivo (Cresti, 2005a; 2005b; 2010)
unità interazionali	
turno	misura applicabile a situazioni dialogiche: "stream of speech bounded by speech of another, usually an interlocutor" (cfr. Sacks <i>et al.</i> , 1974; Ford <i>et al.</i> , 1996; 2002; Selting, 1996)
unità sintattiche	
clausola (o <i>s-node</i> )	"indicated by tensed or untensed verbs" (Ellis <i>et al.</i> , 1994: 483); fondata sulla struttura argomentale del verbo
<i>T-unit</i>	<i>minimal terminal unit</i> , formata dalla principale e dalle subordinate a essa collegate (Hunt, 1970)

<i>c-unit</i>	predicazioni grammaticali indipendenti o risposte cui manca la ripetizione di un elemento della domanda per soddisfare il criterio di indipendenza (Loban, 1966, in Ellis e Barkhuizen, 2005), proposizioni, parole, sintagmi o clausole, grammaticali e agrammaticali, che hanno un significato referenziale o pragmatico (Pica <i>et al.</i> , 1989: 72)
<i>AS-unit</i>	proposizione prodotta da un solo parlante fatta di una clausola indipendente o di una unità indipendente inferiore alla clausola e tutte le clausole subordinate a essa associate (Foster <i>et al.</i> , 2000)

---

scrivente, con una forma superficiale collegata alla realtà psicologica del ricevente (Kroll, 1977: 85) o come singola unità informativa espressa con una clausola, una frase o un'unità di testo più ampia (Malik, 1990: 229).

Se a un primo sguardo la segmentazione informativa del parlato può apparire intuitiva, nella pratica non è facile riconoscere il confine di un'idea o di un argomento, in particolare nel parlato interattivo e non pianificato, o caratterizzato da un impiego significativo di espressioni di routine. Il fatto poi che le unità semantiche vengano brevemente definite, ma non esemplificate, rende la loro applicazione tendenzialmente poco affidabile (Kobayashi, 2009: 28). Nel tentativo di rendere le definizioni meno ambigue, alcuni ricercatori hanno integrato le descrizioni proposte con criteri sintattici, generando così unità spurie: è il caso, ad esempio, della *Idea unit* ridefinita in Johns (1985: 500) e Johns e Mayes (1990: 85) con criteri sintattici che tentano di mettere in relazione un'unità principalmente comunicativa con elementi quali frase, sintagma, clausola dipendente o clausola ridotta. Per tutti questi motivi, le definizioni semantiche risultano problematiche per stabilire le unità fondamentali di segmentazione del parlato, e possono essere pertanto considerate rilevanti solo per eventuali indagini che abbiamo come oggetto il contenuto dei testi.

Il secondo tipo di unità proposto riguarda le unità tonali, impiegate più frequentemente in ricerche dedicate alla L1.

Esse hanno come elementi chiave l'intonazione e le pause e si basano fondamentalmente sull'assunto che la prosodia sia uno degli strumenti con cui il parlante può delimitare il dominio di progettazione sintattica, discorsiva o informativa (Halliday, 1985; Chafe, 1988; Voghera, 1992a; 2017; Cresti, Moneglia, 2005). Tra quelle maggiormente impiegate in letteratura possiamo identificarne tre: *tone unit*, *Idea unit* (con focus intonativo), e unità basate sull'enunciato.

La *tone unit* è definita come una configurazione distinta di accenti tonali con un chiaro centro o nucleo. Il nucleo è la sillaba (o in alcuni casi una serie di sillabe) che ha la maggior preminenza nell'unità tonale. In genere ai confini dell'unità si ritrovano brevi pause, variazioni in lunghezza o aspirazioni (Crystal, Davy, 1975: 16). Nelle applicazioni dell'unità possiamo ritrovare alcuni studi che ne valorizzano l'aspetto più strutturale e si basano dunque anche su indizi sintattici per la sua identificazione, altri invece che ne enfatizzano l'aspetto funzionale, dando maggior spazio a un'interpretazione semantica. Queste due prospettive portano però a interpretazioni sostanzialmente diverse per unità che portano lo stesso nome (Crookes, 1990: 186).

L'*Idea unit* con focus intonativo, proposta da Chafe (1988: 13-14; 1994), è descritta invece come unità contenuta in un unico profilo intonativo, delimitata da pause, in genere espressa da un'unica clausola formata da un predicato verbale e da altri elementi ad esso associati. Chafe (1994) in studi successivi sviluppa ulteriormente la definizione di *Idea unit* descrivendola come "unità del processing mentale e linguistico". Basandosi su criteri in parte intonativi e in parte semantici, distingue tra unità intonative che trasmettono un significato e unità intonative che regolano la comunicazione.

Infine vi sono le unità basate sull'enunciato. Manteniamo qui il termine inglese *utterance* per riferirci all'unità proposta

per l'inglese da Crookes, così da distinguerla dall'*enunciato*, come definito per l'italiano da Cresti. L'*utterance* come unità è stata inizialmente sviluppata per indagini relative all'acquisizione dell'inglese L1 (vedi ad esempio Brown, Bellugi-Klima, 1964; Scollon, 1976), in particolare la misura ad essa associata, MLU *mean length of utterance* (Brown, 1973), è generalmente accettata come indice capace di rilevare lo sviluppo morfosintattico del linguaggio infantile. Crookes (1990: 187) nel proporla per l'indagine del parlato in L2 definisce l'*utterance* come porzione di parlato con almeno una delle seguenti caratteristiche: contenuta entro un unico profilo intonativo, delimitata da pause e costituita di un'unica unità semantica. Per l'italiano Cresti (2005a;b; 2010; 2018; cfr. Cresti, Gramigni, 2004; Cresti, Moneglia, 2005) propone come unità di produzione l'*enunciato*, definito come ogni espressione linguistica interpretabile pragmaticamente poiché esprime un'ilocuzione, legata: a) a una condizione semantica di piena significanza dell'espressione in questione; b) alla sua realizzazione intonata secondo un pattern melodico di valore illocutivo. L'*enunciato* come unità è impiegato in indagini relative all'italiano parlato in L1 e per quanto ci risulta non è stato applicato a indagini relative alla L2.

Nonostante la rilevanza per il parlato di unità di questo tipo, in particolare in relazione alla loro capacità di riflettere i processi di pianificazione, il loro impiego per la ricerca acquisizionale pone alcuni problemi. Nelle produzioni interlinguistiche, in misura maggiore rispetto a quanto accade per quelle di nativi, l'intonazione può essere influenzata fortemente dalla lingua madre, risultando a volte di difficile interpretazione, mentre le pause sono frequentemente determinate dalla necessità di attivare il lessico utile per trasmettere un determinato significato e non solo impiegate per indicare il termine di un'unità (Foster *et al.*, 2000: 359).

Un terzo tipo di unità è costituito da quelle che si fondano sull'osservazione dei comportamenti interazionali, e in particolare sui meccanismi di alternanza dei turni tra parlanti (Sacks *et al.*, 1974; Fele, 2007). Il turno è in genere definito come porzione di parlato delimitata dal parlato di un altro partecipante ed è considerata l'unità di base di organizzazione dell'interazione, costituita appunto dall'alternarsi delle voci degli interlocutori (Schegloff, 1996; 2000; 2011). Nonostante sia un'unità chiave quando l'obiettivo è indagare le dinamiche interattive, ha il limite di essere applicabile esclusivamente a dati dialogici.

Il quarto tipo di unità si riferisce alle unità sintattiche. In generale queste misure riportano un buon livello di affidabilità, perché nella maggior parte dei casi i criteri per la loro identificazione sono più oggettivi ed espliciti di quelli relativi ad altre unità. Esse possono essere inoltre considerate valide su un piano psicolinguistico, in quanto riflettono i processi di pianificazione del parlato (Haiman, Thompson, 1988; Crookes, 1990; Foster *et al.*, 2000; Voghera *et al.*, 2004; Voghera, 2017). Gli studi empirici hanno ampiamente mostrato infatti come nella produzione entrino in gioco sia processi di macro-pianificazione sia processi di micro-pianificazione, che si realizzano attraverso unità centrate intorno alla clausola (Beattie, 1980; Halliday, 1985; Chafe, 1988; Biber *et al.*, 1998; Rastelli, 2013; 2014; Skehan, 2018; Ellis, Wulff, 2019).

Le unità sintattiche proposte in letteratura sono le seguenti: la clausola (o *s-node*), costituita da "tensed or untensed verbs" e fondata sostanzialmente sulla struttura argomentale del verbo (Ellis *et al.*, 1994: 483); la *T-unit* (*Terminal unit*, Hunt, 1965: 37; 1970), formata dalla principale e dalle dipendenti ad essa collegate; la *c-unit* (*communication unit*), definita come predicazione grammaticale indipendente o, nei dialoghi, come risposta a cui manca la ripetizione di un elemento della domanda per soddisfare il criterio di indipendenza (Loban, 1966, in Ellis, Bar-



khuizen, 2005); l'AS-unit (*Analysis of Speech unit*), proposizione prodotta da un solo parlante composta da una clausola indipendente, o da una unità indipendente inferiore alla clausola, e tutte le clausole dipendenti a essa associate (Foster *et al.*, 2000).

La distinzione principale nella selezione di un'unità sintattica è tra quelle basate sulla clausola e altre più ampie della clausola. Se l'analisi basata sulla clausola è spesso meno problematica, le unità più ampie offrono una maggiore validità. Esse permettono infatti di rendere conto delle porzioni di parlato caratterizzate da processi di pianificazione più complessi (Foster *et al.*, 2000: 360). Sia l'AS-unit che la T-unit riguardano costruzioni di livello superiore alla clausola, e presentano diversi elementi in comune.

La T-unit come unità è stata inizialmente concepita per l'analisi delle produzioni scritte. La sua applicazione alla segmentazione sintattica del parlato pone alcune questioni teoriche e metodologiche. La definizione tradizionale di clausola come espressione di senso compiuto contenente una predicazione, alla base di questa misura, non soddisfa infatti le richieste di analisi del parlato, dove, oltre alle strutture sintattiche più o meno complesse, hanno un ruolo rilevante elementi sintattici con scarsa struttura (Voghera, 1992b; Giordano, Voghera, 2009). Consideriamo a titolo esemplificativo un estratto dal corpus prodotto da una parlante nativa.

- (1) 1 INV: volevo semplicemente sentire:: da te [un un racconto+  
 2 ELI: [comunque mh:: #o\_2 carino  
 3 INV: +degli eventi che succedono  
 4 ELI: carino. cioè # bello. che poi si rincontrano.

In questo esempio vi sono almeno tre diverse strutture sintattiche che mettono in crisi l'applicazione della definizione tradizionale di clausola, che sta alla base di quella di *T-unit*: (a)

clausole senza verbo, o con verbo sottinteso o presupposto (1.2; 1.4); (b) unità sintattiche distribuite su più turni (1.1-1.3); (c) clausole dai confini incerti (1.4). Poiché nella definizione standard di *T-unit* non si prevede come trattare questi elementi, la loro interpretazione rischia di essere arbitraria, compromettendo così l'affidabilità dell'analisi. Per questo motivo alcuni ricercatori hanno addirittura suggerito di eliminarli dall'analisi (per es. Biber *et al.*, 1998). Un approccio di questo tipo però limiterebbe la validità dell'unità quando applicata al parlato, poiché si rischierebbe di escludere dall'indagine circa il 40% delle locuzioni (Scarano, 2003; Moneglia, 2006). La proposta avanzata da Loban (1966) di integrare l'analisi in *T-unit* con la *c-unit*, oltre a richiedere al ricercatore di combinare due unità distinte, ha il rischio di sostenere un'interpretazione dei fenomeni tipici della sintassi parlata come strutture "secondarie" o mancanti di qualche elemento, analizzabili esclusivamente in quanto "ricostruibili come frasi". In questo senso l'*AS-unit*, proprio perché pensata specificamente per il parlato, compreso quello in L2, permette di superare i limiti a cui si è accennato sopra, offrendo uno strumento psicolinguisticamente valido in grado di dar conto di tutte le combinazioni di parole che rappresentano unità di elaborazione e produzione del linguaggio orale. La proposta di Foster *et al.* (2000) ha avuto una notevole risonanza nelle ricerche sulla seconda lingua ed è attualmente considerata una delle principali unità di analisi del parlato (Ellis, Barkhuizen, 2005; Loewen, Reinders, 2011). Per queste motivazioni teoriche si è scelta l'*AS-unit* come unità di produzione in questo studio.

Al di là degli aspetti più strettamente teorici sopra discussi, molte delle unità presentate evidenziano criticità anche dal punto di vista metodologico, in quanto in genere manca un'attenzione per la loro validazione (Foster *et al.*, 2000; Lu, 2011). Almeno tre elementi limitano infatti la loro applicazione a nuove ricerche:

- a) la definizione: unità uguali vengono chiamate in modo diverso nei vari studi, spesso con definizioni imprecise, ma anche quando le descrizioni sono precise, di solito sono troppo semplici per rispondere alle esigenze dei dati reali;
- b) gli esempi: se esemplificate, le definizioni sono accompagnate da uno o due esempi soltanto, e in genere poco rappresentativi del parlato, in particolare di quello interattivo;
- c) l'affidabilità: raramente le ricerche riportano un'analisi dell'affidabilità dell'applicazione della misura.

Tali limiti rendono non solo poco affidabile la riapplicazione dell'unità in questione, ma anche incerta la comparabilità tra studi o le meta-analisi (Wolfe-Quintero *et al.*, 1998; Ortega, 2003; Ellis, Barkhuizen, 2005; Norris, Ortega, 2006). Le incertezze definitorie fanno sì che, talvolta, nell'applicazione pratica ci si trovi a prendere decisioni anche arbitrarie, senza avere la possibilità, principalmente per ragioni di spazio, di argomentare le proprie scelte. Per migliorare complessivamente la qualità metodologica degli studi a favorire il confronto tra ricerche diverse, sarebbe invece importante impiegare non solo unità teoricamente valide, ma anche chiaramente definite.

Nel tentativo di risolvere almeno in parte questi limiti metodologici connessi all'affidabilità, Foster *et al.* (2000) offrono una descrizione molto dettagliata dell'AS-unit. Nonostante ciò la sua applicazione a dati monologici e interattivi in interlingua e in italiano ha comunque richiesto una parziale ridefinizione di alcuni aspetti. Per la discussione sulla sua applicazione si rimanda a § 2.4 e a Ferrari (in stampa).

#### 1.2.2.2. Le misure analitiche

Si è già visto nelle sezioni precedenti come complessità, accuratezza e fluenza rappresentino diverse dimensioni della perfor-

mance. La complessità è vista in genere come capacità di produrre strutture varie e complesse, l'accuratezza è data dall'assenza di errori nelle produzioni e infine la fluenza è legata alla velocità. Quando osserviamo complessità e accuratezza osserviamo gli effetti di *representation* e *restructuring*, quando osserviamo la fluenza vediamo l'automaticità delle produzioni (§ 1.1.2.).

La misurazione CAF dell'uso linguistico dovrebbe dunque permettere riflessioni sui processi di apprendimento, nonché aiutare a distinguere studenti a diversi livelli di proficiency (per es. Housen *et al.* 2012; Lambert, Kormos, 2014; De Jong *et al.*, 2012). Larsen–Freeman (1978: 440) in un articolo ormai di diversi anni fa sottolinea come una misura ideale per l'abilità linguistica dovrebbe «increase uniformly and linearly as learners proceed towards full acquisition of a target language». Tale preferenza sembra giustificata dalla pratica didattica e dal testing: in genere le scale di proficiency sono lineari. Come riconoscerà la stessa Larsen–Freeman (1997), l'attesa che il ritmo di sviluppo sia uniforme in uno stesso individuo e nel gruppo non è stata giustificata dalle ricerche condotte da allora ad oggi (De Bot *et al.*, 2007; De Bot, 2008). Ad esempio il ritmo varia sostanzialmente da un individuo all'altro (per es. Perdue, Klein, 1992; Skehan, 1989; Slavoff, Johnson, 1995; Larsen–Freeman, 2006) e il legame tra lo sviluppo di una proprietà *x* e il conseguente sviluppo di una proprietà *y*, tipico dell'apprendimento della L1, non è sempre riscontrabile nell'acquisizione della L2 (per es. Clahsen, Muysken, 1986; 1989; Meisel, 1997; Tonkyn, 2007). Ne consegue che la richiesta di un indice che cresca linearmente non è né necessaria né tantomeno giustificata. Un obiettivo più realistico sarebbe identificare un gruppo di misure che, combinate insieme, produca un profilo dell'apprendente che può poi essere classificato in modo affidabile a un dato livello di una scala predeterminata (Larsen–Freeman, 2009; Norris, Ortega, 2009; Lambert, Kormos, 2014; Housen *et al.*, 2019).

Nel selezionare tale gruppo di misure, così come per l'unità di segmentazione del parlato, dobbiamo sottometerci alla validità, affidabilità e praticabilità. Wolfe–Quintero *et al.* (1998: 117–18) e Norris, Ortega (2003) hanno individuato tre fonti di informazioni che permettono di valutare la *construct validity* di una particolare misura di L2 proficiency: validità predittiva (*predictive validity*), *concurrent validity*, e *repeated sampling reliability*. Validità predittiva si otterrebbe se gli stadi di sviluppo di un particolare aspetto della grammatica avessero valori corrispondenti a quelli della misura in questione. Questa era stata ad esempio la motivazione iniziale che aveva spinto a considerare la lunghezza media dell'enunciato come misura dello sviluppo linguistico nei bambini (Brown, 1973). La *concurrent validity* si ottiene invece nei casi in cui la misura correla con differenze tra livelli di proficiency o con i livelli di proficiency o con entrambi. Infine, la *repeated sampling reliability* si otterrebbe quando una misura dimostra un andamento lineare tra i diversi livelli di proficiency. Wolfe–Quintero *et al.* (1998) notano come la validità predittiva delle misure non è stata indagata dai ricercatori di SLA e attraverso la loro meta-analisi si concentrano sulla valutazione di *concurrent validity* e *repeated sampling reliability*. L'affidabilità delle misure potrà invece essere ottenuta tramite la formazione dei valutatori e il double-marking. Quest'ultimo tipo di verifica sull'affidabilità è raramente applicato in letteratura. Infine la praticabilità è legata alla possibilità di impiegare la misura in questione senza un eccessivo impiego di tempo e fatica.

Wolfe–Quintero *et al.* (1998: 119) attraverso la loro meta-analisi di misure CAF indicano quali tra quelle utilizzate nella ricerca valga la pena implementare. Nelle sezione che segue si discutono le diverse misure CAF proposte dalla letteratura, in § 2.4.1. si presentano invece le misure selezionate per questo studio.

La prima dimensione è la complessità. Foster e Skehan (1996) definivano lo sviluppo della complessità come «progressi-

vely more elaborated language» e «a greater variety of syntactic patterning» (p. 303). Tale costrutto è rilevante in SLA poiché implica che l'apprendimento linguistico richieda, tra altri processi, la crescita del repertorio sintattico in L2 e l'abilità nell'utilizzarlo in modo appropriato nelle diverse situazioni comunicative. Inoltre le validazioni a posteriori dei test linguistici provano come la grammatica sia un fattore significativo nel differenziare tra punteggi o tra livelli di proficiency globale (cfr. Ellis, 2001). La complessità sintattica ha a che fare con la gamma di strutture sintattiche impiegate e riguarda la possibilità per l'apprendente di accedere rapidamente a strutture sia semplici che sofisticate. L'analisi della complessità sintattica dovrebbe quindi rilevare quanto varie o complesse sono le unità di produzione, o il lessico o determinate strutture grammaticali e discorsive: lunghezza dell'unità di produzione, quantità di subordinazione e gamma di tipi di strutture sono gli elementi presi in considerazione dalle diverse misure analitiche presenti in letteratura (per una discussione critica vedi anche Norris, Ortega, 2009; Gunnarson, 2012; Polio, 2017; Bulté, Housen, 2012; 2014; Housen *et al.* 2019).

Le misure oggettive della complessità sintattica devono il loro sviluppo in buona parte agli studi sulla L1, in particolare al lavoro di Hunt (1965; 1970), che occupandosi di produzioni scritte ha dimostrato come la lunghezza della frase (individuata sulla base della punteggiatura) non sia un indicatore sufficiente della proficiency linguistica, poiché non vi è sempre corrispondenza tra livello di proficiency e lunghezza della frase. Hunt propone quindi un'unità di produzione alternativa, la *minimal terminal unit*, T-unit (cfr. § 1.2.2.1.), definita come clausola principale e tutte le clausole a essa collegate. La lunghezza media delle T-unit (MLU), combinata con misure che calcolano l'incidenza di clausole subordinate e indipendenti, offre indicazioni affidabili sulla maturità sintattica delle produzioni scritte di apprendenti di L1.

Misure analoghe sono state proposte anche per la L2, dove la complessità sintattica viene misurata attraverso il calcolo della lunghezza media dell'unità di produzione (per es. AS-unit) o con misure di tipo grammaticale che rilevano il grado di subordinazione. Quest'ultime sono di due tipi: misure che analizzano la clausola, la T-unit o la AS-unit in rapporto l'una all'altra (per es. parole/AS-unit; clausole dipendenti/AS-unit) e misure che analizzano la distribuzione di una determinata struttura sintattica (per es. passivi/AS-unit) (cfr. Rimmer, 2006; Foster, Skehan, 1996). Wolfe-Quintero *et al.* (1998) attraverso la loro meta-analisi concludono come lunghezza media di T-unit o clausola, clausole per T-unit e clausole subordinate per T-unit o clausole correlino bene con gruppi di livello e valutazioni olistiche. La tabella 1.5. riassume le principali misure impiegate.

Quando si selezionano le misure da impiegare in uno studio è necessario stabilire di quante e quali di queste misure si ha bisogno per rilevare correttamente il costrutto complessità: non tutte le misure sopra presentate sono tra loro equivalenti, alcune misure potrebbero essere ridondanti se utilizzate contemporaneamente, altre potrebbero essere complementari, in quanto rilevano aspetti diversi del costrutto.

**Tabella 1.5.** Le misure CAF della complessità sintattica.

<i>Lunghezza</i>	lunghezza media della clausola, lunghezza media della T-unit/AS-unit. Misura proposta per la L1, frequentemente usata in SLA (per es. Brown, 1973; Hunt, 1965; Loban, 1976; Scott, 1988)
<i>Grado di subordinazione</i>	clausole dipendenti per T-unit/AS-unit, clausole dipendenti per clausole, clausole dipendenti per clausole complete (per es. Elder, Iwashita, 2005; Skehan, Foster, 2005; Michel <i>et al.</i> , 2007)
<i>Grado di coordinazione</i>	misura di coordinazione, misura suggerita da Bardovi-Harlig (1992) per i livelli elementari
<i>Frequenza d'uso di alcune strutture</i>	conteggio della frequenza di alcune strutture della morfologia verbale (per es. passivi), della classi di verbi (per es. modali) o di strutture sintattiche (per es. relative), cfr. Ellis, Yuan (2005), Robinson (2007b)

Si discutono ora più in dettaglio quali diversi aspetti della complessità sintattica dovrebbero essere colti dalle misure impiegate. È generalmente condiviso in letteratura che la complessità sintattica si sviluppi in due modi: un apprendente può realizzare clausole più complesse, aggiungendo sintagmi e parole, oppure può produrre unità più complesse, realizzando subordinate. Poiché misurare la complessità sintattica significa osservare questi due aspetti è necessario chiedersi che tipo di complessificazione viene rilevata dalle diverse misure. Considererò in particolare lunghezza media dell'unità di produzione e grado di subordinazione mettendo in evidenza la loro complementarità. Non si considereranno invece qui in dettaglio le misure legate alla rilevazione dell'incidenza di strutture morfosintattiche specifiche impiegate in alcune ricerche più recenti, poiché tali indici non sono stati impiegati nell'analisi dei dati di questo studio (per una discussione cfr. Berman, 2008; Ravid, Berman, 2010).

Le misure di lunghezza sono due: lunghezza media dell'unità di produzione e lunghezza media della clausola. La prima è ampiamente usata negli studi sullo sviluppo delle competenze di produzione scritta. Un aumento della lunghezza media della T-unit è generalmente considerato un incremento della complessità sintattica. La misura lunghezza media della T-unit (numero di parole per T-unit), come sottolinea lo stesso Hunt (1965), coincide con il numero di parole per clausola moltiplicato per il numero di clausole per T-unit, di conseguenza la misura parole per T-unit (o AS-unit) può aumentare sia come conseguenza della produzione di clausole più lunghe che di una maggior quantità di subordinate. Sulla base di queste osservazioni risulta chiaro come la misura della lunghezza media della AS-unit potrebbe essere una misura globale della complessità sintattica, ma non permette certo di distinguere la qualità della complessificazione. Consideriamo ora invece la misura



lunghezza media della clausola. Il suo incremento nel numero di parole è il risultato dell'aggiunta di modificatori o dell'uso di strutture complesse, come nel caso della nominalizzazione. Questa misura rileva dunque l'aumento della complessità a livello di clausola. Le due misure di lunghezza sono spesso usate in letteratura come equivalenti, ma la discussione appena riportata mette in luce come siano in realtà diverse (cfr. la discussione in Norris, Ortega, 2009). Il conteggio delle parole per AS-unit può essere considerato una misura globale, il conteggio delle parole per clausola rileva invece un aspetto specifico della complessificazione. La lunghezza non è unanimemente considerata un efficace indicatore della complessità sintattica. Ad esempio, Wolfe-Quintero *et al.* (1998: 25), ammettendo di esprimere una posizione minoritaria, preferiscono addirittura considerare la lunghezza una misura di fluenza piuttosto che di complessità. A questo proposito è bene sottolineare come le due diverse interpretazioni possano essere entrambe giustificate dalla definizione stessa di AS-unit e di clausola. Queste infatti sono legate, oltre che alla sintassi, a fenomeni intonativi e pause. In questo senso possono assumere un doppio valore, riflettendo sia aspetti della fluenza che della complessità.

Si discutono ora le misure che rilevano il grado di subordinazione. Queste misurano la complessità sintattica in termini di aumento dell'uso della subordinazione. In letteratura sono stati proposti due tipi di misure: il calcolo delle clausole subordinate per unità di produzione e il conteggio delle clausole subordinate sul totale delle clausole. Entrambe le misure rilevano il grado di subordinazione e sono usate alternativamente o insieme nei diversi studi. Sulla base di queste considerazioni, in § 3.4.1. si verificherà sperimentalmente l'ipotizzata equivalenza tra le due misure. Le misure di subordinazione si sono dimostrate in generale efficaci nel distinguere livelli di proficiency (cfr. Wolfe-Quintero *et al.*, 1998; Ellis, Barkhuizen, 2005). Wolfe-Quintero

*et al.* (1998) e Ortega (2003) nelle loro meta-analisi mettono comunque in luce alcune debolezze di queste misure e dei risultati che offrono. Da un punto di vista metodologico abbiamo già visto come il confronto tra i diversi studi sia limitato dal fatto che i diversi autori non sempre definiscono con precisione l'unità di produzione impiegata e non ne controllano validità e affidabilità. Dal punto di vista dell'andamento delle misure a diversi livelli di proficiency è stato osservato come complessità sintattica in termini di aumento del grado di subordinazione e proficiency potrebbero non sempre correlare in modo lineare.

L'andamento non lineare del processo di complessificazione sintattica è stato osservato anche nel caso della L1. Scott (1988) ad esempio, analizzando le produzioni scritte di alcuni alunni, osserva come il grado di subordinazione si stabilizzi entro l'ottavo anno scolastico, mentre la lunghezza dell'unità di produzione continua a crescere, come risultato della complessificazione a livello di sintagma. Negli studi sull'acquisizione della L2 tale andamento non lineare è stato giustificato a partire da due ipotesi: l'Ipotesi di Predizione dello Sviluppo (*Developmental Prediction*) e l'Ipotesi di Transfer Retorico dalla L1 (*Cross-rhetorical transfer*).

L'Ipotesi di Predizione dello Sviluppo, proposta da Cooper (1976), Monroe (1975), Sharma (1980), Halliday, Martin (1993), Halliday, Mattiessen (1999), ripresa da Wolfe-Quintero *et al.* (1998) e da Norris, Ortega (2009), sostiene una complessificazione non lineare della subordinazione, sulla base dell'assunto appunto che nella lingua scritta gli apprendenti avanzati tendano a complessificare a livello sintagmatico piuttosto che frasale. In altre parole, se nel passaggio dal livello elementare al livello intermedio gli apprendenti tendono ad ampliare l'uso della subordinazione, nel passaggio a livello avanzato si riduce l'uso di subordinate a favore di nominalizzazioni. Da un punto di vista sperimentale tale predizione sarà supportata se la ricerca

dimosterrà una relazione inversa tra grado di subordinazione e lunghezza delle unità di produzione ai livelli più alti. L'Ipotesi di Transfer Retorico dalla L1 è suggerita invece sulla base di uno studio condotto da Neff *et al.* (1998) in cui spagnoli apprendenti intermedio-avanzati d'inglese tendono ad avere valori di subordinazione sorprendentemente alti rispetto ai parlanti nativi dell'inglese. Ciò potrebbe dipendere da una maggiore tendenza a subordinare dello spagnolo scritto rispetto all'inglese, che verrebbe così trasferita alla seconda lingua. È probabile che entrambi i fattori, Ipotesi di Predizione dello Sviluppo e Ipotesi di Transfer Retorico dalla L1 possano giocare un ruolo importante nel determinare l'andamento dello sviluppo della complessità sintattica. A queste due ipotesi è possibile aggiungere un'ulteriore ragione per la non linearità rilevata dalle misure di subordinazione: l'influenza del genere e della modalità, scritto vs orale. Il tipo di discorso o di testo richiederà l'attivazione di strutture diverse, così come la lingua parlata e la lingua scritta si differenziano per incidenza di frasi complesse (cfr. Nippold, 2004; Berman, Nir-Sagiv, 2007; Berman, 2008; Biber *et al.*, 2011; Norris, Manchon, 2012). Se ciò fosse vero, ci si dovrebbe attendere una graduale diminuzione da parte degli apprendenti a livello avanzato delle misure di subordinazione, una volta apprese le strategie retoriche appropriate della seconda lingua. Ne consegue che sarebbe probabilmente inappropriato considerare le misure di complessità sintattica come misure assolute di sviluppo o indicatori diretti della competenza linguistica: un aumento o una diminuzione dei valori non corrisponde automaticamente a un cambiamento in positivo o in negativo, tutt'altro, una riduzione dei valori di subordinazione potrebbe essere comunque il risultato di un aumento della complessità globale della produzione. "Più complesso" può significare più sviluppato in modi diversi e soprattutto non significa forzatamente meglio (cfr. Pallotti, 2009; 2015). Il progresso nel-

la L2 può includere i due tipi di complessificazione sintattica, ma riguardare anche lo sviluppo di un repertorio discorsivo e sociolinguistico che il parlante può utilizzare in diverse situazioni comunicative. Se queste ipotesi sono corrette, il grado di subordinazione potrebbe essere una misura significativa ai livelli intermedi, mentre a livelli avanzati la lunghezza media della clausola potrebbe essere più rilevante. Si tratta di ipotesi suggestive, ma non ci risulta che nessuno studio abbia finora empiricamente dimostrato in modo sistematico questo tipo di flessione dei valori di subordinazione nei parlanti non nativi di livello più avanzato. Pochi studi hanno poi indagato longitudinalmente lo sviluppo della complessità sintattica nel parlato di apprendenti avanzati (per es. Polat, Kim, 2014; Spoelman, Verpoor, 2010; Vercellotti, 2017). Attraverso l'analisi condotta in questa ricerca si cercherà di contribuire alla verifica sperimentale di ipotesi di questo tipo.

Riassumendo, la discussione appena illustrata dimostra come la complessità sintattica sia un costrutto multidimensionale, per misurarla efficacemente sarà utile quindi combinare diverse misure. Si è visto come lunghezza media dell'unità di produzione può rappresentare una misura globale, mentre lunghezza media della clausola e grado di subordinazione offrono rispettivamente informazioni sulla complessificazione ottenuta attraverso l'elaborazione sintagmatica e quella dovuta all'aumento dell'uso della subordinazione, pertanto sono complementari. Le diverse misure del grado di subordinazione (cioè clausole subordinate/AS-unit e clausole subordinate/clausole) invece rilevano tutte la complessità in termini di aumento della subordinazione e si possono pertanto considerare tra loro equivalenti. Per l'applicazione di queste misure ai dati considerati per questa ricerca si veda § 3.4.

Passiamo ora alla seconda dimensione CAF, l'accuratezza. Foster e Skehan (1996) definiscono l'accuratezza come "free-

dom from error”, ossia come uso target delle strutture linguistiche in contesto. In letteratura sono state impiegate diverse misure di accuratezza. Un primo tipo è rappresentato dal conteggio dell’incidenza di unità senza errori sul totale delle unità, un secondo tipo dal conteggio del numero e del tipo di errori.

La percentuale di unità senza errori sul totale delle unità come misura è stata criticata da alcuni studiosi. Il conteggio delle unità senza errori, dando lo stesso peso a un’unità di produzione contenente uno o più errori, non rivela molto sulla loro distribuzione, non distinguendoli in categorie non permette osservazioni sulla loro tipologia (Bardovi-Harlig, Bofman, 1989; Foster, Wigglesworth, 2016). Wolfe-Quintero *et al.* (1998) nella loro meta-analisi della validità e affidabilità delle diverse misure impiegate in letteratura mostrano però come il rapporto *unità di produzione senza errori sul totale delle unità* prodotte risulti essere una misura che meglio di altre correla con i progressi di breve periodo e con i giudizi olistici a diversi livelli di competenza. Queste osservazioni suggeriscono come una misura basata sul conteggio delle unità corrette possa essere un indicatore globale utile, anche se non esaustivo.

Alcuni ricercatori hanno proposto un secondo tipo di misura, basato sul conteggio del *numero* e del *tipo* di errori per proposizione o per clausola. In questo caso due sono le modalità di classificazione proposte, una basata sul tipo di errore — lessicale, morfo-sintattico, pragmatico o di pronuncia — l’altra basata sulla gravità dell’errore, definita in termini di comprensibilità. La classificazione degli errori per tipo ha messo in evidenza poi come, con il progredire della competenza linguistica, l’accuratezza non si sviluppi linearmente e come errori di diverso tipo possano avere evoluzioni più o meno lineari. Anche questa misura dimostra dunque alcuni punti di debolezza: nelle produzioni degli apprendenti non è sempre univoca l’assegnazione di un errore a una categoria, il

conteggio indifferenziato di tutti gli errori inoltre non è sensibile alla quantità di contesti offerti. Nell'applicazione a dati spontanei e semi-spontanei si aggiungono difficoltà rispetto al controllo di una sufficiente occorrenza di contesti obbligatori: alcune strutture potrebbero essere opzionali in quel determinato contesto comunicativo.

Un terzo tipo di unità cerca di superare questi ultimi limiti conteggiando e identificando le occorrenze corrette di determinate strutture. Le ricerche condotte in questa direzione si sono concentrate su un numero limitato di strutture. L'interesse qui è il possibile punto di contatto tra misure e studi sulle sequenze acquisizionali.

Wolfe–Quintero *et al.* (1998) nella loro meta-analisi evidenziano come per le misure di accuratezza considerate sopra non vi sia una chiara relazione con il livello di proficiency globale, ma piuttosto con valutazioni olistiche o progressi nel breve periodo. I risultati dei diversi studi sono spesso contraddittori. I ricercatori ipotizzano che le misure di accuratezza non siano collegate al costrutto sviluppo, ma piuttosto a quello di errore. Larsen–Freeman (1983) mostra come la maggior parte degli errori grammaticali riguardi aspetti della morfologia, articoli, preposizioni e tempi verbali, sostenendo che le strutture morfologiche risultino più complesse di quelle sintattiche. Bardovi–Harlig e Bofman (1989) dimostrano come sia gli studenti che hanno superato con successo il test che quelli che non l'hanno superato mostrano la stessa distribuzione di errori per tipo (57% errori morfologici, 27% errori lessicali, 15% errori sintattici), nonostante il gruppo meno competente commettesse in generale più errori. Lo studio suggerisce che possa esserci una differenza tra le strutture sintattiche più universali e centrali di una lingua, acquisite prima, e quegli aspetti più *language-specific*, periferici, locali della morfologia. Bardovi–Harlig e Bofman (1989: 28) definiscono

questo come «strong syntax–weak morphology stage». La definizione di come l'errore si relazioni allo sviluppo linguistico e la spiegazione delle ragioni per cui le misure di accuratezza (al contrario di complessità e fluenza) non correlano con i livelli di proficiency sono aspetti che la ricerca in futuro dovrà approfondire. In particolare, così come suggerito anche in Pallotti (2009; 2015), è importante tenere a mente la netta distinzione tra accuratezza e sviluppo: è possibile infatti essere molto accurati e impiegare strutture che emergono ai primi stadi e viceversa essere poco accurati, ma impiegare strutture più avanzate. Thewissen (2013) in uno studio recente ha osservato come ai livelli intermedi la frequenza di errori non è legata al livello di proficiency. Attraverso l'analisi condotta in questo studio si cercherà di offrire un contributo su questi temi. Nella sezione § 2.4. si illustrano più in dettaglio le misure scelte per l'analisi.

La terza dimensione CAF presa in considerazione è la fluenza. In ambito acquisizionale il termine fluenza è interpretato in due modi. Secondo una definizione ampia indica in generale la proficiency orale. In questo senso, con fluenza ci si riferisce al punto più alto di una scala che descrive la conoscenza della L2. Secondo una definizione più ristretta con fluenza ci si riferisce a una componente specifica della proficiency orale. Lennon (1990), riprendendo il lavoro di Mohle (1984), considera la fluenza in senso stretto, misurandola in termini di velocità e lunghezza della produzione. In questo senso la fluenza riguarda «the processing of language in real time» (Schmidt, 1992: 358) con attenzione primaria verso il significato (Foster, Skehan, 1996: 304). Essa può essere caratterizzata in due modi. Il primo riguarda l'uso appropriato di routine, dove per routine si intende sia una formula pragmatica (House, 1966) che espressioni automatizzate (Ellis, 1996). L'uso di routine genera un aumento nelle misure di fluenza

in quanto facilita la realizzazione di unità di produzione. Il secondo riguarda l'automatizzazione dell'uso della lingua. Towell, Hawkins e Bazergui (1996: 85) affermano che «fluent second language production requires that knowledge (from whatever source) be proceduralised».

Seguendo la definizione ristretta le misure di fluenza sono di due tipi (Wiese, 1984; Lennon, 1990): variabili temporali, in relazione alla velocità dell'eloquio e alle pause, fenomeni di esitazione in relazione alle disfluenze. Queste misure indicano il grado di successo di un apprendente nel gestire il *processing* in tempo reale tipico dell'interazione. Le variabili temporali si riferiscono alla velocità e lunghezza nella realizzazione di unità. I fenomeni di esitazione invece sono «connected to moment-by-moment decisions during performance, reflecting adjustments and improvements that are feasible within the pressure of real-time communication» (Skehan, Foster, 1999: 230). I fenomeni di esitazione possono essere collegati a differenze individuali, con apprendenti che monitorano costantemente le loro produzioni e studenti che non monitorano per nulla. Skehan (1998) dimostra attraverso un'analisi fattoriale l'indipendenza dei due tipi di misure di fluenza, definendo fluenza legata alla velocità (*speed fluency*) e fluenza legata alle pause (*breakdown fluency*) le dimensioni della fluenza caratterizzata dalle variabili temporali e fluenza legata alle riparazioni (*repair fluency*) quella definita dai fenomeni di esitazione. Così come per la complessità sintattica (vedi sopra) il costrutto fluenza è multidimensionale. Un'analisi per essere efficace dovrà utilizzare misure rappresentative sia delle variabili temporali che di fenomeni di esitazione, così da riuscire a indagare le sotto-dimensioni della fluenza identificate (cfr. Kormos, Denes, 2004; De Jong *et al.*, 2012). La tabella 1.6. riporta le principali misure impiegate distinguendole in rapporto alle dimensioni di fluenza: fluenza legata



alla velocità e fluenza legata alle pause, entrambe rilevate con misure temporali; fluenza legata alle riparazioni, basata sulle misure che rilevano auto-correzioni e riformulazioni.

Le misure della fluenza legata alla velocità più frequentemente usate in letteratura sono velocità d'eloquio e lunghezza tra due pause. Velocità d'eloquio è una misura che combina due aspetti temporali della produzione orale, il tempo di pianificazione *on-line* e la velocità di articolazione. La misura lunghezza tra due pause ci dà informazioni invece su quanto il parlante è in grado di produrre segmenti di messaggi senza pause. Questa misura rileva la capacità del parlante di accedere velocemente a "pezzi" di lingua già pronti (sistema basato sugli esempi) e contemporaneamente la capacità di organizzare il parlato per unità discrete. Le misure solitamente impiegate per osservare la fluenza legata alle pause sono due, numero di pause e durata delle pause, e rilevano quanto il parlante deve interrompere il discorso per pianificare il messaggio. Gli apprendenti che fanno meno pause sono in questo senso più fluenti. Fluenza legata alla velocità e fluenza legata alle pause sono considerate indipendenti: un apprendente può ottenere una certa velocità d'eloquio e allo stesso tempo fare lunghe pause. Per i fenomeni di esitazione le misure sono quattro: false partenze, ripetizioni, riformulazioni e sostituzioni e indicano le capacità del parlante di sostenere efficacemente la comunicazione in tempo reale.

Wolfe-Quintero *et al.* (1998), ammettendo di esprimere una posizione minoritaria, preferiscono considerare lunghezza della T-unit o della clausola e lunghezza delle T-unit senza errori come migliori misure per la fluenza nella composizione scritta. La lunghezza media delle T-unit è interpretata come misura di fluenza piuttosto che di complessità sulla base della constatazione che la misura parole

per T-unit non distingue le cause dell'incremento della lunghezza dell'unità di produzione, rilevando pertanto la lunghezza e non la complessità. La lunghezza della T-unit senza errori è in genere trattata come misura di accuratezza, proprio perché la correttezza è parte della formula. Wolfe-Quintero *et al.* (1998) suggeriscono come tale misura rileva la fluenza delle unità senza errori, proprio perché misura la lunghezza di tali unità, ma non il numero di errori o il

**Tabella 1.6.** Le misure di fluenza, adattato da Ellis e Barkhuizen (2005).

<i>variabili temporali</i>	
<i>fluenza legata alla velocità</i>	
velocità d'eloquio	solitamente misurata in termini di numero di sillabe o parole prodotte al secondo o al minuto. Il numero delle sillabe disfluenze escluse ( <i>pruned syllables</i> ) e/o il numero delle sillabe disfluenze incluse ( <i>unpruned syllables</i> ) viene conteggiato e diviso per il numero totale di sillabe o parole prodotte
lunghezza tra due pause	numero medio delle sillabe o delle parole comprese tra due pause di una lunghezza predefinita (per es. 1 secondo). Questa misura non prende in considerazione le disfluenze
<i>fluenza legata alle pause</i>	
numero di pause	il numero totale delle pause, piene e vuote
durata delle pause	misurata in due modi: (a) la durata totale delle pause di una lunghezza predefinita (per es. 1 secondo); (b) la durata media delle pause di una lunghezza predefinita (per es. 1 secondo). La durata delle pause misura il silenzio durante un task
<i>fenomeni di esitazione</i>	
<i>fluenza legata alle riparazioni</i>	
false partenze	proposizioni incomplete (per es. frammenti) o piani abbandonati. Possono o meno essere seguite da riformulazioni
ripetizioni	parole, sintagmi o clausole ripetuti
riformulazioni	parole, sintagmi o clausole ripetute parzialmente o con auto-correzioni
sostituzioni	elementi lessicali sostituiti da altri elementi lessicali

grado di correttezza delle unità stesse (cfr. Ishikawa, 1995). La loro meta-analisi mostra come queste misure crescano linearmente in relazione ai livelli di proficiency nei vari studi, indipendentemente da task, lingua target o definizione del livello di proficiency. L'unica eccezione è riportata nello studio di Larsen–Freeman (1983).

La letteratura sulla fluenza ha spesso messo enfasi sulla rapidità dell'eloquio dei parlanti nativi. È chiaro che la stessa fluenza dei nativi non è un costrutto uniforme e monolitico, ma vi è un certo grado di variabilità individuale legata all'argomento della conversazione, alla situazione, alle caratteristiche individuali del parlante o altri possibili fattori. A questo proposito Filmore (1979: 93) distingue almeno quattro tipi di fluenza nativa (1) «the ability to fill time with talk» o «disc-jockey fluency,» (2) «the ability to talk in coherent, reasoned and “semantically dense” sentences,» (3) «the ability to have appropriate things to say in a wide range of contexts,» e (4) «the ability [...] to be creative and imaginative in language use». Raupach (1983) evidenzia come il parlato di nativi apparentemente corretti e fluenti sia comunque caratterizzato spesso da differenze nelle scelte sintattiche o false partenze ed enunciati incompleti. Questi aspetti del parlato non sempre dunque influenzano l'impressione generale di fluenza e spesso sono evidenti solo tramite un'osservazione attenta o una trascrizione dettagliata. Poiché questi fenomeni sono presenti anche nelle produzioni fluenti di parlanti nativi, sarà la distribuzione della loro frequenza piuttosto che la loro presenza o assenza l'elemento che offre informazioni. Davies (2003) sostiene che ciò che distingue le pause di un nativo da quelle di un non nativo è più facilmente la posizione piuttosto che la semplice durata. Nello studio condotto da Foster e Skehan (1996), ad esempio, nativi e non-nativi si differenziano proprio nella distribuzione delle pause, con i primi che le inseriscono prevalentemente in corrispondenza del

confine di AS-unit. Risultati di questo tipo suggeriscono come un approccio che si limiti a un semplice conteggio delle pause possa avere dei limiti nella sua applicazione.

La discussione presentata in § 1.2.1.2. dimostra come le tre dimensioni CAF rilevino diversi aspetti della performance e come nella produzione esista in alcuni casi una correlazione tra le dimensioni CAF, in altri si generi un complesso sistema di trade-off. Quando l'obiettivo è individuare misure capaci di rilevare i livelli di proficiency e lo sviluppo longitudinale, è fondamentale conoscere non solo l'andamento delle singole misure prese individualmente, ma anche le loro interconnessioni. Peter Skehan e Peter Robinson nei loro modelli, seppure con le differenze che abbiamo già visto, identificano una tensione tra forma (complessità e accuratezza) e significato (fluenza). Skehan (1998) in particolare ha descritto queste tensioni sotto forma di ipotesi di trade-off, predicendo come l'attenzione a una dimensione abbia un'influenza sulle altre (cfr. § 1.2.1.2.). La ricerca sembra supportare l'idea di un trade-off tra complessità e accuratezza, sebbene non sia semplice trarre generalizzazioni a partire dagli studi condotti fino ad oggi a causa delle differenze metodologiche. Così come evidenziato da Ellis (2005) nella sua presentazione di ricerche a proposito, i diversi tipi di planning favoriscono complessità o accuratezza, raramente entrambe. Analoghi i risultati relativi agli effetti della ripetizione del task o alle caratteristiche dei task, con uniche eccezioni gli studi di Gilabert (2007) e Tavokoli, Skehan (2005). La fluenza aumenta invece nella maggior parte degli studi, le differenze nelle misure impiegate non permettono però generalizzazioni in proposito.

Interessanti i risultati sui possibili trade-off tra dimensioni CAF ottenute da Wolfe-Quintero *et al.* (1998) nella loro meta-analisi. I ricercatori osservano infatti differenze tra misure di

accuratezza da un lato e complessità e fluenza dall'altro nel modo in cui si relazionano ai livelli di proficiency. Le misure di accuratezza risultano in genere linearmente coerenti con giudizi olistici e cambiamenti sul breve periodo in classi intatte, mentre complessità e fluenza sono significativamente correlate con programmi, livello di scuola o livello di proficiency globale. Wolfe–Quintero *et al.* sulla base di questo dato concludono che le misure di fluenza e complessità si rifanno allo stesso costrutto, quello di sviluppo, mentre l'accuratezza potrebbe essere legata ad un diverso costrutto, quello di errore. Lo studio qui condotto si pone l'obiettivo di approfondire queste questioni.

### 1.3. L'approccio del testing

I test linguistici sono sempre più diffusi in contesto scolastico e professionale, i risultati vengono utilizzati per valutare le competenze di un individuo, ma soprattutto per prendere decisioni, non solo educative, su di lui. Ad esempio, un test linguistico potrebbe essere utilizzato per analizzare i bisogni degli studenti immigrati di una scuola o essere prerequisito per l'accesso a un determinato percorso di studi o ancora per ottenere la cittadinanza. I test pertanto hanno la potenzialità di aiutarci a raccogliere informazioni utili a una varietà di individui, ma per realizzare appieno questa loro funzione è necessario che il loro uso e la conseguente interpretazione dei risultati siano affidabili e validi (cfr. Bachman, 2004). In altre parole la comprensione del rapporto tra la performance di un candidato e le inferenze, predizioni o generalizzazioni fatte sulla base della performance stessa è il problema centrale del testing (*validity problem*). Una serie di variabili possono infatti influire sulla validità delle inferenze fatte a partire da un test:

il task, il valutatore, la scala di proficiency, le caratteristiche del candidato e infine le possibili interazioni tra le variabili stesse (cfr. McNamara, 1996).

Se le produzioni di un apprendente variano, è essenziale determinare quale performance testare. Ecco che lo studio della variazione è un tema centrale anche nel testing, poiché può contribuire in modo sostanziale all'elaborazione di criteri efficaci per la selezione di task e di metodi di valutazione. I modelli di Bachman e Palmer (1996) e Skehan (1998) presentati in § 1.1., descrivendo i fattori che possono influire sulla performance determinando variazione, hanno stimolato nell'ambito del testing una varietà di studi che hanno cercato di approfondire il tema. Questa sezione presenta la letteratura sulla variazione nel testing (§ 1.3.1.) e sulla sua misurazione (§ 1.3.2.).

### 1.3.1. *La variazione*

L'attenzione rivolta alla competenza comunicativa in campo acquisizionale e didattico ha avuto un impatto importante anche nella valutazione. A partire dagli anni '90 la letteratura sul testing si è infatti sempre più occupata del cosiddetto *language performance testing*, ossia la valutazione del candidato sulla base di performance elicitate, che hanno un certo grado di autenticità rispetto ai bisogni comunicativi del candidato stesso (per una presentazione dettagliata, vedi McNamara, 1997; Norris *et al.*, 1998). Se, come dimostra la letteratura acquisizionale, diversi task realizzano diversi gradi di accuratezza, fluenza, grammaticalità o comprensibilità, offrendo di conseguenza diverse prospettive sul grado di sviluppo dell'interlingua, quali strumenti possono permettere la creazione di un profilo affidabile e rappresentativo? Una soluzione riguarda l'uso di task più autentici, ossia più vicini agli usi reali della lingua. L'abilità linguistica così può essere misurata valutando la ca-

pacità dello studente di cavarsela in quella situazione. Chiaramente un approccio di questo tipo aumenterebbe la validità della prova, a scapito però di affidabilità e generalizzabilità (Hoekje, Linnel, 1994).

La ricerca ha tentato di risolvere questa tensione adottando contemporaneamente un approccio basato sui task (*task-centred approach*), per mantenere il legame con il mondo reale, e un approccio basato sul costrutto (*construct-centered approach*), per garantire comunque un certo grado di astrazione e generalizzazione. L'interesse per questa forma di testing ha determinato infatti il bisogno da un lato di comprendere meglio la natura dei task usati per elicitarne la performance, dall'altro di sviluppare strumenti adeguati per valutarne l'utilità (Bachman, 2000). Proprio per la dimensione sociale che il testing assume, la valutazione è luogo di possibili conflitti politici o scontro di valori; attraverso la validazione, chi si occupa di testing ha l'obbligo di considerare le possibili conseguenze di un test ed evitare quelle negative. Di qui l'attenzione nel testing per definizione del costrutto e validazione (Cronbach, 1971; Messick, 1989; Kane, 1992; Kunnan, (1993); Mislavy *et al.*, 2003; per una presentazione dettagliata del concetto di validazione si vedano Bachman, 1990; Weir, 2005a e McNamara, 2006). Chi elabora un test deve seguire pertanto alcune tappe: innanzitutto deve specificare cosa vuole rilevare circa l'abilità del candidato sulla base di una performance; poi deve determinare che tipo di performance può supportare le osservazioni che si vogliono fare. Questo dipende chiaramente da una teoria che descrive le caratteristiche di una performance di successo, un costrutto che sarà chiaramente riflesso nelle scale di valutazione utilizzate; successivamente deve definire in termini generali i tipi di task in cui il candidato dovrà cimentarsi; e infine deve procedere alla valutazione, e conseguente elaborazione di un punteggio, sulla base della performance elicitata. Infine un aspetto fon-

damentare del processo sarà accompagnare il risultato di un test a dati che ne dimostrano la rilevanza rispetto alle osservazioni che vogliamo fare su quel test (*validity argument*). In altre parole è essenziale che un tester sia in grado di collegare una performance da un lato all'abilità linguistica del candidato, dall'altro all'uso di quell'abilità in un contesto reale. Per fare ciò sarà necessario capire:

- a) come le variabili interne al candidato possano influenzare la performance;
- b) come le variabili esterne al candidato possono influenzare la performance (per es. caratteristiche del task);
- c) come la performance in un task può essere costruita e interpretata in relazione a contesti d'uso reale (per es. utilizzando scale di proficiency o selezionando i task secondo criteri specifici).

Diversi studi si sono dunque occupati del legame tra elementi quali task, valutatore, scala di proficiency e caratteristiche del candidato con la variabilità della performance, identificando una serie di fattori rilevanti. Si illustra la letteratura in proposito.

Per quel che riguarda le caratteristiche dei candidati, O'Sullivan (2000: 71-72) propone un quadro di riferimento e le raggruppa in tre macro-categorie: caratteristiche fisiche (es. sesso, età o disabilità), caratteristiche psicologiche (per es. personalità, motivazione o stile cognitivo) e caratteristiche esperienziali (es. educazione, conoscenza del mondo, paese di residenza, ecc.). Nonostante l'importanza delle caratteristiche del candidato nel processo di validazione, la letteratura sul tema è piuttosto limitata. Per una presentazione a proposito si rimanda a O'Sullivan, (2000); O'Sullivan, Green, (2011); Wigglesworth, (2005).



Un'altra serie di studi si è occupata invece del processo di test-taking e delle strategie usate dai candidati nello svolgere un test. In questi studi si utilizzano spesso metodologie introspettive (per es. Buck, 1991; Wu, 1998; Weigle, 1999; Cohen, 2000; Weigle *et al.*, 2013) e interviste (per es. Purpura, 1993; 1997; 1999; Cheng, Deluca, 2011; Smagorinski, 2001; Sasaki, 2008). In altri più innovativi si applicano le tecniche dell'analisi conversazionale o strumenti analitici di analisi dell'interlingua per approfondire il legame tra intervista orale e tipologia di discorso (per es. van Lier, 1989; Berwick, Ross, 1996; Lazaraton, 1996; Lumley, O'Sullivan, 2005) o caratteristiche linguistiche specifiche e livelli di proficiency (Banerjee *et al.*, 2007; Iwashita *et al.*, 2007; cfr. § 1.4.).

Per quanto riguarda l'impatto del task un buon numero di studi ha mostrato che diversi task possono generare variazione nella performance e di conseguenza nella valutazione del test (cfr. Henning, 1983; Shohamy 1983; Shohamy, Reves, Bejarano, 1986). Così come Tarone (1983) ipotizzava che la variabilità non fosse un semplice fenomeno della performance, ma che piuttosto la conoscenza dell'apprendente possa essere variabile, anche nell'ambito del testing alcuni autori (Chalhoub-Deville, 1995a; 1995b) suggeriscono che la natura stessa del costrutto della L2 sia dipendente del contesto e pertanto vari da task a task. Chalhoub-Deville sostiene la sua proposta elaborando empiricamente task-specific dimensions di proficiency orale su tre diversi task, intervista, racconto di una storia e lettura ad alta voce. Anche nel caso in cui il task sia mantenuto costante, l'argomento può determinare variazioni nelle produzioni: Smith (1989) mostra come pronuncia, grammatica, fluenza e comprensibilità sono soggetti a variazione in relazione all'argomento, specifico di un campo disciplinare o generale, determinando differenze notevoli nei risultati del test. Nel suo studio infatti buona parte dei par-

tecipanti non supera entrambi i test. Risultati analoghi sono stati ottenuti rispetto a fattori quali concretezza, specificità dell'argomento, livello di vocabolario, complessità sintattica, carico cognitivo o quantità di processing richiesto (Alderson, Urquhart, 1985; Anderson *et al.*, 1991; Hale, 1988; Freedle, Kostin, 1993; Perkins *et al.*, 1995; Bachman *et al.*, 1996).

Un ulteriore filone di studi si è occupato dell'impatto che l'interlocutore ha sulle produzioni (cfr. Wiseman, 2012; Winke *et al.*, 2013). Hoekje e Linnel (1994) valutano e confrontano tre strumenti per la valutazione orale, the SPEAK test, the OPI (*oral proficiency interview*) e IP test (*interaction performance*), che si differenziano per presenza o meno dell'interlocutore, numero degli interlocutori, potere dell'interlocutore rispetto a presa dei turni e gestione dell'argomento di conversazione, dimostrando come i tre strumenti mostrino differenze sostanziali riguardo all'uso linguistico considerato, in particolare rispetto a lunghezza dei turni e all'uso dei segnali discorsivi. Brown (2003; 2005) indaga l'influenza sulla performance di elementi quali stile interazionale e identità dell'intervistatore, mostrando come genere, identità professionale, esperienza pongano difficoltà diverse al candidato (per una presentazione dettagliata di queste ricerche vedi anche McNamara, Hill, May, 2002).

Anche rispetto a valutatore e strumenti di valutazione la letteratura evidenzia una serie di possibili cause di variazione. Diversi studi hanno mostrato l'impatto di variabili quali background del valutatore, familiarità, eventuale conoscenza nativa della lingua target, livello di formazione (Brown, 1995; Chalhoub–Deville, 1996; McNamara, 1996; Weigle, 1994; 1998; Schoonen *et al.*, 1997; Major *et al.* 2002; Carey *et al.* 2011; Winke *et al.*, 2013). In alcuni casi il processo di valutazione è stato studiato in dettaglio, così da identificare come il valutatore seleziona una certa valutazione. Linacre (1989) offre una discussione interessante sulla variazione associata ai valuta-

tori e individua i principali punti di possibile divergenza: due valutatori possono differenziarsi per severità generale, oppure due valutatori possono essere più o meno severi rispetto a determinate strutture o compiti, o ancora due valutatori possono distinguersi nell'interpretazione di uno stesso criterio di valutazione, infine i valutatori possono distinguersi in coerenza. In genere gli studi condotti mostrano come, sebbene non sia né possibile e forse nemmeno auspicabile, eliminare del tutto le differenze, la formazione può per lo meno ridurle. Diverse ricerche mostrano infatti come la formazione renda i valutatori più consistenti e riduca le differenze in severità (Lunz, Stahal, 1990; McIntyre, 1993; Weigle, 1994; Shohamy *et al.*, 1992; Isaacs, Thomson, 2013).

Ricapitolando, una delle implicazioni della variazione sul testing riguarda il cosiddetto *validity problem*, ossia la standardizzazione del processo di valutazione stesso, con conseguenze in particolare per l'elaborazione di task e prove di valutazione. La questione è riassumibile nella seguente domanda: che tipo di elicitazione è meglio utilizzare, vista la sensibilità che l'interlingua dimostra verso minime differenze contestuali? E soprattutto, come misurarla? Questo studio attraverso l'esplorazione della variazione situazionale e longitudinale in diversi task cercherà di offrire un contributo a queste questioni. La sezione che segue è dedicata alla presentazione della letteratura relativa alla misurazione della performance attraverso scale di proficiency.

### 1.3.2. *La misurazione della variazione*

Attualmente in Europa il punto di riferimento ufficiale per l'insegnamento e la valutazione degli apprendimenti linguistici è il *Quadro Comune Europeo di Riferimento per le lingue* (d'ora in poi QCER, Consiglio d'Europa 2001; 2007; 2018). Tale documen-

to è il risultato del lavoro trentennale del Consiglio d'Europa nell'elaborazione di scale e descrittori dedicati alle competenze e agli usi linguistici con l'obiettivo di agevolare la cooperazione e il confronto nell'ambito dell'insegnamento/apprendimento delle lingue moderne, proponendo un linguaggio comune e favorendo la trasparenza e la coerenza in termini di definizioni e descrizioni di obiettivi di apprendimento, contenuti e metodi.

Sin dalla sua pubblicazione il QCER ha stimolato l'interesse di chi si occupa di insegnamento o valutazione, e la sua applicazione è stata oggetto di sperimentazione in diversi contesti di insegnamento e valutazione a livello europeo (ad esempio, Progetto Lingue Moderne; The Dutch CEF Construct Project, Dialang, in Alderson, 2002; Alderson *et al.*, 2004; Cefling Project, in Alanen *et al.* 2010; Young Learners Project; EuroPASS Project; AYLLIT Project in Hasselgreen 2013) e in ambito italiano (ad es. Progetto Lingue 2000; Grego Bolli, 2006a; 2006b; vedi anche RILA, 2013, numero monografico). Questi studi riconoscono le potenzialità e il valore del documento, e mettono in luce alcuni dei limiti nella sua applicazione al testing, evidenziando la necessità di ulteriori ricerche ai fini della sua implementazione. In questa sezione dopo una breve presentazione delle scale di proficiency, si illustrano le principali caratteristiche del QCER e le potenzialità e i limiti delle sue applicazioni al testing.

### 1.3.2.1. Le scale di proficiency

In risposta ai bisogni di trasparenza dei sistemi educativi che danno sempre più valore alle descrizioni di cosa significhi esattamente essere a un determinato livello di proficiency, nell'ambito della valutazione si sono sempre più diffuse le scale di proficiency linguistica. Nella letteratura relativa al testing ci si riferisce spesso a tali scale con una varietà di termini «band scores,

band scales, profile bands, proficiency levels, proficiency scales, proficiency ratings» (Alderson, 1991: 71) o «guidelines, standards, levels, yardsticks, stages, scales or grades» (De Jong, 1992: 43). I diversi nomi impiegati per indicare le scale di proficiency linguistica riflettono il fatto che le diverse scale prodotte abbiano origini diverse. Oltre a una differenziazione di origine, le scale possono essere poi ulteriormente distinte sulla base del loro uso. A tal proposito Alderson (1991: 72-4) introduce una loro classificazione funzionale distinguendo tre tipologie di scale:

- a) *user-oriented*, con la funzione di riportare informazioni sul comportamento di un apprendente a un determinato livello;
- b) *assessor-oriented*, con la funzione di guidare il processo di valutazione offrendo una descrizione della qualità della produzione;
- c) *constructor-oriented*, con la funzione di guidare la costruzione di prove offrendo una descrizione dei task comunicativi che l'apprendente dovrebbe affrontare.

L'elemento comune di questi strumenti è l'obiettivo che perseguono: offrire appunto una descrizione per livelli successivi della competenza linguistica. Tali scale possono pertanto contribuire alla trasparenza e alla coerenza nell'insegnamento linguistico ed essere utili per:

- a) fornire uno "stereotipo" con cui l'apprendente può confrontare la sua posizione (Trim, 1978; 2001; Oscarson, 1978; 1984);
- b) fornire livelli che riflettono il curriculum scolastico e che possono essere formulati in termini di livello di proficiency raggiunta (Hargreaves, 1992; Hasselgreen, 2013);

- c) aumentare l'affidabilità di valutazioni soggettive, in particolare rispetto alle competenze di produzione, fornendo uno standard comune (Alderson, 1991; Shohamy, 2017);
- d) fornire linee guida per la costruzione di prove di valutazione (Alderson, 1991);
- e) offrire punti di riferimento per chi deve gestire test d'ingresso, programmazioni, preparazione di materiale didattico, test di progresso e certificazioni finali (North, 1991; 2007);
- f) permettere il confronto tra sistemi e gruppi di apprendenti utilizzando un metro comune (Bachman, Savignon, 1986; Carrol, West, 1989; cfr. North, 2000; 2002; Morrow, 2004).

Diverse critiche sono state mosse alle scale di proficiency linguistica prodotte, in particolare rispetto al fatto che esse siano per la maggior parte scritte intuitivamente e quasi sempre manchino di una dimostrata validità (cfr. Brindley, 1998; Alderson, 2007; Harsh, 2018). Costruire una scala di proficiency non è certo un'operazione semplice. Si pongono infatti due tipi di problemi: uno descrittivo — le categorie utilizzate devono rifarsi ad un modello di proficiency — l'altro relativo alla misurazione — la scala di proficiency dovrà essere collegata a un modello di misurazione.

Dal punto di vista descrittivo, le scale di proficiency, in quanto composte da una serie di livelli ascendenti, sono strettamente collegate alle ricerche acquisizionali. I risultati ottenuti in letteratura non sono però ancora sufficientemente chiari per permettere una descrizione dettagliata dei percorsi di apprendimento: non c'è un'interpretazione univoca di competenza, proficiency e performance, né è stato elaborato un modello descrittivo condiviso della lingua in uso (cfr. § 1.1.). Ne consegue, come eviden-

zia McNamara (1995), che spesso le scale tendono a offrire un modello semplificato, per non dire impoverito, della lingua in uso, che con difficoltà può essere collegato a un modello teorico. Dal punto di vista della misurazione poi nella maggior parte dei casi sono scarsi o nulli i tentativi di dimostrare la validità della formulazione verbale dei descrittori prima della loro pubblicazione. Quando vengono fornite informazioni rispetto alla validità si tratta generalmente di analisi statistiche dell'affidabilità dei giudizi tra valutatori che utilizzano la scala dopo che essa è stata prodotta (cfr. McNamara, 1995; Harsh, Martin, 2012). Il fatto che diversi utilizzatori possono interpretare coerentemente le scale di language proficiency non garantisce infatti che la scala sia una valida descrizione delle diverse tappe dell'acquisizione di una seconda lingua (Brindley, 1998: 22).

Per questa tensione esistente tra modelli teorici e modelli operativi sviluppati da esperti del settore e l'assenza di una descrizione empirica adeguatamente validata della complessità della proficiency linguistica, alcuni autori sono estremamente critici verso tali strumenti, suggerendo come la costruzione di scale di proficiency o quadri di riferimento non sia operazione da intraprendere (cfr. Lantolf, Frawley, 1985; 1988; 1992). Come evidenzia North (2000: 13) nonostante i problemi connessi al descrivere un fenomeno complesso con poche parole e sulla base di una teoria incompleta, le scale di proficiency linguistica hanno la potenzialità di avere un'influenza positiva sull'orientamento, l'organizzazione e la valutazione dell'apprendimento linguistico e per questo meritano di essere sviluppate. Secondo l'autore infatti i limiti delle scale possono essere ridimensionati, facendo una distinzione fondamentale tra quello che è un modello teorico di descrizione della natura della proficiency e quello che può essere un modello operativo utile a insegnanti e valutatori. Un modello operativo chiaramente è sempre una semplificazione di un modello teorico: fa riferimento alla teoria,

ma ne reinterpreta gli elementi per renderli più accessibili in un particolare contesto. D'altronde, come suggeriscono anche altri autori, chi si occupa di valutazione degli apprendimenti linguistici non può certo attendere lo sviluppo di un modello di proficiency linguistica validato sperimentalmente prima di elaborare criteri per la valutazione delle produzioni nella seconda lingua (Brindley, 1998: 56), ma piuttosto dovrà lavorare «con tassonomie che appaiono sensate anche se non completamente riconducibili a una descrizione teorica» (Hulstijn, 1985: 277).

L'arbitrarietà che dunque caratterizza gli standard di proficiency può essere limitata:

- a) prendendo in considerazione teoria e ricerca;
- b) assicurandosi che le decisioni prese si basino su un largo consenso per i contesti rilevanti;
- c) tenendo conto della soggettività delle valutazioni individuali e di gruppo;
- d) chiarendo esplicitamente quali sono i limiti della dichiarata validità empirica: anziché limitarsi ad affermare che una scala è valida, bisognerebbe specificare per cosa è valida (Henning, 1990: 379).

Il QCER tenta di superare i limiti discussi: le scale di descrittori non sono state formulate in maniera intuitiva, ma sono state sviluppate seguendo una metodologia rigorosa, effettuando analisi di tipo sia qualitativo che quantitativo (cfr. North, 2000). Dopo un'attenta revisione delle scale di proficiency esistenti rispetto alle categorie rilevanti per il modello teorico assunto nel QCER, gli autori hanno selezionato e prodotto una serie di descrittori. Questi sono stati prima sottoposti a una revisione con il coinvolgimento di insegnanti nel ruolo di informatori, poi calibrati utilizzando una versione scalare del *Rasch Model* (per una presentazione del modello vedi ad esempio McNama-



ra, 1996, capitolo 6), ottenendo così una scala comune di descrittori della proficiency linguistica. Gli autori hanno dunque tentato di identificare categorie che potessero essere adeguate alla teoria, comprensibili a valutatori, insegnanti e studenti, e capaci di un'interpretazione sufficientemente condivisa in diversi contesti educativi (cfr. North, 2000). Nella sezione che segue il QCER viene presentato più in dettaglio.

### 1.3.2.2. Il Quadro Comune Europeo di Riferimento

Il QCER è riconosciuto come una delle pubblicazioni più influenti nel campo dell'apprendimento e della valutazione, è il prodotto finale di un progetto europeo iniziato negli anni '70 e coordinato dal Consiglio d'Europa (1988; 1992; 1996; 2002; 2003; 2005; 2007; 2018). In una prima fase è stata elaborata una serie di descrittori dei livelli di competenza, prima il *Threshold level* e poi i successivi *Waystage*, *Breakthrough Levels*, *Vantage*, *Effective* e infine *Mastery levels* (cfr. van Ek, 1975; 1977; Trim, 1978; van Ek, Trim, 1984; 1990; 1991; 1995). In una seconda fase si è mirato a riunire i diversi livelli in un unico quadro coerente (cfr. North, 2000). Tale quadro, il QCER appunto, sviluppato da John Trim, Brian North, Daniel Coste e Joseph Sheils è stato poi oggetto di un'intensa revisione a livello europeo da parte di esperti del settore. La complessità del lavoro ha portato alla pubblicazione di due versioni intermedie (Consiglio d'Europa, 1996; 1998), prima di arrivare alla versione finale del 2001 (Consiglio d'Europa, 2001). Inoltre lo sviluppo del QCER è stato documentato nei dettagli in diverse pubblicazioni (North, Schneider, 1998; North, 2000; Schneider, North, 2000) e in una serie di studi di caso (Alderson, 2002; *Language Testing, Special Issue*, 2005). Per l'italiano, si vedano ad esempio Vedovelli (2002), Mezzadri (2004), Bosisio (2005), Grego Bolli, (2006a), Jafrancesco, (2004) e RILA, *Special Issue* 2013.

Il QCER è un documento complesso e ambizioso, che nella propria dichiarazione di intenti al capitolo 1 si offre come strumento di lavoro a tutti coloro che si occupano di apprendimento linguistico: docenti e studenti naturalmente, ma anche autori di programmi e curricula, responsabili delle politiche linguistiche, esponenti del mondo dell'editoria e enti certificatori. Gli autori del QCER descrivono così il documento e le sue finalità:

*Il Quadro Comune Europeo di Riferimento per le lingue* fornisce una base comune in tutta l'Europa per l'elaborazione di programmi, linee guida curriculari, esami, libri di testo per le lingue moderne ecc. Descrive in modo esaustivo ciò che chi studia una lingua deve imparare per usarla per comunicare e indica quali conoscenze e abilità deve sviluppare per agire in modo efficace. La descrizione riguarda anche il contesto culturale nel quale la lingua si situa. Inoltre il *Quadro di Riferimento* definisce i livelli di competenza che permettono di misurare il progresso dell'apprendimento ad ogni stadio del percorso, nella prospettiva dell'educazione permanente. (QCER, p. 1)

Il QCER descrive dunque la competenza linguistica attraverso una serie di scale accompagnate da introduzioni che forniscono utili informazioni su elementi rilevanti nell'insegnamento, nell'apprendimento e nella valutazione delle lingue. L'approccio di fondo che caratterizza il QCER è l'orientamento all'azione. La competenza, pertanto, viene formulata con la descrizione di cosa l'apprendente è in grado di fare nello svolgimento di un compito comunicativo in un determinato contesto utilizzando un certo repertorio di risorse linguistiche. Il documento permette di descrivere la competenza dell'apprendente secondo due dimensioni: una verticale quantitativa, data dalle scale globali, e una orizzontale qua-

litativa, costituita dai parametri per le attività comunicative e la competenza linguistico-comunicativa (QCER, capp. 4 e 5).

I descrittori, formulati in termini di “dichiarazioni di capacità” (*Can-do statements*), sono raggruppati in tre livelli principali: A (livello elementare), B (livello intermedio) e C (livello avanzato). A sua volta ciascuno dei tre livelli è suddiviso in altri due sotto-livelli, dando origine a una scala a 6 bande distinte secondo le due modalità, dello scritto e del parlato, e secondo i tre parametri, della ricezione, della produzione e dell’interazione. Questi vengono poi dettagliati in descrittori applicabili ad attività ben definite quali *Intervistare ed essere intervistati* o *Prendere la parola in una discussione*. A questo si affianca poi una descrizione dei parametri ricezione, produzione e interazione secondo le categorie che catturano aspetti delle competenze di un apprendente, quali la *fluenza nel parlato*, l’*accuratezza grammaticale* e l’*appropriatezza sociolinguistica*. I descrittori così ottenuti sono stati selezionati, validati e ordinati lungo la scala (cfr. North, Schneider, 1998; North, 2000; 2014).

La procedura impiegata dai redattori del QCER per ordinare i descrittori sulla scala è estremamente raffinata, ma non esente da limiti. Per chi utilizza il QCER è importante capire il modello statistico impiegato, l’estensione del *Rasch Model* (cfr. Figueras *et al.*, 2005). Caratteristica fondamentale di tale modello è l’assunzione dell’unidimensionalità del fenomeno. Questo significa che, secondo il modello, se il descrittore X sulla scala sta allo stesso punto del descrittore Y, allora un apprendente che riesce a fare X con un certo livello di successo riuscirà anche a fare Y. Questa intercambiabilità presupporrebbe che la proficiency nel parlato, il nostro descrittore Y ad esempio, potrebbe essere osservata attraverso un test di ascolto, il nostro descrittore X, e tale generalizzazione non sarebbe accettabile dal punto di vista teorico. Come riconoscono

gli stessi autori del documento, un modello unidimensionale non può soddisfare pienamente le richieste di un sistema multidimensionale come quello rappresentato dalla competenza linguistica (cfr. North, Schneider, 1998; North 2000; 2014; cfr. § 1.1.2.). Per questo il QCER stesso sottolinea in più punti di essere un documento di riferimento e mette in guardia su come un suo uso troppo rigido possa addirittura essere controproducente:

La costruzione di un quadro esaustivo, trasparente e coerente per l'apprendimento e l'insegnamento delle lingue non implica l'adozione di un sistema unico e uniforme. Al contrario, il QCER dovrebbe essere aperto e flessibile, in modo da poter essere applicato, con gli opportuni adattamenti, alle singole situazioni (QCER, p. 9).

Di conseguenza il QCER si definisce uno strumento multifunzionale, flessibile, aperto, dinamico, amichevole e soprattutto non dogmatico:

Gli esempi sono semplicemente indicativi, è possibile prenderne in considerazione alcuni, scartarne altri, e aggiungerne di nuovi. Il lettore deve sentirsi libero di farlo, dal momento che è suo compito definire gli obiettivi e i risultati (QCER, p. xiv).

Il lettore è invitato a usare le scale e i relativi descrittori in modo critico (QCER, p. xv).

È importante sottolineare come il QCER non sia stato concepito per indicare come devono essere apprese, insegnate o valutate le lingue seconde, ma piuttosto per far emergere la consapevolezza, stimolare la riflessione e migliorare la comunicazione tra esperti del settore, ponendosi appunto come punto di riferimento. Come sottolineano gli stessi autori nell'introduzione, obiettivo del QCER è far scaturire

domande e non dare risposte preconfezionate. In quanto “di riferimento” il QCER intende promuovere riflessioni e discussioni e offrirsi come strumento per descrivere la diversità, non certo indirizzare verso un sistema pan-europeo rigido e uniforme (Figueras *et al.*, 2005; Figueras, 2012).

Il QCER pone un’attenzione particolare alla valutazione e alla certificazione ed esplicita chiaramente le sue possibili funzioni in questi due ambiti:

Le possibilità d’uso [rispetto alla valutazione della competenza linguistica] sono tre:

- a) per specificare il contenuto di test ed esami;
- b) per definire i criteri in base ai quali un obiettivo di apprendimento si può considerare raggiunto, con riferimento sia alla valutazione di una particolare prestazione orale o scritta, sia alla valutazione continua dell’insegnante, dei pari e all’autovalutazione;
- c) per descrivere i livelli di competenza rilevati in test ed esami esistenti, in modo da fare confronti tra sistemi di certificazione diversi. (QCER, p. 24).

L’accettazione del QCER da parte di specialisti dell’insegnamento e della valutazione in diversi contesti europei ha motivato la realizzazione del Manuale *Relating language examinations to the Common European Framework of Reference for languages: learning, teaching, assessment* (Consiglio d’Europa, 2003, 2005; cfr. Figueras *et al.*, 2005). Obiettivo della pubblicazione è rispondere ai bisogni degli enti di valutazione che devono mettere in relazione i propri esami con i livelli del QCER, si presenta pertanto come una guida all’applicazione di buone prassi del testing e dei codici professionali del settore.

Poiché il QCER si pone come risorsa per specificare cosa valutare, come interpretare la performance e come confron-

tare diversi sistemi di certificazione, per riuscire a realizzare appieno queste sue funzioni deve soddisfare alcuni criteri: deve cioè essere esaustivo<sup>2</sup>, trasparente e coerente. Ecco come vengono definiti tali criteri nel QCER:

- *esaustivo*: dovrebbe cercare di specificare nel modo più completo possibile la gamma di conoscenze, capacità e usi linguistici (QCER, p. 8);
- *trasparente*: l'informazione deve essere formulata in modo chiaro ed esplicito, in modo da essere accessibile e facilmente comprensibile alle persone interessate (QCER, p. 8);
- *coerente*: la descrizione deve essere scevra da contraddizioni interne (QCER, p. 9).

Il QCER ha avuto un impatto positivo sulla valutazione in Europa e sicuramente è al momento l'elenco di descrittori di proficiency meno arbitrario. Nonostante i punti di forza però, l'uso del QCER attuale ai fini della valutazione pone alcuni problemi (per es. Fulcher, 2004; Little, 2011; Hulsijin, 2011; 2014; Simons, Colpaert, 2015). Il QCER è certamente più esplicito e preciso di documenti che lo hanno preceduto, ma per sua natura rimane un sistema descrittivo astratto. Questo genera una discrepanza tra le descrizioni disponibili e i bisogni concreti di chi si occupa di valutazione (Green, 2010; Harsh, 2014; Díez-Bedmar, 2017). Bisogna chiedersi infatti se e in quale misura sia

2. Nella versione inglese del QCER si usa l'aggettivo *comprehensive*: «It describes in a comprehensive way what language learners have to learn to do in order use a language» (CEFR, p. 1). A questo proposito Figueras *et al.* (2005: 264) sottolineano che «The CEFR has aimed to be *comprehensive*, but this has to be understood in a very practical way, as not being synonymous with *exhaustive*». Mi sembra opportuno tenere conto di questa precisazione anche a proposito dell'italiano *esaustivo*, forse più cautamente sostituibile con *ampio*.

realmente possibile per gli operatori cui il QCER è destinato, applicare concretamente il documento o parti di esso per il proprio lavoro di tutti i giorni, poiché il QCER non è né di facile interpretazione, né di facile utilizzo, ma richiede piuttosto che gli operatori siano accompagnati nella sua applicazione. Due sono i principali punti critici da tenere in considerazione: interpretazione e quantificazione (*linking problem*):

Per quel che riguarda l'interpretazione il problema è generato dal fatto che i *Can-do statements* non sempre coincidono con comportamenti direttamente osservabili. Essi devono essere intesi come categorie astratte e richiedono pertanto interpretazione da parte di chi li usa per costruire prove o valutare la performance. È necessario chiedersi quindi fino a che punto il QCER può essere uno strumento utile per la valutazione di diverse attività linguistiche e competenze. I problemi riscontrati nell'uso delle scale sembrano riguardare in particolare quattro aspetti: terminologia, lacune nelle specificazioni necessarie per il testing, incoerenze nei descrittori e assenza di definizione per alcune strutture. Vediamo un esempio a titolo illustrativo. Il descrittore relativo a *coerenza-coesione* per il livello A2+ è così formulato:

[l'apprendente] è in grado di collegare frasi semplici usando i connettivi più usuali per raccontare una storia o descrivere qualcosa, realizzando un semplice elenco di punti.

Quello del livello B1 recita invece:

è in grado di collegare una serie di elementi relativamente brevi e semplici in una sequenza lineare per punti.

È evidente che un valutatore rischia di trovarsi in seria difficoltà nel momento in cui deve determinare se l'apprendente

che sta valutando dal punto di vista della coerenza–coesione si trova a livello A2+ o a livello B1: come definire le *frasi semplici*? E come distinguerle dagli *elementi relativamente brevi e semplici*? E quali sono i *connettivi più usuali*? Curiosamente, il descrittore del livello precedente (A2) risulta invece più preciso nelle indicazioni, giacché specifica a quali indicatori linguistici fa riferimento: *È in grado di collegare gruppi di parole con connettivi semplici quali “e”, “ma” e “perché”*. Perché tale esemplificazione è presente solo per questo livello e non per i due successivi? Casi di questo tipo sono tutt’altro che isolati nel QCER (cfr. Alderson *et al.*, 2004; Ferrari, Nuzzo, 2010; Moonen *et al.*, 2013). Per rendere il QCER adatto ai bisogni del testing sarà necessario, da parte della ricerca, approfondire meglio in futuro il legame tra i diversi livelli e la varietà del repertorio di strutture linguistiche e forme d’uso (Harsh, Martin, 2012; Hulstijn *et al.*, 2012; Moonen *et al.*, 2013; Hutha *et al.*, 2014). Diversi autori sottolineano a questo proposito l’utilità ad esempio dell’analisi di corpora (cfr. Hawkey, Barker, 2004), il raffronto con le sequenze di acquisizione individuate nelle diverse lingue (cfr. Brindley, 1998) o il raffronto tra misure analitiche e olistiche (Wisniewsky, 2017a; 2017b; Harsh, Martin, 2013; Gyllstad *et al.*, 2014; Treffers–Daller *et al.*, 2018). Nel tentativo di ridimensionare i limiti sopra descritti il Manuale sottolinea in più parti l’importanza della formazione dei valutatori e della verifica della validità e coerenza delle loro interpretazioni e auspica che la valutazione degli apprendimenti linguistici sulla base del QCER diventi parte integrante del curriculum degli insegnanti: una buona conoscenza del QCER è prerequisito essenziale per il raggiungimento di un maggior grado di affidabilità nei giudizi. La formazione specifica degli insegnanti nell’uso di questo strumento dovrebbe prevedere sia una fase iniziale di familiarizzazione con il QCER sia dei momenti di esercitazione e di confronto tra valutatori (*benchmarking*), con l’obiettivo di uniformare il più possibile le



interpretazioni dei descrittori e delle performance. A questo proposito il Consiglio d'Europa ha messo a punto ulteriori materiali di supporto: oltre a una banca dati europea di esempi di performance e di item rappresentativi (cfr. [www.coe.int/portfolio](http://www.coe.int/portfolio)), e un Companion Volume (Consiglio d'Europa, 2018).

Per quel che riguarda il problema di quantificazione tre sono gli aspetti da considerare: la quantificazione dei *Can-do statements*, la quantificazione della performance, la messa in relazione delle due quantificazioni. Qual è la procedura con cui una persona può essere collocata in uno dei livelli descritti sulla base della performance, in modo che il livello assegnato corrisponda effettivamente al livello descritto nel QCER? I *Can-do statements* sono caratterizzati da una dimensione qualitativa e una quantitativa (cfr. Hulstijn, 2007). Se nell'applicazione dovessimo limitarci alla dimensione qualitativa, il risultato della valutazione sarebbe un profilo qualitativo fatto di un elenco di *Can-do statements*. Un approccio di questo tipo sarebbe sicuramente appropriato da un punto di vista teorico e psicometrico, ma poco economico da un punto di vista pratico. Diventa quindi necessario rifarsi all'aspetto quantitativo del QCER che permette di ordinare le celle di descrittori lungo un continuum. Chi utilizza le griglie avverte la necessità di sapere quale sia il livello di qualità della performance che ci si aspetta dagli apprendenti nello svolgere un determinato compito a un determinato livello. Occorrerebbe cioè che fossero chiaramente esplicitati i criteri di valutazione di *come* un apprendente sa fare qualcosa. Tali criteri dovrebbero essere strettamente connessi con i parametri contestuali e teorici che influenzano l'abilità che si sta misurando. In altre parole per applicare le griglie in modo efficace sarà necessario rispondere ad almeno due domande: qual è la percentuale di item a un determinato livello che l'apprendente deve superare, e/o qual è la porzione di descrittore che un apprendente deve soddisfare per essere assegnato a un

determinato livello? Quanti campioni scritti o orali sono necessari per assegnare un apprendente a un particolare livello?

Sulla base di osservazioni di questo tipo, diversi autori che l'hanno esaminato criticamente, come ad esempio Weir (2005b: 281–282), sostengono che il QCER nella sua forma attuale non è sufficientemente ampio, coerente e trasparente per un uso efficace nel testing linguistico. Poiché il quadro non è stato realizzato per rispondere a tutte le esigenze del testing, «it will require considerable, long-term research, much reflective test development by providers, and prolonged critical interaction between stakeholders in the field to address these deficiencies» (Weir, 2005b: 283; cfr. Alderson, 2007; North, 2014). In questo studio pertanto si cercherà di avere un approccio sperimentale all'uso del QCER, come del resto fortemente auspicato anche dal Consiglio d'Europa stesso.

#### 1.4. L'interfaccia acquisizione–testing

I due campi della linguistica acquisizionale e del testing hanno molto in comune, tuttavia diversi elementi li mantengono piuttosto distinti: il focus, la metodologia di analisi e infine i vincoli entro i quali operare. Rispetto al focus, se le ricerche di tipo acquisizionale tendono a concentrarsi su dettagli o aspetti specifici della performance con l'obiettivo di descrivere il processo di apprendimento linguistico e individuare gli elementi in grado di favorirlo, le ricerche relative al testing cercano di assumere uno sguardo più globale e ottenere gli strumenti per misurare la proficiency in modo affidabile e generalizzabile a contesti reali, sulla base di una performance elicitata; rispetto alla metodologia di analisi, se le ricerche acquisizionali hanno sviluppato strumenti descrittivi analitici e protocolli etnografici o sperimentali per l'analisi dell'interlingua, le ricerche

sul testing utilizzano principalmente procedure statistiche da applicare ai risultati di un test; e infine rispetto ai vincoli, nonostante SLA offra indicazioni utili all'insegnamento, l'applicazione pratica dei risultati non è tra gli obiettivi primari, il testing invece è più orientato alla soluzione di richieste concrete, dovendo arrivare alla produzione di test.

Diversi studiosi hanno però riconosciuto come tali differenze possano costituire il punto di partenza per sviluppare interazioni fruttuose, capaci di avvantaggiare entrambi gli ambiti. In questa sezione si presenterà la letteratura relativa all'esplorazione delle differenze e dei punti di contatto tra acquisizione e testing (§ 1.4.1.), poi si illustreranno i vantaggi di una maggiore scambio tra i due campi in particolare rispetto al tema oggetto di questo studio, ossia la descrizione e misurazione della proficiency linguistica. Si farà riferimento ai temi della validità dei costrutti, dell'affidabilità degli strumenti di misurazione e infine dello sviluppo di scale di proficiency (§ 1.4.2.).

#### 1.4.1. *L'acquisizione e il testing*

La natura della relazione tra acquisizione e testing viene discussa per la prima volta in modo specifico in un testo del 1998, dal titolo *Interfaces Between Second Language Acquisition and Language Testing Research*, curato da Lyle F. Bachman e Andrew D. Cohen. I vari contributi affrontano il tema da prospettive diverse: Carol A. Chapelle discute della definizione dei costrutti e del tema della loro validità nelle ricerche acquisizionali, Elaine Tarone esamina le ricerche sulla variazione interlinguistica e evidenzia le possibili implicazioni per il testing, Geoff Brindley si concentra sulla relazione tra scale di valutazione e ricerca acquisizionale, Dan Douglas discute il tema del metodo di valutazione nella ricerca acquisizionale di

tipo context-based, Elana Shohamy usa il caso dell'analisi del discorso per evidenziare come i due campi non tengano conto dei rispettivi risultati e infine viene ripubblicato un contributo di Lyle F. Bachman del 1988 che riassume in prospettiva generale le possibili interazioni tra i due campi. La pubblicazione, evidenziandone i diversi potenziali benefici, sostiene l'importanza di una maggiore interazione tra i due ambiti. Questa prima discussione è stata seguita da altre pubblicazioni (Shohamy, 1998; 2000; Ellis, 2001; Douglas, 2001; Laufer, 2001; Hulstijn, Schoonen, 2006; Read, 2007; Bartning *et al.*, 2010; Alanen *et al.*, 2011; Alderson, 2010; Alderson, Huhta, 2011; Alderson *et al.*, 2017) e ha portato alla costituzione di un gruppo di ricerca europea sul tema (SLATE, con il coordinamento di Jan Hulstijn). Vediamo ora in dettaglio le principali differenze e somiglianze tra i due ambiti.

È possibile identificare almeno cinque macro-differenze tra acquisizione (SLA) e testing (LT). La prima è di tipo strutturale: ricercatori di SLA e ricercatori di LT appartengono a comunità distinte, ciascuna con i suoi convegni e le sue pubblicazioni specialistiche. Pochi i ricercatori che sono membri attivi di entrambe le comunità.

La seconda riguarda invece l'orientamento generale. SLA è più legata alla teoria, in quanto mira alla costruzione o validazione di teorie in grado di spiegare il processo di apprendimento. Pertanto, nonostante la disciplina sia collegata all'insegnamento, raramente l'obiettivo è elaborare proposte didattiche o ottenere applicazioni immediate dei risultati. Al contrario il testing ha un orientamento più applicativo: la maggior parte degli studi LT sono infatti condotti entro progetti legati a test di proficiency e mirano a dimostrare che il sistema di valutazione impiegato è valido, funzionale e affidabile. La teoria in questo caso ha l'obiettivo di mettere in relazione la performance in un test con quella reale.

La terza differenza può essere riassunta dalla distinzione tra competenza e proficiency. In qualche modo SLA è più orientata alla descrizione della rappresentazione mentale, ossia della competenza e dei processi a essa collegati, mentre LT è orientato alla descrizione dell'abilità a usare tale conoscenza linguistica.

La quarta differenza riguarda gli strumenti di misurazione. I ricercatori di SLA hanno messo a punto sistemi di analisi dell'interlingua legati al conteggio delle occorrenze di determinate strutture in contesti obbligatori o misure analitiche più globali, come quelle di complessità, accuratezza e fluenza, che offrono descrizioni estremamente dettagliate della performance, a costo però di una loro applicazione piuttosto dispendiosa in termini di tempo. I ricercatori LT dovendo sviluppare o validare strumenti di misurazione da utilizzare con popolazioni numerose valutano la performance a partire da giudizi olistici, solitamente realizzati da valutatori esperti a partire da scale di proficiency. In altre parole, anche se i task di elicitazione possono essere simili, le misure applicate si differenziano in modo sostanziale.

Infine la quinta differenza riguarda aspetti più tecnici della ricerca. Gli studi SLA solitamente coinvolgono un numero limitato di partecipanti su base volontaria, mentre le ricerche LT si basano su test che hanno solitamente un impatto rilevante sulla vita dei candidati. Di conseguenza, se i ricercatori SLA possono elaborare con una certa flessibilità i loro task di elicitazione, i ricercatori LT sono più spesso legati a vincoli imposti dagli *stake-holders* che richiedono un certo tipo di valutazione.

Le differenze appena descritte sono bilanciate da una serie di possibili interazioni. Shohamy (1998) ne identificava sei: tre rappresentano possibili contributi di LT a SLA e tre di SLA a LT.

Le aree in cui LT può contribuire a SLA sono:

- a) definizione del costrutto abilità linguistica. Una componente importante delle ricerche sul testing riguarda la definizione dell'abilità linguistica, LT può offrire modelli utili per verificare la validità dei risultati;
- b) applicazione di risultati LT per testare ipotesi SLA. I ricercatori LT ottengono diversi tipi di dati circa il comportamento linguistico durante un test, analizzando ad esempio come specifiche strutture possano variare nelle condizioni di testing;
- c) scelta di criteri validi per la selezione di test e task. Un importante contributo di LT a SLA riguarda lo sviluppo di criteri per sviluppare task in grado di elicitare dati acquisizionali validi e affidabili.

Le aree in cui SLA può contribuire a LT sono:

- a) identificazione delle componenti linguistiche da elicitare e dei criteri per la loro misurazione. La ricerca SLA può identificare i costrutti necessari a chi si occupa di testing e offrire pertanto indicazioni utili per la costruzione di test. Ad esempio, le ricerche su variazione e task danno informazioni importanti sul ruolo di variabili quali interlocutore, argomento o status sociale sulla produzione, con una ricaduta importante sulla selezione dei task. Analogamente, rispetto alla definizione dei costrutti da misurare, un tester può basarsi sugli studi SLA;
- b) costruzione di task per la valutazione. I ricercatori SLA hanno utilizzato una varietà di strumenti di elicitazione. Uno dei possibili contributi di SLA a LT potrebbe riguardare proprio l'ampliamento di tale repertorio;

- c) informare i tester sui fattori che influenzano l'apprendimento. I ricercatori SLA possono suggerire ai tester possibili aree problematiche. Ad esempio, se gli studi SLA concludessero che la L1 ha un'influenza importante, allora i tester potrebbero adattare i test al background linguistico dei candidati.

Nonostante l'utilità reciproca di una proficua collaborazione, a tutt'oggi le interazioni tra i due campi sono ancora scarsamente indagate. Di qui, il tentativo, con questo studio, di considerare entrambi gli approcci nello studio della variazione. In particolare ci interessano tre possibili aree di contatto: validità dei costrutti, qualità degli strumenti di misurazione, e infine sviluppo di scale di proficiency. Nella sezione che segue si illustrano più in dettaglio questi temi.

#### 1.4.2. *I costrutti, la misurazione e le scale di proficiency*

Nelle ricerche acquisizionali gli strumenti di misurazione quantitativa e i test linguistici o i task di elicitazione hanno un ruolo piuttosto rilevante, fornendo le basi per la raccolta e la codifica dei dati. Ciononostante, come evidenziavano già Norris, Ortega (2003) nella loro presentazione della letteratura sulla misurazione in SLA, uno dei limiti principali riguarda proprio l'attenzione alle possibili fonti di errore di misurazione di tali strumenti. Questo è testimoniato dall'assenza nella maggior parte delle ricerche di dati relativi all'affidabilità degli strumenti impiegati. Questa lacuna sarebbe difficilmente comprensibile nella prospettiva LT, che controlla con attenzione l'affidabilità degli strumenti attraverso specifiche procedure statistiche.

L'affidabilità comunque è solo un aspetto di una buona misurazione. Di importanza ancora maggiore è la corri-

spondenza tra base teorica e dato empirico, ossia il ricercatore deve poter specificare cosa una data misura rileva. Il concetto di validità del costrutto ha un ruolo centrale in LT (cfr. Cronbach, Meehl, 1955; Messick, 1989; Bachman, Palmer, 1996) e riguarda proprio l'interazione tra motivazione teorica (costrutto) e evidenza empirica (punteggio nel test o altro dato). Messick (1996) mette in luce due possibili rischi: sottorappresentazione del costrutto (*construct underrepresentation*) e varianza irrilevante al costrutto (*construct-irrelevant variance*). La sottorappresentazione riguarda i casi in cui la misura applicata non richiede al candidato di utilizzare tutta la gamma di abilità o conoscenze necessarie per svolgere il task nella vita reale. La varianza irrilevante al costrutto riguarda invece i casi in cui la misura è influenzata da fattori che non hanno nulla a che vedere con ciò che si intende misurare.

Nel caso di SLA vi è rischio di sottorappresentazione dei costrutti in almeno due casi. Il primo è legato all'operativizzazione: i diversi approcci teorici SLA — generativo, interazionista e emergentista — hanno identificato costrutti rilevanti, ma senza offrire né definizione del tipo di dato empirico che può sostenere una certa interpretazione, né validazioni sperimentali (cfr. Pallotti, 2007). Il secondo caso è legato alla metodologia e all'uso di protocolli cross-sectional, che portano inevitabilmente a generalizzazioni, piuttosto che protocolli longitudinali, più affidabili nella descrizione del processo di sviluppo. Ciò è ancora più evidente per i molti casi in cui i livelli dei diversi gruppi di apprendenti da confrontare sono stabiliti in modo intuitivo o attraverso strumenti non adeguati (cfr. Norris, Ortega, 2003).

In entrambi i casi SLA può rafforzare la propria analisi e l'affidabilità dei risultati appoggiandosi alla letteratura e alle consolidate procedure di LT. Vediamo ora un'area in cui è LT



che può ricavare elementi utili da SLA: la creazione di scale di proficiency.

Il *proficiency movement*, richiedendo l'osservazione degli apprendenti in diversi contesti comunicativi, ha reso necessario lo sviluppo di scale di valutazione in grado di descrivere i diversi livelli di performance e assistere dunque i valutatori nell'applicazione coerente dei criteri di valutazione. Abbiamo già presentato in § 1.3.2. le scale di proficiency e il QCER e discusso di alcuni limiti che li caratterizzano. Le scale di proficiency nella maggior parte dei casi non si appoggiano a basi teoriche e i costrutti che esse descrivono non sono né definiti esplicitamente né validati. Quando una forma di validazione è stata applicata, come per il QCER, il punto di riferimento è dato dai giudizi degli insegnanti (cfr. North, Schneider, 1998). Inoltre la formulazione dei descrittori in termini di *Can-do statements* non fa riferimento a strutture linguistiche specifiche associate a quel determinato livello (cfr. § 1.3.2.2.). Le ricerche di SLA potrebbero certo contribuire a colmare queste due lacune.

Esempi in questo senso sono rappresentati da alcuni studi che hanno tentato di esplorare le possibili interazioni tra i due campi (Hulstijn, 2010; Hulstijn *et al.*, 2010); Brindley (1998) ad esempio, riprendendo la proposta di Pienemann e Mackey (1992) suggerisce la possibilità di applicare la Teoria della Processabilità alla valutazione della proficiency orale. Banerjee *et al.* (2007) hanno studiato la relazione tra bande di livello IELTS e features linguistiche nella produzione scritta, misurando quindi dimensioni della performance scritta quali uso di strumenti di coesione, ricchezza lessicale, complessità sintattica e accuratezza grammaticale a diversi livelli, constatando come tutte le dimensioni (a eccezione della complessità sintattica) sono informative rispetto ai progressi in proficiency. In uno studio analogo Iwashita *et al.* (2007) con-

frontano la relazione tra misure analitiche di accuratezza, complessità, vocabolario, pronuncia e fluenza e giudizi olistici in cinque diversi task del TOEFL iBT e a cinque diversi livelli di proficiency. Anche in questo caso le categorie selezionate sono significative nella distinzione tra livelli. Studi di questo tipo non possono che contribuire alla descrizione e misurazione della proficiency linguistica, mostrando l'utilità dell'applicazione in un contesto LT di misure sviluppate in ambito SLA.

Ricapitolando, chi si occupa di SLA può imparare dal testing l'importanza di una verifica di validità e affidabilità degli strumenti di elicitazione e misurazione, oltre che di una più adeguata definizione operativa dei costrutti teorici e ritrovare quindi nella letteratura LT procedure e modelli operativi da utilizzare a questi scopi. Chi si occupa di LT può migliorare i propri strumenti di descrizione della performance includendo descrizioni dettagliate di aspetti linguistici caratteristici dei diversi livelli di proficiency. In questo studio si è cercato di cogliere tali stimoli e prestare attenzione a questi aspetti. Da un lato la metodologia di ricerca include sia discussioni circa il legame tra costrutti e relative misure sia controlli sperimentali circa l'affidabilità dell'applicazione delle misure stesse. Dall'altro, attraverso una delle domande di ricerca, che riguarda proprio il confronto tra misurazioni CAF e valutazioni QCER, si cercherà di ricavare implicazioni utili per una formulazione più dettagliata di descrittori di profili di proficiency.

### **1.5. I task**

Con il diffondersi dell'approccio comunicativo i task e la loro relazione con performance e acquisizione hanno ri-

cevuto una crescente attenzione in ambito didattico, ma anche in campo acquisizionale e nel testing. La vitalità di tali ricerche è testimoniata da un numero di pubblicazioni legate appunto a task per la didattica, la ricerca e il testing (per es. Bygate *et al.*, 2001; Ellis, 2003; 2018b; Long, Crookes, 1992; Skehan, 2003; 2011a; 2011,b; van den Branden, 2006; Skehan, 2011; Long, 2015; 2016). Chiaramente il tipo di approccio ai task sviluppato in ciascuno dei tre ambiti è strettamente legato al contesto d'uso: la linguistica acquisizionale esplora i task come strumento per indagare il processo di apprendimento, la didattica come possibile attività per lo sviluppo linguistico, e infine il testing come modalità di valutazione delle competenze. In questa sezione si offrirà una definizione di task (§ 1.5.1.) e si discuterà più in dettaglio dell'uso del task nei tre ambiti: didattica, ricerca e testing (§ 1.5.2.).

### 1.5.1. *La definizione di task*

I task sono interpretati e definiti in modo diverso dai vari gruppi di persone che li utilizzano. Lo scopo di questa sezione è illustrarne le diverse interpretazioni e definizioni per poi selezionare quella utile per questo studio.

I task sono utilizzati in almeno tre contesti: didattica, ricerca e testing. Nel primo caso l'enfasi è rivolta all'efficacia nell'uso in classe, nel secondo invece è posta su ciò che accade all'apprendente durante lo svolgimento di un task (o di una serie di task) o sui cambiamenti che avvengono nell'interlingua e infine l'ultimo indaga la natura della performance elicitata. La tabella 1.7. proposta da Bygate *et al.* (2001) illustra i tre possibili contesti d'uso dei task e mostra come in tutti i casi l'interesse possa essere di tipo sia pedagogico che sperimentale.

**Tabella 1.7.** Lo studio dei task, da Bygate, Skehan e Swain, 2001.

	pragmatic/pedagogic	research
teachers and teaching	<ul style="list-style-type: none"> <li>– task as a unit of work in a scheme of work</li> <li>– interlinked activity sequences developing thematic unit</li> <li>– methods of involving learners</li> <li>– deliberate starting point for unknown direction or explorations</li> </ul>	<ul style="list-style-type: none"> <li>– task as researchable unit</li> <li>– neat, cross sectional approach</li> <li>– relatively brief time interventions</li> <li>– focus on the isolation of variables</li> <li>– search for “effects” through manipulation</li> </ul>
learners and learning	<ul style="list-style-type: none"> <li>– learners orientation and autonomy</li> <li>– task reinterpretability</li> <li>– interactive development through collaboration of groups of learners</li> <li>– authenticity of response</li> </ul>	<ul style="list-style-type: none"> <li>– extent to which learning processes are catalysed</li> <li>– identification of theorised methods of operationalising constructs and measuring dependent variables</li> <li>– research design to probe:               <ul style="list-style-type: none"> <li>– salient task variables</li> <li>– salient task conditions</li> </ul> </li> </ul>
testing	<ul style="list-style-type: none"> <li>– formative evaluation</li> <li>– provision of structured feedback on communication</li> <li>– reactive, unstandardised and individual based</li> </ul>	<ul style="list-style-type: none"> <li>– summative evaluation</li> <li>– task as a format</li> <li>– comparability and standardisation</li> <li>– issues in performance assessment</li> </ul>

Sulla base del contesto in cui viene affrontata — insegnamento, ricerca o testing — la definizione di task può essere elaborata così da mettere in luce aspetti diversi. In ambito didattico i task possono essere intesi sia come attività contenute (Nunan, 1989) o sequenze più lunghe e composite che occupano diverse lezioni (ad es. il *task cycle* proposto da Willis, 1996) o addirittura temi da cui partire per sviluppare una serie di attività (ad es. *project work*). In ogni caso l’obiettivo è quello di produrre attività pedagogiche efficaci. Nel caso

della ricerca acquisizionale invece il task è visto come attività che mira a elicitare dati utili per il ricercatore, l'interesse è sempre motivato da domande di tipo teorico. Di conseguenza di estrema rilevanza saranno aspetti quali standardizzazione e validità ecologica. In un approccio quantitativo si cercherà di ottenere valutazioni statistiche circa la robustezza dei dati raccolti e in un approccio più qualitativo si condurranno analisi a partire dalle trascrizioni delle interazioni. In entrambi i casi l'obiettivo è proporre metodi d'indagine che permettano lo studio del processo di apprendimento in modo affidabile e replicabile. Infine nell'ambito del testing l'attenzione è rivolta a come rendere il task uno strumento valido, affidabile e praticabile ai fini della valutazione della proficiency. L'obiettivo qui è l'esplorazione di come funzionano i task e quali influenze sistematiche hanno sulla produzione linguistica.

Nei tre diversi contesti sono state proposte varie definizioni di task, le principali sono riportate qui sotto. Alcune di queste sono più ampie ed inclusive, altre definiscono il task secondo criteri più restrittivi. Un elemento di distinzione riguarda la definizione del legame tra task e comunicazione autentica: in alcuni casi l'autenticità è definita attraverso la corrispondenza con attività quotidiane (per es. Long, 1985a), in altri è data dalla natura della risposta dell'apprendente e dalla sua somiglianza, diretta o indiretta, con la lingua usata in contesti di comunicazione autentica (per es. Skehan, 1998).

A task is a piece of work undertaken for oneself or for others, freely or for some reward. Thus examples of tasks include painting a fence, dressing a child ... In other words, by "task" is meant the hundred and one things people do in everyday life, at work, at play, and in between. (Long, 1985a)

A piece of work or an activity, usually with a specific objective, undertaken as part of an educational course, at work, or used to elicit data for research. (Crookes, 1986)

One of a set of differentiated, sequencable, problem-posing activities involving learners and teachers in some joint selection from a range of varied cognitive and communicative procedures applied to existing and new knowledge in the collective exploration and pursuance of foreseen or emergent goals within a social milieu. (Candlin, 1987)

An activity which required learners to arrive at an outcome from given information through some process of thought and which allowed teachers to control and regulate that process was regarded as “task”. (Prabhu, 1987)

Any structured language learning endeavour which has a particular objective, appropriate content, a specified working procedure, and a range of outcomes for those who undertake the task. “Task” is therefore assumed to refer to a range of workplans which have the overall purpose of facilitating language learning — from the simple and brief exercise type, to more complex and lengthy activities such as group problem-solving or simulations and decision-making. (Breen, 1987)

A piece of classroom work which involves learners in comprehending, manipulating, producing or interacting in the target language while their attention is principally focused on meaning rather than on form. (Nunan, 1989)

A task [is] any activity in which a person engages, given an appropriate setting, in order to achieve a specifiable class of objectives. (Carroll, 1993)

We define a *language use task* as an activity that involves individuals in using language for the purpose of achieving a particular goal or objective in a particular situation. (Bachman, Palmer, 1996)

Tasks are always activities where the target language is used by the learner for a communicative purpose (goal) in order to achieve an outcome. (Willis, 1996)

A task is an activity in which

- meaning is primary
- learners are not given other people's meanings to regurgitate
- there is some sort of relationship to comparable real-world activities
- task completion has some priority
- the assessment of the task is in term of the outcome. (Skehan, 1998)

Selezionare una definizione appropriata di task, significa rifarsi all'uso per cui esso è inteso. Poiché in questo studio l'obiettivo è misurare la variabilità della performance, ci si riferirà al task riprendendo la seguente definizione proposta da Bygate *et al.* (2001:12).

A task is a contextualised, standardised activity, relatable to learner choice or to learning processes, which requires learners to use language, with emphasis on meaning, to attain an objective, and which elicits data which may be the basis for research.

In questo studio il task è usato come strumento di elicitazione in grado di favorire un'osservazione replicabile ed ecologicamente valida del processo di apprendimento. L'analisi dei dati mira però ad offrire elementi di riflessione utili anche a chi utilizza i task per scopi di insegnamento o testing.

### 1.5.2. *I task per insegnamento, SLA e testing*

L'approccio comunicativo è una sorta di *post-method approach* all'insegnamento (Brumfit, Johnson, 1979; Brumfit, 1984; 1988), che favorisce l'uso di una varietà di tecniche, piuttosto che offrire all'insegnante pacchetti preconfezionati. Riprendendo idee sviluppate in ambito pragmatico, sociolinguistico o dell'analisi del discorso, pone al centro dell'attenzione l'importanza di un uso contestualizzato e autentico della lingua nella classe di L2 (Brumfit, Johnson, 1979). A partire dagli anni '70 si sono sviluppate dunque diverse proposte didattiche in questo senso. Le prime interpretazioni hanno portato alla creazione di attività in grado di promuovere un uso più autentico della lingua, quali *information gap activities* (Harmer, 1991) o *jigsaw activities* (Geddes, Sturtridge, 1978; per una discussione più recente vedi Wesche, Skehan, 2002). Inizialmente ci si riferiva a queste attività con il termine generico "attività comunicative", poi sempre più spesso si è utilizzata la parola *task*. Questi sviluppi pedagogici hanno influenzato notevolmente la glottodidattica, portando sia all'elaborazione di diverse proposte di approcci didattici o sillabi basati interamente sul *task* (per es. Long, Robinson, 1998; Nunan, 1989; Prabhu, 1997; Willis, Willis, 1988; 2013; Adams, 2009; van den Branden, 2006; Walker *et al.* 2018), sia di procedure di valutazione (per es. Morrow, 1979; Long, Norris, 2000; Norris, 2002; 2009; 2016; Brindley, 2013; Purpura, Turner, 2014). Contemporaneamente la ricerca acquisizionale si è interessata al *task* come strumento di elicitazione, oltre che costruito pedagogico da indagare (Ellis 2003; Long, 2015; Skehan, 2011; Bygate, 2015; Plonsky, Kim, 2016).

Un approccio all'insegnamento basato sul *task* parte dal presupposto che una lingua si apprende per comunicare, pertanto poiché il significato è primario rispetto alla forma, le attività



devono essere in grado di favorire un'interazione più autentica. Il punto di criticità di questo approccio però è che attirando l'attenzione verso il significato, si rischia di distrarre l'apprendente dalla forma. La sfida dunque è riuscire a stimolare una quantità sufficiente di *focus on form* così da permettere lo sviluppo del sistema interlinguistico senza compromettere la naturalezza dello scambio comunicativo generato dal task (cfr. Long, 2016; Ellis, 2018b). Ottenere ciò significa sviluppare gli strumenti per comprendere meglio almeno tre elementi: l'impatto del task design e delle condizioni di somministrazione sulla performance, l'impatto della selezione dei task sull'uso linguistico e infine le relazioni tra task e fattori di processing psicolinguistico. Rispetto al primo elemento, la priorità è l'identificazione delle caratteristiche di un task in grado di avere un impatto positivo sulla qualità e complessità del processing linguistico. Il secondo elemento ha a che vedere con l'impatto della selezione dei task sull'uso linguistico, qui l'attenzione è posta sulle modalità con cui la performance porta a cambiamenti nella proficiency. Infine, il terzo elemento si rifà alla relazione tra task e fattori di processing, includendo pertanto temi come la natura del processo di comprensione, i modi in cui l'interazione porta l'apprendente a concentrarsi sulle relazioni forma-significato o come le caratteristiche dei task possano influenzare l'elaborazione linguistica.

L'interesse pedagogico per i task ha pertanto stimolato ricerche di vario tipo, raggruppabili almeno secondo quattro approcci principali:

- a) l'approccio interazionale, influenzato principalmente dal lavoro di Long (1985b; 1989);
- b) l'approccio socioculturale, rappresentato dal lavoro di Lantolf (2000) e Swain (Swain, 1998; Swain, Lapkin, 2001);

- c) l'approccio focalizzato all'elicitazione di determinate strutture (*structure focused approach*), in cui i task sono impiegati per stimolare la produzione di forme linguistiche specifiche (Loschky, Bley-Vroman, 1993; Van Patten, 1996);
- d) l'approccio cognitivo di Skehan (per es. Skehan, 1998; 2014; 2018) e Robinson (per es. 2001a; 2001b; 2011; 2015).

A questo proposito è bene evidenziare come le ricerche basate sui task per l'insegnamento condotte fino ad oggi risentono di un limite: raramente inquadrano i task entro un contesto pedagogico più ampio, più frequentemente li studiano a prescindere dal contesto classe. Se questo ha permesso di approfondire le conoscenze rispetto allo sviluppo delle L2, in futuro sarà necessario porre maggior attenzione ai contesti pedagogici d'uso (Bygate *et al.*, 1998; Samuda, 2001; Ellis, 2012).

La ricerca si è interessata ai task non solo come costrutto, ma anche come strumento di indagine. In questo secondo caso si affronta il problema di come il task possa essere impiegato per osservare in modo ecologicamente più valido il processo di apprendimento (Plonsky, 2015; Plonsky, Kim, 2016). L'elicitazione dei dati e la loro analisi sono il punto di interesse principale e la selezione dei task è strettamente legata alle domande di ricerca. In questo caso vengono definiti task attività di vario tipo, in molti casi riferendosi a varie tecniche di elicitazione semi-spontanea (cfr. Gass, Mackey, 2007).

Anche nell'ambito del testing, in modo simile alla ricerca, l'interesse è rivolto verso task intesi come attività ben definite e limitate che rappresentino interazioni comunicative con una corrispondenza con i bisogni linguistici dell'apprendente, che possano essere valutate in modo valido e coerente e quindi raggiungere un certo grado di standardizzazione. Come sottolineano anche Bachman, Palmer (1996), se le caratteri-

stiche di un task influenzano il risultato del test, allora diventa essenziale capire gli effetti di tale influenza (Walker *et al.*, 2018). L'interesse è dunque rivolto in primo luogo alle diverse caratteristiche dei task che possono avere impatti significativi sulla performance, ed essere pertanto elementi fondamentali per l'interpretazione del test e la sua standardizzazione. La validazione d'altronde è un aspetto chiave nel testing (cfr. 1.3.). Si è già visto come i risultati ottenuti da un test possono infatti essere influenzati da una serie di elementi: valutatore e criteri di valutazione, condizioni di somministrazione e caratteristiche dei task. Le ricerche condotte in questo ambito hanno pertanto studiato fattori quali valutatori e scale di valutazione (per es. Lumley, McNamara, 1995; North, 1991; 2000), o gli effetti delle relazioni di potere tra partecipanti (Van Lier, 1989) o le difficoltà di un task (Skehan, 2001). Quest'ultimo aspetto è stato affrontato secondo due punti di vista: utilizzando scale di valutazione (e.g. Wigglesworth, 1997; 2001; Chalhoub–Deville, 2001) o misure oggettive di sviluppo CAF (e.g. Skehan, 1998; Robinson, 2001). Ricerche di questo tipo possono offrire indicazioni estremamente utili a chi utilizza i task ai fini della valutazione.

Interessante notare come l'attenzione per la competenza comunicativa negli anni '70, il *proficiency movement* degli anni '80 e i più recenti stimoli nei confronti del *performance testing* abbiano portato chi si occupa di testing a condividere sempre più elementi con la didattica: il task rappresenta un esempio di questa maggior connessione tra valutazione e insegnamento (per es. Alderson, 2005a; 2005b; Lee, 2015; Alderson *et al.*, 2017). Dal punto di vista della metodologia di ricerca inoltre, le domande poste dall'impiego dei task richiedono una maggior interazione tra testing e ricerche acquisizionali (cfr. § 1.4.; Alderson, 2010; Hulstijn, 2011). In questo studio si è cercato di coniugare le due prospettive, il task è qui usato

come strumento di elicitazione, con l'obiettivo però di offrire elementi utili sia rispetto al tema della selezione di task per l'insegnamento o il testing, sia rispetto a quello della selezione di criteri affidabili per la misurazione della performance.

## Metodologia di ricerca

Questo capitolo presenta prima i soggetti dello studio, le modalità di raccolta dati e le caratteristiche del corpus, poi illustra le unità impiegate per l'analisi della produzione linguistica.

I dati considerati provengono dal corpus V.I.P., *Variabilità dell'Interlingua Parlata* realizzato con l'obiettivo di indagare la variabilità situazionale dell'interlingua di apprendenti avanzati di italiano L2 (Pallotti *et al.* 2010). La raccolta è stata realizzata nell'ambito di due progetti di ricerca: COFIN 2003, co-finanziato dal MIUR e dall'Università di Verona (coordinatore locale Camilla Bettoni), e COFIN 2006, co-finanziato dal MIUR e dall'Università di Modena e Reggio Emilia (coordinatore locale Gabriele Pallotti), con il coordinamento in entrambi i casi a livello nazionale di Giuliano Bernini.

### 2.1. Gli obiettivi generali della ricerca

Il presente studio si propone da un lato di investigare in prospettiva acquisizionale la variabilità situazionale e longitudinale dell'interlingua di apprendenti avanzati impegnati nello svolgimento di task comunicativi, e dall'altro di verificare sperimentalmente con prospettiva legata al testing le relazioni che intercorrono tra due tipi di misurazione: quella analitica CAF che rileva la complessità, l'accuratezza e la fluenza della produ-

zione orale, impiegata nei cosiddetti *developmental index studies* di stampo acquisizionale, e quella soggettiva basata sui descrittori del QCER, maggiormente legata all'ambito del testing.

Questa ricerca ha un carattere essenzialmente esplorativo, nel senso che si pone l'obiettivo di rispondere ad alcune domande-guida facendo scaturire eventuali ipotesi piuttosto che di verificare ipotesi già formulate e/o di convalidare teorie specifiche. In particolare attraverso l'analisi quantitativa e qualitativa del parlato di apprendenti a livello avanzato confrontato con quello di parlanti nativi si cercherà di rispondere alle seguenti domande:

- Come variano complessità, accuratezza e fluenza CAF nel corso del tempo man mano che progredisce l'interlingua? Come variano con il variare dell'attività comunicativa per gli apprendenti rispetto ai parlanti nativi? Come si sviluppa il percorso individuale di ciascun soggetto? Quali sono le interazioni tra queste tre dimensioni dell'interlingua?
- I descrittori del QCER sono efficaci nella valutazione delle competenze orali di apprendenti di italiano L2 di livello avanzato? In particolare sono sensibili allo sviluppo longitudinale e alla variazione tra task?
- Quali relazioni intercorrono tra le misure analitiche CAF e le valutazioni soggettive della proficiency basate sul QCER?

## 2.2. I soggetti

I soggetti presi in considerazione per questo studio sono sei. Si tratta di quattro apprendenti di italiano L2 a livello avanzato e due studentesse italiane. Al momento della prima rileva-

**Tabella 2.1.** I Soggetti.

	Pandita	Catherine	Eden	Shirley	Elisa	Valentina
codice	PAN	CAT	EDE	SHI	ELI	VAL
paese d'origine	India	Ghana	Eritrea	Nigeria	Italia	Italia
L1	punjabi	twi	tigrigno	inglese	italiano	italiano
età al t1	17	19	19	15	15	15
anni in Italia al t1	4	6	6	6		
livello QCER al t1*	B1	B1	B2	B2	C1	C1

\* La valutazione del livello di competenza al t1 per i 6 soggetti è stata effettuata dall'equipe di ricerca COFIN utilizzando i descrittori globali del QCER

zione nel 2005 tutte e sei sono iscritte a istituti professionali, e hanno un'età compresa tra i 15 e i 19 anni. Le apprendenti risiedono in piccoli comuni della provincia tra Modena e Bologna, sono di diversa nazionalità e hanno frequentato con buon rendimento la scuola italiana per un periodo di 4–6 anni. Nei primi mesi in Italia tutte avevano partecipato nelle rispettive scuole di arrivo a corsi di italiano L2 che hanno visto il ricercatore coinvolto come docente. La conoscenza e il rapporto di familiarità con i soggetti ha favorito la partecipazione attiva alla ricerca nel corso dei tre anni. La tabella 2.1. riassume i dati dei sei soggetti. I paragrafi che seguono riportano invece un profilo delle apprendenti e delle studentesse italiane ricostruito sulla base della conoscenza diretta dei soggetti e dei racconti emersi durante le interviste.

Pandita, 17 anni al momento della prima raccolta dati, è arrivata dall'India nell'estate del 2001 insieme alla sorella, al fratello e alla madre per ricongiungersi con il padre, che viveva già in provincia di Modena da una quindicina d'anni. Al suo arrivo Pandita viene iscritta alla scuola media del comune di residenza e le viene richiesto di ripetere la classe terza, appena conclusa in India. Frequenta per tutto il primo anno scolastico corsi specifici di italiano L2 e gli insegnanti realizzano diversi progetti per facilitare la sua partecipazione alle

lezioni disciplinari. In generale Pandita ricorda positivamente l'impatto con gli insegnanti e con la scuola, nonostante le difficoltà oggettive dell'apprendimento dell'italiano e dell'incontro con la nuova realtà. Nella fase iniziale segnala come poco significativi i rapporti con i compagni, giustificando tale situazione sulla base delle difficoltà di comunicazione. Dopo aver superato con successo l'esame di licenza media, Pandita su consiglio degli insegnanti si iscrive ad un Istituto Tecnico Industriale a Modena. Frequenta la scuola con successo, anche se non nasconde le difficoltà affrontate nei primi anni: la lingua per lo studio continua a richiedere impegno e i rapporti con i compagni non sono molto intimi. Al momento della prima registrazione Pandita frequenta il terzo anno, in questa fase si sente più sicura rispetto alla lingua italiana e più autonoma nello studio, racconta inoltre di aver allacciato diversi rapporti con i compagni di scuola e valuta la sua esperienza scolastica come soddisfacente. Il recente viaggio in India sembra aver dato motivazioni nuove rispetto alla permanenza in Italia e Pandita si vede orientata verso una continuazione degli studi. Gli insegnanti la descrivono come una studentessa diligente e precisa. Al t4 frequenta con soddisfazione e successo il primo anno di università.

Catherine, 19 anni al momento della prima raccolta dati, è arrivata dal Ghana all'inizio del 1999 per ricongiungersi con il resto della famiglia che viveva già da una quindicina d'anni in Italia. Catherine è la più piccola di cinque fratelli. Al suo arrivo viene iscritta alla scuola media, per età avrebbe dovuto frequentare la terza, ma viene inserita in prima. I primi mesi di scuola frequenta un laboratorio di italiano L2. Catherine ricorda il primo anno come estremamente difficile, l'impatto con la nuova realtà e in particolare la lingua italiana non sono stati del tutto positivi. Nel giro di un paio d'anni riesce però a superare le difficoltà iniziali e comincia ad interagire



più efficacemente con insegnanti e compagni. Al momento della prima registrazione, frequenta la terza superiore di un Istituto Professionale ad indirizzo aziendale della provincia di Bologna. In questa fase giudica positivamente la sua partecipazione alla vita scolastica, segue nelle diverse discipline percorsi facilitati e dedica molto tempo allo studio. Catherine è piuttosto soddisfatta della sua vita in Italia, dove vede il suo futuro. A t<sub>4</sub> frequenta con discreto successo il primo anno di università.

Eden, 19 anni al momento della prima raccolta dati, è arrivata dall'Eritrea nel 1999 insieme alla sorella per raggiungere la madre, che viveva da una decina d'anni in Italia. Fin da subito l'arrivo in Italia viene vissuto come rottura con il passato: la madre richiede l'uso della lingua italiana in casa e spinge entrambe le figlie a inserirsi il più efficacemente possibile a scuola e nel territorio. Al suo arrivo Eden viene iscritta in prima media, anche se per età avrebbe dovuto frequentare la terza. Ricorda i primi anni di scuola come difficili, l'apprendimento della lingua italiana è stato inizialmente lento e impegnativo e i rapporti con i compagni non sono stati sempre facili, in alcuni casi addirittura conflittuali. Per Eden ha sempre costituito una grossa difficoltà la differenza di età, e di conseguenza di interessi e obiettivi, che la distingue dai compagni. Questo elemento ha caratterizzato l'intero percorso scolastico. A t<sub>1</sub> Eden frequenta la classe terza di un Istituto Professionale ad indirizzo aziendale della provincia di Bologna. Gli insegnanti la descrivono come una studentessa seria e capace. Tra t<sub>2</sub> e t<sub>3</sub> Eden accetta una proposta di lavoro e continua la scuola seguendo i corsi serali. Il cambiamento è vissuto in modo estremamente positivo, in particolare rispetto al rapporto con compagni e colleghi. A t<sub>4</sub> Eden ha conseguito il diploma e continua con soddisfazione la sua attività lavorativa.

Shirley ha 15 anni alla prima raccolta dati. È arrivata dalla Nigeria all'inizio del 1999 insieme alla sorella, per riunirsi con la famiglia, genitori e due fratelli, che risiedevano in Italia da più di 12 anni. Al suo arrivo viene inserita in seconda media. Durante i primi due anni di scuola frequenta corsi di italiano L2 e gli insegnanti di classe organizzano diversi progetti educativi per favorire il suo inserimento. Shirley ricorda positivamente il primo periodo a scuola, in particolare il rapporto con gli insegnanti. Terminata la scuola media si iscrive ad un Istituto Professionale ad indirizzo turistico di Modena: è contenta della scelta fatta e partecipa con entusiasmo a stage o proposte formative extracurricolari offerte dalla scuola. A t1 Shirley frequenta la classe terza con buon rendimento, ha ottimi rapporti con i compagni e una buona rete di relazioni sociali. Frequenta assiduamente la chiesa e le attività della comunità ghanese di Modena. A t4 Shirley frequenta con successo il primo anno di università.

Elisa al momento della prima registrazione ha 15 anni ed è iscritta al secondo anno di un Istituto Professionale ad indirizzo aziendale. Vive in provincia di Bologna con i genitori ed un cagnolino. La scelta della scuola superiore si è basata più sulla posizione geografica della scuola che sulle reali propensioni di Elisa, che avrebbe preferito seguire un percorso di studi più legato alle arti. Nonostante questo il rendimento scolastico è discreto. Elisa ha gusti e preferenze molto precisi, sia nella scelta degli amici che delle attività da svolgere. Nel tempo libero si dedica alla musica, è inserita in una larga rete sociale di relazioni e si ritiene piuttosto soddisfatta di quello che fa. Gli insegnanti la descrivono come una studentessa riservata, ma capace.

Valentina al momento della prima registrazione ha 15 anni e frequenta il secondo anno di un Istituto Professionale ad indirizzo aziendale. Vive in provincia di Bologna con i genitori

e un fratello più piccolo. L'esperienza scolastica di Valentina, nonostante risultati discreti, non è del tutto positiva. Gli insegnanti la descrivono come una studentessa capace, ma svogliata e demotivata. Vedremo che anche nella partecipazione a questo studio, la seconda rilevazione della sua produzione linguistica non è esente da problemi.

### 2.3. Le rilevazioni

I dati impiegati in questo studio sono stati raccolti utilizzando un protocollo di ricerca sperimentale per l'elicitazione di parlato semi-spontaneo elaborato dal gruppo di ricerca coinvolto nei due progetti COFIN 2003 e COFIN 2006 (Pallotti *et al.*, 2010).

Tale protocollo è costituito di due momenti, indicati come giornata 1 e giornata 2. I task impiegati nella giornata 1 si rifanno a procedure tradizionalmente usate nella ricerca sull'acquisizione nella seconda lingua e coinvolgono l'intervistatore e un apprendente; i task utilizzati nella giornata 2 sono task comunicativi inseriti all'interno di un gioco a coppie a cui partecipano l'intervistatore, due apprendenti e diversi esperti adolescenti o adulti. La giornata 1 è costituita di quattro attività. La prima è un'intervista informale tra intervistatore e apprendente che segue una traccia tematica, anche se con una certa flessibilità. La seconda attività è il task *il racconto di un film*. La terza attività è il *map task* con adulti (Pallotti *et. al.*, 2014). In questo caso si è selezionato quello utilizzato nel progetto API (Archivio del Parlato Italiano, [www.cirass.unina.it](http://www.cirass.unina.it)). L'ultima attività è un task di *racconto di storia per immagini*. La giornata 2 è organizzata in tre task comunicativi: il *map task* con compagno, e due task in cui le studentesse per svolgere il compito assegnato

— ad esempio, selezionare il miglior cellulare da un catalogo, organizzare la gita di classe, scegliere un libro o un dvd da regalare — devono raccogliere informazioni consultando telefonicamente esperti e negozi per concordare poi insieme una possibile soluzione.

Le produzioni delle apprendenti sono state registrate in quattro rilevazioni: t1 nel 2005, t2 nel 2006, t3 nel 2007, e infine t4 nel 2008. In t1 e t4 sono stati utilizzati gli stessi task, in t2 e t3 sono stati utilizzati task confrontabili. Le studentesse italiane sono state osservate invece in due rilevazioni: t1 nel 2005 e t2 nel 2007.

#### 2.4. I task

Il corpus di questo studio è costituito da quattro task: due task monologici, il racconto di un film e il racconto di una storia ad immagini, e due task interattivi, l'intervista e l'apertura delle telefonate di servizio, realizzati in quattro registrazioni effettuate nell'arco di tre anni.

Nel task racconto di un film, alle studentesse viene richiesto di guardare un video, della durata di 10 minuti circa e raccontarne poi la storia all'intervistatore, che non ne conosce il contenuto. Le studentesse guardano l'estratto una sola volta. Il materiale non è loro familiare e non è stato imposto un limite di tempo. Per questo task sono stati utilizzati il breve estratto *Alone and hungry* tratto dal film *Modern Times* di Charlie Chaplin in t1, due episodi della serie animata *Pink Panther* intitolati *Slink Pink* e *The Island* in t2 e t3 e nuovamente l'estratto dal video di *Modern Times* in t4.

Nel task racconto di una storia ad immagini, alle studentesse viene richiesto di narrare la storia all'intervistatore. Il materiale non è loro familiare e anche in questo caso non è

posto un limite di tempo. In t<sub>1</sub> è stato utilizzato il racconto ad immagini *Frog, where are you?* di Mayer, in t<sub>2</sub> il racconto *Ho trovato un pettirosso* di Blanch, in t<sub>3</sub> *One frog too many* di Mayer, e in t<sub>4</sub> nuovamente il racconto *Frog, where are you?*.

Il task intervista consiste in una conversazione informale. Sono stati considerati due estratti per ciascun soggetto. Nel primo le studentesse parlano di se stesse e della propria famiglia, nel secondo raccontano del loro primo giorno di scuola in Italia o all'istituto professionale, o ancora dell'esperienza all'università o al lavoro. Il contenuto della conversazione viene mantenuto il più possibile costante tra t<sub>1</sub> e t<sub>4</sub>. In ogni rilevazione però è stata coinvolta come intervistatore una persona diversa, non conosciuta dalle ragazze. Questa è sempre un'insegnante di italiano L2, di età compresa tra i 34 e i 38 anni, con una certa esperienza in attività di questo tipo. In questo modo le studentesse non hanno percepito il task come ripetitivo, ed è stato possibile mantenere sia un buon grado di spontaneità sia un buon livello di comparabilità tra rilevazioni.

L'apertura delle telefonate di servizio è un task incluso in diversi task comunicativi più ampi del corpus che prevedono, tra l'altro, una serie di telefonate per raccogliere informazioni ed accordarsi sulla scelta del miglior cellulare, libro, cd musicale o film e sull'organizzazione di una gita scolastica o un viaggio per sé o altri. Questi task sono dunque a soluzione aperta: per svolgere il compito assegnato le apprendenti devono raccogliere informazioni consultando telefonicamente varie persone sconosciute, tra cui coetanei, adulti e personale di negozi specializzati e concordare infine la soluzione ottimale. In questo studio sono state considerate da cinque a sette telefonate per soggetto, analizzandone esclusivamente l'apertura e l'atto di richiesta perché rappresentano eventi comunicativi relativamente uniformi e nel quale le partecipanti producono turni di una certa lunghezza.

**Tabella 2.2.** Il Corpus.

	<i>apprendenti italiane</i>				<i>totale</i>		
	t1	t2	t3	t4	t1	t2	
	parole						
film	1982	1480	1562	2955	844	552	9375
storia	2209	1869	1846	2595	1057	661	10237
intervista	3014	3238	3946	1290	1254	662	13404
telefonata	361	461	614	691	332	235	2694
totale	7566	7048	7968	7531	3487	2110	35710
	AS-unit						
film	362	253	248	515	124	75	1577
storia	419	312	281	420	167	91	1690
intervista	650	630	735	225	248	259	2747
telefonata	95	111	159	169	73	43	650
totale	1526	1306	1423	1329	612	468	6664

## 2.5. I dati

Le produzioni realizzate dai sei soggetti nei quattro task eseguiti per quattro volte dalle apprendenti e due dalle italiane, sono state videoregistrate, trascritte e segmentate in AS-unit, l'unità di analisi del parlato ritenuta più adatta agli scopi dello studio (cfr. § 1.4.1.). Il corpus è pertanto costituito da 35.710 parole, equivalenti a 6.664 AS-unit (cfr. tabella 2.2.).

Per la trascrizione si è fatto riferimento al sistema jeffersiano (Atkinson, Heritage, 1984) e a CHAT ([www.childes.psy.cmu.edu](http://www.childes.psy.cmu.edu)), con alcuni adattamenti necessari per il trattamento informatico dei dati: £parola£ indica una parola pronunciata ridendo; .hhh e hhh indicano rispettivamente inspirazione ed espirazione; paro:::la segnala un prolungamento di suono; f\*parola\*f e s\*parola\*s segnalano rispettivamente parlato veloce e parlato lento; il punto fermo indica intonazione finale

discendente, la virgola intonazione finale sospesa ed il punto interrogativo intonazione finale ascendente; # segnala le micropause, #2\_o indica la durata della pausa, in questo caso 2 secondi.

Per la rappresentazione della segmentazione in unità di produzione, le AS-unit sono contenute tra barre verticali doppie (|| xxxxx ||), mentre le clausole o unità inferiori alla clausola sono racchiuse tra barre verticali semplici (| xxxxx |). Infine, false partenze o ripetizioni vengono racchiuse tra parentesi graffe.

## 2.6. L'unità e le misure dell'analisi

Questa sezione illustra gli strumenti di misurazione utilizzati per l'analisi dei dati, prima quelli del CAF (§ 2.6.1.), poi quelli del QCER (§ 2.6.2.). Poiché qualsiasi strumento di misurazione, oltre che valido, deve essere anche affidabile, la presentazione dettagliata delle unità di analisi è accompagnata quando possibile da un controllo empirico della loro affidabilità.

### 2.6.1. *Le misure CAF*

Per misurare CAF nelle produzioni sono state proposte diverse misure quantitative (vedi § 1.2.2.2.). In questa sezione si illustrano quelle scelte per questo studio e se ne giustifica l'utilità. Dapprima si presenta l'unità di produzione, specificando gli adattamenti resi necessari per la sua applicazione a dati in italiano L1 e L2 (cfr. Ferrari, in stampa), poi si presentano le misure quantitative CAF.

Foster *et al.* (2000) definiscono l'AS-unit, ovvero l'*Analysis of Speech Unit* come "a single speaker's utterance consisting of an *independent clause*, or *sub-clausal unit*, together with any

*subordinate clause(s)* associated with either.” In (2) ne diamo un esempio:

- (2) SH1: || | e::: la pantera? #o\_4 {e::;} ha sempre fatto di tutto.  
 | in modo che::: {il ca} il cacciatore s:::incavoli con e:::  
 il cane. | || #o\_2

Questi autori distinguono poi le clausole che costituiscono le AS-unit in:

- clausole indipendenti
- unità indipendenti inferiori alla clausola
- clausole subordinate.

Una clausola indipendente è definita come «a clause including a finite verb»:

- (3) ELI: || | e::: {va::;} ruba {un pezzo di legno} # insomma del  
 fuoco::: dalla casa del cacciatore # | || o\_6.hhh

Un'unità indipendente inferiore alla clausola «will consist of either one or more phrases which can be elaborated to a full clause by means of recovery of ellipted elements from the context of the discourse or situation [...] or a minor utterance, which will be defined as one of the class of *Irregular sentences* or non sentences identified by Quirk *et al.*»:

- (4) SH2: sì  
 SH1: || | allora Danielle Steel | ||  
 SH2: sì  
 SH1: || | il ranch | ||  
 SH2: vediamo



Una clausola subordinata «will consist minimally of a finite or non finite verb element plus at least one other clause element (subject, object, complement or adverbial)»:

- (5) PAN: || | c'è un uomo | che dice | di prendere una: # legna  
 {n} come {co;} co:no | ||

Le false partenze, le ripetizioni e le autocorrezioni vengono escluse dalla segmentazione in AS-unit e pertanto figurano tra parentesi nella trascrizione. Con falsa partenza Foster *et al.* si riferiscono sia ai casi in cui il piano discorsivo viene abbandonato sia ai casi di riformulazione. Se una AS-unit o una clausola completa è prodotta prima che il piano venga abbandonato, essa viene invece inclusa nel conteggio:

- (6) ELI: || | lui s- muove un: e::: | || #I\_6  
 ELI: || | non lo so. | ||  
 ELI: || | comunque lui lavora in un cantiere tipo | ||

Nel caso di ripetizioni si distingue tra ripetizione come disfluenza (7), quindi esclusa dall'analisi, e ripetizione come strategia retorica (8), quindi inclusa nel conteggio:

- (7) PAN: || | e- adesso quando {il >polisiotto} anche il polisiotto < è alzato? | || #o\_2  
 PAN: || | ha dato una #o\_4 botta nella testa? | ||  
 INV: ℓhhℓ  
 PAN: || | .h poi {sono cs scappati? #o\_3 e: #o\_6} s:ono scappati | ||  
 PAN: || | e sono .h seduti una vicina casa? | ||  
 (8) PAN: || | poi Charlot- entra in <caffettiera> | ||  
 PAN: || | e mangia tutto tutto | ||  
 PAN: || | non paga niente? | ||

Si parla di autocorrezione quando un parlante identifica un errore durante o subito dopo la sua produzione e intende correggerlo riformulando. Quando ciò accade si conteggia l'ultima formulazione:

- (9) VAL: || | ed era {dafra-} davanti {a un::} a #o\_2 un fornaio  
così | ||

Si escludono dal conteggio *tch*, *mhm* o segnali discorsivi come *cioè*:

- (10) VAL: || | {cioè} si capisce | || #o\_5  
VAL: || | {cioè} non lo dice {cc-} | ||  
INV: da cosa si capisce h  
VAL: || | {boh} non lo so | ||

Nell'interazione sono frequenti anche fenomeni di *scaffolding* e interruzione. Nei casi in cui una subordinata sia collegata alla produzione precedente dell'apprendente, questa verrà considerata come clausola della stessa AS-unit, anche se conclusa dopo un turno dell'interlocutore:

- (11) PAN: || | [scappa dietro: #o\_5 alla ragassa #o\_5 | ||  
PAN: || | e: la ragassa #o\_3 cade sopra: h #o\_7 il vuoto .h |  
che l- era andato | a lavorare {in: °n na- # } in nave per la°  
INV: .h ne nel cantiere  
PAN: cantiere | ||

L'AS-unit è stata definita e operativizzata da Foster *et al.* (2000) con notevole precisione e tenendo specificamente conto delle dinamiche interazionali in L1 e L2: nella maggior parte dei casi la sua applicazione non ha generato ambiguità interpretative. Tuttavia, nella sperimentazione sono

emersi alcuni ambiti problematici, che in parte riguardano elementi specifici della lingua italiana, in parte toccano aspetti più generali. Due sono stati i tipi di questioni affrontati, il primo rispetto ai confini esterni dell'unità, "una o due AS-unit?", il secondo rispetto ai confini interni, "una o due clausole?". Vediamo in dettaglio le precisazioni operative che è stato necessario apportare nell'applicare l'AS-unit al corpus e la procedura impiegata per verificarne l'affidabilità nell'applicazione.

La prima questione riguarda l'identificazione del confine dell'unità, in particolare ci si riferisce a due strutture specifiche: la coordinazione e alcuni tipi di clausole introdotte da *perchè* o *che*.

Per quanto riguarda la coordinazione, gli stessi Foster *et al.* (2000) riscontrano alcuni problemi nel riconoscimento della coordinazione tra sintagmi verbali:

There is the problem of coordinated verb phrases. We argue that in many cases, the second VP, even if the subject is omitted, constitutes a fresh start for the speaker. To address this problem we have identified conditions for determining where the unit boundary lies. In cases where coordination of verbs occurs, the coordinated phrases will normally be considered to belong to the same AS-unit, unless the first phrase is marked by falling or rising intonation or is followed by a pause of at least 0.5 seconds.

Vediamo un esempio prodotto da una parlante nativa:

- (12) VAL: ehm poi:: ehm una ragazza:: entra in un negozio e::  
prende:: {un pani} hh una baguette? #o\_6 e:: {scappa::}  
e scappa via

Applicando la definizione di Foster *et al.* (2000) l'estratto (16) sarebbe costituito di 2 AS-unit, la prima di due clausole coordinate (*una ragazza entra in un negozio e prende una baguette*), la seconda di una sola clausola (*e scappa via*). In altri termini, il criterio intonativo-articolatorio servirebbe a distinguere la coordinazione tra due clausole entro la stessa AS-unit e tra due AS-unit.

Tuttavia, in una lingua come l'italiano le frasi senza soggetto sono di gran lunga più numerose, il tentativo di differenziarle solo sulla base di criteri intonativi o sulle pause rischia di essere piuttosto arbitrario. Infatti il ruolo delle pause e dell'intonazione non è sempre univoco: perciò non è sempre possibile decidere ad esempio se una pausa sia conseguenza della ricerca di una parola o della pianificazione di un nuovo enunciato (per una discussione del legame tra prosodia e sintassi nel parlato in L1 cfr. Savy, Voghera, 2010; Cresti, 2016). Ne consegue che determinare l'esistenza di un legame sintattico affidandosi a criteri di tipo intonativo o alle pause potrebbe limitare l'affidabilità della misura.

Dal punto di vista sintattico, poi, nonostante sia chiaro che la coordinazione è una relazione di non dipendenza, rispetto alla dipendenza della subordinazione, come viene sottolineato nella letteratura sulla tipologia dei legami inter-clausali (cfr. Talmy, 1978; Longacre, 1985; Lehmann, 1988; Telve, 2009), coordinazione e subordinazione possono essere descritte come i due poli di un continuum. A un estremo non troviamo nessuna relazione gerarchica tra le clausole che compongono una frase complessa, dall'altro troviamo un chiaro rapporto di dipendenza, dove la clausola subordinata è un costituente ben definito della frase. La posizione di una clausola subordinata rispetto alla reggente è legata a vincoli di tipo grammaticale. Questo non avviene nel caso della coordinazione: se invertiamo la posizione di due clausole coordinate possiamo

forse confondere la chiarezza semantica dell'enunciato, ma la sintassi rimane indifferente all'ordine delle clausole.

Nella ridefinizione dell'AS-unit si è stabilito, alla luce di queste riflessioni, di scostarsi dalle indicazioni di Foster *et al.* (2000) e fissare la linea di demarcazione tra coordinazione e subordinazione piuttosto presto, considerando le clausole coordinate come AS-unit indipendenti. In questo modo l'AS-unit risulta più affidabile nella sua applicazione e maggiormente legata a principi sintattici.

Un secondo punto potenzialmente problematico rispetto alla definizione dei confini esterni all'AS-unit riguarda l'interpretazione di alcuni tipi di clausola introdotti da *perché* o *che*, a volte chiaramente interpretabili come subordinate, a volte come nuovo inizio. In questi casi, ci è parso preferibile seguire l'impostazione proposta da Foster *et al.* (2000), che propone di usare criteri intonativi per distinguere tra clausole dipendenti incluse in un progetto sintattico unitario e clausole che paiono dipendenti in quanto introdotte da connettivi quali *perché* o *che*, che nell'analisi del linguaggio scritto sono considerati senza discussione come subordinanti ma che nel parlato spesso hanno una funzione di congiunzione semanticamente e sintatticamente "lasca" tra diversi elementi del testo.

(13) VAL: || | prendete me | perché son stata io. | || #o\_4

(14) PAN: || | [era f-]cane era fuori. | ||

PAN: || | perché::::: dopo che l'hanno cacciato fuori. |

INV: eh.

PAN: | guardava dalla finestra. | || # o\_4

Su un piano psicolinguistico, che ci pare più rilevante per la nostra trattazione ed è anche quello invocato da Foster *et al.* (2000), si può dire che in un enunciato come (13) è

piuttosto evidente il legame di dipendenza sintattica oltre che logico-concettuale tra le due clausole: tale appartenenza della clausola dipendente a un unico progetto sintattico-discorsivo è manifestata in particolare dal suo essere prodotta all'interno dello stesso profilo intonativo della reggente, in accordo con la prospettiva secondo cui ci sarebbe un forte parallelismo tra intonazione e sintassi nella lingua parlata (cfr. Chafe, 1988; Savy, Voghera, 2010). In (14), invece, *perché* sembra avere più le funzioni di un riempitivo o di un segnale discorsivo, che aiuta il parlante a regolare la strutturazione del discorso e la gestione del testo, dandogli modo di pianificare i passaggi successivi pur mantenendo un effetto di discorso vagamente coeso. Le due clausole sono distribuite su due profili intonativi distinti, *perché* è caratterizzato da un allungamento della vocale finale tipico dell'introduzione di nuovi topic ed è seguito dal connettore *dopo* che chiaramente appartiene alla struttura gerarchica seguente *dopo che l'hanno cacciato ... guardava*. Anche da un punto di vista discorsivo-funzionale i due casi sono differenti. In (13) il *perché* introduce una subordinata avverbiale con scope sulla predicazione, mentre in (14) lo scope è sull'unità precedente in quanto enunciato: [*dico così*] *perché*. L'intonazione in questi casi aiuta la scelta del ricercatore.

Si possono fornire esempi simili con *che* (15) e (16) sono tratti dalle produzioni di due apprendenti di italiano L2.

(15) SHI: || | trova: {il: # e:il: .h #o\_5 } l'uomo | che sta facendo #  
il pane. | che dal furgone va dentro {a a} al negozio? | ||

(16) SHI: || poi: e:: mentre .h # stava ancora pensando | sono tornati poi nella:: .hhh realtà | dove: # arriva # il poliziotto. | ||

SHI: || | che:::: cominciano ancora a scappare. | ||

Anche in questo caso *che* può rappresentare il connettore di una relativa (15) o essere un connettore discorsivo generico (16) e avere dunque un effetto di subordinazione solo semantica con quanto precedentemente detto. Nel caso del *che* relativo, la subordinata è sotto lo stesso profilo intonativo della reggente (o con profilo intonativo parallelo), mentre nel caso del *che* sintatticamente inattivo, le due clausole cadono sotto un diverso contorno intonativo e il *che* è caratterizzato da un allungamento vocalico (cfr. Scarano, 1996; 2003). In entrambi i casi dunque il ricercatore può basarsi sull'esistenza di relazioni sintattiche ed essere poi guidato a scegliere da elementi intonativi. Da un punto di vista operativo solo in presenza di chiari indizi intonativi si potrà stabilire l'indipendenza dei due progetti di pianificazione linguistica, ascrivendoli a AS-unit distinte. Nei casi dunque in cui clausole introdotte da *perchè* o *che* siano precedute da una pausa uguale o superiore a 0,5 secondi e caratterizzate da un diverso profilo intonativo, o ancora siano presenti esitazioni sul connettivo subordinante che portano a interpretarlo come l'inizio di un nuovo piano sintattico-intonativo piuttosto che come il compimento di uno preesistente, si indicherà l'inizio di una nuova AS-unit.

Il secondo tipo di problema che emerge nell'applicazione della AS-unit così come definita da Foster *et al.* (2000) riguarda la determinazione dei confini interni alla AS-unit, con particolare riferimento alla determinazione delle caratteristiche che una clausola deve avere per essere considerata subordinata (cfr. Giordano, Voghera, 2009; Cresti, 2016).

Siccome per Foster *et al.* (2000), come abbiamo visto, una clausola subordinata “will consist minimally of a finite or not finite verb element plus at least one other clause element”, clausole avverbiali o relative non pongono particolari difficoltà:

- (17) VAL: || | ha fatto di tutto | che così il pa#drone lo mandasse  
via. | ||
- (18) VAL: || | tch h e:m dopo che::: l'hanno licenziato | in pratica:  
sta passeggiando lungo la strada h | quando vede: una  
ragazza #o\_2 | che: {sta ruba cioè} sta rubando un filon-  
cino di pane | ||

Ma come interpretare “plus at least one other element” nei casi che seguono?

- (19) PAN: || | è cominciato a mangiare: | ||
- (20) PAN: || | mh # e::: quindi comincia: a boiare# ancora? | ||  
#o\_6
- (21) VAL: || | la voleva portare con sé in prigione | ||
- (22) VAL: || | vallo a cercare | || #o\_2

Seguendo la definizione di Foster *et al.* gli estratti in (19) e (22) dovrebbero essere analizzati come AS-unit di una clausola, quelli in (20) e (21) come AS-unit di due clausole. Questi esempi richiedono una definizione di clausola subordinata più precisa. Come devono essere considerati i verbi che dimostrano un alto grado di integrazione, come i modali e i fasali seguiti dall'infinito? La letteratura sull'italiano L1 e L2 mostra come non ci sia un accordo generale nelle modalità di analisi di queste strutture. Seguendo Giacalone Ramat (1999a; 1999b), gli esempi precedenti sarebbero clausole indipendenti seguite da una subordinata implicita. Seguendo Renzi *et al.* (1988-1995) e Chini (2003) sarebbero invece clausole indipendenti costituite da un verbo complesso.

Nella definizione di criteri applicativi per l'AS-unit si è pertanto stabilito di definire linee di demarcazione precise per distinguere i casi in cui il verbo reggente e l'infinito formano



una specie di verbo composto da quelli in cui l'infinito costituisce un membro indipendente del sintagma verbale. Secondo Renzi e Salvi (2001: cap. 9) e Graffi (1994; 2001), un verbo reggente seguito da un'infinito costituisce un verbo complesso se sono soddisfatte le seguenti condizioni: (a) i clitici, pur dipendendo dall'infinito si trovano accanto ai verbi reggenti; (b) il complemento oggetto del verbo all'infinito, può diventare il soggetto della costruzione con il *si* passivo; (c) l'infinito non può essere negato. Pertanto l'infinito non può essere considerato il predicato di una subordinata nei seguenti casi: (a) in costruzioni fattitive o causative, come *fare*, *lasciare* e alcuni verbi percettivi, *intendere*, *sentire*, *udire* e *vedere*:

(23) VAL: || | e ci\_h # aveva fatto scappare la ragazza prima h | ||

(24) SHI: || | pantera rosa. | #o\_4 che: ho:: sentito parlare. | non  
ho mai avuto ehm guardato ne[ssuna puntata del cartone  
ne | ||

e (b) in costruzioni a ristrutturazione, con verbi modali (*dovere*, *potere*, *volere* e *sapere*), verbi aspettuali (*cominciare*, *finire*) e verbi di moto (*andare*, *venire*):

(25) PAN: || | poi: #o\_5 la: ragazza ha cominciato d- a scappare |  
||

(26) PAN: || | e:: la ragazza #o\_3 cade sopra: h #o\_7 il vuoto .h |  
che l- era andato a lavorare {in: °n na- # } in nave per la°

INV: .h ne nel cantiere

PAN: cantiere | ||

(27) VAL: || | [e quindi vorremmo sapere | se lei ci può consiglia-  
re | ||

Anche le perifrasi verbali vengono considerate un'unica clausola:

- (28) SHI: || | poi: e:: mentre .h # stava ancora pensando | sono  
tornati poi nella: .hhh realtà | dove: # arriva # il poli-  
ziotto | ||

Un ulteriore aspetto critico della segmentazione in AS-unit riguarda alcune costruzioni marcate tipiche del parlato, che in particolare se dialogico è caratterizzato dall'effetto di sintassi frammentaria, con una struttura per periodi giustapposti. In primo luogo, il parlato si sviluppa attraverso microprogettazione tematica e sintattica: ne sono la manifestazione pause, esitazioni, piani abbandonati, frasi incomplete, nonché le repentine e frequenti fratture di continuità tematica non segnalate esplicitamente da connettivi, ma solo da pause e da demarcativi riempitivi e fatici (Voghera, 2017). In secondo luogo, le riprese e i mutamenti di topic spesso non vengono esplicitate, mentre le varie parti del discorso vengono cucite con segnali di articolazione, la catena anaforica si realizza attraverso riprese ad hoc, rimandi all'universo di discorso, realizzando così la minima codifica morfosintattica. Infine frequente è l'uso di dislocazioni o frasi segmentate (Bazzanella, 1991; 1994; 2002).

Uno dei fenomeni più caratteristici della sintassi del parlato riguarda proprio la variazione nell'ordine delle parole, che in questa modalità segue principi di strutturazione diversi rispetto all'ordine canonico dell'italiano, in cui il soggetto è tema e dato e viene a sinistra del verbo, e l'oggetto è rema e nuovo e viene a destra. Particolari realizzazioni marcate dell'ordine dei costituenti sono connesse sia con la struttura spezzettata e disarticolata, sia con l'accentuazione empatica, tipiche del parlato. I costrutti marcati presenti

nel corpus sono principalmente di tre tipi: posposizione del soggetto, frase segmentata (con dislocazione) e frasi scisse (o pseudoscissa e presentativa) (per una trattazione dei costrutti di sintassi marcata dell'italiano si rimanda ad esempio a Sabatini, 2011; Berruto, 2012; De Cesare *et al.*, 2016). Se i primi due non pongono particolari difficoltà nell'identificazione dei confini interni alle unità, le frasi scisse in alcuni casi possono prestarsi ad interpretazioni discordanti. In questo studio si è pertanto stabilito che esse costituiscono una AS-unit formata da una sola clausola:

(29) ELI: || | è lui che stanno cercando | || #

Un'altra costruzione funzionalmente simile è rappresentata dalle presentative del tipo *esserci x che*:

(30) PAN: || | c'era u:n bastone | #o\_3 che {c'è} > sotto < c'era una  
.h cono | #o\_5 che ha fermato nave? | ||

Questo tipo di costrutto è particolarmente utile e pertanto frequente nel parlato spontaneo con gittata di pianificazione ridotta, poiché il parlante trova vantaggio nell'enunciare un oggetto di discorso, lasciando al momento processuale successivo la predicazione su di esso. Allo stesso modo l'informazione distribuita su due frasi risulta di maggior agio ricettivo anche per l'ascoltatore. Berruto (1986) considera queste costruzioni pseudo-relative e assegna loro un valore fondamentalmente esplicativo e di commento. La frase introdotta da *che* non è infatti un modificatore di *x*, ma contiene la proposizione principale del costrutto, di cui il sintagma scisso è argomento necessario. Nonostante il valore semantico sia ben diverso da quello delle relative standard, esse condividono almeno in parte, certe caratte-

ristiche di superficie. Per questa ragione, in sede di analisi le costruzioni del tipo *esserci x che y* vengono interpretate come una AS-unit di due clausole. Interessanti però i due esempi (31) e (32), dove il verbo *essere* è sottinteso e il sintagma nominale modificato dalla relativa viene codificato come unità indipendente inferiore alla clausola:

- (31) SHI: || | allora mia sorella grande | che non c'è | ||  
 (32) SHI: || | il mio fratello | che: si chiama m[artin. | ||

Per verificare l'affidabilità della AS-unit è stato effettuato un controllo sperimentale su una parte del corpus, che ha coinvolto due valutatori. In una prima fase i valutatori hanno potuto familiarizzarsi con la definizione di AS-unit discutendo alcuni estratti dal corpus e esercitandosi nella segmentazione. In una seconda fase hanno, l'uno indipendentemente dall'altro, segmentato l'intervista, il racconto di un film e l'apertura delle telefonate realizzati da Pandita e da Elisa a T1 (1278 parole). Si è proceduto poi al confronto delle due codifiche (*inter-coder agreement*). Poiché in letteratura si considerano affidabili i giudizi quando il valore dell'accordo è pari o superiore al 75% (Bachman, 2004; Stemler, 2004), l'applicazione della AS-unit così ridefinita ha dimostrato un ottimo grado di affidabilità, con un accordo tra i due valutatori nell'identificazione sia di clausole che di AS-unit pari al 98%.

La tabella 2.3. riassume le riflessioni esposte sopra proponendo una parziale ridefinizione dell'AS-unit per l'italiano. La descrizione dei fenomeni specifici è accompagnata da esempi tratti dal corpus.

**Tabella 2.3.** L'AS-unit, Analysis of Speech Unit per la lingua italiana.

unità	esempio	definizione
clausola indipendente	ELI:      e::: {va:::} ruba {un pezzo di legno} # insomma del fuoco:: dalla casa del cacciatore	clausola che include un verbo finito, anche complesso: costruzioni fattive e causative; alcuni verbi percettivi come <i>intendere</i> , <i>sentire</i> o <i>vedere</i> , con verbi modali, aspettuali, di moto; perifrasi verbali, come <i>stare</i> + gerundio
	VAL:      la voleva portare con sé in prigione	
	PAN:      è cominciato a mangiare:	
unità indipendente inferiore alla clausola	SHI:      allora Danielle Steel      SH2: sì SHI:      il ranch      SH2: vediamo	clausola costituita da una o più parole rielaborabili come clausola completa recuperando elementi sottintesi dal contesto del discorso o dalla situazione, frasi senza verbo
clausola dipendente	PAN:      quando arriva lì   gli dà una lettera.	clausole argomentali, avverbiali, relative e infinitive. Un verbo reggente seguito da infinito non dà luogo a due clausole nelle costruzioni fattive o causative, con verbi modali, aspettuali e di moto o con perifrasi verbali.
	VAL:      la pantera:: ha sempre fatto di tutto   di modo che il cacciatore si incavoli con il cane	
	SHI:      vorrei sapere dei informazioni su un libro  da regalare a una mia amica	
	PAN:      {,h} c'è un:: # vuoto   che dice   di prendere una: # legna {n} come {co:} co:no	
fenomeno	esempio	analisi
clausola coordinata	PAN:      e prende quella lettera.           e va a cercare lavoro #1_5.           e lo prendono	si conteggia come AS-unit indipendente

fenomeno	esempio	analisi
clausola introdotta da <i>che</i> o <i>perché</i>	VAL:      prendete me   <b>perché son stata io.</b>      PAN:      [era f-]cane era fuori.           <b>perché:::: dopo che l'hanno cacciato fuori.</b>      INV: eh. PAN:   guardava dalla finestra.      ELI:      la pantera è sempre dentro l'acqua con un ombrello #0_5           <b>che:: non ho capito perché</b>	si distinguono i connettori sintatticamente attivi dai casi in cui <i>che</i> o <i>perché</i> sono segnali discorsivi. Criteri: cambi del profilo intonativo, connettivo caratterizzato da allungamento vocalico, preceduto da una pausa uguale o superiore a 0,5 sec
frasi scisse	ELI:      volevo chiederle un'informazione      SH2: sì ELI:      lei non è che avrebbe per caso dei cd in offerta dei pink floyd      #	si considera come AS-unit formata da una clausola
presentative (esserci X che)	PAN:      c'è una ragazza   che aveva fame	si considerano come una AS-unit formata di due clausole
<i>scaffolding</i> e interruzione	PAN:      e: la ragazza #0_3 cade sopra: h #0_7 il vuoto .h   che l- era andato   a lavorare {in: °n na- # } <b>in nave per la°</b> INV: .h ne nel cantiere PAN: <b>cantiere</b>	se una clausola è collegata alla produzione precedente del parlante, viene considerata parte della stessa AS-unit
ripetizione	PAN:      e- adesso quando { <b>il &gt;polisiotto</b> } anche il polisiotto < è alzato?      PAN:      poi Charlot- entra in <caffetteria>           e mangia <b>tutto tutto</b>           non paga niente?	si distingue tra ripetizione come disfluenza, non inclusa nell'analisi (e inserita quindi tra parentesi graffe), e ripetizione come strategia retorica, inclusa
autocorrezione	VAL:      ed era {dafra-} davanti {a un::} a #0_2 un fornaio così	si conteggia l'ultima versione

falsa partenza	ELI:      {che dice che ehm} # {cioè} ho visto   che va a trovare un lavoro	La clausola è esclusa dall'analisi, a meno che una AS-unit o una clausola completa sia prodotta prima che il piano venga abbandonato
<i>tch</i> , <i>mhm</i> o segnali discorsivi puramente faticosi come <i>cioè</i>	VAL:     {mhm cioè} non lo dice {cc-}      INV: da cosa si capisce h VAL:     {boh} non lo so	si escludono dal conteggio del numero di parole per AS-unit

Definita così l'AS-unit e stabilita la sua affidabilità, vediamo ora le misure CAF. La presentazione è limitata alle misure utili per questo studio. Si presentano prima le misure relative alla complessità sintattica, poi quelle relative ad accuratezza, e infine si illustrano le misure di fluenza.

Si è visto in § 1.2.2.2. come le misure di complessità sintattica siano numerose e alternative. La tabella 2.4. riporta quelle principali. Esse rilevano due aspetti del processo di complessificazione: il grado, ossia la capacità di usare strutture più complesse sia a livello di clausola che di AS-unit; e la gamma, ossia la capacità di usare una varietà di strutture. Applicare tutte queste misure al corpus sarebbe oltre che oneroso in termini di tempo anche poco pratico, soprattutto in funzione dell'interpretazione dei risultati. Se in alcuni casi le informazioni fornite da queste misure certamente non sono uguali, in altri sono ridondanti. Per questo è stato necessario fare una cernita.

La misura globale di grado è rappresentata dalla lunghezza delle unità di produzione. Abbiamo visto in § 1.2.2.2. come non ci sia generale consenso in letteratura sulla funzione di questa misura. In questo studio essa viene considerata un indicatore globale della complessità sintattica. Le misure globali e specifiche di grado si distinguono tra di loro, in primo luogo per elemento al numeratore, ossia il numero

**Tabella 2.4.** Le principali misure di complessità sintattica proposte dalla letteratura.

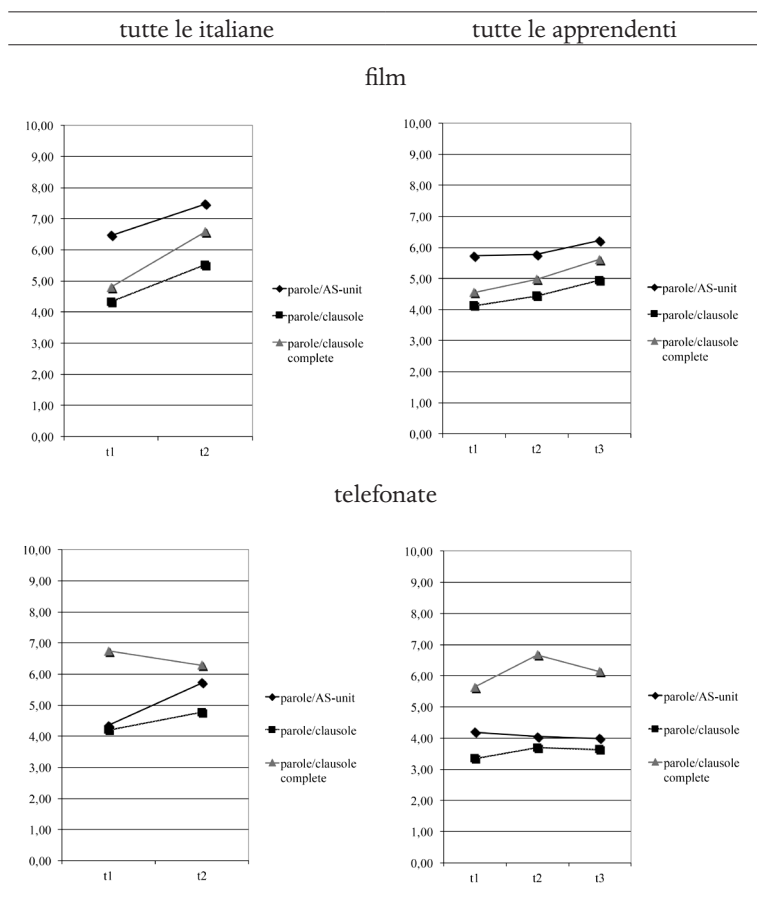
<i>misura globale di grado</i>	
lunghezza della AS-unit	numero medio di parole per AS-unit
<i>misura specifica di grado</i>	
lunghezza della clausola	numero medio di parole per clausola
	numero medio di parole per clausola completa (escludendo cioè le unità inferiori alla clausola)
grado di subordinazione	numero medio di clausole subordinate per AS-unit
	numero medio di clausole subordinate per clausola
	numero medio di clausole subordinate per clausola completa
<i>misura qualitativa di gamma</i>	
tipo di clausola	percentuale dei 3 tipi di clausole: indipendenti, unità indipendente inferiore alla clausola, dipendenti sul totale delle clausole
	percentuale dei tipi di subordinate: argomentali, avverbiali e relative sul totale delle subordinate

di parole e di subordinate, e in secondo luogo per l'elemento al denominatore, ossia se si tratta di AS-unit, clausola o clausola completa. Possiamo infatti ipotizzare che una differenza nel numeratore generi misure complementari, una differenza nel denominatore generi probabilmente misure ridondanti (cfr. Norris, Ortega, 2009). Per verificare questo sperimentalmente si è considerato il parlato prodotto dai sei soggetti nelle prime tre rilevazioni in due task del corpus: film e apertura delle telefonate. I risultati ottenuti dai due



gruppi applicando le misure di grado sono stati poi confrontati. Il grafico 2.1. riporta le medie dei risultati ottenuti sulle misure di lunghezza per i due gruppi delle quattro apprendenti e delle due studentesse italiane, mentre la tabella 2.5. riporta la variazione tracciata dalle misure.

**Grafico 2.1.** CAF — Il confronto tra le misure di lunghezza — Italiane e Apprendenti (valori di gruppo).



**Tabella 2.5.** Le misure di lunghezza, variazione.

	<i>film</i>		<i>telefonate</i>	
	apprendenti	italiane	apprendenti	italiane
parole/AS-unit	8,70%	15,44%	-4,74%	31,55%
parole/clausole	19,87%	27,40%	8,73%	7,14%
parole/clausole complete	23,38%	30,39%	8,92%	7,12%

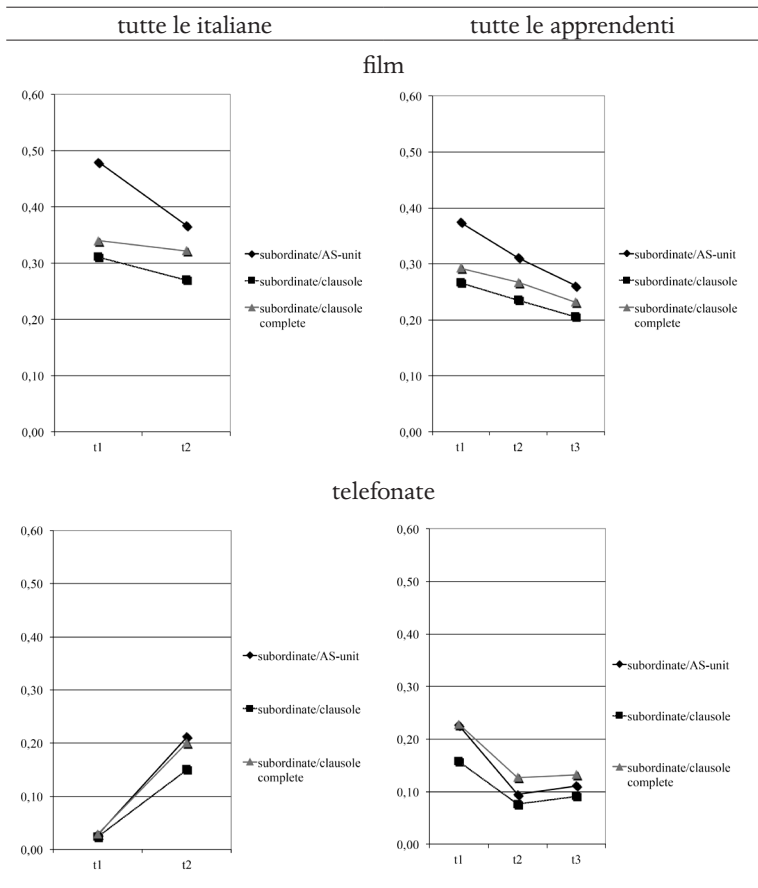
Confrontando gli andamenti di parole/AS-unit e parole/clausole o parole/clausole complete si osserva come l'andamento di parole/AS-unit sia diverso rispetto a quello delle altre due misure. Se si considerano ad esempio i valori ottenuti nel task apertura delle telefonate realizzato dalle apprendenti, è evidente come la misura parole/AS-unit abbia non solo un andamento diverso, ma addirittura inverso rispetto agli altri due: la misura parole/AS-unit rileva in questo caso una variazione in negativo del 4,74% tra  $t_1$  e  $t_3$ , mentre le misure parole/clausole e parole/clausole complete variano in positivo dell'8,73% e dell'8,92% rispettivamente.

Questo supporta quanto già ipotizzato in letteratura rispetto alla relazione tra questi indici (cfr. § 1.2.2.2.): la misura basata sulla AS-unit e le misure basate sulla clausola rilevano aspetti diversi. La prima rileva la complessificazione generata sia a livello di clausola sia tramite l'uso di subordinate, le seconde sono sensibili invece solo al primo tipo di complessificazione. Confrontando invece le due misure parole/clausole e parole/clausole complete vediamo che hanno andamenti molto simili e quindi possiamo ritenerle analoghe.

Il grafico 2.2. riporta le medie dei risultati ottenuti per le misure del grado di subordinazione dai due gruppi delle studentesse italiane e delle apprendenti, e la tabella 2.6. mostra la variazione tra misure. Qui constatiamo che le tre misure

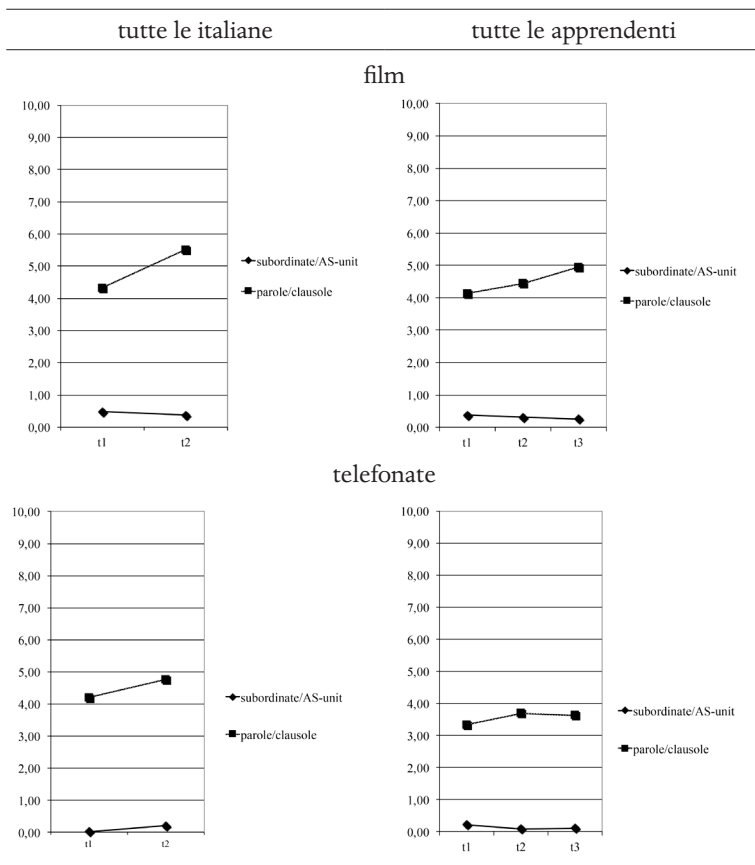
(subordinate/AS-unit vs subordinate/clausole e subordinate/clausole complete) hanno andamenti molto simili, con l'unica differenza che subordinate/AS-unit sembra un po' più sensibile al mutamento, dimostrando in tutti i contesti maggior variazione. Possiamo pertanto ritenere che le tre misure siano equivalenti.

**Grafico 2.2.** CAF — Il confronto tra le misure di subordinazione — Italiane e Apprendenti (valori di gruppo).



**Tabella 2.6.** Le misure di subordinazione, variazione.

	<i>film</i>		<i>telefonate</i>	
	apprendenti	italiane	apprendenti	italiane
subordinate/AS-unit	-30,36%	-23,46%	-50,86%	11,34%
subordinate/clausole	-22,83%	-13,12%	-42,14%	8,92%
subordinate/clausole complete	-20,60%	-9,97%	-41,87%	12,04%

**Grafico 2.3.** CAF — Il confronto tra le misure di subordinazione e di lunghezza — Italiane e Apprendenti (valori di gruppo).

Sulla base di questi risultati si è deciso di utilizzare solo due misure per l'analisi: per la lunghezza la misura parole/clausole e per il grado di subordinazione la misura subordinate/AS-unit. Il grafico 2.3. e la tabella 2.7. mostrano come queste due misure hanno andamenti diversi, in alcuni casi addirittura inversi.

**Tabella 2.7.** Le misure di subordinazione e di lunghezza, variazione.

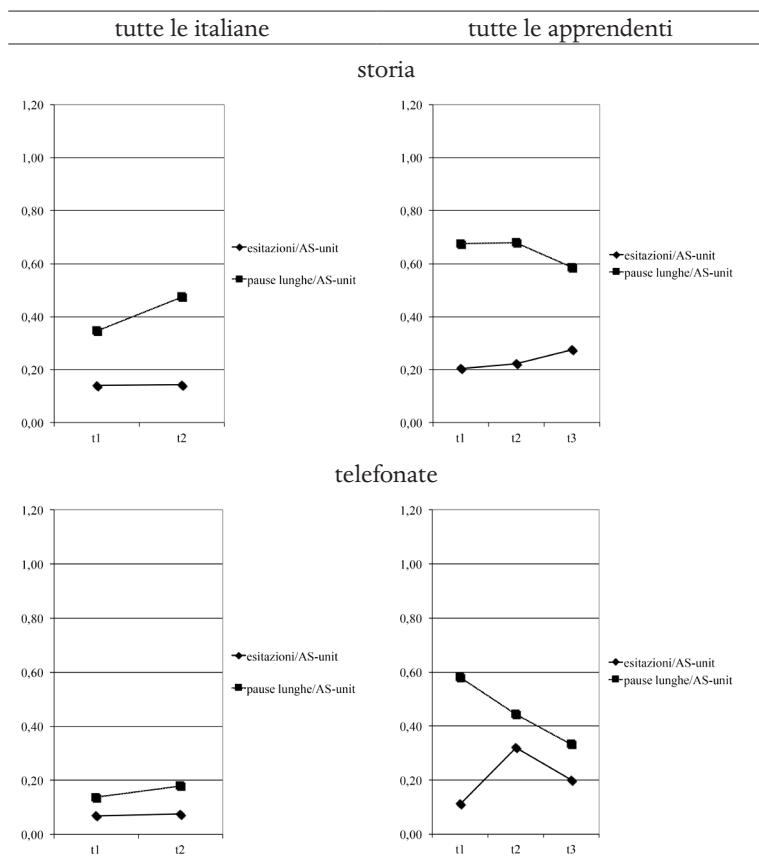
	<i>film</i>		<i>telefonate</i>	
	apprendenti	italiane	apprendenti	italiane
parole/clausole	19,87%	27,40%	8,73%	7,14%
subordinate/AS-unit	-30,36%	-23,46%	-50,86%	11,34%

Per quanto riguarda le misure qualitative di gamma, sono state mantenute tutte le misure riportate nella tabella 2.4.

Vediamo ora le misure per rilevare l'accuratezza dell'interlingua orale. In §1.2.2.2. abbiamo visto le principali impiegate in letteratura, in questa sezione si illustra la misura scelta per questo studio e se ne giustifica l'utilità.

Il rapporto tra il numero di unità di produzione senza errori e il numero totale delle unità prodotte in letteratura è considerata una misura globale utile (anche se non esaustiva) per rilevare l'accuratezza delle produzioni a livello avanzato come sono i soggetti coinvolti in questo studio. Pertanto è stata qui utilizzata pur nella piena consapevolezza — è ovvio — che i risultati di un'analisi di questo tipo permetteranno di fare solo considerazioni globali sull'accuratezza.

In sede di analisi sono stati considerati errori le inesattezze di tipo morfo-sintattico o lessicale, come in (33-36), mentre sono state conteggiate come corrette quelle strutture che, pur deviando dallo standard, risultano corrette nell'italiano parlato, come in (37):

**Grafico 2.4.** CAF — Il confronto tra le misure di fluenza pause lunghe e esitazioni — Italiane e apprendenti (valori di gruppo).**Tabella 2.8.** Le misure di fluenza, variazione.

	<i>storia</i>		<i>telefonate</i>	
	apprendenti	italiane	apprendenti	italiane
pause lunghe/AS-unit	-13,20%	37,25%	-42,77%	50,45%
esitazioni/AS-unit	33,99%	1,31%	-11,63%	20,46%

- (33) PAN: || | un: negoziere | ||
- (34) PAN: || | {mh .h} e # lo pane prende vuomo #o\_2 >quel de<  
cantie:re | ||
- (35) SHI: || | .hhh ha fatto: di tutto | cosi la polizia gli: # venisse  
a prend ere, | ||
- (36) VAL: || | {e: tch cioè: h} lui prende dei sigari | ||
- VAL: || | lo offre anche a dei £bambini£ | || #o\_5
- (37) VAL: || | e si vede::: l'ultima scena | che si ritrovano su un  
pezzo di legno | a galleggiare::: | ||

Inoltre nel caso in cui lo studente correggeva un proprio errore, questo non veniva conteggiato:

- (38) SHI: || | e poi {ha arriva} è arrivato sull'isola::: | ||

Infine, nell'identificare le AS-unit senza errori sono stati esclusi gli errori di pronuncia:

- (39) PAN: || | poi loro due contrano | ||

Per verificare l'affidabilità di questi criteri di misurazione dell'accuratezza è stato effettuato un controllo sperimentale su una parte del corpus: i task intervista, racconto di un film e apertura delle telefonate realizzati da Pandita ed Elisa a tr (1278 parole, 446 AS-unit). Il controllo ha coinvolto due parlanti nativi nel ruolo di valutatori, cui è stato richiesto di identificare le AS-unit senza errori. Si è proceduto poi al confronto delle due codifiche, che ha dato un buon grado di affidabilità, con un accordo tra i due valutatori pari al 98%.

Si presenta ora la misurazione della fluenza. Le diverse misure di fluenza proposte in letteratura cercano di rilevarne tre aspetti: la fluenza legata alla velocità e la fluenza legata alle

pause, caratterizzate dalle variabili temporali, e la fluenza legata alle riparazioni definita dai fenomeni di esitazione.

Il conteggio delle misure di fluenza ha un grosso limite pratico, poiché richiede prima una trascrizione di altissima qualità e dunque estremamente impegnativa in termini di tempo. Ai fini di questo studio, tra le misure proposte in letteratura (cfr. § 1.2.2.2., tabella 1.6.), non ne sono state prese in considerazione tre: velocità d'eloquio, durata delle pause e lunghezza media tra due pause. Il tipo di corpus e gli strumenti a disposizione del ricercatore non hanno infatti permesso il conteggio della prima misura, la velocità dell'eloquio. La seconda è stata scartata dopo un controllo sperimentale su una parte del corpus: il task racconto di una storia realizzato in t3 da Catherine, Elisa e Valentina. Qui i risultati prodotti dai due ricercatori che si sono occupati delle trascrizioni ha mostrato come possa esserci un buon grado di affidabilità nell'individuazione delle pause e nella distinzione tra micro-pause e pause lunghe (con un grado di accordo dell'87,5%) ma risulta meno affidabile il conteggio della durata delle pause, con differenze anche molto marcate (accordo del 32% dei casi, con variazione media di #o\_3, e una deviazione standard di #o\_2/#o\_6). Si è pertanto ritenuto che il conteggio della durata media delle pause non potesse essere applicato ai dati di questo corpus, in quanto avrebbe portato a risultati poco affidabili. Infine, l'esclusione della terza misura, lunghezza media tra due pause, è dovuta alla difficoltà di applicazione a task interattivi, dove la definizione del confine dell'unità non è scontata. Poiché la conversazione è costruita con l'interlocutore, il flusso del parlato si interrompe non solo nei casi in cui l'apprendente fa una pausa, piena o vuota, ma anche nel caso dei cambi di turno. Da un lato considerare i cambi di turno come confine dell'unità significherebbe, nei task interattivi, ottenere unità molto brevi, dunque interpretabili



come poco fluenti, dall'altro non considerarli confine di unità porterebbe a unità anche troppo lunghe, rendendo pertanto difficile l'interpretazione dei risultati.

In questo studio sono state scelte due misure di fluenza: numero di pause lunghe per AS-unit e numero di fenomeni di esitazione per AS-unit. Così come per la complessità, è stato effettuato un confronto tra le due misure su due task, il racconto di una storia ad immagini e l'apertura delle telefonate, realizzati dalle apprendenti e dalle italiane tra t1 e t3.

Il grafico 2.4. riporta i risultati ottenuti dai due gruppi sulle due misure di fluenza considerate e la tabella 2.8. ne riporta la variazione. Il confronto degli andamenti sui tre anni mostra come i risultati ottenuti siano piuttosto diversi gli uni dagli altri, indicando in alcuni casi tendenze addirittura opposte. Questo dato supporta l'idea suggerita in letteratura secondo la quale le misure considerate rilevano diversi e complementari aspetti della fluenza. Entrambe sono pertanto state impiegate per l'analisi del corpus oggetto di questo studio.

Ricapitolando, in questo studio, riguardo al CAF vengono usate le seguenti misure: per la complessità sintattica due specifiche di grado (lunghezza e subordinazione); tre qualitative di gamma relative al tipo di clausola (percentuale di clausole indipendenti, unità indipendenti inferiori alla clausola e clausole subordinate sul totale delle clausole), e infine tre qualitative di gamma relative alla subordinazione (percentuale di avverbiali, argomentali e relative sul totale delle subordinate).

Per l'accuratezza viene usata la misura percentuale delle AS-unit senza errori sul totale delle AS-unit.

Infine per la fluenza sono state scelte due misure: il numero di pause lunghe per AS-unit, e quello di fenomeni di esitazione per AS-unit.

### 2.6.2. Le valutazioni QCER

In questa sezione si presentano le griglie del QCER utilizzate in questo studio, si discute un'indagine di validazione empirica dell'affidabilità delle valutazioni e infine si illustra la procedura impiegata per la valutazione dei dati.

I descrittori sono tratti dalle griglie relative agli aspetti qualitativi dell'uso della lingua parlata e si riferiscono alle tre dimensioni della coerenza-coesione, della correttezza grammaticale e della fluenza. Questa selezione è dovuta al fatto che questi descrittori sono confrontabili con ciò che viene misurato attraverso le misure analitiche CAF esposte in § 2.6.1. Infatti alla complessità del CAF corrisponde la coerenza-coesione (QCER, p. 154), all'accuratezza corrisponde la correttezza grammaticale (QCER, p. 140), mentre — ovviamente — la fluenza del CAF coincide con quella del QCER (p. 158). La tabella 2.9. riporta i descrittori del QCER selezionati.

**Tabella 2.9.** QCER — Gli aspetti qualitativi dell'uso della lingua parlata.

	<i>coerenza-coesione</i>	<i>correttezza grammaticale</i>	<i>fluenza</i>
C2	È in grado di realizzare un discorso coerente e coeso usando in modo appropriato una grande varietà di schemi organizzativi e un'ampia gamma di connettivi e di meccanismi coesivi di altro tipo.	Mantiene costantemente il controllo grammaticale di forme linguistiche complesse, anche quando la sua attenzione è rivolta altrove (ad es. nella pianificazione di quanto intende dire e nell'osservazione delle reazioni altrui).	È in grado di esprimersi in discorsi lunghi con scioltezza naturale, senza sforzi né esitazioni. Fa pause solo per cercare le parole per esprimere con precisione il proprio pensiero o per trovare un esempio o una spiegazione adatti.

	<i>coerenza-coesione</i>	<i>correttezza grammaticale</i>	<i>fluenza</i>
C1	È in grado di realizzare un discorso chiaro, sciolto e ben strutturato, mostrando un uso controllato degli schemi organizzativi, di connettivi ed espressioni coesive.	Mantiene costantemente un livello elevato di correttezza grammaticale; gli errori sono rari e poco evidenti.	È in grado di esprimersi con scioltezza e spontaneità, quasi senza sforzo. Solo un argomento concettualmente complesso può ostacolare la naturale scioltezza del discorso.
B2+	È in grado di usare in modo efficace diversi connettivi per esplicitare i rapporti tra i concetti.	Ha una buona padronanza grammaticale; nella struttura delle frasi possono ancora verificarsi sbagli occasionali, errori non sistematici e difetti minori, che sono per lo più rari e corretti a posteriori.	È in grado di comunicare con spontaneità, dando per lo più prova di notevole scioltezza e uso disinvolto dei mezzi espressivi anche in discorsi piuttosto lunghi e complessi.
B2	È in grado di usare un numero limitato di elementi di coesione per collegare i propri enunciati in un discorso chiaro e coerente. In un intervento lungo possono presentarsi dei "salti" logici.	Mostra una padronanza grammaticale piuttosto buona. Non fa errori che possano provocare fraintendimenti.	È in grado di produrre sequenze discorsive con un ritmo abbastanza uniforme; anche se può avere delle esitazioni quando cerca strutture ed espressioni, fa poche pause evidenti. È in grado di interagire con spontaneità e scioltezza tali da consentire una normale interazione con parlanti nativi, senza sforzi per entrambe le parti.
B1+		Comunica con ragionevole correttezza in contesti familiari; la padronanza grammaticale è generalmente buona anche se si nota l'influenza della lingua madre. Nonostante gli errori, ciò che cerca di esprimere è chiaro.	È in grado di esprimersi con relativa disinvoltura. Nonostante alcuni problemi di formulazione che possono sfociare in pause e blocchi, è in grado di portare avanti il discorso efficacemente senza aiuto.

	<i>coerenza-coesione</i>	<i>correttezza grammaticale</i>	<i>fluenza</i>
B1	È in grado di collegare una serie di elementi relativamente brevi e semplici in una sequenza lineare per punti.	Usa in modo ragionevolmente corretto un repertorio di formule di <i>routine</i> e strutture d'uso frequente, relative alle situazioni più prevedibili.	È in grado di mantenere il discorso comprensibile, anche se sono evidenti pause per cercare parole e forme grammaticali e per riparare agli errori, specialmente nelle sequenze di produzione libera di una certa lunghezza.
A2+	È in grado di collegare frasi semplici usando i connettivi più usuali per raccontare una storia o descrivere qualcosa, realizzando un semplice elenco di punti.		Riesce a farsi comprendere con enunciati molto brevi, nonostante che pause, false partenze e riformulazioni siano molto evidenti.
A2	È in grado di collegare gruppi di parole con connettivi semplici quali "e", "ma" e "perché".	Usa correttamente alcune strutture semplici, ma continua sistematicamente a fare errori di base — per esempio tende a confondere i tempi verbali e a dimenticare di segnalare gli accordi; ciononostante ciò che cerca di dire è solitamente chiaro.	Su argomenti familiari è in grado di formulare espressioni e ha sufficiente capacità di portare a termine scambi comunicativi brevi, nonostante le esitazioni e le false partenze siano evidenti.
A1	È in grado di collegare parole o gruppi di parole con connettivi molto elementari quali "e" o "allora".	Ha solo una padronanza limitata di qualche struttura grammaticale e di semplici modelli sintattici, in un repertorio memorizzato.	È in grado di cavarsela con enunciati molto brevi, isolati, solitamente memorizzati, facendo molte pause per cercare le espressioni, per pronunciare le parole meno familiari e per riparare agli errori di comunicazione.

Per rilevare l'affidabilità nell'applicazione dei descrittori selezionati, ci si rifa all'indagine sperimentale condotta da Ferrari e Nuzzo (2010). Tale indagine consiste in una validazione empirica dell'affidabilità delle valutazioni espresse da due gruppi di valutatori, tutti insegnanti, tra cui 5 esperti valutatori e 5 non esperti, su tre attività del corpus (il racconto di un film, l'intervista e le telefonate) realizzate da 3 soggetti (due apprendenti, Pandita e Shirley, e una studentessa italiana, Elisa). Gli insegnanti esperti sono docenti di un Centro Territoriale Permanente, formati come esaminatori e regolarmente coinvolti come valutatori della prova orale negli esami di certificazione CELI dell'Università per Stranieri di Perugia. Gli insegnanti non esperti hanno diverse esperienze nel mondo della scuola, e sono tutti impegnati come volontari in scuole di italiano L2 per adulti.

Ciascun gruppo di insegnanti è stato impegnato in una sessione di valutazione della durata di 3 ore circa durante la quale si sono valutate le videoregistrazioni dei 3 task, riferendosi a griglie basate sia sui descrittori degli aspetti qualitativi dell'uso della lingua parlata che sui descrittori specifici riferiti ai tipi di task. L'attribuzione del livello avveniva subito dopo la visione di ciascuna produzione ed era suddivisa in due momenti: una prima valutazione individuale e una seconda valutazione condivisa da raggiungere attraverso una discussione di gruppo. L'analisi dell'affidabilità dei giudizi espressi (statistica descrittiva e correlazione) è stata condotta sulle valutazioni individuali e sulle valutazioni condivise.

La tabella 2.10. riporta i risultati relativi ai tre descrittori impiegati per questo studio. Ricordando che in letteratura in genere si considerano affidabili i giudizi quando il valore della correlazione è pari o superiore a 0,80 (Alderson *et al.*, 1995), osserviamo che i descrittori relativi alla correttezza grammaticale hanno ottenuto il maggior accordo. Rispetto

a coerenza-coesione le discrepanze più rilevanti sono relative al task telefonate, poiché in diversi casi le insegnanti non sono state in grado di esprimere una valutazione ritenendo i descrittori inadeguati per un task interattivo. I risultati relativi alla fluenza sono stati invece quelli meno soddisfacenti.

**Tabella 2.10.** L'affidabilità dei giudizi. Spearman rank-order (da Ferrari, Nuzzo, 2010).

	tutti gli insegnanti	insegnanti esperti	insegnanti non esperti	giudizi comuni
coerenza-coesione	0,83	0,89	0,83	0,79
correttezza grammaticale	0,90	0,88	0,81	0,87
fluenza nel parlato	0,68	0,87	0,55	0,83

Gli insegnanti esperti dimostrano maggior accordo rispetto ai colleghi non esperti. Questo risultato è in linea con ciò che viene da più parti evidenziato nella letteratura sul testing, e cioè che la formazione aumenta la coerenza e l'affidabilità delle scelte dei valutatori (Lunz *et al.*, 1990) oltre che ridurre le differenze di severità tra esaminatori (Wigglesworth, 1993; Weigle, 1994). Lumley e McNamara (1995) suggeriscono che i vantaggi della formazione non sono però di lunga durata, dimostrando come sia necessario un momento di formazione prima di ogni sessione di valutazione. Interessanti inoltre i risultati ottenuti nelle valutazioni condivise: la discussione nei gruppi per il raggiungimento di un accordo sembra portare gli insegnanti ad allinearsi maggiormente nelle loro scelte. Questo dato sostiene l'utilità della pratica diffusa nelle certificazioni di impiegare due valutatori per ciascuna valutazione (cfr. Alderson *et al.*, 1995). Nonostante non ci sia sempre una correlazione ottimale dei giudizi per i 10 insegnanti, i risultati di questa indagine permettono alcune osservazioni in posi-

tivo. Sugli 11 livelli disponibili<sup>1</sup> per ciascuna competenza, lo scarto tra le scelte effettuate oscilla tra 2 e 4 livelli, con una preferenza per 2 valutazioni diverse. Le diverse valutazioni inoltre si riferiscono nella maggior parte dei casi a livelli contigui sulla scala. Questo suggerisce come il QCER possa essere un buon punto di riferimento, anche se sembra non riuscire a guidare in modo chiaro e preciso le scelte degli insegnanti nella distinzione tra livelli contigui, in particolare per la fascia C. Del resto difficoltà per i livelli più alti sono state evidenziate anche in altre ricerche: Kaftandjieva e Takala (2002: 113), ad esempio, sottolineano come il livello C2 non sia sempre distinto dal C1; Cassandro e Maggini (2004: 71), conducendo un'esercitazione con un gruppo di insegnanti, osservano come non manchino divergenze vistose nell'attribuzione dei livelli, specialmente per quanto riguarda la parte alta del continuum. Questo fa presupporre che anche se la formazione appare come un elemento centrale nel garantire l'affidabilità delle valutazioni, non si possono escludere fattori interni al QCER per la scarsa affidabilità dei giudizi, dal momento che anche tra i valutatori esperti si riscontrano problemi di allineamento.

Poiché insegnanti esperti e valutazioni condivise ottengono i risultati migliori, nel definire la procedura di valutazione per questo studio si è cercato di soddisfare entrambe le condizioni. Sono stati coinvolti due valutatori esperti, entrambi regolarmente impegnati come esaminatori in esami di certificazione, oltre che buoni conoscitori del QCER. Ai valutatori è stato richiesto di esprimere un giudizio su complessità, accuratezza e fluenza delle produzioni di tutte le apprendenti del corpus in tre task (intervista, storia e telefonate) realizzati

1. Calcolando livelli e sotto-livelli, dal momento che in molti casi i livelli delle fasce A e B sono ulteriormente suddivisi (per esempio, A1, A1+, A2, A2+).

a t1 e a t4, utilizzando le griglie del QCER predisposte. Sono state così valutate in 3 sessioni di 2 ore ciascuna 36 videoregistrazioni, randomizzate per task, apprendenti e momento di rilevazione. Ai valutatori è stato richiesto di esprimere prima una valutazione individuale e di concordare poi una valutazione condivisa. I risultati della valutazione individuale sono stati utilizzati per verificare l'*inter-rater reliability* dei due valutatori. I risultati sono stati piuttosto buoni: 0,82 per coerenza-coesione, 0,88 per correttezza grammaticale, e 0,84 per fluenza. I risultati della valutazione condivisa sono stati invece utilizzati per osservare la variazione longitudinale e situazionale di complessità, accuratezza e fluenza, oggetto di indagine di questo studio.



Il capitolo è dedicato alla presentazione dei risultati delle analisi effettuate sul corpus. Prima verranno presentati i risultati dell'analisi CAF (§§ 3.1.–3.4.), poi quelli delle valutazioni QCER (§ 3.5.). Infine verranno discusse alcune interazioni tra i due approcci (§ 3.6.).

### **3.1. La complessità sintattica CAF**

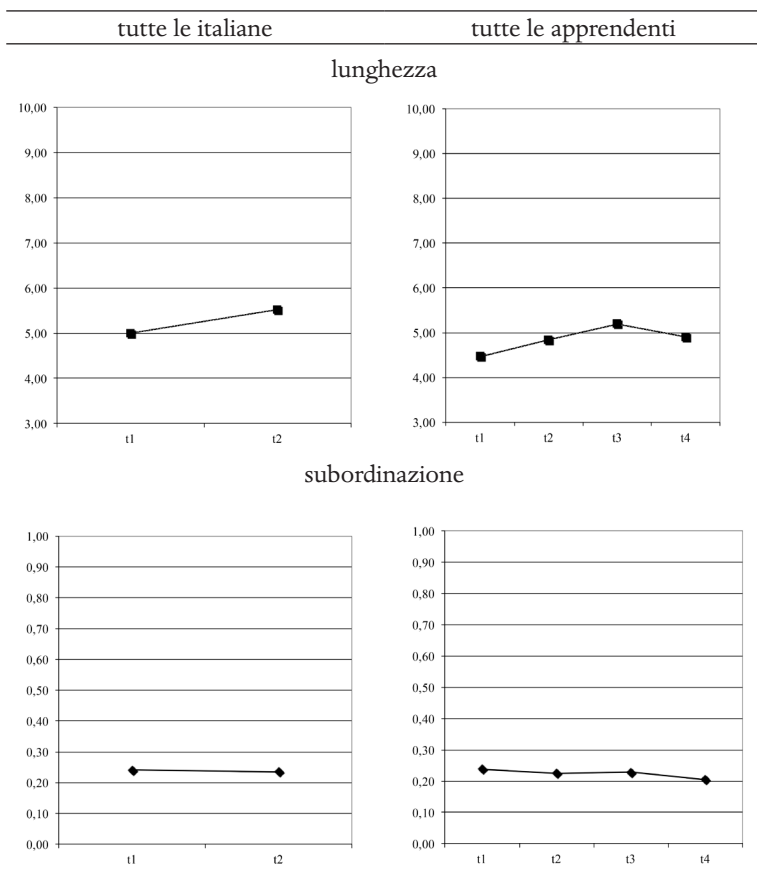
La sezione presenta i risultati relativi all'analisi della complessità sintattica effettuata attraverso l'applicazione delle due misure globali: lunghezza (parole / clausole) e subordinazione (subordinate / AS-unit); e delle sei misure specifiche: tipo di clausola (percentuale di clausole indipendenti, di unità indipendenti inferiori alla clausola e di subordinate sul totale delle clausole) e tipo di subordinata (percentuale di avverbiali, di argomentali e di relative sul totale delle subordinate).

Si inizia ogni volta con l'analisi della variazione longitudinale della complessità sintattica considerando complessivamente i due gruppi di soggetti (studentesse italiane e apprendenti) cumulativamente in tutti i task (§ 3.1.1.); si prosegue con quella della variazione situazionale individuando le differenze tra i quattro task (§ 3.1.2.). Si conclude illustrando, con esempi oltre che con grafici, il percorso individuale di ogni studentessa (§ 3.1.3.).

3.1.1. *La variazione longitudinale*

Il grafico 3.1. mostra i valori delle due misure di lunghezza e di subordinazione relativamente allo sviluppo longitudinale, mentre la tabella 3.1. riporta quelli delle sei misure relative al tipo di clausola e di subordinata. In entrambi i casi i dati si riferiscono ai valori medi dei due gruppi considerati cumulativamente per tutti i task.

**Grafico 3.1.** CAF — Complessità — Variazione longitudinale — Tutti i task — Italiane e apprendenti (valori di gruppo) — Lunghezza e subordinazione.



I dati mostrano come già a t1 le apprendenti sono piuttosto avanzate rispetto alla complessità sintattica. Infatti la misura di subordinazione ottiene gli stessi valori delle compagne italiane, e la misura di lunghezza è di poco inferiore. Alcune differenze si registrano però rispetto alla distribuzione qualitativa di tipi di clausole e di subordinate, poiché le apprendenti producono più unità indipendenti inferiori alla clausola rispetto alle italiane, e utilizzano subordinate relative e argomentali piuttosto che avverbiali. In prospettiva longitudinale si realizzano due tipi di progresso. In primo luogo, rispetto alla lunghezza vi è un aumento dei valori, e in secondo luogo, nonostante per la subordinazione ci sia una lievissima flessione della misura, la distribuzione qualitativa del tipo di clausole e di subordinate si avvicina man mano a quella nativa.

Questo dato è in linea con quanto suggerito in letteratura con l'Ipotesi di Predizione dello Sviluppo proposta da Wolfe–Quintero *et al.* (1998; cfr. § 1.2.2.2.), secondo la quale ai livelli più avanzati lo sviluppo della complessità non riguarda tanto l'aumento della misura di subordinazione, bensì la complessificazione a livello di clausola, con un conseguente aumento della lunghezza media delle clausole, accompagnato in questi dati,

**Tabella 3.1.** CAF — Complessità — Variazione longitudinale — Tutti i task — Italiane e apprendenti (valori di gruppo) — Tipi di clausole e di subordinate.

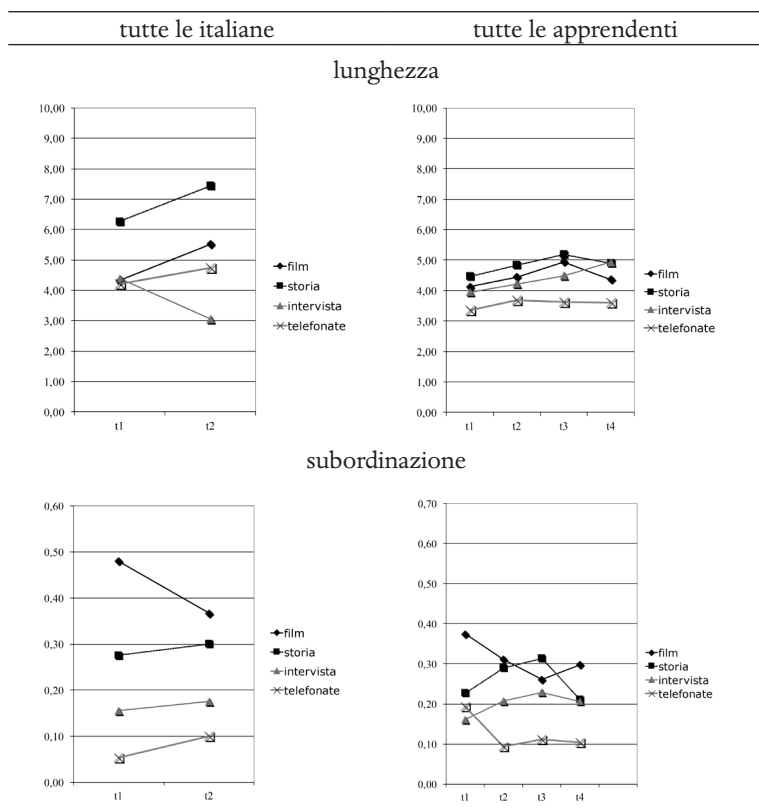
	t1	t2	t1	t2	t3	t4
	<i>italiane</i>		<i>apprendenti</i>			
clausole indep.	61,89%	61,62%	58,30%	61,00%	60,88%	64,39%
unità inf. clausola	19,55%	19,56%	23,58%	21,54%	22,59%	20,76%
clausole sub.	18,56%	18,82%	18,12%	18,62%	17,94%	16,15%
avverbiali/sub.	28,84%	44,44%	42,97%	42,78%	48,23%	54,28%
argomentali/sub.	44,40%	30,27%	28,47%	32,51%	28,60%	24,42%
relative/sub.	26,77%	40,11%	33,74%	26,96%	32,27%	25,45%

anche da un uso più vario di tipi di subordinate. Tali risultati giustificano l'utilità della combinazione dei diversi tipi di misure nell'analisi della complessità sintattica a livello avanzato.

### 3.1.2. La variazione situazionale

Il grafico 3.2. mostra i valori della complessità sintattica delle produzioni dei due gruppi di soggetti relativi alle misure di

**Grafico 3.2.** CAF — Complessità — Variazione situazionale — Distribuzione per task — Italiane e apprendenti (valori di gruppo) — Lunghezza e subordinazione.



**Tabella 3.2.** Complessità — Variazione situazionale — Distribuzione per task — Italiane e apprendenti (valori di gruppo) — Tipi di clausole e di subordinate\*.

	<i>t1</i>	<i>t2</i>	<i>t1</i>	<i>t2</i>	<i>t3</i>	<i>t4</i>
	<i>italiane</i>		<i>apprendenti</i>			
<i>film</i>						
clausole indep.	58,83%	56,63%	63,91%	65,93%	68,07%	66,33%
unità inf. clausola	10,05%	16,35%	9,45%	10,52%	11,37%	10,93%
clausole sub.	31,12%	27,04%	26,65%	23,56%	20,57%	22,74%
avverbiali/sub.	20,19%	28,57%	39,41%	44,37%	59,03%	39,79%
argomentali/sub.	50,97%	36,54%	29,70%	39,41%	24,80%	27,78%
relative/sub.	28,85%	49,18%	30,89%	21,63%	49,50%	32,43%
<i>storia</i>						
clausole indep.	68,25%	72,03%	68,17%	70,26%	66,78%	71,71%
unità inf. clausola	10,69%	4,89%	13,60%	8,04%	12,95%	11,25%
clausole sub.	21,07%	23,09%	18,23%	21,70%	23,51%	17,04%
avverbiali/sub.	35,55%	34,71%	27,34%	35,31%	35,09%	43,11%
argomentali/sub.	38,00%	32,65%	30,43%	27,44%	34,28%	25,19%
relative/sub.	26,46%	32,65%	42,24%	37,25%	30,64%	31,71%
<i>intervista</i>						
clausole indep.	60,54%	51,86%	53,72%	52,42%	56,31%	64,47%
unità inf. clausola	26,12%	32,99%	32,65%	30,50%	25,12%	26,13%
clausole sub.	13,35%	15,16%	13,64%	17,08%	18,58%	15,93%
avverbiali/sub.	30,77%	44,73%	62,16%	48,67%	53,27%	79,96%
argomentali/sub.	44,23%	41,22%	24,21%	30,67%	25,76%	18,93%
relative/sub.	25,00%	14,06%	26,23%	20,67%	20,97%	7,79%
<i>telefonate</i>						
clausole indep.	59,95%	64,42%	47,41%	55,39%	52,37%	55,04%
unità inf. clausola	31,36%	21,66%	38,62%	37,10%	38,50%	36,08%
clausole sub.	8,70%	13,93%	13,97%	10,02%	9,13%	8,89%
avverbiali/sub.	-	-	x	x	x	x
argomentali/sub.	x	x	x	x	x	x
relative/sub.	x	x	x	x	x	x

\* I valori riportati nella tabella si riferiscono ai valori medi delle diverse telefonate del corpus. Poiché questo task non richiede una produzione linguistica consistente, i dati qui forniti devono essere considerati con una certa cautela. Viste poi le scarse occorrenze di subordinate nel corpus, indicare la percentuale dei tipi di subordinate sarebbe stato poco significativo e avrebbe dato una lettura falsata dei risultati, per questo nella tabella si è indicata esclusivamente la presenza (x) o meno (-) della struttura.

lunghezza e di subordinazione. La tabella 3.2. riporta invece i dati relativi alle misure del tipo di clausola e di subordinata nei quattro task. In entrambi i casi vengono forniti, task per task, i valori medi dei due gruppi.

I dati disaggregati per task offrono un quadro in parte diverso da quello descritto con la sola longitudinalità in § 3.1.1. Vediamo nel dettaglio la variazione situazionale segnata da italiane e apprendenti.

Le studentesse italiane dimostrano un certo grado di variazione tra task: i task monologici film e storia stimolano la produzione di clausole più lunghe e di AS-unit più complesse, con un uso di diversi tipi di clausole e di subordinate, quest'ultime soprattutto avverbiali, mentre i task più interattivi intervista e telefonate fanno realizzare AS-unit più semplici e brevi, con un maggior uso di unità indipendenti inferiori alla clausola e un uso ridotto di subordinate, soprattutto argomentali o relative. Questo dato è in accordo con gli studi di approccio psicolinguistico (§ 1.1.2.), secondo cui i diversi task pongono diverse richieste ai parlanti in termini di complessità della lingua prodotta. Nella ricostruzione di un racconto il flusso della conversazione è gestito quasi esclusivamente dal soggetto e l'intervistatore si limita a segnali di assenso. Quindi la narrazione procede con frasi piuttosto lunghe e maggiore subordinazione. Nei due task interattivi, seppure a livelli diversi, la conversazione richiede invece una maggior collaborazione tra soggetto e interlocutore, favorendo quindi una maggior alternanza di turni e l'uso di AS-unit più brevi e semplici.

Longitudinalmente, le studentesse italiane mostrano alcuni cambiamenti. Nei task monologici si riduce l'uso di subordinate a favore di clausole un po' più lunghe, mentre nei task interattivi non si verifica alcun cambiamento sostanziale, se non una riduzione in lunghezza per l'intervista e un lieve aumento in complessità per le telefonate. Si vedrà in § 3.1.3.

come per le telefonate l'incremento sia legato all'uso di formule più complesse di richiesta.

Vediamo ora i dati relativi alle apprendenti. A T1, differenziandosi dalle compagne italiane, dimostrano non solo molta meno variazione tra task, ma anche una cruciale differenza nella direzione di tale variazione. Infatti l'apertura delle telefonate ottiene valori di subordinazione pari a quelli delle due narrazioni. Se nei task monologici e nell'intervista le differenze rispetto alle italiane sono minime, in quello di apertura delle telefonate la distanza dalle compagne italiane è piuttosto evidente. Vedremo in dettaglio attraverso estratti del corpus in § 3.1.3. come le apprendenti in questo task producono sia un maggior numero che una maggior varietà di tipi di subordinate rispetto alle compagne italiane: esse tendono a condensare diverse informazioni in una stessa unità sintattica piuttosto che gestire gli scambi della routine dell'apertura di una telefonata di servizio con enunciati rapidi e brevi. In altre parole, le apprendenti dimostrano di aver già acquisito la capacità di produrre costruzioni complesse, sanno realizzare subordinate e riescono a condensare un alto numero di parole in un unico progetto sintattico. Tuttavia, esse paiono meno in grado di scegliere quando utilizzare tali costruzioni. Gestire efficacemente l'apertura di una telefonata non implica solamente il controllo delle regole sintattiche, ma anche la capacità di saper produrre in rapida successione turni che siano contingenti rispetto alle produzioni altrui. Questo tipo di abilità è ancora carente a T1 per le nostre apprendenti, per quanto di livello relativamente avanzato.

In prospettiva longitudinale, le apprendenti realizzano diversi progressi differenziando come le native tra task: la misura di lunghezza registra lievi aumenti con traiettorie lineari in tutti e quattro i task, confermando dunque il quadro definito

dai valori di gruppo, mentre la misura di subordinazione e la distribuzione del tipo di subordinate varia, seguendo traiettorie che si differenziano per intensità e direzione da task a task, ma ottenendo a  $t_4$  in tutti i casi un avvicinamento al quadro di variazione definito per le italiane. Sebbene le apprendenti differenzino meno delle compagne italiane, all'ultima rilevazione lo fanno allo stesso modo. Vediamo in dettaglio la variazione segnata da ogni task.

Così come per le italiane, nei task monologici la traiettoria di sviluppo non è lineare, nel film la misura di subordinazione decresce tra  $t_1$  e  $t_3$ , per poi aumentare nuovamente, mentre nella storia l'andamento è opposto. I cambiamenti quantitativi sono accompagnati nel task storia da avvicinamenti qualitativi della distribuzione dei tipi di clausole di subordinate; nel task film la riduzione in subordinazione è invece controbilanciata da un aumento della lunghezza delle clausole prodotte.

Diverso l'andamento nei task interattivi. Nel task intervista lo sviluppo è più lineare, le apprendenti dimostrano un generale aumento tra  $t_1$  e  $t_4$ , sia della lunghezza che della subordinazione, oltre che un uso più vario di tipi di subordinate. È bene sottolineare come in questo task le apprendenti parlino addirittura di più rispetto alle compagne italiane, dando più informazioni e richiedendo un minor numero di domande o interventi da parte dell'intervistatore.

Infine nel task telefonate l'avvicinamento al quadro di variazione nativo è ottenuto attraverso una riduzione lineare della misura di subordinazione e dei tipi di subordinate impiegate.

Ricapitolando, in § 3.1.1 abbiamo visto come la complessità misurata in termini di subordinazione possa svilupparsi secondo percorsi non lineari a livelli avanzati (cfr. Ipotesi di Pre-dizione dello Sviluppo, § 1.2.2.2.). I dati disaggregati per task confermano questa tendenza nei task monologici, ma non in



quelli interattivi. Questo risultato suggerisce quindi una seconda possibile ragione dell'andamento non lineare di questa misura: il tipo di attività svolta. Come abbiamo visto, nel task interattivo l'alto valore di subordinazione evidenzia incompetenza sul piano interazionale: trovandosi a gestire uno scambio rapido e contingente come quello richiesto dall'apertura di una telefonata, le nostre apprendenti, ansiose di dire subito tutto il dovuto, a t1 preferiscono condensare le informazioni in unità complesse, che probabilmente reputano di potere gestire meglio, piuttosto che realizzare turni brevi e incalzanti rispetto a quelli dell'interlocutore. A t2 e t3 il controllo più nativo del task è ottenuto di conseguenza attraverso una riduzione della complessità sintattica. Il processo di apprendimento di una lingua non è infatti caratterizzato semplicemente dalla crescita del repertorio sintattico dell'apprendente, ma anche dall'abilità ad utilizzarlo appropriatamente in una varietà di situazioni.

### 3.1.3. *La variazione individuale*

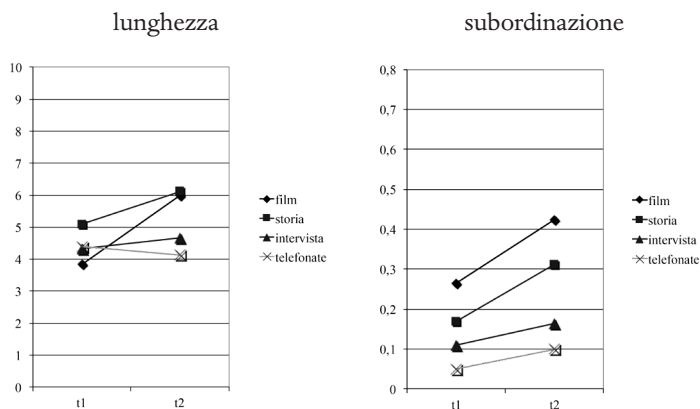
I grafici 3.3. e 3.4. mostrano soggetto per soggetto i valori della complessità relativi alle misure di lunghezza e di subordinazione, e le tabelle 3.3.–3.8. quelli relativi alle sei misure del tipo di clausola e di subordinata. La discussione delle cifre è accompagnata dall'illustrazione di esempi dal corpus.

Il grafico 3.3. e le tabelle 3.3. e 3.4. riportano i valori individuali ottenuti dalle due studentesse italiane. Pur confermando il quadro di variazione identificato in § 3.1.2., si registra variazione individuale soprattutto nei task monologici. In questi task alla prima rilevazione Elisa sviluppa narrazioni brevi e semplici, organizzate attraverso la giustapposizione di clausole indipendenti, Valentina, al contrario, produce racconti già lunghi e sintatticamente complessi, in particolare nella narrazione del task relativo al film. Le differenze quantitative sono

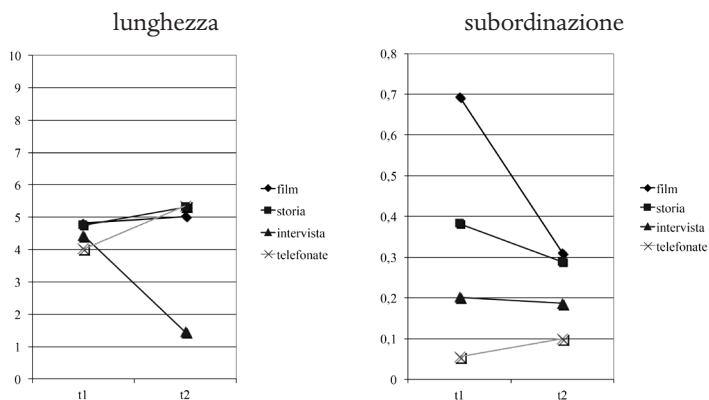
accompagnate da differenze qualitative della distribuzione di tipi di clausole e di subordinate. In prospettiva longitudinale si riduce la variazione individuale.

**Grafico 3.3.** CAF — Complessità — Variazione individuale— Distribuzione per task — Italiane — Lunghezza e subordinazione.

*Elisa*



*Valentina*



**Tabella 3.3.** CAF — Complessità — Variazione individuale — Distribuzione per task — Elisa — Tipi di clausole e di subordinate.

	<i>t1</i>		<i>t2</i>		<i>t1</i>		<i>t2</i>	
	<i>film</i>				<i>storia</i>			
clausole indep.	39	62,90%	27	58,70%	59	77,63%	29	69,05%
unità inf. clausola	10	16,13%	5	10,87%	6	7,89%	3	7,14%
clausole sub.	13	20,97%	14	30,43%	11	14,47%	10	23,81%
avverbiali/sub.	1	7,69%	4	28,57%	5	45,45%	4	40,00%
argomentali/sub.	7	53,85%	7	50,00%	3	27,27%	3	30,00%
relative/sub.	5	38,46%	3	21,43%	3	27,27%	3	30,00%
	<i>intervista</i>				<i>telefonate</i>			
clausole indep.	80	61,07%	40	56,34%	3,25	60,41%	4,00	59,26%
unità inf. clausola	38	29,01%	21	29,58%	1,63	30,30%	1,75	25,93%
clausole sub.	13	9,92%	10	14,08%	0,50	9,29%	1,00	14,81%
avverbiali/sub.	3	23,08%	3	30,00%	-	-	-	-
argomentali/sub.	5	38,46%	5	50,00%	x	-	x	-
relative/sub.	5	38,46%	2	20,00%	-	-	x	-

**Tabella 3.4.** CAF — Complessità — Variazione individuale — Distribuzione per task — Valentina — Tipi di clausole e di subordinate.

	<i>t1</i>		<i>t2</i>		<i>t1</i>		<i>t2</i>	
	<i>film</i>				<i>storia</i>			
clausole indep.	69	54,76%	30	54,55%	83	58,87%	57	75,00%
unità inf. clausola	5	3,97%	12	21,82%	19	13,48%	2	2,63%
clausole sub.	52	41,27%	13	23,64%	39	27,66%	17	22,37%
avverbiali/sub.	17	32,69%	-	-	10	25,64%	5	29,41%
argomentali/sub.	25	48,08%	3	23,08%	19	48,72%	6	35,29%
relative/sub.	10	19,23%	10	76,92%	10	25,64%	6	35,29%
	<i>intervista</i>				<i>telefonate</i>			
clausole indep.	93	60,00%	108	47,37%	3,67	59,48%	4,00	69,57%
unità inf. clausola	36	23,23%	83	36,40%	2,00	32,41%	1,00	17,39%
clausole sub.	26	16,77%	37	16,23%	0,50	8,10%	0,75	13,04%
avverbiali/sub.	10	38,46%	22	59,46%	-	-	-	-
argomentali/sub.	13	50,00%	12	32,43%	x	-	x	-
relative/sub.	3	11,54%	3	8,11%	x	-	x	-

Vediamo attraverso alcuni esempi come le italiane realizzano i quattro task e quali sono le tendenze della variazione individuale. Nei task monologici, alla prima rilevazione le due studentesse italiane utilizzano stili e strategie diverse. Elisa sviluppa una narrazione sintetica, riporta solo i punti principali della trama attraverso la giustapposizione o la coordinazione di AS-unit brevi e semplici, così come illustrato dall'esempio (40) relativo al task film a tr.

- (40) ELI: || | {eh. che dice che:: # lui e: tch vedo che} cioè ho visto  
| che va:: {a trovar un va} a cercare un lavoro? | ||
- INV: mhm #I\_4
- ELI: || | e però va a finir male. | || #
- ELI: || | fa::: un casino | ||
- ELI: || | diciamo | ||
- ELI: || | c'è una nave | {che:: ℒhhhℒ}
- INV: ℒhhhℒ
- ELI: .hh ℒhhℒ che:: se ne va | ||
- ELI: || | diciamo: | ||
- ELI: || | lui s- muove un: e:: | || #I\_6
- ELI: || | non lo so. | ||
- ELI: || | comunque lui lavora in un cantiere tipo | ||
- INV: mhm
- ELI: || | insomma alla fine lo cacciano via? | || #o\_6
- ELI: || | poi::: per la strada incontra questa ragazza {che è:  
#o\_7} | ||
- ELI: || | {non ha: mh} cioè è una ragazza povera:: | ||
- ELI: || | non ha da mangiare:: | ||
- ELI: || | è da sola:: | ||
- ELI: || | così. | ||
- ELI: || | che:: ruba:: un filoncino di pane::? | || #o\_4

Qui Elisa racconta una delle scene iniziali del video: la studentessa riporta informazioni generiche sugli eventi, senza scendere nei dettagli. Realizza due subordinate, la prima argomentale retta da *vedere* e la seconda realizzata da una relativa introdotta dal presentativo *c'è*, non certo considerabili esempi di subordinazione sofisticata. La lunghezza media delle clausole è di 4,5 parole. L'utilizzo di segnali discorsivi come *comunque*, *insomma* e *diciamo* o l'uso del *che* come connettore generico compensano però la brevità delle clausole e l'uso ridotto di subordinazione, generando frasi in cui si estende l'uso di indicatori di dipendenza da un livello sintattico a un più ampio livello discorsivo e pragmatico.

Diversamente dalla compagna, Valentina realizza i due task monologici producendo resoconti lunghi il doppio. La sua narrazione è piuttosto dettagliata, caratterizzata da un uso importante della subordinazione come evidente dall'estratto (41), realizzato nel task film a T1. Qui la studentessa riporta la stessa scena presa ad esempio per la compagna realizzando però una narrazione in cui le informazioni sono impacchettate entro AS-unit più lunghe e complesse.

- (41) VAL: || | e::m tch come lavoro gli danno::: | quello di::: andare::: mh tch in una specie {di::: h #o\_5 cioè di me di::: h} di porto | dove ci son tutte delle navi::: #o\_4 così ||  
 VAL: || | e c'è anche un uomo | ||  
 INV: in un cantiere  
 VAL: || {un cant-} sì | ||  
 INV: cantiere navale tipo  
 VAL: || | sì un cantiere navale così | || h h  
 VAL: || | e c'è un uomo? | in pratica che:: gli dice | che deve iniziare a lavorare | per vedere | se {era} appunto::: h riusciva::: | a trovarsi bene | ||  
 VAL: || | doveva portargli un cuneo | || #o\_4

INV: m m  
 VAL: || | poi #o\_4 subito:..... gli prende un cuneo: lungo | ||  
 #o\_3  
 VAL: || | che in pratica {era::} il suo datore gli aveva detto | di  
 portargliene uno: triangolare così | ||  
 INV: mh mh #o\_8  
 VAL: || | hh e:..... dice | ||  
 VAL: || | vallo a cercare | || #o\_2  
 VAL: || | e lui va a cercarlo | || #o\_6  
 VAL: || | ne prende uno | ||  
 VAL: || | solo che dopo ℒparte la naveℒ he ?ℒhℒ h | || # 1\_o  
 VAL: || | quindi in pratica ha fatto un guaio h | || #o\_8  
 VAL: || | e quindi:: l'hanno licenziato | ||  
 INV: ?ℒhℒ h? ah#o\_6  
 VAL: || | tch h e:m dopo che::: l'hanno licenziato | in pratica:  
 sta passeggiando lungo la strada h | quando vede: una  
 ragazza #o\_2 | che: {sta ruba cioè} sta rubando un filon-  
 cino di pane | ||

La codifica del tipo di clausole e l'analisi della loro frequenza confermano le tendenze osservare finora (tabelle 3.3.–3.4.). Nel racconto di Elisa, la ragazza che parla meno e con meno subordinate, viene preferito l'uso di clausole indipendenti (62,90%) e unità indipendenti inferiori alla clausola (16,13%) rispetto alle clausole subordinate, il cui uso è ridotto a poche occorrenze (20,97%), prevalentemente argomentali e relative. Abbiamo visto precedentemente come lo scarso uso di subordinate sia compensato dall'uso di clausole indipendenti introdotte dal connettore *che* con la funzione di congiunzione semanticamente e sintatticamente “lasca” e da segnali discorsivi. A questi si aggiunge l'uso frequente di argomentali, utile comunque a dare una certa coesione al racconto. Nell'esempio (42) l'indipendente *si vedono in questa casa* seguito da due AS-unit intro-

dotte dal complementatore generico *che*: *che stanno bene* e *che c'è la mucca*. Le due AS-unit sono state prodotte come aggiunte a un enunciato che avrebbe potuto terminare con *tutta la roba da mangiare e tutto*, il loro iniziare con il complementatore generico *che* mostra comunque l'intenzione di Elisa di dare una certa impressione di complessità, collegando alla narrazione principale l'enunciazione di elementi ad essa correlati pragmaticamente e semanticamente, ma non sintatticamente:

- (42) ELI: con: tutta: la roba da mangiare e tutto #o\_5 | ||  
 ELI: || | che stanno bene | ||  
 ELI: || |  $\mathcal{L}h\mathcal{L}$  #che c'è la mucca | che fa il latte | || [he

Diversamente, Valentina, la studentessa italiana che realizza il racconto più lungo e complesso, fa un uso ridotto di unità inferiori alla clausola (3,97%), a favore di clausole indipendenti (54,76%) e clausole subordinate (30,43%), queste ultime di vario tipo, argomentali, avverbiali e relative. Prevalgono le subordinate argomentali e avverbiali, ridotto l'uso di relative.

A t2, emergono nelle due studentesse italiane alcuni cambiamenti, con una riduzione delle differenze individuali. Valentina ottiene valori vistosamente più bassi nella misura relativa alla subordinazione, mentre Elisa aumenta la complessità delle produzioni sia rispetto a lunghezza che subordinazione. La riduzione ottenuta da Valentina è dovuta ad alcune difficoltà della studentessa nel seguire il filmato video. Di conseguenza il suo racconto è poco preciso, poco dettagliato e soprattutto influenzato da interventi costanti dell'intervistatore che cerca di aiutarla a concludere il task. Questo dato mostra come la variazione possa essere legata in alcuni casi ad elementi esterni al task. A t2 Elisa dimostra invece diversi cambiamenti dal punto di vista sia quantitativo che qualitativo: pur mantenendo una certa essenzialità, il suo racconto è

un po' più lungo e dettagliato e raggiunge valori di complessità maggiori che a T1. La studentessa produce clausole più lunghe e AS-unit più complesse, come illustrato dall'esempio (43). I cambiamenti quantitativi sono accompagnati da variazioni qualitative della distribuzione dei tipi di subordinate: si riduce l'uso di relative, a favore di avverbiali e argomentali.

- (43) ELI: || | e:: # solo che alla fine:: si:: si vede che:: | ||  
 ELI: || | e::hm praticamente il cacciatore e la pantera rosa  
 si ritrovano insieme ancora | a scappare dal cane | che  
 c'ha {la::} # la dinamite in bocca ℒhhhℒ | ||  
 INV: ℒhhℒ ℒokeyℒ  
 ELI: || | {e poi:: e:: # e::hm ah} sì | || #  
 ELI: || | poi dopo:: alla fine:: ehm {il cacciatore} # il cane  
 riesce | a raggiungerlo | ||#  
 ELI: || | solo che:: gli mette la dinamite davanti | ||  
 ELI: || | lui cerca | di coprirlo con della terra? | || #  
 ELI: || | e:: solo che esplose tutto | ||  
 ELI: || | e viene un buco nell'isola? | || #o\_6  
 ELI: || | e {sprofonda::} ℒhhhℒ sprofonda nel mare | ||  
 ELI: || | e si vede:: l'ultima scena | che si ritrovano su un  
 pezzo di legno | a galleggiare:: | ||  
 INV: tutti insieme  
 ELI: || | e:: la pantera rosa è sempre dentro l'acqua con un  
 ombrello | ||  
 INV: okey  
 ELI: || | che non ho capito perché. | ||  
 ELI: || | e poi ci sono:: # il cacciatore e il cane su un pezzo  
 di legno | ||

Il task intervista è per sua natura più interattivo di quelli fin qui considerati. In questo caso intervistatore e apprendente sono coinvolti in una conversazione a due. In genere



l'intervistatore guida l'interazione introducendo domande e lasciando alla studentessa lunghi turni di risposta, così come in (44) e in (45), dove si riportano due estratti prodotti rispettivamente da Elisa e Valentina a T1.

(44) INV: mh. #o\_6 e che tipo di persone sono. se dovessi descrivermele?

ELI: || | allora; #o\_4 mio padre è una persona un po': in molto introverso, f\*sempre\*f zitto, | || #

INV: mh mh.

ELI: || | tranquillo. | ||

ELI: || | non mi: f\*cioè\*f insomma andiam d'accor. | ||

ELI: || | mia madre invece è tutto il contrario | ||

ELI: || | è, # sempre {di f estrov estrov} estroversissima. | ||

ELI: || | parla con tutti, | || #

ELI: || | son due persone oppostissime | ||

ELI: || | ma: van d'accordo. | || #o\_6

(45) INV: cosa gli fai fare?

VAL: || | di solito:/ guardiam la televisione insieme: | || #

VAL: || | oppure a volte vuole | che gli faccio fare degli esercizi # tipo problemi: # dettati: così | || #

INV: !a::h! # come si chiama #

VAL: || | giuseppe | || #

INV: °giuseppe° # e:: quando guardate la televisione che cosa guardate insieme #

VAL: || | guardiamo: i ℒcartoni animatiℒ | || #

INV: i cartoni animati tu devi adattarti ℒalleℒ

VAL: || | sì: | || #

VAL: || | oppure # e perché per esempio a volte mettono i canali della musica così | || #

VAL: || | solo che a lui non gli piace molto | ||

VAL: || | quindi: # guardiamo i cartoni animati. | || #

In diversi punti però la conversazione porta a scambi più rapidi, è il caso in cui ad esempio l'intervistatore richiede approfondimenti, chiarimenti o commenta quanto detto dalla studentessa. Dal punto di vista sintattico il task favorisce l'uso di AS-unit semplici costituite di clausole indipendenti o unità indipendenti inferiori alla clausola e l'introduzione di un argomento per volta. La distribuzione qualitativa dei tipi di subordinate mostra come, nonostante l'uso quantitativamente ridotto, siano realizzati vari tipi di subordinazione.

In prospettiva longitudinale per Elisa non si registrano cambiamenti sostanziali rispetto alle misure di complessità, Valentina riporta invece una rapida riduzione della lunghezza media delle clausole. Questo è dovuto all'andamento del flusso conversazionale. Valentina a t2 risulta meno disposta a raccontarsi rispetto a t1, spesso risponde alle domande in modo vago, con una o due AS-unit semplici, richiedendo diversi interventi da parte dell'intervistatore e dunque maggiore interazione, turni più brevi, fatti di unità indipendenti inferiori alla clausola. Anche in questo caso le produzioni di Valentina mostrano come fattori esterni al task, quali la motivazione o l'interesse verso l'attività, possono causare variazione individuale.

Infine, il task di apertura delle telefonate è quello che mostra maggior variazione situazionale rispetto alla complessità, ma minima variazione individuale: le due studentesse italiane realizzano in modo sostanzialmente simile il task. Vediamone alcuni esempi.

Le telefonate di servizio rappresentano un genere ben definito tra le conversazioni telefoniche, avendo lo scopo di richiedere e fornire dei servizi, quali informazioni, prenotazioni o assistenza. L'apertura è caratterizzata da una forte compressione delle mosse comunicative, che sono spesso ridotte all'essenziale e si basano su alcune formule di routine.

La letteratura sulle aperture delle telefonate (ad es. Schegloff, 1979; 1986; per l'italiano Bercelli, Pallotti, 2002) ha evidenziato la presenza ricorrente di alcune mosse:

- a) apertura del canale: chi chiama manda un segnale (lo squillo del telefono) per verificare se c'è qualcuno dall'altro capo del filo; chi risponde, dicendo *pronto* o fornendo la propria identità, mostra che qualcuno c'è ed è pronto a interagire;
- b) identificazione: il chiamante controlla se ha effettivamente raggiunto la persona desiderata ricevendo una risposta positiva; passa poi all'auto-identificazione, non sempre presente nelle chiamate di servizio a negozi, dove invece è spesso inclusa nella risposta del ricevente;
- c) saluti: possono essere prodotti dal chiamante, dal ricevente o da entrambi;
- d) interessamenti: si trovano esclusivamente nelle telefonate tra intimi o in alcuni programmi radio-televisivi, mentre mancano nelle telefonate di servizio;
- e) entrare nel merito: esauriti i cerimoniali di apertura, i partecipanti si orientano al motivo della telefonata, introducendo il primo topic della conversazione.

La richiesta di informazioni è generalmente introdotta da pre-richieste e mitigata attraverso forme indirette o modificatori.

A TI Elisa segue la struttura prima delineata: dopo l'apertura del canale, saluta, si auto-identifica ed entra nel merito mediante una serie di turni molto brevi, che introducono un argomento per volta e sono contingenti rispetto a quelli dell'interlocutore, creando così una punteggiatura piuttosto rapida e staccata del flusso conversazionale, come illustra-

to dall'esempio (46). Qui dopo le mosse iniziali, Elisa entra nel merito introducendo prima della richiesta vera e propria pre-richieste distribuite su diverse AS-unit e diversi turni. La densità informativa di ciascuna AS-unit è molto bassa e anche quando si introduce la richiesta vera e propria, essa è formulata nel modo più generico possibile.

- (46) ELI: ((squillo))  
 AG4: pronto?  
 ELI: || | {e;} buongiorno. | ||  
 AG4: sì #  
 ELI: || | siamo dieci persone | ||  
 ELI: || | e dovremmo andare a fare: un viaggio a londra. | || #  
 AG4: sì #  
 ELI: || | volevo alcune informazioni? | ||  
 AG4: sì.  
 ELI: || | {hh allora e:: mh: come: cioè } in aereo {come:} #  
 come:: siete organizzati | || #o\_2

Valentina, realizza l'apertura delle telefonate di servizio in modo analogo, come illustrato nell'esempio (47). Dopo l'apertura del canale e i saluti, l'entrata in merito viene formulata in 4 turni e 4 AS-unit, di cui 2 costituite da unità indipendenti inferiori alla clausola, 1 da una clausola indipendente con verbo a costruito complesso e 1 AS-unit che realizza una frase scissa con topicalizzazione. La richiesta è poi realizzata in modo molto generico.

- (47) VAL: ((squillo))  
 AG2: moito viaggi #  
 VAL: || | buonasera | ||  
 AG2: sì  
 VAL: || | volevo chiederle informazione | || #

AG2: sì  
 VAL: || | e: per Londra | || #  
 AG2: sì  
 VAL: || | una classe di: ragazzi | ||  
 AG2: mh mh  
 VAL: || | e: qualcosa di conveniente {che::} # che c'è | ||  
 AG2: 「allora intanto mi dovresti dire le date # il quantitativo  
 di persone

L'analisi qualitativa della distribuzione del tipo di clausole (tabelle 3.3.-3.4.) conferma il quadro descritto finora. Le due studentesse utilizzano in prevalenza AS-uni brevi e semplici, fatte di clausole indipendenti e unità indipendenti inferiori alla clausola, fanno un uso limitato della subordinazione, e anche nel caso in cui una clausola subordinata viene prodotta, questa in genere non compare mai nei primi turni ed è sempre preceduta dalle mosse iniziali, apertura, identificazione e saluti, realizzate con turni brevi e incalzanti, che introducono gradualmente l'atto di richiesta.

In prospettiva longitudinale si conferma il quadro delineato fin ora. Si registra un lieve aumento della misura di subordinazione in tre telefonate del corpus. Queste si differenziano dalle altre per lunghezza ed elaborazione, così come esemplificato in (48) prodotto da Elisa a t2.

- (48) ELI: ((squillo))  
 SHI: mediastore buongiorno sono stefano  
 ELI: || | buongiorno | ||  
 ELI: || | volevo chiederle un'informazione | ||  
 SHI: mi dica  
 ELI: || | dovrei:: {consigliare} cioè regalare un cd ad una signora di quarant'anni | ||  
 SHI: sì

- ELI: || | tch che::: comunque ama la musica # | ||  
 ELI: || | ascolta musica # da viaggio soprattutto | ||  
 ELI: || | e io avrei una lista di di di gruppi | ||  
 ELI: || | però non saprei::: | ||  
 SHI: ma una lista lunga  
 ELI: || | no:: | ||  
 ELI: || | son quattro gruppi | ||  
 ELI: || | però non so |{cosa:::} cosa sia meglio | || #  
 SHI: beh bisogna vedere che cosa ascolta lei generalmente  
 ELI: ma::: musica::: lei viaggia molto quindi:::

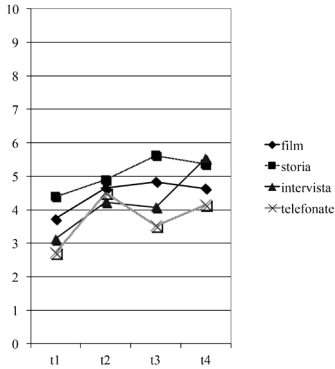
Dopo l'apertura del canale e i saluti, Elisa entra nel merito e prepara la realizzazione della richiesta in 4 turni e 9 AS-unit. La richiesta è realizzata in modo generico e indiretto *però non so* |{cosa:::} *cosa sia meglio*, con una AS-unit di due clausole, di cui una costituita da una subordinata argomentale. Interessante osservare come anche in questo caso l'AS-unit complessa compaia solo al quarto turno.

Vediamo ora la variazione disegnata dalle apprendenti. Il grafico 3.4. e le tabelle 3.5.-3.8. riportano i valori ottenuti dalle apprendenti nelle misure considerate. Una prima osservazione globale dei grafici ci permette da un lato di confermare il quadro di variazione identificato per il gruppo, dall'altro di osservare variazione individuale anche tra le apprendenti, soprattutto rispetto alla misura di subordinazione e a quelle relative alla distribuzione di tipi di clausole e di subordinate.

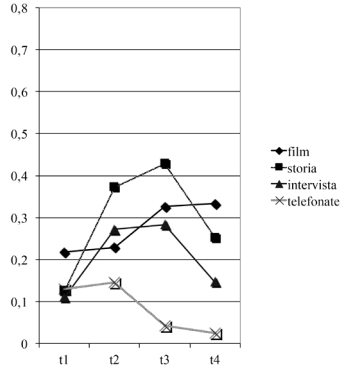
**Grafico 3.4.** CAF — Complessità — Variazione individuale— Distribuzione per task — Apprendenti — Lunghezza e subordinazione.

*Pandita*

lunghezza

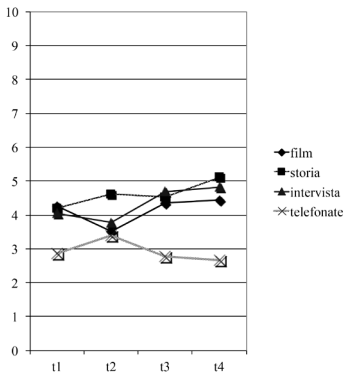


subordinazione

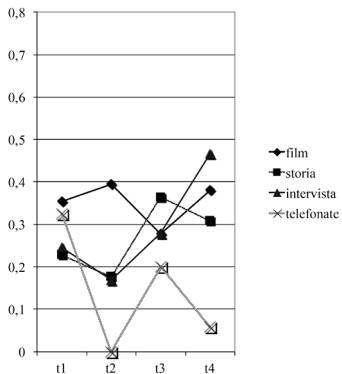


*Catherine*

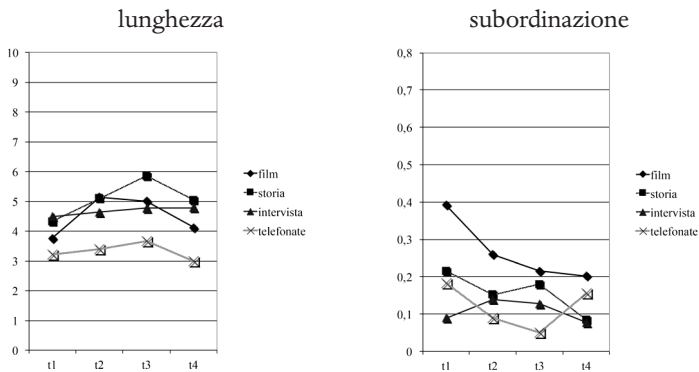
lunghezza



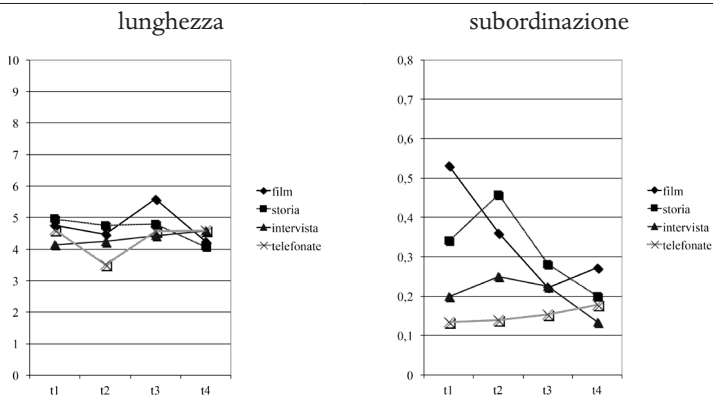
subordinazione



## Eden



## Shirley



**Tabella 3.5.** CAF — Complessità — Variazione individuale — Distribuzione per task — Pandita — Tipi di clausole.

	t1		t2		t3		t4	
	<i>film</i>							
clausole indep.	127	66,49%	102	76,12%	134	68,72%	154	69,68%
unità inf. clausola	30	15,71%	7	5,22%	13	6,67%	11	4,98%
clausole sub.	34	17,80%	25	18,66%	48	24,62%	56	25,34%
avverbiali/sub.	6	17,65%	10	40,00%	25	52,08%	25	44,64%
argomentali/sub.	12	35,29%	11	44,00%	10	20,83%	11	19,64%
relative/sub.	16	47,06%	4	16,00%	13	27,08%	20	35,71%



<i>storia</i>								
clausole indep.	122	75,78%	58	71,60%	49	70,00%	82	66,13%
unità inf. clausola	21	13,04%	1	1,23%	-	-	17	13,71%
clausole sub.	18	11,18%	22	27,16%	21	30,00%	25	20,16%
avverbiali/sub.	5	27,78%	11	50,00%	13	61,90%	9	36,00%
argomentali/sub.	4	22,22%	6	27,27%	4	19,05%	3	12,00%
relative/sub.	9	50,00%	5	22,73%	4	19,05%	13	52,00%
<i>intervista</i>								
clausole indep.	80	49,38%	126	57,53%	179	52,19%	34	87,18%
unità inf. clausola	66	40,74%	46	21,00%	87	25,36%	-	-
clausole sub.	16	9,88%	47	21,46%	77	22,45%	5	12,82%
avverbiali/sub.	6	37,50%	15	31,91%	43	55,84%	5	100%
argomentali/sub.	3	18,75%	21	44,68%	16	20,78%	-	-
relative/sub.	7	43,75%	11	23,40%	18	23,38%	-	-
<i>telefonate</i>								
clausole indep.	2,17	48,22%	5,25	61,63%	3,43	52,15%	5,80	74,22%
unità inf. clausola	1,83	40,67%	2,50	27,12%	3,00	44,22%	1,60	23,56%
clausole sub.	0,50	11,11%	0,75	11,25%	0,29	3,63%	0,20	2,22%
avverbiali/sub.	x	-	-	-	-	x	-	-
argomentali/sub.	x	-	x	-	-	x	-	-
relative/sub.	-	-	x	-	-	-	-	x

**Tabella 3.6.** CAF — Complessità — Variazione individuale — Distribuzione per task — Catherine — Tipi di clausole e di subordinate.

	<i>t1</i>	<i>t2</i>	<i>t3</i>	<i>t4</i>
<i>film</i>				
clausole indep.	44	67,69%	32	60,38%
unità inf. clausola	4	6,15%	6	11,32%
clausole sub.	17	26,15%	15	28,30%
avverbiali/sub.	3	17,65%	3	20,00%
argomentali/sub.	9	52,94%	8	53,33%
relative/sub.	5	29,41%	4	26,67%
<i>storia</i>				
clausole indep.	105	67,31%	97	72,93%
unità inf. clausola	22	14,10%	16	12,03%
clausole sub.	29	18,59%	20	15,04%
avverbiali/sub.	8	27,59%	7	35,00%
argomentali/sub.	16	55,17%	9	45,00%
relative/sub.	5	17,24%	4	20,00%

<i>intervista</i>								
clausole indep.	90	54,22%	58	46,40%	86	65,15%	38	57,58%
unità inf. clausola	43	25,90%	49	39,20%	17	12,88%	7	10,61%
clausole sub.	33	19,88%	18	14,40%	29	21,97%	21	31,82%
avverbiali/sub.	19	57,58%	9	50,00%	14	48,28%	16	76,19%
argomentali/sub.	6	18,18%	6	33,33%	11	37,93%	2	9,52%
relative/sub.	8	24,24%	3	16,67%	4	13,79%	3	14,29%
<i>telefonate</i>								
clausole indep.	1,75	29,58%	4,00	80,00%	3,17	38,67%	3,60	46,03%
unità inf. clausola	2,50	49,31%	1,00	20,00%	3,50	44,43%	3,60	48,89%
clausole sub.	1,50	21,11%	-	-	1,50	16,90%	0,40	5,08%
avverbiali/sub.	x	-	-	-	-	-	-	-
argomentali/sub.	x	-	-	x	-	-	-	-
relative/sub.	x	-	-	x	-	x	-	x

**Tabella 3.7.** CAF — Complessità — Variazione individuale — Distribuzione per task — Eden — Tipi di clausole e di subordinate.

	<i>t1</i>		<i>t2</i>		<i>t3</i>		<i>t4</i>	
<i>film</i>								
clausole indep.	66	59,46%	75	74,26%	59	74,68%	84	69,42%
unità inf. clausola	14	12,61%	5	4,95%	6	7,59%	17	14,05%
clausole sub.	31	27,93%	21	20,79%	14	17,72%	20	16,53%
avverbiali/sub.	18	58,06%	13	61,90%	7	50,00%	6	30,00%
argomentali/sub.	8	25,81%	8	38,10%	4	28,57%	7	35,00%
relative/sub.	5	16,13%	-	-	3	21,43%	7	35,00%
<i>storia</i>								
clausole indep.	80	64,52%	94	77,69%	36	78,26%	51	79,69%
unità inf. clausola	22	17,74%	11	9,09%	3	6,52%	8	12,50%
clausole sub.	22	17,74%	16	13,22%	7	15,22%	5	7,81%
avverbiali/sub.	5	22,73%	5	31,25%	2	28,57%	2	40,00%
argomentali/sub.	7	31,82%	4	25,00%	2	28,57%	1	20,00%
relative/sub.	10	45,45%	7	43,75%	3	42,86%	2	40,00%
<i>intervista</i>								
clausole indep.	50	51,02%	115	55,02%	87	55,06%	38	55,07%
unità inf. clausola	40	40,82%	68	32,54%	53	33,54%	26	37,68%
clausole sub.	8	8,16%	26	12,44%	18	11,39%	5	7,25%
avverbiali/sub.	8	100,00%	16	61,54%	13	72,22%	4	80,00%
argomentali/sub.	-	-	4	15,38%	1	5,56%	1	20,00%
relative/sub.	-	-	6	23,08%	4	22,22%	0	0,00%
<i>telefonate</i>								
clausole indep.	3,00	52,67%	3,80	52,00%	6,00	66,67%	4,20	40,40%
unità inf. clausola	2,00	35,33%	3,00	40,67%	2,50	29,17%	4,80	46,23%
clausole sub.	0,60	12,00%	0,40	7,33%	0,50	4,17%	1,40	13,37%
avverbiali/sub.	x	-	x	-	-	-	x	-
argomentali/sub.	x	-	x	-	-	-	x	-
relative/sub.	-	-	-	-	x	-	x	-

**Tabella 3.8.** CAF — Complessità — Variazione individuale — Distribuzione per task — Shirley — Tipi di clausole e di subordinate.

	<i>t1</i>		<i>t2</i>		<i>t3</i>		<i>t4</i>	
<i>film</i>								
clausole indep.	75	61,98%	18	52,94%	14	63,64%	90	62,07%
unità inf. clausola	4	3,31%	7	20,59%	4	18,18%	24	16,55%
clausole sub.	42	34,71%	9	26,47%	4	18,18%	31	21,38%
avverbiali/sub.	27	64,29%	5	55,56%	3	75,00%	11	35,48%
argomentali/sub.	2	4,76%	2	22,22%	1	25,00%	9	29,03%
relative/sub.	13	30,95%	2	22,22%	-	-	11	35,48%
<i>storia</i>								
clausole indep.	41	65,08%	30	58,82%	52	63,41%	68	70,83%
unità inf. clausola	6	9,52%	5	9,80%	12	14,63%	12	12,50%
clausole sub.	16	25,40%	16	31,37%	18	21,95%	16	16,67%
avverbiali/sub.	5	31,25%	4	25,00%	4	22,22%	6	37,50%
argomentali/sub.	2	12,50%	2	12,50%	5	27,78%	7	43,75%
relative/sub.	9	56,25%	10	62,50%	9	50,00%	3	18,75%
<i>intervista</i>								
clausole indep.	203	60,24%	104	50,73%	140	52,83%	54	58,06%
unità inf. clausola	78	23,15%	60	29,27%	76	28,68%	28	30,11%
clausole sub.	56	16,62%	41	20,00%	49	18,49%	11	11,83%
avverbiali/sub.	30	53,57%	21	51,22%	18	36,73%	7	63,64%
argomentali/sub.	20	35,71%	12	29,27%	19	38,78%	3	27,27%
relative/sub.	6	10,71%	8	19,51%	12	24,49%	1	9,09%
<i>telefonate</i>								
clausole indep.	3,20	59,17%	3,25	27,92%	4,40	51,99%	7,20	59,51%
unità inf. clausola	1,80	29,17%	5,75	60,61%	2,80	36,19%	3,00	25,62%
clausole sub.	0,60	11,67%	1,25	11,47%	1,00	11,82%	1,80	14,87%
avverbiali/sub.	-		x		x		x	
argomentali/sub.	x		-		x		x	
relative/sub.	x		-		-		-	

A *t1* Pandita e Catherine dimostrano poca variabilità nella complessità sintattica tra i quattro task. Catherine in particolare realizza una variazione con orientamento inverso rispetto alle italiane: il maggior valore nella misura di subordinazione è ottenuto nelle telefonate. Shirley e Eden al contrario dimostrano una maggiore variabilità tra task e già a *t1* ottengono valori molto più vicini a quelli delle compagne italiane. Le differenze più notevoli riguardano anche in questo caso il task interattivo di apertura delle telefonate di servizio, dove

le apprendenti utilizzano AS-unit più complesse rispetto alle italiane. In prospettiva longitudinale le traiettorie di sviluppo delle due coppie di apprendenti sono piuttosto diverse.

Pandita e Catherine, le meno avanzate a T1, dimostrano progressi significativi: c'è più variabilità tra task e la direzione di tale variazione è in sintonia con le produzioni delle studentesse italiane. Tale sviluppo è ottenuto attraverso una crescita della misura di subordinazione e una più varia distribuzione di tipi di subordinate nei task monologici, e una riduzione sia quantitativa che qualitativa nel task interattivo di telefonate. L'avvicinamento ai valori delle studentesse italiane avviene gradualmente, sebbene non sempre segua traiettorie lineari. Nei task monologici la complessificazione in una prima fase riguarda la clausola, con la produzione di un maggior numero medio di parole, in una seconda fase aumenta l'incidenza delle subordinate e la distribuzione per tipi è più simile alle modalità native. Nei task interattivi invece si assiste prima all'introduzione di più turni e poi alla distribuzione delle unità informative su diverse AS-unit semplici, con una conseguente riduzione della misura di subordinazione e della varietà dei tipi di subordinate impiegate.

Le traiettorie di sviluppo di Eden e Shirley, relativamente più avanzate già a T1, hanno un andamento diverso: per entrambe si registra nel tempo una graduale riduzione della misura di subordinazione in tutti i task, con una conseguente diminuzione della variazione tra task. Nel caso di Eden questo è controbilanciato nei task monologici da un uso più vario e simile a quello nativo dei diversi tipi di subordinate, oltre che da una complessificazione a livello di clausola, ottenuta grazie ad un aumento della misura di lunghezza, per Shirley invece è legato a scelte individuali rispetto alla strategia narrativa, basata più sull'intonazione calzante ed espressiva che sulla complessità. Per entrambe la riduzione per i task inte-

rattivi rileva invece diversi progressi a livello interazionale, in particolare nella gestione dell'apertura delle telefonate. Tra t1 e t4 Eden e Shirley realizzano scambi comunicativi man mano più rapidi e contingenti, diluendo la mossa di richiesta su più turni e distribuendo le informazioni su diverse AS-unit. In questo caso a t4 le apprendenti risultano un po' più complesse delle studentesse italiane, questo però non è da interpretarsi come un regresso, ma così come per le italiane a t2, l'analisi qualitativa mostra come esse utilizzino in realtà una maggior varietà di formule di richiesta o pre-richiesta contenenti in alcuni casi esempi di subordinazione.

Ricapitolando, i dati appena discussi confermano come la complessità sintattica possa aumentare in modo non lineare come conseguenza dell'interazione con livello, con task, tra misure e infine in relazione alla variazione individuale. Vediamo ora più in dettaglio il percorso di ciascuna apprendente attraverso alcuni esempi dal corpus. I paragrafi che seguono riportano, apprendente per apprendente, esempi delle produzioni realizzate nei diversi task. Si presentano prima i percorsi di Pandita e Catherine, poi quelli di Eden e Shirley.

A t1 la produzione di Pandita si discosta sotto diversi aspetti da quelle delle compagne italiane, con differenze sostanziali tra task monologici e interattivi.

Nei task monologici, Pandita realizza una narrazione molto lunga, ma caratterizzata da un andamento frammentario, per dettagli giustapposti. L'esempio (49) prodotto a t1 nel task film illustra come Pandita ricostruisce la stessa sequenza di eventi presa in considerazione per le compagne italiane.

- (49) PAN: || | e loro danno lavoro. | ||  
 INV: mh mh  
 PAN: || | {h} c'è un.: # uomo | che dice | di prendere una:  
 # legna {n} come {co;} co:no #o\_6 | ||

- INV: mh mh
- PAN: || | fatta come un cono | ||
- INV: .h come un cuneo
- PAN: || | sì | || #o\_5
- PAN: || | e li {pre-} da {LunL LhL}
- INV: .h
- PAN: °un {bisto°} #I\_o {°pis no. pistone}
- PAN: || | no.° | || #o\_5
- PAN: || | {h} c'è {una-} # hh
- PAN: || | n[on so | ||
- PAN: || | {u} una legna #o\_4 lunga | ||
- INV: [dimmi cosa vuoi dire # o\_6 ah ah # un bastone
- PAN: un bastone | ||
- INV: .h oka[y
- PAN: || | [e >ho detto< | ||
- PAN: || | no. | ||
- PAN: || | {lu} # l'uomo {dic} dice # | ||
- PAN: || | no. | ||
- PAN: || | non ho detto questo. | || #
- PAN: || | ho detto come un ↑cono? | ||
- INV: ah ah
- PAN: || | {eh.} # in va | {dove ci s} dove {c'era} # fanno il nave | ||
- INV: mh mh
- PAN: || | c'era u:n bastone # o\_5 | ||
- PAN: | || che {c'è} >sotto< c'era una .h cono | #o\_2 che ha fermato nave? | ||
- INV: Lh[hL
- PAN: || | [L a prende da sotto?L | ||
- INV: LhL
- PAN: || | {LhL L} la nave lo parteL | ||
- INV: LnoL ho LhL hL .H
- PAN: || | [(va via) | ||

- PAN: || | >ha detto< | ||  
 PAN: || | {o-} vado in {>piro<} prigionie £adesso£ | ||  
 INV: .h  
 PAN: || | l[ui £pensa£ | ||  
 INV: [£hhh£ ÷h  
 PAN: || | {poi} #I\_o poi c'era una ragazza? | # che aveva fa:me  
 | ||#o\_7  
 INV: e quindi cosa fa, se ne va?  
 PAN: || | sì. | ||  
 PAN: || | se ne va.  
 INV: a  
 PAN: da là. # {cioè.} | || #o\_7  
 PAN: || | poi c'era un'altra {u} una ragazza | che aveva fame  
 #I\_o | ||  
 PAN: || | loro due contrano | ||  
 INV: e dove si trovano?  
 PAN: || | in una: città | ||  
 INV: ah ah  
 PAN: || | { in: città #o\_3 vicino a un negozio #o\_2 del pane}  
 no-. | ||  
 PAN: || | la ragazza | che aveva {fa fame} | .h ruba un pane #  
 da panetteria cioè £h£ | ||  
 INV: £h[h£  
 PAN: || | [e scappa | ||

La sequenza è lunga il doppio di quella delle compagne italiane, ma con un uso ridotto della subordinazione e AS-unit più brevi. A questa diversa distribuzione quantitativa corrisponde anche una diversa distribuzione qualitativa dei tipi di subordinate: diversamente dalle compagne italiane Pandita utilizza prevalentemente relative (47,06%) o argomentali (35,29%) e in modo assai più limitato le subordinate avverbiali (17,65%). La maggior parte delle relative impie-

gate è introdotta dal presentativo *c'è*. L'effetto di frammentarietà è poi amplificato dalla quasi totale assenza di segnali discorsivi.

Tra  $t_1$  e  $t_4$  Pandita progredisce mostrando un aumento complessivo delle misure di complessità, sia rispetto alla lunghezza che alla subordinazione. Tra  $t_1$  e  $t_3$  il progresso riguarda sia la lunghezza, con un incremento significativo del numero medio di parole per clausola, sia la misura di subordinazione. Tra  $t_3$  e  $t_4$  non ci sono incrementi in lunghezza e la misura di subordinazione decresce nel task storia. Le misure relative alla distribuzione del tipo di clausole e di subordinate mostrano come il cambiamento non sia solo di tipo quantitativo, ma soprattutto qualitativo, con una riduzione dell'uso di relative e una redistribuzione delle clausole subordinate tra avverbiali e argomentali, con un avvicinamento al comportamento nativo. A  $t_4$  pertanto la narrazione è ricca di dettagli, più fluida e con un andamento meno frammentario, l'apprendente produce clausole più lunghe e AS-unit più complesse. Tale complessificazione quantitativa è supportata dall'uso di diversi tipi di subordinazione. Se a  $t_1$  la forma più rilevante di complessificazione sintattica era rappresentata dall'uso del *c'è* presentativo più relativa, a  $t_4$  Pandita utilizza una maggior varietà di tipi di subordinate, avvicinandosi alle scelte linguistiche delle compagne italiane, così come esemplificato in (50) realizzato a  $t_4$  nel task film.

- (50) PAN: || | quindi il poliziotto ha scritto una lettera | # che è una persona (più) molto responsabile molto #o\_8 fidabile | || #  
 PAN: || | quindi:: # così trovi lavoro | ||  
 INV: mh mh #o\_2  
 PAN: || | e poi prende quella lettera | ||



- PAN: || | e va cercare un lavoro | || #1\_5
- PAN: || | e va a cercare un lavoro {dove fanno} # praticamente | dove fanno #0\_5 le {barche:::} le navi:: | ||
- INV: si. #0\_2
- PAN: || | .hh e lo trovano | || #
- PAN: || | e lo danno lavoro | || #0\_2
- INV: mh mh #1\_5
- PAN: || | il quale ha#nno detto | di iniziare subito | ||
- PAN: || | {che lui} che lui però non voleva uscire da prigione | || #
- PAN: || | che lavoro devo farle? | || #
- PAN: || | dice | ||
- PAN: || | devi cercare {un} #0\_5 un legno come un triangolo | ||
- INV: mh mh #0\_2
- PAN: || | e lui comincia a cercare | ||
- PAN: || | ma non lo trova | ||
- INV: mh::: #0\_2
- PAN: || | no | || #
- PAN: || | vai a cercare da quella parte | || #
- PAN: || | da un'altra parte va cercare | || #
- PAN: || | lo trova {quel li} # quel legno | che è fatto di triangolo però #0\_2 | che viene usato per sostegno della nave | ||
- INV: ah #0\_5
- PAN: || | se lo toglieva | andava via la nave | ||
- PAN: || | e andava nell'acqua! | ||
- INV: ah ah #0\_2
- PAN: || | è quello | ||
- INV: ah ah #0\_2
- PAN: || | ha cercato di toglierlo | ||
- PAN: || | e e ci riesce # a tirarlo | ||
- INV: ah ah #0\_2

- PAN: || | e poi non guarda dietro | ||  
 PAN: || | e porta quello | ||  
 INV: ah ah #o\_2  
 PAN: || | però quando lo guarda | la nave va giù | ||  
 INV: ah ah #o\_2  
 PAN: || | e::: #o\_6 e va nell'acqua # sotto l'acqua | ||  
 PAN: || | e tutti guardano così | ||  
 INV: haha #o\_2  
 PAN: || | {e lui} # lui scappa da là prima | che lo mettevano  
 di nuovo in prigione | || #  
 PAN: || | perché lui voleva andare di nuovo in prigione | || #  
 PAN: || | però #o\_5 scappa da là | ||

Nel task intervista a T1 la produzione di Pandita si discosta sia qualitativamente che quantitativamente dalla produzione delle studentesse italiane, in particolare rispetto alla misura di lunghezza. La studentessa riporta nei dettagli le sue esperienze, parlando sicuramente più delle compagne italiane, l'effetto generale della sua produzione è però, come per i task monologici, di estrema frammentarietà. Le AS-unit sono brevi e semplici, giustapposte tra loro, senza nessuna esplicitazione linguistica dei legami tra le diverse unità informative, così come illustrato dall'esempio (51).

- (51) INV: e cosa avete fatto il primo giorno di scuola. raccontami il  
 primo giorno di scuola.# in italia.  
 PAN: || |{mh.}#o\_7 {da} inisio da casa. hh | ||  
 INV: s[ì.]  
 PAN: || | [de]vo prendere auto. | || #  
 PAN: || | e la non conoscevo. | quale devo prendere. | || #  
 PAN: || | f\*ho detto\*f? | ||  
 PAN: || | bomporto? | ||  
 PAN: || | f\*ho detto\*f | ||

- PAN: || | sì. | ||  
 PAN: || | allora son salita | ||  
 PAN: || | .h f\*son venuta\*f qua. |# perché ho fatto questa  
 scuola. | ||  
 INV: sì:.#  
 PAN: || | e son entrata. | ||#  
 PAN: || | io avevo solo foglio scritto.#o\_6 terza effe. | ||#  
 PAN: || | f\*ho fatto vedere.\*f | || #  
 PAN: || | dov'è ℒterza?[effe.ℒ] | ||  
 INV: [ℒhhℒ]  
 PAN: || | [mi hanno detto | ||  
 PAN: || | su. su. su. | ||  
 PAN: || | sono andata avanti. | ||  
 PAN: || | sonio salita. | ||#o\_2

Significativi i progressi in prospettiva longitudinale. La misura di lunghezza registra un aumento lineare della lunghezza media delle clausole: a t<sub>4</sub> le unità sono lunghe il doppio rispetto a t<sub>1</sub>. Meno lineare lo sviluppo della traiettoria registrata dalla misura di subordinazione: tra t<sub>1</sub> e t<sub>3</sub> si osserva un aumento importante dell'uso di subordinate, qui Pandita registra valori superiori addirittura a quelli delle compagne, a t<sub>4</sub> si osserva però una riduzione, con un allineamento sui valori ottenuti dalle italiane. A t<sub>4</sub> se da un lato continua la complessificazione sintattica a livello di clausola, si riduce l'incidenza di subordinate. Come esemplificato in (52), tale riduzione è compensata dall'uso di strutture pseudo-coordinanti introdotte da *che* o *perché* che le consentono di dare un effetto di dipendenza semantica tra gli enunciati. La conversazione pertanto risulta coesa e fluente, senza quell'effetto di frammentarietà che caratterizzava t<sub>1</sub>. Tale riduzione in complessità la avvicina quantitativamente e qualitativamente alle scelte delle italiane.

- (52) PAN: || | studio secondo me fino adesso abbiamo...cioè | ||  
 PAN: || | perché stiamo facendo dei ripassi per me | ||  
 PAN: || | perché ho fatto chimica da 5 anni | ||  
 PAN: || | sono le stesse cose | ||  
 PAN: || | stiamo facendo dei ripassi | ||  
 PAN: || | magari solo la matematica e fisica | ||  
 PAN: || | che in superiore abbiamo fatto in primi 2 anni di  
 fisica | non sapevo la lingua | ||  
 PAN: || | quindi non # cioè # non ho studiato tanto bene | ||  
 PAN: || | poi ai {superiori devo stud/università} devo studia-  
 re quella | ||  
 PAN: || | ho studiato, | ||  
 PAN: || | ho già passata quella | ||  
 PAN: || | poi è andata tanto bene rispetto a altre materie | ||  
 PAN: || | sono andate più bene della chimica. | ||  
 INV: ah! allora sei proprio brava!

Infine, anche rispetto alle telefonate, a T1 Pandita si differenzia notevolmente dalle compagne. Se sono presenti le mosse fondamentali che caratterizzano l'apertura delle telefonate, i saluti, l'entrata in merito, a volte l'autoidentificazione mediante appartenenza categoriale si realizzano però in uno o due turni composti spesso da una sola AS-unit lunga e complessa. Le pre-richieste che le studentesse native realizzano con una serie di turni o frasi indipendenti, vengono più spesso incorporate in un'unica costruzione insieme alla formulazione della richiesta vera e propria, così come esemplificato in (53). Qui Pandita realizza l'entrata in merito formulando la richiesta in un'unica AS-unit di due clausole piuttosto dense da un punto di vista informativo: introduce subito un argomento specifico, *i costi dell'aereo*, dopo una pausa di #1\_3 completa l'unità sintattica e semantica aggiungendo *su andare a Londra*. La pausa in questo caso è co-costruita con

l'interlocutore, i *costi dell'aereo* è infatti una clausola sintatticamente completa, ma non semanticamente, per questo l'operatore rimane in attesa.

- (53) PAN: ((squillo))  
 AG2: vacanza for you. sono paola?  
 PAN: || | .hh buonasera. | || hh .hh m-  
 AG2: pronto?  
 PAN: || | buonasera. | || #o\_8  
 PAN: || | io vorrei sapere #o\_4 i costi dei aereo. |  
 #I\_3 su andare a Londra. | || #I\_6  
 AG2: per partire quando?

In questa, così come nelle altre telefonate realizzate a t1, Pandita utilizza non solo più subordinate delle compagne italiane, ma anche una maggior varietà di tipi.

In prospettiva longitudinale assistiamo invece a diversi progressi. La traiettoria di sviluppo segue un andamento lineare, aumenta un po' la lunghezza media delle clausole e si riduce la misura di subordinazione, con un graduale allineamento ai valori delle studentesse italiane. Nell'esempio (54), realizzato a t2, dopo l'apertura del canale e i saluti, Pandita entra nel merito, utilizza forme di pre-riciesta e realizza enunciati con una minor densità informativa rispetto a t1. A differenza delle parlanti native qui la richiesta vera e propria è ancora specifica e condensa due informazioni molto precise, numero di pagine e prezzo. Il progresso più evidente è dato però dalla distribuzione delle unità su più turni, con la richiesta che non compare più nei primi, ma viene bensì preparata da altre mosse. Cambia anche la distribuzione qualitativa dei tipi di subordinate impiegati, non più principalmente avverbiali come nelle rilevazioni precedenti, ma piuttosto argomentali o relative, con un avvicinamento

anche rispetto a questo parametro alle scelte delle compagnie italiane.

- (54) PAN: ((squillo))  
 SHI: pronto  
 PAN: || | pronto | ||  
 PAN: || | buonasera | ||  
 SHI: buonasera  
 PAN: || | io volevo chiedere # un'informazione su un libro | ||  
 SHI: dica  
 PAN: || | è un libro di:: autore di Danielle Steel | ||  
 SHI: sì::  
 PAN: || | il titolo è il {renz}  
 SHI: °il ranch°  
 PAN: {ranch} # ranch | ||  
 SHI: ah  
 PAN: || | e # volevo sapere::: | quante pagine ci sono | ||  
 PAN: || | e:: quanto costa | ||  
 SHI: no un titolo del genere Danielle Steel non lo conosco

A t3 e t4, Pandita realizza la telefonata mantenendo uno stile sintatticamente semplice, con uno scambio molto più vicino a quello delle parlanti native: le informazioni sono introdotte una per volta attraverso AS-unit semplici, le mosse tipiche dell'apertura distribuite su più turni, così come esemplificato in (55) realizzato a t4.

- (55) PAN: ((squillo))  
 AG2: buonasera agenzia (dila)ventura::?  
 PAN: || | buonas::era | || #o\_8  
 PAN: || | scusi | ||  
 PAN: || | vorrei sapere informazioni::: | || #o\_5  
 PAN: || | perché vorrei fare una vacanza a milano | || #I\_5

AG2: no # non abbiamo mica niente # su una vacanza a milano #o\_5  
 PAN: || sì. | ||  
 AG2: no. #o\_5  
 PAN: || | cioè non conosce degli alberghi:: | || #o\_5  
 AG2: no. # no signora non conosciamo alberghi #o\_3  
 PAN: niente::?  
 AG2: no no no  
 SHI: °allora cosa fate°  
 AG2: facciamo l'estero no # milano  
 PAN: ah::: okey grazie  
 AG2: prego

Vediamo ora il percorso individuale di Catherine. Nei task monologici, a t1 Catherine produce un racconto che per lunghezza totale e valori di complessità si situa entro il quadro di variazione delineato dalle due studentesse italiane. Emergono però alcune differenze da un punto di vista qualitativo, così come illustrato dall'esempio (56) realizzato nel task film a t1. Catherine racconta il video con una certa imprecisione, fornendo piuttosto un elenco di eventi più o meno principali dove i fatti non sono sempre esplicitamente collegati tra loro. L'analisi della distribuzione dei tipi di clausole mostra come le subordinate utilizzate siano soprattutto argomentali e relative, piuttosto che avverbiali.

- (56) CAT: || | un giorno >è andato< in un negossio | ||  
 CAT: || | (ha) incontrato una: {una} donna #3\_o | che {i signore/ # rubato u::n} signore rubato: #o\_4 un panino  
 CAT: || | non lo so | ||  
 CAT: un baguette | || #o\_2  
 CAT: || | e hanno da- {hann} hanno dato la colpa alla signora | || #1\_5

INV: mhmm\ #3\_0

CAT: || | {e::} #1\_6 e:: {quell che ve} quella che vende il panino, | .hh ha chiamato la: polisiotta | per venire a arrestarlo | | #1\_2

In prospettiva longitudinale dimostra anche lei diversi progressi, ma diversamente da Pandita le misure di lunghezza e di subordinazione aumentano con traiettorie non lineari a zig zag, con effetti di trade-off nelle due misure e andamenti opposti nel film e nella storia. A t4 Catherine ottiene i progressi più significativi. Le sue produzioni nei due task sono lunghe il doppio o il triplo delle precedenti, il racconto risulta più coerente, preciso e dettagliato e Catherine sembra svolgere con più agio il task. Quantitativamente aumentano entrambe le misure di complessità: subordinazione e lunghezza. Progressi anche più significativi si ottengono rispetto ad alcuni aspetti qualitativi della produzione. Qui la narrazione è più fluida, precisa e gli eventi vengono ricostruiti in modo più chiaro che a t1. La distribuzione del tipo di subordinate è ora più vicina a quella delle studentesse italiane: aumenta in modo significativo l'uso di subordinate avverbiali, come esemplificato in (57), realizzato nel task film a t4.

(57) CAT: || | {allora lui #2\_5 tch uscendo da lavoro:: oh} mentre lui {f\*andava andava} andava | {c'era una\*f signorina #0\_5} c'era una signorina #0\_5 .hh | che ha visto mh un:: panettiere

INV: mh::

CAT: || | si dice un panettiere | || #0\_2

CAT: {(un panettiere} | che tirava delle pane dal:: sua macchina | {per} per portare:: # mh:: {negossio} [[suo negozio | |

INV: [si sì



- INV: [[negozio mh mh #  
 CAT: || | .hhh e la ragazza avendo fame:: ehm:: | quando  
 ha visto | che uomo è entrato in negozio | poi aveva:: il  
 furgone aperto | .hh lei l'ha preso il panino | ||  
 CAT: || | e ha iniziato a correre | ||  
 INV: mh mh  
 CAT: || | {essendo che lei ℓhhℓ ehm:: #o\_6 cioè} quando  
 {lui} lei ha iniziato a correre | è entrato {una signori-}  
 una signora | || #o\_5  
 CAT: || | {l'ha vist} l'ha vista | che:: la ragazza ha preso il  
 pane | per scappare lì | ||  
 CAT: || | {l per:: mh::} #o\_5 ha preso il pane: | ||  
 CAT: || | e:: #o\_5 sta scappando | ||  
 INV: mh mh  
 CAT: || | .hh appe- allora appena è arrivato:: uomo # | pa-  
 nettiera l'ha detto | ||  
 CAT: || | guarda | che c'è ragazza lì | che sta correndo | ||  
 CAT: || | ha preso {panino} # .hhh sua baguette | ||  
 CAT: || | era baguette | || #o\_2  
 CAT: || | ha preso tua baguette | ||

Nel task intervista a T1 Catherine realizza questa attività senza discostarsi in modo significativo dalle scelte delle compagne italiane, così come evidente da (58).

- (58) INV: cosa hai pensato quando hai saputo che saresti venuta in  
 italia #o\_9 a stare con i tuoi #o\_6  
 CAT: || | cioè io ero contenta | perché:: #o\_2 finalmente  
 adesso vivo con miei genitori #o\_8 |||  
 CAT: || | perché non sono mai {abitu} cioè abitata con loro?  
 | || #o\_8  
 CAT: || | sono sempre l- qua | || #o\_6  
 CAT: || | vengono in ghana | || #o\_3

- CAT: || | fanno un mese | ||  
 CAT: || | e poi tornano qua | || #o\_6  
 CAT: || | allora #2\_1 ehmm #2\_0 era ehm #o\_3 ero un po'  
 contenta #o\_5 di #o\_7 | ||  
 INV: un po' o molto #o\_3  
 CAT: ||| | °ℒmolto anche°ℒ | || #I\_4

Il suo sviluppo longitudinale si differenzia invece da quello realizzato dalle altre apprendenti. Se infatti la misura di lunghezza non dimostra variazioni sostanziali, quella di subordinazione cresce in modo evidente tra t2 e t4, dove raggiunge valori due volte superiori a quelli delle compagne italiane. A differenza delle compagne, Catherine complessifica gradualmente le sue produzioni, usando più subordinate e riducendo la quantità di unità indipendenti inferiori alla clausola, come esemplificato dai due estratti (59) e (60) realizzati rispettivamente a t3 e t4.

- (59) CAT: || | allora ero molto contenta | di vivere coi miei | perché: là abitavo con la mia zia | || #  
 CAT: || | e:: # ero molto contenta | quando venivo qua | || #  
 CAT: || | almeno se non c'è nessuno | che conoscevo | {av-} i miei # e:h {co- conos-} conoscevano | ||
- (60) CAT: || | ma il vantaggio dalla scuola all'università è | che puoi fare l'esame | quando ti pare | ||  
 CAT: || | vai a lezione quando ti pare | ||  
 CAT: || | però qua sei obbligato a fare l'esame | che ci sono | ||  
 CAT: || | e poi se passi o non passi | voto ti mettono lì | ||  
 CAT: || | e poi qua sei obbligato a venire a scuola sempre | ||  
 CAT: || | se no se fai troppi assenti | ti chiamano a casa i tuoi: | ||

- CAT: || | gli dicono | ||  
 CAT: || | però all'università niente | ||  
 CAT: || | e poi qua non è che devi studiare dalla mattina fino  
 alla sera: | ||  
 CAT: || | ti basta 4 ore | per studiare | ||  
 CAT: || | però all'università ci vuole molto, molto, molto,  
 molto da studiare, molto | ||  
 CAT: || | non devi neanche perdere un secondo | per fare  
 ogni cosa | ||

Rispetto al task telefonate, a T1 invece le produzioni di Catherine differiscono notevolmente da quelle realizzate dalle compagne italiane, in particolare rispetto alla misura di subordinazione: così come Pandita, anche qui tutte le informazioni sono introdotte in un solo pacchetto che risulta così composto da un cospicuo numero di parole e clausole. In (61) possiamo vedere un esempio delle aperture realizzate da Catherine.

- (61) CAT: ((squillo))  
 SH3: gigatel buonasera sono [clau in che cosa  
 CAT: || | [{e;} pronto | || #  
 CAT: || | {buon-} buonasera | ||  
 SH3: buonasera # [mi dica  
 CAT: || | {[vorrei- eh} vorrei avere informazioni | se avete  
 ericson zeta seicento | || #  
 CAT: || | il telefono di ericson | || #2\_2

Qui è interessante notare l'assenza di sincronia nei primissimi turni in cui si realizzano l'apertura del canale e i saluti: Catherine non lascia che l'addetta finisca la sua frase di routine *sono Claudia in che cosa posso esserle utile?*, ma entra a metà con il suo *pronto buonasera*, causando così un piccolo scompenso nel flusso conversazionale. Per entrare nel merito, in

questa così come nelle altre telefonate a t1, Catherine usa la formula *vorrei avere informazioni se avete*, non solo scorretta da un punto di vista grammaticale, ma anche del tutto atipica sul piano pragmatico: abbiamo visto come le studentesse italiane di solito producono *vorrei avere un'informazione* come unità a sé, a cui segue una nuova frase indipendente, frequentemente sotto forma di domanda diretta. Catherine, così come le altre apprendenti, realizza la richiesta con una AS-unit complessa, composta di due clausole e sei parole (considerando una parola unica marca e modello del cellulare).

In prospettiva longitudinale Catherine ottiene diversi progressi. La sua traiettoria di sviluppo non ha un andamento lineare come quello della compagna Pandita, ma procede piuttosto con un'alternanza di progressi e di regressi. A t3 l'analisi quantitativa pare mostrare regresso in quanto assistiamo a un aumento delle misure di complessità, l'analisi qualitativa mostra però come tale regresso sia solo apparente. L'aumento dell'incidenza di clausole subordinate è controbilanciato da un avvicinamento alle modalità native nella scelta del tipo di clausola da utilizzare. Possiamo osservare nell'esempio (62) come Catherine realizzi l'apertura rispettando le modalità native, cioè producendo AS-unit brevi e contingenti con una bassa densità informativa. Le due AS-unit complesse realizzate a t3 sono infatti ben diverse da quelle impiegate a t1. Si tratta in questo caso di due relative che hanno la funzione di specificare l'entità dell'oggetto su cui verterà la richiesta. Abbiamo visto in precedenza come esempi di questo tipo siano presenti anche nel corpus delle italiane. Interessante notare come l'AS-unit complessa sia distribuita su due turni, mantenendo quindi una certa rapidità nello scambio comunicativo.

(62) CAT: ((squillo))

SH3: ricordi mediastore buongiorno sono marco

CAT: || | pronto | ||  
 CAT: || | buongiorno | || # .hhh  
 SH3: prego  
 CAT: || | eh mi scusi | ||  
 CAT: || | mi devo chiedere un un'informazione | ||  
 SH3: prego  
 CAT: || | l'ultimo cd di norah jones |  
 SH3: sì  
 CAT: | che si intitola not too [late | ||  
 SH3: [too late # sì sì ce l'[abbiamo  
 CAT: || | quanto costerà secondo lei | ||  
 SH3: e::: uf dovrebbe essere venti euro e novanta

A t<sub>4</sub> le produzioni di Catherine sono sia qualitativamente che quantitativamente più vicine a quelle delle studentesse italiane. L'apertura è realizzata con AS-unit brevi, la richiesta distribuita su più turni, mitigata da forme di pre-richiasta e formulata in modo generico.

- (63) CAT: ((squillo))  
 EA1: pronto? #o\_5  
 CAT: || | pronto | || #  
 CAT: || | buongiorno | || #I\_5  
 EA1: buongiorno  
 CAT: || | è la [signo- | ||  
 EA1: [chi parla?  
 CAT: || | la catherine | || #I\_8  
 EA1: ah ciao #o\_5 dimmi  
 CAT: || | allora #o\_5 {volevo} # volevo chiederti un'informazione | || #I\_o  
 EA1: sì  
 CAT: || | su un viaggio a verona | || #I\_o  
 EA1: sì. #o\_2 da dove devi partire::?

Vediamo ora le produzioni realizzate da Eden. Nei task monologici, a T1 produce racconti già vicini al quadro di variazione identificato per le studentesse italiane, dimostrando dunque di essere già piuttosto avanzata fin dalla prima raccolta dati. Le sue produzioni sono lunghe e dettagliate, caratterizzate da un uso importante della subordinazione, soprattutto nel film, dove Eden riporta il racconto ricostruendo in modo coerente e lineare gli eventi più importanti. L'estratto in (64) prodotto nel task film a T1 riporta una delle sequenze centrali del video, qui Eden produce diverse AS-unit complesse, utilizzando subordinate di vario tipo, in particolari avverbiali e argomentali. Anche la distribuzione qualitativa del tipo di clausole è vicina a quella delle italiane.

- (64) EDE: || | mentre va in strada così | #o\_2 c'è {una;} #o\_2  
 {una signora}  
 EDE: || | non so | ||  
 EDE: {una;} #o\_3 (u)na ragazza, .hh | che: {va} #o\_8 entra in  
 un negozio | ||  
 EDE: || | prende il pane | ||  
 EDE: || | inizia a pagare | ||  
 EDE: || | però scappa. .h | ||  
 EDE: || | mentre sta scappando | incontra charlie chaplin:  
 #o\_2 sul marciapiede | ||  
 EDE: || | si incontrano | ||  
 EDE: || | cadono insieme. .h | ||  
 EDE: || | allora la signora va- a chiamare: .hh quel signore il  
 proprietario, # | ||  
 EDE: || | però il proprietario #o\_4 {cioè} prende lui | dicen-  
 do | che è stato lui # | ||  
 EDE: || | {f\*cioè} # insomma {la d-} \*f charlie chaplin la di-  
 fende .h | dicendo | che sono stato io | che {ho ruba;} ho  
 rubato questo [pane qua] | ||

EDE: || | dice. .hh | || #o\_2  
 EDE: || | va beh alla fine? #o\_7 la signora | che ha visto: #o\_4  
 questa ragazza #o\_3 | dice | ||  
 EDE: || | è stata lei | ||  
 EDE: || | è stata lei | ||  
 EDE: || | non è stato lui. | || #o\_3  
 EDE: || | va beh | ||  
 EDE: || | alla fine prendono lei, | || #i\_o

In prospettiva longitudinale il progresso è realizzato attraverso una graduale riduzione della misura di subordinazione e un aumento della lunghezza media delle clausole, anche in questo caso con una traiettoria di sviluppo non lineare, con un effetto di trade-off tra le due misure. La riduzione quantitativa della misura di subordinazione è qui bilanciata da un uso qualitativamente più vario di tipi di subordinate che aiutano l'apprendente a strutturare con più precisione gli eventi, così come esemplificato in (65).

- (65) EDE: || | poi dopo # mentre va nella strada così #i\_o | incontra una ragazza | || #o\_8  
 EDE: || | cioè no che incontra:: | || #2\_o  
 EDE: || | nel frattempo fanno vedere una ragazza #i\_2 | che cammina per la strada |  
 INV: si:: #o\_2  
 EDE: che arriva {dal:: #o\_8} davanti al forno:: | che vede:: | || #2\_o  
 EDE: || | insomma ha fame | ||  
 INV: mh mh #o\_5  
 EDE: || | allora ruba questo pane | ||  
 EDE: || | e:: comincia a scappare | ||  
 EDE: || | allora una signora l'ha vista:: | || #o\_8  
 EDE: || | va dal fornaio | ||

- EDE: || | e dice | che quella:: ragazza (niente) ha rubato il pane | ||
- EDE: || | allora #o\_5 si incontrano con il charlie chaplin | ||
- INV: ah ah #o\_2
- EDE: || | è cadono per terra | ||
- INV: la ragazza? #o\_2
- EDE: || | la ragazza | ||
- EDE: || | sì:: | ||
- INV: mh mh
- EDE: || | charlie chaplin {rimane::} #o\_8 gli rimane il pane in mano | ||
- INV: sì::
- EDE: || | allora arrivano i poliziotti | ||
- EDE: || | e £prendono lui£ | || hehehe
- EDE: || | allora la signora cerca {di con- ehm} di convincere il fornaio | dicendo | che è lei | || #o\_8
- EDE: || | allora corrono tutti quanti | || #o\_5
- EDE: || | {dei poliziotti poi #o\_5} un poliziotto:: #o\_2 prende la ragazza | ||
- EDE: || | e lasciano andare lui | || #

Anche nel task intervista, a t1 la produzione di Eden in termini di complessità è equivalente, sia qualitativamente che quantitativamente, a quelle delle compagne italiane. Rispetto alle altre apprendenti però la sua produzione risulta più semplice nella misura di subordinazione. Eden in questo task preferisce utilizzare AS-unit semplici, ma di una certa lunghezza, diluendo così la densità informativa dello scambio comunicativo. In prospettiva longitudinale non ci sono cambiamenti sostanziali tra t1 e t4 rispetto alle misure considerate, così come evidente dagli estratti (66) e (67), realizzati rispettivamente a t1 e a t4.



- (66) INV: e com'è stato il primo giorno di scuola  
 EDE: || | bruttissimo | || [hehehe]  
 INV: [ɕdai raccontameloɕ]  
 EDE: || | bruttissimo | perché: f\*c'era il prof di matema | ||  
 EDE: || | avevamo la prima ora | ||  
 EDE: || | avevamo matematica\*f cioè | || #  
 EDE: || | tutta la giornata è finita {a me} per me e per mia  
 sorella | per spiegare tutte le cose insomma | ||  
 EDE: || | come funziona cioè | || .h  
 EDE: || | almeno per creare un po' di: #I\_o di confidenza, | ||  
 INV: !certo!  
 EDE: || | sì | ||  
 EDE: || | è finita tutta l'ora | ||  
 EDE: || | {a noi} cioè e noi siamo arrivate a casa che [hehehe]  
 | || #  
 INV: [eravate distrutte hh] #  
 EDE: || | sì sì | || #o\_2  
 EDE: || | troppo. | ||
- (67) INV: ho capito # ah #2\_4 va bene # senti #2\_2 e::: dimmi un  
 po' tu adesso lavori quand'è che hai finito la scuola? #  
 EDE: || | l'anno scorso | || #  
 INV: l'anno scorso hai appena finito ah quindi hai trovato su-  
 bito il lavoro  
 EDE: || | no no io::: # lavoravo | || #o\_2  
 EDE: || | e poi venivo a scuola serale | ||  
 INV: ah facevi il serale quindi  
 EDE: || | perché io sono andata::: a fare lo stage là  
 INV: ah ah  
 EDE: e::: a novembre del 2006 | || #  
 INV: sì  
 EDE: || | poi mi hanno chiesto di lavorare con loro? | || #o\_2  
 EDE: || | e sono rimasta | || #o\_2

EDE: || | {e faccio cioè lavoro} # lavoravo di giorno | || #  
 EDE: || | facevo otto ore e poi la sera | || #o\_2  
 INV: che brava  
 EDE: || | facevo il corso | ||

Infine, rispetto alle telefonate, anche per Eden a t1 le produzioni in questo task differiscono da quelle delle compagne italiane: l'entrata in merito, con la pre-richiesta e la successiva richiesta, sono realizzate anche qui attraverso AS-unit complesse dal punto di vista sintattico e dense dal punto di vista informativo, con la richiesta che indirizza un'informazione molto precisa, come esemplificato in (68). Le subordinate impiegate non solo sono quantitativamente più rilevanti, ma anche qualitativamente più varie.

- (68) EDE: ((squillo))  
 AG2: centro viaggi !  
 EDE: || | buonasera | ||  
 EDE: || | volevo chiedere u(n)' informazione | per andare a barcellona . | ||  
 AG2: s:i ! un att(imo) , ma che tipo di informazioni lei voleva ?  
 EDE: || | volevo: chiedere cio(è) | se c' è il tre::no e::h | || +/-  
 AG2: →treno? no # non facciamo treno noi eh?

In prospettiva longitudinale anche Eden realizza diversi progressi. Tra t1 e t3 i valori delle misure di complessità raggiungono livelli identici a quelli delle parlanti native, grazie a un lieve aumento della lunghezza media delle clausole e un'importante riduzione quantitativa e qualitativa dell'uso di subordinate. Tra t3 e t4 sembra invece esservi un'inversione di tendenza delle due misure di complessità che ritornano verso i valori iniziali. Vediamo più in dettaglio come si sviluppano nel tempo le produzioni di Eden utilizzando alcuni esempi dal corpus.

A t2 Eden non ottiene cambiamenti vistosi. Questo è dovuto alle diverse modalità con cui realizza le richieste. In alcuni casi l'apertura della telefonata non sembra essere molto diversa da t1. Dunque, a differenza delle studentesse italiane, non solo condensa più informazioni in un'unica AS-unit complessa contenente la pre-richiesta e la richiesta, ma come nel caso di Catherine, realizza le richieste con strategie formalmente e pragmaticamente inusuali: *volevo chiedere:: informazione per un libro | se c'avete o no*. Si è già precedentemente visto come in questi casi le studentesse italiane preferiscano produrre due AS-unit e soprattutto realizzare la richiesta con un'interrogativa diretta. In altri casi, Eden è in grado di condurre lo scambio di routine di apertura con enunciati rapidi e distaccati, dunque più vicini alle modalità delle studentesse italiane. A t3 non si riscontra più tale variabilità nelle modalità di apertura e come esemplificato in (69) l'apertura viene realizzata con modalità simili a quelle delle native. Lo scambio comunicativo è rapido e contingente, la mossa di richiesta è distribuita su più turni e realizzata con AS-unit semplici. Anche la distribuzione qualitativa dei tipi di clausole è ora più in linea con i valori delle studentesse italiane.

- (69) EDE: ((squillo))  
 SH4: libreria giunti # buongiorno  
 EDE: || | sì | ||  
 EDE: || | buongiorno | ||  
 EDE: || | sono eden | || #  
 EDE: || | vendete cd | || #  
 SH4: e::hm # qualcuno  
 EDE: || | qualcuno | || #  
 EDE: || | allora le dico i titoli:: | || #  
 EDE: || | non lo so | || #  
 EDE: || | elton john | || #  
 SH4: e::: vediamo # xxxx un po'::: xxx guardo un attimo eh

A t4 le misure di complessità ritornano su valori vicini a quelli iniziali. L'aumento della misura di subordinazione è legato, in questo caso, all'uso di 4 AS-unit complesse. Un'osservazione qualitativa dei dati mostra però come il regresso indicato da tale aumento sia solo apparente. L'aumento dell'incidenza di clausole subordinate è infatti legato all'uso di una maggior varietà di formule di richiesta nelle chiamate ai negozi, così come esemplificato in (70), e all'uso di AS-unit complesse nelle pre-richieste delle telefonate indirizzate agli esperti, così come esemplificato in (71) e (72). Così come per le italiane le AS-unit complesse non compaiono più ai primissimi turni e sono anticipate da mosse preparatorie.

- (70) EDE: ((squillo))  
 SH3: libreria (memoli) buonasera sono tatiana #o\_2  
 EDE: || | eh:: buonasera | || #o\_5  
 EDE: || | allora # voi vendete divudi::? | || #o\_5  
 SH3: sì:::  
 EDE: || | sì:: | || #  
 EDE: || | okey | || #  
 EDE: || | volevo sapere | se avete:: # la finestra sul cortile | || #8\_0  
 EDE: || | °l'altro qual'era?° | || #o\_8  
 ST4: °o l'assassinio sull'oriente express° #i6\_0  
 SH3: sì. #o\_5 ce ne sono un po' di copie .....
- (71) EDE: || | io l'ho chiamata | perché dovrei fare un regalo ad una signora {che ha} # di quarant'anni | ||
- (72) EDE: || | allora io avrei bisogno di una mano:: mh # | da regalare un film giallo a una signora | || #o\_6

Vediamo infine il percorso di sviluppo seguito da Shirley. Nei task monologici, a T1 Shirley racconta la storia con un resoconto breve, poco denso di particolari, ma con valori di complessità piuttosto alti sia rispetto alla lunghezza che alla subordinazione. Nella sua produzione è importante l'uso di clausole subordinate: le informazioni tendono ad essere impacchettate in AS-unit lunghe e complesse. Qui i valori dell'analisi quantitativa tendono a raggiungere i livelli di Valentina, la studentessa italiana che parla di più.

Nell'esempio (73), prodotto a T1 Shirley racconta la sequenza del video considerata per le italiane:

- (73) SHI: || | {h # gli: hanno::: preso. # un lav-} a lavoro il primo  
giorno gli hanno detto | di cercare un pezzo di: #o\_2  
legno. | ||
- SHI: || | .h stava cercando cercando cercando | ||
- SHI: || | non l'ha trovato. | ||
- SHI: || | ne ha trovati: #I\_o una? | ||
- SHI: || | che:: .h {e::} appena arrivato lì | fa, | ||
- SHI: {uh. #o\_5} || | caspita. | ||
- SHI: || | guardalo qui. | || #o\_6
- SHI: || | invece? # non era solo: #I\_o un legno. | ||
- SHI: || | ma serviva? | .h a tener ferma la::: b-barca | per non  
partir {da:: .h} #o\_6 da solo. | ||
- INV: mhm
- SHI: || | appena: # l'ha tolto? | parte la: #o\_2 barca. | ||
- SHI: || | e::: # da lì l'hanno licenziato. | ||

Rispetto alla distribuzione qualitativa dei tipi di subordinate, nella sua produzione è significativamente superiore la presenza di avverbiali (64,29%) e relative (30,95%), a scapito delle argomentali (4,76%).

In prospettiva longitudinale, i resoconti del video prodotti a t2 e t3 risultano essere tre volte più brevi e meno dettagliati del racconto prodotto a t1. A t4 la narrazione torna a essere un po' più precisa, ma mantiene sempre un maggiore grado di semplicità rispetto a t1. In modo più evidente a t2 e t3, Shirley si limita a riportare sommariamente la trama, senza scendere negli dettagli delle azioni, cercando piuttosto di portare a termine il compito il prima possibile e avvicinandosi così alle modalità di narrazione tipiche di Elisa, la studentessa italiana che produceva il resoconto più semplice. La scelta stilistica di una narrazione snella ed essenziale si riflette nella progressiva riduzione delle misure di complessità, in particolare di quella relativa alla subordinazione. La misura di lunghezza non dimostra cambiamenti importanti, dopo una lieve riduzione a t2, un aumento a t3, a t4 Shirley ritorna agli stessi valori ottenuti a t1. Più evidenti invece i cambiamenti legati alla misura di subordinazione: tra t1 e t3 si assiste a una riduzione evidente dell'uso di subordinate, seguito poi da una nuova lieve crescita a t4. All'ultima rilevazione il racconto risulta essere realizzato favorendo l'uso di AS-unit semplici, la coerenza nella narrazione è ottenuta attraverso la giustapposizione di clausole indipendenti e frammenti sostenuti da un'intonazione calzante ed espressiva. I cambiamenti ottenuti da Shirley, già avanzata a t1, sono dunque legati non tanto ad uno sviluppo della complessità sintattica intesa come acquisizione di forme di subordinazione più complessa o unità di produzioni più lunghe, ma sembrano piuttosto dovuti alla sperimentazione di diversi stili discorsivi.

- (74) SHI: || | e subito # gli:: #o\_8 prendono | || #  
 SHI: || | e:: #o\_5 comincia subito | || #o\_5  
 SHI: || | solo che gli:: # dicono | di andare a prendere un::  
 tipo di:: # {le} legno | ||

- INV: mh mh
- SHI: || | .hh che lui comunque | ||
- SHI: || | sì [ʎokeyʎ | || #
- INV: [che lavoro deve fare? #o\_5 qual è 'sto lavoro?
- SHI: || | eh che:: # mh:: # sono dei- de:: #I\_2 costruttori  
di:: navi # tipo | ||
- INV: ah
- SHI: || | allora lui:: dovrà eh:: ehm:: ra::ccogliere pezzi di::  
legno mh | {per::} per poterla fare | || #
- SHI: || | gli hanno detto | di prenderne un tipo #o\_8 specifi-  
co | || #
- SHI: || | va | ||
- SHI: || | prende questo | ||
- SHI: || | non è quello | || #
- SHI: || | va | ||
- SHI: || | prende questo | ||
- SHI: || | è l'unico | che ha #o\_8 trovato | || # .hh
- SHI: || | era l'unico:: | che:: teneva la:: #o\_5 nave #I\_o fer-  
ma | || #o\_5
- SHI: || | ah. lui. visto bello | || #I\_o
- SHI: || | toglie | ||
- SHI: || | non esce | ||
- SHI: || | secondo te # se è là | è perché magari | c'è un::  
motivo | || #o\_5
- SHI: || | martello pun pun pun | || #
- SHI: || | toglie {la me:} # la nave | che se ne va per i cavoli  
suoi | ||
- SHI: || | tutti li guardano | ||
- INV: he va nel mare? [hehe
- SHI: || | [sì sì sì | ||
- SHI: || | se ne va per cavoli suoi lui zitto zitto | || #
- SHI: || | giacca | || #
- SHI: || | cappello: | ||

SHI: || | e va #o\_8 via | perché sa | che #o\_2 ormai è licenziato dopo questa | || #2\_2

Il confronto tra gli esempi (75) e (76) realizzati rispettivamente a t1 e t4 mostrano poi come la riduzione dell'incidenza di subordinate sia legata in alcuni casi a una complessificazione a livello di clausola o all'acquisizione di un lessico più specifico. Qui l'uso della parola *panettiere* porta alla riduzione della quantità di subordinate necessarie a t1.

(75) SHI: || | trova: # {il: # e:il: .h #o\_5 } l'uomo | che sta facendo # il pane | che dal furgone va dentro {a a} # al negozio | ||

(76) SHI: || | e c'era giusto giusto il:: # panettiere | che stava ehm:: # scaricando il:: ehm il pane:: # dal camion | || #

La traiettoria di sviluppo registrata da Shirley mostra come ai livelli più avanzati, una volta raggiunto il grado di complessità dei parlanti nativi, la variazione longitudinale può riguardare non tanto un'ulteriore complessificazione, ma piuttosto la sperimentazione di diversi stili o strategie retoriche. Nel caso di Shirley lo stile narrativo si fa più semplice e colloquiale, questo porta alla realizzazione di una minima codifica morfosintattica, con le varie parti del discorso che vengono cucite con riprese anaforiche ad hoc, rimandi all'universo del discorso o un uso espressivo e incalzante dell'intonazione. Le AS-unit impiegate risultano più brevi e semplici o caratterizzate da frasi ellittiche e strutture pseudo-subordinanti, in ragione delle caratteristiche di questo tipo di parlato, in cui, per ottenere un maggior agio ricettivo le informazioni sono distribuite su più frasi e distribuite in modo che in prima posizione compaia l'elemento di interesse.



Nel task intervista, come per Eden, anche la produzione di Shirley a T1 equivale a quella delle studentesse italiane, sia rispetto alle misure di complessità che alla distribuzione di clausole per tipo. In prospettiva longitudinale non ci sono cambiamenti sostanziali tra T1 e T4 rispetto alle due misure di complessità, così come esemplificato rispettivamente in (77) e (78).

- (77) INV: [che cosa è successo il primo gi[[orno.  
 SHI: || | [[.h all'inizio sono andata | || h.  
 INV: e cosa hai pensato quando dovevi andare.  
 SHI: || | h male | || #  
 SHI: || | stavo tremando. | ||  
 INV: !stavi tremando!?  
 SHI: || | sì: con tutti | ||  
 SHI: || | perché #o\_2 eravamo appena venuti | ||  
 SHI: || | una settimana dopo ave abbiamo cominciato. | ||  
 SHI: || | .h perciò vedere tutta sta gente bianca davanti a me:,  
 | || #o\_2 .h  
 SHI: || | poi all'i:nizio sono venuti tutti | perché credevano  
 | che fossimo dall'america. | ||  
 INV: okay  
 SHI: || | visto che {hanno v crede ha} hanno sentito | che  
 parlavo bene l'in#o\_2glese | sono venuti tutti | ||.h  
 SHI: || | appena hanno sentito dall'africa | tutti si son[o riti-  
 rati indietro | ||  
 SHI: || | ho detto | ||  
 SHI: || | guardate | che a me proprio no. .h non me ne frega  
 niente | ||  
 SHI: || | potete starci | ||  
 SHI: || | o non starci perciò .h | ||  
 SHI: || | io infatti in seconda: ho fatto le mie cose | ||

- (78) SHI: || | abbiamo cominciato le superiori | || #  
 SHI: || | tanto per non staccarci #o\_2 | lei è andata a fare i i i  
 i servizi sociali  
 INV: mh mh  
 SHI: e io al cattaneo | || #o\_2  
 SHI: || | e solo che sono due scuole {e e e in in} in un'unica  
 # r\_4 struttura | ||  
 INV: mh mh  
 SHI: || | perciò:: stavamo:: #o\_6 insieme | || #o\_8  
 SHI: || | ha fatto:: tre anni | || #o\_4  
 SHI: || | e solo che:: è stata:: bocciata #r\_8 | perché poi  
 {l'ingle} # lei aveva perso già::: voglia | || #  
 SHI: || | aveva già venti anni | || #2\_4  
 SHI: || | sai | ||  
 INV: quindi era un po' grande forse per stare a scuola  
 SHI: || | grande | || #o\_4  
 SHI: || | sì | ||  
 SHI: || | poi:: voleva dare una mano:: anche a mamma e  
 papà | ||  
 SHI: || | perciò:: voleva già smettere per un # lavoro | ||

Infine, rispetto alle telefonate di servizio, anche Shirley a tr si discosta dalle modalità di apertura delle studentesse italiane: le sue interazioni sono brevi e complesse, le informazioni tendono ad essere impacchettate in un'unica unità. Ad esempio, in (79) dopo l'apertura del canale e i saluti, Shirley realizza una AS-unit di 2 clausole che contiene la mossa di autoidentificazione e la pre-richiasta, seguite dalla richiesta vera e propria, anche in questo caso di due informazioni molto precise, *perciò volevamo chiedere il costo dell'aereo e gli # orari*. A differenza di quanto accade per le studentesse italiane qui l'interazione non è fluida, alla densità informativa si accompagnano pause medio lunghe tra turni e tra mosse comunicative.

- (79) SHI: ((squillo))  
 AG3: moito viaggi?  
 SHI: || | {.hhh} #o\_8 buonasera | || #o\_5  
 AG3: buonasera  
 SHI: || | allora noi siamo un # gruppo di quattro # amici |  
 che # vorremmo far un viaggio # a londra | || #o\_4  
 AG3: sì  
 SHI: || | perciò volevamo chiedere il costo [dell'aerico e gli  
 #orari | || #o\_5  
 AG3: ho [capito e: quando vorreste andare a londra

In prospettiva longitudinale Shirley non sembra ottenere cambiamenti sostanziali rispetto alle misure di complessità, rimanendo su valori superiori a quelli delle studentesse italiane. Shirley dimostra però diversi progressi sul piano interazionale, poiché già a t<sub>2</sub> non condensa più l'entrata in merito in un'unica AS-unit come a t<sub>1</sub>, e produce diversi turni e diverse AS-unit. A differenza delle italiane però continua a realizzare una maggiore quantità di subordinate, producendo formule del tipo *allora vorrei sapere::: dei informazioni se avete xxx* (80) o *allora volevo:: sapere | se avete già {un nome di un} un titolo di [# {un} un album* (81), e dunque continuando a discostarsi sul piano pragmatico dalle studentesse italiane, che realizzano invece pre-richieste del tipo *volevo chiederle un'informazione* come unità a sé e realizzando poi la richiesta vera e propria sotto forma di domanda.

- (80) SHI: ((squillo))  
 SH2: (libreria muratori buonasera)  
 SHI: || | buonasera | ||  
 SH2: (buonasera)  
 SHI: || | allora vorrei sapere::: dei informazioni | se avete  
 alcuni # libri | ||  
 SH2: sì

SHI: || | allora Danielle Steel | ||

SH2: sì

SHI: || | il ranch | ||

SH2: vediamo

(81) SHI: ((squillo))

SH3: ricordi mediastore buongiorno sono cristina #

SHI: || | buongiorno | || #

SHI: || | a[llora e:: | ||

SH3: [prego

SHI: || | vorrei sapere {delle:: e} e delle informazioni sul #  
gruppo latino americano a#ventura | || #

SHI: || | allora volevo:: sapere | se avete già {un nome di un}  
un titolo di [# {un} un album | ||

PAN: [°un cd°

L'esempio (82) realizzato a t<sub>4</sub> mostra come dal punto di vista interazionale Shirley segua ormai il pattern nativo. Le informazioni sono introdotte una alla volta per favorire la graduale messa a fuoco del topic da parte dell'interlocutore. Dal punto di vista delle scelte linguistiche però sembra preferire l'uso di AS-unit più complesse. Tale scelta risulta più elaborata, ma comunque naturale e fluente.

(82) SHI: ((squillo))

EPI: pronto #o\_2

SHI: || | pronto | || #

SHI: || | buonasera | || #

SHI: || | parlo con francesco? | || #o\_5

EPI: sì # sono io #o\_5

SHI: || | salve | || #

SHI: || | allora eh sono la eh #o\_5 benedetta | || #

SHI: || | volevo sapere delle informazioni per:: milano | ||

SHI: || | perché devo organizzare:: un #o\_2 viaggio | ||  
 SHI: || | allora volevo sapere | se lei eh mi poteva dare delle  
 informazioni | su:: # cosa {vede- eh} visitare | ||  
 SHI: || | non so | || #  
 SHI: || | {gli al:: .hh} gli alberghi .hh e [quanto costano | ||  
 SHI: || | e::: qualche altro informazione che # [può essere utile  
 | || #o\_8

### 3.1.4. Riassunto

In § 3.1. è stata riportata l'analisi della variazione della complessità sintattica delle produzioni dei soggetti rilevata utilizzando le due misure quantitative di lunghezza e di subordinazione e le sei misure qualitative relative alla distribuzione di tipi di clausole e di subordinate. I dati mostrano variazione longitudinale e situazionale, sia a livello di gruppo che individuale. Vediamo con quali tendenze.

Longitudinalmente, le apprendenti mostrano progressi in termini di lunghezza e di distribuzione di clausole e subordinate per tipo, mentre cumulativamente in tutti i task la misura di subordinazione segna una lieve riduzione. Questo dato è in linea con quanto evidenziato in letteratura (cfr. 1.2.2.2.): a livelli più avanzati il progresso in complessità sintattica riguarda non tanto l'uso di una maggior quantità di subordinate, ma piuttosto di una maggior varietà e di una complessificazione a livello di clausola.

Situazionalmente, le studentesse italiane dimostrano variazione tra task, con quelli monologici che risultano più complessi di quelli interattivi. Anche per le apprendenti c'è variazione, ma alla prima rilevazione con una polarità opposta, le ragazze usano una sintassi tendenzialmente più complessa nei task interattivi. Nel tempo i valori di gruppo mostrano un avvicinamento ai valori delle italiane con la traiettoria di sviluppo di ciascuna

misura che cambia da task a task, evidenziando come il tipo di attività svolta possa costituire un'ulteriore causa dell'andamento non lineare della complessità sintattica. Nei task interattivi il progresso è ottenuto invece per tutte le apprendenti attraverso una riduzione delle misure di complessità e un conseguente allineamento ai valori delle compagne italiane. I dati confermano l'importanza della valutazione di questa dimensione della proficienza in diversi contesti comunicativi.

Anche individualmente c'è variazione. Per le studentesse italiane questo è evidente nei task monologici, dove una risulta più complessa. Per le apprendenti, i dati mostrano come due di loro siano più avanzate già alla prima rilevazione, nel tempo ciascuna segue un proprio andamento e ritmo, con cambiamenti più o meno lineari e più o meno repentini. Nei task monologici le apprendenti più avanzate realizzano progressi soprattutto in termini di lunghezza, le studentesse meno avanzate in termini di sviluppo della subordinazione, con un effetto di trade-off tra le due misure, accompagnato da un avvicinamento qualitativo nelle misure relative alla distribuzione di clausole e di subordinate per tipo. I dati disaggregati per soggetto confermano l'andamento non sempre lineare della complessità nel tempo, mostrando come ciò sia legato, oltre che al livello e al task considerato, anche ad un'interazione di questi fattori con la variazione individuale. Il gruppo definisce il quadro di variazione generale entro il quale poi si sviluppano i percorsi individuali di ciascun apprendente. Il confronto tra gruppo e individuo risulta dunque utile nell'elaborare profili più affidabili.

Infine, rispetto alle misure, i dati confermano come queste non siano sempre indicatori diretti della competenza, nel senso che un loro aumento non implica automaticamente progresso. Infatti i valori vanno sempre considerati come il risultato di un'interazione con la variazione situazionale e individuale. Attraverso la combinazione di diverse misure e il confronto con

il comportamento dei parlanti nativi e con i valori del gruppo è però possibile ottenere profili individuali di proficiency affidabili.

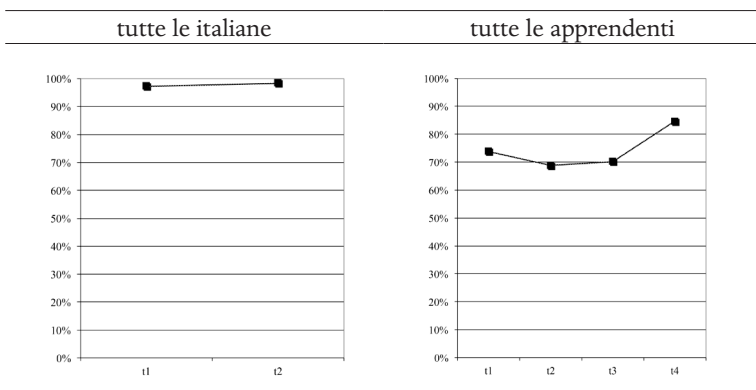
### 3.2. L'accuratezza CAF

Vengono qui presentati i risultati relativi all'analisi dell'accuratezza effettuata calcolando la percentuale di AS-unit senza errori sul totale delle AS-unit prodotte. Si presenterà prima, per i due gruppi di soggetti, l'analisi della variazione longitudinale cumulativamente in tutti i task (§ 3.2.1.), si procederà con quella della variazione situazionale tra task (§ 3.2.2.), e si concluderà illustrando il percorso individuale di ogni studentessa (§ 3.2.3.).

#### 3.2.1. La variazione longitudinale

Il grafico 3.5. mostra i valori medi dell'accuratezza per i due gruppi di soggetti relativamente a tutti i task.

**Grafico 3.5.** CAF — Accuratezza — Variazione longitudinale — Tutti i task — Italiane e apprendenti (valori di gruppo).



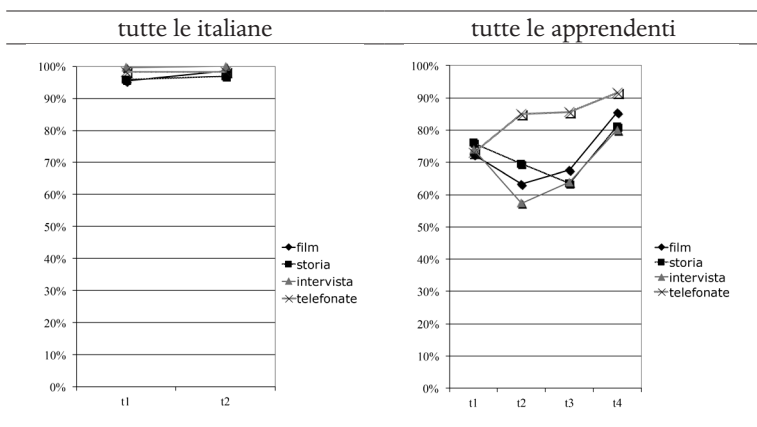
Alla prima rilevazione le apprendenti, come atteso, sono meno accurate delle compagne italiane. In prospettiva longitudinale però esse ottengono un aumento dei valori di accuratezza, anche se la traiettoria di sviluppo non è lineare. Tra t1 e t3 si assiste a una lieve riduzione, seguita poi da un rapido incremento tra t3 e t4. Vedremo in § 3.2.2. a cosa è dovuta questa non linearità.

### 3.2.2. La variazione situazionale

Il grafico 3.6. mostra le medie delle studentesse italiane e delle apprendenti mettendo a confronto i quattro task. I dati disaggregati per task riconfermano solo parzialmente il quadro generale descritto in § 3.2.1.

Per le studentesse italiane la variazione tra task è minima in entrambe le rilevazioni, poiché il livello di accuratezza è sempre vicino al 100%. Solo i task monologici risultano un po' meno accurati di quelli interattivi. Abbiamo già evidenziato in § 3.1.2.1. come il task racconto di un film, richiedendo

**Grafico 3.6.** CAF — Accuratezza — Variazione situazionale — Distribuzione per task — Italiane e apprendenti (valori di gruppo).





ai soggetti di ricordare le sequenze video viste una sola volta e di pianificare l'esposizione, genera alcune incertezze nelle produzioni delle studentesse italiane. Questo dato è in linea con i risultati condotti nell'ambito dei modelli psicolinguistici dell'*information processing view* che mostrano come esista un legame stretto tra accuratezza e richieste del task, con effetti di trade-off (cfr. § 1.2.1.2.).

Per le apprendenti non c'è variazione tra task a t<sub>1</sub>. In prospettiva longitudinale le telefonate si discostano notevolmente dalle altre attività. Mentre in questi ultimi la traiettoria non è lineare, nell'apertura delle telefonate lo è, con un graduale aumento nel tempo. Inoltre a t<sub>4</sub> questo è il task in cui le apprendenti risultano più accurate.

I dati disaggregati per task permettono pertanto di rilevare due ragioni per la non linearità della traiettoria: l'interazione tra task e accuratezza e tra task e sviluppo. Rispetto alla prima ragione, per le apprendenti, come per le studentesse italiane, i due task monologici e, anche se in misura minore, il task intervista richiedono l'uso di strutture più complesse rispetto al task interattivo di apertura delle telefonate, come pronomi clitici, discorso indiretto o alcune forme di subordinazione complessa, non sempre perfettamente acquisite dalle apprendenti. Di conseguenza in questi task si realizza un effetto di trade-off tra richieste del task e accuratezza, con una conseguente traiettoria di sviluppo non lineare. Nel task relativo alle telefonate invece, dove si favorisce l'uso di poche strutture semplici, la traiettoria di sviluppo dell'accuratezza è lineare e la misura impiegata segna un progresso graduale.

Rispetto alla seconda ragione, nei task monologici l'incremento di errori è dovuto all'aumento delle sperimentazioni di forme complesse o alla ristrutturazione di forme già conosciute (così come verrà esemplificato in dettaglio in § 3.2.3.). Questo dato è in linea con i risultati di altre ricerche acquisi-

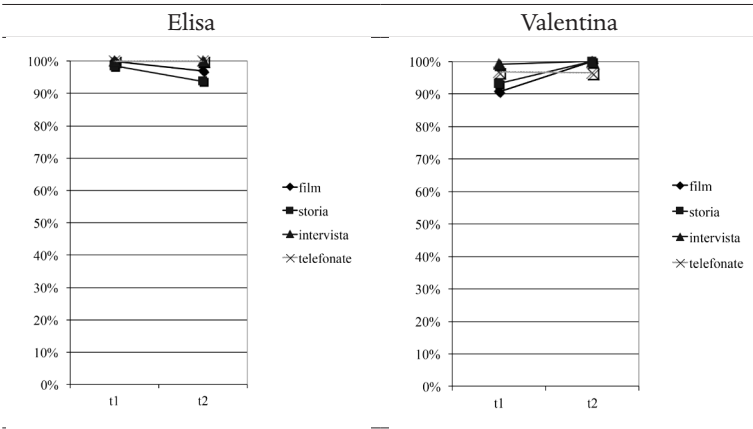
zionali. Già in uno studio del 1983, Larsen–Freeman dimostra come l'accuratezza tenda a diminuire passando dal livello intermedio a quello più avanzato. L'apprendente in una fase iniziale non produce la struttura, e molto spesso ne evita anche i contesti d'uso, mostrandosi dunque poco evoluto ma accurato. In una successiva inizia a sperimentarne l'uso, creandosi più contesti, impiegandola in modo corretto in alcuni casi e errato in altri o generalizzandola a contesti in cui non sarebbe richiesta, risultando pertanto più avanzato da un punto di vista interlinguistico ma meno accurato. Infine la struttura viene impiegata in modo più vicino a quello nativo. In altre parole l'accuratezza è determinata dall'interazione tra elementi quali: grado di accuratezza della rappresentazione stessa, forza di eventuali rappresentazioni in competizione tra loro e grado di automatizzazione, così come suggerito in diversi studi acquisizionali (cfr. § 1.1.2.; § 1.2.2.2.). Questo dato permette di fare un'ulteriore riflessione: accuratezza e sviluppo interlinguistico rappresentano due costrutti diversi, pertanto un apprendente può essere accurato ma poco evoluto, o più evoluto ma poco accurato. Poiché l'accuratezza interagisce con le ristrutturazioni che ricorrono nelle fasi in cui si stanno acquisendo nuove strutture, lo sviluppo in questa dimensione può seguire andamenti a U, con momentanei regressi, seguiti da progressi (cfr. Wolfe–Quintero *et al.*, 1998; § 1.2.2.2.).

### 3.2.3. *La variazione individuale*

Le studentesse italiane (grafico 3.7.) pur riconfermando il quadro delineato dai dati di gruppo dimostrano una leggera variabilità a livello individuale, con Valentina che risulta anche un po' meno accurata della compagna.

Abbiamo visto in § 3.2.2. come questo sia legato alle richieste dei due task: dovendo ricostruire la trama di una sto-

**Grafico 3.7.** CAF — Accuratezza — Variazione individuale — Distribuzione per task — Italiane (valori individuali).



ria le studentesse non sempre riescono a prestare attenzione contemporaneamente alla pianificazione del discorso e al controllo della forma, in alcuni casi pertanto producono alcune imperfezioni (83–85). In altri casi invece le incertezze sono legate all'uso di forme dell'italiano popolare (91). L'analisi qualitativa degli errori mostra come questi siano oltre che ridotti in numero, 5 casi in tutto, limitati ad alcune strutture specifiche. In tre casi si tratta dell'uso o dell'accordo del pronome clitico. Valentina ad esempio utilizza frequentemente il pronome clitico *lo* per riferirsi alla rana protagonista nella storia, accordando quindi il pronome per senso, la rana protagonista è probabilmente di sesso maschile, e non grammaticalmente, la rana è un sostantivo femminile. Elisa in un caso non realizza invece la ripresa, così come esemplificato in (83).

- (83) ELI: || | solo che:: ehm # durante una notte la pantera rosa |  
 che era fuori al freddo | aveva bisogno | di riscaldarsi |  
 ||# .hhh

- ELI: || | e::: {va:::} ruba {un pezzo di legno} #o\_3 insomma  
del fuoco:: dalla casa del cacciatore # | || o\_3.hhh
- ELI: || | il cane corre dietro | per {pre} riprendendosi il ba-  
stoncini {con} con il fuoco? | || #o\_8

Le altre occorrenze riguardano un caso di accordo del participio passato, così come esemplificato in (84) realizzato da Elisa, due casi di uso errato della forma di subordinazione, così come in (85) e (86), realizzati da Valentina.

- (84) VAL: || | hh solo che::: appunto vede questo::: #o\_2 cioè  
Charlie Chaplin | {che: dice # cioè} che: h dice | ||
- VAL: || | {pre} non prendete lei | ||
- VAL: || | ma prendete me | perché son stata io | || #o\_4
- (85) VAL: || | quindi la polizia: | vedendo | che non paga #o\_4 |  
{lo a} lo prende | da portare in galera | ||
- (86) VAL: || | sembra | che si sta facendo il bagno qua ℒh[hℒ .h  
sullo stagno{col cane in te- e} col cane in testa | ||

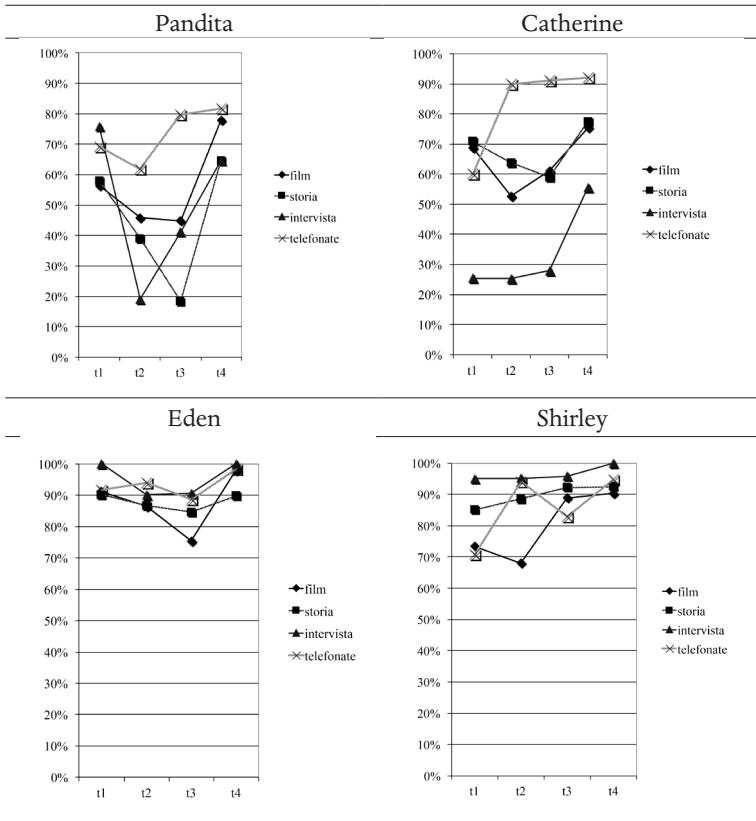
Per le apprendenti i percorsi individuali dimostrano notevole variabilità. Il grafico 3.8. raffina il quadro tracciato con i dati di gruppo.

Pandita e Catherine sono meno accurate delle altre due compagne, con una certa variazione tra task; così come per le compagne italiane, anche nel loro caso il task interattivo di apertura delle telefonate risulta più accurato degli altri tre. In prospettiva longitudinale le due studentesse confermano il quadro di variazione identificato per il gruppo: sono entrambe più accurate in tutti i task, mentre nei due task monologici, e per Pandita anche nel task interattivo intervista, le traiettorie di sviluppo non sono lineari, mentre lo sono nel

task interattivo di apertura telefonate, e nel task intervista per Catherine.

Eden e Shirley, già piuttosto avanzate fin da  $t_1$ , diversamente dalle altre compagne italiane mostrano meno variazione tra task e con tendenze opposte, il task di apertura delle telefonate risulta meno accurato degli altri. In prospettiva longitudinale, Eden segue lo stesso percorso di sviluppo di Pandita e Catherine mentre per Shirley invece l'accuratezza

**Grafico 3.8.** CAF — Accuratezza — Variazione individuale — Distribuzione per task — Apprendenti (valori individuo).



crece in tutti i task. Il livello iniziale può dunque costituire una terza causa di variazione che va ad aggiungersi a quelle identificate in § 3.2.1. e § 3.2.2., cioè lo sviluppo interlinguistico e il task. Vediamo ora in dettaglio attraverso esempi dal corpus.

Per Pandita a T1, così come per le italiane, i task interattivi sono un po' più accurati di quelli monologici e in prospettiva longitudinale ci sono differenze tra task nelle traiettorie di sviluppo. Nei task monologici gli errori commessi da Pandita sono riconducibili ad alcune strutture in particolare, così come esemplificato dagli estratti che seguono, tratti dal task film. Oltre ad alcune incertezze lessicali, legate soprattutto alla scelta delle preposizioni, si ritrovano errori nell'uso dei clitici (87), nella maggior parte delle occorrenze assenti, con qualche sperimentazione di *lo*, dell'articolo, anch'esso assente in buona parte dei contesti obbligatori (88), del passato prossimo, sia per quel che riguarda la scelta dell'ausiliare che l'accordo del participio passato, delle forme riflessive, con l'omissione della particella *si* nella maggior parte dei casi, e infine la scelta delle preposizioni (89).

- (87) PAN: || | rubato  $\mathcal{L}$ io il pane | ||  
 PAN: || | ce l'ho io pa[n $\mathcal{L}$  | ||

- (88) PAN: || | allora lui dà: # suo posta alla ragazza? | ||

- (89) PAN: || | e:: la ragazza #o\_3 cade sopra: h #o\_7 il vuoto .h |  
 che l- era andato a lavorare {in: °n na- # } in nave per la°  
 INV: .h ne nel cantiere  
 PAN: cantiere | ||

In prospettiva longitudinale l'andamento della traiettoria di sviluppo non è lineare, dopo una riduzione del grado di

accuratezza a t2 e a t3, a t4 Pandita raggiunge valori ben superiori a quelli iniziali. Se a t2 e t3 Pandita risulta quantitativamente meno accurata, l'analisi qualitativa del tipo di errori mostra come tale riduzione sia legata alla sperimentazione di nuove strutture o alla ristrutturazione di alcune forme. È il caso, ad esempio, dei verbi riflessivi. A t2 l'uso è corretto in alcuni contesti (90), in altri si assiste a generalizzazioni, come esemplificato in (91) e in (92).

(90) PAN: || |  $\mathcal{L}hh\mathcal{L}$  # il cane si accorge | che:: #o\_8 {e un;} non è una statuo. |||

PAN: || | è un #I\_o un liono.# riale. | ||

(91) PAN: || | e si dorme | ||

(92) PAN: || | {e::} e pink panther.# {e si::} si sente | che arriva padrone. | || #

In modo analogo procede l'uso di AS-unit con dipendenti complesse. Nell'esempio (93) realizzato a t3 notiamo come la sperimentazione di relative complesse sia accompagnato da incertezze grammaticali e produca forme scorrette.

(93) PAN: || | però un certo punto {il il liono} # il pink panther f\*ha sentito la voce::\*f de:l {s;} qualcuno | ha {sp} sparato | ||

Infine, il confronto tra gli esempi (94) e (95), realizzati a t2 e t3 rispettivamente, illustra chiaramente come l'acquisizione di forme più complesse generi un andamento non lineare della misura di accuratezza.

(94) PAN: || | { e poi f\*questa\*f statua era:::? } | invece lui è:#o\_4 come una statuo. | ||

- (95) PAN: || | mette sopra la sua testa | # facendo finta che | se  
come se lui è un albero. | || #o\_6

A t<sub>4</sub> la produzione di Pandita risulta in generale più accurata rispetto alle strutture sopra descritte, come esemplificato in (96)

- (96) PAN: || | prima li ha svegliati {il} #o\_2 il ragazzo | ||  
PAN: || | e poi ha fatto svegliare la ragazza | ||  
PAN: || | {gli ha e} gli ha detto | di scappare:: | ||  
INV: mh mh  
PAN: || | pertanto:: # in quel momento c'era anche il poliziot-  
to | che diceva | di svegliarla | ||

Nel task intervista, Pandita a t<sub>1</sub> risulta essere più accurata che in quelli monologici, con valori vicini all'80%. Le sue interazioni sono realizzate per frasi brevi e semplici, in questa prima fase Pandita è poco evoluta ma sufficientemente accurata. In prospettiva longitudinale, così come per i due task visti sopra, si assiste a una rapida riduzione della misura di accuratezza, con un livello addirittura inferiore al 20% tra t<sub>2</sub> e t<sub>3</sub>, seguita da un ulteriore cambiamento in positivo a t<sub>4</sub>. Nonostante tale riduzione, nel tempo la conversazione risulta man mano più fluida, e almeno per alcuni task fatta di AS-unit più lunghe e complesse. Proprio questo sviluppo qualitativo dell'interazione offre da un lato, più contesti per strutture più complesse, dall'altro maggiori occasioni di commettere errori, come evidente dal confronto tra le produzioni a t<sub>1</sub> (97) e le produzioni a t<sub>2</sub> (98) e a t<sub>3</sub> (99–101). In questa seconda fase, Pandita è quindi più evoluta ma meno accurata. A t<sub>4</sub> si confermano i progressi qualitativi, a cui segue un aumento quantitativo del livello di accuratezza rispetto a t<sub>1</sub>.

- (97) PAN: || | f\*ho detto\*f? | ||  
PAN: || | bonporto? | ||



[...]

PAN: || | allora son salita | ||

PAN: || |.h f\*son venuta\*f qua. | # perché ho fatto questa  
scuola. | ||

INV: sì:. #

PAN: || | e son entrata. | || #

PAN: || | io avevo solo foglio scritto.#o\_6 terza effe. | || #

PAN: || | f\*ho fatto vedere.\*f | # dov'è ℒterza?[effe.ℒ] | ||

INV: [ℒhhℒ]

PAN: || | [mi hanno detto | ||

PAN: || | su. su. su. | ||

PAN: || | sono andata avanti. | ||

PAN: || | son salita. | || #o\_2

PAN: || | c'era prof paola.# di italiano. | che aspettava. | || #

(98) PAN: || | però,#o\_2 secondo me è stato bene.|# perdere un  
anno. | || #

PAN: || | quello terza media,# mi ha aiutato molto. | per im-  
parare lingua. | || #

[...]

PAN: || | {se} magari se. # andavo. # subito a superiore. | era  
molto difficile per me. | || #

INV: perché.#

PAN: || | perché {tes com} la chimica | che faccio adesso | è  
molto difficile. | || #

(99) PAN: || | quindi {non} lei non ha mai eh bisogno di dirmi |  
che stu:dio | ||

(100) PAN: || | mio padre voleva | che mh magari lui faccia una  
scuo:la | così imparasse bene lingua .hh | ||

(101) PAN: || | l'unico posto # in cui # con la quale si può lavorare  
senza anche scuola | ||

Anche il task interattivo di apertura delle telefonate è a t1 più accurato di quelli monologici, ma in prospettiva longitudinale ci sono differenze rispetto alla traiettoria di sviluppo che, contrariamente a quanto avviene negli altri task, qui è lineare.

A t1 le AS-unit di richiesta realizzate da Pandita sono nella maggior parte dei casi poco standard nella forma (102).

(102) PAN: || | {h} vorrei sapere informazioni sul telefono | ||  
#0\_5

In prospettiva longitudinale i cambiamenti seguono una traiettoria lineare. A t2 forme ora più accurate (103) si alternano a sperimentazioni che portano a errori formali (104), a t3 (105-106) e a t4 (107) Pandita evidenzia progressi in due direzioni, da un lato usa forme accurate, dall'altro differenzia le modalità di richiesta.

(103) PAN: || | io volevo chiedere # un'informazione su un libro | ||

(104) PAN: || | perché io volevo comprare un libro per mio amica. ||  
PAN: || | #0\_5 che:: lei piace libri di romanzo. | || #2\_0

(105) PAN: || | posso chiedere {un'infor} un'informazione | ||

(106) PAN: || | io vorrei sapere | se magari lei sa # un titolo di un  
cd | || #

(107) PAN: || | mi chiamo pandita, | ||

EAr: sì::

PAN: || | e vorrei sapere di alcuni divudi | || #0\_8

A t4 Pandita ottiene valori di accuratezza molto alti, nonostante permangano ancora alcune inapproprietezze formali di tipo morfosintattico in due delle formule di richiesta impiegate, come esemplificato in (108) e (109).

(108) PAN: || | vorrei sapere su informazioni su::: # un viaggio a milano | || #I\_2

(109) PAN: || | vorrei sapere delli costi {sul} # di alcuni film | || #I\_5

Catherine mostra un quadro di variazione vicino a quello di Pandita. A t1 è meno accurata nell'intervista, ma più accurata negli altri task. In prospettiva longitudinale anche per lei c'è variazione nelle traiettorie di sviluppo tra task monologici e interattivi, con differenze tra i due anche più significative rispetto alla compagna. Vediamo alcuni esempi dal corpus.

Nei task monologici Catherine non si discosta molto dal percorso di Pandita, nonostante a t1 risulti un po' più accurata. Anche nel suo caso gli errori sono limitati ad alcune strutture: l'uso delle preposizioni, l'accordo del participio passato e l'uso dell'ausiliare nel passato prossimo, così come evidente dall'estratto (110) realizzato nel task film a t1.

(110) CAT: || | con il signore | invece {di sedere su} # di sedere sulla sedia | (si) è {sedu} è andata a sedere {su una donna} # sopra una donna | || .hh

INV ah ah?

CAT: || | quella donna {ha:: uff} ha spinto | || #

[..]

CAT: || | a certo punto seduto su una sedia | ||

CAT: || | è venuto ancora | || #

CAT: || | spinto | || #

CAT: || | quella donna {ha:: uff} ha spinto | || #

In prospettiva longitudinale, anche lei, come Pandita, segue una traiettoria di sviluppo non lineare. Anche in questo caso tra t1 e t3 la riduzione è legata alla sperimentazione di forme più complesse o alla ristrutturazione di forme note, come esemplificato in (I11) e (I12) realizzati a t3 nel task film, dove gli errori sono legati al tentativo di utilizzare subordinate più complesse.

(I11) CAT: || | però::: il rano grande #I\_2 {ha salto} # ha salito su #o\_2 | seguendo mh | per cacciare il piccolo rano | || #o\_2

(I12) CAT: || | mi sembra | che {sta-} cioè col nervoso sta dicendo | di non seguirli::: # l'indicazione | dove vanno | #o\_8 di prendere::: la sua parte::: | || #o\_6

A t4 lo sviluppo qualitativo è accompagnato da un livello di accuratezza globale maggiore rispetto a t1, così come esemplificato in (I13).

(I13) CAT: || | {è andato a cercare un} # è andato a cercare un lavoro da un uomo | || # .hhh

CAT: || | pero lui non lo ha preso | ||

CAT: || | {l'ha} # l'ha scritto una lettera #o\_5 | per andare a presentare l'altra azienda | {se::: # la::: # se::: #} se la prendono | || .hhh

CAT: || | lui andato a fare a presente un'azienda | || #

CAT: || | l'han preso | || #o\_5

Nei task interattivi Catherine segue traiettorie di sviluppo più lineare rispetto alla compagna. Nell'intervista tra t1 e t3 non ci sono variazioni sostanziali nella misura di accuratezza, ma a t4 si ottengono progressi notevoli. Come esemplificato

dagli estratti in (114) e (115), se la misura di accuratezza non registra cambiamenti sostanziali tra  $t_1$  e  $t_3$ , l'analisi qualitativa mostra evoluzioni qualitative dell'interlingua, con la sperimentazione ad esempio di forme più complesse, che porta a imprecisioni formali.

- (114) CAT: || | diciottanni che #o\_3 sono qui | || #o\_4  
 INV: !caspita! #o\_3 e tu #o\_3 quando eri in ghana con chi stavi? #o\_3  
 CAT: || | con mia zia | || #o\_4  
 INV: con tua zia #o\_9 .hhh #o\_5  
 INV: e come è stato rit- #o\_4 venire in italia e stare di nuovo con #o\_3 con i tuoi genitori #2\_1  
 CAT: || | niente | || #o\_8
- (115) CAT: || | allora ero molto contenta | di vivere coi miei | perché: là abitavo con la mia zia | || #  
 CAT: || | e:: # ero molto contenta | quando venivo qua | || #  
 CAT: || | almeno se non c'è nessuno | che conoscevo | {av-} i miei # e:h {co- conos-} conoscevano | ||

A  $t_4$  invece i progressi interlinguistici sono accompagnati da un incremento dei valori della misura di accuratezza, così come esemplificato in (116).

- (116) CAT: || | 18, 19, il mio voto è quello | ||  
 CAT: || | {ma io non ho} all'inizio non ho mai accettato un voto così | ||  
 CAT: || | perché se io inissio a accettare un voto così | poi dopo alla fine, | quando mi laureo | avrò un voto bassissimo | ||  
 INV: certo!  
 CAT: || | allora io # cioè # rifiuto | ||

CAT: || | poi riprovo a farlo | ||

CAT: || | e poi piano piano, | essendo che l'università è primo  
anno | io ci sono alcuni termini che non conosco, allora | ||

CAT: || | se conosco poco | provo a far l'esame | ||

In modo analogo, per il task di apertura delle telefonate, Catherine registra progressi lineari. Se a t<sub>1</sub>, l'apprendente utilizza una sola forma di richiesta caratterizzata da incorrettezze formali (117–118), nelle rilevazioni successive non è solo più accurata, ma usa anche una maggior varietà di formule, così come esemplificato da (119) e (120–121) realizzati rispettivamente a t<sub>2</sub> e t<sub>4</sub>.

(117) CAT: || | e:: vorrei avere informazione sul nek trecento trentotto | || #o\_8

(118) CAT: || | {e;} vorrei avere informazioni | ||

(119) CAT: || | vorrei chieder una cosa #

ST6: °sui libr°

CAT: e::: sui libri | ||

(120) CAT: || | allora #o\_5 {volevo} # volevo chiederti un'informazione | || #1\_o

(121) CAT: || | [vorrei avere alcune informazioni su un viaggio | ||

Vediamo ora i percorsi individuali di Eden e Shirley, le due apprendenti più avanzate alla prima rilevazione. Nel loro caso i progressi sono più lineari, sebbene vi siano alcune differenze a livello individuale. Vediamo con quali tendenze.

Alla prima rilevazione Eden ottiene nei task monologici e nell'intervista livelli di accuratezza molto alti, dimostrandone

do così di essere già piuttosto avanzata. In prospettiva longitudinale le sue traiettorie di sviluppo sono simili a quelle di Pandita e Catherine, sebbene la portata del cambiamento sia quantitativamente meno significativa. A t<sub>1</sub> gli errori di Eden riguardano le preposizioni, l'uso dei clitici e alcune forme di subordinazione, come nel caso dell'uso del discorso indiretto (122) o di relative indirette (123).

(122) EDE: || | { f\*cioè} # insomma {la d-} \*f charlie chaplin la difende .h | dicendo | che sono stato io | che {ho ruba;} ho rubato questo [pane qua] | ||

(123) EDE: || | e gli scrive una lettera | dove: può andare a lavorare | ||

Lo sviluppo longitudinale viene realizzato anche nel suo caso attraverso un'iniziale riduzione dell'accuratezza: a t<sub>2</sub> e t<sub>3</sub> la ristrutturazione di forme complesse provoca maggiori imperfezioni. È il caso dell'accordo del passato prossimo con ripresa del clitico: se a t<sub>1</sub> Eden tende a esplicitare l'oggetto, a t<sub>3</sub> utilizza il pronome clitico, generando però incorrettezze rispetto all'accordo del participio passato, così come nell'esempio riportato in (124), dove il participio dovrebbe essere al femminile, perché riferito a una bomba.

(124) EDE: || | allora il cane invece l'ha portato # | mentre era svenuto il padrone | || #

EDE: || | cioè il cane l'ha appoggiato vicino a lui | || #

A t<sub>4</sub> la produzione di Eden risulta più accurata, ormai ai livelli delle studentesse italiane, così come esemplificato in (125)

(125) EDE: || | quando arriva lì: | gli dà quella lettera lì | ||

EDE: || | e lo accettano | ||

- INV: mh mh #o\_2  
 EDE: || | poi {si::} subito comincia a lavorare | || #  
 EDE: || | gli dice | di trovare una tavoletta | || #  
 EDE: || | {cioè gli fa vedere come::} # gli dice | di cercare  
 un'altra come questa | || #o\_8  
 EDE: || | lui va a prendere::  $\mathcal{L}$ quella hehehehe $\mathcal{L}$  | {dove}  
 dov'è #o\_8 appoggiata la barca no | ||

Anche Shirley è già piuttosto accurata fin da t1 e come per le altre apprendenti gli errori sono legati ad alcune strutture specifiche: l'accordo dei clitici (126), l'accordo del participio passato nel passato prossimo, in particolare in contesti in cui ricorre il clitico (127), la formazione di alcune AS-unit complesse al congiuntivo (128).

- (126) SHI: || | appena::? # l'ha preso. | {c'è} c'era la signora? .h |  
 che le stava fissando::, | ||
- (127) SHI: || | ne ha trovati: #I\_o una? | ||
- (128) SHI: || | ma serviva? | .h a tener ferma la::: b-barca | per non  
 partir {da:: .h} #o\_6 da solo. | ||

Si differenzia però dalle compagne rispetto alla traiettoria di sviluppo longitudinale, nel suo caso lineare anche nei task monologici e nell'intervista. L'analisi permette di osservare come tale aumento sia il risultato, da un lato di una narrazione più breve e semplice, dall'altro dall'uso più accurato di varie strutture linguistiche, come illustrato dall'evoluzione che intercorre tra l'esempio (129) prodotto a t2 e (130) a t3 nel task film.

- (129) SHI: || | che:: lui ha fatto di tutto | {per far m} di sostì .hhh  
 tuire questo cane qua:::, | ||



SHI: || | ha fatto di tutto | che così il pa # drone lo mandasse  
via. | così lui restava in .hhh casa. | ||

[...]

SHI: || | ha fatto finta | che {l'av} l'aveva mangiato il cane. | ||

(130) SHI: || | e::: la pantera? #o\_4 {e:::} ha sempre fatto di tutto.  
| di modo che::: {il ca} il cacciatore s::i::: incavoli con e:::  
il cane. | || #o\_2

Nel task di apertura delle telefonate, sia Eden che Shirley seguono traiettorie di sviluppo abbastanza lineari.

Per Eden non ci sono cambiamenti molto significativi tra  $t_1$  e  $t_4$ , poiché il livello di accuratezza è piuttosto alto fin dalla prima rilevazione. Se a  $t_1$  e  $t_2$  Eden produce diverse occorrenze di formule di richiesta formalmente scorrette, come quella riportata in (131), a  $t_3$  (132) e a  $t_4$  (133) l'apprendente realizza richieste attraverso formule più varie e accurate. A  $t_4$  raggiunge un livello di accuratezza molto vicino al 100%. L'unica AS-unit contenente un errore è esemplificata in (134). In questo caso si tratta di una AS-unit di richiesta piuttosto complessa. L'imprecisione qui riguarda la scelta del connettore che introduce la subordinata implicita.

(131) EDE: || | volevo chiederti informazione sui telefonini. | || #o\_2

(132) EDE: || | allora volevo chiederle alcune informazioni:: {per  
#o\_8 per una::} per un # corso | che stiamo facendo | ||

(133) EDE: || | noi:: s:: volevamo chiedere un'informazione per una  
gita | che dobbiamo organizzare | || #1\_0

(134) EDE: || | allora io avrei bisogno di una mano:: mh # | da  
regalare un film giallo a una signora | || #o\_6

Anche Shirley a t<sub>1</sub> utilizza una sola formula di richiesta scorretta dal punto formale, come esemplificato in (135). La ripetizione di questo stesso errore incide in modo significativo sul livello di accuratezza globale. Tra t<sub>1</sub> e t<sub>2</sub> l'apprendente ottiene una repentina crescita, superando il 90% (136).

(135) SHI: || | allora vorrei sapere {delle} dei: informazioni sul sul telefono nek trecentotrentotto | || #

(136) SHI: || | a:: allora vorrei sapere delle informazioni su un libro | da consigliare ad una mia amica. | || #

A t<sub>3</sub> si assiste a una lieve riduzione: così come per Eden gli errori sono dovuti a tentativi di produrre AS-unit di richiesta più elaborate, così come in (137) e (138).

(137) SHI: || | volevo delle # informazioni e: | perché dovrei re # galare un cidi per una mia amica # di #o\_4 quarant'anni | || #

(138) SHI: || | allora e:: volevo sapere delle informazioni a ri # guardo del #o\_4 gruppo a # ventura | || #

A t<sub>4</sub> si riconferma il generale progresso in termini di accuratezza, così come esemplificato in (139).

(139) SHI: || | allora # eh:: volevo sapere delle informazioni # | perché # eh:: dovrei organizzare un # vi::aggio a: # milano | || # .hh

SHI: || | allora volevo sapere | se lei mi poteva dare {dei::: in in} delle eh informazioni utili | || #o\_8

Ricapitolando, i dati dunque mostrano come il livello di accuratezza e l'andamento della sua traiettoria di sviluppo siano

legati a tre elementi: il tipo di task, le interazioni con lo sviluppo interlinguistico, e il livello iniziale. I task interattivi tendono ad essere più accurati di quelli monologici e soprattutto il task di apertura delle telefonate ha una traiettoria di sviluppo lineare per tutte le apprendenti. I task monologici seguono traiettorie di sviluppo con andamenti che si differenziano sulla base del livello iniziale: Shirley, più accurata a  $t_1$ , segue traiettorie di sviluppo lineare, le altre apprendenti seguono una traiettoria di sviluppo dell'accuratezza con un andamento ad U, effetto dell'interazione tra accuratezza e sviluppo interlinguistico.

La misura di accuratezza scelta si conferma dunque un buon indicatore globale dello sviluppo di questa dimensione, sensibile alla variazione longitudinale e situazionale segnata dal gruppo e dall'individuo. I dati evidenziano anche però come a livello avanzato non sia opportuno fare generalizzazioni sulle competenze in base a un unico risultato, poiché un aumento o una diminuzione non corrisponde automaticamente a un progresso o un regresso della proficiency globale.

#### 3.2.4. *Riassunto*

In questa sezione si è discussa l'analisi della variazione dell'accuratezza misurata attraverso la misura quantitativa percentuale di AS-unit senza errori sul totale delle AS-unit. I dati mostrano variazione longitudinale e situazionale, sia a livello di gruppo che individuale.

Longitudinalmente, le quattro apprendenti tra  $t_1$  e  $t_4$  dimostrano progresso, le traiettorie di sviluppo non sono lineari e hanno un andamento a U, determinato dall'interazione tra accuratezza e sviluppo interlinguistico. Questo dato è in sintonia con quanto osservato in altre ricerche acquisizionali che mostrano come gli apprendenti più avanzati, poiché

sperimentano forme più complesse, tendano a produrre più errori (cfr. Larsen–Freeman, 1983; Skehan, 1998; Wolfe–Quintero *et al.*, 1998).

Situazionalmente, per le italiane c'è variazione tra task, i task interattivi sono più accurati che quelli monologici. Questo dato conferma quanto evidenziato dalla letteratura sui modelli psicolinguistici. Diversi task pongono diverse richieste al parlante: i task monologici richiedono sia di memorizzare una serie di eventi, che pianificarne la narrazione, generando quindi un effetto di trade-off con l'accuratezza, i task interattivi, richiedendo l'uso di formule o strutture poco complesse permettono invece una maggior attenzione alla forma. Le apprendenti dimostrano meno variazione delle compagne a T1, ma nel tempo le traiettorie di sviluppo dell'accuratezza si differenziano sulla base del task, con il task di apertura delle telefonate che segna incrementi lineari e quelli monologici e l'intervista che seguono un andamento a U.

Anche individualmente c'è variazione. Due delle apprendenti sono meno accurate a T1 delle compagne e l'accuratezza nel tempo segue traiettorie più o meno lineari in relazione al livello iniziale. Se nel task apertura delle telefonate tutte le apprendenti ottengono aumenti lineari dell'accuratezza, nei task monologici e nell'intervista si registrano alcune differenze. Per tre delle apprendenti questi task disegnano traiettorie non lineari, mentre per una delle apprendenti, tra le più avanzate del gruppo a T1, i progressi sono sempre lineari.

Infine, rispetto alla misura selezionata, i dati mostrano come l'andamento delle traiettorie di sviluppo sia legata a tre elementi: task considerato, sviluppo interlinguistico e livello iniziale. Pertanto, così come osservato per la complessità, la percentuale di AS-unit corrette sul totale delle AS-unit non deve essere intesa come misura diretta della proficiency linguistica, poiché un aumento o una diminuzione non sempre

corrisponde automaticamente a un miglioramento. Se da un lato quindi è importante valutare l'accuratezza in diversi contesti, dall'altro per avere profili più affidabili sarebbe utile, oltre che osservare l'accuratezza in diversi task, combinare la misura selezionata con indicatori capaci di rilevare lo sviluppo interlinguistico. In questo senso potrebbe essere più adeguato rilevare la presenza di alcune strutture con valore diagnostico rispetto agli stadi dell'interlingua.

### 3.3. La fluenza CAF

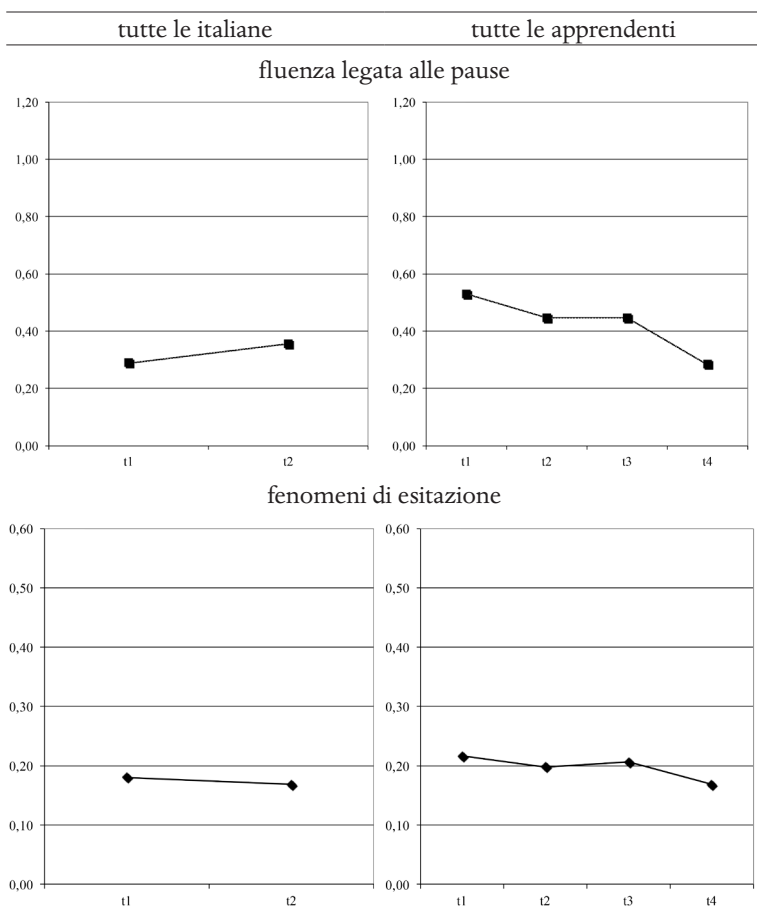
Di seguito si presentano i risultati relativi all'analisi della fluenza effettuata attraverso l'applicazione di due misure analitiche: il numero di pause superiori a #0\_5, e la percentuale dei fenomeni di esitazione, entrambi calcolati sul totale delle AS-unit. Si ricorda che nel caso di queste misure il progresso è realizzato attraverso una riduzione dei valori, poiché viene considerato più fluente il parlante che produce un numero minore di pause lunghe e di fenomeni di esitazione. Si inizierà presentando per i due gruppi l'analisi della variazione longitudinale cumulativamente in tutti i task (§ 3.3.1.), si procederà con quella della variazione situazionale tra task (§ 3.3.2.), e si concluderà illustrando il percorso individuale di ogni studentessa (§ 3.3.3.).

Nell'analisi di questa variabile, nel task intervista sono stati considerati solo i primi cinque minuti di conversazione in ogni rilevazione. Le conversazioni di ciascuna ragazza infatti variano notevolmente per durata, e questo task, al contrario degli altri, non ha delimitazioni ben precise. Per arrivare a un confronto più preciso è stato quindi necessario uniformare il campione, e cinque minuti sono stati considerati un tempo sufficiente per ottenere risultati validi (cfr. per es. Foster, Skehan, 1996).

3.3.1. *La variazione longitudinale*

Il grafico 3.9. mostra la media dei valori ottenuti tra t1 e t4 dall'applicazione delle due misure di fluenza alle produzioni dei due gruppi delle studentesse italiane e delle apprendenti, globalmente in tutti i task.

**Grafico 3.9.** CAF — Fluenza — Variazione longitudinale — Tutti i task — Italiane e apprendenti (valori di gruppo) — Pause ed esitazioni.

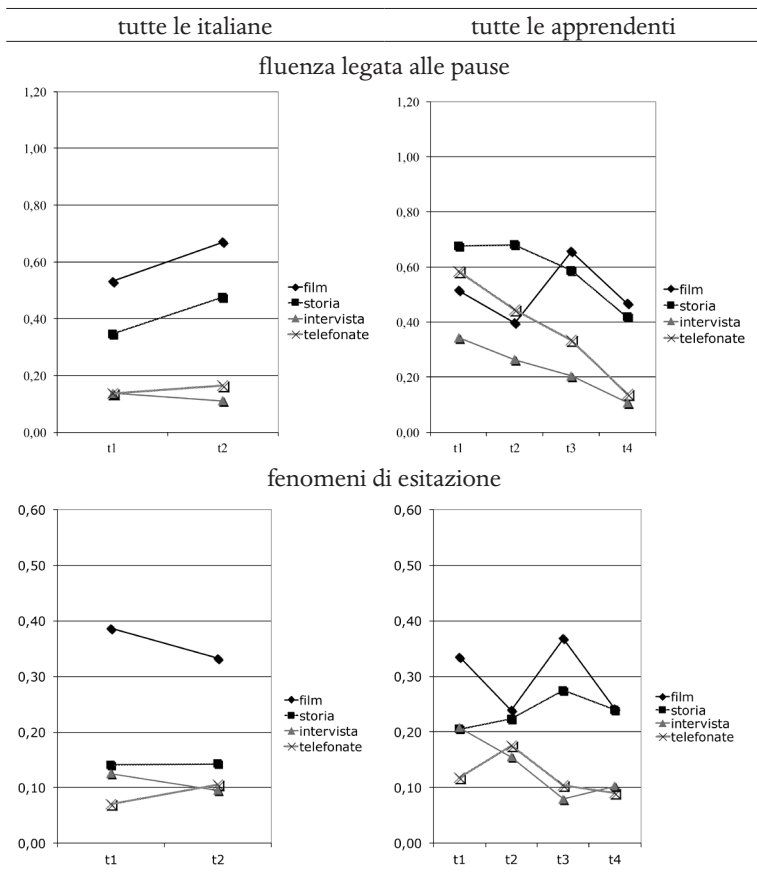


A  $t_1$  le apprendenti sono complessivamente meno fluenti delle italiane rispetto al numero delle pause, ma non per numero di esitazioni. Tra  $t_1$  e  $t_4$  si allineano ai valori delle compagne italiane anche nella prima misura.

### 3.3.2. *La variazione situazionale*

Il grafico 3.10. mostra la media dei valori ottenuti dai due gruppi di soggetti nelle due misure di fluenza disaggregati per task. Questi valori rivelano notevole variazione situazionale, sia tra le italiane sia tra le apprendenti e riguardo sia alle pause sia alle esitazioni.

Per le studentesse italiane i task monologici risultano meno fluenti degli interattivi. Questo dato è in linea con i modelli psicolinguistici (§ 1.2.2.2.), secondo i quali le richieste comunicative dei diversi task favoriscono diverse dimensioni della produzione, con i task interattivi che richiedono la produzione di turni rapidi e brevi, con un uso ridotto di pause lunghe ed esitazioni. Nei dati oggetto dello studio questo è evidente soprattutto per le pause, poiché le italiane producono il maggior numero di pause lunghe nel raccontare il film. La misura dei fenomeni di esitazione rileva un quadro in parte diverso, poiché il racconto di una storia presenta valori simili a quelli dei task interattivi. Anche in questo caso la differenza potrebbe essere legata alle caratteristiche del task stesso, poiché il supporto delle immagini favorisce lo sviluppo di una narrazione spezzettata per didascalie successive. In prospettiva longitudinale entrambe le misure mostrano variazione nei task monologici: tra  $t_1$  e  $t_2$  la misura legata alle pause registra un lieve aumento, mentre quella dei fenomeni di esitazione si riduce, mantenendo così il livello di fluenza globale sostanzialmente invariato. Nei task interattivi non ci sono cambiamenti sostanziali nel tempo.

**Grafico 3.10.** CAF — Fluenza — Variazione situazionale — Distribuzione per task — Italiane e apprendenti (valori di gruppo) — Pause ed esitazioni.

Anche le apprendenti mostrano notevole variazione tra task, con tendenze diverse però rispetto a quelle delle compagne. Nel loro caso, le differenze non sono marcate per quelli monologici, ma sono piuttosto notevoli in quelli interattivi. Il task relativo alle telefonate registra le differenze maggiori rispetto agli altri tre: in questo, in cui i turni di conversazione



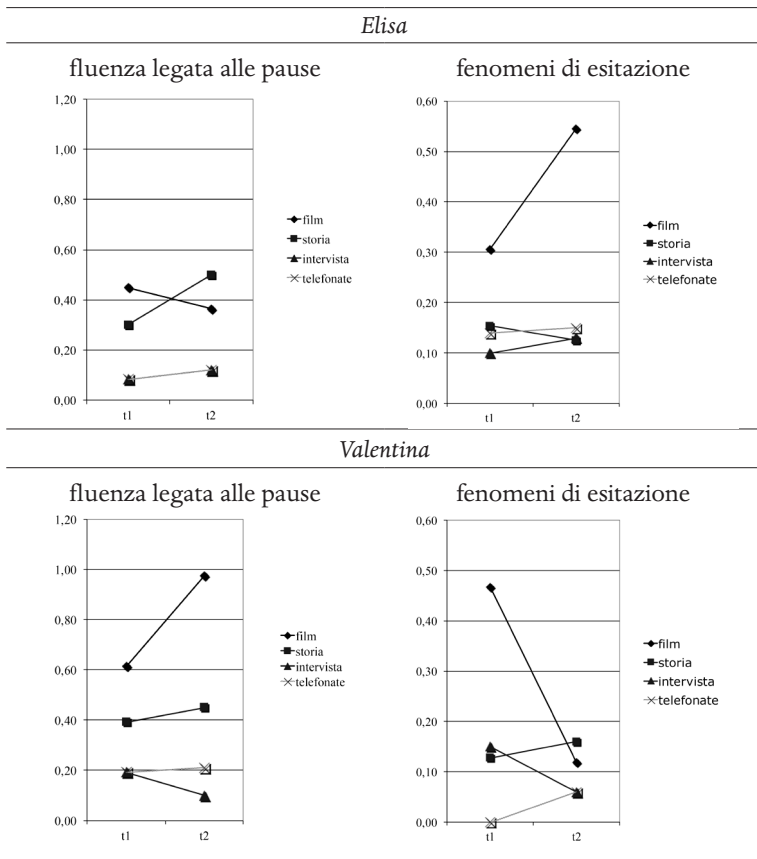
dovrebbero alternarsi in modo rapido e contingente, le apprendenti ottengono a  $t_1$  valori piuttosto alti, con differenze più evidenti nella misura relativa alle pause. Nell'arco dei tre anni di osservazione mostrano progressi, con un graduale avvicinamento ai valori delle compagne italiane in entrambe le misure, e in modo più evidente per i task interattivi, ossia quelli in cui a  $t_1$  mostravano maggiori differenze.

Insomma l'analisi longitudinale e situazionale mostra come l'andamento delle traiettorie di sviluppo delle due misure di fluenza sia legato a due fattori: il task e l'interazione tra le due misure. Nei task interattivi le traiettorie sono piuttosto lineari, mentre in quelli monologici, dove peraltro già da  $t_1$  le apprendenti sono in linea con la variazione nativa, seguono un andamento a zig zag. Nei task monologici, per le apprendenti come per le native, gli aumenti in una misura sono spesso bilanciati da lievi riduzioni nell'altra.

### 3.3.3. *La variazione individuale*

La sezione riporta i risultati dell'analisi della variazione individuale della fluenza. Si presentano, task per task, prima i dati delle due studentesse italiane, poi quelli delle apprendenti.

Il grafico 3.11. riporta i percorsi individuali di ciascuna studentessa italiana. A  $t_1$  si riconferma così il quadro di variazione delineato dai valori di gruppo, e cioè che i due task monologici risultano meno fluenti dei task interattivi nella misura legata alle pause, mentre il task racconto di un film è meno fluente degli altri tre rispetto alle esitazioni. La variazione che riscontriamo tra le due ragazze è dovuta al fatto che Valentina nel raccontare il film a  $t_2$  risulta poco fluente. Abbiamo già evidenziato in § 3.1.3. come questa ragazza abbia diverse difficoltà a svolgere questo task nella seconda rilevazione. Il video le risulta poco chiaro, e pertanto la ricostruzione che ne segue è, oltre

**Grafico 3.11.** CAF — Fluenza — Variazione individuale — Distribuzione per task — Italiane — Pause ed esitazioni.

che imprecisa, estremamente faticosa, e richiede diversi interventi da parte dell'intervistatore. Questo caso mostra come la variazione della performance in un test dipenda anche da fattori esterni alla competenza linguistica o ai fattori di processing.

Il grafico 3.12. riporta i dati relativi alle quattro apprendenti. Anche per loro possiamo osservare una certa variabilità individuale. Alla prima rilevazione, la scarsa fluenza delle apprendenti rispetto alle italiane, già notata in § 3.3.1., non

è uniformemente distribuita tra tutte e quattro, ma dovuta principalmente a Pandita e Catherine. Infatti queste due ragazze sono meno fluenti di Eden e Shirley in tutti i task, in particolare nelle telefonate.

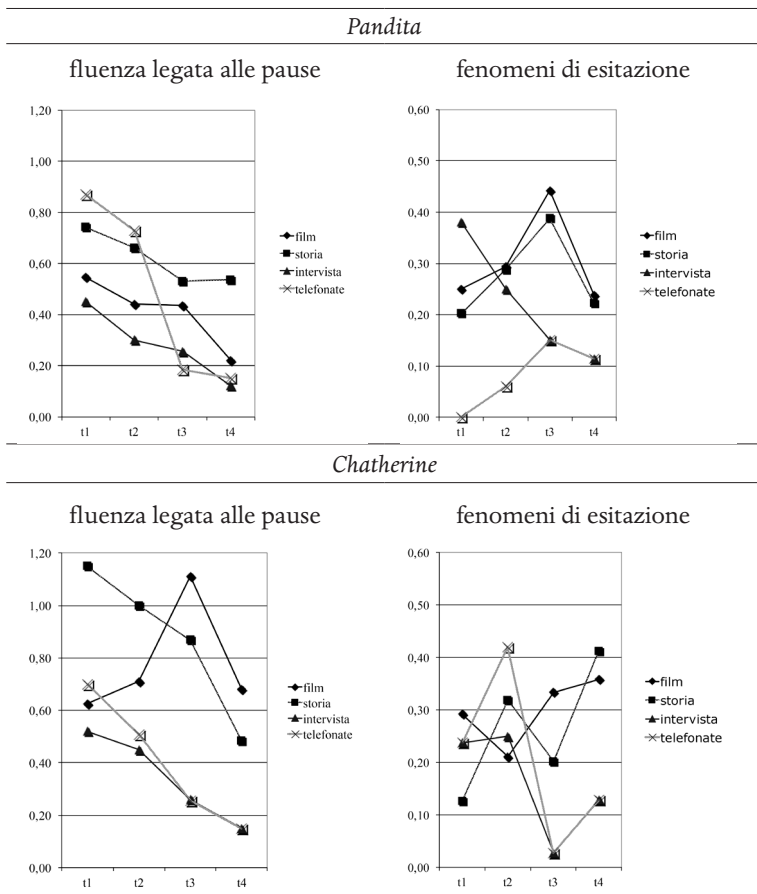
Anche le traiettorie di sviluppo mostrano differenze individuali tra le apprendenti: Pandita e Catherine, meno fluenti a  $t_1$ , ottengono cambiamenti più evidenti e più lineari rispetto a Eden e Shirley. Il livello iniziale costituisce dunque un terzo elemento di variazione, che va ad aggiungersi ai due identificati in § 3.3.2., e cioè il task e le interazioni tra le due misure di fluenza.

Volendo adesso differenziare ulteriormente tra Pandita e Catherine, vediamo come Pandita progredisce linearmente in tutti i task, mentre Catherine è meno lineare nel racconto del film. In questo caso non si registrano sostanziali cambiamenti tra  $t_1$  e  $t_4$ , e la sua traiettoria di sviluppo non è lineare poiché lievi riduzioni sono seguite da aumenti quantitativamente equivalenti. A  $t_4$  Catherine si conferma come la meno fluente del gruppo.

Poiché Eden e Shirley a  $t_1$  sono già entro il quadro di variazione delle italiane, per loro nel tempo i cambiamenti sono meno evidenti di quelli delle altre due compagne e le traiettorie di sviluppo delle misure di fluenza sono poco lineari, senza variazioni sostanziali. Unica eccezione il task relativo alle telefonate di servizio per Eden: se a  $t_1$  si registravano alcune differenze con le italiane, nel tempo questa è l'attività in cui l'apprendente ottiene cambiamenti, allineandosi a  $t_4$  con le tendenze native.

#### 3.3.4. *Riassunto*

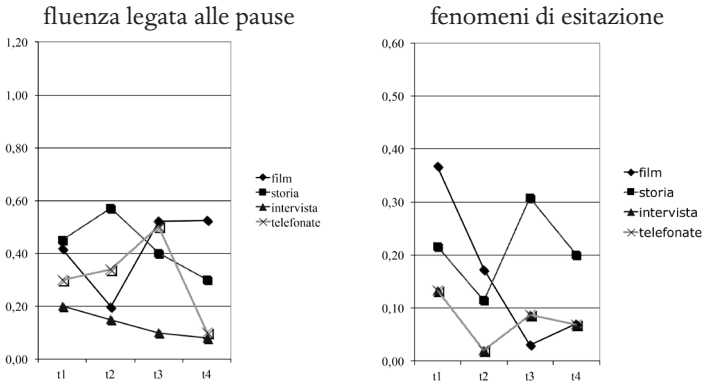
In questa sezione è stata riportata l'analisi della variazione della fluenza delle produzioni dei sei soggetti, rilevata utilizzando le due misure relative alle pause e ai fenomeni di esitazione. I dati mostrano variazione longitudinale e situazionale, sia a livello di gruppo che individuale. Vediamo con quali tendenze.

**Grafico 3.12.** CAF — Fluenza — Variazione individuale — Distribuzione per task — Apprendenti — Pause ed esitazioni.

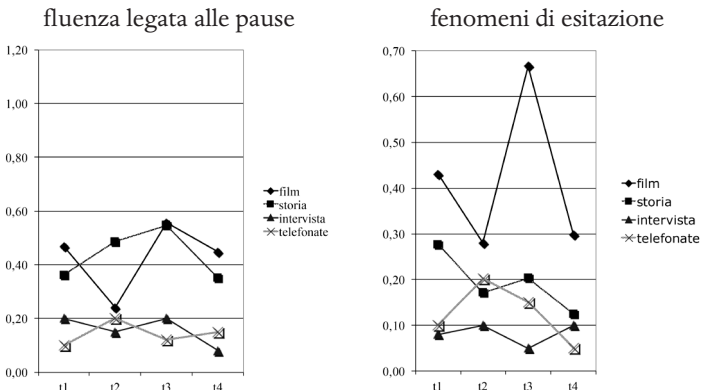
Longitudinalmente, tutte le apprendenti dimostrano cambiamenti in positivo, con un allineamento tra t1 e t4 ai valori delle compagne italiane.

Situazionalmente, le studentesse italiane mostrano differenze tra task, con quelli monologici che risultano un po' meno fluenti degli interattivi. Anche nel caso delle apprendenti c'è variazione, ma con polarità opposta. Se le differenze rispetto

## Eden



## Shirley



alle italiane sono poco notevoli per i task monologici, in quelli interattivi le apprendenti risultano meno fluenti. Nel tempo i progressi più significativi riguardano dunque senza sorprese questi ultimi task, in particolare le telefonate di servizio.

Anche individualmente c'è variazione: Pandita e Catherine, che a t1 si discostano maggiormente dalle compagne italiane rispetto a Eden e Shirley, in prospettiva longitudinale

fanno i progressi maggiori e ottengono risultati più evidenti e lineari.

Infine, rispetto alle due misure considerate, i dati mostrano come l'andamento delle traiettorie di sviluppo della fluenza sia legato ad almeno tre fattori: il tipo di task, il livello di fluenza alla prima rilevazione, e infine il tipo di misura considerata. Gli andamenti infatti sono più lineari nei task interattivi, per le apprendenti meno avanzate, e per la misura delle pause che non nei task monologici, per le apprendenti più avanzate e per la misura dei fenomeni di esitazione.

In diversi casi poi le due misure della fluenza si controbilanciano, con aumenti delle pause seguiti da diminuzioni nelle esitazioni. Questo dato, in linea con la letteratura (cfr. § 1.2.2.2.), conferma l'ipotesi che esse rilevino diverse dimensioni di fluenza e pertanto sia utile combinarle. È anche vero però che, per questi dati, la misura legata alle pause ottiene meno variazione individuale tra le due studentesse italiane, e per questo si può supporre sia tra le due la più precisa nella misurazione complessiva di variazione longitudinale e situazionale.

### **3.4. Complessità, accuratezza e fluenza CAF a confronto**

Viene qui illustrato il confronto tra gli andamenti delle tre dimensioni CAF, esaminate separatamente in §§ 3.1.–3.3. Poiché, come abbiamo visto laddove sono state usate più misure alcune risultano più incisive di altre, qui ne viene selezionata una per ciascuna dimensione: la complessità sintattica viene misurata attraverso la misura subordinate per AS-unit, l'accuratezza come percentuale di AS-unit senza errori sul totale delle AS-unit prodotte, e infine la fluenza come numero medio di pause lunghe per AS-unit. Si inizierà

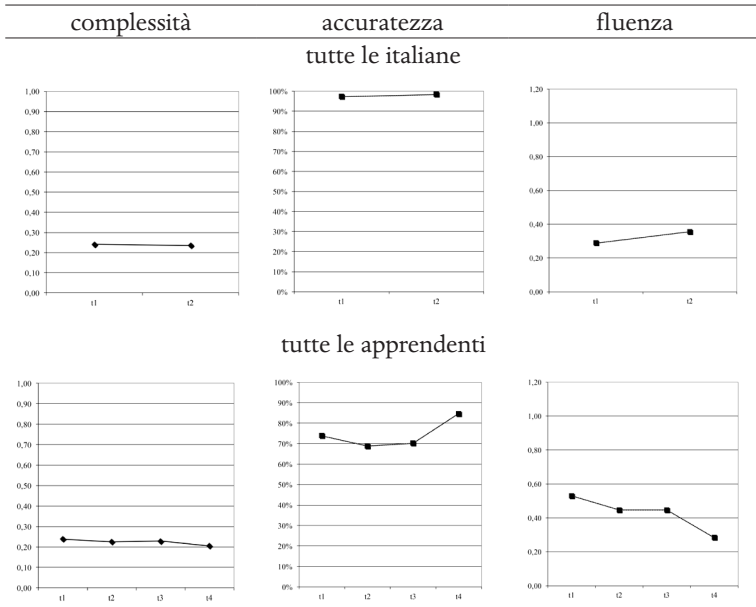
illustrando le interazioni tra le tre dimensioni nello sviluppo longitudinale cumulativamente in tutti i task (§ 3.4.1.), si continuerà con la variazione tra task (§ 3.4.2.), e si concluderà con il percorso individuale di ogni soggetto (§ 3.4.3.).

### 3.4.1. La variazione longitudinale

Il grafico 3.13. mostra i valori medi nelle tre misure scelte per le tre dimensioni considerando i due gruppi di soggetti cumulativamente in tutti i task.

A T1 si registrano differenze tra italiane e apprendenti, in misura diversa però per ogni dimensione: le apprendenti sono complessivamente meno fluenti e accurate delle compagne italiane,

**Grafico 3.13.** CAF — Variazione longitudinale — Tutti i task — Italiane e apprendenti (valori di gruppo).



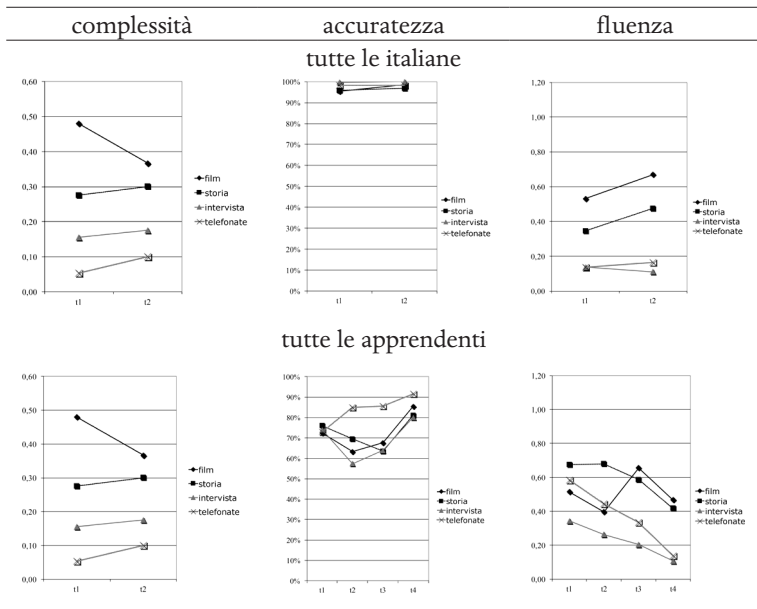
mentre non ci sono differenze sostanziali in complessità. Nel corso dei tre anni, senza sorprese, accuratezza e fluenza sono le dimensioni in cui si registra il maggior progresso. A t<sub>4</sub> le apprendenti si allineano ai valori delle compagne italiane rispetto alla fluenza, mentre rimangono alcune differenze in accuratezza.

### 3.4.2. La variazione situazionale

Il grafico 3.15 mostra le medie dei valori ottenuti dai due gruppi in ciascun task. I dati così disaggregati rivelano una certa variazione situazionale, arricchendo il quadro descritto in § 3.4.1.

Il gruppo di studentesse italiane dimostra variazione situazionale in tutte e tre le dimensioni, con differenze maggiori

**Grafico 3.14.** CAF — Variazione situazionale — Distribuzione per task — Italiane e apprendenti (valori di gruppo).





per complessità e fluenza: i task monologici risultano sintatticamente più complessi, ma un po' meno accurati e fluenti di quelli interattivi. Abbiamo già evidenziato in §§ 3.1.–3.3. come tale variazione rifletta le diverse richieste dei task (cfr. § 1.2.1.), con quelli monologici che favoriscono appunto la complessità e quelli interattivi la fluenza.

Nel tempo per le italiane si registra variazione: come già evidenziato dalle analisi condotte sulle singole misure (§§ 3.1.–3.3.) le traiettorie di sviluppo sono più o meno lineari e più o meno ascendenti in relazione al task considerato. Il confronto tra dimensioni CAF aggiunge poi un'ulteriore possibile interpretazione dell'andamento non lineare delle traiettorie di sviluppo: i possibili effetti di trade-off tra le tre dimensioni CAF. Nei task interattivi i progressi possono coinvolgere tutte e tre le dimensioni contemporaneamente, mentre nei task monologici si registrano interazioni tra le dimensioni: nel task storia un aumento della complessità è accompagnato da una riduzione della fluenza, mentre nel task film un aumento in accuratezza è accompagnato da una riduzione sia in fluenza che in complessità.

Vediamo ora la variazione situazionale definita dalle apprendenti. Alla prima rilevazione le apprendenti si discostano dal quadro di variazione disegnato dalle italiane, confermando quanto osservato in §§ 3.1.–3.3.: non solo dimostrano meno variazione tra task, ma in alcuni casi addirittura di polarità opposta. Se nei task monologici le apprendenti alla prima rilevazione sono già piuttosto vicine ai valori delle compagne, con differenze un po' più evidenti in accuratezza, nei task interattivi, soprattutto nelle telefonate, le differenze con le italiane sono più evidenti, in tutte e tre le dimensioni. I task interattivi hanno livelli di complessità più alti delle italiane, allo stesso livello di quelli monologici, oltre che essere caratterizzati da una minore fluenza. Come già sottolineato nell'analisi relativa

a ciascuna dimensione (§§ 3.1.–3.3.) i dati confermano dunque l'importanza della valutazione della proficiency in diverse attività e contesti, con la variabile monologico–interattivo che risulta essere particolarmente significativa.

Nel tempo le apprendenti ottengono notevoli progressi e a t<sub>4</sub> nel complesso sono molto più vicine ai valori e al quadro di variazione definito per le italiane, con i progressi più evidenti nel task in cui mostravano più difficoltà a t<sub>1</sub>, cioè quello interattivo di apertura delle telefonate. Così come per le italiane, i dati da un lato confermano quanto mostrato nell'analisi di ogni singola dimensione, cioè che i cambiamenti nel tempo seguono traiettorie diverse sulla base del task considerato, dall'altro aggiungono un ulteriore elemento che giustifica gli andamenti più o meno lineari delle traiettorie di sviluppo: i possibili effetti di trade–off tra dimensioni CAF. Vediamone alcuni esempi.

Nei task monologici i cambiamenti nel complesso riguardano soprattutto l'accuratezza, con lievi progressi in complessità e fluenza. Nel film tra t<sub>1</sub> e t<sub>2</sub> complessità e accuratezza si riducono contemporaneamente, con un progresso in fluenza; tra t<sub>2</sub> e t<sub>4</sub> si assiste ad una lieve riduzione in complessità, accompagnata da un evidente progresso in accuratezza. La fluenza dopo l'incremento iniziale, ha un andamento non lineare senza sostanziali cambiamenti nel tempo, controbilanciando gli andamenti nelle altre due dimensioni. Diversamente, nella storia, tra t<sub>1</sub> e t<sub>3</sub> aumentano complessità sintattica e fluenza, mentre l'accuratezza si riduce tra t<sub>3</sub> e t<sub>4</sub>, con un'inversione di tendenze: i resoconti risultano sintatticamente meno complessi, ma più accurati e molto più fluenti. Quest'ultima dimensione è quella che ottiene i progressi più evidenti.

Nei task interattivi invece le apprendenti si allineano ai valori delle compagne attraverso una riduzione della misura di complessità, accompagnata da un aumento in accuratezza e fluenza. Nell'intervista, senza sorprese, nel tempo i progressi

più evidenti sono ottenuti in fluenza, dove i valori si allineano con quelli delle compagne italiane, con effetti di trade-off invece tra complessità e accuratezza. Se la fluenza progredisce linearmente nel tempo, complessità, accuratezza si sviluppano secondo traiettorie non lineari, spesso con polarità opposta. Infine, nel task telefonate le apprendenti realizzano i progressi più importanti, la complessità decresce allineandosi a valori più vicini a quelli nativi, mentre accuratezza e fluenza progrediscono: le apprendenti a t4 producono meno errori e meno pause. Anche in questo caso è possibile osservare un'interazione tra complessità e accuratezza e fluenza, con una riduzione dei valori nella prima dimensione controbilanciata da progressi nelle altre due.

Ricapitolando, il quadro definito da italiane e apprendenti è chiaramente in accordo con i modelli psicolinguistici presentati in § 1.2.1.2., in cui si sostiene che le capacità di attenzione siano limitate e che pertanto nello svolgimento di un compito comunicativo, poiché il parlante decide di gestire le proprie risorse, si generano effetti di trade-off tra dimensioni. I dati mostrano come questo si rifletta anche nello sviluppo longitudinale: le traiettorie segnate dalle tre dimensioni variano non solo in relazione ai task ma anche ai conseguenti trade-off tra misure.

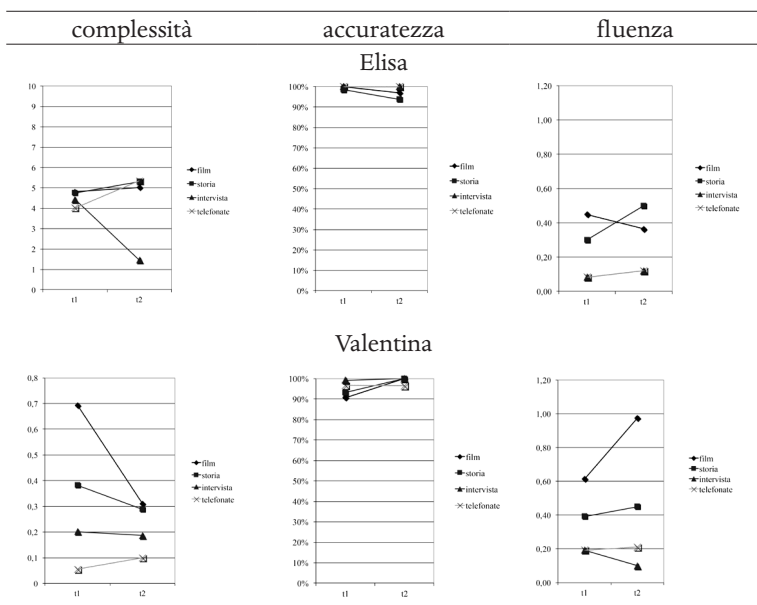
### 3.4.3. *La variazione individuale*

I grafici 3.15.–3.16. riportano i valori ottenuti in ciascun task disaggregando i dati per soggetto. I percorsi delle singole studentesse, pur mostrando variazione individuale riflettono il quadro delineato dai valori di gruppo, con i task monologici che risultano più complessi, ma meno fluenti e un poco meno accurati di quelli interattivi. Vediamo in dettaglio i percorsi delle due studentesse italiane e delle quattro apprendenti.

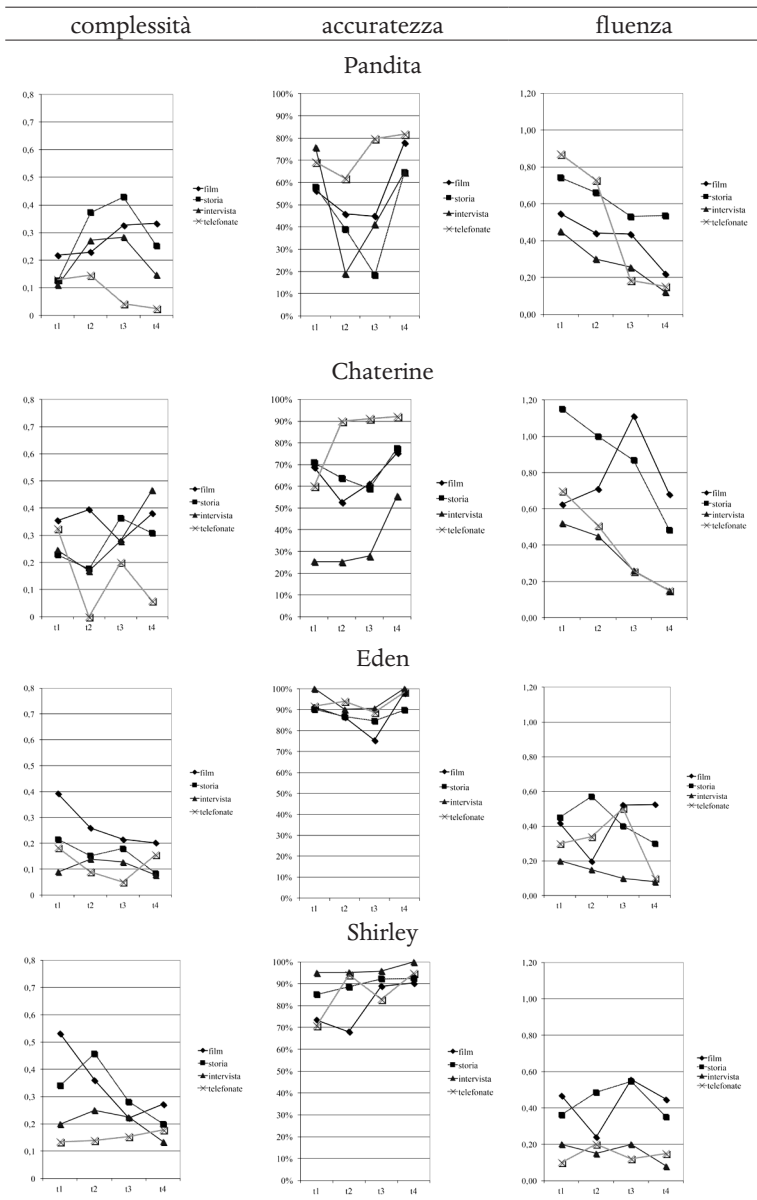
Elisa e Valentina dimostrano variazione individuale nei task monologici, a t<sub>1</sub> Elisa produce narrazioni meno complesse della compagna, risultando pertanto più fluente ed accurata. In prospettiva longitudinale le due studentesse convergono su valori simili. Da un lato Elisa complessifica la sua narrazione, con una conseguente riduzione in fluenza, Valentina, al contrario semplifica la sua narrazione, risultando anche più accurata.

I dati mostrano come in alcuni casi il parlante possa decidere di favorire diverse dimensioni CAF nello svolgere un determinato compito. Tale scelta può essere legata, come nel caso appena descritto, a scelte individuali rispetto alla strategia narrativa: poiché le capacità di attenzione sono limitate,

**Grafico 3.15.** CAF — Variazione situazionale — Distribuzione per task — Italiane.



**Grafico 3.16. CAF — Variazione individuale — Distribuzione per task — Apprendenti.**



se il parlante sceglie di realizzare un racconto complesso, ci saranno riduzioni in fluenza e accuratezza, come nel caso di Valentina; se sceglie invece di realizzare un racconto più semplice, la riduzione in complessità sarà accompagnata da un aumento di fluenza e di accuratezza, come nel caso di Elisa. In altri casi ancora, come nel task film a t2 per Valentina, la variazione può essere legata a fattori esterni al task: abbiamo già visto in § 3.1.–3.3. come la studentessa incontri diverse difficoltà nello svolgere questo compito, di conseguenza la sua narrazione è semplice, frammentaria, caratterizzata da pause ed esitazioni.

Il grafico 3.16., che riporta i risultati relativi alle quattro apprendenti, mostra variazione individuale, con Pandita e Catherine che a t1 dimostrano meno variabilità tra task rispetto a Eden e Shirley, risultando nel complesso meno avanzate delle compagne. Per tutte le apprendenti comunque le differenze più evidenti riguardano i task interattivi, in particolare quello relativo alle telefonate. Nel tempo tutte le ragazze realizzano diversi progressi, anche se variano da individuo a individuo dimensioni coinvolte e andamenti delle traiettorie di sviluppo. Vediamo nel dettaglio i percorsi individuali delle quattro apprendenti. Si illustreranno task per task prima i dati relativi a Pandita e Catherine, poi quelli di Eden e Shirley.

Pandita e Catherine, le meno avanzate a t1, dimostrano i progressi più significativi, con un sostanziale allineamento ai valori nativi a t4. Nel tempo c'è più variabilità tra task e la direzione di tale variazione è in sintonia con le produzioni delle studentesse italiane.

Nei task monologici per Pandita la crescita in complessità sintattica e l'acquisizione di strutture linguistiche generano un effetto di trade-off con l'accuratezza, mentre la fluenza cresce linearmente. Analogo il percorso di sviluppo di Catherine, anche se meno lineare. Nel suo caso possiamo osserva-

re trade-off tra tutte e tre le dimensioni. Anche per lei una riduzione in accuratezza è il risultato di uno sviluppo interlinguistico, diversamente da Pandita è però accompagnato da cambiamenti poco lineari anche in fluenza.

Nel task intervista Pandita e Catherine seguono invece percorsi di sviluppo fra loro diversi. Per Pandita non ci sono sostanziali differenze rispetto agli andamenti segnati dai task monologici: in una prima fase l'apprendente cresce soprattutto in complessità e fluenza, con un trade-off rispetto all'accuratezza, in una seconda fase, riavvicinandosi ai valori nativi, decresce in complessità, avendo imparato a diluire meglio le unità informative e mantenere lo scambio più fluido, mentre aumentano accuratezza e fluenza. Catherine in questo task segue al contrario un andamento più lineare, segnando un graduale progresso in tutte e tre le dimensioni.

Nel task relativo alle telefonate, sia Pandita che Catherine a T1 si discostano notevolmente dalle italiane, in prospettiva longitudinale i loro percorsi di sviluppo si differenziano. Pandita progredisce in tutte e tre le dimensioni seguendo traiettorie lineari: riduce la complessità sintattica e progredisce in accuratezza e fluenza. In questo task l'effetto di trade-off è legato a un avvicinamento anche pragmatico alle modalità di apertura delle telefonate tipiche dei nativi. Per Catherine invece le traiettorie di sviluppo CAF non sono lineari, nel suo caso i progressi in complessità e accuratezza generano un trade-off con la fluenza.

Le traiettorie di sviluppo di Eden e Shirley, relativamente più avanzate già a T1, hanno un andamento diverso da quello delle due compagne.

Nei task monologici, si assiste ad una riduzione della complessità sintattica in termini di subordinazione: l'analisi qualitativa presentata in § 3.1.1. mostra però come tale riduzione sia il risultato, nel caso di Eden dell'uso di clausole più lunghe,

con un progresso in complessità in termini di lunghezza, in quello di Shirley della scelta di uno stile narrativo più breve e semplice. Le diverse cause di riduzione dei valori di complessità in termini di subordinazione è controbilanciata da diversi andamenti della misura di accuratezza: nel caso di Eden la complessificazione in termini di lunghezza è accompagnata tra  $t_1$  e  $t_4$  da una parziale riduzione in accuratezza, nel caso di Shirley, la riduzione della complessità sintattica come conseguenza di una strategia narrativa più semplice porta ad un aumento dell'accuratezza. Tale sviluppo longitudinale genera anche nella fluenza un andamento non lineare e piuttosto simile per le due apprendenti, indicando a  $t_4$  un lieve progresso rispetto alla prima rilevazione, già in linea comunque con i valori delle compagne italiane.

Nel task intervista sia Eden che Shirley seguono traiettorie di sviluppo lineari in tutte le dimensioni, senza effetti di trade-off.

Nel task relativo alle telefonate invece le due apprendenti seguono percorsi meno lineari, oltre che diversi tra loro. Eden, così come Catherine, riduce la complessità sintattica con un lieve effetto di trade-off rispetto alla fluenza. Shirley segue invece traiettorie di sviluppo non lineari in tutte e tre le dimensioni, con la complessità che registra un lieve aumento.

I dati fin qui discussi ci permettono di riconfermare alcune osservazioni relative agli andamenti delle misure CAF e alle interazioni che si realizzano durante la realizzazione di un determinato task o come conseguenza dello sviluppo longitudinale. Come nel caso delle nostre apprendenti, la traiettoria di sviluppo di queste dimensioni può procedere in modo non lineare sulla base di almeno quattro elementi: livello iniziale, task considerato, variazione individuale e trade-off tra dimensioni. Gli andamenti non lineari nelle traiettorie di sviluppo e le interazioni tra dimensioni suggeriscono come



la combinazione delle misure CAF permetta una descrizione più affidabile del profilo di proficiency di un apprendente in un determinato task.

#### 3.4.4. *Riassunto*

In questa sezione è stata riportata l'analisi delle interazioni tra le dimensioni CAF. Per la complessità sintattica è stata considerata la misura subordinate per AS-unit, per l'accuratezza la percentuale di AS-unit senza errori sul totale delle AS-unit e infine per la fluenza il numero di pause lunghe per AS-unit.

Longitudinalmente, per tutte e quattro le apprendenti i dati dimostrano progresso, nel senso di un avvicinamento alle italiane in tutte e tre le dimensioni, con cambiamenti più evidenti in accuratezza e fluenza, le dimensioni che dimostravano le maggiori differenze a t1.

Situazionalmente, c'è variazione tra task. Per le studentesse italiane c'è variazione in tutte e tre le dimensioni: i task monologici risultano sintatticamente più complessi, ma un po' meno accurati e fluenti dei task interattivi. Anche per le apprendenti c'è un po' di variazione tra task a t1, ma le differenze sono meno evidenti che per le italiane, in diversi casi con una polarità inversa: ad esempio, rispetto a complessità e fluenza spesso le apprendenti risultano più complesse e meno fluenti nei task interattivi che in quelli monologici. Nel tempo tutte le apprendenti realizzano progressi, più evidenti nei task interattivi e rispetto a accuratezza e fluenza, e a t4 sono più vicine al quadro di variazione disegnato dalle compagne italiane. L'andamento non lineare delle misure CAF è legato in alcuni casi a interazioni tra le dimensioni stesse: ad esempio l'accuratezza e la complessità nei task monologici, all'aumentare dell'una si riduce l'altra.

Anche individualmente vi sono differenze quantitative e qualitative nelle traiettorie di sviluppo percorse da ciascun soggetto. Per le italiane, nei task monologici la scelta personale di uno stile narrativo più complesso, informativamente denso e coeso, piuttosto che uno stile più conciso e veloce, genera un'interazione tra complessità, accuratezza e fluenza, con la narrazione più complessa che risulta un po' meno accurata e fluente. Anche per le apprendenti c'è variazione individuale, con alcune differenze rispetto alle italiane: i dati individuali mostrano infatti come la variazione dipenda, oltre che da task e interazioni tra dimensioni, anche dal livello iniziale. Questo può determinare infatti andamenti più o meno lineari nelle misure, oltre che favorire un tipo di cambiamento piuttosto che un altro.

Infine, rispetto alle misure considerate, gli andamenti non lineari nelle traiettorie di sviluppo e le interazioni tra dimensioni suggeriscono come queste misure non possano essere considerate indicatori diretti della competenza, poiché un aumento o una diminuzione è un cambiamento in positivo e in negativo sulla base del livello di partenza e del target nativo, né possono essere usate singolarmente, poiché è la loro combinazione che permette la descrizione affidabile del profilo di proficiency di un apprendente in un determinato task.

### 3.5. Le valutazioni QCER

La sezione riporta i risultati delle valutazioni espresse dai due valutatori sulla base di tre descrittori del QCER: coerenza-coesione, correttezza grammaticale e fluenza. Sono qui considerati i dati individuali di tre task del corpus (storia, intervista e apertura delle telefonate di servizio) prodotti alla prima e all'ultima rilevazione (t1 e t2 per le italiane, t1 e t4 per le apprenden-

ti). Si illustreranno prima i risultati della variazione longitudinale per ciascun descrittore (§ 3.5.1.), quindi si presenteranno alcune osservazioni sulla variazione situazionale (§ 3.5.2.).

### 3.5.1. La variazione longitudinale

La tabella 3.9. mostra le valutazioni ottenute dalle due studentesse italiane. Elisa e Valentina non ottengono sempre il punteggio massimo, in particolare per il descrittore coerenza-coesione, dove si collocano in fascia B. Longitudinalmente registrano alcuni progressi, con differenze individuali rispetto ai descrittori coinvolti: Elisa mostra alcuni cambiamenti in termini di accuratezza e Valentina soprattutto in termini di fluenza.

La tabella 3.10. riporta i dati relativi alle apprendenti. Come atteso ricevono valutazioni più basse rispetto alle compagne italiane, Pandita e Catherine risultano meno avanzate in tutti e tre i descrittori e si collocano al livello B<sub>1</sub> o B<sub>1+</sub>, Eden e

**Tabella 3.9.** QCER — Variazione longitudinale — Italiane.

	Elisa		Valentina	
	t1	t2	t1	t2
coerenza-coesione				
storia	C1	C1	B2+	B2
intervista	B2+	B2+	B2+	C1
telefonate	C2	C1	C1	C1
accuratezza				
storia	C2	C2	C2	C2
intervista	C1	C2	C1	C2
telefonate	C2	C2	C2	C2
fluenza				
storia	C2	C2	C1	B2
intervista	C2	C1	C1	C2
telefonate	C2	C2	C1	C2

**Tabella 3.10.** QCER — Variazione longitudinale — Apprendenti.

	Pandita		Catherine		Eden		Shirley	
	t1	t4	t1	t4	t1	t4	t1	t4
coerenza-coesione								
storia	B1	C1	B1	B1	B2	C1	B2+	C1
intervista	B2	B2+	B1+	C1	?	C1	B2+	C1
telefonate	B1	B2+	?	B2	?	B2+	B2+	C1
accuratezza								
storia	B1+	B2+	B1+	B1+	B2+	C1	C1	C1
intervista	B2	B2	B1+	B2+	C1	C2	C1	C2
telefonate	B2	C1	B2	C1	C1	C1	C1	C1
fluenza								
storia	B1+	C1	B1	B1	B1+	B2+	C1	C2
intervista	B2+	C1	B1+	C1	C1	C1	B2+	C2
telefonate	B1+	C1	B1+	B1+	C1	C2	C1	C2

Shirley sono più avanzate e si collocano tra i livelli B2 e C2 in tutti i descrittori, con le valutazioni migliori in accuratezza. In prospettiva longitudinale tutte le apprendenti registrano progressi, anche se con una certa variazione individuale rispetto all'intensità e ai descrittori considerati.

Pandita è l'apprendente che ottiene i cambiamenti maggiori in tutte le dimensioni, soprattutto in fluenza, dove si assiste a un passaggio dal livello B al livello C. Catherine mostra meno variazione longitudinale della compagna, con i cambiamenti più evidenti in termini di accuratezza. Eden e Shirley progrediscono in tutte le dimensioni, con i cambiamenti maggiori in accuratezza per Eden e in coerenza-coesione e fluenza per Shirley.

Rispetto ai descrittori selezionati, si sono evidenziati alcuni limiti nell'uso del QCER. I valutatori infatti hanno avuto difficoltà a stabilire una valutazione per il descrittore relativo a coerenza-coesione: le discussioni per un accordo erano più lunghe

e in tre casi non è stato espresso alcun giudizio. Vedremo in § 3.5.2. come le difficoltà maggiori riguardino i task interattivi.

### 3.5.2. *La variazione situazionale*

Per studentesse italiane e apprendenti a t1 c'è variazione tra task, con i task interattivi che ottengono valutazioni in genere migliori. In prospettiva longitudinale i progressi si differenziano nei diversi task e coinvolgono diverse dimensioni, con una certa variazione individuale.

La tabella 3.11. riporta le valutazioni espresse per le studentesse italiane. Elisa e Valentina ottengono le valutazioni migliori nel task interattivo di apertura delle telefonate, mentre nei task storia e intervista il descrittore coerenza-coesione registra valutazioni più basse. Il task di apertura delle telefonate di servizio ottiene le valutazioni migliori in tutte e tre le dimensioni considerate.

**Tabella 3.11.** QCER — Variazione situazionale — Italiane.

	Elisa		Valentina	
	t1	t2	t1	t2
<i>storia</i>				
coerenza-coesione	C1	C1	B2+	B2
accuratezza	C2	C2	C2	C2
fluenza	C2	C2	C1	B2
<i>intervista</i>				
coerenza-coesione	B2+	B2+	B2+	C1
accuratezza	C1	C2	C1	C2
fluenza	C2	C1	C1	C2
<i>telefonate</i>				
coerenza-coesione	C2	C1	C1	C1
accuratezza	C2	C2	C2	C2
fluenza	C2	C2	C1	C2

La tabella 3.12. riporta le valutazioni per le quattro apprendenti. Come per le italiane, possiamo osservare variazione situazionale, con i task interattivi che ottengono valutazioni in genere più alte del task monologico.

A t1 Pandita raggiunge le valutazioni migliori per l'intervista in tutti i descrittori e per le telefonate nel descrittore accuratezza (B2). Così come per le italiane, la storia è il task che ottiene le valutazioni più basse (B1 e B1+). In prospettiva longitudinale, l'apprendente mostra un generale miglioramento, il più significativo riguarda le telefonate, dove per accuratezza e fluenza passa da B2 e B1+ a C1.

Catherine invece non mostra variazione situazionale a t1, collocandosi tra B1 e B1+ in tutti i task e per tutti i descrittori. In prospettiva longitudinale dimostra progressi soprattutto nel task intervista, mentre in quello relativo alle telefonate ci sono alcuni cambiamenti in termini di accuratezza. Non ci sono progressi nel task monologico.

**Tabella 3.12.** QCER — Variazione situazionale — Apprendenti.

	Pandita		Catherine		Eden		Shirley	
	t1	t4	t1	t4	t1	t4	t1	t4
<i>storia</i>								
coerenza-coesione	B1	C1	B1	B1	B2	C1	B2+	C1
accuratezza	B1+	B2+	B1+	B1+	B2+	C1	C1	C1
fluente	B1+	C1	B1	B1	B1+	B2+	C1	C2
<i>intervista</i>								
coerenza-coesione	B2	B2+	B1+	C1	?	C1	B2+	C1
accuratezza	B2	B2	B1+	B2+	C1	C2	C1	C2
fluente	B2+	C1	B1+	C1	C1	C1	B2+	C2
<i>telefonate</i>								
coerenza-coesione	B1	B2+	?	B2	?	B2+	B2+	C1
accuratezza	B2	C1	B2	C1	C1	C1	C1	C1
fluente	B1+	C1	B1+	B1+	C1	C2	C1	C2

A t1 Eden dimostra una certa variazione tra task, con la storia che ottiene valutazioni nella fascia B e i task interattivi nella fascia C. In prospettiva longitudinale, i progressi riguardano soprattutto il task storia, mentre nell'intervista progredisce solo l'accuratezza, e nelle telefonate la fluenza.

Shirley è la studentessa che a t1 ottiene le valutazioni migliori, con valutazioni un po' più basse nel task intervista e per il descrittore coerenza-coesione. In prospettiva longitudinale l'apprendente migliora in tutti i task, senza sorprese, soprattutto nell'intervista e per i descrittori coerenza-coesione e fluenza.

Così come anticipato in §3.5.1. il QCER mostra però alcuni limiti nella valutazione della coerenza-coesione nei task interattivi. In tre casi i valutatori faticano a raggiungere un accordo. In pratica, i valutatori hanno la sensazione che l'apprendente si trovi a un certo livello, ma le definizioni del descrittore non siano sempre adatte a descrivere la qualità della performance. Nel nostro caso infatti la descrizione della coerenza-coesione è formulata in termini di complessificazione e non è pertanto adatta a descrivere le richieste di un compito altamente interattivo. I dati suggeriscono dunque il bisogno di specificare la formulazione di questo descrittore in relazione al task considerato, monologico o interattivo. Limiti di questo tipo per l'applicazione del QCER al testing sono già stati evidenziati dalla letteratura (*linking problem*, cfr. Alderson *et al.*, 2004; Weir, 2005; Harsh, Martin, 2012; 2013; Wisniewski, 2017a; 2017b; Gyllstad *et al.*; 2014 § 1.3.2.; § 2.5.).

### 3.5.3. Riassunto

Riguardo alla descrizione delle valutazioni espresse da due valutatori a partire dai descrittori QCER selezionati (coerenza-coesione, accuratezza e fluenza) su tre task (storia, intervista

sta e telefonate) realizzati alla prima e all'ultima rilevazione, i dati mostrano variazione.

Longitudinalmente, tutte le apprendenti realizzano progressi. A T1 le apprendenti, come atteso, ottengono valutazioni inferiori rispetto a quelle delle compagne italiane. Nel tempo accuratezza e fluenza ottengono i cambiamenti più significativi.

Situazionalmente, si registrano a T1 per tutti i soggetti differenze tra task, con quello monologico che riceve le valutazioni più basse. Interessante notare che le studentesse italiane non raggiungono sempre il livello C. Nel tempo le apprendenti non progrediscono allo stesso ritmo in tutti i task, i cambiamenti più significativi riguardano infatti quelli interattivi.

Anche individualmente c'è variazione. A T1 Pandita e Catherine si collocano nella fascia B e Eden e Shirley più vicine alla fascia C. Nel tempo poi i progressi si differenziano da apprendente a apprendente per intensità e per task e descrittori coinvolti.

Infine, per quel che riguarda i descrittori QCER, quelli selezionati risultano nel complesso piuttosto efficaci, anche se per coerenza-coesione non sono sempre sensibili alla variazione tra task. I valutatori hanno difficoltà nell'esprimere la loro valutazione su questa dimensione nei task interattivi, in modo più evidente per le telefonate. Il descrittore QCER risulta poco sensibile alle caratteristiche di questa attività. Sarebbe dunque opportuno per chi utilizza il QCER ai fini della valutazione avere informazioni più dettagliate sulla qualità della produzione attesa per ogni livello differenziando per task, con la variabile monologico-interattivo che si rivela particolarmente interessante.

### 3.6. Misure CAF e valutazioni QCER a confronto

La sezione mette a confronto i risultati dell'analisi della variazione di complessità, accuratezza e fluenza dei sei soggetti,



ottenuti dall'applicazione delle misure CAF e dalle valutazioni soggettive QCER, prima esaminate separatamente (§§ 3.1.–3.4. e § 3.5.). Sono qui considerati i dati individuali di tre task del corpus, racconto di una storia ad immagini, intervista e apertura delle telefonate di servizio, realizzati a t<sub>1</sub> e t<sub>2</sub> dalle studentesse italiane e a t<sub>1</sub> e t<sub>4</sub> dalle apprendenti. Si discuterà prima la variazione longitudinale (§ 3.6.1.), poi la variazione situazionale (§ 3.6.2.). L'obiettivo è indagare le relazioni che intercorrono tra i due tipi di misurazioni e rilevare somiglianze, differenze ed eventuali punti di complementarità.

### 3.6.1. *La variazione longitudinale*

Nel complesso, i quadri di variazione longitudinale definiti dai due tipi di misurazione sono coerenti. Le studentesse italiane, senza sorprese, sono più avanzate delle apprendenti, sia CAF che QCER mostrano poi come non sempre raggiungono il punteggio massimo e come ci sia variazione individuale. In prospettiva longitudinale entrambi i tipi di misurazione rilevano lievi cambiamenti nel tempo. Anche per le apprendenti CAF e QCER disegnano quadri nel complesso coerenti. A t<sub>1</sub> entrambe le misurazioni mostrano come Pandita e Catherine siano meno avanzate di Eden e Shirley in tutte le dimensioni e come, nel tempo, tutte le ragazze registrano progressi.

Ciononostante, nell'analisi della variazione longitudinale il confronto tra CAF e QCER mette in luce alcune differenze. La prima riguarda l'ampiezza dello sviluppo delle singole dimensioni: le misure quantitative CAF mostrano come le differenze principali tra studentesse italiane e apprendenti riguardino accuratezza e fluenza, mentre QCER a t<sub>1</sub> non rileva questa differenza. In prospettiva longitudinale però i due strumenti concordano maggiormente, così come CAF anche QCER registra i progressi più evidenti in fluenza. Que-

sta prima differenza costituisce un punto di reciproca complementarità: l'analisi CAF infatti permette di identificare quali dimensioni ottengano il maggior cambiamento a livello avanzato, offrendo spunti utili per una miglior calibratura dei descrittori QCER.

La seconda differenza riguarda invece la direzione e l'andamento dello sviluppo delle dimensioni. Le traiettorie di sviluppo CAF non sono sempre lineari, abbiamo visto in §§ 3.1.–3.4. come questo sia il risultato dell'interazione di una serie di elementi, livello interlinguistico, sviluppo, task considerato e interazione tra dimensioni. Queste misure pertanto non sono indicatori diretti della competenza, ma devono essere interpretate a partire dalle reciproche relazioni e dal confronto con parlanti nativi. Le valutazioni QCER confermano l'idea che il progresso non sia indiscriminatamente un incremento in complessità o accuratezza: in tutti questi casi infatti segnano un progresso, guidando dunque l'interpretazione di CAF e confermando le osservazioni fatte nell'analisi delle singole misure. Approfondiremo meglio questa seconda differenza in § 3.6.2. descrivendola in rapporto ai singoli task.

### 3.6.2. *La variazione situazionale*

Il confronto tra misure CAF e valutazioni QCER nella descrizione della variazione situazionale evidenzia come i due tipi di misurazione, nonostante descrivano un quadro nel complesso coerente, mostrino alcune differenze.

Innanzitutto il QCER è poco sensibile alla variazione situazionale. I risultati dell'analisi CAF mostrano come in genere le apprendenti siano più distanti dal comportamento nativo nel task interattivo di apertura delle telefonate, e più vicine alle compagne italiane nell'intervista e nel task monologico. Il QCER non riflette esattamente lo stesso quadro, le appren-

denti ottengono valutazioni equivalenti in tutti i task e spesso migliori in quelli interattivi. Anche longitudinalmente, se CAF mostra maggior progresso in alcuni task e dimensioni, QCER rileva un miglioramento complessivo in tutti i task e tutte le dimensioni. In questo senso è evidente come QCER e CAF differiscono nel registrare le variazioni tra task. Vediamo più in dettaglio le differenze che emergono dal confronto tra CAF e QCER nei singoli task.

Nei task monologici le differenze riguardano gli andamenti di complessità e accuratezza. Nel caso dell'accuratezza, per Pandita, Catherine e Eden la misura CAF segue traiettorie ad U, mentre per Shirley lo sviluppo è lineare. Abbiamo visto in § 3.2. come questo sia legato all'interazione tra misura scelta e sviluppo interlinguistico. Il valutatore, al contrario, più attento al tipo di errore, piuttosto che al numero, probabilmente combina in qualche modo nella sua interpretazione i costrutti sviluppo e accuratezza, registrando così progressi in accuratezza per tutte e quattro le apprendenti. In questo caso dunque QCER e CAF si integrano e possono guidare l'uno l'interpretazione dell'altro. Rispetto alla complessità invece, abbiamo osservato in § 3.1.3. come le traiettorie di sviluppo CAF delle apprendenti si differenziano sulla base del livello di partenza iniziale, per Pandita e Catherine, le meno avanzate, i progressi sono in termini di subordinazione, mentre per Eden e Shirley in lunghezza o dovuti alla sperimentazione di una più ampia varietà di stili discorsivi. Abbiamo più volte sottolineato come questo dato confermi l'ipotesi di Predizione dello Sviluppo, secondo la quale a livelli più avanzati la complessificazione non riguarda tanto la quantità di subordinate usate, ma piuttosto la varietà o l'uso di strutture più complesse a livello di clausola. QCER segna progressi per tutte le ragazze, mostrandosi dunque sensibile ad entrambi i tipi di complessificazione. Pure in questo caso dunque QCER e CAF integrano l'uno l'interpretazione dell'altro.

Anche nei task interattivi le differenze riguardano la dimensione complessità sintattica. In questi task i valutatori hanno avuto alcune difficoltà nell'applicazione del descrittore QCER coerenza-coesione, in tre casi addirittura non sono stati in grado di arrivare a un accordo. Questo sembra legato a una percepita discrepanza tra la descrizione del QCER e le richieste del task interattivo. L'analisi CAF mostra infatti come questo compito favorisca la produzione di AS-unit semplici con una bassa densità informativa, il descrittore QCER al contrario definisce la coerenza-coesione in termini di complessificazione, risultando pertanto poco adatto alla valutazione del task telefonate. In questo caso i profili di variazione tra italiane e apprendenti definiti dalle misure CAF potrebbe essere un utile punto di partenza per migliorare la formulazione del descrittore QCER coerenza-coesione.

### 3.6.3. *Riassunto*

Messe a confronto le analisi CAF e le valutazioni QCER, entrambe rilevano variazione nelle produzioni di studentesse italiane e apprendenti. A volte i due tipi di misurazione evidenziano le stesse tendenze, altre i due approcci risultano complementari, con uno che guida l'interpretazione dell'altro. Il QCER complementa CAF nell'interpretazione dell'andamento delle misure di complessità sintattica e di accuratezza in quelli monologici. Sono questi i casi, ad esempio, dell'andamento non lineare delle misure CAF di complessità sintattica nei task interattivi o di accuratezza in quelli monologici. CAF invece complementa QCER nella descrizione della variazione situazionale. Sono questi i casi del descrittore coerenza-coesione nel task telefonate e delle indicazioni che CAF può dare rispetto alle dimensioni che variano a livello avanzato in relazione ai task.

## Conclusioni

Questa ricerca è stata condotta con lo scopo di osservare da un lato la variabilità longitudinale e situazionale dell'interlingua di apprendenti avanzati impegnati nello svolgimento di quattro task comunicativi, e dall'altro di verificare la relazione che intercorre tra due tipi di misurazione: quella analitica che rileva la complessità, l'accuratezza e la fluenza del parlato (CAF), impiegata nei cosiddetti *developmental index studies* di stampo acquisizionale, e quella più soggettiva basata sui descrittori del *Quadro Comune Europeo di Riferimento* (QCER) maggiormente legata all'ambito del testing.

Il corpus analizzato è costituito dalle produzioni di quattro studentesse immigrate di diversa nazionalità, raccolte in quattro rilevazioni nell'arco di tre anni, integrato con due rilevazioni di due parlanti native con simili caratteristiche, che costituiscono la pietra di paragone per il confronto con le apprendenti.

Nel presente capitolo si tratterà un quadro sintetico dei risultati ottenuti dall'analisi, nel tentativo di rispondere alle domande che hanno guidato la ricerca. Si proverà infine a individuare spunti di applicazione dei risultati agli ambiti della ricerca acquisizionale, del testing e della didattica.

Rivediamo dunque le domande alla base della ricerca, per poi procedere alla presentazione delle risposte ricavate dall'analisi dei dati.

*Come variano complessità, accuratezza e fluenza nel corso del tempo man mano che progredisce l'interlingua?*

Alla prima rilevazione, senza sorprese, l'analisi cumulativa su tutte le produzioni rileva differenze importanti tra i due gruppi delle studentesse italiane e delle apprendenti, in misura diversa però per ogni dimensione: se in complessità non sono sostanziali, in fluenza e accuratezza sono piuttosto evidenti, poiché le apprendenti ottengono valori inferiori a quelli delle compagne italiane. Nel tempo il gruppo delle apprendenti registra variazione longitudinale in tutte e tre le dimensioni, con progresso più evidente nell'accuratezza e nella fluenza, rispetto alla complessità. All'ultima rilevazione le ragazze sono perciò più vicine ai valori ottenuti dalle compagne, nonostante permangano alcune differenze in accuratezza.

Le traiettorie di sviluppo di CAF non sono però lineari per tutte le dimensioni. Nel caso della complessità e dell'accuratezza i dati di gruppo rivelano come il livello di competenza linguistica avanzato delle quattro apprendenti possa costituire una prima spiegazione degli andamenti più o meno lineari. Nel caso della complessificazione sintattica, in accordo con quanto suggerito dall'Ipotesi di Predizione dello Sviluppo proposta per esempio da Cooper (1976), Monroe (1975), Sharma (1980) e Wolfe–Quintero *et al.* (1998), le apprendenti non progrediscono più tanto in quantità di subordinate impiegate, quanto in distribuzione di clausole per tipo e in lunghezza. In modo analogo procede l'accuratezza: le apprendenti mostrano traiettorie di sviluppo a U, con riduzioni legate a sperimentazioni o ristrutturazioni di forme linguistiche più avanzate. Questo dato è in linea con quanto evidenziato in diversi studi acquisizionali (e.g. Larsen–Freeman, 1983) che mostrano come dal livello intermedio all'avanzato possa aumentare l'incidenza di errori, con la minor accuratezza accompagna-

ta da evoluzioni interlinguistiche. D'altronde anche i modelli psicolinguistici (Skehan, 1998; cfr. § 1.1.2.) mostrano come l'accuratezza sia determinata dall'interazione di tre elementi: il grado di accuratezza della rappresentazione stessa, la forza di eventuali rappresentazioni in competizione tra loro e infine il grado di automatizzazione della produzione stessa. In altre parole, in questo caso l'andamento non lineare è legato all'interazione tra i due costrutti, accuratezza e sviluppo, con apprendenti che possono essere molto accurati ma poco evoluti in una prima fase, e meno accurati ma più evoluti in una successiva, fino ad ottenere un riallineamento sui valori nativi. In questo caso con evoluto si intende sia la vicinanza della rappresentazione mentale alla norma della lingua target sia il grado di automatizzazione.

L'analisi della variazione longitudinale conferma dunque quanto già suggerito in letteratura; il progresso non è caratterizzato da un aumento in tutte le dimensioni contemporaneamente. Anzi, a livello avanzato il progresso può riguardare alcune dimensioni più di altre ed essere ottenuto sia tramite aumenti che riduzioni dei valori delle misure CAF.

*Come variano le tre dimensioni CAF con il variare dell'attività comunicativa per gli apprendenti rispetto ai parlanti nativi?*

L'analisi della variazione situazionale effettuata disaggregando i dati di gruppo per task, permette di raffinare ulteriormente il quadro sopra delineato.

Alla prima rilevazione, le studentesse italiane dimostrano variazione tra task, nel senso che i loro task monologici sono sintatticamente più complessi, e un po' meno accurati e fluenti di quelli interattivi. Questo dato è in linea con i modelli psicolinguistici dell'*information processing theory* secondo i quali i parlanti gestiscono le proprie risorse per risponde-

re alle richieste della situazione comunicativa, favorendo in alcuni casi la fluenza, in altri la complessità o l'accuratezza (cfr. Skehan, 1998; 2011; 2014; 2015; 2019 e Robinson, 2001a; 2011a; 2011b; 2015; 2017). Anche per le apprendenti c'è variazione tra task fin dalla prima rilevazione, con differenze poco evidenti rispetto alle italiane, a volte di polarità inversa. Nel primo caso si tratta dei due task monologici, nel secondo dei task interattivi. Qui le apprendenti risultano più complesse e fluenti delle italiane.

In prospettiva longitudinale tutte le ragazze registrano notevoli progressi, senza sorprese nel task in cui differivano maggiormente dalle italiane alla prima rilevazione. Gli andamenti nel tempo delle tre dimensioni variano con i diversi task, permettendo di identificare così una seconda ragione degli andamenti più o meno linearmente ascendenti delle misure considerate. Poiché i task monologici richiedono maggiore controllo sulla complessità, e quelli interattivi sulla fluenza, le apprendenti devono acquisire questo quadro di variazione. Pertanto nei task interattivi il progresso si realizza attraverso una riduzione della complessità; saper gestire un evento altamente interattivo come l'apertura delle telefonate di servizio non richiede un mero controllo delle regole sintattiche, ma piuttosto la capacità di produrre una rapida serie di turni. Il tipo di task può quindi generare andamenti non lineari in virtù di un riallineamento con i pattern di variazione situazionale dei nativi. I dati dunque, in accordo con quanto evidenziato dalla letteratura con approccio sociolinguistico, evidenziano come apprendere una L2 significa acquisire anche i pattern di variazione tipici dei nativi (cfr. Rehner, Mougeon, 1999; Dewaele, Regan, 2002; Donaldson, 2017). Di qui l'importanza di un'osservazione della performance in diversi task e contesti comunicativi, con la variabile monologico-interattivo che si rivela particolarmente fruttuosa.



*Come si sviluppa il percorso individuale di ciascun soggetto?*

L'analisi dei percorsi individuali mostra variazione individuale per tutti i soggetti. Le studentesse italiane registrano variazione individuale soprattutto nei task monologici. Alla prima rilevazione Elisa produce narrazioni più snelle e meno complesse della compagna, risultando pertanto più fluente e accurata. Questo dato conferma quanto suggerito dai modelli psicolinguistici. Un parlante infatti dovendo gestire le proprie risorse d'attenzione, nel realizzare un racconto può decidere di volta in volta se adottare uno stile narrativo che favorisce la complessità, a scapito di accuratezza e fluenza, o viceversa. Nel tempo le due apprendenti convergono su valori simili. I cambiamenti registrati da Valentina nel task film permettono un'ulteriore riflessione: la variazione individuale può essere legata a fattori esterni sia al task che alla competenza. In questo caso infatti Valentina produce una narrazione semplice, frammentaria e poco fluente perché incontra alcune difficoltà nella comprensione del video e si mostra poco coinvolta nell'attività.

Anche le apprendenti mostrano variazione individuale, Pandita e Catherine al t<sub>1</sub> risultano meno avanzate di Eden e Shirley. Per le prime due ragazze il progresso nel corso dei tre anni è prevedibilmente maggiore che per le seconde due. L'analisi mostra poi come l'avvicinamento al comportamento linguistico delle studentesse italiane procede in modo e in misura diversa da apprendente a apprendente. Questo è dovuto in alcuni casi a scelte individuali, è il caso di Shirley, una delle apprendenti più avanzate. Tra t<sub>1</sub> e t<sub>4</sub> Shirley nei task monologici riduce la complessità, cambiando stile narrativo, scegliendo una narrazione caratterizzata da clausole indipendenti e frammenti giustapposti sostenuti da un'intonazione calzante ed espressiva. In altri casi ciò dipende dal

livello alla prima rilevazione. Le apprendenti meno avanzate nei task monologici aumentano la complessità e seguono una traiettoria di sviluppo a U in accuratezza, nei task interattivi invece si riduce la complessità, a favore di accuratezza e fluenza. Le apprendenti più avanzate riducono la complessità in tutti i task, con progressi tendenzialmente più lineari in accuratezza e fluenza.

Dati di gruppo e dati individuali risultano così complementari: il gruppo definisce le tendenze generali, il caso individuale osservato longitudinalmente permette di ricostruire ciò che accade durante il percorso di apprendimento. Importante sottolineare come nei dati la variazione individuale si collochi in genere entro lo stesso quadro di variazione del gruppo, sostenendo così l'idea suggerita dalla letteratura (cfr. Pienemann, 2008; Larsen–Freeman, 2006), secondo cui l'interlingua procede per traiettorie individuali e aderisce contemporaneamente a una sequenzialità più universale.

#### *Quali sono le interazioni tra queste tre dimensioni dell'interlingua?*

Le tre dimensioni CAF possono avere traiettorie di sviluppo non lineari, oltre che per i fattori visti sopra, in virtù delle interazioni che si creano tra loro. È questo il caso delle produzioni delle studentesse italiane e di Pandita e Catherine nei task monologici. Per le studentesse italiane, le scelte stilistiche individuali che abbiamo appena commentato rispetto alle modalità narrative generano un trade-off tra le tre dimensioni. Se realizzare un racconto complesso significa ridurre l'attenzione alla forma e dunque all'accuratezza, realizzarne uno più semplice e fluente porta a una riduzione della complessità. In accordo con i modelli psicolinguistici dell'*information processing view* (cfr. § 1.2.1.2.), nello svolgimento di un com-

pito comunicativo il parlante può gestire le proprie risorse, e scegliere in alcuni casi di favorire il significato affidandosi al sistema basato sugli esempi, con un conseguente aumento della fluenza, in altri di favorire il sistema basato sulle regole, ottenendo così una maggiore complessità. Per Pandita e Catherine, la produzione di narrazioni più complesse e l'uso di forme di subordinazione più avanzate generano in alcune fasi compensazione sull'accuratezza e sulla fluenza. Questo effetto non si realizza invece nei task interattivi, proprio perché questi ultimi richiedono la produzione di AS-unit semplici, favorendo pertanto uno sviluppo più lineare di accuratezza e fluenza.

L'andamento non lineare di alcune misure, dovuto a una serie di fattori di natura alquanto diversa — la diversa competenza linguistica delle apprendenti alla prima rilevazione, il tipo di task che svolgono, lo stile discorsivo che adottano, e infine le interazioni tra dimensioni, cui conseguono effetti di trade-off — suggerisce dunque una nota di cautela nel generalizzare conclusioni sulla competenza in base a un unico dato, poiché a un aumento o a una diminuzione in una misura non corrisponde automaticamente un progresso o un regresso. Di qui l'importanza di una loro combinazione e di un confronto con la variazione nativa quando l'obiettivo è arrivare all'elaborazione di profili di proficiency più affidabili.

*I descrittori selezionati dal QCER sono efficaci nella valutazione delle competenze orali di studentesse italiane e apprendenti di italiano L2 di livello avanzato? In particolare sono sensibili allo sviluppo longitudinale e alla variazione tra task?*

Nel complesso i descrittori del QCER si sono rivelati efficaci nella valutazione delle competenze orali tanto delle studen-

tesse italiane quanto delle apprendenti e hanno rilevato variazione longitudinale e situazionale.

Le italiane, pur non ottenendo sempre il punteggio massimo, si collocano nella fascia C, e vengono dunque valutate come più avanzate delle apprendenti. Queste ottengono valutazioni più basse in tutti i task, con differenze più evidenti per Pandita e Catherine, che sono peraltro meno avanzate di Eden e Shirley. In prospettiva longitudinale, le apprendenti dimostrano progressi, registrando nel corso dei tre anni, seppur in misura diversa da studentessa a studentessa, il passaggio dal livello B al livello C.

Nonostante l'efficacia complessiva dei descrittori QCER, sono emerse alcune difficoltà, in particolare rispetto al descrittore coerenza-coesione per i task interattivi. Questo evidenzia come il QCER non sia sempre sensibile alla variazione tra task. Non a caso, su questo aspetto le discussioni tra valutatori per raggiungere un accordo sono state più lunghe e a volte infruttuose. Come già evidenziato nelle pubblicazioni in merito (per es. Alderson *et al.*, 2004; Weir, 2005a; 2005b), il QCER nella sua formulazione attuale non è sempre in grado di rispondere ai bisogni concreti di chi si occupa di valutazione. Chi utilizza le griglie avverte la necessità di sapere qual è il livello di qualità della performance che ci si aspetta nello svolgere un determinato compito a un determinato livello, come nel caso delle telefonate. Superare questi limiti richiederà di approfondire il legame tra i diversi livelli e le forme d'uso, in particolare per i livelli più alti, che pongono appunto le maggiori difficoltà (cfr. Alderson *et al.*, 2004; Ferrari, Nuzzo, 2010). Su questo tema è auspicabile una maggiore interazione tra la linguistica acquisizionale e il testing.

*Che relazione intercorre tra le misure analitiche CAF e le valutazioni soggettive della proficiency basate sul QCER?*

Il confronto tra i risultati ottenuti con le misure CAF e quelli ottenuti con le valutazioni QCER ha permesso di verificare come i due sistemi di misurazione rilevano nel complesso tendenze di variazione longitudinale e situazionale tra loro coerenti. In alcuni casi essi disegnano le stesse tendenze, in altri casi invece l'interpretazione di uno complementa quella dell'altro.

Si ritrovano le stesse tendenze nella variazione individuale nel tempo, in entrambe le misurazioni infatti le studentesse italiane e due delle apprendenti risultano più avanzate, con conseguenti differenze nei percorsi di sviluppo. Si tratta invece di interpretazioni complementari tra CAF e QCER nel caso degli andamenti non lineari delle misure CAF per complessità e accuratezza. Da un lato, il QCER può guidare l'interpretazione di CAF. Infatti abbiamo visto come il tipo di task, monologico o interattivo, il livello di partenza, le scelte individuali del parlante e il trade-off tra dimensioni CAF generano andamenti non sempre lineari nelle misure considerate, mentre il QCER registra ovunque progresso, guidando pertanto l'interpretazione degli andamenti delle misure stesse. Dall'altro lato, è invece il CAF che potrebbe contribuire a raffinare la il QCER. L'analisi CAF mostra come i task interattivi favoriscano le produzioni di AS-unit semplici (ossia con subordinazione ridotta e limitata ad alcune forme specifiche), con una bassa densità informativa e l'introduzione di un elemento per volta. Questi elementi potrebbero costituire un'utile base per una ridefinizione del descrittore coerenza-coesione, che risulta invece poco adatto a rilevare la formazione situazionale.

## **Implicazioni per la ricerca acquisizionale, il testing e la didattica**

I risultati ottenuti in questo lavoro, per quanto riferiti ad un campione limitato e pertanto da considerarsi con cautela, ci permettono di trarre alcune implicazioni per la ricerca acquisizionale, per il testing e per la didattica, in parte confermando quanto già evidenziato da altri, in parte contribuendo qualcosa di nuovo.

Per quanto riguarda l'acquisizione, i risultati confermano ciò che è emerso anche in studi precedenti, ovvero che alcuni aspetti dell'interlingua variano sistematicamente secondo i task e le situazioni comunicative, e che pertanto ogni task fornisce un'immagine parziale dell'interlingua. Molti studi acquisizionali invece traggono le loro conclusioni usando dati elicitati per mezzo di un solo tipo di compito, spesso monologico. Una ricerca come questa mostra che attraverso la combinazione di più task è possibile creare profili più attendibili delle competenze di un parlante e avere informazioni più precise circa gli andamenti nel tempo. L'analisi ha evidenziato inoltre come il percorso evolutivo non corrisponda sempre ad un aumento lineare e generalizzato della complessità sintattica, dell'accuratezza o della fluenza.

Per quanto riguarda il testing linguistico, i risultati di questo studio sottolineano che per ottenere una valutazione precisa dello sviluppo della proficiency è necessario osservare le produzioni degli apprendenti in diverse attività e contesti, rappresentativi di tutte le competenze che si vogliono testare. I dati evidenziano inoltre che è utile non separare rigidamente la competenza linguistica dalla competenza interazionale, proprio perché esse sono strettamente connesse. Infine lo studio conferma l'efficacia di una valutazione basata sui task, come sostengono da tempo molti autori (Brown *et al.*, 2002;

Language Testing, 2002). L'uso dei task, oltre a garantire una maggiore validità esterna o ecologica, consente appunto di esaminare come la competenza linguistica e quella comunicativa interagiscono nello svolgimento di specifiche attività.

Per quanto riguarda l'interfaccia tra misure CAF e valutazioni QCER, l'indagine svolta conferma l'utilità in questo senso di una maggiore interazione tra ricerca acquisizionale e testing linguistico. Chi si occupa di testing può trovare nell'acquisizione spunti per una migliore definizione degli elementi linguistici corrispondenti ai livelli di proficiency, necessari per un maggior affinamento delle scale di proficiency. D'altro canto, chi si occupa di acquisizione può trovare nel testing strumenti per una più rigorosa verifica dell'affidabilità degli strumenti di elicitazione e misurazione, elemento indispensabile per ottenere non solo risultati più robusti, ma permettere anche una maggiore replicabilità degli studi.

Per l'insegnamento, nella costruzione di un syllabo linguistico-comunicativo la variabilità situazionale riscontrata nella nostra analisi mostra quanto è importante tener conto dei dettagli dell'interazione e delle routine ad esse collegate. Ma per conoscerli, come nel caso dell'apertura della telefonata, non basta ricorrere all'intuizione dell'insegnante, ma è doveroso partire dai dati della ricerca (Bettoni, 2006; Boxer, 2010). Anche sul piano della didattica occorre dunque approfondire il legame tra competenza linguistica e competenza interazionale. I dati di questa ricerca mostrano che esiste una sintassi *sui generis* per l'interazione, di conseguenza la glottodidattica dovrà tenere conto delle ricerche sulla *grammar in interaction* (Ochs *et al.*, 1996; Ford *et al.*, 2002), contestualizzando l'uso delle varie forme linguistiche all'interno di pratiche comunicative. Fermo restando che l'obiettivo generale dovrebbe essere quello di stimolare negli apprendenti l'attenzione e la consapevolezza per certi fenomeni, l'apertura di una telefo-

nata può costituire un buon esempio di situazione didatticizzabile: poche mosse, relativamente predeterminate quanto al contenuto, possono essere realizzate in una grande varietà di modi, ciascuno dei quali avrà conseguenze sui dettagli della dinamica interazionale di un particolare frammento (cfr. Pallotti, Ferrari, 2008). Si aprono così interessanti prospettive per una didattica della grammatica focalizzata sui micro–dettagli dell’interazione, che può condurre all’elaborazione di sillabi linguistico–comunicativi innovativi.



## Bibliografia

- ADAMS R. (2009), *Recent publications on task-based language teaching: a review*, in «International Journal of Applied Linguistics» 19, pp. 339–55.
- ADAMSON H.D. (2008), *Interlanguage Variation in Theoretical and Pedagogical Perspective*, Routledge, New York.
- ADAMSON H.D., REGAN V. (1991), *The acquisition of community norms by Asian immigrants learning English as a second language: A preliminary study*, in «Studies in Second Language Acquisition» 13, pp. 1–22.
- AHMADIAN M.J. (2012), *The effects of guided careful online planning on complexity, accuracy and fluency in intermediate EFL learners' oral production: The case of English articles*, in «Language Teaching Research» 16, 1, pp. 129–49.
- AHMADIAN M.J., TAKAVOLI M. (2010), *The effects of simultaneous use of careful online planning and task repetition on accuracy, complexity and fluency in EFL learners' oral production*, in «Language Teaching Research» 15, 1, pp. 35–59.
- ALANEN R., HUHTA A., TARNANEN M. (2010), *Designing and assessing L2 writing tasks across CEFR proficiency levels*, in I. Bartning, M. Martin, I. Vedder (eds.), *Communicative Proficiency and Linguistic Development: intersections between SLA and language testing research*, Eurosla monographs, pp. 21–56.

- ALANEN R., HUHTA A., TARNANEN M., JARVIS S., MARTIN M. (2011), *Issues and challenges in combining SLA research and language testing*, in D. TSAGARI, I. Csepes (eds.), *Collaboration in language testing and assessment*, Peter Lang, Frankfurt, pp. 5–30.
- ALBERT A. (2011), *When individual differences come into play: The effect of learner creativity on simple and complex task performance*, in P. Robinson (ed.), *Second Language Task Complexity. Researching the Cognition Hypothesis of language learning and performance*, John Benjamins Publishing Company, Amsterdam, pp. 239–66.
- ALDERSON J.C. (1991), *Bands and scores*, in J.C. Alderson, B. North (eds.), *Language Testing in the 1990s*, Macmillan, London, pp. 71–86.
- (2002) (ed.), *Common European Framework of Reference for Languages: Learning, Teaching, Assessment. Case studies*, Council of Europe, Strasbourg.
- (2005a), *Diagnosing foreign language proficiency: The interface between learning and assessment*, Continuum, London.
- (2005b), *Editorial*, in «Language Testing» 22, pp. 257–60.
- (2007), *The CEFR and the Need for More Research*, in «The Modern Language Journal» 91, 4, pp. 659–63.
- (2010), *Language testing–informed SLA? SLA–informed language testing?*, in I. Bartning, M. Martin, I. Vedder (eds.), *Communicative Proficiency and Linguistic Development: intersections between SLA and language testing research*, Eurosla monographs, pp. 239–47.
- ALDERSON J.C., BRUNFAUT T., HARDING L. (2017), *Bridging assesment and learning: a view from second and foreign language assessment*, in «Assessment in Education: Principle, Policy and Practice» 24, 3, pp. 379–87.
- ALDERSON J.C., CLAPHAM C., WALL D. (1995), *Language test construction and evaluation*, Cambridge University Press, Cambridge.

- ALDERSON J.C., FIGUERAS N., KUIJPER H., NOLD G., TAKALA S., TARDIEU C. (2004), *The development of specifications for item development and classification within the Common European Framework of Reference for Languages, learning, teaching, assessment. Reading and listening: final report of the Dutch Construct Project*. Available on request from the Project Coordinator, J.C. Alderson, c.alder-son@lancaster.ac.uk.
- ALDERSON J.C., HUTHA A. (2011), *Can research into the diagnostic testing of reading in a second or foreign language contribute to SLA research?*, in L. Roberts, G. Pallotti, C. Bettoni (eds.), «Eurosla Yearbook» 11, *Annual Conference of the European Second Language Association*, John Benjamins Publishing Company, Amsterdam, pp. 30–52.
- ALDERSON J.C., URQUHART A.H. (1985), *The effect of students' academic discipline on their performance on ESP reading tests*, in «Language Testing» 2, pp. 192–204.
- ANDERSON J.R. (1995), *Learning and Memory: An Integrated Approach*, Wiley, New York.
- ANDERSON N., BACHMAN L.F., COHEN A.D., PERKINS K. (1991), *An exploratory study into the construct validity of a reading comprehension test, triangulation of data sources*, in «Language Testing» 8, pp. 41–66.
- Applied Linguistics* (2006), «Special Issue» 27, G. Cook, G. Kasper (eds.).
- Applied Linguistics* (2009), «Special Issue» 30, 4, A. Housen, F. Kuiken, I. Vedder (eds.).
- ATKINSON J.M., HERITAGE J. (1984), *Transcription Notation*, in J.M. Atkinson, J. Heritage (eds.), *Structure of Social Interaction: Studies in Conversation Analysis*, Cambridge University Press, Cambridge, pp. IX–XVI.

- BACHMAN L.F. (1988), *Language testing—SLA research interfaces*, in «Annual Review of Applied Linguistics» 9, pp. 193–209, ristampato in L.F. Bachman, A. Cohen (1998) (eds.).
- (1990), *Fundamental considerations in language testing*, Oxford University Press, Oxford.
- (1998), *Language testing—SLA research interfaces*, in L.F. Bachman, A.D. Coehen (eds.), *Interfaces between second language acquisition and language testing research*, Cambridge University Press, Cambridge, pp. 177–95.
- (2000), *Modern Language Testing at the turn of the century: assuring that what we count counts*, in «Language Testing» 17, pp. 1–41.
- (2004), *Statistical analysis for language assessment*, Cambridge University Press, Cambridge.
- BACHMAN L.F., COHEN A. (1998a) (eds.), *Interfaces between second language acquisition and language testing research*, Cambridge University Press, Cambridge.
- BACHMAN L.F., DAVIDSON F., MILANOVIC M. (1996), *The use of test methods in the content analysis and design of EFL proficiency tests*, in «Language Testing» 13, pp. 125–50.
- BACHMAN L.F., PALMER A.S. (1981), *A multitrait multimethod investigation into the construct validity of six tests of speaking and reading*, in A.S. Palmer, J.M. Groot, G.A. Trosper (eds.), *The construct validation of tests of communicative competence*, TESOL, Washington, pp. 149–63.
- (1996), *Language testing in practice*, Oxford University Press, Oxford.
- (2010), *Language Assessment in Practice*, Oxford University Press, Oxford.

- BACHMAN L.F., SAVIGNON S.J. (1986), *The evaluation of communicative language proficiency: A critique of the ACTFL oral interview*, in «The Modern Language Journal» 70, pp. 380–90.
- BADDELEY A. (2007), *Working Memory, Thought and Action*, Oxford University Press, Oxford.
- BANERJEE J., FRANCESCHINA F., SMITH A.M. (2007), *Documenting features of written language production typical at different IELTS band score levels*, in «IELTS Research Report» 7.
- BARDOVI-HARLIG K. (1992), *A second look at T-unit analysis: Reconsidering the sentence*, in «TESOL Quarterly» 26, pp. 390–5.
- BARDOVI-HARLIG K., BOFMAN T. (1989), *Attainment of syntactic and morphological accuracy by advanced language learners*, in «Studies in Second Language Acquisition» 11, pp. 17–34.
- BARRON A., SCHNEIDER K.P. (2009), *Variational pragmatics: Studying the impact of social factors on language use in interaction*, in «Intercultural Pragmatics» 6, 4, pp. 425–42.
- BARTNING I., MARTIN M., VEDDER I. (eds.) (2010), *Communicative Proficiency and Linguistic Development: intersections between SLA and language testing research*, Eurosla monographs.
- BAYLEY C.J.N. (1973), *Variation and Linguistic Theory*, Center for Applied Linguistic, Washington D.C.
- BAYLEY R. (1991), *Variation theory and second language learning: Linguistic and social constraints on interlanguage tense marking*, PhD. Dissertation, Stanford University, citato in Preston, Bayley (1996).
- (1994), *Interlanguage variation and the quantitative paradigm: Past-tense marking in Chinese-English*, in E. Tarone, S. Gass, A. Cohen (eds.), *Research methodology in Second Language Acquisition*, Lawrence Erlbaum, Hillsdale, NJ, pp. 157–81.

- (1996), *Competing constraints on variation in the speech of adult Chinese learners of English* (eds.), in R. Bayley, D. Preston, *Second Language Acquisition and Linguistic Variation*, John Benjamins Publishing Company, Amsterdam, pp. 96–120.
- (2007), *Second Language Acquisition: A variationist perspective*, in R. Bayley, C. Lucas (eds.), *Sociolinguistic Variation. Theories, Methods, and Applications*, Oxford University Press, Oxford, pp. 133–44.
- BAYLEY R., LANGMAN J. (2004), *Variation in the group and the individual: Evidence from second language acquisition*, in «IRAL International Review of Applied Linguistics» 42, pp. 303–18.
- BAYLEY R., LUCAS C. (2007) (eds.), *Sociolinguistic Variation. Theories, Methods, and Applications*, Oxford University Press, Oxford.
- BAYLEY R., PRESTON D. (1996) (eds.), *Second Language Acquisition and Linguistic Variation*, John Benjamins Publishing Company, Amsterdam.
- (2009), *Variationist linguistics and second language acquisition*, in W.C. Ritchie, T.K. Bathia, *The new handbook of second language acquisition*, Emerald Group, London, pp. 89–113.
- BAYLEY R., TARONE E. (2012), *Variationist perspective*, in S. Gass, A. Mackey, *The Routledge Handbook of Second Language Acquisition*, Routledge, London/New York.
- BAZZANELLA C. (1991), *Tratti substandard nel parlato colloquiale*, in C. Lavinio, A. Sobrero, *La lingua degli studenti universitari*, La Nuova Italia, Firenze, pp. 123–63.
- (1994), *Le facce del parlare. Un approccio pragmatico all'italiano parlato*, La Nuova Italia, Firenze.
- (2002) (a cura di), *Sul dialogo. Contesti e forme di interazione verbale*, Guerrini e Associati, Milano.

- BEATTIE G. (1980), *The role of language production in the organisation of behaviour in face-to-face interaction*, in B. Butterworth (ed.), *Language production*, Vol. I, Academic Press, London, pp. 69–109.
- BEEBE L.M. (1981), *Social and Situational Factors Affecting the Communicative Strategy of Dialect Code-Switching*, in «International Journal of the Sociology of Language» 31, pp. 139–49.
- (1985), *Reservations about the Labovian paradigm of style shifting and its extension to the study of interlanguage*, ERIC document.
- (1987), *Sociolinguistic variation and style shifting in second language acquisition*, in «Language Learning», 2, 1980, 433–48, reprinted in G. Ioup, S. Weinberger (eds.), *Interlanguage Phonology*, Newbury House Publishers.
- BEEBE L.M., ZUENGLER J. (1983), *Accommodation theory: An explanation for style shifting in second language dialects*, in N. Wolfson, E. Judd (eds.), *Sociolinguistics and language acquisition*, Newbury, Rowley MA, pp. 195–213.
- BERCELLI F., PALLOTTI G. (2002), *Conversazioni telefoniche*, in C. Bazzanella (a cura di), *Sul dialogo. Contesti e forme di interazione verbale*, Guerrini e Associati, Milano, pp. 177–192.
- BERMAN R. (2008), *The psycholinguistics of developing text construction*, in «Journal of Child Language» 35, pp. 735–71.
- BERMAN R., NIR-SAGIV B. (2007), *Comparing narrative and expository text construction across adolescence: A developmental paradox*, in «Discourse Processes» 43, pp. 79–120.
- BERRUTO G. (1986), *Un tratto sintattico dell'italiano parlato: il c'è presentativo*, in K. Lichem, E. Mara, S. Knaller (a cura di), *Parallela 2. Aspetti della sintassi dell'italiano contemporaneo*, Tübingen, Narr, 61–73.
- (2012), *Sociolinguistica dell'italiano contemporaneo*, Carocci editore, Roma.

- BERWICK R., ROSS S. (1996), *Cross-cultural pragmatics in oral proficiency interview strategies*, in M. Milanovic, N. Saville (eds.), *Performance testing, cognition and assessment*, University of Cambridge Local Examinations Syndacate and Cambridge University Press, Cambridge, pp. 34–54.
- BETTONI C. (2006), *Usare un'altra lingua. Guida alla pragmatica interculturale*, Laterza, Bari–Roma.
- BIBER D., CONRAD S., REPPEN R. (1998), *Corpus Linguistics: Investigating Language Structure and Use*, Cambridge University Press, Cambridge.
- BIBER D., GRAY B., POONPON K. (2011), *Should we use characteristics of conversation to measure grammatical complexity in L2 writing development?*, in «TESOL Quarterly» 45, pp. 5–35.
- BICKERTON D. (1971), *Inherent variability and variable rules*, in «Foundations of Language» 7, pp. 457–92.
- Bilingualism: Language and Cognition* (2007), «Special Issue» 10, 1.
- BOSISIO C. (2005), *Dagli approcci tradizionali al Quadro comune europeo di riferimento. Riflessioni glottodidattiche e applicazioni per l'insegnante di italiano L2*, ISU Università Cattolica, Milano.
- BOXER D. (2010), *Complaints: How to gripe and establish rapport*, in A. Martínez–Flor, E. Usó–Juan (a cura di), *Speech act performance. Theoretical, empirical and methodological issues*, John Benjamins, Amsterdam, pp. 163–178.
- BREEN M. (1987), *Contemporary paradigms in syllabus design (Parts 1 and 2)*, in «Language Teaching» 20, pp. 91–92 e pp. 157–74.
- BRINDLEY G. (1998), *Describing language development? Rating scales and SLA*, in L.F. Bachman, A.D. Cohen (eds.), *Interfaces between second language acquisition and language testing research*, Cambridge University Press, Cambridge, pp. 112–40.



- (2013), *Task-based assessment*, in C. Chapelle (ed.), *The encyclopedia of applied linguistics*, Blackwell, London.
- BROCK C. (1986), *The effect of referential questions on ESL classroom discourse*, in «TESOL Quarterly» 20/1, pp. 47–59.
- BROWN A. (1995), *The effect of rater variables in the development of an occupation-specific language performance test*, in «Language Testing» 12, pp. 1–15.
- (2003), *Interviewer variation and the co-construction of speaking proficiency*, in «Language Testing» 20, pp. 1–25.
- (2005), *Interviewer variability in language proficiency interviews*, Peter Lang, Frankfurt.
- BROWN J.D., HUDSON T.D., NORRIS J.M., BONK W. (2002), *An investigation of second language task-based performance assessments*, University of Hawaii Press, Honolulu, HI.
- BROWN R. (1973), *A First Language: the Early Stages*, Harvard University Press, Cambridge, Mass.
- BROWN R., BELLUGI-KLIMA U. (1964), *Three processes in the child's acquisition of syntax*, in «Harvard Educational Review» 34, pp. 133–151.
- BRUMFIT C. (1984), *Communicative Methodology in Language Teaching*, Cambridge University Press, Cambridge.
- (1988), *Applied linguistics and communicative language teaching*, in W. Grabe (ed.), *Annual Review of Applied Linguistics: Communicative Language Teaching*, Cambridge University Press, New York, 1988.
- BRUMFIT C., JOHNSON K. (1979), *The Communicative Approach to Language Teaching*, Cambridge University Press, New York.
- BRUNER J.S. (1973), *The relevance of education*, Norton, New York.
- BUCK G. (1991), *The testing of listening comprehension: an introspective study*, in «Language Testing» 8, pp. 67–91.

- BULTÉ B., HOUSEN A. (2012), *Defining and operationalizing L2 complexity*, in A. Housen, F. Kuiken, I. Vedder (eds.), *Dimensions of L2 Performance and proficiency: Complexity, Accuracy and Fluency in SLA*, John Benjamins Publishing Company, Amsterdam, pp. 21–46.
- (2014), *Conceptualizing and measuring short-term changes in L2 writing complexity*, in «Journal of Second Language Writing» 26, 4, pp. 42–65.
- BYGATE M. (1996), *Effects of task repetition: appraising the developing language of learners*, in J. Willis, D. Willis (eds.), *Challenge and Change in Language Teaching*, Heinemann, London, pp. 136–46.
- (1999), *Task as context for framing, reframing and unframing language*, in «System» 27, pp. 33–48.
- (2001), *Effects of task repetition on the structure and control of oral language*, in M. Bygate, P. Skehan, M. Swain (eds.), *Researching Pedagogic tasks, Second Language Learning, Teaching and Testing*, Longman, London, pp. 23–48.
- (2009), *Effects of task repetition on the structure and control of oral language*, in K. Van den Branden, M. Bygate, J. Norris (eds.), *Task-based language teaching: A reader*, John Benjamins Publishing Company, Philadelphia, pp. 249–74.
- (2015) (ed.), *Domains and Directions in the Development of TBLT. A decade of plenaries from the international conference*, John Benjamins Publishing Company, Amsterdam.
- BYGATE M., DÖRNYEI Z., CSÖLLE A., KIRÁLY K., KORMOS J., NEMETH N. (1998), *Perspective from a task-based instruction project*, Colloquium, 32nd Annual TESOL Convention, Seattle.
- BYGATE M., SKEHAN P., SWAIN M. (2001) (eds.), *Researching Pedagogic Tasks: Second Language Learning, Teaching and Testing*, Longman, London.

- BYRNE D., CALLAGHAN G. (2014), *Complexity Theory and the Social Sciences: The State of the Art*, Routledge, Abingdon.
- CAMERON L., DEIGNAN A. (2006), *The emergence of metaphor in discourse*, in «Applied Linguistics», Special Issue 27, pp. 671–90.
- CANALE M. (1983), *On some dimensions of language proficiency*, in J.W. Oller (ed.), *Issues in language testing research*, Newbury House, Rowley, Mass., pp. 333–42.
- CANALE M., SWAIN M. (1980), *Theoretical basis of communicative approaches to second language teaching and testing*, in «Applied Linguistics» 1, pp. 1–47.
- (1981), *A theoretical framework for communicative competence*, in A.S. Palmer, J.M. Groot, G.A. Trostler (eds.), *The construct validation of tests of communicative competence*, TESOL, Washington, pp. 31–6.
- CANDLIN C.N. (1987), *Towards task-based language learning*, in C.N. Candlin, D.F. Murphy, *Language Learning Tasks*, in «Lancaster Practical Papers in English Language Education» vol. 7, Prentice Hall, Cambridge, pp. 5–22.
- CAREY M.D., MANNELL R.H., DUNN P.K. (2011), *Does a rater's familiarity with a candidate's pronunciation affect the rating of oral proficiency interviews?*, in «Language Testing» 28, 2, pp. 201–19.
- CARROLL J.B. (1993), *Human Cognitive Abilities*, Cambridge, New York.
- CARROLL J.B., WEST R. (1989), *ESU Framework. Performance Scales for English Language Examinations*, Longman, Harlow.
- CASSANDRO M., MAGGINI M. (2004), *Osservazione e valutazione di apprendenti stranieri in relazione ai livelli comuni di riferimento del Quadro comune europeo*, in E. Jafrancesco (a cura di), *Le tendenze innovative del quadro comune europeo di riferimento per le lingue ed il portfolio*. «Atti del XII convegno nazionale ILSA», Edilingua, Firenze, 64–107.

- CHAFE W. (1988), *Linking Intonation Units in Spoken English*, in Hai-man J., Thompson S.A. (eds.), *Clause Combining in Grammar and Discourse*, John Benjamins Publishing Company, Amsterdam, pp. 1–27.
- (1994), *Discourse, consciousness, and time: The flow and displacement of conscious experience in speaking and writing*, University of Chicago Press, Chicago.
- CHALHOUB-DEVILLE M. (1995a), *Deriving oral assessment scales across different tests and rater groups*, in «Language Testing» 12, pp. 16–33.
- (1995b), *A contextualized approach to describing oral language proficiency*, in «Language Learning» 45, pp. 251–81.
- (1996), *Performance assessment and the components of the oral construct across different tests and rater groups*, in M. Milanovic, N. Saville (eds.), *Performance testing, cognition and assessment*, University of Cambridge Local Examinations Syndicate and Cambridge University Press, Cambridge, pp. 55–73.
- (2001), *Task-based assessments: Characteristics and validity evidence*, in M. Bygate, P. Skehan, M. Swain (eds.), *Researching pedagogic tasks: second language learning, teaching, and testing*, Pearson Education, New York, pp. 210–28.
- CHAMBERS J.K., SCHILLING N. (2013) (eds.), *The Handbook of Language Variation and Change*, 2nd edition, Wiley-Blackwell.
- CHENG L., DE LUCA C. (2011), *Voices from Test-Takers: Further Evidence for Language Assessment Validation and use*, in «Educational Assessment» 16, 2, pp. 104–22.
- CHINI M. (1998), *La subordinazione in testi narrativi di apprendenti tedescofoni: forma e funzione*, in «Linguistica e filologia» 7, pp. 121–59.
- (2002), *Fra sintassi e pragmatica in italiano L2: gli ordini marcati in testi di apprendenti tedescofoni*, in P. Cordin, R. Franceschini, G.

- Held (a cura di), *Parallela 8. Lingue di confine, confini di fenomeni linguistici/Grenzsprachen.Grenzen von linguistischen Phänomenen*, Bulzoni, Roma, pp. 311–333.
- (2003), *Le phénomène de la jonction interpropositionnelle dans la narration en italien L2: agrégation et intégration*, in «AILE Acquisition et Interaction en Langue Etrangère» 17, pp. 71–106.
- CHOMSKY N. (1965), *Aspects of the Theory of Syntax*, The MIT Press, Cambridge (Mass.).
- (1975), *Reflections on Language*, Temple Smith, London.
- (1980), *Rules and Representations*, Blackwell, Oxford.
- CLAHSEN H. (1980), *Psycholinguistic aspects of L2 acquisition*, in S.W. Felix (ed.), *Second language development, trend and issue*, Tübingen, pp. 57–79.
- CLAHSEN H., MUYSKEN P. (1986), *The availability of UG to adult and child learners. A study of the acquisition of German word order*, in «Second Language Research» 2, pp. 93–119.
- (1989), *The UG paradox in L2 acquisition*, in «Second Language Research» 5, pp. 1–29.
- COHEN A.D. (2000), *Exploring strategies in test-taking: Fine-tuning verbal reports from respondents*, in G. Ekbatani, H. Pierson (eds.), *Learner-directed assessment in ESL*, Lawrence Erlbaum, Mahwah, NJ, pp. 127–150.
- CONNOR-LINTON J., POLIO C. (2014), *Comparing perspectives on L2 writing: Multiple analyses of a common corpus*, in «Journal of Second Language Writing» 26, 4, pp. 1–9.
- CONSIGLIO D'EUROPA (1988), *Final report of the project report of project 12: learning and teaching modern languages for communication*, Council for Cultural Co-operation, Strasbourg, CC-GP12, p. 11.

- (1992), *Transparency and coherence in language learning in Europe. Objectives, evaluation, certification*, Report on the Rüsclikon Symposium.
- (1996), *Modern languages: learning, teaching, assessment: a Common European Framework of Reference*, Draft 2 of a framework proposal, Council for Cultural Co-operation, Education Committee, Strasbourg.
- (2001), *Quadro comune europeo di riferimento per le lingue*, La Nuova Italia, Scandicci, Firenze.
- (2002), *Relating language examinations to the Common European Framework of Reference for Languages: learning, teaching, assessment* (CEF), DGIV/EDU/LANG, 15, Language Policy Division, Strasbourg.
- (2003), *Relating language examinations to the Common European Framework of Reference for Languages: learning, teaching, assessment* (CEF), Manual: Preliminary Pilot Version, DGIV/EDU/LANG 2003, 5, Language Policy Division, Strasbourg.
- (2005), *Reference supplement to the preliminary version of the manual for relating examinations to the Common European Framework of Reference for Languages: learning, teaching, assessment*, DGIV/EDU/LANG 2005, 13, Language Policy Division, Strasbourg.
- (2007), *Tools for Language Teaching: Common European Framework of Reference and Portfolios*, Francis Goullier. Les éditions Didier, Paris / Council of Europe, Strasbourg.
- (2018), *Common European Framework of Reference for Languages: Learning, Teaching, Assessment. Companion volume with new descriptors*, [www.coe.int/lang-cefr](http://www.coe.int/lang-cefr).
- COOPER T.C. (1976), *Measuring written syntactic patterns of second language learners of German*, in «The Journal of Educational Research» 69, pp. 176–83.

- CORDER S.P. (1981), *Error analysis and interlanguage*, Oxford University Press, Oxford.
- CRESTI E. (2005a), *Enunciato e frase: teoria e verifiche empiriche*, in M. Biffi, O. Calabrese, L. Salibra (a cura di), *Italia Linguistica: discorsi di scritto e di parlato. Scritti in onore di Giovanni Nencioni*, Prolagon, Siena, pp. 249–60.
- (2005b), *Per una nuova classificazione dell'illocuzione*, in E. Burr (a cura di), *Tradizione e innovazione*, Atti del VI convergno della Società di Linguistica e Filologia Italiana (SILFI), Cesati, Firenze, pp. 233–46.
- (2010), *La Stanza: un'unità di costruzione testuale del parlato*, in A. Ferrari (ed.), *Sintassi storica e sincronica dell'italiano. Subordinazione, coordinazione e giustapposizione. Atti del X Congresso della Società Internazionale di Linguistica e Filologia Italiana*, (Basilea, 30 giugno–3 luglio 2008), Cesati Editore, Firenze, pp. 713–32.
- (2012), *L'unité de suffixe: identification et interprétation des unités de la langue parlé*, in Caddéo, Sandrine et al. (eds.), *Penser les langues avec Claire Blanche-Benveniste*, Presses Universitaires de Provence, Aix-en-Provence, pp. 201–13.
- (2016), *Dalla struttura informativa (alla prosodia) alla sintassi: dati sulla subordinazione nell'italiano parlato*, in A. Elia, C. Iacobini, M. Voghera (eds.), *Livelli di Analisi e Fenomeni di Interfaccia. Atti del LXVII Congresso Internazionale SLI*, Bulzoni, Roma, pp. 53–73.
- (2018), *The illocution–prosody relationship and the Information Pattern in spontaneous speech according to the Language into Act Theory (L–Act)*, in «Linguistick online» 88, 1, p. 18.
- CRESTI E., GRAMIGNI P. (2004), *Per una linguistica corpus based dell'italiano parlato: le unità di riferimento*, in F.A. Leoni, F. Cutugno,

- M. Pettorino, R. Savy (a cura di), *Il parlato italiano*, Atti del Convegno Nazionale, D'Auria Editore, Napoli, pp. 1–23.
- CRESTI E., MONEGLIA M. (a cura di) (2005), *C-ORAL-ROM, Integrated Reference Corpora for Spoken Romance Language*, full edition in 9DVDs, ELRA, Paris.
- CRONBACH L.J. (1971), *Test validation*, in R.L. Thorndike (ed.), *Educational Measurement*. 2nd edition, American Council on Education, pp. 443–507.
- CRONBACH L.J., MEEHL P.E. (1955), *Construct validity in psychological tests*, in «Psychological Bulletin» 52, pp. 281–302.
- CROOKES G. (1986), *Task classification: A cross-disciplinary review*, Technical Report 4, Center for Second Language Classroom Research, University of Hawaii, Honolulu.
- (1989), *Planning and interlanguage variation*, in «Studies in Second Language Acquisition» 11, pp. 367–83.
- (1990), *The utterance and other basic units for second language discourse analysis*, in «Applied Linguistics» 11, pp. 183–199.
- CROOKES G., GASS S. (eds.) (1993), *Tasks in language learning: Integrating theory and practice*, Multilingual Matters, Clevedon, Avon.
- CRYSTAL D., DEVI D. (1975), citato in Ellis, Barkhuizen (2005).
- DAVIDSON F., FULCHER G. (2007), *The Common European Framework of Reference (CEFR) and the design of language tests: A matter of effect*, in «Language Teaching» 40, 3, pp. 231–41.
- DAVIES A. (1990), *Principles of language testing*, Basil Blackwell, London.
- (2003), *Three heresis of language testing research*, in «Language Testing» 20, pp. 355–68.
- DE BOT K. (2008), *Introduction: second language development as a dynamic process*, in «Modern Language Journal» 92, 2, pp. 166–78.



- DE BOT K., LOWIE W., VERSPOOR M. (2007), *A dynamic systems theory to second language acquisition*, in «Bilingualism: Language and Cognition» 10, pp. 7–21.
- DE CAMP D. (1973), *Implicational scales and sociolinguistic linearity*, in «Linguistics» 73, pp. 30–43.
- DE CESARE A.M., GARASSINO D., ROCÌO AGAR M., ALBOM A., CIMMINO D. (2016), *Sintassi marcata dell'italiano dell'uso medio in prospettiva contrastiva con il francese, lo spagnolo, il tedesco e l'inglese. Uno studio basato sulla scrittura dei quotidiani on line*, Peter Lang, Berlino.
- DE JONG J. (1992), *Assessment of language proficiency in the perspective of the 21st century*, in «AILA Review» 9, pp. 39–45.
- DE JONG N., STEINEL M., FLORIJN A., SCHOONEN R., HULSTIJN J. (2012), *Linguistic skills and speaking fluency in a second language*, in «Applied Psycholinguistics» 34, pp. 893–916.
- «Developmental Review» (2005), Special Issue 3–4, *Development as Self-organization: New Approaches to the Psychology and Neurobiology of Development*, edited by Mark L. Howe, Marc D. Lewis.
- DEWAELE J.M. (1992), *L'omission du ne dans deux styles oraux d'interlangue française*, in «Interface: Journal of Applied Linguistics» 7, pp. 3–17.
- (2007), *Diachronic and/or synchronic variation? The acquisition of sociolinguistic competence in L2 French*, in D. Ayoun (ed.) *Handbook of French Applied Linguistic*, John Benjamins Publishing Company, Amsterdam, pp. 208–36.
- (2008), *Variation in advanced interlanguage: The effect of proficiency on style choice in Pragmatics and Corpus Linguistics*, in J. Romero (ed.), *Pragmatics and Corpus linguistics*, Mouton De Gruyter, Berlin, pp. 65–90.

- DEWAELE J.M., REGAN V. (2002), *Maîtriser la norme sociolinguistique en interlangue français: le cas de l'omission variable*, in «Journal of French Language Studies» 12, pp. 123–48.
- DICKERSON L.J. (1975), *The learners' interlanguage as a system of variable rules*, in «TESOL Quarterly» 9, pp. 401–7.
- DÍEZ-BEDMAR M.B. (2017), *Fine-tuning descriptors for CEFR B1 level: insights from learner corpora*, in «ELT Journal» 72, 2, pp. 199–209.
- DONALDSON B. (2017), *Negation in near-native French: Variation and Sociolinguistic Competence*, in «Language Learning» 67, 1, pp. 141–70.
- DOUGLAS D. (2001), *Performance consistency in second language acquisition and language testing research: A conceptual gap*, in «Second Language Research» 17, pp. 442–56.
- ECKMAN F.R. (1994), *The Competence–Performance Issue in Second Language Acquisition Theory: A Debate*, in E. Tarone, S.M. Gass, A.D. Cohen (1994), *Research Methodology in Second Language Acquisition*, Lawrence Erlbaum, Hillsdale, pp. 3–15.
- ELDER C., IWASHITA N. (2005), *Planning for test performance: does it make a difference?*, in R. Ellis (ed.), *Planning and Task Performance in a Second Language*, John Benjamins Publishing Company, Amsterdam, pp. 219–37.
- ELLIS N.C. (1996), *Sequencing in SLA: Phonological memory, chunking and points of order*, in «Studies in Second Language Acquisition» 18, pp. 91–126.
- (1998), *Emergentism, connectionism and language learning*, in «Language Learning» 48, pp. 631–64.
- (2006a), *Language acquisition as rational contingency learning*, in «Applied Linguistics» 27, pp. 1–24.
- (2006b), *Selective attention and transfer phenomena in SLA: Contingency, cue competition, salience, interference, overshadowing*,

- blocking, and perceptual learning*, in «Applied Linguistics» 27, pp. 1–31.
- (2006c), *Cognitive perspectives on SLA: The Associative Cognitive CREED*, in «AILA Review» 19, pp. 100–21.
- ELLIS N.C., LARSEN-FREEMAN D. (2006), *Language emergence: Implications for applied linguistics (introduction to the special issue)*, in «Applied Linguistics» 27, pp. 558–89.
- (2009), *Constructing a Second Language: Analyses and Computational Simulations of the Emergence of Linguistic Constructions From Usage*, in «Language Learning» 59, 1, pp. 90–125.
- ELLIS N.C., WULFF S. (2019), *Cognitive approaches to L2 acquisition*, in J.W. Schwieter, A. Benati (eds.), *The Cambridge Handbook of Language Learning*, Cambridge University Press, Cambridge, pp. 41–61.
- ELLIS R. (1985a), *Sources of variability in interlanguage*, in «Applied Linguistics» 6, pp. 118–31.
- (1985b), *Understanding second language acquisition*, Oxford University Press, Oxford.
- (1987a), *Interlanguage variability in narrative discourse: Style shifting in the use of past tense*, in «Studies in Second Language Acquisition» 9, pp. 1–20.
- (1987b), *Contextual variability in second language acquisition and the relevancy of language teaching*, in R. Ellis (ed.), *Second Language Acquisition in Context*, Prentice–Hall International, London, pp. 179–94.
- (1994), *The study of second language acquisition*, Second edition, Cambridge University Press, Cambridge.
- (1996), *SLA and language pedagogy*, in «Studies in Second Language Acquisition» 19, pp. 69–92.

- (1999), *Item versus system learning: Explaining free variation*, «Applied Linguistics» 20, pp. 460–80.
- (2001), *Some thoughts on testing grammar: An SLA perspective*, in C. Elder, A. Brown, N. Iwashita, E. Grove, K. Hill, T. Lumley, T. McNamara, K. O’Loughlin (eds.), *Experimenting with uncertainty. Essays in honour of Alan Davies*, Cambridge University Press, Cambridge, 251–63.
- (2003), *Task-based Language Learning and Teaching*, Oxford University Press, Oxford.
- (ed.) (2005), *Planning and task performance in a second language*, John Benjamins Publishing Company, Philadelphia.
- (2009), *The differential effects of three types of task planning on the fluency, complexity and accuracy in L2 oral production*, in «Applied Linguistics» 19, pp. pp. 221–246.
- (2012), *Language Teaching Research and Language Pedagogy*, Wiley-Blackwell.
- (2018a), *Taking the critics to task: The case for task-based teaching*, in I. Walker, D. Chan, K.G. Nagami, M. Bourguignon (eds.), *New Perspectives on the Development of Communicative Related Competence in Foreign Language Education*, De Gruyter Mouton, Berlin, pp. 23–41.
- (2018b), *Reflections on task-based language teaching*, Multilingual Matters, Cleveland.
- ELLIS R., BARKHUIZEN G. (2005), *Analysing learner language*, Oxford University Press, Oxford.
- ELLIS R., TANAKA Y., YAMAZAKI A. (1994), *Classroom interaction, comprehension, and the acquisition of L2 word meanings*, in «Language Learning» 44, pp. 449–491.

- ELLIS R., YUAN F. (2005), *The effects of careful within-task planning on oral and written task performance* in R. Ellis (ed.), *Planning and task performance in a second language*, John Benjamins Publishing Company, Amsterdam, pp. 167–92.
- FASOLD R.W., PRESTON D.R. (2007), *The psycholinguistic unity of inherent variability: Old Occam whips out his razor*, in R. Bayley, C. Lucas (eds.), *Sociolinguistic Variation. Theories, Methods, and Applications*, Oxford University Press, Oxford, pp. 49–69.
- FELE G. (2007), *L'analisi della conversazione*, il Mulino, Bologna.
- FÉLIX-BRASDEFER J.C., KOIKE D. (eds.) (2012), *Pragmatic Variation in First and Second Language Contexts*, John Benjamins Publishing Company, Amsterdam.
- FERRARI S. (in preparazione), *L'AS-unit come unità di produzione per l'italiano parlato in L1 e in L2. Problemi di applicazione e possibili soluzioni*.
- FERRARI S., NUZZO E. (2010), *La valutazione delle competenze orali in italiano L2. Una verifica sperimentale dei criteri suggeriti dal Quadro Comune Europeo*, in «Atti del XV Convegno Nazionale Giscel», Milano, 6–8 marzo 2008, pp. 279–293.
- FIGUERAS N. (2012), *The impact of the CEFR*, in «ELT Journal» 66, 4, pp. 477–85.
- FIGUERAS N., NORTH B., TAKALA S., VERHELST N., VAN AVERMAET P. (2005), *Relating examinations to the Common European Framework: A manual*, in «Language Testing» 22, pp. 261–79.
- FILLMORE C.J. (1979), *On fluency*, in C. Fillmore, D. Kempler, W.S.Y. Wang (eds.), *Individual differences in language ability and language behaviour*, Academic Press, New York, pp. 85–101.
- FORD C.E., FOX B.A., THOMPSON S.A. (1996), *Practices in the construction of turns: The "TCU" revisited*, in «Pragmatics» 6, pp. 427–54.

- FORD C., FOX B.A., THOMPSON S.A. (eds.) (2002), *The language of turn and sequence*, Oxford University Press, Oxford.
- FOSTER P., SKEHAN P. (1996), *The influence of planning on performance in task-based learning*, in «Studies in Second Language Acquisition» 18, pp. 299–324.
- (1997), *Modifying the task: the effects of surprise, time and planning type on task based foreign language instruction*, in Thames Valley University Working Papers in English Language Teaching 4, pp. 38–50.
- (1999), *The influence of source of planning and focus of planning on task-based performance*, in «Language Teaching Research» 3, pp. 215–47.
- FOSTER P., WIGGLEWORTH G. (2016), *Capturing Accuracy in Second Language Performance: The case for a weighted clause ratio*, in «Annual Review of Applied Linguistics» 36, pp. 98–116.
- FOSTER P., WIGGLESWORTH G., TONKYN A. (2000), *Measuring spoken language: a unit for all reasons*, in «Applied Linguistics» 21, pp. 354–75.
- FREED B. (2000), *Is fluency in the eyes (and ears) of the beholder?*, in H. Rigggenbach (ed.), «Perspectives on Fluency», The University of Michigan Press, Ann Arbor, pp. 243–65.
- FREEDLE R., KOSTIN I. (1993), *The prediction of TOEFL reading item difficulty: implications for construct validity*, in «Language Testing» 10, pp. 133–70.
- FULCHER G. (2004), *Deluded by Artifices? The Common European Framework and Harmonization*, in «Language Assessment Quarterly» 1, 4, pp. 253–266.
- FULCHER G., MÁRQUEZ REITER R. (2003), *Task difficulty in speaking tests*, in «Language Testing» 20, 3, pp. 321–344.
- GASS S.M., MACKEY A. (2007), *Data Elicitation for Second and Foreign Language Research*, Lawrence Erlbaum Associates, Mahwah, NJ.

- GASS S.M., MACKEY A., FERNANDEZ M., ALVAREZ-TORRES M. (1999), *The effects of task repetition on linguistic output*, in «Language Learning» 49, pp. 549–80.
- GEDDES M., STURTRIDGE G. (1978), *Listening links*, Heinemann, London.
- GIACALONE RAMAT A. (1999a), *Le strategie di collegamento tra proposizioni nell'italiano di germanofoni*, in N. Dittmar, A. Giacalone Ramat (eds.), *Grammatik und Diskurs/Grammatica e discorso*, Tübingen, Stauffenberg.
- (1999b), *Functional typology and strategies of clause connection in second-language acquisition*, in «Linguistics» 37, pp. 519–48.
- GILABERT R. (2007), *The simultaneous manipulation along the planning time and +/- Here-and-Now dimensions: effects on oral L2 production*, in P. Garcia Mayo (ed.), *Investigating tasks in formal language learning*, Multilingual Matters, Clevedon, pp. 44–68.
- GILES H. (ed.) (1977), *Language, ethnicity and intergroup relations*, American Press, London–New York.
- GILES H., GASIOREK J., SOLIZ J. (2015), *Recent developments in Communication Accomodation Theory: Innovative contexts and application*, in «Language and Communication» 41, Special Issue.
- GIORDANO R., VOGHERA M. (2009), *Frase senza verbo: il contributo della prosodia*, in A. FERRARI (a cura di), *Sintassi storica e sincronica dell'italiano. Subordinazione, coordinazione, giustapposizione*, Atti del X Congresso della Società Internazionale di Linguistica e Filologia Italiana, vol. II, Cesati editore, Firenze, pp. 1005–1024.
- GOLDSCHNEIDER J., DEKEYSER R. (2001), *Explaining the “Natural Order of L2 Morpheme Acquisition” in English: A meta-analysis of multiple determinants*, in «Language Learning» 51, pp. 1–50.

- GRABOWSKI J. (2005), *Speaking, writing, and memory span performance: Replicating the Bourdin and Fayol results on cognitive load in German children and adults*, in L. Allal, J. Dolz (eds.), *Proceedings Writing 2004*, Adcom Productions (CD-ROM), Geneva.
- (2007), *The writing superiority effect in the verbal recall of knowledge: Sources and determinants*, in G. Rijlaarsdam (Series ed.), M. Torrance, L. van Waes, D. Galbraith (Volume eds.), *Writing and cognition: Research and applications*, Elsevier, Amsterdam, pp. 165–179.
- GRAFFI G. (1994), *Sintassi*, il Mulino, Bologna.
- (2001), *200 Years of Syntax. A Critical Survey*, John Benjamins Publishing Company, Amsterdam.
- GRANFELDT J. (2008), *Speaking and writing in L2 French: Exploring effects on fluency, complexity and accuracy*, in S. Van Daele, A. Housen, F. Kuiken, M. Pierrard, I. Vedder (eds.), *Complexity, accuracy and fluency in second language use, learning and teaching*, KVAB Universa Press, Wetteren, pp. 87–98.
- GREEN A. (2010), *Requirements for Reference Level Descriptions for English*, in «English Profile Journal» 1, 1.
- GREGG K. (1990), *The Variable Competence Model of second language acquisition and why it isn't*, in «Applied Linguistics» 11, pp. 364–83.
- (2010), *Review Article: Shallow draughts: Larsen–Freeman and Cameron on Complexity*, in «Second Language Research» 26, 4, pp. 549–60.
- GREGO BOLLI G. (2006a), *Il quadro e la valutazione*, in M. Mezzadri (a cura di), *Integrazione linguistica in Europa*, UTET Università, Torino, pp. 220–46.
- (2006b), *Progetti europei. Nuove prospettive sulla scia del Quadro Comune Europeo di Riferimento*, <http://www.cvcl.it/canale.asp?id=198>.



- GUNNARSSON C. (2012), *The development of complexity, accuracy, and fluency in the written production of L2 French*, in A. Housen, F. Kuiken, I. Vedder (eds.), *Dimensions of L2 Performance and proficiency: Complexity, Accuracy and Fluency in SLA*, John Benjamins Publishing Company, Amsterdam, pp. 247–276.
- GYLLSTAD H., GRANFELDT J., BERNARDINI P., KÄLLKVIST M. (2014), *Linguistic correlates to communicative proficiency levels of the CEFR: The case of syntactic complexity in written L2 English, L3 French and L4 Italian*, in «EUROSLA Yearbook» 14, 1.
- HAIMAN J., THOMPSON S.A. (1988), *Clause Combining in Grammar and Discourse*, John Benjamins Publishing Company, Amsterdam.
- HALE G. (1988), *Student major field and text content: Interactive effects on reading comprehension in the Test of English as a Foreign Language*, in «Language Testing» 5, 49–61.
- HALLIDAY M.A.K. (1985), *Spoken and Written Language*, Oxford University Press, Oxford.
- HALLIDAY M., MARTIN J. (1993), *Writing Science: Literacy and Discursive Power*, The Falmer Press, Bristol/London.
- HALLIDAY M., MATTHIESSEN C. (1999), *Construing Experience through Meaning: A Language-Based Approach to Cognition*, Cassel, London/New York.
- HARGREAVES A. (1992), *Cultures of teaching: A focus for change*, in A. Hargreaves, M. Fullan (eds.), *Understanding teacher development*, Teacher College Press, New York, 216–40.
- HARMER J. (1991), *The Practice of English Language Teaching*, Longman, London.
- HARSCH C. (2014), *General Language Proficiency Revisited: Current and Future Issues*, in «Language Assessment Quarterly», 11, 2, pp. 152–169.

- (2018), *How Suitable Is the CEFR for Setting University Entrance Standards?*, in «Language Assessment Quarterly» 15, 1, pp. 102–108.
- HARSH C., MARTIN G. (2012), *Adapting CEF-descriptors for rating purposes: Validation by a combined rater training and scale revision approach*, in «Assessing Writing» 17, 4, pp. 228–50.
- (2013), *Comparing holistic and analytic scoring methods: issues of validity and reliability*, in «Assessment in Education: Principles, Policy, Practice» 20, 3, pp. 281–307.
- HASSELGREEN A. (2013), *Adapting the CEFR for the Classroom Assessment of Young Learners' Writing*, in «Canadian Modern Language Review» 69, 4, pp. 415–35.
- HAWKEY R., BARKER F. (2004), *Developing a common scale for the assessment of writing*, in «Assessing Writing» 9, pp. 122–59.
- HEIL C. (2014), *Learning oriented-assessment: The proficiency dimension*, in «Working Papers in TESOL, Applied Linguistics» 14, 2 pp. 44–6.
- HENNING G. (1983), *Oral proficiency testing: Comparative validities of interview, imitation, and completion methods*, in «Language Learning» 33, pp. 315–32.
- (1990), *Priority issues in the assessment of second language abilities*, in «Foreign Language Annals» 23, pp. 379–84.
- HERDINA P., JESSNER U. (2002), *A dynamic model of multilingualism: Perspectives of change in psycholinguistics*, Multilingual Matters, Clevedon.
- HOEKJE B., LINNELL K. (1994), «Authenticity» in language testing: *Evaluating spoken language tests for international teaching assistants*, in «Studies in Second Language Acquisition» 15, pp. 461–93.
- HORGAN J. (1995), *The new social Darwinist*, in «Scientific America» 273, pp. 174–81.

- HOUSE J. (1966), *Development of pragmatic fluency in English as a foreign language*, in «Studies in Second Language Acquisition» 18, pp. 225–252.
- HOUSEN A., DE CLERCQ B., KUIKEN F., VEDDER I. (2019), *Multiple approaches to complexity in second language research*, in «Second Language Research», volume 35, Special Issue in Linguistic Complexity.
- HOUSEN A., KUIKEN F. (2009), *Complexity, Accuracy and Fluency in Second Language Acquisition*, in «Applied Linguistic» 30, 4, Special Issue, pp. 461–473.
- HOUSEN A., KUIKEN F., VEDDER I. (eds.) (2012), *Dimensions of L2 Performance and proficiency: Complexity, Accuracy and Fluency in SLA*, John Benjamins Publishing Company, Amsterdam.
- HOWARD M. (2004), *On the interactional effect of linguistic constraints on interlanguage variation: The case of past marking*, in «IRAL, International Review of Applied Linguistics» 42, pp. 319–34.
- HOWARD M., LEMÉE I., REGAN V. (2006), *The L2 acquisition of a phonological variable: The case of the /l/ deletion in French*, in «Journal of French Language Studies» 16, 1, pp. 1–24.
- HOWARD M., MOUGEON R., DEWAELE J.M. (2013), *Sociolinguistics and second language acquisition*, Oxford University Press, Oxford.
- HUHTA A., ALANEN R., TARNANEN M., MARTIN M., HIRVELÄ T. (2014), *Assessing learners' writing skills in a SLA study: Validating the rating process across tasks, scales and languages*, in «Language Testing» 31, 3, pp. 307–28.
- HULSTIJN J.H. (1985), *Testing second language proficiency with direct procedures. A response to Ingram*, in K. Hyltenstam, M. Pienemann (eds.), *Modelling and Assessing Second Language Development*, Multilingual Matters, Clevedon, pp. 277–82.

- (2007), *The Shaky Ground Beneath the CEFR: Quantitative and Qualitative Dimensions of Language Proficiency*, in «The Modern Language Journal» 91, 4, pp. 663–67.
- (2010), *Linking L2 proficiency to L2 acquisition: Opportunities and challenges of profiling research*, in I. Bartning, M. Martin, I. Vedder (eds.), *Communicative Proficiency and Linguistic Development: intersections between SLA and language testing research*, Eurosla monographs, pp. 233–37.
- (2011), *Language Proficiency in Native and Nonnative Speakers: An Agenda for Research and Suggestions for Second–Language Assessment*, in «Language Assessment Quarterly» 8, 3, pp. 229–49.
- (2014), *The Common European Framework of Reference for Languages: A challenge for applied linguistics*, in «ITL–International Journal of Applied Linguistics» 165, 1, pp. 3–18.
- HULSTIJN J.H., SCHOONEN R. (2006), *Scientific report of the ESF sponsored exploratory workshop: Bridging the gap between research on second–language acquisition and research on language testing*, University of Amsterdam, Amsterdam Center for Language and Communication, [http://www.esf.org/fileadmin/be\\_user/ew\\_docs/05–208\\_Report.pdf](http://www.esf.org/fileadmin/be_user/ew_docs/05–208_Report.pdf).
- HULSTIJN J.H., SCHOONEN R., ALDERSON J.C. (2010), *Developmental stages in second language acquisition and levels of second language proficiency: Are there links between them?*, in I. Bartning, M. Martin, I. Vedder (eds.), *Communicative Proficiency and Linguistic Development: intersections between SLA and language testing research*, Eurosla monographs, pp. 11–19.
- HULSTIJN J.H., SCHOONEN R., DE JONG N.H., STEINEL M.P., FLORIJN A. (2012), *Linguistic competences of learners of Dutch as a second language at the B1 and B2 levels of speaking proficiency of the Com-*

- mon *European Framework of Reference for Languages (CEFR)*, in «Language Testing» 29, 2, pp. 203–21.
- HUNT K.W. (1965), *Grammatical structures written at three grade levels*, The National Council of Teachers of English, Urbana IL.
- (1970), *Recent measures in syntactic development*, in M. Lester (ed.), *Reading in applied transformational grammar*, Holt. Rinehart, New York, pp. 187–200.
- HYMES D. (1971), *Competence and performance in linguistic theory*, in R. Huxley, E. Ingram (eds.), *Language Acquisition, Models and Methods*, American Press, New York, pp. 3–28.
- (1972), *On communicative competence*, in J.B. Pride, J. Holmes (eds.), *Sociolinguistics*, Penguin, Harmondsworth, pp. 269–93.
- INGRAM D.E. (1985), *Assessing proficiency: An overview of some aspects of testing*, in K. Hyltenstam, M. Pienemann (eds.), *Modelling and assessing second language acquisition*, Multilingual Matters, Clevedon, Avon, pp. 215–76.
- ISAACS T., THOMSON R.I. (2013), *Rater experience, rating scale length, and judgments of L2 pronunciation: Revisiting research conventions*, in «Language Assessment Quarterly» 10, pp. 135–59.
- ISHIKAWA T. (1995), *Objective measurement of low-proficiency EFL narrative writing*, in «Journal of Second Language Writing» 4, pp. 51–70.
- (2007), *The effect of increasing task complexity along along the +/– Here-and-Now dimension*, in P. Garcia Mayo (ed.), *Investigating tasks in formal language learning*, Multilingual Matters, Clevedon, pp. 136–56.
- IWASHITA N., BROWN A., MCNAMARA T., O'HAGAN S. (2007), *Assessed levels of second language proficiency: How distinct?*, in «Applied Linguistics Advance Access» published October, 26, pp. 1–26.

- IWASHITA N., MCNAMARA T., ELDER C. (2001), *Can we predict task difficulty in an oral proficiency test? Exploring the potential of an information processing approach to task design*, in «Language Learning» 51, pp. 401–36.
- JAFRANCESCO E. (a cura di) (2004), *Le tendenze innovative del quadro comune europeo di riferimento per le lingue ed il portfolio*. Atti del XII convegno nazionale ILSA, Edilingua, Firenze.
- JOHNS A.M. (1985), *Summary protocols of under-prepared and adept university students: Replications and distortion of the original*, in «Language Learning» 35, pp. 497–517.
- JOHNS A.M., MAYES P. (1990), *An analysis of summary protocols of university ESL students*, in «Applied Linguistics» 11, pp. 253–272.
- KAFTANDJIEVA F., TAKALA S. (2002), *Council of Europe scales of language proficiency: a validation study*, in ALDERSON (ed.), *Common European Framework of Reference for Languages: Learning, Teaching, Assessment. Case studies*, Council of Europe, Strasbourg, 106–29.
- KANE M. (1992), *An argument-based approach to validity*, in «Psychological Bulletin» 112, pp. 527–35.
- KAWAUCHI C. (2005), *The effects of strategic planning on the oral narratives of learners with low and high intermediate proficiency*, in R. Ellis (ed.), *Planning and Task-Performance in a Second Language*, John Benjamins Publishing Company, pp. 143–64.
- KIM Y., TRACY-VENTURA N. (2011), *Task complexity, language anxiety, and the development of the simple past*, in P. Robinson (ed.), *Second Language Task Complexity. Researching the Cognition Hypothesis of language learning and performance*, John Benjamins Publishing Company, Amsterdam, pp. 287–306.
- (2013), *The role of task repetition in L2 performance development: What needs to be repeated during task-based interaction?*, in «System» 41, 3, pp. 829–40.

- KOBAYASHI M. (2009), *Hitting the mark. How can text organization and response format affect reading test performance*, Peter Lang, Berlin.
- KORMOS J. (2011), *Task complexity and linguistic and discourse features of narrative writing performance*, in «Journal of Second Language Writing» 20, 2, pp. 148–61.
- (2014), *Differences across modalities of performance*, in H. Byrnes, R. Machòn (eds.), *Task-Based Language Learning — Insights from and for L2 Writing*, John Benjamins Publishing Company, Amsterdam, pp. 193–216.
- KORMOS J., DENES M. (2004), *Exploring measures and perceptions of fluency in the speech of second language learners*, in «System» 32, pp. 145–64.
- KORMOS J., TREBITS A. (2011), *Working memory capacity and narrative task performance*, in P. Robinson (ed.), *Second Language Task Complexity. Researching the Cognition Hypothesis of language learning and performance*, John Benjamins Publishing Company, Amsterdam, pp. 267–86.
- (2012), *The role of task complexity, modality and aptitude in narrative task performance*, in «Language Learning» 62, pp. 439–72.
- KROLL B. (1977), *Combining ideas in written and spoken English: a look at subordination and coordination*, in E. Ochs Keenan, T.L. Bennet (eds.), *Discourse across Time and Space. Southern California Occasional Papers in Linguistics 5*, Los Angeles, California, Department of Linguistics, University of South California.
- KUIKEN F., VEDDER I. (2004a), *Il Bed and Breakfast più bello d'Italia. Cognitieve taakcomplexiteit en tekstkwaliteit in Italiaans T2 [Il Bed and Breakfast più bello d'Italia. Cognitive task complexity and text quality in Italian L2]*, in «Incontri» 19, pp. 31–39.
- (2004b), *De relatie tussen cognitieve taak complexiteit en linguïstische performance bij het schrijven in T1 en T2 [The relation between*

- en cognitive task complexity and linguistic performance in writing in L1 and L2*], in «Toegepaste Taalwetenschap in Artikelen» 72, pp. 23–32.
- (2007a), *Task complexity and linguistic complexity in L2 writing: a discussion*, in «IRAL Special Issue on task complexity and instructed SLA», pp. 261–84.
- (2007b), *Cognitive task complexity and linguistic performance in French L2 writing*, in M.P. Garcia Mayo (ed.), *Investigating tasks in formal language learning*, Multilingual Matters, Clevedon, UK, pp. 117–35.
- (2007c), *Task complexity and measures of linguistic performance in L2 writing*, in «IRAL» 45, pp. 261–84.
- (2008), *Cognitive task complexity and written output in Italian and French as a second language*, in «Journal of Second Language Writing» 17, pp. 48–60.
- (2011), *Task Complexity and Linguistic Performance in L2 Writing and Speaking: The Effect of Mode*, in P. Robinson (ed.), *Second Language Task Complexity. Researching the Cognition Hypothesis of language learning and performance*, John Benjamins Publishing Company, Amsterdam, pp. 91–104.
- KUIKEN F., VEDDER I., MOS M. (2005), *Cognitive task complexity and second language writing performance*, in S. Foster–Cohen, M.P. Garcia–Mayo, J. Cenoz (eds.), «Eurosla yearbook», Vol. 5, John Benjamins Publishing Company, Amsterdam, pp. 195–222.
- KUNNAN A.J. (1993), *Approaches to validation in language assessment*, in A.J. Kunnan (ed.), *Validation in language assessment*, Lawrence Erlbaum, Mahwah, pp. 1–18.
- LABOV W. (1970), *The study of language in its social context*, in «Studium Generale» 23, pp. 30–87.



- (2001), *Principles of Linguistic Change*, Vol. II, «Social Factor», Wiley–Blackwell.
- LAMBERT C., KORMOS J. (2014), *Complexity, Accuracy, and Fluency in Task-based L2 Research: Toward More Developmentally Based Measures of Second Language Acquisition*, in «Applied Linguistics» 35, 5, pp. 607–14.
- LANGMAN J., BAYLEY R. (2002), *The Acquisition of Verbal Morphology by Chinese Learners of Hungarian*, in «Language Variation and Change» 14, pp. 55–77.
- Language and Communication* (2015), «Special Issue» 41, H.Giles, J. Gasiorek, Jordan Soliz (eds.).
- Language Learning* (2009), «Special Issue» 59, 1, Ellis N.C., Larsen-Freeman D. (eds.).
- Language Testing* (2002), *Special issue on task-based language testing*.
- Language Testing* (2005), *Special issue on the Common European Framework of Reference*.
- LANTOLF J.P. (2000), *Sociocultural Theory and Second Language Learning*, Oxford University Press, Oxford.
- LANTOLF J.P., FRAWLEY W. (1985), *Oral proficiency testing: a critical analysis*, in «The Modern Language Journal» 69, pp. 337–45.
- (1988), *Proficiency: Understanding the construct*, in «Studies in Second Language Acquisition» 10, pp. 181–96.
- (1992), *Rejecting the OPI — again: a response to Hagen*, in «ADFL Bulletin» 23, pp. 34–7.
- LARSEN-FREEMAN D. (1978), *An ESL index of development*, in «TESOL Quarterly» 12, pp. 439–48.
- (1983), *Assessing global second language proficiency*, in H.W. SELIGER, M. LONG (eds.), *Classroom-oriented research in se-*

- cond language acquisition*, Newbury House, Rowley, MA, pp. 287–304.
- (1997), *Chaos/complexity science and second language acquisition*, in «Applied Linguistics» 18, pp. 141–65.
- (2002), *Language acquisition and language use from a chaos/complexity theory perspective*, in C. Kramsch (ed.), *Language Acquisition and Language Socialization*, Continuum, London.
- (2006), *The emergence of complexity, fluency and accuracy in the oral and written production of five chinese learners of English*, in «Applied Linguistics», Special Issue 27, pp. 590–619.
- (2007), *On the complementarity of Chaos/Complexity Theory and Dynamic Systems Theory in understanding the second language acquisition process*, in «Bilingualism: Language and Cognition» 10, pp. 35–7.
- (2009), *Adjusting expectations: The study of Complexity, Accuracy and Fluency in Second Language Acquisition*, in «Applied Linguistics» 30, 4, Special Issue, pp. 579–589.
- (2011a), *A Complexity Theory Approach to Second Language Development/Acquisition*, in D. Atkinson, *Alternative approaches to second language acquisition*, Routledge, London/New York, pp. 48–72.
- (2011b), *Complexity Theory*, in S. Gass, A. Mackey, *The Routledge Handbook of Second Language Acquisition*, Routledge, London/New York, pp. 73–87.
- LARSEN-FREEMAN D., CAMERON L. (2008), *Complex Systems and Applied Linguistics*, Oxford University Press, Oxford.
- LAUFER B. (2001), *Quantitative evaluation of vocabulary*, in C. Elder, A. Brown, N. Iwashita, E. Grove, K. Hill, T. Lumley, T. McNamara, K. O’Loughlin (eds.), *Experimenting with uncertainty. Es-*

- says in honour of Alan Davies, Cambridge University Press, Cambridge, pp. 241–50.
- LAZARATON A. (1996), *Interlocutor support in oral proficiency interviews: the case of CASE*, in «Language Testing» 13, pp. 151–72.
- LEE Y. (2015), *Future of diagnostic language assessment*, in «Language Testing» 32, 3, pp. 295–8.
- LEHMANN C. (1988), *Towards a typology of clause linkage*, in J. Haiman, S.A. Thompson (eds.), *Clause combining in grammar and discourse*, John Benjamins Publishing Company, Amsterdam.
- LENNON P. (1990), *Investigating fluency in EFL: A quantitative approach*, in «Language Learning» 40, pp. 387–417.
- LEVELT W. (1989), *Speaking: From Intention to Articulation*, Cambridge University Press, Cambridge.
- LINACRE J.M. (1989), *Many-faceted Rasch measurement*, MESA Press, Chicago.
- LITTLE D. (2011), *The Common European Framework of Reference for Languages: A research agenda*, in «Language Teaching» 44, 3, pp. 381–93.
- LIU G. (1991), *Interaction and second language acquisition: a case study of a Chinese child's acquisition of English as a second language*, Unpublished PhD thesis, La Trobe University, citato in E. Tarone, G. Liu (1995).
- LLURDA E. (2000), *On competence, proficiency, and communicative language ability*, in «International Journal of Applied Linguistics» 10, 1, pp. 85–94.
- LOBAN W. (1976), *Language development, Kindergarten through Grade 12*, Research Report N. 18, NCTE Illinois — syntactic structures citato in E. Barkhuizen (2005).

- LOEWEN S., REINDERS H. (2011), *Key Concepts in Second Language Acquisition*, Palgrave Macmillan International Higher Education, New York.
- LONG M. (1985a), *A role for instruction in second language acquisition: Task-based language training*, in K. Hyltenstam, M. Pienemann (eds.), *Modelling and assessing second language acquisition*, Multilingual Matters, Clevedon, Avon, pp. 77–100.
- (1985b), *Input in second language acquisition theory*, in S.M. Gass, C. Madden (eds.), *Input and Second Language Acquisition*, Newbury House, Rowley, MA, pp. 377–93.
- (1989), *Task, group, and task-group interaction*, in «University of Hawaii Working Papers in English as a Second Language» 8, pp. 1–26.
- (2015), *Second language acquisition and task based language teaching*, Wiley-Blackwell.
- (2016), *In Defense of Tasks and TBLT: NONissues and Real Issues*, in «Annual Review of Applied Linguistics» 36, pp. 5–33.
- LONG M., CROOKES G. (1992), *Three approaches to task based syllabus design*, in «TESOL Quarterly» 26, pp. 27–56.
- LONG M., NORRIS J.M. (2000), *Task-based teaching and assessment*, in M. Byram (ed.), *Encyclopedia of language teaching*, Routledge, London, pp. 597–603.
- LONG M., ROBINSON P. (1998), *Focus on form: theory, research and practise*, in C. Doughty, J. Williams (eds.), *Focus on Form in Classroom SLA*, Cambridge University Press, Cambridge, pp. 15–41.
- LONGACRE R. (1985), *Sentences as combinations of clauses*, in T. Shopen (ed.), *Language typology and syntactic description*, vol. II, Cambridge University Press, Cambridge.

- LOSCHKY L., BLEY-VROMAN R. (1993), *Grammar and task-based methodology*, in G. Crookes, S.M. Gass, *Tasks and Language Learning: Integrating Theory and Practice*, Multilingual Matters, Clevedon, Avon, pp. 123–67.
- LU X. (2011), *A corpus-based evaluation of syntactic complexity measures as indices of college-level ESL writers' language development*, in «TESOL Quarterly» 45/1, pp. 36–62.
- LUMLEY T., MCNAMARA T. (1995), *Rater characteristics and rater bias: implications for training*, in «Language Testing» 12, pp. 54–71.
- LUMLEY T., O'SULLIVAN B. (2005), *The effect of test-taker gender, audience and topic on task performance in tape-mediated assessment of speaking*, in «Language Testing» 22, 4, pp. 415–437.
- LUNZ M.E., STAHAL J. (1990), *Judge consistency and severity across grading periods*, in «Evaluation and the Health Professions» 13, pp. 425–44.
- LUNZ M.E., WRIGHT B.D., LINACRE J.M. (1990), *Measuring the impact of judge severity on examination scores*, in «Applied Measurement in Education» 3, pp. 331–45.
- LYNCH T., MCLEAN J. (2001), *A case of exercising: effects of immediate task repetition on learners' performance*, in M. Bygate, P. Skehan, M. Swain (eds.), *Researching Pedagogic Tasks, Second Language Learning, Teaching and Testing*, Longman, London, pp. 141–61.
- MACINTYRE P., LEGATTO J.J. (2011), *A dynamic system approach to willingness to communicate: Developing an idiodynamic method to capture rapidly changing affect*, in «Applied Linguistics» 32, 2, pp. 149–71.
- MACWHINNEY B. (1998), *Models of emergence of language*, in «Annual Review of Psychology» 49, pp. 199–227.
- (ed.) (1999), *The emergence of language*, Lawrence Erlbaum, Mahwah, NJ.

- (2001), *The gradual evolution of language*, in B. Malle, T. Givón (eds.), *The evolution of language*, John Benjamins Publishing Company, Philadelphia, pp. 1–28.
- (2006), *Emergentism — Use often and with care*, in «Applied Linguistics», Special Issue 27, pp. 729–40.
- MACWHINNEY B., O'GRADY W. (2015), *The Handbook of Language Emergence*, John Wiley, Sons.
- MALIK A.A. (1990), *A psycholinguistic analysis of the reading behaviour of EFL-proficient readers using culturally familiar and culturally nonfamiliar expository texts*, in «American Educational Research Journal» 27/ 1, pp. 205–223.
- MAJOR R.C. (2004), *Gender and stylist variation in second language phonology*, in «Language Variation and Change» 16, pp. 169–188.
- MAJOR R.C., FITZMAURICE S.F., BUNTA F., BALASUBRAMANIAN C. (2002), *The effects of nonnative accents on listening comprehension: Implication for ESL assessment*, in «TESOL Quarterly» 36, 2, pp. 173–90.
- MANCHÓN R.M. (ed.) (2012), *L2 writing development: Multiple perspectives*, Walter de Gruyter, Boston.
- MARTÍNEZ–FLOR A. (2006), *Task effects on EFL learners productions of suggestions: A focus on elicited phone messages and emails*, in «Miscelánea: A journal of English and American Studies» 33, pp. 47–64.
- MARTINI A. (a cura di) (2017), *Rilevazioni Nazionali degli Apprendimenti 2016–2017*, «Rapporto Risultati», [http://www.invalsi.it/invalsi/doc\\_eventi/2017/Rapporto\\_Prove\\_INVALSI\\_2017.pdf](http://www.invalsi.it/invalsi/doc_eventi/2017/Rapporto_Prove_INVALSI_2017.pdf).
- MCINTYRE P.N. (1993), *The importance and effectiveness of moderation training on the reliability of teacher assessments of ESL writing samples*, Unpublished MA Thesis, University of Melbourne, citata in McNamara (1996).

- MCLAUGHLIN B. (1990), *Restructuring*, in «Applied Linguistics» 11, pp. 113–28.
- MCMNAMARA T.F. (1995), *Modelling performance: Opening Pandora's box*, in «Applied Linguistics» 16, pp. 159–79.
- (1996), *Measuring Second Language Performance*, Longman, London.
- (1997), *Performance testing*, in C. Clapham, D. Corson (eds.), *Language testing and assessment, Encyclopedia of language and education*, vol 7, Kluwer Academic, Dordrecht, pp. 131–9.
- (2006), *Validity in language testing: The challenge of Sam Messick's legacy*, in «Language Assessment Quarterly» 3, pp. 31–51.
- MCMNAMARA T.F., HILL K., MAY L. (2002), *Discourse and assessment*, in «Annual Review of Applied Linguistics» 22, pp. 221–42.
- MEHNERT U. (1998), *The effects of different lengths of time for planning on second language performance*, in «Studies in Second Language Acquisition» 20, pp. 52–83.
- MEHRANG F., RAHIMPOUR M. (2010), *The impact of task structure and planning conditions on oral performance of EFL learners*, in «Procedia Social and Behavioral Sciences»: World Conference on Educational Sciences 2010, Istanbul.
- MEISEL J.M. (1997), *The acquisition of the syntax of negation in French and German: Contrasting first and second language development*, in «Second Language Research» 13, pp. 227–63.
- MEISEL J.M., CLASHEN H., PIENEMANN M. (1981), *On determining developmental stages in second language acquisition*, in «Studies in Second Language Acquisition» 3, pp. 109–35.
- MESSICK S. (1989), *Validity*, in R.L. LINN (ed.), *Educational measurement*, American Council in Education and Macmillan, pp. 13–103.

- (1996), *Validity and washback in language testing*, in «Language Testing» 13, pp. 241–56.
- MEZZADRI M. (2004), *Il Quadro comune europeo a disposizione della classe. Un percorso verso l'eccellenza*, Guerra, Perugia.
- MICHEL M.C., KUIKEN F., VEDDER I. (2007), *The influence of complexity in monologic versus dialogic tasks in Dutch L2*, in «IRAL, International Review of Applied Linguistics» 45, pp. 241–59.
- MILLER J., WEINERT R. (1998), *Spontaneous spoken language: syntax and discourse*, Clarendon, Oxford.
- MISLEVY R.J., STEINBERG L.S., ALMOND R.G. (2003), *On the structure of assessment arguments*, in «Measurement: Interdisciplinary Research and Perspective» 1, pp. 3–62.
- MIYAKE A., SHAH P. (1999), *Models of working memory: Mechanism of active maintenance and executive control*, Cambridge University Press, Cambridge.
- MOCHIZUKI N., ORTEGA L. (2008), *Balancing communication and grammar in beginner-level foreign language classrooms: A study of guided planning and relativization*, in «Language Teaching Research» 12, 1, pp. 11–37.
- MÖHLE D. (1984), *A comparison of the second language production of different native speakers*, in H.W. Dechert, D. Möhle, M. Raupach (eds.), *Second Language Productions*, Gunter Narr, Tübingen, pp. 26–49.
- MONEGLIA M. (2000a), *Corpus LABLITA, corpus italien de langue parlée*, in M. Bilger (ed.), *Corpus. Méthodologie et applications linguistiques*, Honoré Champion, Paris.
- (2000b), *Units of analysis of spontaneous speech and speech variation in a cross-linguistic perspective*, in <http://lablita.dit.unifi.it/italian/preprint.html>.
- (2006), *Units of analysis of spontaneous speech and speech variation in a cross-linguistic perspective*, in Y. Kawaguchi, S. Zaima, T. Ta-



- kagaki (eds.), *Spoken language corpus and linguistics informatics*, John Benjamin Publishing, Amsterdam, pp. 153–179.
- MONROE J.H. (1975), *Measuring and enhancing syntactic fluency in French*, in «The French Review» 48, pp. 1023–31.
- MOONEN M., STOUTJESDIJK E., GRAAFF R., CORDA A. (2013), *Implementing CEFR in secondary education: impact on FL teachers' educational and assessment practice*, in «International Journal of Applied Linguistics» 23, 2, pp. 226–46.
- MORROW K. (1979–81), *Communicative language testing: Revolution or evolution?*, in C.J. Brumfit, K. Johnson (eds.), *The communicative approach to language teaching*, Oxford University Press, Oxford, reprinted in J.C. Alderson, A. Huges (eds.) (1981), *Issues in language testing*, ELT Document 111, British Council, London, pp. 9–25.
- (ed.) (2004), *Insight from the Common European Framework*, Oxford University Press, Oxford.
- (2018), *Communicative Language Testing*, in *The TESOL Encyclopedia of English Language Teaching*.
- MOUGEON R., REHNER K. (2001), *Acquisition of sociolinguistic variants by French immersion students: The case of restrictive expressions, and more*, in «Modern Language Journal» 85, pp. 398–415.
- NEFF J., DAFOUZ E., DÌEZ M., PRIETO R., CHAUDRON C. (1998), *Contrastive discourse analysis: Argumentative text in English and Spanish*, Paper presented at the Twenty-Fourth Linguistics Symposium: Discourse across Languages and Cultures, University of Wisconsin–Milwaukee, September, citato in ORTEGA (2003).
- NEMETH N., KORMOS J. (2001), *Pragmatic aspects of task performance: the case of argumentation*, in «Language Teaching Research» 5, pp. 213–40.

- NEWELL A., SIMON H. (1972), *Human Problem Solving*, Prentice Hall, Englewood Cliffs, NJ.
- NICOLÁS-CONESA F., ROCA DE LARIOS J., COYLE Y. (2014), *Development of EFL students' mental models of writing and their effects on performance*, in «Journal of Second Language Writing» 24, 2, pp. 1–19.
- NIPPOLD M. (2004), *Research on later language development: the school-age and adolescent years*, in R. Berman (ed.), *Language Development Across Childhood and Adolescence*, John Benjamins Publishing Company, Amsterdam, pp. 1–8.
- NIWA (2000), citato in P. Robinson (ed.) (2001a), *Cognition and second language instruction*, Cambridge University Press, Cambridge.
- NORRIS J.M. (2002), *Interpretations, intended uses and designs in task-based language assessment*, in «Language Testing», 19, 4, pp. 337–346.
- (2009), *Task-based teaching and testing*, in M. Long, C. Doughty (eds.), *Handbook of language teaching*, Blackwell, Cambridge, MA, pp. 578–94.
- (2016), *Current uses of task-based language assessment*, in «Annual Review of Applied Linguistics» 36, pp. 230–44.
- NORRIS J.M., MANCHON R. (2012), *Investigating L2 writing development from multiple perspectives: issues in theory and research*, in R. Manchon (ed.), *L2 Writing: Multiple Perspectives*. Walter de Gruyter, Berlin, pp. 221–44.
- NORRIS J., ORTEGA L. (2003), *Defining and measuring SLA*, in C.J. Doughty, M.H. Long (eds.), *The Handbook of Second Language Acquisition*, Blackwell Publishing, Malden, MA, pp. 717–61.
- (2006), *Synthesising Research on Language Learning and Teaching*, John Benjamins Publishing Company, Amsterdam.

- (2009), *Towards an organic approach to investigating CAF in instructed SLA: The case of complexity*, in «Applied Linguistics» 30, 4, pp. 555–78.
- NORRIS J.M., YOSHIOKA J., BROWN J.D., HUDSON T. (1998), *Designing second language performance assessments*, University of Hawaii at Manoa, Honolulu.
- NORTH B. (1991), *Standardisation of continuous assessment grades*, in C.J. Alderson, B. North (eds.), *Language Testing in the 1990s*, Macmillan, London, pp. 167–77.
- (2000), *The development of a common european framework scale of language proficiency*, Peter Lang, New York.
- (2002), *A CEF-based self assessment tool for university entrance*, in J.C. Alderson (ed.), *Common European Framework of Reference for Languages: Learning, Teaching, Assessment. Case studies*, Council of Europe, Strasbourg, pp. 146–66.
- (2007), *The CEFR illustrative descriptive scales*, in «The Modern Language Journal» 91, 4, pp. 656–59.
- (2014), *Putting the Common European Framework of Reference to good use*, in «Language Teaching» 47, 2, pp. 228–49.
- NORTH B., SCHNEIDER G. (1998), *Scaling descriptors for language proficiency scales*, in «Language Testing» 15, pp. 217–62.
- NUNAN D. (1989), *Designing Tasks for the Communicative Classroom*, Cambridge University Press, Cambridge.
- OCHS E., SCHEGLOFF E., THOMPSON S.A. (eds.) (1996), *Interaction and Grammar*, Cambridge University Press, Cambridge.
- OECD (2018), *PISA 2015. Results in Focus*, <https://www.oecd.org/pisa/pisa-2015-results-in-focus.pdf>.
- OLLER J.W. (1979), *Language tests at school. A pragmatic approach*, Longman, London.

- (1983), *A consensus for the Eighties?*, in J.W. Oller (ed.), *Issues in language testing research*, Newbury House, Rowley (Mass.), pp. 351–56.
- ORTEGA L. (1999), *Planning and focus on form in L2 oral performance*, in «*Studies in Second Language Acquisition*» 21, pp. 108–48.
- (2003), *Syntactic complexity measures and their relationship to L2 proficiency: A research synthesis of College-level L2 writing*, in «*Applied Linguistics*» 24, pp. 492–518.
- OSCARSON M. (1978), *Approaches to Self-assessment in Foreign Language Learning*, Council of Europe, Strasbourg.
- (1984), *Self-assessment of Foreign Language Skills: a Survey of Research and Development Work*, Council of Europe, Strasbourg.
- OSSERVATORIO NAZIONALE (2007), *La via italiana per la scuola interculturale e l'integrazione degli alunni stranieri*, in [www.istruzione.it](http://www.istruzione.it).
- O'SULLIVAN B. (2000), *Exploring Gender and Oral Proficiency Interview Performance*, in «*System*» 28, 3, pp. 373–86.
- O'SULLIVAN B., GREEN A. (2011), *Test taker characteristics*, in L. Taylor (ed.), *Examining Speaking: Research and Practice* in «*Assessing Second Language Speaking*», Cambridge University Press, Cambridge, pp. 36–65.
- PALLOTTI G. (2007), *An operational definition of the emergence criterion*, in «*Applied Linguistics*» 28, pp. 361–82.
- (2009), *CAF: Defining, refining and differentiating constructs*, in «*Applied Linguistics*» 30, 4, pp. 590–601.
- (2015), *A simple view of linguistic complexity*, in «*Second Language Research*» 31, 1, pp. 117–34.
- PALLOTTI G., FERRARI S. (2008), *La variabilità situazionale dell'interlingua: implicazioni per la ricerca acquisizionale e il testing*, in G. Bernini, L. Spreafico, A. Valentini (a cura di), *Competenze*

- lessicali e discorsive nell'acquisizione di lingue seconde*, Guerra, Perugia, pp. 437–462.
- PALLOTTI G., FERRARI S., NUZZO E., BETTONI C. (2010), *Una procedura sistematica per osservare la variabilità nell'interlingua*, in «Studi Italiani di Linguistica Teorica e Applicata» 39, 2, pp. 215–41.
- PALLOTTI G., FRONTINI G., AFFÉ F., MONACHINI M., FERRARI S. (2014), *Presenting a System of Human–Machine Interaction for Performing Map Tasks*, in N. Calzonari (a cura di) *Proceedings of the Ninth International Conference on Language Resources and Evaluation (LREC '14)*, European Language Resources Association (ELRA), Reykjavik, Iceland, pp. 3963–66.
- PERDUE C., KLEIN W. (1992), *Why does the production of some learners not grammaticalize?*, in «Studies in Second Language Acquisition» 14, pp. 259–72.
- PERKINS K., GUPTA L., TAMMANA R. (1995), *Predicting item difficulty in a reading comprehension test with an artificial neural network*, in «Language Testing» 12, pp. 34–53.
- PICA T., HALLIDAY L., LEWIS N., MORGENTHALER L. (1989), *Comprehensible outputs as an outcome of linguistic demands on the learner*, in «Studies in Second Language Acquisition» 11/1, pp. 63–90.
- PIENEMANN M. (1985), *Learnability and syllabus construction*, in K. Hyltenstam, M. Pienemann (eds.), *Modelling and assessing second language acquisition*, Multilingual Matters, Clevedon, Avon, pp. 23–75.
- (1987), *Determining the influence of instruction on L2 speech processing*, in «Australian Review of Applied Linguistics» 10, pp. 83–113.
- (1988), *Language Processing and Second Language Development: Processability Theory*, John Benjamins Publishing Company, Amsterdam.

- (1992), *COALA: a computational system for interlanguage analysis*, in «Second Language Research» 8, pp. 59–92.
- (1998), *Language Processing and Second-Language Development: Processability Theory*, John Benjamins Publishing Company, Amsterdam.
- (2007), *Variation and dynamic systems in SLA*, in «Bilingualism: Language and Cognition» 10, pp. 43–5.
- (2008), *A brief introduction to processability theory*, in J.U. Keßler (ed.), *Processability Approaches to Second Language Development and Second Language Learning*, Cambridge University Press, Cambridge, pp. 9–30.
- PIENEMANN M., MACKAY A. (1992), *An empirical study of children ESL development and Rapid Profile*, in P. McKay (ed.), *ESL development: Language and Literacy in Schools project*, Vol. II, National Languages and Literacy Institute of Australia, pp. 115–259.
- PLONSKY L. (ed.) (2015), *Advancing quantitative methods in second language research*, Routledge, New York.
- PLONSKY L., KIM Y. (2016), *Task-based Learner Production: A substantive and Methodological Review*, in «Annual Review of Applied Linguistics» 36, pp. 73–97.
- POLAT B., KIM Y. (2014), *Dynamics of Complexity and Accuracy: A Longitudinal Case Study of Advanced Untutored Development*, in «Applied Linguistics» 35, 2, pp. 1–25.
- POLIO C. (1997), *Measures of linguistic accuracy in second language writing research*, in «Language Learning» 47, pp. 101–43.
- (2001), *Research Methodology in Second Language Writing Research: The Case of Text Based Studies*, in T. Silva, P.K. Matsuda (eds.), *On second language writing*, Lawrence Erlbaum, Mahwah, NJ, pp. 91–115.

- (2017), *Second language writing development: a research agenda*, in «Language Teaching» 50, 2, pp. 261–275.
- POLIO C., SHEA M.C. (2014), *An investigation into current measures of linguistic accuracy in second language writing research*, in «Journal of Second Language Writing» 26, pp. 10–27.
- PRABHU N.S. (1987), *Second Language Pedagogy*, Oxford University Press, Oxford.
- PRESTON D. (1993), *Variation linguistics and SLA*, in «Second Language Research» 9, pp. 153–72.
- (1996), *Variationist linguistics and second language acquisition*, in W.C. Ritchie, T.K. Bhatia (eds.), *Handbook of Second Language Acquisition*, Academic Press, New York, pp. 229–65.
- (2000), *Three kinds of sociolinguistics and SLA: a psycholinguistics perspective*, in B. Schwierzbin, F. Morris, M. Anderson, C. Klee, E. Tarone (eds.), *Social and Cognitive Factors in Second Language Acquisition*, Cascadilla Press, Somerville, MA, pp. 3–30.
- (2002), *A variationist perspective on SLA: Psycholinguistic concerns*, in R. Kaplan (ed.), *Oxford Handbook of Applied Linguistics*, Oxford University Press, Oxford, pp. 141–159.
- PURPURA J.E. (1993), *The development and construct validation of an instrument designed to investigate selected cognitive background characteristics of test-takers*, in A.J. Kunnan (ed.), *Validation in language assessment*, Lawrence Erlbaum, Mahwah, pp. 114–40.
- (1997), *An analysis of the relationships between test takers' cognitive and metacognitive strategy use and second language test performance*, in «Language Learning» 47, pp. 289–325.
- (1999), *Modelling the relationships between test takers' reported cognitive and metacognitive strategy use and performance on language*

- tests, University of Cambridge Local Examinations Syndacate and Cambridge University Press, Cambridge.
- (2004), *Assessing grammar*, Cambridge University Press, Cambridge.
- (2016), *Second and Foreign Language Assessment*, in «The Modern Language Journal» 100, pp. 190–208.
- PURPURA J.E., TURNER C.E. (2014), *Learning-oriented assessment in language classrooms: Using assessment to gauge and promote language learning*, Routledge, New York.
- RAHIMPOUR M. (1997), *Task condition, task complexity and variation in oral discourse*, unpublished PHD dissertation, citato in P. Robinson (2001).
- (1999), *Task complexity and variation in interlanguage*, in N. Jungheim, P. Robinson (eds.), *Pragmatics and Pedagogy: Proceedings of The 3rd Pacific Second Language Research Forum*, vol. 2, PacSLRF, Tokyo, pp. 115–34.
- Rassegna Italiana di Linguistica Applicata* (2012), «Special Issue», S. Cacchiani, S. Morgan, M. Silver.
- RASTELLI S. (2013), *Il processing nella seconda lingua. Teoria, dati sperimentali, didattica*, Carocci editore, Roma.
- (2014), *Discontinuity in Second Language Acquisition. The Switch between Statistical and Grammatical Learning*, Multilingual Matters, Cleveland.
- RAUPACH M. (1983), *Analysis and evaluation of communication strategies*, in C. Faerch, G. Kasper (eds.), *Strategies in «Interlanguage Communication»*, Longman, London/New York, pp. 199–209.
- RAVID D., BERMAN R. (2010), *Developing noun phrase complexity at*



- school age: A text-embedded cross-linguistic analysis*, in «First Language» 30, pp. 3–26.
- READ J. (2007), *Towards a new collaboration: research in SLA and language testing*, in «New Zealand Studies in Applied Linguistics» 13, pp. 22–35.
- REGAN V. (1995), *The acquisition of sociolinguistic native speech norms: Effects of a year abroad on second language learners of French*, in B. Freed (ed.), *Second Language Acquisition in a Study Abroad Context*, John Benjamins Publishing Company, Amsterdam/Philadelphia, pp. 245–67.
- (1996), *Variation in French interlanguage: A longitudinal study of sociolinguistic competence*, in R. Bayley, D. Preston (eds.), *Second Language Acquisition and Linguistic Variation*, John Benjamins Publishing Company, Amsterdam, pp. 177–201.
- (1997), *Les apprenants avancés, la lexicalisation et l'acquisition de la compétence sociolinguistique: une approche variationniste*, in «AILE: Acquisition et Interaction en Langue Etrangère» 9, pp. 193–210.
- (2004), *The relationship between the group and the individual and the acquisition of native speaker variation patterns: A preliminary study*, «IRAL International Review of Applied Linguistics» 42, pp. 335–48.
- REHNER K., MOUGEON R. (1999), *Variation in the spoken French of immersion students: to ne or not to ne, that is the sociolinguistic question*, in «Canadian Modern Language Review» 56, pp. 124–56.
- REHNER K., MOUGEON R., NADASDI T. (2003), *The learning of sociolinguistic variation by advanced FLS learners: The case of nous versus on in immersion French*, in «Studies in Second Language Acquisition» 25, pp. 127–56.
- RENZI L., SALVI G. (a cura di) (2001), *Grande grammatica italiana di consultazione*, vol. II: *I sintagmi verbale, aggettivale, avverbale. La subordinazione*, il Mulino, Bologna.

- RENZI L., SALVI G., CARDINALETTI A. (1988–1995), *Grande grammatica italiana di consultazione*, vol I–III, il Mulino, Bologna.
- RIMMER W. (2006), *Measuring grammatical complexity: The Gordian knot*, in «Language Testing» 23, pp. 497–519.
- ROBINSON P. (2000), *Task complexity and reasoning demands: effects on dyadic NNS–NNS interaction, fluency, accuracy and incorporation of input*. Unpublished data, Aoyama Gakuin University, Tokyo, citato in Robinson (2001a).
- (2001a) (ed.), *Cognition and second language instruction*, Cambridge University Press, Cambridge.
- (2001b), *Task complexity, cognitive resources, and syllabus design: A triadic framework for examining task influences on SLA*, in P. Robinson (ed.), *Cognition and Second Language Instruction*, Cambridge University Press, Cambridge, pp. 287–318.
- (2001c), *Task complexity, task difficulty, and task production: Exploring interaction in a componential framework*, in «Applied Linguistics» 22, pp. 27–57.
- (2003a), *Attention and memory*, in C.J. DOUGHTY, M.H. LONG (eds.), *The Handbook of Second Language Acquisition*, Blackwell, Oxford, pp. 631–78.
- (2003b), *The Cognition Hypothesis, task design and adult task-based second language learning*, in «Second Language Studies», 21, pp. 45–105.
- (2005a), *Cognitive complexity and task sequencing: A review of studies in a componential framework for second language task design*, in «IRAL International Review of Applied Linguistics» 43, pp. 1–33.
- (2005b), *Aptitude and second language acquisition*, in «Annual Review of Applied Linguistics» 25, pp. 45–73.

- (2007a), *Criteria for classifying and sequencing pedagogic tasks*, in M. García Mayo (ed.), *Investigating tasks in formal language settings*, Multilingual Matters, Clevedon, pp. 7–25.
- (2007b), *Task complexity, theory of mind, and intentional reasoning: effects of L2 speech production, interaction, uptake and perceptions of task difficulty*, in «IRAL International Review of Applied Linguistics» 45, pp. 193–213.
- (2007c), *Aptitude, abilities, contexts, practice*, in R. DeKeyser (ed.), *Practice in a Second Language: Perspectives from Applied Linguistics and Cognitive Psychology*, Cambridge University Press, New York, pp. 256–286.
- (2011a), *Task-based Language Learning: A Review of Issues*, in «Language Learning», 61, 1, pp. 1–36.
- (ed.) (2011b), *Second Language Task Complexity. Researching the Cognition Hypothesis of language learning and performance*, John Benjamins Publishing Company, Amsterdam.
- (2012), *Individual differences, Aptitude Complexes, SLA Processes, and Aptitude test Development*, in M. Pawlak (ed.), *New Perspectives on Individual Differences in Language Learning and Teaching*, Springer-Verlag, Berlin/Heidelberg, pp. 57–75.
- (2015), *The Cognition Hypothesis, second language task demands, and the SSARC model of pedagogic task sequencing*, in M. Bygate (ed.), *Domains and Directions in the Development of TBLT. A decade of plenaries from the international conference*, John Benjamins Publishing Company, Amsterdam, pp. 87–122.
- (2017), *Attention and Awareness*, in J. Cenoz, D. Gorter, S. May (eds.), *Language Awareness and Multilingualism*, Springer, pp. 125–134.
- ROBINSON P., ELLIS N.C. (eds.) (2008), *Handbook of Cognitive Linguistics and Second Language Acquisition*, Routledge, London.

- SABATINI F. (2011), *L'italiano dell'uso medio: una realtà tra le varietà linguistiche italiane* in V. Coletti, R. Coluccia, P. D'Achille, N. De Blasi, D. Proietti, R. Cimaglia (a cura di), *L'italiano nel mondo moderno. Saggi scelti dal 1968 al 2009*. Tomo II, Liguori, Napoli, pp. 3–36.
- SACKS H., SCHEGLOFF E.A., JEFFERSON G. (1974), *A Symplest Systematics for The Organization of Turn-Taking*, in «*Conversation, in Language*» 50, pp. 696–735.
- SAMPLE E., MICHEL M. (2014), *An Exploratory Study into Trade-off Effects of Complexity, Accuracy, and Fluency on Young Learners' Oral Task Repetition*, in «*TESL Canada Journal*» 31, Special Issue, pp. 23–46.
- SAMUDA V. (2001), *Guiding relationship between form and meaning during task performance: the role of the teacher*, in M. Bygate, P. Skehan, M. Swain (eds.), *Researching Pedagogic Tasks: Second Language Learning, Teaching, and Testing*, Longman, London, pp. 119–40.
- SASAKI M. (2004), *A multiple-data analysis of the 3.5-year development of EFL student writers*, in «*Language Learning*» 54, 3, pp. 525–82.
- SASAKI T. (2008), *Concurrent think-aloud protocol as a socially situated construct*, in «*International Review of Applied Linguistics*» 46, pp. 349–74.
- SATO C.J. (1988), *Origins of Complex Syntax in Interlanguage Development*, in «*Studies in Second Language Acquisition*» 10/3, pp. 371–395.
- SAVY R., VOGHERA M. (2010), *A corpus-based study on syntactic and phonetic prosodic phrasing boundaries in spontaneous Italian speech*, in M. Hasegawa-Johnson (ed.), *Proceedings of Speech Prosody 2010*, <http://speechprosody2010.illinois.edu/papers/100077.pdf>.
- SCARANO A. (1996), *Fraasi relative e pseudorelative in italiano*, in «*Studi di grammatica italiana*» XVI, pp. 377–423.

- (2003), *Relative appositive e aggettivi appositivi. Tra sintassi e articolazione dell'informazione*, in P. D'Achille (a cura di), *Atti del convegno internazionale della SILFI, Generi, architetture e forme testuali*, Roma 1–5 ottobre 2002, in preprint LABLITA 1, 1–12, in <http://lablita.dit.unifi.it/italian/preprint.html>.
- SCHEGLOFF E.A. (1979), *Identification and recognition in telephone conversation openings*, in G. Psathas (ed.), *Everyday Language: studies in ethnomethodology / conversation analysis*, Irvington Publishers, New York.
- (1986), *The Routine as Achievement*, in «Human Studies» 9, pp. 111–51.
- (1996), *Turn Organization: One Intersection of Grammar and Interaction*, in E. Ochs, E.A. Schegloff, S. Thompson (eds.), *Interaction and Grammar*, Cambridge University Press, Cambridge, pp. 52–133.
- (2000), *Overlapping Talk and the Organization of Turn-Taking for Conversation*, in «Language in Society» 29, 1, pp. 1–63.
- (2011), *Word repeats as unit ends*, in «Discourse Studies» 13, 3, pp. 367–380.
- SCHMIDT R.W. (1992), *Psychological mechanism underlying second language fluency*, in «Studies in Second Language Acquisition» 14, pp. 357–85.
- SCHNEIDER G., NORTH B. (2000), *Fremdsprachen Können: was heisst das? Skalen zur Beschreibung, Beurteilung und Selbsteinschätzung der fremdsprachlichen Kommunikationsfähigkeit*, in *Nationales Forschungsprogramm 33: Wirksamkeit unserer Bildungssysteme*, Verlag Rüegger, Chur, Zürich.
- SCHNEIDER K.P., BARRON A. (2008), *Variational Pragmatics. A focus on regional varieties in pluricentric languages*, John Benjamins Publishing Company, Amsterdam.

- SCHOONEN R., VERGEER M., EITING M. (1997), *The assessment of writing ability: expert readers versus lay readers*, in «Language Testing» 14, pp. 157–84.
- SCOLLON R.T. (1976), *Conversations with a One-year Old: A Case Study of the Developmental Functions of Syntax*, University of Hawaii Press, Honolulu.
- SCOTT C.M. (1988), *Spoken and written syntax*, in M.A. Nippold (ed.), *Lateral language development. Ages 9 through 19*, College Hill Press, Boston, pp. 49–95.
- SEGALOWITZ N. (2010), *Cognitive bases of second language fluency*, Routledge, New York.
- SELTING M. (1996), *On the interplay of syntax and prosody in the constitution of turn-constructive units and turns in conversation*, in «Pragmatics» 6, pp. 357–88.
- SHARMA A. (1980), *Syntactic maturity: Assessing writing proficiency in a second language*, in R. Silverstein (ed.), *Occasional Papers in Linguistics* 6, Southern Illinois University, Carbondale.
- SHOHAMY E. (1983), *The stability of the oral proficiency trait on the Oral Interview Speaking Test*, in «Language Learning» 33, pp. 527–40.
- (1994), *The role of language tests in the construction and validation of second language acquisition theories*, in E. Tarone, S.M. Gass, A.D. Cohen (eds.), *Research methodology in second language acquisition*, Lawrence Erlbaum, Hillsdale, NJ, pp. 115–31.
- (1998), *How can language testing and SLA benefit from each other? The case of discourse*, in L.F. Bachman, A.D. Cohen (eds.), *Interfaces between second language acquisition and language testing research*, Cambridge University Press, Cambridge, pp. 156–76.
- (2000), *The relationship between language testing and second language acquisition, revisited*, in «System» 28, pp. 541–53.

- (2017), *Critical Language Testing*, in «Language Testing and Assessment», pp. 441–454.
- SHOHAMY E., GORDON C.M., KRAEMER R. (1992), *The effect of raters' background and training on the reliability of direct writing tests*, in «Modern Language Journal» 76, pp. 27–33.
- SHOHAMY E., REVES T., BEJARANO Y. (1986), *Introducing a new comprehensive test of oral proficiency*, in «English Language Teaching Journal» 40, pp. 212–22.
- SIMONE R. (1990), *Fondamenti di linguistica*, Bari, Laterza.
- SIMONS M., COLPAERT J. (2015), *Judgmental evaluation of the CEFR by stakeholders in language testing*, in «Revista de Lingüística y Lenguas Aplicadas» 10, 1, pp. 66–7.
- SKEHAN P. (1989), *Individual Differences in Second–Language Learning*, Edward Arnold, London.
- (1995b), *Analysability, accessibility and ability for use*, in G. COOK, S. Seidlhofer (eds.), *Principle and practice in applied linguistics: Studies in honour of H.G. Widdowson*, Oxford University Press, Oxford, pp. 91–106.
- (1996a), *A framework for the implementation of task-based instruction*, in «Applied Linguistics» 17, pp. 38–62.
- (1998), *A cognitive approach to language learning*, Cambridge University Press, Cambridge.
- (2001), *Task and language performance assessment*, in M. Bygate, P. Skehan, M. Swain (eds.), *Researching Pedagogic Tasks, Second Language Learning, Teaching and Testing*, Pearson Education, London, pp. 167–85.
- (2003), *Task-based instruction*, in «Language Teaching» 36, pp. 1–14.
- (2008), *Models of speaking and assessment of second language pro-*

- iciency*, in P. Skehan, *Second Language Task Based Performance. Theory, Research, Assessment*, Routledge, Oxford.
- (2009), *Modelling Second Language Performance: Integration Complexity, Accuracy, Fluency and Lexis*, in «Applied Linguistics» 30, 4, 1, pp. 510–32.
- (2011), *Researching Tasks: Performance, Assessment, and Pedagogy*, Shanghai Foreign Language Education Press.
- (ed.) (2014), *Processing perspectives on task performance*, John Benjamins Publishing Company, Amsterdam.
- (2015), *Limited Attention Capacity and Cognition: Two hypothesis regarding second language performance on tasks*, in M. Bygate (ed.) *Domains and Directions in the Development of TBLT. A decade of plenaries from the international conference*, John Benjamins Publishing Company, Amsterdam, pp. 123–56.
- (2016), *Tasks Versus Conditions: Two Perspectives on Task Research and Their Implications for Pedagogy*, in «Annual Review of Applied Linguistics» 36, pp. 367–95.
- (2018), *Second Language Task-based Performance, Theory, Research, Assessment*, Routledge.
- SKEHAN P., FORSTER P. (1997), *Task type and task processing conditions as influences on foreign language performance*, in «Language Teaching Research» 1, pp. 185–211.
- (1999), *The influence of task structure and processing conditions on narrative retellings*, in «Language Learning» 49, pp. 93–120.
- (2001), *Cognition and tasks*, in P. Robinson (ed.), *Cognition and second language instruction*, Cambridge University Press, Cambridge, 183–205.
- (2005), *Strategic on-line planning: The influence of surprise information and task time on second language performance*, in R. Ellis (ed.),



- Planning and Task Performance in a Second Language*, John Benjamins Publishing Company, Amsterdam/Philadelphia, pp. 193–218.
- SLAVOFF G.R., JOHNSON J.S. (1995), *The effects of age on the rate of learning a second language*, in «Studies in Second Language Acquisition» 17, pp. 1–16.
- SMAGORINSKY P. (2001), *Rethinking protocol analysis from a cultural perspective*, in «Annual Review of Applied Linguistics» 21, pp. 233–45.
- SMITH J. (1989), *Topic and variation in ITA oral proficiency: SPEAK and field specific test*, in «English for Specific Purposes» 8, 155–67.
- SORNICOLA R. (1981), *Sul parlato*, il Mulino, Bologna.
- SPOELMAN M., VERSPOOR M. (2010), *Dynamic Patterns in Development of Accuracy and Complexity: A Longitudinal Case Study in the Acquisition of Finnish*, in «Applied Linguistics» 31, 4, 1, pp. 532–53.
- STEMLER S.E. (2004), *A comparison of consensus, consistency, and measurement approaches to estimating interrater reliability. Practical Assessment, Research & Evaluation* 9/4, <https://pareonline.net/getvn.asp?v=9&n=4>.
- STEPHAN A. (1999), *Varieties of emergentism*, in «Evolution and Cognition» 5, pp. 49–59.
- STERN H.H. (1983), *Fundamental concepts of language teaching*, Oxford University Press, New York.
- SWAIN M. (1998), *Focus on form through conscious reflection*, in G. Doughty, J. Williamson (eds.), *Focus on Form in Classroom Second Language Acquisition*, Cambridge University Press, Cambridge, pp. 64–81.
- SWAIN M., LAPKIN S. (2001), *Focus on Form through collaborative dialogue: exploring task effects*, in M. Bygate, P. Skehan, M. Swain (eds.), *Researching Pedagogic Tasks: Second Language Learning, Teaching, and Testing*, Longman, London, pp. 99–118.

- TALMY L. (1978), *Relations between subordination and coordination*, in J. Greenberg (ed.) *Universals of human language*, vol IV: «Syntax», Stanford University Press, Stanford, pp. 487–513.
- TARONE E. (1983), *On the variability of interlanguage systems*, in «Applied Linguistics» 4, pp. 143–63.
- (1985), *Variability in Interlanguage use: A study of style-shifting in morphology and syntax*, in «Language Learning» 35, pp. 373–403.
- (1988), *Variation in Interlanguage*, Edward Arnold, London.
- (1990), *On variation in interlanguage: A response to Gregg*, in «Applied Linguistics» 11, pp. 392–400.
- (2000), *Still wrestling with “context” in interlanguage theory*, in «Annual Review of Applied Linguistics» 20, pp. 182–98.
- (2007), *Sociolinguistic Approaches to Second Language Acquisition Research, 1997–2007*, in «The Modern Language Journal» 91, 1, pp. 837–48.
- (2010), *Social context and cognition in SLA: A variationist perspective*, in R. Batstone (ed.), *Sociocognitive perspectives on language use and language learning*, Oxford University Press, Oxford, pp. 54–72.
- TARONE E., LIU G. (1995), *Situational context, variation, and second language acquisition theory*, in G. Cook, B. Seidlhofer (eds.), *Principle and practice in applied linguistics. Studies in honour of H.G. Widdowson*, Oxford University Press, Oxford, pp. 107–24.
- TAVAKOLI P., SKEHAN P. (2005), *Strategic planning, task structure, and performance testing*, in R. Ellis (ed.), *Planning and task performance in a second language*, John Benjamins Publishing Company, Amsterdam, pp. 239–73.
- TAYLOR D.S. (1988), *The meaning and use of the term “competence” in linguistics and applied linguistics*, in «Applied Linguistics» 9, pp. 148–68.

- TELVE S. (2009), *Coordinazione e negazione. Sulla sintassi di né nell'italiano contemporaneo*, in *Riflessioni sull'Europa, Quaderni DiSCom, Dipartimento di Studi sulla Comunicazione, Università della Tuscia*, 1, Nutrimenti, Roma, pp. 60–79.
- THELEN E., BATES E. (2003), *Connectionism and dynamic systems: Are they really different?*, in «*Developmental Science*» 6, pp. 378–91.
- THELEN E., SMITH L.B. (1994), *A Dynamic Systems Approach to the Development of Cognition and Action*, MIT Press, Cambridge MA.
- THEWISSEN J. (2013), *Capturing L2 accuracy developmental patterns: Insights from an error-tagged EFL learner corpus*, in «*Modern Language Journal*» 97, pp. 77–101.
- TONKYN A. (2007), *Short-term changes in complexity, accuracy and fluency: Developing progress-sensitive proficiency measures*, in S. Van Daele, A. Housen, F. Kuiken, M. Pierrard, I. Vedder (eds.), *Complexity, accuracy and fluency in second language use, learning and teaching*, Brussels: KVAB, pp. 263–84.
- TOWELL R., HAWKINS R., BAZERGUI N. (1996), *The development of fluency in advanced learners of French*, in «*Applied Linguistics*» 17, pp. 84–119.
- TREFFERS-DALLER J., PARSLow P., WILLIAMS S. (2018), *Back to Basics: How Measures of Lexical Diversity Can Help Discriminate between CEFR Levels*, in «*Applied Linguistics*» 39, 3, pp. 302–27.
- TRIM J. (1978), *Some Possible Lines of Development of an Overall Structure for a European Unit/Credit Scheme for Foreign Language Learning by Adults*, Council of Europe, Strasbourg.
- (2001), *Common European Framework of Reference for Languages: Learning, teaching, assessment. A Guide for users*, Council of Europe, Strasbourg, <https://rm.coe.int/1680697848>.

- TURNER C.E., PURPURA J.E. (2015), *Learning-oriented assessment in the classroom*, in D. Tsigari, J. Banerjee (eds.), *Handbook of second language assessment*, DeGruyter Mouton, Berlin.
- VAN DAELE S., HOUSEN A., KUIKEN F., PIERRARD M., VEDDER I. (eds.) (2007), *Complexity, accuracy and fluency in second language use, learning and teaching*, Brussels: KVAB.
- VAN DEN BRANDEN K. (ed.) (2006), *Task-based language education: From theory to practice*, Cambridge University Press, Cambridge.
- VAN EK J. (1975), *The Threshold level in a European Unit/Credit System for Modern Language Learning by Adults*, Council of Europe, Strasbourg, republished in 1977 as *The Threshold level for Modern Language Learning*, Longman, London.
- (1977), *The Threshold Level for modern language learning in schools*, Longman, London.
- VAN EK J., TRIM J.L.M. (eds.) (1984), *Across the Threshold*, Pergamon Press, Oxford.
- (1990), *Waystage 1990*, Cambridge University Press, Cambridge.
- (1991), *Threshold 1990, a Revisited and Extended Version of "The Threshold level by J.A. van Ek"*, Council of Europe, Strasbourg.
- (1995), *Vantage Level*, Council of Europe, Strasbourg.
- VAN GEERT P. (1991), *A dynamic system mode of cognitive and language growth*, in «Psychological Review» 98, pp. 3–53.
- (1994), *Vygotskian Dynamics of Development*, in «Human Development» 37, pp. 346–65.
- (2005), *Explaining after by before: Basic aspects of a dynamic system approach to the study of development*, in «Developmental Review» 25, 3–4, pp. 408–42.
- VAN GEERT P., STEENBEEK H. (2005), *Explaining after by before: Basic*

- aspects of a dynamic systems approach to the study of development*, in «Developmental Review» 25, pp. 408–42.
- VAN GEERT P., VAN DIJK M. (2002), *Focus on variability: New tools to study intra-individual variability in developmental data*, in «Infant Behaviour, Development» 25, pp. 340–74.
- VAN LIER L. (1989), *Reeling, writhing, drawling, stretching, and fainting in coil: Oral proficiency interviews as conversation*, in «TESOL Quarterly» 23, pp. 489–508.
- VAN PATTEN B. (1990), *Attending to content and form in the input: an experiment in consciousness*, in «Studies in Second Language Acquisition» 12, pp. 287–301.
- (1996), *Input Processing and Grammar Instruction*, Ablex, New York.
- VASYLETS O., GILABERT R., MANCHÓN R.M. (2017), *The Effects of Mode and Task Complexity on Second Language Production*, in «Language Learning» 67, 2, pp. 394–430.
- VEDOVELLI M. (2002), *Guida all'italiano per stranieri: la prospettiva del quadro comune europeo*, Carocci, Roma.
- VERCELLOTTI M.L. (2017), *The Development of Complexity, Accuracy, and Fluency in Second Language Performance: A Longitudinal Study*, in «Applied Linguistics» 38, 1, pp. 90–111.
- VERSPOOR M., DE BOT K., LOWIE W. (eds.) (2011), *A Dynamic Approach to Second Language Development*, John Benjamins Publishing Company, Amsterdam.
- VERSPOOR M., SCHMID M.S., XU X. (2012), *A dynamic usage based perspective on L2 writing*, in «Journal of Second Language Writing» 21, 3, pp. 239–63.
- VOGHERA M. (1992a), *Sintassi e intonazione dell'italiano parlato*, il Mulino, Bologna.
- (1992b), *La nozione di semplificazione come categoria interpretativa*

- del parlato?*, in L. Brasca, M.L. Zambelli (a cura di), *Grammatica del parlare e dell'ascoltare a scuola*, Quaderni del Giscel, La Nuova Italia, Firenze, pp. 79–98.
- (2006), *Progettare la grammatica del parlato*, in M. Pettorino, A. Giannini, M. Vallone, R. Savy (a cura di), *La comunicazione parlata*, “Comunicazione parlata”, Napoli 23–25 febbraio 2006, Tomo III, Liguori, Napoli, pp. 1696–1714.
- (2017), *Dal parlato alla grammatica. Costruzione e forma dei testi spontanei*, Carocci editore, Roma.
- VOGHERA M., BASILE G., CERBASI D., FIORENTINO G. (2004), *La sintassi della clausola nel dialogo*, in A. Leoni, F. Cutugno, M. Pettorino, R. Savy (a cura di), *Il parlato italiano, Atti del Convegno Nazionale*, D’Auria Editore, Napoli, cd-rom.
- VOLLMER H.J. (1981), *Why are we interested in ‘general language proficiency’*, in J.C. Alderson, A. Huges (eds.), *Issues in language testing*, ELT Document 111, British Council, London, pp. 152–76.
- VOLLMER H.J., SANG F. (1983), *Competing hypothesis about second language ability: A plea for caution*, in J.W. OLLER (ed.), *Issues in language testing research*, Newbury House, Rowley (Mass.), pp. 29–79.
- WALKER I, CHAN D., NAGAMI KG, BOURGUIGNON M. (eds.) (2018), *New Perspectives on the Development of Communicative Related Competence in Foreign Language Education*, De Gruyter Mouton, Berlin.
- WANG W. (2009), *Interlanguage Theory and Emergentism: Reconciliation in Second Language Developmental Index Studies*, in «English Language Teaching» 2, pp. 31–36.
- WEIGLE S.C. (1994), *Effects of training on raters of ESL compositions*, in «Language Testing» 11, 2, pp. 197–223.
- (1998), *Using FACETS to model rater training effects*, in «Language Testing» 15, pp. 263–67.

- (1999), *Investigating rater/prompt interactions in writing assessment: Quantitative and qualitative approaches*, in «Assessing Writing» 6, 2, pp. 145–78.
- WEIGLE S.C., YANG W., MONTEE M. (2013), *Exploring reading processes in an academic reading test using short-answer questions*, in «Language Assessment Quarterly» 10, 1, pp. 28–48.
- WEIR C.J. (2005a), *Language Testing and Validation: An Evidence-based Approach*, Palgrave, McMillan, London.
- (2005b), *Limitations of the Common European Framework for developing comparable examinations and tests*, in «Language Testing» 22, pp. 281–300.
- WENDEL J.N. (1997), *Planning and second language narrative production*, unpublished doctoral dissertation, temple University, Japan, citato in ELLIS (2003).
- WESCHE M., SKEHAN P. (2002), *Communicative teaching, content-based instruction, and task-based learning*, in R. Kaplan (ed.), *Handbook of Applied Linguistics*, Oxford University Press, Oxford, pp. 207–228.
- WIDDOWSON H.G. (1989), *Knowledge of language and ability of use*, in «Applied Linguistics» 10, pp. 128–37.
- WIEMANN J.M., BACKLUND P. (1980), *Current theory and research in communicative competence*, in «Review of Educational Research» 50, pp. 185–99.
- WIESE R. (1984), *Language production in foreign and native languages. Same or different?*, in H. Dechert, D. Möhle, Raupach M. (eds.), *Second Language Productions*, Gunter Narr, Tübingen, pp. 11–25.
- WIGGLESWORTH G. (1993), *Exploring bias analysis as a tool for improving rater consistency in assessing oral interaction*, in «Language Testing» 10, pp. 305–35.

- (1997), *An investigation of planning time and proficiency level on oral test discourse*, in «Language Testing» 14, pp. 85–106.
- (2001), *Influences on performance in task-based oral assessments*, in M. Bygate, P. Skehan, M. Swain (eds.), *Researching Pedagogic Tasks: Second Language Learning, Teaching, and Testing*, Longman, London, pp. 186–209.
- (2005), *Current approaches to researching second language learner processes*, in «Annual Review of Applied Linguistics» 25, pp. 98–111.
- WILLIS J. (1996), *A Framework for Task-based learning*, Longman, London.
- WILLIS J., WILLIS D. (1988), *COBUILD Book 1*, Collins, London.
- WINKE P, GASS S., MYFORD C. (2013), *Raters' L2 background as a potential source of bias in rating oral performance*, in «Language Testing» 30, 2, pp. 231–52.
- WISEMAN C.S. (2012), *Rater effects: Ego engagement in rater decision-making*, in «Assessing Writing» 17, 3, pp. 150–73.
- WISNIEWSKI K. (2017a), *The Empirical Validity of the Common European Framework of Reference Scales. An Exemplary Study for the Vocabulary and Fluency Scales in a Language Testing Context*, in «Applied Linguistics» 39, 6, pp. 933–59.
- (2017b), *Empirical Learner Language and the Levels of the Common European Framework of Reference*, in «Language Learning» 67, S1, pp. 232–53.
- WOLFE-QUINTERO K., INAGAKI S., KIM H.Y. (1998), *Second Language Development in writing: measures of fluency, accuracy, and complexity*, University of Hawaii, Second Language Teaching and Curriculum Center, Honolulu.
- WOLFRAM W. (1985), *Variation in tense marking: A case for the obvious*, in «Language Learning» 35, pp. 229–53.



- WU Y. (1998), *What do tests of listening comprehension test? A retrospective study of EFL test-takers performing a multiple-choice task*, in «Language Testing» 15, pp. 21–44.
- YOUNG R. (1988), *Variation and the interlanguage hypothesis*, in «Studies in Second Language Acquisition» 10, pp. 281–302.
- (1991), *Variation in Interlanguage Morphology*, Peter Lang, New York.
- (1996), *Form–function relations in articles in English interlanguage*, in R. Bayley, D. Preston (eds.), *Second Language Acquisition and Linguistic Variation*, John Benjamins Publishing Company, Amsterdam, pp. 135–75.
- (1999), *Sociolinguistic approaches to SLA*, in «Annual Review of Applied Linguistics» 19, pp. 105–32.
- YUAN F., ELLIS R. (2003), *The effects of pre-task and on-line planning on fluency, complexity and accuracy in L2 monologic oral production*, in «Applied Linguistics» 24, pp. 1–27.
- ZUENGLER J. (1989), *Assessing an interactive paradigm: how accommodative should we be?*, in M. Eisenstein (ed.), *The Dynamic Interlanguage: Empirical studies in second language variation*, Plenum Press, New York.



AREE SCIENTIFICO–DISCIPLINARI

AREA 01 — Scienze matematiche e informatiche

AREA 02 — Scienze fisiche

AREA 03 — Scienze chimiche

AREA 04 — Scienze della terra

AREA 05 — Scienze biologiche

AREA 06 — Scienze mediche

AREA 07 — Scienze agrarie e veterinarie

AREA 08 — Ingegneria civile e architettura

AREA 09 — Ingegneria industriale e dell'informazione

AREA 10 — **Scienze dell'antichità, filologico–letterarie e storico–artistiche**

AREA 11 — Scienze storiche, filosofiche, pedagogiche e psicologiche

AREA 12 — Scienze giuridiche

AREA 13 — Scienze economiche e statistiche

AREA 14 — Scienze politiche e sociali

AREA 15 — Scienze teologico–religiose

AREA 16 — Discipline musicali

*Il catalogo delle pubblicazioni di Aracne editrice è su*

[www.aracneeditrice.it](http://www.aracneeditrice.it)

Finito di stampare nel mese di dicembre del 2019  
dalla tipografia «The Factory S.r.l.»  
00156 Roma – via Tiburtina, 912  
per conto della «Giacchino Onorati editore S.r.l. – unipersonale»  
di Canterano (RM)