



UNIVERSITÀ  
DEGLI STUDI  
DI PADOVA

## Università degli Studi di Padova

Dipartimento di Studi Linguistici e Letterari

Corso di Laurea Magistrale in Linguistica

Classe LM-39

Tesi di Laurea

### *Lo sviluppo della complessità lessicale nell'italiano L2: uno studio di caso*

Relatore  
Prof. Matteo Santipolo

Laureando  
Claudia Lasagni  
n° matr.1155577 / LMLIN

Anno Accademico 2017 / 2018



# INDICE

<i>Introduzione</i> .....	5
---------------------------	---

## CAPITOLO 1

<i>La complessità linguistica</i> .....	9
1.1 Definizione del costrutto.....	9
1.2 Tassonomia della complessità.....	12
1.2.1 Complessità assoluta e relativa.....	12
1.2.2 Complessità palese e nascosta .....	14
1.2.3 Localizzare la complessità linguistica.....	16
1.3 La complessità linguistica nelle ricerche SLA.....	18
1.3.1 Variabile indipendente o dipendente.....	18
1.3.2 CAF: complessità, accuratezza e fluenza.....	19
1.3.3 Gli obiettivi della ricerca SLA sulla complessità.....	21
1.4 Complessità e task pedagogico.....	22
1.4.1 Una definizione operativa di "task".....	22
1.4.2 La complessità del task nella ricerca linguistica.....	23
1.4.3 Il Triadic Componential Framework di Robinson.....	24
1.4.4 Il Limited Attentional Capacity Model di Skehan .....	26
1.4.5 Complessità o difficoltà?.....	28
1.5 Complessità e sviluppo .....	30
1.6 Complessità e competenza.....	32
1.7 Misurare la complessità linguistica.....	34
1.7.1 Un arduo compito.....	34
1.7.2 La complessità morfologica.....	36
1.7.3 La complessità sintattica .....	40

## **CAPITOLO 2**

<i>La complessità lessicale</i> .....	43
2.1 L'importanza del lessico.....	43
2.2 La complessità lessicale e le sue dimensioni.....	45
2.3 La diversità lessicale: come si misura? .....	50
2.4 La diversità lessicale: perché si misura?.....	55
2.5 I fattori che influenzano la complessità lessicale.....	57
2.6 Le domande di ricerca.....	60

## **CAPITOLO 3**

<i>Lo studio</i> .....	63
3.1 Le peculiarità dello studio.....	63
3.2 Il corpus di dati.....	66
3.3 I soggetti.....	68
3.4 I task.....	68
3.4.1 Task prevalentemente monologici.....	70
3.4.2 Task prevalentemente dialogici.....	71
3.5 La procedura di analisi.....	74
3.5.1 La trascrizione delle registrazioni .....	74
3.5.2 La preparazione dei dati all'analisi.....	75
3.5.3 L'analisi tramite il tool informatico.....	77

## **CAPITOLO 4**

<i>I risultati</i> .....	81
4.1 I risultati individuali.....	81
4.1.1 I risultati medi: un confronto fra i soggetti .....	90
4.2 I risultati per task.....	96
4.2.1 I risultati medi: un confronto fra i task.....	106

## **CAPITOLO 5**

<i>Considerazioni didattiche</i> .....	112
5.1 Il ruolo dell'interazione nello sviluppo dell'interlingua .....	113
5.2 La difficoltà interazionale dei task .....	117
5.3 L'uso dei task nell'insegnamento delle lingue.....	120
5.4 Verso una maggiore complessità linguistica.....	124
<i>Conclusioni</i> .....	127
<i>Appendice</i> .....	129
<i>Bibliografia</i> .....	137



## *Introduzione*

In questo elaborato si prenderà in esame la complessità linguistica, un costrutto che è stato assorbito da diversi settori della ricerca e con le intenzioni più svariate: cercare di dimostrare la superiorità di una lingua su un'altra, diagnosticare i disturbi del linguaggio, studiare il processo di apprendimento di una lingua straniera.

È con quest'ultimo obiettivo che il presente studio si propone di indagare lo sviluppo della complessità lessicale nell'interlingua di quattro adolescenti straniere, parlanti di diverse L1 e studentesse di italiano L2. Attraverso l'analisi di campioni di lingua orale, raccolti mediante compiti comunicativi (*task*), si cercherà di ricostruire come sia evoluta nell'arco di quattro anni la capacità delle apprendenti di produrre un linguaggio lessicalmente diversificato. Il materiale linguistico che analizzeremo è stato raccolto all'interno di un progetto di ricerca pluriennale sull'acquisizione di competenze linguistiche e pragmatiche, finanziato dal Ministero dell'Istruzione, dell'Università e della Ricerca, che ha visto la collaborazione di tre università italiane: l'Università di Bergamo in quanto capofila, nonché l'Università di Verona e l'Università di Modena e Reggio Emilia come unità locali.

Il corpus di dati oggetto di questa indagine presenta almeno tre caratteristiche che lo distinguono dalla maggior parte delle ricerche empiriche nel campo della *Second Language Acquisition* (SLA): in primo luogo, la sua estensione consente la realizzazione di studi longitudinali, da sempre trascurati rispetto a quelli trasversali, per ovvie ragioni di comodità, ma decisamente più significativi nell'illustrare i processi di acquisizione; in secondo luogo, la competenza linguistica delle apprendenti si colloca ad un livello avanzato, poco studiato rispetto alle prime fasi dell'apprendimento; infine, la produzione orale viene elicitata tramite lo svolgimento di cinque attività comunicative che, lungi dal pretendere di dare un quadro completo del sistema interlinguistico, offrono uno spettro di situazioni e atti comunicativi più variato di quanto solitamente si osservi negli studi empirici in questo ambito (Pallotti et al., 2010).

I task utilizzati si differenziano sotto diversi aspetti socio-psico-linguistici, come la richiesta di un dialogo o di un monologo, il tipo di interlocutore (adulto/pari, conosciuto/estraneo, nativo/non nativo), la presenza o meno del contatto visivo, o il coinvolgimento della sfera personale; queste variabili possono influenzare in molteplici modi la performance delle apprendenti, inoltre l'osservazione degli effetti della struttura del compito sulla produzione linguistica è da sempre un terreno privilegiato per l'elaborazione di strategie didattiche.

Lo scopo non secondario della ricerca è quello di validare una procedura informatizzata di misurazione della diversità lessicale dei testi, implementata attraverso un *tool* ideato da Brezina e Pallotti (2015). L'esponente che viene generato, lo *Standardized Type/Token Ratio*, vuole essere un contributo ad un'animata discussione che dagli anni '40 in poi ha coinvolto diverse discipline, dalla linguistica all'ecologia, circa i metodi di misurazione della diversità e i problemi ad essi correlati. L'indice in questione, adoperato nell'analisi quantitativa dei dati, può rivelarsi un valido strumento nelle mani dell'insegnante per valutare lo stato dell'interlingua dei singoli studenti e calibrare gli interventi didattici, oppure, nel caso di un corpus di dati più esteso, può indirizzare il ricercatore nella comprensione più ampia dei processi di apprendimento linguistico.

Dunque, nella prima parte dell'elaborato si tenterà di descrivere lo stato dell'arte sul tema della complessità linguistica, mentre in un secondo momento si parlerà delle sue applicazioni operative, per introdurre poi lo studio di caso.

Nel primo capitolo viene presentato il concetto di complessità linguistica e una breve rassegna tassonomica descrive le diverse prospettive teoriche da cui può essere considerato. Si ripercorrono le origini del costrutto e la sua rilevanza per il campo di ricerca SLA, specialmente in relazione a tre obiettivi: la descrizione della performance, l'analisi dello sviluppo linguistico e la valutazione della competenza. Si discute, infine, la possibilità di individuare delle misure matematiche della complessità contenuta nei livelli morfologico, lessicale e sintattico di un sistema linguistico o di un testo.

Il secondo capitolo tratta più nel dettaglio la complessità in campo lessicale, una nozione multidimensionale che viene resa operativa in termini di *diversity*. Si passeranno in rassegna varie proposte per la misurazione della diversità lessicale e si



valuteranno le potenziali applicazioni del costrutto nella ricerca linguistica e più nello specifico in quella acquisizionale. Infine, vengono presentate le domande di ricerca alle quali si tenterà di dare risposta durante lo studio.

Il terzo capitolo introduce lo studio di caso, descrivendo il protocollo di raccolta dati e la procedura utilizzata per l'analisi; mentre il quarto espone i risultati dell'indagine, operando un confronto sia fra i soggetti partecipanti che fra i task proposti.

L'ultimo capitolo contiene alcune implicazioni didattiche formulate sulla base dei risultati emersi dallo studio. Tali considerazioni riguardano essenzialmente l'utilizzo strategico dei task pedagogici per lo sviluppo della complessità nell'interlingua degli apprendenti di italiano L2, rientrando in un quadro di *Task-based language teaching*.



# CAPITOLO 1

## *La complessità linguistica*

In questo capitolo si tenterà di introdurre la “complessità linguistica”, un concetto difficile da definire in maniera univoca, a cui sono state attribuite molteplici etichette nel corso del tempo, che rispecchiano le altrettante angolazioni teoriche da cui è possibile osservare l’argomento.

Dopo una breve rassegna terminologica, che farà luce sulle sfumature di significato interne alla nozione di complessità, vedremo come essa sia oggetto di interesse da parte dei ricercatori, sia nel campo della tipologia linguistica, che in quello della *Second Language Acquisition*, a cui si allinea il presente studio.

Passeremo in analisi gli obiettivi centrali della ricerca SLA sulla complessità linguistica e sulla base di ognuno di essi costruiremo un approfondimento, al fine di indagare tre dimensioni strettamente legate al costrutto preso in esame: la performance, lo sviluppo e la competenza linguistica.

Concluderemo, poi, discutendo sulla possibilità di individuare una serie di metodi di misurazione matematici della complessità linguistica contenuta in un testo, che rendano esplicite le diverse facce della complessità, con particolare attenzione a tre livelli del sistema linguistico: la morfologia, la sintassi e il lessico.

### **1.1 Definizione del costrutto**

Gli apprendenti di una lingua seconda devono confrontarsi con l'arduo compito di imparare a padroneggiare la moltitudine di aspetti di cui si compone una lingua, come la sua fonologia, morfologia, sintassi, lessico e le regole pragmatiche ad essa associate. Al fine di cogliere queste molteplici sfaccettature del processo di apprendimento, i ricercatori hanno fatto riferimento alle lingue come a sistemi complessi e

multidimensionali, composti da vari sottosistemi interconnessi tra loro.

Nel corso dell'apprendimento, lo studente realizza i suoi obiettivi comunicativi facendo affidamento su un vero e proprio sistema linguistico caratterizzato da regole che in parte coincidono con quelle della L2<sup>1</sup>, in parte sono riconducibili alla L1 e in parte sono indipendenti da entrambe. Questo sistema, chiamato “interlingua”, è in continua evoluzione verso la L2, perciò aumenta gradualmente di complessità, come dimostrano, fra le altre cose, l'uso di una più vasta varietà di lessico, frasi più lunghe e un sistema morfologico più elaborato (Pallotti, 1998; Bettoni, 2001).

È facile intuire come il concetto di complessità sia di particolare rilevanza nel settore della ricerca acquisizionale sulle lingue seconde (SLA), ma si è dimostrato altrettanto allettante per molte discipline scientifiche.

La complessità quale elemento di studio è approdata negli anni '90 nell'ambito delle scienze naturali quali la biologia, la fisica e la chimica (Mitchell, 2009) e solo qualche anno più tardi la riflessione è maturata in ambito linguistico, dove la complessità si è rivelata un'importante variabile in diversi campi di studio, dalla ricerca applicata ai disturbi del linguaggio (Fergadiotis et al., 2013) alle discussioni teoriche sulla complessità come proprietà formale del linguaggio (Miestamo et al., 2008; Sampson, 2009).

Le ricerche sulla complessità sono state per lungo tempo al centro degli studi di tipologia, branca della linguistica che studia le similarità e le differenze fra le lingue (Ramat, 1987). Tale interesse nasceva dalla volontà dei linguisti di contrastare un'ideologia molto diffusa nell'Ottocento, secondo la quale la maggiore complessità (morfologica) di una lingua -con particolare riferimento alle lingue classiche- corrispondeva a una sorta di “superiorità” linguistica. Nell'ambiente culturale dell'epoca, la complessità era quindi presa come valore-pretesto per la creazione di gerarchie a sfondo eurocentrico (Sampson, 2009). *L'equi-complexity hypothesis* nasce come reazione a tale assunto e afferma che tutte le lingue sono ugualmente complesse, dal

---

1 Con l'espressione "seconda lingua" (abbreviato in L2) ci riferiamo a ogni lingua che viene appresa dopo la lingua materna (L1), cioè anche alle terze, quarte lingue ecc. Spesso la "seconda lingua" viene contrapposta alla "lingua straniera" (abbreviato in LS): la L2 è una lingua appresa nel paese dove essa viene parlata abitualmente (come l'italiano appreso in Italia dagli immigrati), invece la LS è una lingua appresa, tipicamente in contesti scolastici, in un paese dove non viene parlata abitualmente (come l'inglese appreso a scuola dagli studenti italiani) (Pallotti, 1998:13).

momento che devono soddisfare simili bisogni comunicativi. Queste necessità comunicative sono da considerarsi universali in quanto le comunità di parlanti sono composte da esseri umani dotati delle stesse capacità biologiche, mentali e culturali (Szmrecsany e Kortmann, 2012).

Tuttavia, le differenze di complessità interlinguistiche sono evidenti, anche solo in modo istintivo, a tutti i parlanti. I non-linguisti, infatti, tendono a classificare le lingue non native in una scala di complessità, che solitamente si basa sulla distanza percepita fra la propria lingua madre e quella straniera (Kusters, 2003). Questa valutazione è influenzata dalle esperienze individuali, oltre che dal proprio ambiente culturale, e generalmente si basa su aspetti linguistici come la fonetica (suoni difficili da pronunciare) o il tipo di alfabeto utilizzato nella scrittura, ma riguarda anche l'intuizione che alcune lingue possiedano di fatto componenti linguistici meno sofisticati in comparazione con i corrispondenti elementi di altre lingue (Shosted, 2006).

A questo proposito, senza smentire l'ipotesi dell'equi-complessità, la *negative correlation hypothesis* sostiene che la complessità dei diversi livelli linguistici (fonetica, morfologia, sintassi e lessico) vari all'interno di una stessa lingua, e che ognuno sopperisca quantitativamente alle mancanze dell'altro: ad esempio, è possibile che una lingua sia morfologicamente più complessa di un'altra, ma tale complessità sarà compensata laddove la stessa lingua sarà sintatticamente più semplice, dando luogo a un grado di complessità linguistica globale identico per tutte le lingue (Shosted, 2006).

Più recentemente, l'ipotesi dell'equi-complessità, un tempo ritenuta quasi dogmatica, è tornata al centro del dibattito. Ne è un iniziatore John McWhorter, che nel 2001 ha affermato che le grammatiche delle lingue creole, nate dal contatto tra due o più lingue, sono le più semplici del mondo, specialmente in relazione alle lingue più antiche. Secondo l'autore (2001), le lingue hanno bisogno di tempo per creare gradualmente strategie ed elementi espressivi funzionali alla comunicazione, perciò aumentano di complessità nel corso della loro storia.

Il discorso sulla complessità si è sviluppato copiosamente, come dimostrano le numerose pubblicazioni sull'argomento (Robinson, 2001; Dahl, 2004; Shosted 2006; Miestamo et al., 2008; Sampson et al., 2009; Bulté & Housen, 2012; Bisang, 2014; Pallotti, 2015a), nonostante ciò, tale concetto è ancora poveramente definito e manca

una continuità terminologica e di operazionalizzazione (Pallotti, 2009). La polisemia intrinseca al costrutto stesso ha dato vita a una vasta gamma di termini che riflettono diversi approcci all'argomento, alcuni dei quali verranno annoverati nel prossimo paragrafo.

## **1.2 Tassonomia della complessità**

### ***1.2.1 Complessità assoluta e relativa***

Sebbene la nozione di complessità linguistica possa sembrare intuitiva, si è rivelata estremamente ambigua nel campo delle scienze del linguaggio. Se prendiamo una definizione dizionariale (Devoto-Oli, 2000), il termine “complesso” è caratterizzato da una polisemia che si polarizza intorno a due significati principali: un oggetto è definito “complesso” quando risulta dall'unione di più parti, oppure quando è difficile da comprendere.

È sulla base di questa polisemia che poggia la distinzione tra *complessità assoluta* (chiamata anche “intrinseca”, “oggettiva” o “strutturale”) e *complessità relativa* (“estrinseca”, “soggettiva” o “relazionata all'utente”), molto utilizzata per definire il costrutto nel campo della tipologia linguistica e in quello dell'apprendimento delle lingue seconde. I più evidenti tentativi di circoscrivere il concetto di complessità sono da attribuire a Miestamo (2009) e Bulté e Housen (2012; 2015) che introducono questa prima tassonomia.

Per *complessità assoluta* si intende una proprietà intrinseca del sistema linguistico o di una data produzione linguistica, che si definisce in funzione al numero di elementi in un sistema linguistico o testo e il numero e il tipo di relazioni tra questi elementi. Ciò implica che una lingua, per essere definita complessa, dovrebbe essere composta da elementi altamente relazionati fra loro; al contrario di un sistema linguistico semplice, dove gli elementi linguistici tendono ad essere pochi e isolati (Miestamo, 2009; Bulté & Housen, 2012).

La *complessità relativa*, invece, si definisce in relazione al parlante e riguarda la

difficoltà con cui gli elementi linguistici vengono elaborati, in termini di sforzo cognitivo e tempistiche di sviluppo (Bulté & Housen, 2012; Miestamo, 2009). Si tratta, quindi, di una proprietà di quelle strutture linguistiche che vengono acquisite più tardi e con un costo cognitivo importante da parte del soggetto, sforzo che dipende in parte dal background individuale dell'apprendente (atteggiamento, motivazione, livello di competenza nella L2, influenza della L1) (Housen et al., 2012b).

L'approccio oggettivo della complessità assoluta la rende più propensa alla quantificazione e, di conseguenza, più appetibile per le ricerche SLA, che usano le misure della complessità per valutare la produzione degli apprendenti.

Tuttavia, se i termini “intrinseca” e “assoluta” conferiscono un forte senso di oggettività alla complessità, Pallotti (2015a) ricorda che questa proprietà linguistica è, inevitabilmente, frutto di una descrizione teorica e non può essere, in nessun caso, considerata neutrale o assoluta nel senso stretto del termine. Per questo, l'autore preferisce utilizzare il termine “strutturale” per riferirsi alla complessità assoluta; mentre, per quanto riguarda la complessità relativa, propone un'ulteriore scomposizione terminologica:

There are thus three main meanings of ‘complexity’ in linguistic research: 1. *Structural complexity*, a formal property of texts and linguistic systems having to do with the number of their elements and their relational patterns; 2. *Cognitive complexity*, having to do with the processing costs associated with linguistic structures; 3. *Developmental complexity*, i.e. the order in which linguistic structures emerge and are mastered in second (and, possibly, first) language acquisition (Pallotti, 2015a:118, corsivo mio).

La letteratura è ricca di sfumature terminologiche che riflettono la polisemia problematica di questo costrutto. Se dal punto di vista teorico le etichette vengono presentate come dicotomiche, dal punto di vista pratico si intersecano nell'esperienza. Infatti, gli elementi linguistici considerati complessi in senso assoluto/strutturale molto spesso risultano essere anche i più difficili da acquisire in termini di tempo e sforzo cognitivo. Questa correlazione, per quanto stretta possa rivelarsi, si instaura a livello empirico e non comporta che i termini teorici fin'ora elencati siano da considerarsi nomi diversi per uno stesso costrutto. Oltretutto, il fatto che siano costrutti differenti è

evidenziato dall'asimmetria della relazione: "*cognitive difficulty reflects rather than creates complexity*" (Rescher, 1998:17). Ciò significa che la complessità strutturale sarebbe spesso responsabile della complessità cognitiva: da qui l'esigenza di designare in modo diverso la causa e l'effetto (Pallotti, 2015b).

Non in tutti i casi si verifica la suddetta relazione, ad esempio, un Sudoku con 18 cifre è strutturalmente meno complesso ma cognitivamente più complesso, o difficile, di uno con 25 cifre. Allo stesso modo, in certi casi gli elementi linguistici strutturalmente più complessi possono risultare più semplici da produrre rispetto agli altri e l'efficienza comunicativa viene raggiunta tramite una maggiore complessità (Pallotti, 2015a).

Sulla scia di questa deduzione introduciamo un'altra opposizione teorica nella tassonomia della complessità, quella tra complessità palese e nascosta.

### ***1.2.2 Complessità palese e nascosta***

Secondo Bisang (2014), esistono due esigenze comunicative che influenzano la produzione linguistica dei parlanti: il bisogno di esplicitazione e il bisogno di economia. Una comunicazione tra due parlanti, per essere efficace, deve bilanciare il desiderio di comunicare quanta più informazione possibile e il rischio di esprimere un'informazione ridondante perché deducibile dal contesto, o un carico eccessivo di informazioni, fattore che può innalzare il rischio di incomprensioni.

I due bisogni fondamentali individuati da Bisang generano due tipi di complessità: una *complessità palese* e una *complessità nascosta* (Bisang, 2014)<sup>2</sup>.

Il bisogno di esplicitazione stimola il parlante o l'apprendente a esprimere dei significati tramite elementi grammaticali obbligatori e adatti al contesto comunicativo, in modo che gli enunciati prodotti siano ricchi di elementi correlati fra loro. Poiché la complessità linguistica che ne deriva è sensorialmente percepibile, è detta "palese".

Il bisogno di economia, invece, spinge il parlante ad evitare gli elementi grammaticali che veicolano un significato già deducibile dal contesto, producendo così un enunciato ricco di elementi funzionali il cui significato deve essere dedotto

---

2 Libera traduzione dei due corrispondenti nel testo originale: "overt complexity" e "hidden complexity" (Bisang, 2014:127).



pragmaticamente (Bisang, 2014). Dal punto di vista della complessità palese, quindi, il bisogno di economia tende a semplificare la struttura grammaticale dell'enunciato, ma lo rende più complesso dal punto di vista della complessità nascosta, in quanto servirà un grande lavoro inferenziale da parte del ricevente per decodificarlo.

Secondo Bisang (2014), la complessità ha quindi due facce: una esplicita, che si manifesta tramite la grammatica, e un'altra nascosta, sottesa nelle deduzioni pragmatiche attuate dal destinatario dell'enunciato.

Questo tipo di approccio alla complessità ha delle implicazioni sul discorso tipologico accennato in precedenza (vedi paragrafo 1.1). Come abbiamo già detto, i linguisti vedono la complessità palese come risultato di un processo storico di maturazione della lingua, per questo motivo McWhorter (2001) sostiene che le grammatiche delle giovani lingue creole siano anche le più semplici. Questa visione dà per scontata la coevoluzione di forma e significato, per cui ad un alto grado di grammaticalizzazione di una lingua sono associate tipicamente l'inflessione e la presenza di elementi grammaticali obbligatoriamente espressi. Secondo Bisang (2009), questo presupposto è problematico. L'autore apporta l'esempio di alcune lingue asiatiche, come il cinese e il vietnamita, in cui la mancanza di elementi grammaticali obbligatori e la presenza di elementi multifunzionali consentono ai parlanti una maggiore economia linguistica, portandoli a produrre enunciati che richiedono continue inferenze pragmatiche, quindi caratterizzati da un alto grado di complessità nascosta. Se si ammettesse come universale la coevoluzione di forma e significato, sarebbe impossibile considerare la complessità nascosta come risultato di un lento sviluppo storico, eppure queste lingue dimostrano che ciò è effettivamente possibile. Se, al contrario, si intendono le strutture grammaticali come frutto di forze concorrenti -di economia e di esplicitazione- allora si può intuire come la complessità palese e quella nascosta non si escludano reciprocamente, ma ci sia sempre un po' dell'una e un po' dell'altra, anche se in misura diversa, dentro ogni lingua (Bisang, 2009).

### 1.2.3 Localizzare la complessità linguistica

Se parte dell'eterogeneità dei precedenti approcci alla complessità linguistica si può ridurre alla macro-distinzione tra complessità assoluta e relativa, rimane estremamente variabile il luogo di applicazione di tale costrutto.

Alcuni studi hanno analizzato la complessità come proprietà formale del sistema linguistico nel suo insieme, mentre altri hanno limitato l'indagine ad un particolare aspetto di un'area circoscritta del sistema linguistico. Questa divergenza metodologica viene sintetizzata nell'opposizione tra la complessità considerata a un livello globale o locale, distinzione che serve al ricercatore per circoscrivere lo scopo e l'oggetto di studio, dal generale allo specifico.

Con *complessità globale* intendiamo la complessità del sistema linguistico nella sua totalità, un costrutto la cui operazionalizzazione richiederebbe l'identificazione di tutti i sotto-domini del sistema linguistico e una corretta misurazione della loro complessità.

La *complessità locale*, invece, riguarda i diversi ambiti del sistema linguistico, come il livello fonologico, morfologico, sintattico e lessicale, ciascuno dei quali può essere suddiviso in ulteriori sotto-domini (ad esempio: morfologia flessiva vs derivazionale; fonologia segmentale vs soprasegmentale) (Bulté & Housen, 2012; Szmrecsanyi & Kortmann, 2012).

Un'altra prospettiva può essere quella di distinguere tra due dimensioni: la "ampiezza" della complessità e la "profondità" della stessa. A questo proposito Bulté e Housen (2012:25) propongono una terminologia alternativa a quella vista in precedenza: la *complessità sistemica* è definita come "*the number, range, variety or diversity*" degli elementi entro un certo dominio o sistema linguistico, perciò si riferisce all'ampiezza della complessità; invece la *complessità strutturale* riguarda la composizionalità dei singoli elementi linguistici ad un livello locale, quindi ha a che fare con la profondità.

Infine, gli autori (2012:25) proseguono la loro rassegna distinguendo tra *complessità formale* e *complessità funzionale*: la prima coinvolge il numero di componenti discrete di una forma linguistica, mentre la seconda si riferisce al numero dei significati e funzioni rivestiti dalla stessa struttura linguistica.

Quest'ultima distinzione può essere applicata sia a livello sistemico che strutturale. Se prendiamo ad esempio il sistema morfologico, la sua complessità sistemica deriverà in parte dal numero di morfemi flessivi e derivazionali (complessità formale), che consentono di esprimere una serie di distinzioni funzionali come il singolare, il plurale e il duale (complessità funzionale). A livello della complessità strutturale, invece, uno specifico morfema rappresenta una data funzione -il plurale- (complessità funzionale) e può essere realizzato da diversi allomorfi dipendenti dal contesto (complessità formale).

La letteratura sull'argomento non è sempre coerente, poiché diversi autori si basano su una terminologia differente o non applicano le stesse distinzioni teoriche. Ad esempio, l'etichetta di "complessità sistemica" utilizzata da Bulté e Housen (2012:25) è ripresa da Pallotti (2015a:120) come parte di un'altra coppia di termini riguardanti la localizzazione della complessità, che vede da un lato la complessità della *langue* Saussuriana -*complessità sistemica* appunto- e dall'altro la complessità di una determinata realizzazione concreta (*parole*), detta *complessità testuale*.

La rassegna tassonomica qui proposta non è esaustiva, ma vuole solo rendere l'idea della considerevole varietà di approcci possibili al tema della complessità linguistica, che si traduce nella letteratura come una fitta e intricata trama di termini e rispettivi costrutti teorici.

In questa sede, sulla base delle distinzioni teoriche adottate, considereremo dunque la complessità linguistica come una proprietà formale associata ad unità linguistiche "tangibili" a livello della morfologia, sintassi o lessico. Questa proprietà può essere definita in senso assoluto, come una caratteristica intrinseca alla lingua, o in senso relativo, in relazione ai singoli parlanti. Può essere colta nella forma palese della grammatica, o nascosta nella pragmatica necessaria a decodificare i messaggi. Inoltre, può essere studiata a livello globale del sistema linguistico o locale dei suoi sotto-domini interni e può essere rappresentata in termini di ampiezza o profondità, di forma o di funzione.

Nel presente studio empirico ci proponiamo di analizzare la complessità linguistica da un punto di vista *assoluto* e *formale*, in termini di *ampiezza*, così come

appare *palesemente* all'interno della produzione *testuale* degli apprendenti, e limitando la nostra indagine al livello *locale* del lessico.

### **1.3 La complessità linguistica nelle ricerche SLA**

Abbiamo visto come la complessità linguistica sia stata oggetto per molto tempo di studi tipologici; ora invece assumiamo la prospettiva delle ricerche sull'acquisizione di lingue seconde, un campo in cui la complessità si fa una materia tanto importante quanto ambigua, a causa della natura multiforme e imprevedibile delle interlingue degli apprendenti.

#### ***1.3.1 Variabile indipendente o dipendente***

Come affermano Bulté e Housen (2012:21), gli studi sull'acquisizione di lingue seconde hanno indagato la complessità linguistica sia come variabile indipendente, che come variabile dipendente.

La prima linea di ricerche si concentra su come la complessità di certe strutture della lingua target (lingua d'arrivo) possa influenzarne la facilità di apprendimento o insegnamento. A questo scopo risulta particolarmente importante il concetto di "difficoltà", sia in relazione alle capacità del soggetto, sia alle caratteristiche della lingua da apprendere (come la sua opacità).

Nella seconda tipologia di ricerche, la complessità è investigata in quanto variabile dipendente, spesso usata per descrivere la performance nella L2 e indicare il livello di competenza degli apprendenti. In questo caso, la complessità è una proprietà della produzione orale o scritta che viene determinata mediante misure quantitative, con il fine di dimostrare l'effetto di qualche altra variabile indipendente sulle realizzazioni linguistiche nella L2. Gli studi condotti in questo modo hanno evidenziato come la complessità delle realizzazioni superficiali sia influenzata da una vasta gamma di fattori, quali ad esempio l'età dei soggetti, il loro livello di competenza, il contesto di apprendimento, la modalità della produzione e le caratteristiche del compito (Bulté &

Housen, 2012).

### ***1.3.2 CAF: complessità, accuratezza e fluenza***

Nel campo delle ricerche SLA, la complessità è spesso considerata insieme ad altre due dimensioni chiave della conoscenza ed elaborazione del linguaggio: l'accuratezza e la fluenza.

Questa triade di costrutti, che insieme vengono definiti dalla sigla CAF, compare per la prima volta in letteratura negli anni '80, quando i ricercatori propongono una prima dicotomia *accuracy vs fluency*, per investigare lo sviluppo della competenza di parlanti L2 nel contesto scolastico. A questo proposito, Brumfit (1984) opera la distinzione tra le attività mirate allo sviluppo della fluenza, le quali aspirano a sviluppare un tipo di interazione spontanea, che si avvicini a quella praticata nella vita quotidiana dai parlanti altamente competenti, e le attività orientate allo sviluppo dell'accuratezza, ovvero quelle che si concentrano sulla produzione controllata di una forma linguistica che sia grammaticalmente corretta nella L2 (Brumfit 1984 in Housen & Kuiken, 2009:461).

Negli anni '90 viene aggiunta anche la terza componente, la complessità, seguendo il modello di Skehan (1989), che per la prima volta indica la CAF come l'insieme delle tre dimensioni fondamentali della competenza L2.

È a quel punto che i costrutti ricevono la loro tradizionale definizione operativa, che si estende fino ad oggi. La *complessità* è stata comunemente caratterizzata come la misura in cui la produzione elicitata durante un compito comunicativo è elaborata e varia (Ellis, 2003:340); l'*accuratezza* è stata definita come l'abilità di produrre un discorso conforme alle norme della lingua target (Wolfe-Quintero et al., 1998:4); mentre la *fluenza* è la misura in cui la produzione L2 presenta pause, esitazioni e riformulazioni (Ellis, 2003:342), le quali possono segnalare una velocità di elaborazione del linguaggio più o meno *native-like*.

Molti ricercatori sostengono che la performance e la competenza L2 siano costrutti multidimensionali in natura, le cui dimensioni principali possono essere

adeguatamente rappresentate dalle nozioni di complessità, accuratezza e fluenza. Di conseguenza, la CAF è al centro di numerosi studi di linguistica applicata, dove viene utilizzata per descrivere la performance orale o scritta degli apprendenti, o come indicatore della competenza sottostante tale performance, o ancora per misurare il progresso nell'apprendimento linguistico (Housen & Kuiken, 2009). Nella maggior parte dei casi, le dimensioni CAF vengono adottate come variabili dipendenti, per misurare la variazione circa variabili indipendenti, quali il livello di acquisizione o le caratteristiche del compito comunicativo (Pallotti, 2009).

Vari contributi evidenziano come la complessità e l'accuratezza siano strettamente legate allo stato corrente della conoscenza linguistica dell'apprendente, e quindi alla sua competenza. In questo senso, la complessità rappresenta "*the scope of expanding or restructured second language knowledge*" (Wolfe-Quintero et al., 1998:4) e l'accuratezza "*the conformity of second language knowledge to target language norms*" (Wolfe-Quintero et al., 1998:4).

Al contrario, la fluenza è relazionata al controllo che l'apprendente ha sulla sua conoscenza della L2, cioè la misura in cui egli ha automatizzato l'accesso alle rappresentazioni linguistiche. Questo controllo si riflette nella velocità e nella facilità con cui il parlante ha accesso all'informazione L2 per comunicare dei significati in tempo reale (Wolfe-Quintero et al., 1998:4).

In conclusione, l'apprendimento di una qualsiasi struttura linguistica complessa coinvolge le tre dimensioni della CAF: l'esposizione ripetuta a una lingua accurata (*target-like*) e complessa serve a sviluppare la competenza, cosicché la conoscenza linguistica viene proceduralizzata e l'uso della lingua risulta fluente (Towell, 2012).

Tuttavia, è ampiamente rifiutata l'idea semplicistica che vede lo sviluppo della lingua seconda come una crescita lineare lungo le tre dimensioni della CAF, per cui gli apprendenti diventerebbero progressivamente più competenti sui tre assi, complessità-accuratezza-fluenza, simultaneamente.

Sebbene concordiamo che un quadro completo dello sviluppo della competenza linguistica nella lingua seconda debba comprendere complessità, accuratezza e fluenza, restringiamo l'analisi del presente elaborato al solo campo della complessità, il più problematico dei tre costrutti (Pallotti, 2009:5).

### ***1.3.3 Gli obiettivi della ricerca SLA sulla complessità***

Ortega (2012) identifica tre obiettivi centrali della ricerca sulla complessità linguistica in quanto variabile dipendente: la descrizione della performance, l'analisi comparativa dello sviluppo e la valutazione della competenza.

Il primo obiettivo riguarda come la produzione degli apprendenti è influenzata dalle variabili contestuali. In particolare, come approfondiremo nel prossimo paragrafo (1.4), risultano molto rilevanti le caratteristiche del compito comunicativo, anche detto *task*, somministrato all'apprendente (Robinson, 2007b, 2011; Skehan, 2009a). La quantità di complessità linguistica ritrovata nei dati prodotti, poi, diventa un importante criterio per valutare la performance in determinati task linguistici.

Questo filone di ricerca ha evidenziato il fatto che le realizzazioni superficiali della complessità possano essere influenzate dai fattori più svariati, come appunto il tipo di task (dialogico, narrativo ecc...), la quantità di tempo a disposizione per pianificare la produzione linguistica, o perfino le preferenze individuali nell'uso di un linguaggio più o meno complesso (Pallotti, 2009).

Il secondo obiettivo è, nel parere di Ortega (2012), il meno indagato dei tre e tratta la relazione esistente tra la complessità linguistica e le fasi di sviluppo dell'interlingua degli apprendenti. L'autrice sottolinea la necessità di misure quanto più oggettive della complessità che possano tracciare la traiettoria dell'acquisizione linguistica della L2, una ricerca cominciata negli anni '70 e ancora oggi in corso (Ortega, 2012).

Studi recenti hanno messo in luce la natura multifaccettata della complessità e dello sviluppo della lingua seconda, dimostrando che specifiche forme di complessità possono crescere, decrescere o stagnare nel corso dello sviluppo, entrando in relazione di competizione o di supporto con altre forme di complessità. Di conseguenza, anche se la tendenza di un gruppo dovesse indicare una generale crescita della complessità linguistica, sarebbe inesatto presumere che tutti gli aspetti della complessità crescano costantemente, o che tutti gli individui seguano strettamente questa traiettoria (Larsen-

Freeman, 2006).

Infine, l'ultimo obiettivo concerne il modo in cui la complessità linguistica può essere efficacemente utilizzata per valutare la competenza dell'apprendente (livello della *langue*) sulla base della sua performance (livello della *parole*) (Ortega, 2012).

Precisiamo che la nozione di competenza non si limita necessariamente alle caratteristiche del sistema linguistico in sé, ma include anche l'abilità del parlante di raggiungere efficientemente gli obiettivi comunicativi facendo affidamento sul proprio sistema linguistico mentale. Per questo, come vedremo più avanti (paragrafo 1.6), un linguaggio più complesso non indica obbligatoriamente una maggiore competenza o una fase più avanzata dello sviluppo linguistico (Ortega, 2003).

## **1.4 Complessità e task pedagogico**

### ***1.4.1 Una definizione operativa di "task"***

Come accennato nel paragrafo precedente, uno degli obiettivi centrali della ricerca sull'acquisizione di lingue seconde, con particolare riferimento al campo della complessità linguistica, è la descrizione della performance, che comporta un'analisi attenta di tutti i fattori che possono influenzare la produzione linguistica.

Tra le variabili contestuali con questo potere, il *task* è uno dei costrutti più indagati e ci sembra opportuno fornirne una breve definizione operativa, vista la sua centralità nel presente studio.

Generalmente, per *task* si intende una qualsiasi attività appartenente al mondo reale che le persone svolgono quotidianamente, come leggere il giornale, rispondere alle telefonate, o partecipare a un meeting di lavoro. Nelle parole di Michael Long (1985), ricadono sotto la definizione di *task* “*The hundred and one things people do in everyday life, at work, at play, in between. 'Tasks' are the things people will tell you they do if you ask them and they are not applied linguists*” (Long, 1985:89).

Il *task* assume connotati più specifici se considerato nell'ambito della linguistica applicata, come attività pedagogica. In questo senso, un contributo più formale viene da



Peter Skehan (1998):

A task is an activity in which: meaning is primary; there is some communication problem to solve; there is some sort of relationship to comparable real-world activities; task completion has some priority; the assessment of the task is in terms of outcome (Skehan, 1998:95).

Il task pedagogico, quindi, anche se svolto in un contesto scolastico, è un'attività molto simile a quelle della vita reale e quotidiana, che richiede agli apprendenti di utilizzare la L2 con particolare enfasi sulla trasmissione di un significato e il raggiungimento di un obiettivo di natura extra-linguistica (Bygate, Skehan & Swain, 2001).

Da circa vent'anni questo costrutto è utilizzato nelle ricerche sull'acquisizione di lingue seconde per raccogliere la produzione linguistica, nonché come strumento operativo in didattica<sup>3</sup>. Nel contesto del presente elaborato, il task è importante non solo in quanto strumento al servizio del ricercatore per elicitare la produzione dell'apprendente di L2, ma anche perché permette di osservare l'evoluzione della produzione linguistica al variare della complessità del compito comunicativo svolto.

#### ***1.4.2 La complessità del task nella ricerca linguistica***

La complessità, o difficoltà, che caratterizza il task influenza in svariati modi la complessità della performance linguistica, per questo motivo ha ricevuto una notevole attenzione da parte dei ricercatori nel campo della *Second Language Acquisition* (SLA), *Task-Based Language Teaching* (TBLT) e *Task-Based Language Assessment* (TBLA).

Nel primo campo di ricerca, la questione ricorrente è determinare se diversi tipi e livelli di complessità del task sono relazionati a una variazione sistematica della performance L2. Il tema principale nel TBLT riguarda, invece, i criteri per classificare i task, in modo da assegnare agli studenti attività che siano sufficientemente impegnative da promuovere l'apprendimento, ma mai eccessivamente gravose da diventare frustranti.

3 Dagli anni '90 ad oggi proliferano gli studi in merito al *Task-Based Language Teaching* (TBLT), il filone di ricerca che applica il task all'insegnamento delle lingue seconde. A questo proposito si veda Long sul *Focus on Form* (Long, 2015).

Infine, nel campo del TBLA, l'obiettivo è scoprire quali sono i task più appropriati per valutare gli apprendenti a diversi livelli di competenza, o se esiste una correlazione sistematica tra la complessità del task e i punteggi dei test (Pallotti, 2017).

Nonostante la grande mole di studi riguardanti la "complessità" o "difficoltà" dei task, c'è ancora molta controversia circa il significato di questi due costrutti, che vengono spesso utilizzati in modo intercambiabile o con accezioni diverse da parte di autori diversi. Nei prossimi paragrafi riportiamo una breve sintesi dei modelli più influenti nella ricerca SLA che cercano di dar conto della relazione tra la complessità del task e le varie dimensioni della performance linguistica, tra cui accuratezza, fluenza e complessità.

### ***1.4.3 Il Triadic Componential Framework di Robinson***

Nei primi anni 2000 Robinson (2001, 2007b) propone un modello basato su tre macro-fattori per classificare i task: *task complexity*, *task conditions* e *task difficulty*.

I primi due hanno a che fare con le caratteristiche intrinseche del task, che influenzano rispettivamente il piano cognitivo e quello interazionale; mentre il terzo riguarda il rapporto tra il task stesso e l'individuo che lo svolge.

La *task complexity* viene definita come "*the result of the attentional, memory, reasoning, and other information processing demands imposed by the structure of the task on the language learner*" (Robinson, 2001:29).

Come si vede nella tabella 1, ogni macro-fattore raggruppa poi una serie di aspetti più specifici. La prima colonna contiene i fattori che contribuiscono a rendere il task cognitivamente più complesso: alcuni di questi, se presenti nel compito, inducono i parlanti a produzioni linguistiche più complesse (*resource-directing variables*), mentre altri comportano l'utilizzo di maggiori risorse cognitive senza però necessariamente indurre l'apprendente a produrre un output linguistico più complesso (*resource-dispersing variables*). In questo primo senso, la complessità riguarda quindi tutto ciò che rende un task strutturalmente più impegnativo dal punto di vista cognitivo per i partecipanti in generale, senza tener conto delle differenze individuali.

<b>Task complexity</b>	<b>Task conditions</b>	<b>Task difficulty</b>
<i>resource-directing variables making cognitive/conceptual demands</i>  ± here and now ± few elements ± spatial reasoning ± causal reasoning ± intentional reasoning ± perspective-taking	<i>participation variables making interactional demands</i>  ± open solution ± one way low ± convergent solution ± few participants ± few contributions needed ± negotiation not needed	<i>ability variables and task relevant resource differentials</i>  h/l working memory h/l reasoning h/l task switching h/l aptitude h/l ield independence h/l mind reading
<i>resource-dispersing variables making performative/procedural demands</i>  ± planning time ± prior knowledge ± single task ± task structure ± few steps ± independency of steps	<i>participant variables making interactant demands</i>  ± same proficiency ± same gender ± familiar ± shared content knowledge ± equal status and role ± shared cultural knowledge	<i>affective variables and task relevant state-trait differentials</i>  h/l openness h/l control of emotion h/l task motivation h/l processing anxiety h/l willingness to communicate h/l self-efficacy

Tab. 1 – Triadic Componential Framework (Robinson, 2011:6)

Tuttavia, il successo nel completare un task dipende anche dalla quantità di risorse che l'apprendente in questione può e decide di impiegare in esso. Robinson (2001:31) definisce *task difficulty* la prospettiva che tiene conto dell'incontro tra le caratteristiche del task e alcuni tratti personali dello studente, che possono influenzare le risorse a sua disposizione. I fattori che contribuiscono all'aumento o alla diminuzione delle risorse disponibili al parlante per affrontare il task sono di due tipi (terza colonna della tabella 1): le *affective variables*, come la sicurezza, la motivazione o la capacità di controllare l'ansia sono più mutabili e temporanee; mentre le *ability variables*, quali ad esempio lo stile cognitivo, l'indipendenza o l'ampiezza della memoria di lavoro, sono fattori più stabili e perciò diagnosticabili prima dell'implementazione di un determinato

task (Robinson, 2001:32).

La terza prospettiva, quella delle *task conditions*, non riguarda più né il task né il soggetto considerati singolarmente, ma piuttosto le caratteristiche della partecipazione, cioè le condizioni di interazione. I fattori che rendono il task più impegnativo dal punto di vista interazionale possono riguardare l'attività in sé (*participation variables*), come ad esempio la direzione del flusso di informazioni e il tipo di soluzione prevista, oppure le caratteristiche dei partecipanti che concretamente la svolgono (*participant variables*), come il sesso, la familiarità o il ruolo assunto nel task (Robinson, 2001:32).

#### ***1.4.4 Il Limited Attentional Capacity Model di Skehan***

Ugualmente importante è il contributo di Skehan (1998; 2001;2009a), che elaborando il modello di *Limited Attentional Capacity* tenta di prevedere l'influenza della complessità del task sulla performance linguistica, declinata nelle sue dimensioni: complessità, accuratezza e fluenza.

A differenza di Robinson, Skehan e Foster (2001) si riferiscono alla complessità del task utilizzando i termini "complexity" e "difficulty" come sinonimi. Le dimensioni proposte per descrivere tale complessità sono circoscritte entro tre macro-aree (Skehan & Foster, 2001:194-195):

##### **Code complexity**

linguistic complexity and variety  
vocabulary load and variety

##### **Cognitive complexity**

cognitive familiarity  
familiarity of topic  
familiarity of discourse genre  
familiarity of task  
cognitive processing  
information organization  
amount of 'computation'  
clarity of information given  
sufficiency of information given

### **Communicative stress**

time pressure

scale (number of participants, lengths of texts used)

modality

stakes

opportunity for control

La prima macro-area concerne il tipo di strutture linguistiche che il task utilizza/richiede di impiegare, la seconda il tipo di operazioni cognitive da svolgere, la terza le condizioni contestuali che possono avere un effetto sulla performance, inducendo maggiore o minore stress nel partecipante. Entro quest'ultima sono presenti aspetti della complessità non attribuibili strettamente al task, ma piuttosto riguardanti la dimensione dell'interazione tra i partecipanti. Infatti, variabili quali il numero di persone coinvolte o il tempo a disposizione e, quindi, la pressione percepita dall'apprendente, influenzano non solo le tre dimensioni della produzione -complessità accuratezza e fluenza-, ma anche i momenti di negoziazione o, più in generale, l'aspetto interazionale del task (Skehan & Foster, 2001).

Il *Limited Attentional Capacity Model* è stato spesso messo a confronto con il *Multiple Attentional Resource Model* di Robinson (2007b). In maniera molto semplificata, entrambi i modelli tentano di spiegare come le condizioni del task, quali ad esempio la complessità cognitiva che lo caratterizza o la quantità di tempo a disposizione dell'apprendente per pianificare la produzione, influenzino la performance linguistica e nel farlo giungono a conclusioni contrastanti: il modello di Skehan teorizza che la quantità limitata di risorse cognitive a disposizione degli apprendenti li costringa a destinare tali risorse selettivamente alla complessità, all'accuratezza o alla fluenza, alle spese delle altre dimensioni della performance (*trade-off hypothesis*); al contrario, il modello proposto da Robinson, come suggerito dall'aggettivo "multiple", avanza l'ipotesi che gli apprendenti siano in grado di prestare attenzione a queste dimensioni della performance simultaneamente, senza necessariamente dare la priorità ad una di esse.

Come si può facilmente immaginare, le conseguenze pedagogiche tratte da questi due modelli sono differenti: mentre Skehan sostiene che nella realizzazione di un task "prioritization and predisposition (or both) seem to orient performance toward one

(or two) of the three areas of accuracy, fluency and complexity" (Skehan, 1998), Robinson afferma che maggiore è il carico cognitivo associato a un task in termini di contenuto, maggiori saranno la complessità e l'accuratezza ricercate dall'apprendente nella produzione (Robinson, 2001). A questo proposito, avremo modo di approfondire la *cognition hypothesis* di Robinson nell'ultimo capitolo di questo elaborato.

I due modelli presentati riguardano prevalentemente le dimensioni CAF della performance, tuttavia Skehan si interessa anche ai ruoli dei diversi domini della complessità (lessicale, morfologica, sintattica) all'interno della task performance. In particolare, Skehan (2009a) rileva un dato interessante: "*for native speakers, indeed, complexity may be more unidimensional in that lexical complexity and structural [syntactic] complexity go hand in hand, but for non-native speakers, the two areas do not seem to be integrated so well*" (Skehan, 2009a:528). Più precisamente, Skehan e Foster (2008, in Skehan 2009b) trovano correlazioni positive tra la *word difficulty*, presentata come una dimensione della complessità lessicale, e la complessità sintattica nei parlanti nativi, mentre la stessa correlazione risulta negativa nella produzione dei parlanti non-nativi, in task con svariati livelli di complessità. Queste potenziali compensazioni tra la complessità lessicale e quella sintattica nei parlanti L2 sono interessanti in quanto potrebbero chiarire come gli apprendenti assegnano le loro risorse cognitive ai vari domini del sistema linguistico (morfologia, sintassi, lessico), oltre che alle già indagate dimensioni della performance (CAF).

#### ***1.4.5 Complessità o difficoltà?***

Anche limitandoci ai modelli più diffusi abbiamo reso l'impressione dell'ambiguità che spesso circonda i termini "complessità" e "difficoltà" in campo scientifico, specialmente quando applicati al concetto di task: "*it is hard to imagine any other construct could equal task complexity in terms of the level of ambiguity and internal inconsistency achieved over the years*" (Gill & Hicks, 2006:2). Cercheremo quindi di fare un po' di chiarezza e adottare una posizione teorica per il presente studio.

Campbell (1988) elenca tre diverse prospettive interne al concetto generico di complessità del task:

- the psychological experience perspective
- the task-person interaction perspective
- the objective characteristics perspective

A questa scomposizione teorica si possono sovrapporre, poi, diverse scelte terminologiche. Abbiamo visto (paragrafo 1.4.3) come, nel frangente della ricerca SLA, Robinson (2011) proponga di assegnare il termine *complexity* alla prima prospettiva, riferendosi agli aspetti che rendono un task cognitivamente più impegnativo per qualunque individuo, mentre chiama *difficulty* il secondo punto di vista, cioè quello risultante dall'incontro del task con alcune caratteristiche personali dell'individuo.

Come afferma Pallotti (2017), la terza prospettiva, che ha a che fare con la struttura interna di un task, non è stata sistematicamente isolata dalle altre nel campo della ricerca SLA, dove spesso viene incorporata alla prima. Ad esempio, il "numero di elementi", fattore che riguarda la composizione strutturale di un task, appare nel modello di Robinson a fianco di "ragionamento causale, spaziale e intenzionale" o "tempo di pianificazione", tutti aspetti inerenti ai processi cognitivi svolti dal partecipante.

Adotteremo qui la visione di Pallotti (2017), che propone di circoscrivere l'uso del termine "complessità" esclusivamente al livello oggettivo e strutturale, ovvero il terzo significato nell'elenco di Campbell. Le prime due prospettive, invece, vengono definite dal termine "difficoltà" e differenziate tra loro rispettivamente come *difficoltà interindividuale*, inerente al task, e *difficoltà individuale*, inerente al soggetto che compie il task. Ecco un esempio immediato:

Climbing Mount Everest is *objectively more difficult* (i.e. more difficult for everyone) than climbing the Ben Nevis, even though climbing both mountains is for me *subjectively more difficult* than for an experienced mountaineer (Pallotti, 2009:6).

La difficoltà interindividuale corrisponde, dunque, a quella che Robinson chiama complessità. Il termine "difficoltà", tuttavia, ha il vantaggio di essere meno ambiguo poiché rappresenta meglio la vicinanza tra i due concetti di difficoltà per il singolo soggetto (individuale) e per tutti indistintamente (interindividuale), piuttosto che una tassonomia che utilizza due nomi completamente diversi, quali "complessità" e "difficoltà".

Inoltre, questa scelta si allinea con l'idea, presente in altri lavori, che vede la complessità come una proprietà assoluta, mentre intende la difficoltà come "costo" o "impegno" richiesti da un'attività (Dahl, 2009).

## 1.5 Complessità e sviluppo

Nella rassegna stilata da Ortega (vedi paragrafo 1.3.3), il secondo focus della ricerca SLA sulla complessità linguistica riguarda l'evoluzione di quest'ultima nel tempo, attraverso le fasi di sviluppo della L2.

Questo punto è spesso discusso nell'ambito della Teoria dei Sistemi Dinamici (DST)<sup>4</sup> applicata al campo d'indagine dell'acquisizione di lingue seconde. I sostenitori dell'approccio DST alla ricerca SLA considerano il sistema linguistico in sviluppo come un costrutto multidimensionale, in cui la crescita delle diverse parti dipende dalle risorse disponibili (Spoleman & Verspoor, 2010). La relazione tra queste parti può essere di tipo solidale, quando due aspetti del sistema linguistico crescono di pari passo nel tempo (*connected growers*), oppure di tipo competitivo, se la crescita di una componente risulta nel calo di un'altra (*trade-offs*). Queste relazioni, poi, possono essere dinamiche e variabili nel corso dello sviluppo linguistico, così, ad esempio, lo sviluppo lessicale e quello grammaticale possono essere solidali all'inizio ma entrare in competizione in una fase più avanzata. Inoltre, l'approccio DST tende a enfatizzare la variazione personale e la traiettoria individuale dello sviluppo, per cui questo genere di *trade-offs*, ossia

---

4 La Dynamic System Theory (DST) è un settore di studi emerso come branca della matematica e delle scienze naturali, ma con applicazioni spesso interdisciplinari, che si occupa dell'analisi delle proprietà e della costituzione di un sistema in quanto tale. Il linguaggio può essere visto come un sistema dinamico, cioè un insieme di variabili che interagiscono nel tempo, e lo sviluppo linguistico può essere visto come un processo dinamico. Per un approfondimento consultare De Bot, Lowie & Verspoor (2007).



compensazioni tra diverse parti del sistema linguistico, non deve essere necessariamente generalizzato a gruppi di studenti numerosi (Larsen-Freeman, 2006).

Abbiamo già parlato di trade-offs quando abbiamo introdotto il *Limited Attentional Capacity Model* di Skehan (paragrafo 1.4.4): in quel frangente, le relazioni solidali o competitive tra le varie dimensioni della performance erano ugualmente rilevanti nel campo della ricerca sulla complessità del task e venivano osservate da un punto di vista sincronico. L'approccio DST, invece, pone al centro le trade-offs, considerandole come una parte essenziale dello sviluppo linguistico, spesso analizzate dal punto di vista diacronico, in studi longitudinali<sup>5</sup>, e non come risultato delle caratteristiche del task. Eppure, un principio di base accomuna le trade-offs dello sviluppo alle trade-offs relazionate al task: entrambe derivano dalla competizione di risorse cognitive limitate, che vengono assegnate in modo non equo alle diverse dimensioni della performance, in particolare complessità, accuratezza e fluenza (Larsen-Freeman, 2006).

L'approccio DST allo studio dello sviluppo linguistico non si limita alle trade-offs relative alle dimensioni CAF della performance, ma prende esplicitamente in considerazione anche le relazioni tra i diversi domini della complessità, ad esempio dimostrando che la lunghezza della frase (complessità sintattica) entra in competizione con la diversità lessicale (complessità lessicale) nell'inglese accademico scritto (Verspoor et al., 2008).

Tuttavia, non è ancora completamente trasparente il modo in cui le relazioni solidali o competitive, coinvolte nello sviluppo del sistema linguistico della L2, sono legate alle manifestazioni superficiali della complessità (Spoleman & Verspoor, 2010). Poiché tali relazioni tra misure vengono osservate a livello di dati che raccolgono la produzione linguistica, la quantità di complessità ritrovata nel testo non può essere interpretata semplicemente ed esclusivamente come riorganizzazione o espansione del sistema interlinguistico dell'apprendente, dal momento che la produzione è notoriamente influenzata da una vasta gamma di altri fattori, tra i quali la complessità e la difficoltà del task. Diventa difficile, dunque, distinguere tra un tipo di studio che

---

5 Uno studio longitudinale è un tipo di studio in cui si osserva l'apprendente o un gruppo di apprendenti per un certo lasso di tempo documentando i graduali sviluppi dei loro sistemi linguistici (Pallotti, 1998:15).

investiga come diversi domini di complessità si relazionano tra loro nel corso dello sviluppo e un altro che studia come le circostanze della performance influenzano i vari domini della complessità nella produzione a diversi livelli di competenza dei soggetti. Entrambe, infatti, sono indagini che tipicamente implicano l'analisi di dati derivati dalla produzione linguistica degli apprendenti (Spoleman & Verspoor, 2010).

Infine, non è ancora chiaro fino a che punto sia corretto interpretare le relazioni competitive, o trade-offs, che caratterizzano lo sviluppo della L2, come vere correlazioni negative. Un'interpretazione alternativa di queste trade-offs potrebbe essere la "stagnazione" nello sviluppo di un determinato dominio linguistico come reazione compensatoria all'incremento della complessità in un altro dominio. Questo tipo di relazione non implicherebbe di per sé una correlazione negativa statisticamente significativa (Spoleman & Verspoor, 2010).

## 1.6 Complessità e competenza

L'ultimo obiettivo individuato da Ortega (paragrafo 1.3.3) nello studio della complessità linguistica come variabile dipendente, riguarda la sua interpretazione quale indicatore della competenza linguistica dell'apprendente.

Se da una parte la complessità (e l'accuratezza e la fluenza) della produzione di un apprendente è senza dubbio un indicatore rilevante al fine di fornire al ricercatore un quadro sulla sua competenza generale e sull'avanzamento dell'interlingua, dall'altra bisogna astenersi da conclusioni affrettate, poiché, come abbiamo più volte ripetuto, le manifestazioni superficiali della complessità sono ugualmente influenzate da fattori relativi al contesto della produzione o all'individuo, che non riflettono direttamente la competenza dell'apprendente e la sua conoscenza esplicita della L2. In termini Saussuriani, si tratta di ammettere che il solo livello della *parole* non possa bastare al ricercatore per avanzare inferenze circa la *langue* dell'apprendente, visto che la complessità linguistica può emergere in una produzione anche (e soprattutto) in relazione a variabili che possono avere nulla a che fare con la struttura sottostante dell'interlingua. Pertanto, l'istituzione di un legame lineare tra la complessità da un lato,

e la competenza e lo sviluppo dall'altro, diventa in ultima analisi una questione di indagine empirica, al centro di un più ampio movimento critico sullo studio della complessità linguistica (Pallotti, 2009, 2015a; Norris & Ortega, 2009; Bulté & Housen, 2012).

A partire dagli anni 2000, sempre più ricerche hanno sfidato il presupposto che gli apprendenti più avanzati utilizzino necessariamente un linguaggio più complesso, sostenendo che " *'more complex' does not necessarily mean 'better'* " (Ortega 2003:494).

Si è evidenziato come apprendenti più competenti o parlanti nativi tendano, in effetti, ad utilizzare strutture linguistiche più semplici -in senso assoluto- quando ritenute più efficienti, ovvero più adatte agli scopi comunicativi (Ortega, 2003; Pallotti, 2009; Ferrari, 2012; Lambert & Kormos, 2014).

A questo proposito, Pallotti (2009) avanza la proposta di adottare l'adeguatezza comunicativa quale parametro da tenere in considerazione sia come ulteriore dimensione della performance, sia come strumento per interpretare le misurazioni CAF. Ciò aiuterebbe a spiegare perché un apprendente che usa un linguaggio inaccurato, semplice e non fluente può, in certi casi, riuscire comunque a raggiungere gli obiettivi comunicativi prefissati. Allo stesso modo, un linguaggio altamente complesso e normativo potrebbe essere meno appropriato dal punto di vista comunicativo (ad esempio in una conversazione informale) e suscitare nei parlanti nativi un senso di stranezza o inadeguatezza (Pallotti, 2009).

Come vedremo, il nostro studio comprende anche un gruppo di controllo di parlanti native, perciò prende le distanze dal presupposto che i gruppi di parlanti L1 utilizzino in tutti i casi un linguaggio più complesso rispetto ai gruppi di apprendenti L2, inquadrando così lo sviluppo linguistico sulla base di parametri *native-like* (o *norm-like*) di uso della complessità, piuttosto che in un'ottica del "di più è meglio".

Per concludere questa breve panoramica, l'impatto della complessità linguistica nella ricerca SLA ha investito diversi aspetti dell'apprendimento di una lingua seconda. C'è un largo consenso sul fatto che la complessità sia un'importante dimensione dello sviluppo linguistico e molti studi provano che gli apprendenti tendono, in maniera generale, ad usare un linguaggio più complesso man mano che acquisiscono

competenza nella lingua target (Ortega, 2012). Il modo in cui la complessità di vari domini linguistici (morfologico, lessicale, sintattico) e di varie dimensioni della performance (complessità, accuratezza, fluenza) si relaziona alle diverse fasi dello sviluppo o livelli di competenza rimane oggetto di dibattito e investigazione scientifica (Larsen-Freeman, 2006; Spoleman & Verspoor, 2010).

Le discussioni in questo campo enfatizzano che la complessità linguistica osservata a livello della *parole*, può soltanto indirettamente informare i ricercatori sulle effettive conoscenze dell'apprendente a livello della *langue*. Per quanto sarebbe allettante presupporre un legame lineare tra le due dimensioni saussuriane del linguaggio, rimane il ruolo del ricercatore quello di tenere conto di tutti i fattori che si interpongono tra le realizzazioni superficiali della complessità e l'interlingua sottostante dell'apprendente, con particolare attenzione al contesto della produzione e alle caratteristiche del compito comunicativo (Pallotti, 2009; Robinson, 2001; Skehan & Foster, 2001).

## **1.7 Misurare la complessità linguistica**

### ***1.7.1 Un arduo compito***

L'obiettivo che questo paragrafo si pone è quello di lanciare uno sguardo alla letteratura precedente a proposito delle misure utilizzabili per quantificare la complessità linguistica.

La complessità linguistica è stata generalmente descritta tramite misure quantitative, che dovrebbero, almeno idealmente, rappresentare le distinzioni teoriche presentate fino ad ora (paragrafo 1.2). Date le svariate definizioni del concetto di complessità, sono stati elaborati altrettanti metodi di misurazione, che per ragioni di spazio non possiamo elencare in modo esaustivo.

Dal punto di vista degli studi di tipologia, sono stati fatti alcuni tentativi al fine di calcolare la complessità di un sistema linguistico nella sua interezza, per poter mettere a confronto diverse lingue e in ultima analisi smentire o avvalorare l'ipotesi

dell'equi-complessità.

Secondo Nichols (2009), la complessità di una lingua può essere misurata contando gli elementi che compongono ognuna delle sue categorie grammaticali (consonanti e vocali, toni, generi, casi ecc...), il numero delle varianti paradigmatiche (come allofoni, allomorfi, classi di coniugazione), i fenomeni sintagmatici (ad esempio l'accordo di genere e numero) e le regole di combinazione fra tutti questi elementi. Un'operazione che la stessa autrice definisce uno "*huge task*" (Nichols, 2009:112).

Secondo Miestamo (2008), invece, misurare la complessità totale di una lingua è un compito pressoché impossibile, dal momento che nessuna unità di misura può rendere conto di tutti gli aspetti (lessicali, sintattici, morfologici, ma anche fonetici, pragmatici ecc...) di una data lingua; in aggiunta, sarebbe estremamente complicato comparare in maniera oggettiva e quantificare il contributo dato da ciascun livello alla complessità globale della lingua in questione. Per questi motivi Miestamo (2008) propone la distinzione, già nominata in precedenza (paragrafo 1.2), tra la *complessità globale* dell'intero sistema linguistico e la *complessità locale*, riguardante specifiche aree di interesse (lessico, morfologia, sintassi), la quale può essere misurata con più precisione.

Se prendiamo in considerazione il settore di ricerca SLA, l'interesse principale non è più nel calcolare la complessità globale della lingua, ma piuttosto nella misurazione della complessità della produzione orale o scritta dei singoli individui o dei gruppi di apprendenti. In questo campo, dagli anni '90 in poi si sono moltiplicati i metodi di misurazione della complessità linguistica, dando vita a un dibattito sull'abbondanza di questi indici, calcolati sempre più spesso tramite strumenti automatici, e sui limiti del loro utilizzo.

Bulté e Housen (2012), ad esempio, hanno riscontrato, a fronte di un campione di 40 studi sulle lingue seconde, 40 diversi parametri numerici per misurare la complessità linguistica. Molte di queste misure sono da considerarsi ibride, e dunque imprecise, in quanto rappresentano in sé stesse diverse dimensioni della complessità. Inoltre, gli autori rilevano quanto la discrepanza terminologica intorno al costrutto di complessità possa compromettere la chiarezza delle misurazioni: spesso, infatti, gli studi empirici rinunciano a definire la complessità da un punto di vista teorico, così accade

frequentemente che il tipo di complessità che viene misurata nell'indagine rimane implicito. In particolare, pochi studi sono espliciti quando si tratta di definire la misura utilizzata come descrittiva della complessità assoluta o della complessità relativa.

Nel parere di Bulté e Housen (2012), queste problematiche metodologiche unite a una visione semplicistica, sfociano nell'inconsistenza dei risultati ottenuti dalle precedenti ricerche SLA sulla complessità e le dimensioni CAF in generale.

Pallotti (2015a) afferma che per giungere a una valida descrizione della complessità dell'interlingua deve esserci prima un accordo tra gli studiosi sulla definizione di complessità assoluta e su una serie finita di misurazioni per renderla operativa. Questa gamma di misurazioni dovrà investigare in ugual misura i tre principali ambiti della complessità linguistica -morfologico, sintattico e lessicale- poiché essi corrispondono a quelli che Towell (2012) definisce i "*core building blocks*" dello sviluppo della L2, illustrativi di tre diversi tipi di conoscenza linguistica centrali per la ricerca SLA.

Ci sembra, quindi, opportuno presentare brevemente i costrutti di complessità morfologica e sintattica con un accenno ai rispettivi metodi di misurazione utilizzati nella ricerca SLA, per poi passare, nel prossimo capitolo, a una descrizione più accurata della complessità lessicale, oggetto di questo studio.

### ***1.7.2 La complessità morfologica***

Come ogni tipo di complessità, la complessità morfologica può essere considerata una nozione relativa, riguardante lo sforzo cognitivo necessario per elaborare le strutture linguistiche, o una nozione assoluta, definita dalla quantità di elementi e dalla loro combinazione all'interno del sistema linguistico (Miestamo, 2009).

In questo studio empirico approcciamo la complessità linguistica in senso assoluto, come una proprietà essenzialmente quantitativa, perciò possiamo definire la complessità morfologica di un sistema linguistico in funzione del numero di morfemi o esponenti che ne fanno parte e delle loro regole di combinazione e variazione paradigmatica (De Clercq, 2015c).

La complessità morfologica è stata largamente studiata nel campo della ricerca tipologica, specialmente con lo scopo di osservare le relazioni compensative esistenti tra la complessità sintattica e quella morfologica nelle diverse lingue (*negative correlation hypothesis*: vedi paragrafo 1.1). Nello specifico, una lingua viene considerata morfologicamente complessa sulla base della presenza e del numero di marcatori che designano, per esempio, i casi nominali, i tempi verbali, la limitatezza dell'articolo o il genere grammaticale; ma la complessità dipende anche dalle deviazioni allo schema una forma-un significato, per cui ci si imbatte in casi di sincretismo (una forma realizza varie funzioni grammaticali) o di allomorfismo (una funzione grammaticale è espressa attraverso varie forme) (Brezina & Pallotti, 2016).

Nella ricerca SLA, invece, la complessità morfologica può essere usata, con la dovuta cautela, come indicatore dello sviluppo linguistico e della competenza degli apprendenti.

L'acquisizione della morfologia di una lingua seconda rappresenta un grande ostacolo per gli apprendenti, inclusi quelli ad un livello più avanzato, i quali spesso continuano a lottare con le regole morfologiche della L2 anche dopo molti anni di esposizione e studio (De Clercq, 2015c). Il loro processo di apprendimento può essere descritto come uno sviluppo a partire dalle forme basiche, non flesse o con un'unica forma, fino ad arrivare alle forme morfologicamente più varie, dunque la difficoltà cognitiva maggiore spesso scaturisce dalle situazioni di allomorfismo e sincretismo, in cui gli apprendenti devono essere in grado di riconoscere le forme e le loro funzioni, che spesso rendono sottili significati grammaticali, non condivisi da tutte le lingue (Brezina & Pallotti, 2016).

Diventa, perciò, molto utile poter definire quantitativamente la complessità morfologica di un sistema linguistico o di una singola produzione.

Il metodo di misurazione più usato nell'ambito della tipologia consiste nel conteggio degli elementi morfologici che compongono una data lingua. Vengono contati gli esponenti di ogni classe lessicale, cioè le forme assunte dai lessemi<sup>6</sup> per esprimere una categoria grammaticale, una funzione e/o un significato (Pallotti, 2015a); inoltre si identificano i principi che regolano le combinazioni fra i vari elementi. Se si tiene conto

6 Per "lessema" si intende una parola considerata dal punto di vista del significato; l'insieme dei lessemi di una lingua costituisce il suo lessico (Berruto, 2006).

di entrambi gli aspetti, anche una lingua che sembra possedere scarso materiale morfologico potrebbe rivelare segni di complessità relativa, data dal costo cognitivo che richiede l'utilizzo di tale materiale (Baerman et al., 2015).

L'approccio tipologico è difficilmente applicabile al contesto SLA, poiché in quest'ultimo le ricerche sulla complessità morfologica si basano su campioni ristretti di dati che possono essere più rappresentativi di un singolo apprendente a un preciso punto del processo di apprendimento, piuttosto che di un gruppo omogeneo di parlanti (De Clercq, 2015c).

La ricerca SLA è molto carente per quanto riguarda i metodi di misurazione della complessità morfologica. Nella loro rassegna, Bulté e Housen (2012) incontrano solo sei studi all'interno del loro campione che si interessano alla morfologia, utilizzando perlopiù lo stesso metodo di misurazione, mentre riportano ben 31 studi che calcolano la complessità sintattica tramite le più svariate misure. Wolfe-Quintero et al. (1998), in un'altra raccolta, non menzionano nessun tipo di misurazione specifico per la complessità morfologica, mentre sono frequenti le misure relative all'accuratezza morfologica.

Una possibile spiegazione per questa mancanza può ricercarsi nel fatto che le poche misure emerse fino ad oggi sono state ideate per corpora di dati provenienti da apprendenti di inglese L2, la lingua certamente più studiata nella ricerca sull'acquisizione di lingue seconde, e tuttavia una lingua morfologicamente "povera", a tal punto che anche gli apprendenti alle prime armi raggiungono i livelli più elevati secondo gli indici di complessità morfologica (De Clercq, 2015c).

Nonostante ciò, l'inclusione di queste misure è fondamentale poiché il costrutto di "complessità grammaticale" non può essere ridotto alla sola sintassi, specialmente nel caso in cui la lingua target sia morfologicamente "ricca" e quindi possibilmente più semplice dal punto di vista sintattico (Baerman et al., 2015).

Sia gli studi che prendono in considerazione la L2, che quelli relativi alla L1, operazionalizzano la complessità morfologica in termini di *diversity*. La diversità morfologica si può definire come:



“the variety of grammatical word forms in a system or a text. In its most straightforward operationalisation, morphological diversity represents the ratio of unique word forms, inflected or uninflected, to the total number of word forms.” (De Clercq, 2015c:9).

La diversità può essere considerata come indicativa della misura del repertorio morfologico mentale del parlante, mentre le differenze intrinseche alle strutture linguistiche non sono tenute in conto. Ciò significa che, nel calcolare la diversità, ogni forma contribuisce equamente alla complessità del testo, anche se certe strutture possono sembrare più complesse (ad esempio i circonfissi rispetto ai suffissi) o avanzate (come i modi finiti rispetto agli infiniti) di altre.

Secondo questa nozione di diversità, la complessità non consiste solo nell'utilizzo di forme linguistiche complesse ma anche nell'uso di un'ampia varietà di strutture e forme (De Clercq, 2015c).

Brezina e Pallotti (2016) propongono una misura della complessità morfologica che si basa su una concezione puramente strutturale e assoluta della complessità (Pallotti, 2015) e che può essere calcolata automaticamente tramite una procedura informatica: l'Indice di Complessità Morfologica (ICM).

L'ICM calcola la diversità flessiva media di una data classe di parole che occorrono all'interno di un dato testo. Per diversità flessiva si intende quella serie di manipolazioni, di diverso tipo, attuabili su una base lessicale (*default base*), che portano alla formazione di parole diverse. Il termine tecnico che è stato dato a tali manipolazioni è “esponenti” (Brezina & Pallotti, 2016). Ad esempio, il verbo *suona* proviene della default base *suonare* che ha subito un processo flessivo e presenta l'esponente verbale *-a*, il quale indica due diversi significati grammaticali (il tempo presente e la terza persona singolare).

Prima di questa, altre misure della complessità morfologica sono state proposte, non solo basandosi sulla diversità della flessione verbale, ma anche sulla morfologia derivazionale. Per ragioni di spazio, ci limitiamo a menzionare qui il metodo di misurazione ideato da Brezina e Pallotti (2016), che avrà un ruolo importante all'interno nostra indagine empirica.

### 1.7.3 *La complessità sintattica*

La complessità sintattica si manifesta nelle strutture sintattiche di vario tipo che emergono nella produzione linguistica e nel grado di sofisticazione di tali forme (Ortega, 2003).

Quella sintattica è senza dubbio la dimensione più studiata della complessità linguistica nel campo dell'acquisizione di lingue seconde, dove spesso ingloba concetti più generali quali la "complessità grammaticale" o persino la "complessità linguistica" (Bulté & Housen, 2012).

L'attenzione ricevuta da questo costrutto nella ricerca SLA è dovuta al suo utilizzo come indicatore della competenza e dello sviluppo linguistici dell'apprendente L2. Diversi studi hanno mostrato come negli stadi iniziali dell'apprendimento le strutture sintattiche siano molto spesso "*noun-based verbless structures*" (De Clercq, 2015a:10) e solo gradualmente si evolvano in periodi semplici organizzati attorno a verbi finiti, o strutture più complesse come domande e negazioni. Di conseguenza, è comune pensare che lo sviluppo della complessità sintattica nell'interlingua degli apprendenti rappresenti la progressione della loro competenza linguistica col passare del tempo. Bisogna tenere presente, però, che le misure della complessità sintattica potrebbero essere fraintese se utilizzate come indici diretti della competenza linguistica (Ortega, 2003). Infatti, i progressi nell'acquisizione della L2 non si limitano all'emergere di strutture sintattiche più articolate ma implicano anche una crescita del repertorio discorsivo e sociolinguistico che permette al parlante di adeguare la sua produzione alle richieste del contesto in cui si trova a comunicare (Ortega, 2003). Si potrebbe, dunque, affermare che la qualità della complessità sintattica è relativa al contesto e al significato da trasmettere, perciò è la varietà nelle strutture dei periodi a incidere positivamente sulla qualità del testo e non la complessità dei periodi stessi (Beers & Nagy, 2009).

Nonostante questa considerazione, che ha portato la nozione di *diversity* al centro delle ricerche sulla complessità sia in campo lessicale (De Clercq, 2015b; Jarvis, 2013a,b; Malvern et al., 2004) che morfologico (De Clercq, 2015c; Brezina & Pallotti, 2016; Pallotti, 2015a), sono ancora pochi gli studi che hanno incorporato delle misure

della diversità in ambito sintattico.

La complessità sintattica è stata a lungo associata alle misure di lunghezza e di subordinazione. L'unità su cui si basano molte di queste misure è detta *T-unit* e viene definita come “*one main clause plus any subordinate clause or nonclausal structure that is attached to or embedded in it*” (Beers & Nagy, 2009:188). Quest'unità può essere analizzata da diversi punti di vista come la sua lunghezza, il numero di elementi morfologici per periodo, o la proporzione tra proposizioni coordinate e subordinate (Norris & Ortega, 2009).

Ad esempio, è una misura della subordinazione il rapporto ottenuto calcolando la media delle proposizioni subordinate per *T-unit*. Un testo con un alto rapporto presenterà più proposizioni subordinate e relative, mentre un testo con un basso rapporto sarà caratterizzato da periodi più semplici o da una serie di periodi uniti da preposizioni coordinanti. I periodi ipotattici appaiono frequentemente nei livelli intermedi di sviluppo dell'interlingua, quando gli apprendenti sono in grado di esprimere le connessioni logiche tra le idee tramite periodi più articolati; per questo ci si aspetta che un testo strutturato principalmente sulla paratassi venga prodotto da parlanti ad un livello iniziale di apprendimento (Norris & Ortega, 2009). A questo riguardo, alcuni ricercatori invitano ad un'analisi più scrupolosa della subordinazione, che tenga conto dell'insorgere di diversi tipi di subordinazione in diverse fasi dello sviluppo linguistico (Lambert & Kormos, 2014).

Nel caso in cui si debba analizzare la complessità sintattica di una produzione orale, c'è chi preferisce basare le misurazioni non sulla *T-unit*, bensì sulla *C-unit*: il termine sta per “*communication unit*” e include tutti gli enunciati grammaticali indipendenti, compresi quelli privi del verbo (Mehnert, 1998:90). Per esempio, la risposta “nel frigorifero” alla domanda “dov'è il latte?” sarà considerata una *C-unit* dato che, anche se non contiene un verbo, possiede un valore comunicativo relativo al contesto dell'interazione.

Abbiamo detto che la complessità di una produzione linguistica non è data solo dal grado di elaborazione delle strutture, ma anche dalla varietà delle forme padroneggiate dall'apprendente. Visto che i ricercatori hanno sottolineato che un alto grado di complessità strutturale non implica direttamente una buona competenza

(specialmente dal punto di vista dell'efficienza comunicativa), le misure della diversità possono invece rendere più raffinata e precisa la strumentazione utilizzata per descrivere la complessità, considerando la variazione delle strutture linguistiche al variare della situazione comunicativa e della fase di apprendimento. Ad esempio, i parlanti che abusano di alcune strutture sintattiche apparentemente complesse, potrebbero registrare un alto livello di complessità sintattica pur non avendo un repertorio linguistico avanzato; per ovviare a questo problema, è necessario combinare misurazioni che tengono conto di entrambi gli aspetti, della complessità strutturale e della diversità (De Clercq, 2015a).

L'*Index of Productive Syntax* (IPSyn) è un metodo di misurazione computazionale che individua le strutture sintattiche di un testo e le divide automaticamente in quattro categorie sintattiche fondamentali: sintagmi nominali, sintagmi verbali, costruzioni interrogative e negative, e le diverse strutture dei periodi. In totale, sono sessanta le strutture grammaticali che questo programma riesce a identificare. Inoltre, applica a ognuna un valore numerico (da zero a due) che rappresenta il numero delle sue occorrenze nel testo (ad esempio zero se non occorre, oppure due se compare due o più volte) (Hassanali et al., 2014). Grazie all'elevato numero di forme conteggiate, questo metodo di misurazione permette di catturare il livello di varietà e sofisticazione sintattica di una produzione linguistica (Norris & Ortega, 2009).

## CAPITOLO 2

### *La complessità lessicale*

#### **2.1 L'importanza del lessico**

La conoscenza del lessico è un aspetto di primaria importanza nell'acquisizione di una lingua seconda. Le parole sono i mattoni di cui si costruisce il linguaggio, le unità di significato su cui sono strutturate le frasi, i paragrafi e gli interi testi (Read, 2000).

Nei parlanti nativi, sebbene la crescita più rapida avvenga durante l'infanzia, il lessico continua a svilupparsi naturalmente nell'età adulta, in risposta alle nuove esperienze e opportunità di apprendimento vissute dagli individui, così come in relazione alle nuove invenzioni e tendenze della società.

Per coloro che vogliono apprendere una lingua diversa da quella materna, invece, l'acquisizione del lessico è tipicamente un processo più consapevole e impegnativo. Anche ad un livello molto avanzato, i parlanti non-nativi continuano a percepire i loro limiti nella conoscenza dei vocaboli della L2, quando per esempio leggono parole che non riescono a comprendere, o trovano difficoltà nell'esprimere dei concetti con la stessa adeguatezza propria della L1. Per questo, molti studenti credono che l'acquisizione di una lingua seconda sia essenzialmente una questione di apprendimento del lessico e impiegano molte energie e tempo nella memorizzazione di lunghe liste di vocaboli, affidandosi al dizionario bilingue come strumento fondamentale per la comunicazione (Read, 2000). La percezione degli apprendenti è stata per molto tempo distante dalle teorie degli insegnanti, che si preoccupavano principalmente dello sviluppo della competenza grammaticale degli allievi, trascurando il lessico in favore degli aspetti morfosintattici.

Uno dei primi promotori della rivincita del lessico fu Krashen, che durante una conferenza nel 1987 affermò: *“when students travel, they don't carry grammar books,*

*they carry dictionaries*” (Krashen, British Council Conference, Milano, 1987). Questa provocazione ha in sé una forte dose di realismo ed esprime l'esigenza, riconosciuta sempre di più anche dalla linguistica generale e dalla psicolinguistica, di spostare il lessico verso il centro della discussione teorica (Balboni, 2015; Bettoni, 2001).

Per comprendere l'importanza del lessico basti pensare che i primi tentativi di produrre il parlato sono costituiti da formule e da parole; inoltre, anche a livelli più avanzati, gli errori lessicali, rispetto a quelli grammaticali, sono di gran lunga i più comuni, quelli che i parlanti nativi notano di più, e quelli che recano più danno ai fini comunicativi (Bettoni, 2001).

Per vari motivi si può quindi essere d'accordo con Lewis (1993) che vede la lingua come lessico grammaticalizzato, e non come grammatica lessicalizzata. Secondo questa concezione il lessico è centrale in quanto veicola dei significati, in effetti come affermava già Wilkins nel 1974: *“The fact is that while without grammar very little can be conveyed, without vocabulary nothing can be conveyed”*.

Michael Lewis è il padre del Lexical Approach, un metodo di insegnamento delle lingue straniere elaborato nei primi anni novanta, che mette al centro la questione lessicale. La proposta innovativa di Lewis consiste nel tentativo di superare la tradizionale dicotomia tra grammatica e lessico considerando le parole non più come unità isolate da apprendere secondo liste decontestualizzate, bensì come “segmenti lessicali”, o per usare il termine originale “*chunks*”, che comprende tutte le dimensioni intermedie tra la parola e la proposizione<sup>7</sup>. I parlanti nativi fanno un largo uso di queste frasi lessicali, dunque l'obiettivo dell'insegnamento dovrebbe essere quello di mettere gli apprendenti di una lingua seconda nella condizione di avere un rapido accesso a questo linguaggio prefabbricato, permettendogli così di dirigere l'attenzione alla struttura più ampia del discorso e sviluppare una produzione fluente (Lewis, 1993).

Questo approccio non si presenta come un metodo totalmente rivoluzionario rispetto agli orientamenti dell'Approccio Comunicativo, anzi ne assume i principi costitutivi come il primato dell'orale rispetto allo scritto o la priorità della competenza

---

<sup>7</sup> I confini tra la parola, il sintema (“gatto delle nevi”, “peso specifico”), l'espressione idiomatica variabile (“tagliare la testa al toro”, in cui il verbo deve essere coniugato), l'espressione invariabile (“di punto in bianco”), il proverbio, ecc... sono confini molto labili, che non sembrano esistere nell'autocoscienza di chi usa una lingua (Porcelli, 2003).

comunicativa rispetto alla padronanza del codice. La differenza principale è una migliore comprensione della natura del lessico e del suo contributo alla didattica delle lingue (Porcelli, 2003).

Nonostante quanto detto fin'ora, infatti, gli studi sul lessico sono ancora oggi relativamente pochi. Il motivo principale di questa emarginazione, sia in linguistica generale che nel campo dell'acquisizione della L2, è da attribuire al fatto che il lessico è un sistema molto più aperto della morfologia o della sintassi, e le sue unità di base, ovvero le parole, sono numerosissime e difficilmente si prestano alla regolarizzazione. Ne consegue che gli studi sull'apprendimento del lessico della L2 sono più scarsi e non offrono una base di risultati coerenti secondo linee teoriche chiaramente definite (Bettoni, 2001).

La definizione del concetto di complessità lessicale, tema di questo capitolo, e l'individuazione di misure efficaci per rilevarla all'interno della produzione testuale, possono essere validi strumenti nelle mani dell'insegnante per valutare lo sviluppo e la competenza lessicale dei singoli apprendenti, oppure, nel caso di un corpus di dati più esteso, possono indirizzare il ricercatore nella comprensione dei processi di apprendimento e nella formulazione di strategie utili all'insegnamento del lessico L2.

## **2.2 La complessità lessicale e le sue dimensioni**

Come precedentemente asserito, il termine "complessità", quando associato alle scienze linguistiche, conta due connotazioni: una assoluta e una relativa (Miestamo, 2009). Con *complessità lessicale* ci si può quindi riferire al numero di componenti del sistema lessicale di una lingua e/o del lessico mentale attivo di un parlante e alle relazioni fra tali elementi (Jarvis, 2013b); ma allo stesso tempo, si può anche intendere la facilità o difficoltà mentale con cui tali elementi linguistici vengono appresi (Jarvis, 2013b). Possono considerarsi "assolute" quelle forme di complessità che riflettono proprietà oggettive del linguaggio, come l'estensione del lessico e la struttura delle parole, mentre ricadono nel dominio della complessità relativa le forme che sono più difficili da

imparare, come i vocaboli usati meno frequentemente in una lingua (Dahl, 2004; Szmrecsanyi & Kortmann, 2012).

Il concetto di complessità lessicale è stato usato come un termine-ombrello, per raggruppare al suo interno numerosi costrutti, quali la *diversità*, la *sofisticazione* e la *densità* lessicali, ciascuno dei quali riflette una diversa dimensione della conoscenza lessicale.

Non esiste un consenso generale intorno alla definizione e operazionalizzazione di questi costrutti (Pallotti, 2015a), né riguardo alla loro denominazione (si noti che la *diversità* lessicale ha assunto anche il nome di varietà lessicale, variazione lessicale o variabilità lessicale, a seconda dell'autore) (Jarvis, 2013a), tuttavia, tenteremo di darne qui una descrizione il più possibile condivisa.

La *diversità lessicale* si può considerare una forma di complessità assoluta (Pallotti, 2015a), che ha a che fare con l'ampiezza del lessico mentale attivo, cioè il numero di parole che un parlante può effettivamente utilizzare (Jarvis, 2013b). La *diversità lessicale* riflette la capacità del parlante di introdurre nuovi elementi lessicali nella sua produzione (Skehan, 2009b) ed è misurata, nella sua forma più basilare, calcolando il rapporto tra le parole uniche (*types*), cioè quelle presenti una sola volta all'interno del campione, e il numero totale di parole presenti nel testo (*tokens*). Vedremo meglio, nel prossimo paragrafo, quali sono i limiti di questo metodo di misurazione e alcune soluzioni possibili.

Se la *diversità* riflette l'ampiezza del lessico (quante parole conosce il parlante), la *sofisticazione* è invece associata alla dimensione della profondità (quali parole conosce l'individuo) (Bulté & Housen, 2012). In questi termini, non è rilevante il numero assoluto di parole uniche nel testo, bensì quanto le parole utilizzate siano difficili o avanzate.

La *sofisticazione lessicale* è solitamente operazionalizzata sulla base della frequenza delle parole, partendo cioè dal presupposto che le parole meno frequenti siano semanticamente più specifiche o più astratte e di conseguenza più avanzate e difficili da apprendere, perciò la loro presenza in un testo indicherebbe una competenza lessicale più elevata (Laufer & Nation, 1995; Lindqvist et al., 2013).

Le misure della *sofisticazione lessicale*, dunque, si basano su liste di frequenza



delle parole, specifiche per la lingua presa in analisi e compilate raccogliendo le parole più frequenti in quella data lingua nel suo complesso (liste “estrinseche”) oppure selezionando le parole più frequenti nel campione di parlato preso in esame (liste “intrinseche”). Le liste di frequenza intrinseche assicurano una misurazione più precisa della sofisticazione in quanto sono dipendenti dal contesto della comunicazione e dal task, rispettando l'integrità dei campioni analizzati (Housen et al., 2008).

In effetti, la frequenza delle parole varia in relazione a fattori extra-linguistici come il contesto e il registro utilizzato, per questo la sofisticazione così misurata non può essere considerata una proprietà intrinseca del linguaggio, bensì una forma di complessità relativa (De Clercq, 2015b).

Un modo alternativo e meno comune di misurare la sofisticazione lessicale è quello basato sulla lunghezza delle parole. In questo senso, le parole più lunghe sono considerate più sofisticate delle altre. Contrariamente all'approccio basato sulla frequenza, la lunghezza delle parole può essere vista come un tipo di complessità assoluta, poiché misurabile oggettivamente in termini di numero di lettere o fonemi (De Clercq, 2015b).

Infine, la *densità lessicale* si riferisce all'informazione contenuta in un testo. Solitamente, la densità dell'informazione viene misurata attraverso il rapporto tra le parole contenuto e le parole funzione. Pallotti (2015a), tuttavia, considera problematico il ruolo della densità all'interno della complessità lessicale, in quanto non sono emersi elementi oggettivi che permettano di legare una maggior proporzione di parole contenuto o funzione con un maggiore o minore grado di complessità.

Altri metodi di misurazione più specifici sono stati proposti per quantificare la densità, come ad esempio la proporzione di nomi o di verbi all'interno del testo (Wolfe-Quintero et al., 1998).

Le tre dimensioni prese in considerazione, quindi, sono state misurate semplicemente contando il numero di parole diverse, la loro frequenza e la loro densità. Bulté e Housen (2012) ritengono necessario introdurre anche un livello collocazionale, che tenga conto della complessità lessicale contenuta nelle relazioni tra le parole. In questo modo, le nozioni di diversità, sofisticazione e densità possono essere applicate all'uso delle collocazioni e misurate in termini di, rispettivamente, varietà delle

collocazioni, frequenza e numero delle collocazioni nel testo. L'integrazione di questo livello a fianco delle altre dimensioni di analisi della complessità lessicale rimane, però, una questione problematica, dal momento che non è ancora possibile definire e identificare in modo univoco le collocazioni (De Clercq, 2015b).

Riportiamo, in figura 1, il modello descrittivo proposto da Bulté e Housen (2012), che individuano diversi livelli di analisi del costrutto della complessità linguistica. Prima di tutto, la complessità può essere esaminata da un punto di vista astratto e teorico, come una proprietà del sistema (cognitivo), in termini di numero dei suoi componenti e relazioni che intercorrono tra di essi. Da un punto di vista più concreto, invece, le nozioni teoriche di complessità si manifestano nel comportamento linguistico in vari modi, che sono osservabili a livello della performance. Infine, al livello operativo si situano i metodi di misurazione e gli strumenti analitici concepiti per dare un'indicazione quantitativa del grado di complessità di un dato campione testuale.

I due linguisti sostengono che la distinzione fra questi tre livelli sia fondamentale e debba essere resa esplicita, poiché per ottenere delle misure valide della complessità linguistica è necessario prima stabilire chiaramente cosa si intende per complessità (*livello teorico*), come essa si manifesta nei dati linguistici raccolti (*livello di osservazione*), e come queste realizzazioni concrete possono essere catturate e quantificate (*livello operativo*). Molti studi sulla lingua seconda si limitano a definire la complessità al livello più basso, come un costrutto esclusivamente statistico, mentre sono poche le indagini che tentano di analizzarla in quanto costrutto cognitivo e comportamentale (Bulté & Housen, 2012).

Le relazioni fra i diversi livelli di analisi non sono affatto semplici (come indicano le molteplici linee che collegano le caselle dello schema) e rispecchiano le problematiche più comuni nel campo della ricerca sulla complessità linguistica nell'acquisizione di lingue seconde. Ad esempio, accade spesso che nelle ricerche empiriche rimanga vago o implicito il costrutto teorico a cui fanno riferimento le misurazioni utilizzate, sebbene la definizione del legame tra il livello operativo e quello teorico dovrebbe sempre essere trasparente per garantire la validità dei risultati (Bulté & Housen, 2012). In molti casi, i metodi di misurazione usati non rilevano esclusivamente un aspetto della complessità, bensì più di uno, mettendo in crisi la validità del costrutto

(Housen et al., 2008). Ad esempio, le misure che si basano sulla frequenza delle parole, utilizzate per calcolare la sofisticazione, sono in parte indicatori di altre dimensioni della complessità lessicale, come la diversità (questa relazione è visibile nello schema grazie alla doppia linea che parte dalla casella delle *frequency-based measures*) (Housen et al., 2008).

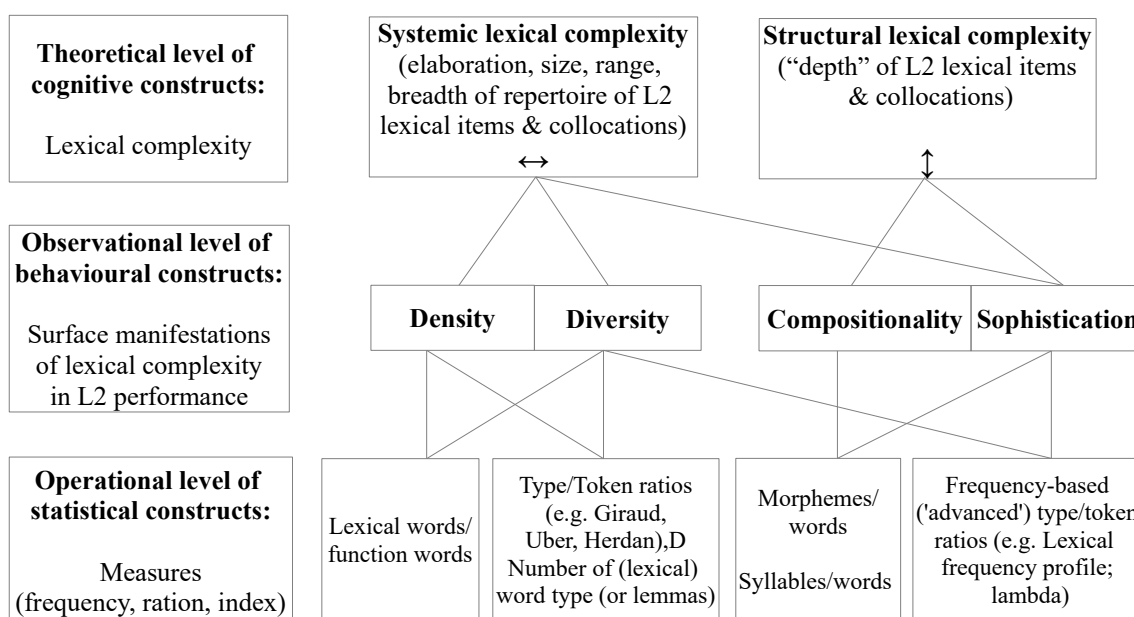


Fig. 1 – Lexical complexity at different levels of construct specification (Bulté & Housen, 2012:28).

Quelle descritte fino ad ora sono le dimensioni salienti della complessità lessicale osservabili nella performance degli apprendenti, ma ne esistono altre, anche rimanendo nell'ambito della produzione degli stessi autori. Bulté e Housen (2008), infatti, parlano anche di *produttività lessicale* e di *fluenza lessicale*. La prima è data dal numero di parole (tokens) che il parlante usa per completare un determinato task: vedremo, ad esempio, come in un dato contesto alcuni parlanti descrivono un'immagine utilizzando 200 parole, mentre altri svolgono lo stesso compito usandone 100. La seconda, invece, si riferisce alla velocità con cui il parlante produce o codifica le parole (specialmente le parole contenute), in base al grado con cui l'informazione lessicale è

stata proceduralizzata. Sia la fluenza che la produttività, come si può intuire, non riflettono soltanto la competenza lessicale del parlante, ma dipendono in buona parte da variabili esterne quali la motivazione stimolata dal task, il desiderio di comunicare e il carattere più o meno estroverso del soggetto (Housen et al., 2008).

La necessità di misure attendibili della complessità lessicale è da tempo dichiarata nel campo della ricerca SLA, poiché senza queste misure diversi quesiti teorici e pratici non possono trovare risposta. Alcuni tra questi interrogativi: la quantità di parole conosciute dai parlanti ad un determinato livello di sviluppo linguistico; la velocità con cui i lessemi vengono acquisiti; l'importanza di fattori come la lunghezza delle parole, la frequenza o la salienza nel determinare la difficoltà/facilità con cui esse vengono apprese; la relazione esistente tra l'ampiezza del vocabolario e altri aspetti come la sua profondità o automatizzazione. Come risposta a questi bisogni, negli ultimi anni si sono moltiplicati i metodi di misurazione relativi al lessico L2 (Housen et al., 2008).

Vediamo ora, più nello specifico, quali sono le misure proposte per la diversità lessicale, oggetto di questo studio, e con quali scopi di ricerca essa viene misurata.

### **2.3 La diversità lessicale: come si misura?**

La complessità lessicale, da un punto di vista assoluto, è generalmente relazionata al numero di componenti lessicali che fanno parte di un sistema linguistico o dell'interlingua di un apprendente. Dal momento che sarebbe impossibile contare in modo esaustivo i lessemi conosciuti da un individuo, ancora una volta questo tipo di analisi deve limitarsi a produzioni testuali concrete. Inoltre, deve necessariamente focalizzarsi sulle forme lessicali, lasciando da parte la complessità semantica, cioè quella derivante dalla polisemia di una parola o dalle restrizioni sulla sua co-occorrenza con altre parole, poiché questi aspetti sembrano impraticabili dal punto di vista operativo. Di conseguenza, la complessità lessicale viene operazionalizzata essenzialmente in termini di diversità: un testo con un'ampia varietà di lessemi sarà da

considerarsi più complesso rispetto ad uno in cui le stesse poche parole vengono ripetute più e più volte (Pallotti, 2015a).

La nozione corrente di diversità lessicale (LD) risale ad un periodo piuttosto intenso di ricerche innovative sul lessico nei campi della linguistica, della psicologia e dell'analisi statistica letteraria, che va dal 1935 fino al 1944. Come afferma Jarvis (2013b), questo periodo è importante poiché è il momento in cui si sviluppano le ricerche fondanti in questo campo, vengono coniat i termini principali e posti gli obiettivi di lavoro, specialmente riguardo al perché i ricercatori hanno bisogno di strumenti per misurare la diversità lessicale, cosa essa può rivelare e quali sono gli ostacoli da superare nella definizione di tali misure. Gli studi sulla diversità lessicale sono continuati fino ad oggi, espandendosi ad altri settori, come l'acquisizione della prima e della seconda lingua, il bilinguismo, il testing e la valutazione linguistica, i disturbi del linguaggio e l'afasia.

Di tutti i metodi proposti per misurare la diversità lessicale dei testi, il *Type/Token Ratio* (TTR) è il più intuitivo e trasparente (Jarvis, 2013a). Come abbiamo visto nel precedente paragrafo, questo indice è dato dal rapporto tra due valori numerici: il termine *token* si riferisce al numero complessivo di parole in un testo, mentre il valore *type* tiene conto di una sola occorrenza per ogni lessema presente nel testo. Per fare un esempio, nella frase “Mi chiamo Laura e mi piace cantare” si contano sette tokens ma sei types, visto che “mi” è presente due volte. Se la differenza tra i due valori è considerevole, significa che il testo è caratterizzato da un alto tasso di ripetizioni e può quindi essere definito lessicalmente poco complesso (Jarvis, 2013b).

Sebbene questa misura sia stata e sia tuttora ampiamente utilizzata per calcolare la diversità lessicale, dobbiamo tenere conto di un grave limite, ovvero la sua sensibilità alla lunghezza del testo. Man mano che un testo si allunga, ovviamente il numero di parole presenti nel testo (tokens) cresce. Se i tokens aumentano in modo lineare (una nuova parola = un nuovo token), non accade lo stesso per i types, che al contrario tendono a diminuire progressivamente. Con l'introduzione di ogni token diminuisce sempre più la probabilità di avere un nuovo type e ciò accade perché “*no text of more than a handful of words can be meaningful without some kind of repetition of tokens*” (McCarthy & Jarvis, 2010:382); per la stessa ragione, i testi con una maggior coesione

interna presentano anche un grado maggiore di ripetizione delle stesse parole (McCarthy & Jarvis, 2010).

La diversità lessicale così calcolata, dunque, decresce in modo sistematico all'aumentare della lunghezza del testo e questo rende impossibile la comparazione di TTR tra testi di lunghezza diversa.

Al fine di ridurre l'incidenza della variabile "lunghezza testuale" sul calcolo della diversità lessicale sono state avanzate numerose proposte nel corso del tempo. La prima risale al 1944, quando lo psicologo e patologo del linguaggio Wendell Johnson suggerì di calcolare il TTR di un numero identico di tokens per ogni testo (le prime 200 parole), o in alternativa, di scomporre i testi analizzati in sottocampioni della stessa lunghezza (50 parole) e calcolare il TTR medio per tutti i campioni. Johnson si riferisce a questa alternativa con il nome di *Mean Segmental Type-Token Ratio* (MSTTR), un indice che nonostante i limiti è tutt'oggi una delle opzioni più valide per calcolare la diversità lessicale in modo indipendente dalla lunghezza del testo (Jarvis, 2013a).

Un'altra ipotesi è quella proposta da Giraud nel 1954. L'*indice di Giraud* si computa ponendo la radice quadrata sul valore che indica i tokens, che si trova al denominatore nel rapporto con i types (dunque:  $\text{type} / \sqrt{\text{token}}$ ) (Giraud, 1954). Questa misura, come del resto molte delle manipolazioni del Type Token Ratio ideate in seguito, non si è dimostrata abbastanza efficace nell'eliminare l'effetto della lunghezza del testo sui risultati (Malvern et al., 2004; Lindqvist et al., 2013).

Un'ulteriore perfezionamento del TTR si realizza nel *tool vocd*, un metodo di misurazione computazionale implementato nell'ambito del programma CLAN (Computerised Language Analysis)<sup>8</sup>. *Vocd* calcola la diversità lessicale tramite un campionamento in modalità sequenziale o random, e la seconda sembra essere la più utilizzata<sup>9</sup>. In un primo passaggio, vengono individuati cento campioni casuali comprendenti dai trentacinque ai cinquanta tokens ciascuno. Successivamente, il

8 Disponibile gratuitamente su <http://childes.psy.cmu.edu> (Malvern et al., 2004).

9 I metodi di misurazione computazionali permettono di analizzare le parole che compongono un testo in modo sequenziale o non sequenziale. Da un lato, l'elaborazione sequenziale sembrerebbe il sistema preferibile in quanto consente di mantenere l'integrità del testo, rendendo l'analisi computazionale più simile ad un'analisi umana. Dall'altro lato, un'elaborazione random offre il vantaggio di evitare i raggruppamenti locali di parole contenuto presenti nel testo, che potrebbero portare ad una visione distorta della diversità lessicale. In certi casi, sono gli stessi ideatori dello strumento di misurazione informatico a suggerire l'approccio ideale. In questo caso, ad esempio, i creatori di *vocd* hanno constatato che il tool è più debole quando utilizzato in modo sequenziale (McCarthy & Jarvis, 2010).

programma calcola il TTR di ogni campione e, infine, esegue la media fra tutti i valori TTR calcolati. Visto che il campionamento del testo non è sequenziale, il risultato varia ogni volta che viene fatta la misurazione, per questo il procedimento viene ripetuto tre volte col fine di ottenere un risultato medio. L'indice così rilevato viene chiamato *D* ed è un numero che oscilla fra i valori 10 e 100, dove un valore più elevato sta ad indicare una maggiore diversità lessicale del testo (Malvern et al., 2004).

Questa procedura è stata successivamente contestata da McCarthy e Jarvis (2010), i quali considerano più efficace la *Measure of Textual Lexical Diversity* (MTLD), un indice elaborato in modo sequenziale, che calcola la lunghezza media delle stringhe di parole nelle quali si mantiene un certo valore di TTR (McCarthy & Jarvis, 2010). Questa misura è speculare rispetto al MSTTR di Johnson: se il MSTTR mantiene la lunghezza del campione costante (50 parole) mentre calcola il TTR medio fra i diversi segmenti del testo, la MTLD fa l'opposto, cioè mantiene il valore del TTR costante (di solito a .72) mentre calcola il numero medio di parole contenute in un segmento di testo che raggiunge quel dato valore.

Come afferma Jarvis (2013a), l'area di studio della linguistica quantitativa lavora a queste problematiche da ormai più di 60 anni, ma questo percorso non ci ha portato molto lontano rispetto ai risultati ottenuti nel '44. Pallotti (2015a:125) aggiunge, a proposito della grande varietà di misurazioni proposte: “*None of them is without problems and there is currently no consensus as regards the best index of lexical diversity*”.

Nuove prospettive sull'argomento potrebbero venire da altri settori della ricerca scientifica: secondo Jarvis (2013a), i linguisti dovrebbero trarre vantaggio dai progressi compiuti nel campo dell'ecologia, la quale si serve di principi simili a quelli utilizzati in linguistica per calcolare la diversità lessicale in un testo, ma con il fine di misurare un altro tipo di diversità, cioè la biodiversità delle specie in un dato ecosistema. In particolare, gli insegnamenti che si possono trarre da questa disciplina sono due: in primo luogo, gli ecologisti hanno adottato un approccio più intuitivo alla diversità, che gli ha evitato l'ossessione della ricerca di un indice di diversità che non vari in base alla lunghezza del testo. A sostegno di questa visione, Jarvis (2013a) ritiene che le misure della diversità non dovrebbero essere valutate solo sulla base della loro capacità di

evitare l'effetto lunghezza, ma principalmente per la loro abilità nel prevedere il giudizio umano sulla diversità. Tuttavia, questo concetto è di difficile applicazione nella pratica, poiché i soggetti che compiono la valutazione dovrebbero prima essere istruiti sul costrutto che stanno giudicando e ciò svierebbe la ricerca dall'obiettivo principale di scoprire se gli umani possiedono già un concetto insito di diversità lessicale condivisibile da tutti.

L'altra importante lezione che l'ecologia può fornire alla linguistica quantitativa riguarda la concezione della diversità come un fenomeno multidimensionale: entrambi i campi vedono la diversità in termini di complessità, ma gli ecologisti si sono spinti oltre nello sviluppo di modelli teorici e metodi di misurazione che permettano di comprendere le diverse dimensioni che costituiscono la diversità. Jarvis (2013a) individua una serie di proprietà della biodiversità, che potrebbero rendere il costrutto più chiaro e approfondito anche nella sua versione lessicale: oltre alle già conosciute *size* e *richness*, che si riferiscono rispettivamente al numero di esemplari/tokens e il numero di specie/types, egli parla di *evenness*, cioè il grado in cui gli esemplari/tokens si distribuiscono in modo equo attraverso le diverse specie/types esistenti nell'ecosistema/testo, e *dispersion*, che riguarda la distribuzione spaziale degli individui e delle specie in un dato ambiente, mentre nel testo si riferisce all'intervallo medio tra due tokens dello stesso type. Inoltre, siccome “*diversity is not just a matter of statistical frequencies and proportions of a given set of categories, but it also fundamentally involves the qualitative nature of and relationship between those categories*” (Jarvis 2013a:98) dobbiamo considerare anche le relazioni che si creano tra gli elementi di un sistema, sia esso linguistico o ecologico. Agli ecologisti è chiaro che non tutti gli elementi appartenenti ad un sistema danno lo stesso contributo alla diversità del sistema stesso: ad esempio, la biodiversità che deriva da tre diverse specie di topo non è la stessa risultante da un elefante, un albero e una formica, nonostante le specie in gioco siano sempre tre. Allo stesso modo, appare evidente ad un giudizio umano che le parole “sedia”, “sedie” e “sgabello” non conferiscono al testo la stessa diversità lessicale di tre parole come “gatto”, “cielo” e “rumore”, distanti dal punto di vista del significante e appartenenti a campi semantici completamente diversi. La *disparity*, quindi, è la proprietà che indica la differenza più o meno marcata tra le specie o types di un sistema.



Infine, gli ecologisti riconoscono che alcune specie sono più importanti di altre all'interno di un ecosistema, ad esempio se ricoprono un ruolo centrale nella catena alimentare (*centrality*) o se sono specie rare, la cui estinzione non potrebbe essere compensata da altre specie (*uniqueness*). L'insieme di questi due aspetti costituisce la proprietà detta *importance*, che si riflette, nel campo del linguaggio, nella centralità di alcune parole, legate alle altre da innumerevoli connessioni, o nella loro sofisticazione, intesa come minor frequenza di una parola in un dato sistema linguistico (Jarvis 2013a).

Supponendo che per ognuna di queste proprietà si sviluppino in futuro misure appropriate e il più possibile indipendenti le une dalle altre, e che si determini il peso di ognuno di questi indici, i linguisti potrebbero arrivare a una misurazione molto più completa e approfondita della diversità lessicale.

## **2.4 La diversità lessicale: perché si misura?**

Perché la ricerca SLA ha impiegato così tante energie e tempo nel tentare di misurare adeguatamente la diversità lessicale? Per rispondere a questa domanda dobbiamo elencare, almeno superficialmente, quali sono le possibili applicazioni del costrutto che ha richiesto un tale sforzo.

Se Jarvis (2013a) afferma che i progressi nella misurazione della diversità lessicale sono stati minimi dagli anni '40 ad oggi, ammette anche che in questo lasso di tempo qualcosa è cambiato:

“Even with imperfect measures, researchers have discovered that indices of lexical diversity serve as useful predictors of other important constructs, such as language proficiency, language complexity, vocabulary knowledge, and lexical proficiency (Berman, Nayditz, & Ravid, 2011; Crossley, Salsbury, McNamara, & Jarvis, 2011; Housen et al., 2011; Munoz, 2010; Sauro & Smith, 2010; Yu, 2010). Measures of lexical diversity have similarly been found to serve as useful gauges of the effects of, inter alia, aphasia (Crepaldi et al., 2011; MacWhinney, Fromm, Holland, Forbes, & Wright, 2010) and nonnativeness (Gui, 2010; Kormos, 2011). Thus, it is recognized that indices of lexical diversity are useful, even though language researchers have neglected the question of what it is that they are actually measuring” (Jarvis, 2013a:94).

Ancora prima di riuscire a definire la diversità lessicale e a misurarla correttamente, i linguisti avevano già scorto gli importanti progressi che questo costrutto avrebbe permesso di compiere nel campo dell'acquisizione di lingue seconde e nella diagnosi dei disturbi del linguaggio, e furono quindi queste scoperte a spronare la ricerca a colmare le lacune teoriche e operative.

Da allora, numerosi studi hanno dimostrato che di tutte le dimensioni della complessità lessicale (analizzate nel paragrafo 2.1), la diversità è senza dubbio quella che mostra una correlazione più forte con la competenza lessicale e con la *proficiency* generale, oltre ad essere il parametro più coerente con l'operazionalizzazione della complessità lessicale (De Clercq, 2015b; Verspoor et al., 2008; Housen et al., 2008; Jarvis, 2013a-b; McCarthy & Jarvis, 2010; Housen et al., 2012a).

In uno di questi studi empirici, Housen et al. (2008) argomentano che le 22 misurazioni da loro utilizzate come indicatori della diversità, della sofisticazione e della produttività, sono tutte sensibili in relazione allo sviluppo linguistico, sia tra gli apprendenti L2 che tra i parlanti nativi. In mezzo a queste però, le misure della diversità lessicale spiccano in quanto ad attendibilità nella discriminazione longitudinale delle diverse fasi dell'acquisizione. Inoltre, contrariamente ai problemi sollevati dalla misurazione della sofisticazione (paragrafo 2.1), uno dei maggiori vantaggi offerti dalla diversità lessicale è che la sua misurazione si basa esclusivamente sulle proprietà interne del testo, di conseguenza è praticabile a partire da testi di qualsiasi lingua, anche nel caso in cui non siano disponibili strumenti specifici per la lingua in questione, come le liste di frequenza (De Clercq, 2015b).

Ma le applicazioni della diversità lessicale non si fermano alla valutazione dello sviluppo linguistico, come affermano McCarthy e Jarvis (2010:381):

"LD has been used by researchers in fields as varied as stylistics, neuropathology, language acquisition, data mining and forensics; and LD indices have been found to be indicative of writing quality, vocabulary knowledge, speaker competence, Alzheimer's onset, hearing variation, and even speaker socioeconomic status".

In merito a quest'ultimo punto, Dillard e Pfau (2002) offrono un interessante punto di vista sulla complessità, che implica un'estensione della prospettiva dal parlante

al ricevente. La diversità lessicale non dovrebbe essere interpretata meramente come un effetto del livello di complessità del sistema linguistico interno di un parlante, poiché la diversità lessicale di un enunciato o di un testo scritto produce di per sé degli effetti sull'ascoltatore o sul lettore. Secondo gli autori, quando la complessità di un messaggio non eccede la capacità di comprensione del ricevente, è il ricevente stesso a preferire la complessità poiché essa stimola il suo interesse, di conseguenza, se il nostro scopo è quello di persuadere utilizzando il linguaggio, la diversità lessicale è uno strumento prezioso in quanto accresce la complessità (e quindi la gradevolezza) del messaggio.

Una serie di studi confermano questa argomentazione, dimostrando che la diversità lessicale è in grado di influenzare la percezione dell'ascoltatore su una serie di aspetti extra-linguistici relativi allo status sociale del parlante, come la credibilità, la competenza professionale, la simpatia o lo status economico (Dillard & Pfau, 2002).

Tutto ciò riconferma ancora una volta la nostra visione della diversità lessicale come un fenomeno percettivo, osservabile nell'esperienza e quantificabile tramite misure oggettive.

## **2.5 I fattori che influenzano la complessità lessicale**

La complessificazione graduale dei vari domini linguistici è un punto centrale nel processo di sviluppo linguistico di un apprendente, tuttavia non è sempre detto che l'aumento o la diminuzione della complessità siano da attribuire alle naturali fasi dell'acquisizione.

Innanzitutto, la complessità linguistica aumenta nel tempo non solo in virtù dei progressi nell'apprendimento, ma anche a causa di un generale sviluppo delle funzioni psico-linguistiche e socio-culturali degli individui. Di conseguenza, il fenomeno non interessa solo gli apprendenti di una lingua seconda, ma anche i parlanti nativi, che come tutti sono soggetti al potenziamento di queste facoltà<sup>10</sup> (Berman & Verhoeven, 2002). Dal punto di vista empirico, in una ricerca che, come la presente, mira ad

---

10 Lo dimostrano gli studi di Berman e Verhoeven (2002) su una varietà di soggetti dalla scuola elementare all'età adulta, parlanti di diverse L1, che registrano un generale aumento della diversità lessicale e della complessità sintattica con l'aumentare dell'età.

analizzare lo sviluppo della complessità linguistica su un corpus di dati longitudinale, si presenta un problema di valutazione: in quale misura l'aumento della complessità nei dati L2 deve essere interpretato come un effetto della continua maturazione cognitiva degli apprendenti o, piuttosto, come risultato dei meccanismi di apprendimento della lingua seconda? Ricerche future dovranno indagare nello specifico il rapporto tra la complessificazione della L1 e quella della L2.

Tra i fattori che influenzano la complessità lessicale i più importanti sembrano essere di natura interlinguistica, cioè quelli che dipendono dalle differenze che caratterizzano le diverse lingue. La diversità lessicale, la sofisticazione e la densità sono sensibili a questi fattori, che possono essere di origine tipologica, basati sulla relazione tra la L1 e la L2, o stilistici (De Clercq, 2015b).

Il già citato studio di Berman e Verhoeven (2002) sullo sviluppo della lingua materna, condotto su un campione di sette lingue (inglese, francese, ebraico, svedese, spagnolo, islandese, nederlandese), ha evidenziato differenze interlinguistiche nelle traiettorie di sviluppo della L1. Ad esempio, la naturale crescita della diversità lessicale legata all'aumento dell'età dei soggetti è risultata più limitata in ebraico e in francese rispetto alle altre lingue del campione. Molti altri studi come questo hanno rilevato differenze interlinguistiche relative alla complessità lessicale, sia in sincronia che in termini di sviluppo linguistico, sia per quanto riguarda la lingua materna che la lingua seconda.

Una possibile spiegazione per queste differenze si può ricercare nell'ambito della tipologia. Le caratteristiche grammaticali di una lingua, ovvero le risorse morfologiche e sintattiche disponibili, potrebbero influenzare il modo in cui vengono utilizzate le risorse lessicali. In questo senso, il fattore prominente nella caratterizzazione di una lingua è il grado in cui essa esprime l'informazione grammaticale, se attraverso elementi lessicali o elementi morfologici. La presenza più o meno marcata di parole contenuto o parole funzione, infatti, influenza i risultati delle misurazioni di diversità, sofisticazione e densità (Malvern et al., 2004). Visto che le parole funzione solitamente appartengono ad una classe chiusa di elementi con un alta frequenza, la loro continua ripetizione in una lingua provoca un generale abbassamento della diversità lessicale. Al contrario, nelle lingue del tipo agglutinante, in cui l'informazione grammaticale è espressa tramite

mezzi morfologici, gli elementi lessicali appartengono per la maggior parte ad una classe aperta di parole con una bassa frequenza, di conseguenza è normale registrare un più alto tasso di sofisticazione lessicale. Infine, dal momento che la densità si misura abitualmente come rapporto tra parole contenuto e parole funzione, è presumibile trovare una minor densità nelle lingue isolanti come l'inglese, che per esprimere i significati grammaticali si appoggiano principalmente a elementi lessicali (come verbi ausiliari e preposizioni) (De Clercq, 2015b).

Nell'ambito dell'acquisizione di lingue seconde spesso le differenze interlinguistiche vengono spiegate sulla base di un concetto chiamato *transfer*, definibile come:

“Un'intera classe di comportamenti, processi e condizionamenti, ciascuno dei quali ha a che fare con l'influenza translinguistica, cioè con l'influenza e l'uso di conoscenze linguistiche precedenti, solitamente ma non esclusivamente della lingua materna. Questa conoscenza contribuisce alla costruzione dell'interlingua interagendo in modo selettivo con l'input della lingua d'arrivo e con proprietà universali di vario genere” (Selinker, 1992:208).

L'influenza della lingua materna può condizionare la complessità lessicale della produzione di un parlante non-nativo: ad esempio, le espressioni lessicali che presentano affinità (nel suono, nella grafia o nel significato) tra la L1 e la L2, possono influenzare la sofisticazione lessicale di un testo. Queste parole, infatti, potrebbero essere molto rare nella lingua target, ma facilmente acquisite dagli apprendenti L2 e utilizzate in modo non appropriato. È il caso, ad esempio, di uno studente italiano o francese che utilizzi l'inglese *respond* al posto del più comune *answer*, a causa della somiglianza del primo con il termine più comune nella propria lingua materna (“rispondere” o *répondre*). Questi casi possono essere sporadici o molto frequenti a seconda del contatto tra le due lingue: in inglese, le parole entrate come prestiti dal francese sono numerosissime e in molti casi corrispondono a termini poco frequenti rispetto agli equivalenti autoctoni; per questo, la produzione in inglese di un apprendente francese può registrare livelli di sofisticazione lessicale non coerenti con la sua competenza linguistica generale (De Clercq, 2015b).

In ultima analisi, Pallotti (2015a) sostiene che la complessità lessicale sia sempre

soggetta alla variazione stilistica. Tutte le lingue, infatti, consentono un margine per questa variazione, basata puramente su preferenze stilistiche individuali. Ad esempio, se le regole per formare una frase subordinata in italiano devono essere necessariamente osservate da tutti i parlanti, il tasso di subordinazione di un testo non è prescritto da nessuna regola e permette quindi un'ampia variazione stilistica, per cui anche le deviazioni più estreme non potrebbero essere giudicate non-grammaticali, ma al massimo insolite. Quali sono, dunque, le differenze interlinguistiche di complessità dipendenti dal fattore stilistico? Diversi studi evidenziano che la variazione stilistica non è soggetta unicamente alle scelte dell'individuo, ma anche a preferenze culturali specifiche per ogni lingua, ovvero valori e modelli retorici tipici di una data comunità di parlanti. A questo punto, ad esempio, la diversità lessicale della lingua russa, consistentemente più elevata in comparazione all'inglese, può essere spiegata non solo come effetto delle differenze tipologiche esistenti tra le due lingue, bensì anche come frutto delle inclinazioni culturali che influenzano lo stile dei parlanti (De Clercq, 2015b).

Vedremo, nel prossimo capitolo dedicato allo studio di caso, quali sono le strategie utilizzate dai ricercatori per tentare di neutralizzare i fattori interlinguistici nella misurazione della complessità lessicale dei testi.

## **2.6 Le domande di ricerca**

Nello studio che ci apprestiamo a descrivere, si osserverà lo sviluppo della diversità lessicale nell'interlingua di quattro apprendenti di italiano L2, in un corpus di dati longitudinali (raccolti dal 2012 al 2015), che vede i soggetti impegnati in una serie di task di natura più o meno dialogica.

L'analisi condotta su questi dati sarà principalmente di tipo quantitativo<sup>11</sup> ma

---

<sup>11</sup> Le analisi quantitative sono quelle in cui i dati vengono descritti mediante variabili numeriche e producono risultati statistici. I dati quantitativi permettono di esprimere in modo esplicito e oggettivo relazioni che altrimenti potrebbero essere lasciate all'impressione del ricercatore. Tuttavia, alcuni fenomeni, come ad esempio le strategie usate dagli studenti per risolvere i problemi comunicativi, non possono essere descritti in modo quantitativo ma necessitano di un'analisi qualitativa. I dati quantitativi e qualitativi quindi non si escludono a vicenda, al contrario si compensano (Pallotti, 1998).

verrà integrata da considerazioni ed esempi qualitativi.

Nello svolgere la ricerca tenteremo di indagare le seguenti questioni:

- Come è mutata la produzione linguistica delle studentesse dal primo anno (2012) di raccolta dei dati all'ultimo (2015)? È possibile tracciare una traiettoria di sviluppo della complessità lessicale sull'asse temporale? I risultati evidenziano un trend comune a tutte le apprendenti?
- Si registrano differenze interindividuali notevoli nella misurazione della diversità lessicale delle apprendenti? Queste differenze rimangono costanti nel tempo?
- Quali sono i compiti comunicativi che generano una produzione linguistica più complessa dal punto di vista lessicale? Sulla base di questo fattore è possibile costruire una scala di complessità dei task? Se sì, questa scala rimane invariata dal 2012 al 2015?
- Quali sono i task che registrano una crescita maggiore della complessità lessicale tra il primo e l'ultimo anno?
- L'elemento  $\pm$ dialogico con riferimento ai task influisce sul valore rilevato dall'indice di complessità lessicale?

Tali riflessioni ci permetteranno di esplorare alcuni degli ambiti di applicazione dello STTR, l'indice che utilizzeremo per la misurazione della diversità lessicale.

Sulla base dei nostri risultati, valuteremo l'utilità di questo strumento nel campo della ricerca sull'acquisizione di lingue seconde, sia come supporto all'analisi testuale nella fase di ricerca scientifica, così come sul piano didattico, con lo scopo di adattare e migliorare le strategie di intervento educativo.





## CAPITOLO 3

### *Lo studio*

Dopo aver fornito un quadro teorico il più possibile articolato della complessità linguistica e più nello specifico di quella lessicale, proseguiamo con la presentazione dello studio di caso. Nel presente capitolo verranno descritti il corpus di dati e la metodologia utilizzata, mentre nel seguente si esporranno i risultati dell'analisi compiuta, insieme ad alcune considerazioni.

#### **3.1 Le peculiarità dello studio**

Circoscrivere il tema della complessità negli studi sulle lingue seconde è fondamentale per via delle sfide che le interlingue, sistemi dalla natura multiforme e imprevedibile, pongono a questo tipo di analisi. Una ricerca sulla complessità che si basi sulla produzione linguistica di utenti nativi consente di fare affidamento su una grammatica che sia perlopiù stabile, condivisa da un'estesa comunità di parlanti e quindi più facilmente descrivibile. Non è altrettanto semplice analizzare i dati provenienti dall'interlingua di un apprendente e tentare di regolarizzare questi sistemi linguistici, per eccellenza instabili e transitori. Ad esempio, è difficile stabilire quali funzioni l'apprendente abbia voluto attribuire ad una data forma, e talvolta non è chiaro nemmeno di quale forma si tratti, se sia utilizzata produttivamente o se sia solo una formula appresa e ripetuta (Brezina e Pallotti, 2016). Risulta quindi complicato quantificare la complessità partendo da dati provenienti dalle interlingue, soprattutto nelle varietà iniziali. Tuttavia, trattandosi di veri e propri sistemi linguistici, non è possibile ricostruire la complessità lessicale, morfologica o sintattica di un'interlingua se non mediante ipotesi puramente inferenziali dal campione osservato (*parole*) al sistema sottostante (*langue*) (Pallotti, 2015b).

Al fine di rendere più praticabile questo compito, terremo a mente l'intenzione, già dichiarata in precedenza (paragrafo 1.2.3), di affrontare questo studio adottando un punto di vista “assoluto” sulla complessità linguistica, che deve essere intesa qui come una proprietà essenzialmente strutturale, definita in base al numero di elementi presenti in un sistema linguistico o in un testo e al numero e il tipo di relazioni tra questi elementi, dunque misurabile tramite un indice matematico (Pallotti, 2015a).

Una volta compiute tutte le misurazioni quantitative, l'analisi sarà essenzialmente di tipo descrittivo, al fine di rendere conto della crescita, decrescita o stagnazione della complessità lessicale attraverso lo sviluppo linguistico degli apprendenti.

Come abbiamo visto nel precedente capitolo, la ricerca sulla complessità lessicale ha una storia travagliata e tuttora manca un consenso universale su alcune questioni come la sua misurazione, nonostante ciò gli studi empirici relativi a questo costrutto sono numerosi, specialmente nel campo dell'acquisizione di lingue seconde, dove le sue applicazioni si sono rivelate essenziali per il progresso scientifico. Tuttavia, per quanto riguarda il presente studio, almeno due caratteristiche lo collocano in un ambito di ricerca specifico ancora scarsamente indagato: ci riferiamo in primo luogo al fatto che si tratti di uno studio longitudinale, e in secondo luogo al fatto che sia condotto su apprendenti L2 ad un livello avanzato di competenza linguistica. Ortega e Byrnes (2008:4) sono stati i primi a lamentare la mancanza di studi che coniugassero questi due aspetti: *“To our knowledge, this is the first time the connection between the two areas of 'longitudinal study' and 'advanced capacities' is made. In fact, the field has only recently begun to explore either theme, and both are currently underresearched”*.

Le prime fasi dell'apprendimento di una lingua seconda hanno attratto l'attenzione di molti ricercatori e sono state ampiamente studiate (Pallotti, 1998), poiché in esse si osserva uno sviluppo linguistico più consistente e rapido, mentre nelle fasi più avanzate lo sviluppo appare più lento ed è necessario più tempo per osservare l'evoluzione dell'interlingua. Tuttavia, lo studio dello sviluppo dell'*advancedness* è estremamente importante per rispondere ad alcune domande che i ricercatori si pongono, specialmente nel campo della complessità. Ad esempio, abbiamo detto che la tensione verso l'efficacia comunicativa può portare gli apprendenti ad utilizzare diversi gradi di complessità a seconda delle esigenze; questo tipo di raffinatezza subentra solo nelle fasi

più avanzate dell'acquisizione ed è quindi osservabile nella produzione linguistica di apprendenti avanzati, per i quali l'obiettivo linguistico diventa una questione di allinearsi alle convenzioni dei parlanti nativi, piuttosto che di acquisire nuove conoscenze.

Per quanto riguarda, invece, la scelta di un corpus longitudinale, sono molti i ricercatori a sottolineare l'importanza di questo tipo di studi per la linguistica applicata, da sempre trascurati rispetto agli studi trasversali<sup>12</sup>, per ovvie ragioni di comodità (Ortega & Byrnes, 2008). Gli studi longitudinali, infatti, presentano un grave inconveniente: i processi di acquisizione possono richiedere in certi casi anche molti anni, mentre di solito i ricercatori desiderano ottenere i dati in tempi assai più brevi (Pallotti, 1998). Tuttavia, l'apprendimento di una lingua seconda prende molto tempo, perciò è solo indagando il fenomeno attraverso il tempo che possiamo raggiungerne una visione più chiara e completa. Housen et alii (2008:1) affermano:

“longitudinal research is needed. Such research, however, is scarce. Most L2 vocabulary studies to date, whether on French or on other L2s, have adopted a synchronic design, analysing lexical data from L2 learners at one point or over a brief period in time. Such studies can only provide information on aspects of lexical use, processing and representation, but not, or only indirectly, on lexical development over time”.

Ortega e Byrnes (2008), oltre accusare la rarità degli studi longitudinali delle nostre carenze nella conoscenza dei ritmi e modelli di sviluppo delle seconde lingue, sostengono la centralità di questi studi nella definizione delle strategie di insegnamento:

“Many researchers believe that longitudinal findings can uniquely bolster our knowledge of language acquisition processes, and many practitioners find longitudinal insights particularly convincing as a basis on which to justify recommendations for educational practice.” (Ortega & Byrnes, 2008:3)

Ci pare dunque un grande vantaggio poter analizzare, nel presente studio, un

---

12 Negli studi trasversali si prendono in considerazione gruppi di soggetti che sono stati esposti per tempi diversi alla lingua d'arrivo (ad esempio un gruppo di studenti del primo anno, uno del secondo, uno del terzo) e si osservano le produzioni linguistiche di ciascun gruppo. Così si ricostruisce artificialmente una sequenza di acquisizione che dovrebbe approssimare quella di ogni singolo apprendente se fosse osservato per un arco di tempo di tre anni (Pallotti, 1998:15).

corpus di dati longitudinali che si estende su un periodo di quattro anni (dal 2012 al 2015), poiché tale prerogativa conferisce maggior spessore ai risultati dell'indagine.

Infine, l'unicità del corpus è data anche dall'estrema multidimensionalità dei compiti comunicativi proposti agli apprendenti, una varietà che permette molteplici confronti e spunti di riflessione per le considerazioni didattiche.

### **3.2 Il corpus di dati**

Il protocollo di raccolta dati che descriveremo è stato elaborato all'interno di un progetto di ricerca pluriennale sull'acquisizione di competenze linguistiche e pragmatiche di giovani apprendenti di sesso femminile di livello intermedio-avanzato. Il progetto in questione – concretizzatosi nei PRIN<sup>13</sup> 2003, 2006 e 2009 – è stato finanziato dal Ministero dell'Istruzione, dell'Università e della Ricerca e ha visto la collaborazione di tre università italiane: l'Università di Bergamo in quanto capofila, nonché l'Università di Verona e l'Università di Modena e Reggio Emilia come unità locali.

I dati linguistici emersi da questo vasto progetto sono confluiti nel corpus intitolato VIP – Variabilità nell'Interlingua Parlata – costituito interamente da produzione orale, registrata in formato audio e video<sup>14</sup>.

La procedura di elicitazione dei dati è stata elaborata appositamente per l'occasione, tenendo conto che il livello delle apprendenti era sufficientemente avanzato da permettere una gamma di attività comunicative anche complesse. Inoltre, i task sono stati selezionati pensando agli interessi e ai gusti di un gruppo giovane e femminile (Pallotti et al., 2010). Alla sua prima edizione, infatti, il progetto prevedeva la partecipazione di sei ragazze immigrate, provenienti da vari paesi con L1 diverse, che al momento della raccolta dati frequentavano un istituto tecnico superiore della provincia di Modena ed erano residenti in Italia da quattro a sei anni. Hanno preso parte al progetto, come gruppo di controllo, anche due studentesse coetanee, parlanti native di

---

13 Progetti di Rilevante Interesse Nazionale

14 L'uso del video in analisi della conversazione rende i dati più ricchi di informazioni rispetto al solo audio. Tuttavia non possiamo considerare i dati video dei dati "completi", poiché anche una videoregistrazione può "tagliare" la realtà a seconda degli scopi di chi compie l'analisi (Pallotti, 1999).

italiano<sup>15</sup>.

È stato chiesto alle partecipanti di svolgere vari task, che si distinguevano per il tipo di interazione richiesta, il grado di confidenza e la differenza di età fra gli interlocutori, la presenza o assenza di contatto visivo o di altri stimoli.

Per quanto riguarda l'interazione tra le partecipanti nell'ambito dei compiti comunicativi dialogici, le coppie erano combinate in modo che le non-native parlassero con altre non-native, avendo cura di impedire che la coppia di studentesse condividesse la stessa L1. In altri casi, invece, le partecipanti comunicavano con interlocutori nativi esterni all'indagine, della stessa età o adulti, conosciuti o estranei, in modo da creare una svariata gamma di situazioni realistiche (Pallotti e Rosi, 2015).

Il protocollo è stato utilizzato una volta all'anno per quattro anni di seguito, con il fine di renderlo utilizzabile per studi longitudinali; inoltre è stato riproposto in due sessioni distinte: una prima dal 2005 al 2008 e un'altra dal 2012 al 2015 (oggetto di questo studio), a cui sono state apportate leggere modifiche sulla base delle osservazioni emerse durante la prima sperimentazione.

Nel complesso, questo corpus di dati è in grado di fornire un'ampia varietà di informazioni sull'interlingua delle apprendenti e sul suo sviluppo longitudinale. Sulla base di questa raccolta è stato possibile condurre studi empirici sull'acquisizione dell'italiano L2, che hanno affrontato diverse dimensioni (tra cui complessità, accuratezza, fluenza, difficoltà interazionale, pragmatica), la loro evoluzione nel tempo e la loro dipendenza da altre variabili linguistiche ed extralinguistiche.

Per il presente studio si è scelto di utilizzare i dati provenienti dal protocollo più recente a nostra disposizione, ovvero l'ultima edizione del progetto svoltasi dal 2012 al 2015. Si è ritenuto opportuno studiare i dati diacronicamente più rilevanti, tralasciando i due anni intermedi e lavorando quindi sulle registrazioni del 2012 e del 2015.

---

15 Il gruppo di controllo è reso utile dal fatto che la variabilità implica una certa opzionalità, per cui l'assenza di determinati elementi nella produzione delle apprendenti non può essere imputata a ignoranza se riscontrata anche nella produzione delle parlanti native (Pallotti et al., 2010).

### 3.3 I soggetti

Il nostro lavoro si baserà sull'analisi della produzione linguistica orale di quattro studentesse non native che hanno partecipato al progetto nella sua ultima edizione. Per questioni di privacy tutte le partecipanti sono state preventivamente coperte da pseudonimo. Nella seguente tabella riportiamo alcune importanti informazioni riguardanti i soggetti e relative al primo anno di indagine (2012).

<b>Pseudonimo</b>	<b>Età</b>	<b>Provenienza</b>	<b>L1</b>	<b>Anni in Italia</b>
Lali	15	India	Punjabi	5
Monika	16	Ucraina	Ucraino	4
Katerina	16	Moldavia	Rumeno	4
Randa	15	Sri Lanka	Tamil	5,5

Tab. 2 – Informazioni sulle partecipanti italofone non native

Per svolgere alcune attività, le quattro apprendenti sono state divise in due coppie, formate in modo tale da riunire le ragazze provenienti da aree geografiche e linguistiche il più possibile distanti tra loro: Lali e Monika formano la prima coppia, Katerina e Randa la seconda.

All'inizio della sperimentazione la loro competenza linguistica in italiano si poteva collocare fra il livello B1 e B2 del Quadro Comune di Riferimento Europeo<sup>16</sup>.

### 3.4 I task

Sul task quale attività comunicativa con un obiettivo extralinguistico si è già detto nel primo capitolo (paragrafo 1.4.1). In questa sede, invece, si intende delineare le caratteristiche dei task selezionati per essere proposti in questo progetto di ricerca e atti

---

<sup>16</sup> Il Quadro Comune di Riferimento Europeo è un sistema descrittivo impiegato per valutare le abilità linguistiche conseguite da chi studia una lingua straniera europea (Consiglio d'Europa, 2001).

a stimolare un ampio spettro di funzioni linguistiche.

La gamma di task che presenteremo consente di controllare una serie di variabili indipendenti di potenziale rilevanza:

- interlocutore pari o adulto
- interlocutore conosciuto o estraneo, sia pari che adulto
- interlocutore nativo o non nativo
- interazione telefonica o faccia-a-faccia
- task monologico o dialogico
- presenza o assenza di contatto visivo tra i parlanti
- presenza o assenza di stimoli visivi
- tipi di funzioni linguistico-comunicative richieste

Per quanto riguarda lo svolgimento dei task, la raccolta dati è stata resa possibile dalla presenza di un'intervistatrice e tecnici operatori della telecamera. Ogni anno il protocollo prevede due momenti di somministrazione, due sessioni pomeridiane separate, a circa una settimana di distanza l'una dall'altra, che si differenziano per il tipo di attività proposte: il primo turno inizia sempre con il task *intervista* e comprende le attività più monologiche, mentre nel secondo si utilizzano procedure più interattive (Pallotti e Rosi, 2015).

I task a carattere monologico sono quelli che prevedono uno scambio interazionale fra i partecipanti a senso unico (*one way*), in cui l'informazione è trasmessa da un parlante all'altro in una sola direzione ed è possibile distinguere il *giver*, colui che possiede l'informazione, dal *follower*, colui che la riceve. Al contrario, i task dialogici prevedono un flusso di informazioni fra i partecipanti in entrambe le direzioni. Il confine tra attività monologiche e dialogiche può essere sfumato, poiché anche le situazioni più monologiche richiedono sempre un'interazione, seppur esigua, con l'intervistatrice: è per questo che classificheremo i task in “prevalentemente monologici” e “prevalentemente dialogici” (Pallotti et al., 2010).

I task proposti alle apprendenti erano gli stessi ogni anno, presupposto fondamentale per la comparabilità dei dati raccolti; tuttavia, varianti leggermente

diverse degli stessi task sono state introdotte da un anno all'altro per scongiurare gli effetti della ripetizione dello stesso compito più volte, come il calo della motivazione dovuto alla noia o il rischio di automatizzazione della performance (Bygate et al., 2001).

Di seguito riportiamo una breve descrizione dei task presi in analisi in questo studio.

### ***3.4.1 Task prevalentemente monologici***

#### STORY RETELLING

In questo task, viene richiesto alle apprendenti di osservare una storia costituita unicamente da immagini e di raccontarla all'intervistatrice. Le studentesse vengono invitate a prendersi qualche minuto per osservare le immagini prima della narrazione, fattore che rende possibile la pianificazione linguistica; inoltre, se lo desiderano, possono continuare a guardare le immagini durante il racconto. La variabile della pianificazione sulla produzione linguistica è stata ritenuta responsabile per l'aumento della fluenza, mentre per quanto riguarda la complessità e l'accuratezza gli effetti sono più limitati (Ellis & Yuan, 2005).

Il testo utilizzato è il libro per immagini di Mercer Mayer dal titolo *Frog, where are you?* del 1969 (le cui immagini sono riportate in appendice), il quale costituisce un interessante stimolo visivo ed già stato impiegato nel settore della ricerca sull'acquisizione della seconda lingua per elicitar narrazioni semi-spontanee. Siccome la "*frog story*" era inizialmente indirizzata ad un pubblico infantile, si presuppone che non generi una difficoltà concettuale o retorica nella maggioranza dei partecipanti.

Durante il task l'intervistatrice rimane in silenzio ed interviene raramente, solo se le apprendenti non conoscono una parola importante nella storia e non sono in grado di parafrasare, o nel caso in cui tralascino delle parti importanti delle scene osservate.

Non è imposto un tempo limite per il monologo, ma di solito esso si conclude entro i sette e i dieci minuti.



## INTERVISTA

Nell'intervista il soggetto è invitato a presentarsi e raccontare di sé, come in una semplice chiacchierata. L'intervistatrice, rappresentata ogni anno da una persona diversa al fine di garantire un reale scambio comunicativo, guida la conversazione in modo che vengano affrontati alcuni nuclei tematici. Nella versioni successive alla prima sperimentazione, l'intervistatrice è stata dotata di un elenco di domande, poiché la precedente traccia si era rivelata inadeguata a restituire dei dati abbastanza simili da permettere un confronto tra le diverse partecipanti.

L'intervista è un importante momento conoscitivo unidirezionale, in cui le ragazze trascorrono quasi un'ora parlando di temi molto personali quali: il viaggio dal paese di origine all'Italia, le difficoltà affrontate durante il primo periodo, l'adattamento alla società d'accoglienza, l'andamento scolastico, gli interessi, le amicizie, il tempo libero, la descrizione caratteriale dei famigliari e un ipotetico ritratto del proprio carattere definito dal punto di vista dei genitori. Il trasporto emotivo con cui si raccontano i fatti personali è considerata una variabile importante in grado di influenzare la produzione linguistica dei soggetti. Altre variabili presenti in questo task sono l'assenza di pianificazione della narrazione e l'interazione del tipo faccia-a-faccia con un parlante nativo, adulto e sconosciuto.

### ***3.4.2 Task prevalentemente dialogici***

I task interattivi qui di seguito sono stati svolti a coppie, con una formazione stabile precedentemente definita (Lali-Monika, Katerina-Randa). Tutte le attività sono state inserite all'interno di una cornice di gioco, presentate come tappe obbligate di una caccia al tesoro a premi. Il compimento di ogni attività nei tempi prestabiliti consente alle partecipanti di ottenere un indizio necessario a risolvere l'enigma che le condurrà al tesoro, costituito da alcune ricariche telefoniche e buoni pasto. L'intervistatrice svolge un ruolo di coordinamento delle attività, fornendo le spiegazioni necessarie e dettando sommariamente le tempistiche. Al termine di ogni prova la guida rivela l'indizio e assegna un punteggio individuale e un punteggio di coppia, premiando la capacità di

collaborare, la precisione nell'esecuzione del compito e l'originalità. Il sistema di valutazione e i premi finali sono stati pensati per creare competitività tra le squadre, favorire la collaborazione interna alla coppia e rendere il contesto comunicativo più motivante.

### MAP TASK

Questo task è stato ideato e sperimentato per la prima volta nell'ambito del progetto AVIP (Archivio Vocale dell'Italiano Parlato), condotto da diverse università italiane<sup>17</sup>. Si tratta di un compito non linguistico di tipo cooperativo, che coinvolge due partecipanti, in questo caso l'apprendente e l'intervistatrice. A ognuno di essi viene consegnata una mappa, ma solo una delle due ha un percorso tracciato. L'obiettivo è quello di trasferire in modo accurato il percorso da una mappa all'altra utilizzando esclusivamente l'interazione verbale. L'apprendente, che possiede la mappa con il percorso, deve guidare l'intervistatrice nel tracciare l'itinerario corretto. Per il superamento della prova si hanno a disposizione venti minuti di tempo e la valutazione finale si basa sulla simiglianza tra il percorso ricostruito e quello originale.

Per complicare il gioco, le due mappe non sono identiche, nel senso che il numero, la natura e la posizione dei punti di riferimento (negozi, fontane, parchi) differiscono; questo genera fraintendimenti tra gli interlocutori, proprio come avviene in situazioni comunicative reali, e rende necessarie le pause esplicative e le richieste di conferma. Inoltre, un ostacolo materiale preclude il contatto visivo tra i due partecipanti, ulteriore elemento di disturbo che limita il linguaggio non verbale.

Il task permette l'utilizzo di diverse funzioni linguistiche: dare e chiedere istruzioni, pianificare e organizzare, risolvere i problemi comunicativi.

### ORGANIZZARE UNA GITA

L'organizzazione di un viaggio ipotetico tra amiche risulta essere un compito comunicativo altamente interattivo e si può considerare un task di *decision making*. Le

---

<sup>17</sup> Nello specifico: la Scuola Normale di Pisa, il Politecnico di Bari, l'Istituto Universitario Orientale di Napoli, l'Università di Napoli Federico II e l'Università di Vercelli. Maggiori dettagli sul progetto sono reperibili sul sito [www.parlaritaliano.it](http://www.parlaritaliano.it).

studentesse hanno a disposizione un budget di cinquecento euro, che devono decidere come spendere, delineando un programma dettagliato delle attività che vorrebbero svolgere durante quattro giornate di vacanza presso una località. Viene fornita loro una guida turistica della città, un elenco di numeri di telefono utili (enti del turismo, agenzie di viaggio, strutture alberghiere, ristoranti locali) e una scheda con i recapiti di sei esperti viaggiatori coetanei e sei adulti, dei quali vengono indicate alcune informazioni (nome, età, professione). Le ragazze possono scegliere chi chiamare ma devono obbligatoriamente consultare almeno quattro esperti prima di prendere le decisioni definitive sul viaggio. Inoltre, per organizzarsi al meglio, devono informarsi telefonicamente sui costi, sugli orari e sui tempi necessari per visitare i luoghi di interesse, calcolando anche i mezzi e i tempi di spostamento tra una tappa e l'altra. Al termine della pianificazione viene richiesto loro di completare una griglia fornendo un elenco preciso del loro programma di viaggio. Per completare il task le partecipanti hanno a disposizione 45 minuti di tempo.

Questo task si compone di due differenti attività, che, nel momento dedicato all'analisi, vengono tenute separate per consentire la corretta interpretazione dei dati linguistici provenienti dai due compiti: la negoziazione e la telefonata.

### NEGOZIAZIONE

È quella parte del task *organizzare una gita* che comprende le interazioni faccia-a-faccia tra le due apprendenti, dunque coetanee e non native. In questo spazio le partecipanti devono confrontarsi sulle diverse possibilità a loro disposizione e operare delle scelte di comune accordo. Per farlo è necessario scambiarsi informazioni, riportare all'altra il contenuto delle telefonate svolte, esprimere i propri gusti e preferenze e raggiungere un compromesso.

### TELEFONATA

Le telefonate costituiscono un compito molto diverso rispetto alla negoziazione, per questo vengono estrapolate dal corpo del testo e analizzate separatamente.

In questo task le funzioni linguistiche riguardano principalmente il pianificare e il

chiedere informazioni. Ovviamente la comunicazione si svolge in assenza del contatto visivo e l'interlocutore non è più un parlante non nativo, bensì un italofono nativo, coetaneo o adulto e, in ogni caso, sconosciuto.

### **3.5 La procedura di analisi**

#### ***3.5.1 La trascrizione delle registrazioni***

Data la vastità del corpus VIP, i dati provenienti dall'ultima edizione del progetto non sono stati ancora trascritti interamente. Abbiamo dunque approfittato per dare un contributo a questo lavoro che richiede il tempo e l'impegno di molti per essere completato. I dati in formato audio relativi alle apprendenti e ai task selezionati per il presente studio ammontano a circa 19 ore di registrazioni, che sono state trasposte in forma scritta per permetterci di esaminare visivamente i fenomeni linguistici.

Anche se i dati audio rimangono l'oggetto primario e più completo di uno studio, la trascrizione è uno strumento indispensabile per il ricercatore, poiché sui dati scritti è possibile compiere operazioni di ricerca, in particolare alcune procedure computerizzate, che non sarebbe possibile attuare sui dati audio. Inoltre, la trascrizione è il primo passo dell'analisi, il momento in cui si familiarizza con i dati e si sviluppano le prime ipotesi teoriche, per questo quando la trascrizione è fatta dal ricercatore stesso la sua comprensione dei fenomeni è molto maggiore. Come afferma Elinor Ochs, “la trascrizione è teoria” poiché trascrivere implica già di per sé fare delle scelte, stabilire delle priorità e fissarsi degli obiettivi (Ochs, 1979, in Pallotti, 1999).

Esistono diversi metodi di trascrizione, nessuno è migliore degli altri, ma ci possono essere tecniche più o meno adatte agli scopi di una ricerca. Il sistema che si decide di adoperare deve rispondere ad alcuni parametri: prima di tutto quello della “coerenza interna”, secondo cui i simboli indicanti i fenomeni del parlato devono sempre essere usati allo stesso modo all'interno di una trascrizione, o all'interno di una ricerca che comprende più trascrizioni; un altro principio fondamentale è la “attendibilità”, secondo cui una trascrizione deve essere il più possibile oggettiva e

lontana dalle interpretazioni dell'autore; infine, la “leggibilità” richiede che un testo venga trascritto in maniera tale da essere letto agevolmente, motivo per cui è bene utilizzare convenzioni notazionali che siano già diffuse (Pallotti, 1999).

I dati qui analizzati sono stati trascritti seguendo una versione riadattata del formato CHAT-CA (*Code for the Human Analysis Transcript*), simile al sistema Jeffersoniano, che rappresenta una delle tecniche più comunemente usate nel campo dell'analisi della conversazione e può essere considerato un modello standard<sup>18</sup>.

Considerato che il corpus VIP può essere analizzato da diverse persone adottando diversi punti di vista teorici, il sistema di trascrizione appositamente scelto<sup>19</sup> risponde all'esigenza di annotare una vasta gamma di fenomeni comunicativi, tra cui quelli pragmatici o extralinguistici, che comprendono:

- sovrapposizioni tra i turni di parola
- suoni non verbali (risate, espirazioni)
- volume della voce
- velocità dell'eloquio
- enfasi
- intonazione
- pause espresse in frazioni di secondo

### ***3.5.2 La preparazione dei dati all'analisi***

Dopo la trascrizione delle registrazioni, il secondo passo è la modifica di tali trascrizioni per renderle adeguate al formato richiesto dal *tool* informatico utilizzato per l'analisi testuale.

Innanzitutto, come già accennato in precedenza, le trascrizioni relative al task *organizzare una gita* sono state scomposte in due distinti testi, uno riguardante le *telefonate* agli esperti e l'altro le *negoziazioni* faccia-a-faccia fra le studentesse, poiché

---

18 Per un approfondimento circa i sistemi di trascrizione dei dati utilizzati nell'ambito della *Conversation Analysis* si faccia riferimento ai lavori di Jefferson (2004, 2015).

19 In appendice si trovano le convenzioni di trascrizione utilizzate per il presente studio, insieme a un esempio di trascrizione tratto dal corpus qui analizzato.

si tratta effettivamente di due task diversi, anche se somministrati all'interno della stessa cornice di gioco.

Per quanto riguarda le trascrizioni di tutti i task, poi, è stato necessario eliminare dal testo i turni di parola dell'intervistatrice, visto che il nostro scopo è analizzare esclusivamente la produzione linguistica delle apprendenti di italiano L2. Nei casi in cui l'interazione si svolgeva fra una coppia di non native, invece, abbiamo separato opportunamente la produzione delle due apprendenti, in modo tale da poter condurre un'analisi individuale su ciascun soggetto.

Dopodiché, le trascrizioni contenenti gli enunciati prodotti da una singola apprendente, sono state sottoposte a una serie di operazioni di sostituzione per eliminare tutto il materiale extralinguistico, costituito dai simboli utilizzati nella trascrizione e dalle risate o esitazioni, molto frequenti nel parlato.

Ogni file è stato ripulito in questo modo fino ad ottenere un puro testo. Per esempio, un frammento dell'interazione tra Katerina (S17) e l'intervistatrice (Investigator) è stato trascritto come segue:

```
S17: ehm:: #1_8 niente: sto cominciando a uscire di più:: anche se i  
miei mi guardano un po' così:: # [però::  
INV: [non sono molto contenti i tuoi che  
esci?  
S17: ehm:: #0_2 [diciamo che io ultimamente stavo sempre:: #0_2  
sempre:: in casa sempre controlla- cioè #0_5 controllata diciamo che::  
sapevano dove ero  
INV: [son preoccupati  
INV: s::i
```

Dopo le modifiche sopra descritte, lo stesso frammento appare così:

```
niente sto cominciando a uscire di più anche se i miei mi guardano un  
po' così però  
diciamo che io ultimamente stavo sempre sempre in casa sempre cioè  
controllata diciamo che sapevano dove ero  
sì
```

A questo punto il testo è pronto per essere immesso nel *tool* di analisi.

### 3.5.3 *L'analisi tramite il tool informatico*

Come abbiamo già visto (paragrafo 2.2), la diversità lessicale si misura abitualmente tramite il *Type/Token Ratio*, un indice in grado di calcolare la varietà o al contrario la ripetizione del lessico all'interno di un testo. Abbiamo anche detto che la sensibilità alla lunghezza del testo rende questa misura inaffidabile e dunque inutilizzabile così com'è, ma i linguisti hanno proposto diverse modifiche del TTR che attenuano o eliminano quasi completamente l'effetto indesiderato. In questo studio si vuole testare una di queste varianti del TTR, chiamata *Standardized Type/Token Ratio* (STTR)<sup>20</sup>. Questa procedura automatizzata opera una frammentazione dei testi in sotto-campioni contenenti 50 lemmi ciascuno e in seguito standardizza la proporzione tra types e tokens calcolando il valore medio del TTR per tutti i sotto-campioni: in questo modo rende due testi di lunghezza diversa comparabili fra loro. Nonostante il calcolo di questa misura sia normalmente affidato ad un computer, la validità dello STTR può essere dimostrata empiricamente ed è quindi controllabile dal ricercatore.

Nella presente ricerca faremo uso di un *tool* informatico sviluppato nel 2015 da due linguisti, Brezina e Pallotti, per l'analisi delle produzioni linguistiche di apprendenti L2 (disponibile all'indirizzo [http://corpora.lancs.ac.uk/vocab/analyse\\_morph.php](http://corpora.lancs.ac.uk/vocab/analyse_morph.php), Brezina e Pallotti, 2015). Il *Morpho complexity tool*, come suggerisce il nome, è stato ideato con lo scopo principale di analizzare la complessità morfologica dei testi, ma si è rivelato uno strumento versatile in grado di fornire un'analisi della complessità anche in campo lessicale. In effetti, abbiamo visto come anche dal punto di vista teorico la complessità linguistica venga spesso affrontata in termini di *diversity*, non stupisce quindi il fatto che una misura della complessità morfologica condivida lo stesso principio sottostante a una misura della complessità lessicale, ovvero il calcolo dell'indice di diversità.

Immettendo nel programma i dati trascritti e precedentemente ripuliti da ogni annotazione di natura extralinguistica, il *tool* restituisce la sua analisi sotto forma di file in formato “foglio elettronico”. In questa prima fase il programma opera una lemmatizzazione dei dati, avvalendosi del sistema *TreeTagger* (Schmid, 1994). La

---

<sup>20</sup> Peraltro già presente nella ricerca di Xanthos e Gillis (2010).

lemmatizzazione consiste nella riduzione di una forma flessa di una parola alla sua forma non marcata, anche detta “lemma” (o *default base*). L'algoritmo alla base del tool rende possibile la scomposizione interna delle parole e sulla base di essa ne riconosce l'appartenenza a una determinata classe (verbi, aggettivi, nomi, avverbi, preposizioni, pronomi, congiunzioni), raggiungendo ottimi livelli di accuratezza. Tuttavia, essendo un processo computazionale rimane comunque suscettibile a errori, che devono essere corretti manualmente dal ricercatore. Trattandosi di dati provenienti dall'interlingua, si possono incontrare nel testo alcune forme erronee che il sistema potrebbe fraintendere. Ne è un esempio la parola “multi” tratta dall'enunciato di Monika “andare diritto fino al giardino delle visite con multi fiori”: il programma l'ha automaticamente analizzata come una forma flessa del verbo “multare”, mentre solo una mente umana potrebbe riconoscere “multi” come una deformazione dell'aggettivo italiano “molti”, basandosi sul contesto comunicativo che la circonda.

Come si evince dalla figura 2, il file prodotto dal *tool* contiene quattro colonne: la prima riporta le parole che compongono il testo, nella seconda si trova la *Part of Speech* (categoria lessicale) corrispondente e nella terza appare il prodotto della lemmatizzazione cioè la *default base* di ogni lessema (ad esempio, il verbo “comincia” viene ricondotto alla base “cominciare”). La colonna “Edit”, invece, è stata aggiunta successivamente e contiene le correzioni apportate manualmente. In questo caso vediamo, ad esempio, che le forme verbali “vai” e “giri” sono state analizzate erroneamente come un aggettivo (vaio) e un nome (giro), il verbo “arrivi” coniugato alla seconda persona singolare è stato confuso con il plurale del nome “arrivo”, mentre il verbo “sei” è stato riconosciuto come aggettivo numerale.



Word_form	POS	Headword	Headword Edit
[S_START]			
partenza	NOM	partenza	partenza
comincia	VER:pres	cominciare	cominciare
vai	ADJ	vaio	andare
giri	NOM	giro	girare
a	PRE	a	a
destra	ADJ	destro	destra
sì	NOM	sì	sì
e	CON	e	e
arrivi	NOM	arrivo	arrivare
vai	ADJ	vaio	andare
avanti	ADV	avanti	avanti
sì	CON	sì	sì
subito	ADV	subito	subito
cioè	ADV	cioè	cioè
sei	ADJ	sei	sei
lì	ADV	lì	lì
e	CON	e	e
poi	ADV	poi	poi
subito	ADV	subito	subito
a	PRE	a	a
destra	NOM	destra	destra

Fig. 2 – Parte di un foglio elettronico prodotto in fase di lemmatizzazione

Sull'importanza della preventiva lemmatizzazione dei dati nel calcolo della diversità lessicale è stato detto molto. Malvern et alii (2004) sostengono che essa sia uno strumento fondamentale per contrastare l'influenza dei fattori interlinguistici (di cui abbiamo parlato nel paragrafo 2.4): pensiamo in particolare alle lingue morfologicamente più ricche, come l'italiano e le altre lingue romanze, in cui lo stesso lemma può comparire sotto moltissime forme flesse (il paradigma di un verbo italiano possiede circa novanta forme), in queste lingue un testo non lemmatizzato può generare

valori altissimi dell'indice di diversità lessicale, senza indicare una reale varietà nell'uso del lessico. La lemmatizzazione, dunque, rende più accurata la comparazione degli studi condotti su lingue tipologicamente distanti come l'italiano e l'inglese (in cui la categoria di genere è praticamente inesistente e il paradigma verbale molto ridotto rispetto a quello dell'italiano) e questa è senza dubbio un'esigenza dei ricercatori perché consente di avanzare nello studio della complessità linguistica sfruttando un'ampia base di dati provenienti da diverse L2.

Dopo aver corretto gli errori, il file viene caricato nuovamente nel software che questa volta calcola il TTR medio di una serie di campioni testuali (fino a 100) composti da 50 lemmi ciascuno<sup>21</sup>.

Ripetendo la procedura per ogni testo del corpus, abbiamo ottenuto gli indici di diversità lessicale relativi ad ogni apprendente e ad ogni task nei due anni di indagine. Infine, avendo a che fare con una grande quantità di valori numerici è stato necessario inserirli in un foglio elettronico, il quale ci ha consentito di lavorare efficientemente sui risultati, calcolando automaticamente medie e variazioni e creando grafici molto utili per comprendere l'entità dei fenomeni di nostro interesse.

---

<sup>21</sup> La dimensione dei segmenti di testo può essere impostata manualmente. Nel calcolo della complessità morfologica solitamente questo valore è impostato automaticamente su 10. In questo caso, trattandosi di lemmi e non di esponenti flessionali, ci sembra opportuno ampliare la quantità di *tokens* nei *samples* da 10 a 50.

## CAPITOLO 4

### *I risultati*

Nelle prossime pagine esporremo i risultati dell'indagine ottenuti tramite l'analisi descritta nel terzo capitolo e alcune annotazioni raccolte in fase di trascrizione.

Si è scelto di organizzare i dati come segue: in un primo momento illustreremo i risultati studentessa per studentessa, mentre nel secondo paragrafo tratteremo gli stessi dal punto di vista dei task. Alla fine di ognuna di queste sezioni si trova un sottoparagrafo che fornisce una visione di insieme grazie all'utilizzo dei valori medi, permettendo di individuare eventuali ricorsività e variazioni emerse dai risultati.

#### **4.1 I risultati individuali**

Il percorso di ogni apprendente viene descritto in primo luogo tramite una tabella, che riporta i valori numerici dell'indice di diversità lessicale registrato nei cinque momenti comunicativi: raccontare il contenuto di una storia per immagini, rispondere a domande personali, interagire per completare una mappa, negoziare con una compagna per prendere delle decisioni e gestire delle telefonate a sconosciuti. Successivamente, un grafico mostra in modo immediato la variazione avvenuta dal primo anno di indagine (2012) all'ultimo (2015).

Passati in rassegna i risultati delle singole studentesse, ci concentreremo sulle differenze o similarità riscontrate fra le apprendenti in merito ai livelli di diversità lessicale impiegati nella produzione linguistica, che come sappiamo è un importante indicatore della competenza lessicale e di quella generale.

## LALI

Nella performance di Lali è evidente una crescita della diversità lessicale dal 2012 al 2015 (indicati rispettivamente come T1 e T2), che interessa in modo generale tutti i task in cui l'apprendente è stata coinvolta.

In particolare, l'indice STTR rilevato in alcuni compiti comunicativi registra un aumento più consistente nel tempo: come appare chiaramente dal grafico, la *telefonata* e il *map task* sono le attività in cui la diversità lessicale, e di conseguenza la complessità, è cresciuta maggiormente dal 2012 al 2015.

Task	T1	T2
Story retelling	33,93	35,63
Intervista	38,09	38,46
Map task	32,33	34,7
Negoziazione	40,21	40,35
Telefonata	36,92	40,29

Tab. 3 – Risultati STTR di Lali nei diversi compiti comunicativi nel 2012 e nel 2015

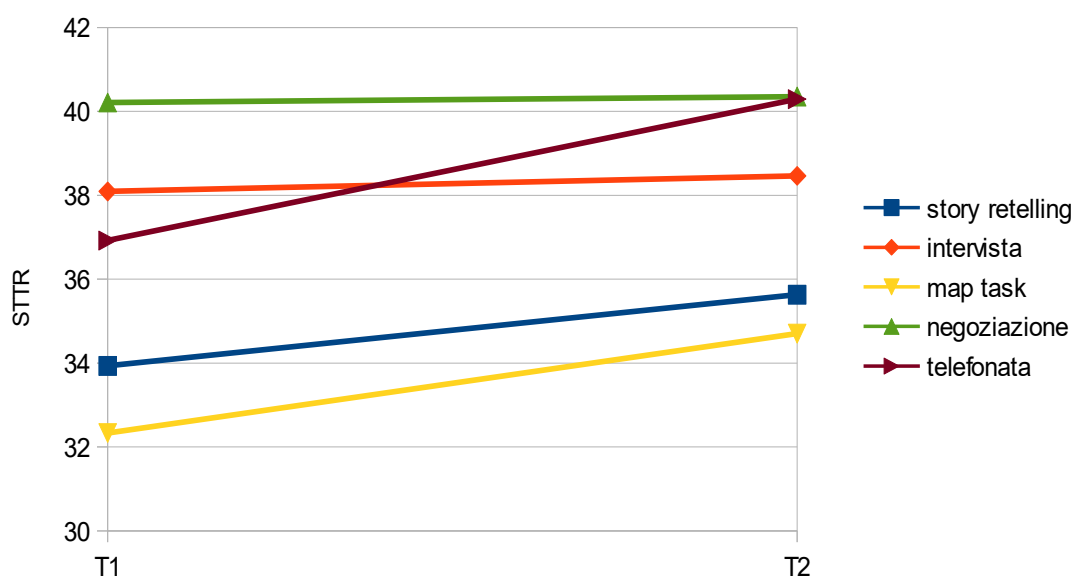


Grafico 1 – Variazione longitudinale dello STTR di Lali per tutti i task

Per identificare quali task abbiano richiesto a Lali un uso più diversificato del lessico, basterà osservare l'asse verticale T1 del grafico, che presenta una sequenza così disposta in ordine decrescente: *negoziazione* > *intervista* > *telefonata* > *story retelling* > *map task*. Lo stesso ordine rimane invariato in prospettiva longitudinale, ad eccezione dei risultati relativi alla *telefonata*, i quali crescono notevolmente nel 2015 arrivando a superare quelli dell'*intervista* ed eguagliare quelli della *negoziazione*, che al contrario si sono mantenuti piuttosto costanti nel tempo.

A questo punto, però, non possediamo ancora dati sufficienti per poter dichiarare se la scala di complessità dei task appena descritta sia da interpretare come frutto delle caratteristiche intrinseche dei task, ovvero *task complexity* e *task condition* per usare le categorie proposte da Robinson (paragrafo 1.4.3), o se al contrario sia da considerarsi una manifestazione della *task difficulty*, dipendente quindi da fattori individuali come lo stile cognitivo, la motivazione o la capacità di controllare l'ansia. Solo il confronto con i risultati ottenuti dalle altre apprendenti potrà fornirci più indizi a questo proposito: ad esempio, se assumiamo che la scala di complessità dei task sia costruita sulla base delle caratteristiche strutturali dei compiti stessi, che richiedono un impiego più o meno consistente di risorse cognitive, allora ci aspetteremmo verosimilmente che la scala dei task fosse identica per ogni apprendente.

Per concludere, in fase di trascrizione delle registrazioni abbiamo avuto l'occasione di ascoltare direttamente l'interlingua di Lali, sviluppando delle considerazioni qualitative. Fin dal 2012, la studentessa indiana possiede una buona padronanza della lingua italiana, in quanto ha trascorso già cinque anni nel paese frequentando le scuole medie e parte delle superiori. Nonostante evidenti difficoltà nella definizione del genere dell'articolo (visibili da esempi come: “il suo rana”, “gli altre rane”, “la proprio dialetto”) o l'ignoranza di alcune irregolarità (“un stagno”, “il hotel”), nel suo linguaggio fluente spiccano espressioni come “tutto sommato”, “oserei dire”, o esclamazioni tipiche dei dialetti emiliani come “oh mama”, ma non sempre vengono utilizzate in modo *native-like*: “c'è stile di vita” per dire che lo stile di vita è buono; “a chi vuoi chiamare?”; “sei stata in India lei?”, dove vediamo un tentativo di utilizzare il

pronomi di cortesia con l'intervistatrice adulta. Notiamo, inoltre, che durante la *negoziazione* compaiono una serie di espressioni del tipo “visitiamo il coso”, “chiama alla cosa” o addirittura “cosare” per sostituire un verbo x (che dimostra tra l'altro un uso produttivo del suffisso -are). Vista la selettività con cui vengono utilizzate queste espressioni, probabilmente l'apprendente è a conoscenza della loro connotazione informale e sceglie consapevolmente di farne uso soltanto durante la negoziazione, quando l'interazione si svolge con una coetanea.

## MONIKA

La produzione linguistica di Monika registra già a partire dal primo anno di indagine un indice di diversità lessicale leggermente superiore rispetto a quello di Lali. Questa “superiorità” si riscontra in modo generale su tutti i task svolti dalle apprendenti nel 2012.

<b>Task</b>	<b>T1</b>	<b>T2</b>
Story retelling	34,81	35,81
Intervista	38,53	39,67
Map task	33,73	34,51
Negoziazione	40,61	40,53
Telefonata	37,09	40,73

Tab. 4 – Risultati STTR di Monika nei diversi compiti comunicativi nel 2012 e nel 2015

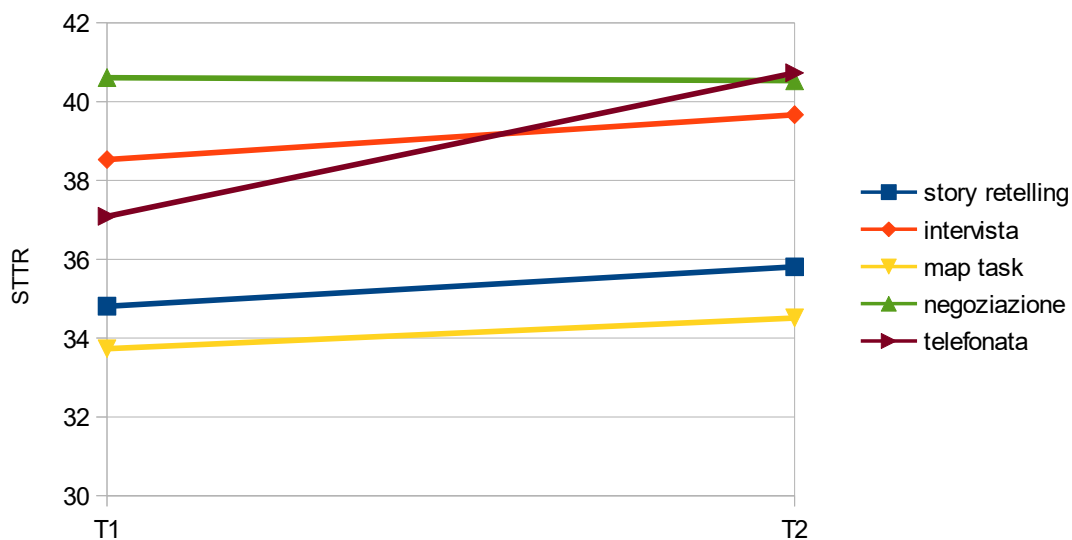


Grafico 2 – Variazione longitudinale dello STTR di Monika per tutti i task

Tuttavia, lo sviluppo longitudinale compiuto da Monika risulta inferiore rispetto alla crescita notevole registrata nei dati di Lali. In questo modo, l'iniziale distanza tra le due compagne viene colmata nel tempo, e nel 2015 il loro risultati appaiono simili.

Comparando il grafico 1 e il grafico 2 è evidente che la sequenza dei task nel 2012, ordinata in modo decrescente in base alla diversità lessicale generata nella produzione linguistica, è identica per le due apprendenti: *negoziiazione* > *intervista* > *telefonata* > *story retelling* > *map task*. Molto interessante è il fatto che, anche in questo caso, la complessità della *telefonata* cresce in modo esponenziale nel 2015, scavalcando l'*intervista* e raggiungendo i livelli registrati dalla *negoziiazione*. Due evoluzioni molto simili, quelle di Lali e Monika, che suggeriscono quindi l'influenza di fattori intrinseci ai task sui livelli di diversità lessicale prodotti dalle apprendenti.

Per quanto riguarda un'analisi più qualitativa del linguaggio, anche Monika rivela qualche difficoltà nell'utilizzo degli articoli italiani: questa volta però la studentessa tende ad ometterli (“sua rana”, “allora alce lo prese su tra le corna e lo buttò in lago”, “invece di rana”), probabilmente a causa dell'influenza della L1, vista l'assenza degli articoli nella lingua ucraina.

Molto ricorrenti nel parlato sono le espressioni “piuttosto” e “più che altro”,

segnali discorsivi di natura pragmatica di cui spesso Monika abusa: “più che altro abito a Guastalla”, “mi saprebbe consigliare qualche albergo con i costi piuttosto bassi?”, “ci potrebbe consigliare dei ristoranti piuttosto convenienti?”.

#### KATERINA

Quello che colpisce nella produzione linguistica di Katerina è un livello molto elevato di diversità lessicale, distribuito su tutti i task indifferenziatamente. Fin dal 2012, infatti, i valori dello STTR sono evidentemente superiori rispetto a quelli ottenuti dalle sue compagne.

Non si può dire lo stesso della crescita longitudinale: come mostra chiaramente il grafico 3, l'inclinazione ascendente delle linee non è paragonabile a quella visibile nei casi precedenti, inoltre, per quanto riguarda i task della *negoziazione* e dello *story retelling* si assiste addirittura a un calo dal 2012 al 2015.

Questo fenomeno potrebbe essere spiegato in diversi modi, ad esempio come conseguenza di un calo della motivazione nello svolgimento del task in un soggetto con una competenza linguistica molto avanzata, oppure come effetto di una più ampia tendenza di sviluppo.

Rimane invariata anche in questo caso la scala di complessità dei task nel 2012: *negoziazione* > *intervista* > *telefonata* > *story retelling* > *map task*.

<b>Task</b>	<b>T1</b>	<b>T2</b>
Story retelling	36,71	35,78
Intervista	38,87	39,65
Map task	35,41	37,33
Negoziazione	40,04	39,42
Telefonata	38,26	38,95

Tab. 5 – Risultati STTR di Katerina nei diversi compiti comunicativi nel 2012 e nel 2015



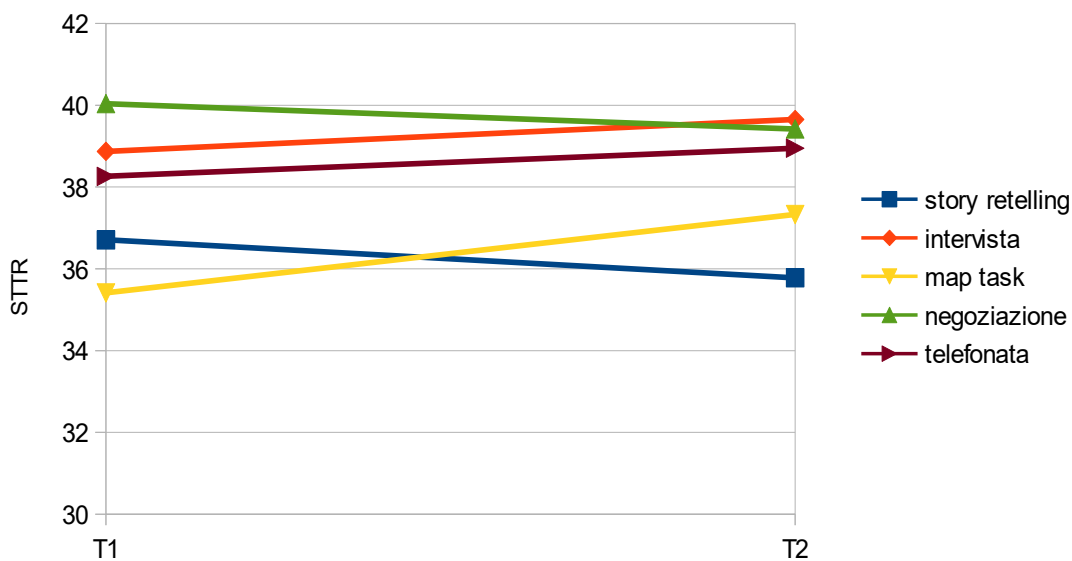


Grafico 3 – Variazione longitudinale dello STTR di Katerina per tutti i task

I dati osservati tramite l'analisi quantitativa vengono perfettamente rispecchiati nella percezione che si ha sentendo parlare Katerina: il linguaggio è molto ricco e fluente e anche dal punto di vista grammaticale le imprecisioni riguardano solo alcune costruzioni ipotetiche (“ogni momento che potrei ci ritornerei”, “se ci tornerei non avrei futuro”), tralaltro frequentemente oggetto degli stessi equivoci anche tra i parlanti nativi di italiano. Il discorso pullula di espressioni idiomatiche (come “rompersi in mille pezzi”, “con la coda tra le gambe”, “ci mancava solo” ecc...) che dimostrano un'ottima padronanza del lessico L2.

## RANDA

La diversità lessicale nei task svolti da Randa registra indici notevolmente più bassi rispetto a quelli delle altre apprendenti.

Anche in questo caso, le misurazioni sono congruenti con l'impressione ricevuta durante l'ascolto delle registrazioni. Pur essendo in Italia da più tempo rispetto alle altre ragazze, Randa presenta un eloquio difficoltoso e lento, forse anche a causa di un

carattere timido e insicuro, che la spinge a parlare con un tono di voce molto basso.

Le lacune lessicali vengono aggirate con l'uso della perifrasi: “la casa dei api” per indicare l'alveare, “la testa di un cervo” per indicarne le corna e così via. Il “cioè” viene ripetuto molto frequentemente nelle pause in cui l'apprendente deve riflettere su cosa dire, il “che” è polivalente (“il cane è qua nascosto che il bambino lo sta chiamando”, “il cane è spaventato che corre”) e il “dove” viene usato impropriamente come connettore (“dove il bambino e il cane cadono dentro a uno stagno e dove il cervo butta dentro lo stagno”).

L'ordine dei task per complessità decrescente nel 2012 è simile a quello delle altre apprendenti, ma vede lo slittamento della telefonata al primo posto: *telefonata* > *negoziazione* > *intervista* > *story retelling* > *map task*.

<b>Task</b>	<b>T1</b>	<b>T2</b>
Story retelling	33,73	35,84
Intervista	34,34	36,27
Map task	29,61	31,31
Negoziazione	35,22	33,81
Telefonata	35,76	34,36

Tab. 6 – Risultati STTR di Randa nei diversi compiti comunicativi nel 2012 e nel 2015

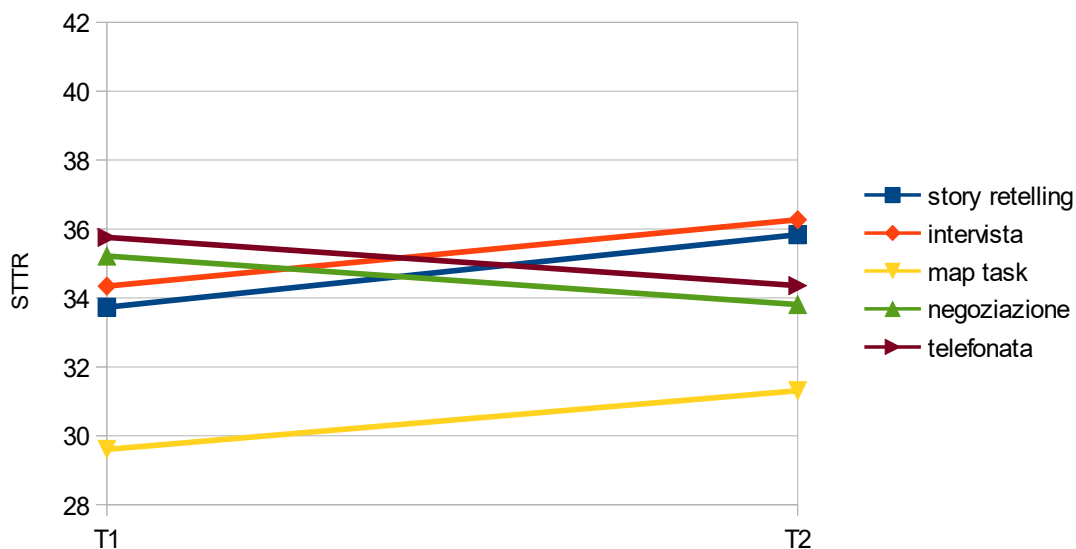


Grafico 4 – Variazione longitudinale dello STTR di Randa per tutti i task

Nell'intervallo temporale T1-T2 la complessità lessicale cresce in tre task su cinque, mentre si registra un calo nei due compiti comunicativi della *telefonata* e della *negoziiazione*. Ricordiamo che queste due attività sono state svolte all'interno dello stesso task *organizzare una gita*, che prevede la collaborazione tra Randa e Katerina. È possibile che nel tempo l'indole più decisa di Katerina abbia preso il sopravvento nell'organizzazione del viaggio, causando un calo motivazionale nell'altra studentessa. A questo proposito ci è utile osservare anche un'altra dimensione della complessità lessicale, ovvero quella che abbiamo chiamato “produttività” (paragrafo 2.1) e che consiste nel numero di tokens utilizzati dalle apprendenti per portare a termine il task. Questo dato, infatti, sottolinea il peso che ha avuto nel 2015 la produzione linguistica di Katerina nel momento della negoziazione, realizzata con un totale di 1573 parole, a discapito dello spazio concesso alla compagna, che ha prodotto solo 273 parole.

Possiamo quindi dedurre che anche la *task difficulty*, riguardante il rapporto tra il compito e le caratteristiche personali dell'individuo che lo svolge, ha in determinati casi una forte influenza.

#### ***4.1.1 I risultati medi: un confronto fra i soggetti***

I risultati individuali sopra esposti riscontrano delle differenze tra le apprendenti sia nel primo anno di indagine che nell'ultimo. Queste differenze possono essere meglio osservate grazie all'utilizzo di dati medi, che offrono una visione complessiva altrimenti difficile da comporre.

Riportiamo nella tabella 7 i valori STTR individuali ottenuti calcolando la media fra i risultati dei diversi task. Oltre ai dati del 2012 e del 2015 abbiamo aggiunto un valore numerico indicante la variazione tra i due momenti, riscontrabile anche, in maniera più immediata, nell'inclinazione delle linee del grafico 5 o nella crescita delle colonne dal primo al secondo istogramma.

<b>ID</b>	<b>T1</b>	<b>T2</b>	<b>variazione</b>
Lali	36,29	37,88	1,59
Monika	36,95	38,25	1,29
Katerina	37,85	38,22	0,36
Randa	33,73	34,31	0,58

Tab. 7 – STTR medio per ogni apprendente nel 2012 e nel 2015 e variazione longitudinale

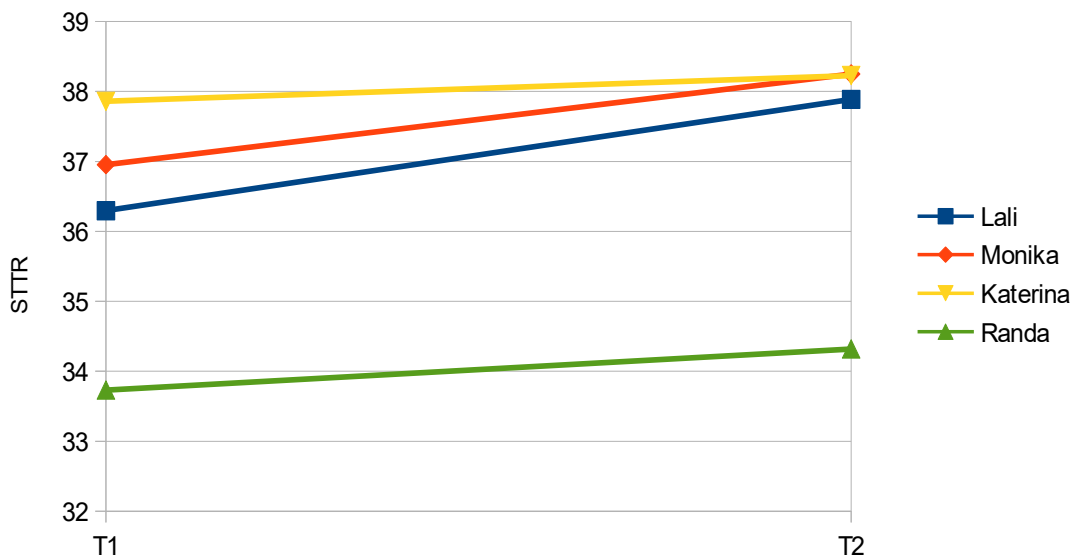


Grafico 5 – Variazione longitudinale dello STTR per ogni apprendente

Come vediamo dal grafico 5, nel momento T1 i livelli di complessità lessicale delle apprendenti si presentano eterogenei e dunque ben distribuiti sull'asse che riporta i valori STTR; tra il momento T1 e il momento T2 si assiste a una crescita della complessità nella produzione di ogni apprendente, infatti le linee del grafico presentano tutte un'inclinazione ascendente; la proporzione di questa crescita però non sembra essere casuale, in quanto nel momento T2 i livelli di complessità di tre apprendenti su quattro tendono a stabilizzarsi sullo stesso punto, come indicato dalla confluenza tra le linee gialla, rossa e blu. Per raggiungere questo punto d'arrivo comune, che si colloca approssimativamente sul valore 38 della scala STTR, le apprendenti non devono percorrere tutte la stessa strada: Katerina, ad esempio, che possiede già nel 2012 un livello di complessità abbastanza elevato, compie una crescita minima (0,36) nel corso degli anni per approdare alla meta; Lali, invece, deve affrontare uno sviluppo notevole (1,59) per arrivare allo stesso risultato; Randa, che già in partenza si distacca sensibilmente dalle altre, non raggiunge l'obiettivo.

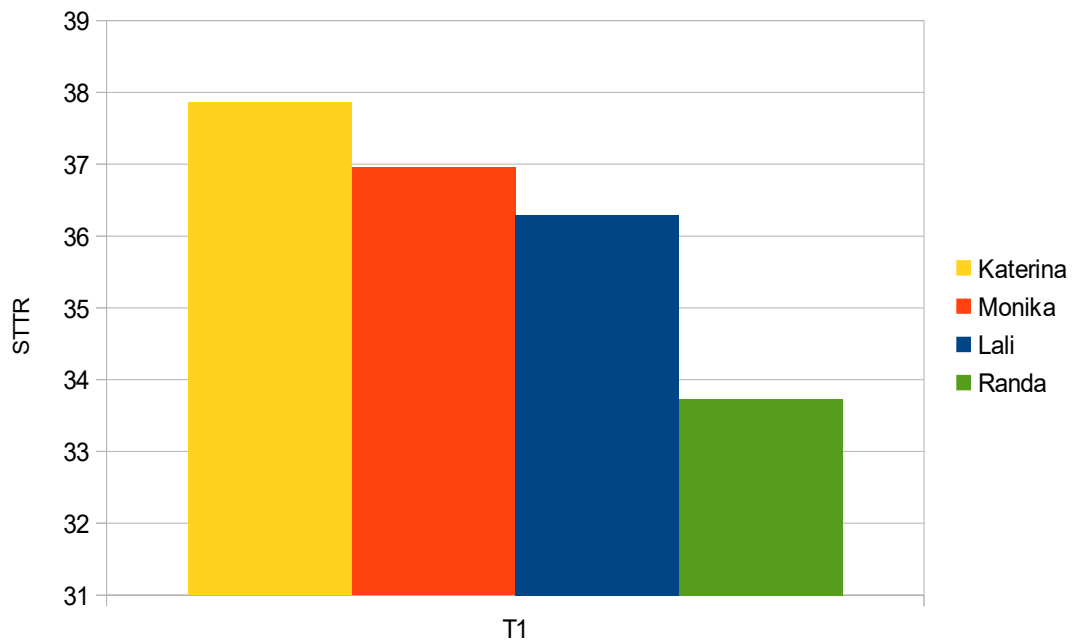
Ma a cosa corrisponde il punto d'arrivo comune? Per rispondere a questa domanda ci serviamo del gruppo di controllo, costituito da parlanti native. L'indice medio di

diversità lessicale della studentessa Giusy, parlante nativa di italiano, si colloca sul valore 37,85. Questo valore è molto vicino a quello raggiunto dalle parlanti L2 nel 2015, di conseguenza possiamo affermare che lo sviluppo lessicale delle apprendenti si muove nella direzione del livello di complessità richiesto dalla lingua target, e che le ragazze straniere tendono ad uniformare la propria performance a quella delle native.

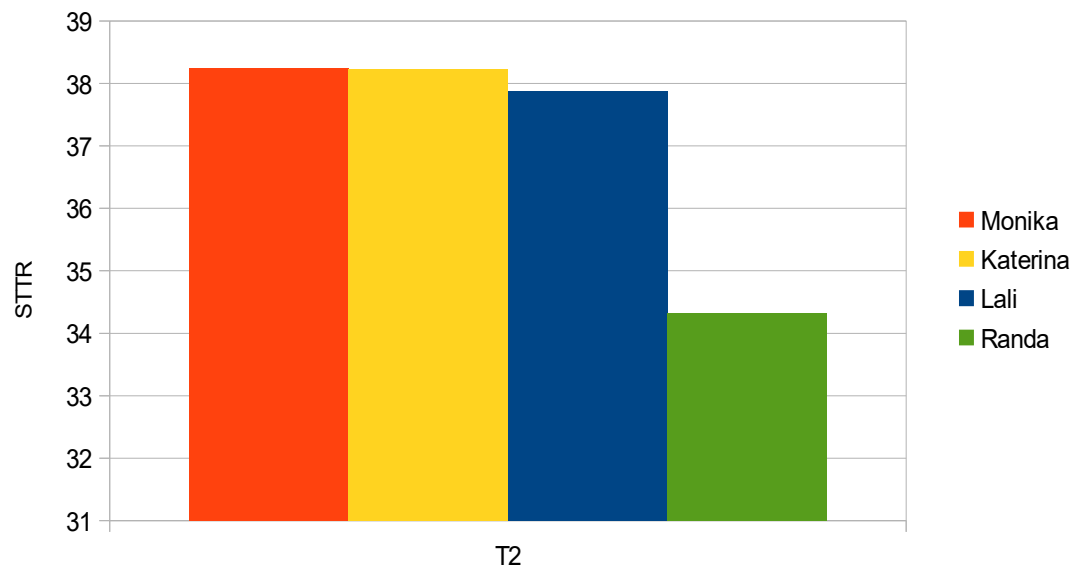
L'istogramma 1, nella pagina seguente, riporta i livelli di diversità lessicale della produzione linguistica delle quattro apprendenti (considerata come media dei risultati ottenuti nei cinque task svolti) nella prima sessione della raccolta dati. Come abbiamo visto nel secondo capitolo, la diversità lessicale è strettamente correlata alla competenza (*proficiency*), dunque la classifica che emerge da questo diagramma ci dà un'idea, anche se non priva di approssimazione, delle differenze di competenza linguistica preesistenti tra le apprendenti all'inizio della sperimentazione.

Non essendovi fra i soggetti un sostanziale scarto nell'età o nel tempo trascorso in Italia, le differenze interindividuali di competenza lessicale (e linguistica in generale) possono essere spiegate come conseguenze del contesto di apprendimento della L2 o come effetto della distanza tra la lingua materna e l'italiano.

La prima ipotesi riguarda il tipo di input linguistico al quale le apprendenti sono state esposte nel contesto in cui hanno appreso la lingua seconda, sia esso un contesto formale (un'istituzione scolastica) o informale (le molteplici situazioni legate alla vita quotidiana). Le caratteristiche linguistiche (e in questo caso specialmente la diversità del lessico) del materiale educativo utilizzato a scuola o dell'ambiente sociale frequentato nel tempo libero potrebbero essere responsabili di uno sviluppo più o meno accentuato della complessità lessicale nei soggetti. In questa ricerca risulta impossibile ricostruire nel dettaglio i singoli contesti di apprendimento delle ragazze negli anni che precedono l'inizio della sperimentazione, tuttavia possiamo supporre che non vi siano differenze importanti, dal momento che tutte le studentesse, una volta arrivate nel paese, si sono integrate nel sistema scolastico italiano, dove hanno seguito i primi corsi di alfabetizzazione, e al momento dell'indagine frequentavano lo stesso istituto tecnico superiore.



Istogramma 1 – Ordine delle apprendenti per complessità lessicale nel 2012



Istogramma 2 – Ordine delle apprendenti per complessità lessicale nel 2015

Rimane quindi da valutare il ruolo della L1 nello sviluppo della complessità lessicale della lingua seconda. Molti linguisti hanno espresso la necessità di approfondire gli studi sullo sviluppo della complessità linguistica includendo dati longitudinali e provenienti da individui parlanti di diverse L1, tipologicamente distanti tra loro. La nostra ricerca offre entrambe le prospettive, anche se l'ampiezza del corpus non ci permette di trarre conclusioni a tale proposito. In ogni caso, possiamo avanzare alcune considerazioni puramente descrittive: prendiamo ad esempio la colonna gialla e quella verde dell'istogramma 1, relative alla complessità della produzione linguistica rispettivamente di Katerina e di Randa; la prima supera nettamente la seconda, questo significa che la performance di Katerina ha raggiunto livelli di STTR considerevolmente superiori in ognuno dei task svolti. Se abbiamo una conoscenza anche solo vaga delle lingue materne parlate da queste ragazze, sappiamo che il rumeno, L1 di Katerina, è una lingua romanza proprio come l'italiano e dunque, in termini genealogici, esse si dicono "lingue sorelle"; il tamil, invece, lingua materna di Randa, non solo non è una lingua romanza ma non appartiene nemmeno alla stessa famiglia linguistica (delle lingue indoeuropee), facendo parte delle lingue dravidiche parlate principalmente nel sud dell'India e in Sri Lanka, paese di origine della ragazza. Questo dato, già di per sé, è da considerare almeno in parte responsabile della distanza tra i risultati delle due apprendenti, poiché Katerina è indubbiamente avvantaggiata dalle numerose similarità, sia nel lessico che nella grammatica, tra la lingua materna e la lingua target. Il discorso è un po' più complicato per quanto riguarda Monika e Lali: dal punto di vista della classificazione genealogica, la lingua ucraina e la lingua punjabi appartengono entrambe alla grande famiglia indoeuropea, più specificamente la prima è una lingua slava e la seconda è una lingua indoiranica, per cui nessuna delle due ha un legame stretto con l'italiano; secondo una classificazione tipologica, l'ucraino è una lingua del tipo flessivo<sup>22</sup> con ordine SVO<sup>23</sup>, esattamente come l'italiano, mentre il punjabi è una lingua agglutinante<sup>24</sup> con ordine SOV<sup>25</sup> e per di più è una lingua tonale,

---

22 Questo tipo morfologico si caratterizza per la presenza di morfemi che esprimono più di una relazione grammaticale (Graffi & Scalise, 2003).

23 La sigla SVO significa "Soggetto, Verbo, Oggetto" e indica l'ordine sintattico al quale abitualmente aderiscono i costituenti in una data lingua.

24 In questo tipo morfologico le parole sono costituite da una radice a cui si aggiungono più morfemi che non vengono mai fusi tra di loro (Graffi & Scalise, 2003).

25 In questo caso l'ordine dei costituenti nella frase è "Soggetto, Oggetto, Verbo".



caratteristiche piuttosto insolite per un idioma indoeuropeo. La vicinanza tipologica che rende più simili l'ucraino e l'italiano di quanto non lo siano l'italiano e il punjabi o il tamil va a sommarsi ad un criterio più strettamente geografico, per cui la lingua ucraina, vista la sua distribuzione geografica, ha avuto storicamente più possibilità di contatto con l'italiano e le altre lingue romanze, rispetto alle lingue parlate in India.

Oltre all'influenza della L1, poi, dobbiamo sempre tenere conto di una serie di caratteristiche individuali che possono aver condizionato la produzione linguistica: come lo stile cognitivo, l'indipendenza, la motivazione, o la capacità di controllare l'ansia. Le difficoltà di Randa, ad esempio, sono dovute in parte alla distanza tra la lingua materna e la lingua da apprendere, ma in parte dipendono dal carattere introverso della ragazza.

Nell'istogramma 2, vediamo gli effetti dello sviluppo della complessità lessicale sulla produzione linguistica del 2015: l'ordine delle apprendenti si è mantenuto sostanzialmente invariato, ma quello che interessa è che, indipendentemente dalle disparità iniziali, Monika, Katerina e Lali raggiungono nel 2015 praticamente lo stesso livello di complessità lessicale. Ciò appare evidente dall'allineamento delle prime tre colonne, come se sopra di esse si appoggiasse una specie di “tetto”, oltre il quale non è possibile progredire. Proprio di *ceiling effect*<sup>26</sup> parlano i ricercatori nel campo dell'acquisizione della lingua seconda: Daller et al. (2013), per esempio, ritengono che lo sviluppo lessicale (e in particolare della diversità) possa essere rappresentato da una curva di potenza che nelle prime fasi dell'apprendimento presenta una ripida crescita, ma raggiunte le fasi più avanzate si appiattisce. In questo senso, l'apprendimento del lessico rispecchia lo stesso percorso osservato nell'apprendimento linguistico più in generale, per cui gli apprendenti con un livello di competenza linguistica più avanzato (come, nel nostro caso, Katerina) mostrano una crescita longitudinale meno significativa. Come affermano Daller et al. (2013), questa constatazione ha delle conseguenze didattiche nella progettazione dei corsi di lingua: una volta che gli apprendenti hanno raggiunto un certo “plateau”, solo un nuovo input, per esempio in forma di lezioni extra, può migliorare l'ulteriore apprendimento del lessico. È molto importante, dunque, per gli operatori individuare il momento in cui questo punto di

<sup>26</sup> In statistica il *ceiling effect* viene definito come quel fenomeno che si osserva quando una variabile indipendente non ha più nessun effetto su una variabile dipendente.

stallo è stato raggiunto, per poter compiere scelte più consapevoli sulle successive strategie di insegnamento.

Tornando al nostro caso, il livello uniforme raggiunto da tre delle studentesse nel 2015, sembra essere un punto di stallo giustificato dal fatto che la diversità lessicale delle apprendenti, a quel punto, coincide con quella delle native. Tuttavia è necessaria una certa cautela, in quanto diversi linguisti hanno messo in guardia a proposito dell'interpretazione dei dati provenienti dai gruppi di controllo. Nel nostro caso, ad esempio, i dati relativi alle native sono stati raccolti una sola volta nel 2012, quando le ragazze avevano quindici anni. Di conseguenza, nel comparare questi dati con quelli delle non native raccolti nel 2015, si mette a confronto la produzione linguistica di ragazze quindicenni (native) con quella di ragazze di diciotto o diciannove anni (non native). Johansson (2008) afferma che la complessità lessicale nella L1 subisce lo sviluppo più consistente tra i tredici e i diciassette anni, perciò dobbiamo tener conto del fatto che le native, al momento della raccolta dati, si trovavano ancora in una fase di transizione dello sviluppo cognitivo e linguistico, ragione per cui il loro indice di complessità lessicale non rispecchia in modo ideale il livello di un parlante nativo adulto e ogni comparazione deve essere fatta tenendo conto di questo aspetto.

## **4.2 I risultati per task**

Per avere una visione quanto più completa dei risultati dell'indagine, assumiamo ora un altro punto di vista, quello dei singoli task proposti alle partecipanti.

Ogni task ha le sue caratteristiche e richiede un determinato livello di complessità lessicale per essere svolto. Una volta esposte le specificità di ogni task e i relativi dati tramite l'ausilio di tabelle e grafici, vedremo quali sono i compiti comunicativi che elicitano una maggiore diversità lessicale, sia in prospettiva sincronica che in termini di sviluppo longitudinale. Queste osservazioni ci condurranno a possibili implicazioni pedagogiche.

## STORY RETELLING

Il racconto di una storia per immagini, essendo un compito prevalentemente monologico, ha permesso alle studentesse di organizzare il discorso in maniera autonoma. Alcuni studi riguardanti l'effetto dei task sull'output L2 hanno dimostrato che le narrazioni monologiche rigorosamente strutturate, come la *frog story*, riducono lo sforzo di elaborazione, generando una produzione linguistica più fluente e accurata. In termini di complessità, invece, l'output risultante da questo tipo di narrazioni risulta meno complesso rispetto a quello elicitato dai task dialogici (Skehan, 2001).

Infatti, i valori STTR relativi a questo compito comunicativo sono inferiori rispetto a quelli registrati in tutti gli altri task, ad eccezione del *map task*.

ID	T1	T2
Lali	33,93	35,63
Monika	34,81	35,81
Katerina	36,71	35,78
Randa	33,73	35,84

Tab. 8 – Risultati STTR nel task "story retelling" nel 2012 e nel 2015

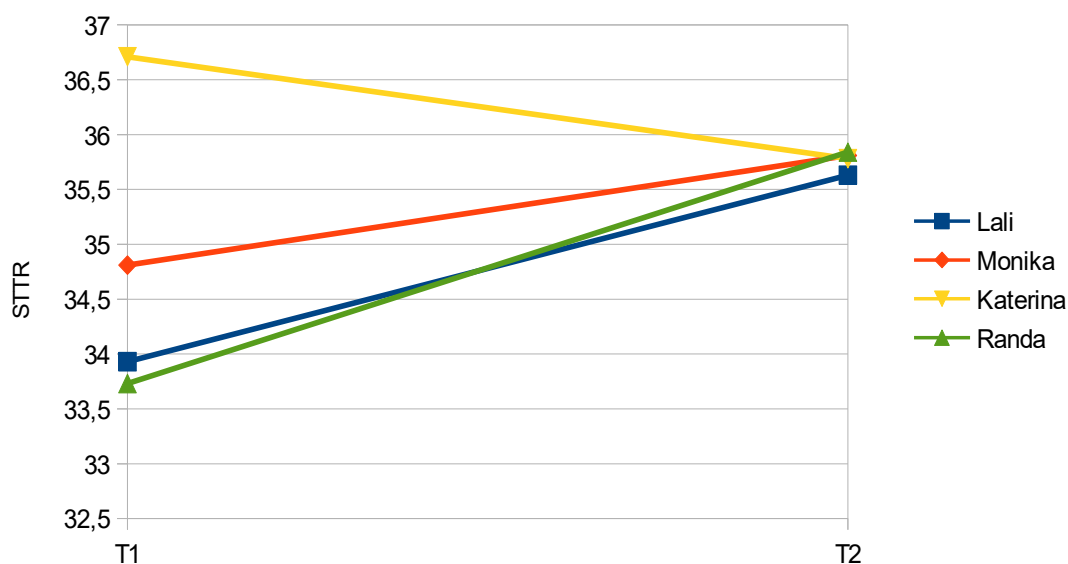


Grafico 6 – Variazione longitudinale dello STTR nel task "story retelling" per le diverse apprendenti

Come appare dal grafico 6, tutte le apprendenti svolgono il task con complessità crescente dal 2012 al 2015, tranne Katerina che, pur presentando l'indice di diversità più elevato nel 2012, mostra un andamento decrescente.

Quello che incuriosisce nel grafico è il fatto che, indipendentemente dall'entità della crescita individuale, o nel caso di Katerina del calo, nel 2015 la complessità lessicale delle apprendenti si stabilizza intorno allo stesso livello, come indicato dalla confluenza di tutte le linee.

Questo livello corrisponde a un valore STTR che si colloca tra 35,5 e 36. Un rapido sguardo ai dati provenienti dal gruppo di controllo ci permette di constatare che le native producono in questo task risultati che si collocano esattamente all'interno di quell'intervallo. Questo dato ci permette di sviluppare nuove ipotesi sulla causa del calo di complessità nella produzione di Katerina, che nella prima analisi dei risultati individuali avevamo attribuito a una mancanza di motivazione: a questo punto, il calo sembra piuttosto un allineamento agli standard nativi. È possibile, infatti, che nella prima somministrazione del task la studentessa fosse impaziente di dimostrare la propria competenza linguistica nella seconda lingua, raggiungendo così livelli di complessità superiori a quelli richiesti per portare a termine il task. Tuttavia, riprendendo la citazione di Ortega (vedi paragrafo 1.6), “più complesso non significa necessariamente migliore”, così nell'ultima performance la studentessa può aver optato giustamente per un linguaggio più semplice ma più consono al task, preferendo alla complessità assoluta l'adeguatezza comunicativa e l'aderenza a parametri *native-like*.

## INTERVISTA

Uno sguardo ai numeri ci permette di osservare che i livelli di diversità lessicale elicitati dal task dell'*intervista* sono considerevolmente superiori rispetto a quelli emersi dallo *story retelling*.

Questa differenza non è attribuibile alla caratteristica  $\pm$ dialogico, in quanto entrambi i task sono prevalentemente monologici. Presentano, tuttavia, alcune differenze strutturali che potrebbero essere responsabili del divario osservato nei risultati. Innanzitutto, l'*intervista* non impone alcuna restrizione al lessico utilizzato per

svolgere il task poiché, a parte qualche domanda posta dall'intervistatrice per guidare la conversazione, le apprendenti sono libere di scegliere cosa raccontare e con quali termini; al contrario, la *frog story* fornisce alla narrazione un tema prestabilito, limitando inevitabilmente le scelte lessicali delle apprendenti. Oltre a ciò, l'*intervista* propone ai soggetti una serie di atti comunicativi (presentarsi, parlare dello stato psichico, esprimere i propri gusti) che realizzano la “funzione personale”, quell'aspetto della competenza socio-pragmatica che entra in gioco quando lo studente rivela la propria personalità e manifesta i propri sentimenti (Balboni, 2015). Questa funzione, stimolando il piano dell'affettività e della motivazione, può elicitare una produzione linguistica più complessa rispetto al racconto di una storia che non ha nessun legame con la vita delle studentesse.

ID	T1	T2
Lali	38,09	38,46
Monika	38,53	39,67
Katerina	38,87	39,65
Randa	34,34	36,27

Tab. 9 – Risultati STTR nel task "intervista" nel 2012 e nel 2015

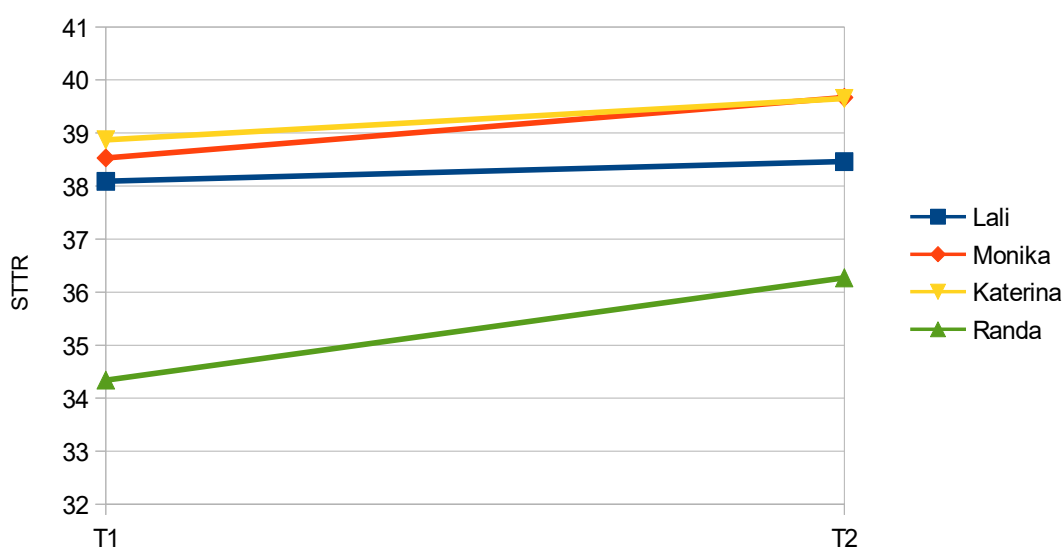


Grafico 7 – Variazione longitudinale dello STTR nel task "intervista" per le diverse apprendenti

Il grafico 7 mostra un aumento uniforme della complessità lessicale dal 2012 al 2015 per tutte le apprendenti. Specialmente Randa cresce nel tempo, anche se non arriva ad eguagliare i risultati delle compagne.

Il livello raggiunto da Monika e Katerina nel 2015 è quello che più si avvicina ai valori riscontrati presso le studentesse native: Giusy presenta nel task dell'*intervista* un indice di diversità lessicale di 39,55, mentre Marta di 40,69.

## MAP TASK

Il map task è stato definito un compito prevalentemente dialogico poiché richiede una partecipazione attiva e un costante scambio di informazioni in entrambe le direzioni: il prodotto è infatti un testo composto da numerosi turni di parola, nel complesso brevi ma anche più lunghi a seconda dell'esigenza informativa di ogni momento. Inoltre, la scelta di eliminare il contatto visivo tra gli interlocutori ha reso più difficile la comunicazione poiché lo sguardo viene usato normalmente come un segnale discorsivo per gestire l'alternanza dei turni (Pallotti e Rosi, 2015): il risultato è un aumento delle sovrapposizioni tra i turni di parola delle due partecipanti.

L'assenza del linguaggio non verbale precluso dall'ostacolo visivo ha spinto gli interlocutori a ricercare frequentemente conferme e chiarimenti verbali tramite l'uso dell'interrogazione e della ripetizione.

Le caratteristiche del task hanno influenzato la produzione linguistica in un altro modo: la descrizione di un percorso disegnato su una mappa implica necessariamente l'utilizzo frequente dei nomi indicanti i punti di riferimento della mappa, ma anche dei verbi di moto, preposizioni e avverbi di luogo, che hanno aumentato il tasso di ripetizione nella produzione linguistica delle apprendenti.

Per questi motivi la diversità lessicale nel *map task* sembra essere inferiore rispetto a quella rilevata negli altri task dell'indagine, collocandosi intorno a un valore STTR medio di 34. Per appurare che la scarsa complessità prodotta dipenda dalla struttura stessa del task più che dalle difficoltà delle apprendenti, operiamo un confronto con i risultati delle parlanti native: Giusy totalizza 34,41, mentre Marta 31,01.

Dal punto di vista diacronico, tutte le studentesse affrontano una crescita consistente nel corso del tempo, di circa due punti della scala STTR.

Se Lali e Monika si muovono intorno a valori congruenti con quelli delle native, Katerina sceglie invece di utilizzare una complessità maggiore, che eccede la quantità richiesta per realizzare l'obiettivo comunicativo prefissato da questo task nella lingua target. Trattando la nostra ricerca uno stadio avanzato dell'apprendimento L2 non è strano incontrare questo tipo di situazioni, poiché il sistema interlinguistico su cui fanno affidamento alcune apprendenti gli concede un ampio margine di variazione stilistica individuale.

ID	T1	T2
Lali	32,33	34,7
Monika	33,73	34,51
Katerina	35,41	37,33
Randa	29,61	31,31

Tab. 10 – Risultati STTR nel task "map task" nel 2012 e nel 2015

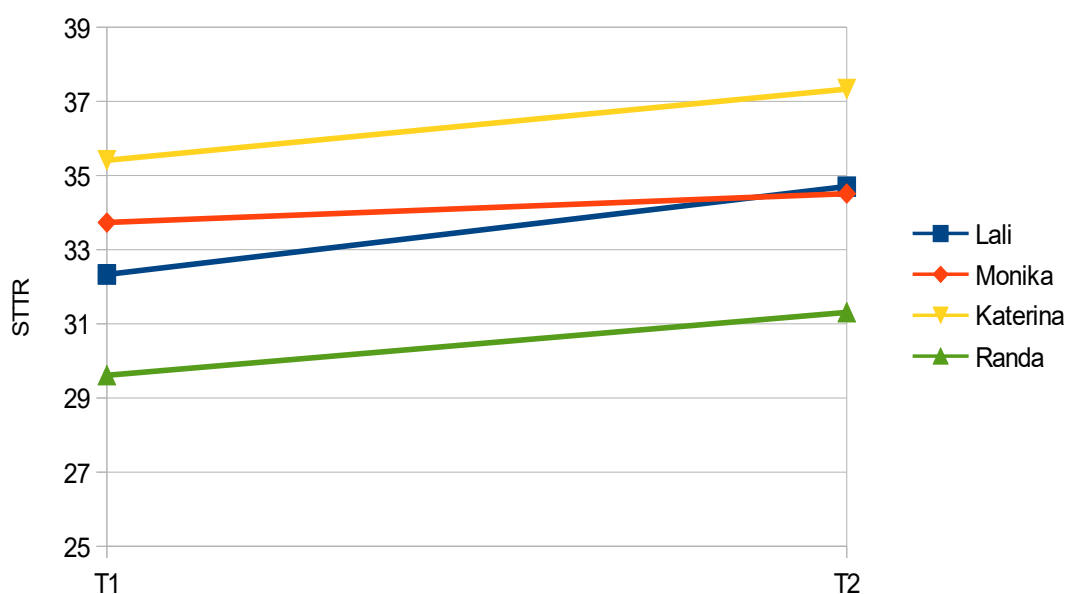


Grafico 8 – Variazione longitudinale dello STTR nel task "map task" per le diverse apprendenti

## NEGOZIAZIONE

Il task *organizzare una gita* è un compito altamente dialogico e la *negoziazione* è la parte che consiste nell'interazione faccia-a-faccia tra le due apprendenti. Lo scopo del dialogo è quello di esprimere le proprie preferenze e fare delle scelte condivise che portino alla stesura di un piano organizzativo della gita. L'interazione, quindi, può riguardare una vasta gamma di argomenti ed è completamente gestita dalle apprendenti senza l'aiuto di un intermediario. Questa libertà, unita al fatto che le agenti della comunicazione sono ragazze coetanee ed entrambe non native, dà vita ad un tipo di interazione estremamente naturale e realistica, probabilmente molto simile al parlato spontaneo delle apprendenti negli scambi quotidiani. L'inserimento della competizione fra le due coppie per la vittoria di un premio ha fornito uno stimolo in più alla collaborazione fra le apprendenti, che erano invogliate così a comprendersi ed esprimersi al meglio al fine di raggiungere il traguardo.

Questi fattori hanno portato le studentesse a produrre un linguaggio molto più complesso in termini di diversità lessicale rispetto a quello utilizzato negli altri task.

Ad esclusione di Randa, i dati relativi alla *negoziazione* mostrano un contesto piuttosto omogeneo. La distanza fra i risultati delle diverse apprendenti è minima sia nel 2012 che nel 2015. Anche l'evoluzione longitudinale non evidenzia particolari movimenti: la complessità rimane pressoché invariata e nel caso della coppia Katerina-Randa diminuisce. Come abbiamo già constatato nell'analisi dei dati individuali, Randa in questo compito ha una produttività scarsissima poiché lascia che sia la compagna più vivace a prendere le decisioni e organizzare la gita. In questo caso, quindi, le differenze caratteriali e di competenza tra le due ragazze sono state un ostacolo alla collaborazione, influenzando l'esito del task.

L'assenza di uno sviluppo diacronico importante si può giustificare con il fatto che già nella prima realizzazione Monika, Lali e Katerina producono un valore STTR simile a quello prodotto dalle parlanti native (40,31 per Giusy e 39,59 per Marta).



ID	T1	T2
Lali	40,21	40,35
Monika	40,61	40,53
Katerina	40,04	39,42
Randa	35,22	33,81

Tab. 11 – Risultati STTR nel task "negoziiazione" nel 2012 e nel 2015

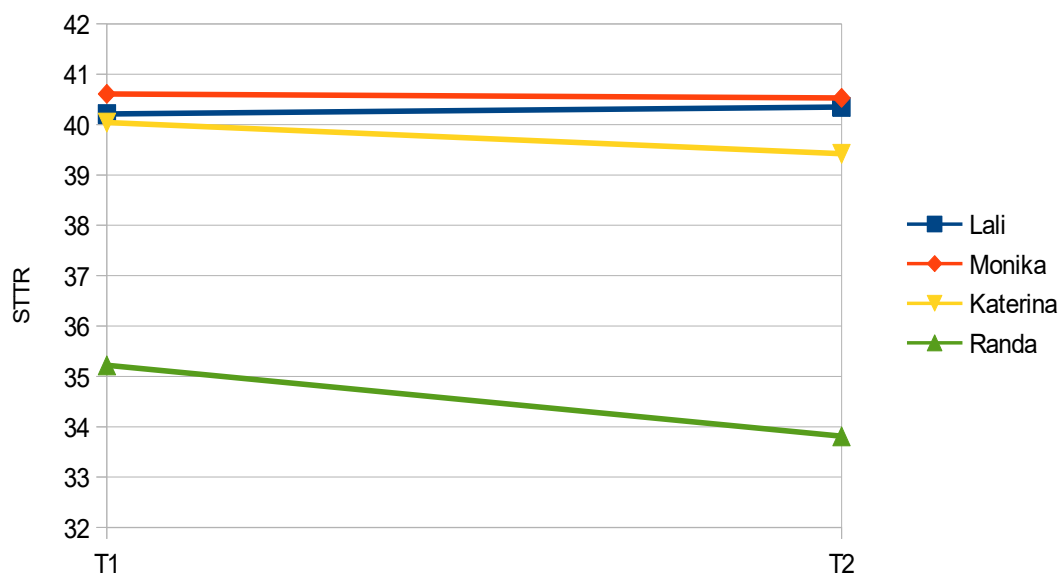


Grafico 9 – Variazione longitudinale dello STTR nel task "negoziiazione" per le diverse apprendenti

## TELEFONATA

Nel task *organizzare una gita* veniva chiesto alle partecipanti di fare circa cinque telefonate a testa per ottenere informazioni utili a definire il programma di viaggio. Gli interlocutori potevano essere coetanei o adulti, maschi o femmine, “esperti viaggiatori” o enti locali come uffici del turismo, musei e ristoranti<sup>27</sup>.

ID	T1	T2
Lali	36,92	40,29
Monika	37,09	40,73
Katerina	38,26	38,95
Randa	35,76	34,36

Tab. 12 – Risultati STTR nel task "telefonata" nel 2012 e nel 2015

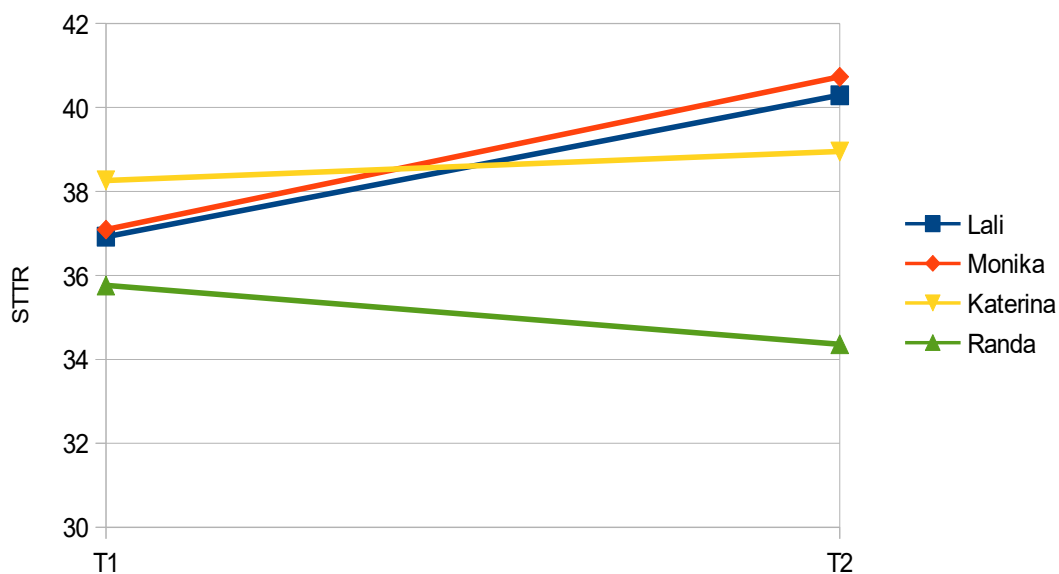


Grafico 10 – Variazione longitudinale dello STTR nel task "telefonata" per le diverse apprendenti

<sup>27</sup> I soggetti che ricevevano le chiamate erano, nel caso degli "esperti", persone preventivamente informate del progetto dall'università, mentre nei restanti casi si trattava di normali telefonate.

La *telefonata* è un task che presenta numerose sfide alle partecipanti: prima di tutto, una telefonata ad uno sconosciuto può essere già di per sé una fonte di stress per molte persone, a maggior ragione per un adolescente e parlante non nativo; all'assenza del contatto visivo in questo caso si sommavano anche i normali disturbi di una telecomunicazione; inoltre, le mete del viaggio da organizzare erano il primo anno La Spezia e il secondo anno Salerno, perciò le telefonate erano indirizzate a parlanti di altre regioni d'Italia che possedendo un accento diverso potevano mettere a rischio la comprensione delle non native. Per quanto riguarda gli “esperti”, invece, le ragazze potevano scegliere liberamente chi chiamare e avevano a disposizione per ciascuno di loro una breve scheda informativa su cui erano riportati il nome, l'età e la professione. Questo ci ha permesso di osservare che alcune apprendenti, con lo scopo di diminuire l'ansia del compito, preferivano optare per un interlocutore di sesso femminile, sia tra i coetanei che tra gli adulti.

La complessità del task ha avuto l'effetto di generare una produzione linguistica molto complessa dal punto di vista lessicale, specialmente nell'ultima somministrazione del task. Lo sviluppo è stato notevole dal 2012 al 2015, per tutte le apprendenti meno che per Randa, i cui risultati sono in linea con quelli della *negoziazione* avvenuta in concomitanza.

### 4.2.1 I risultati medi: un confronto fra i task

Nella tabella 13 si riporta un valore complessivo STTR per ogni task, calcolato operando la media matematica fra i risultati delle singole apprendenti. I dati medi offrono una sintesi di quanto osservato dettagliatamente in questo capitolo, fornendo una visione complessiva dei risultati dell'analisi che ci permette di far emergere le linee di tendenza principali.

<b>Task</b>	<b>T1</b>	<b>T2</b>	<b>variazione</b>
Story retelling	34,79	35,76	0,97
Intervista	37,45	38,51	1,05
Map task	32,77	34,46	1,69
Negoziazione	39,02	38,52	-0,49
Telefonata	37	38,58	1,57

Tab. 13 – STTR medio per ogni task nel 2012 e nel 2015 e variazione longitudinale

Nell'ultima colonna della tabella viene esplicitata la variazione esatta di complessità lessicale che intercorre tra il momento T1 e il momento T2. Lo stesso dato numerico è trasposto nel grafico 11 come inclinazione più o meno accentuata delle linee relative ad ogni task.

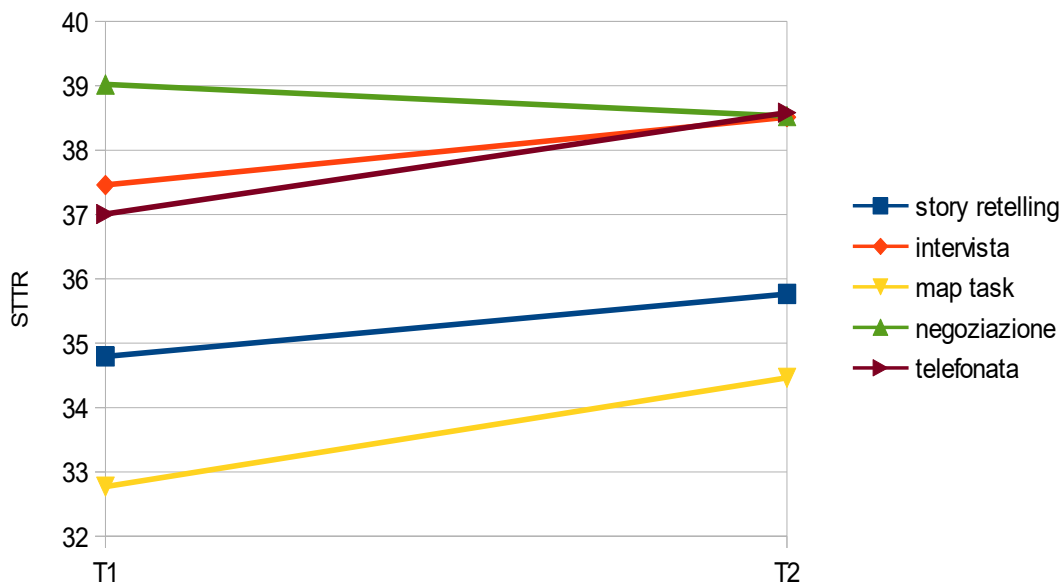


Grafico 11 – Variazione longitudinale dello STTR per ogni task

La conclusione più evidente è che in quasi tutti i compiti comunicativi si rileva un sostanziale incremento dell'indice di diversità lessicale lungo i quattro anni trascorsi dall'inizio della sperimentazione.

Sembra, dunque, che con l'avanzare del tempo e della competenza L2 il lessico delle apprendenti continui a complessificarsi, seguendo uno sviluppo omogeneo e significativo. Questo dato è in parte inaspettato se si considera che diverse ricerche hanno evidenziato un rallentamento della crescita e la presenza di lunghi periodi di stagnazione nell'apprendimento del lessico a livelli avanzati di competenza (Daller et al., 2013). Probabilmente l'osservazione di una tale evoluzione è stata possibile grazie alla considerevole distanza temporale tra i primi e gli ultimi dati analizzati in questa ricerca. Inoltre, non ci è possibile calcolare in quale misura l'aumento della diversità lessicale sia dovuto all'avanzamento della competenza generale delle apprendenti nella lingua target, o sia dipendente dallo sviluppo cognitivo al quale gli individui, nativi e non nativi, sono soggetti in questa fase delicata dell'adolescenza che ha importanti

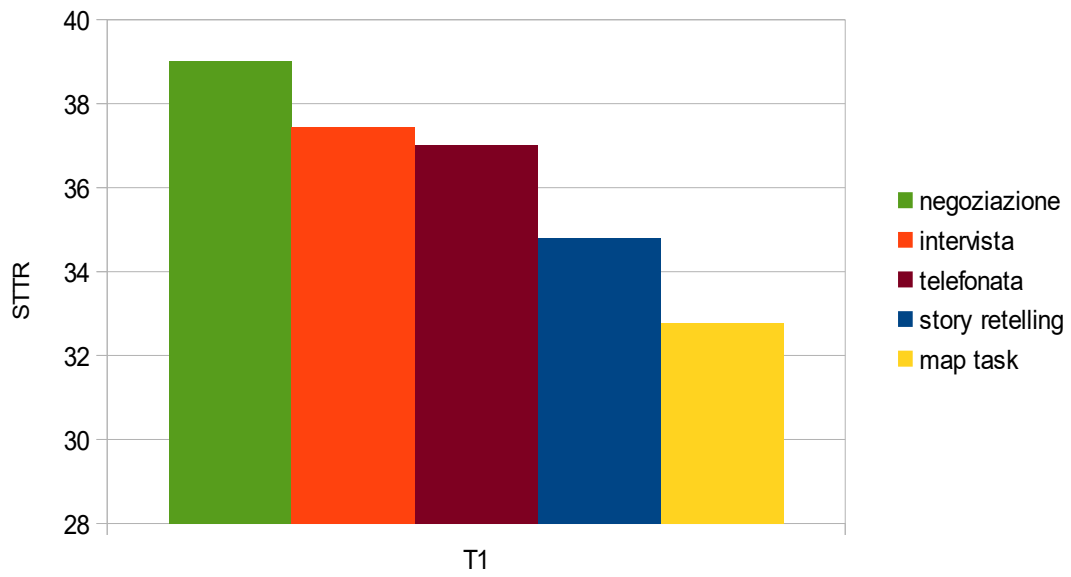
ricadute sullo sviluppo linguistico (vedi paragrafo 2.4).

Parlando in termini di sviluppo longitudinale, la crescita della complessità è particolarmente accentuata nel *map task* e nella *telefonata*, che registrano una variazione rispettivamente di 1,69 e 1,57 sulla scala STTR. Questi due task sono gli unici del nostro corpus ad essere accomunati dalla variabile dell'assenza di contatto visivo tra gli interlocutori.

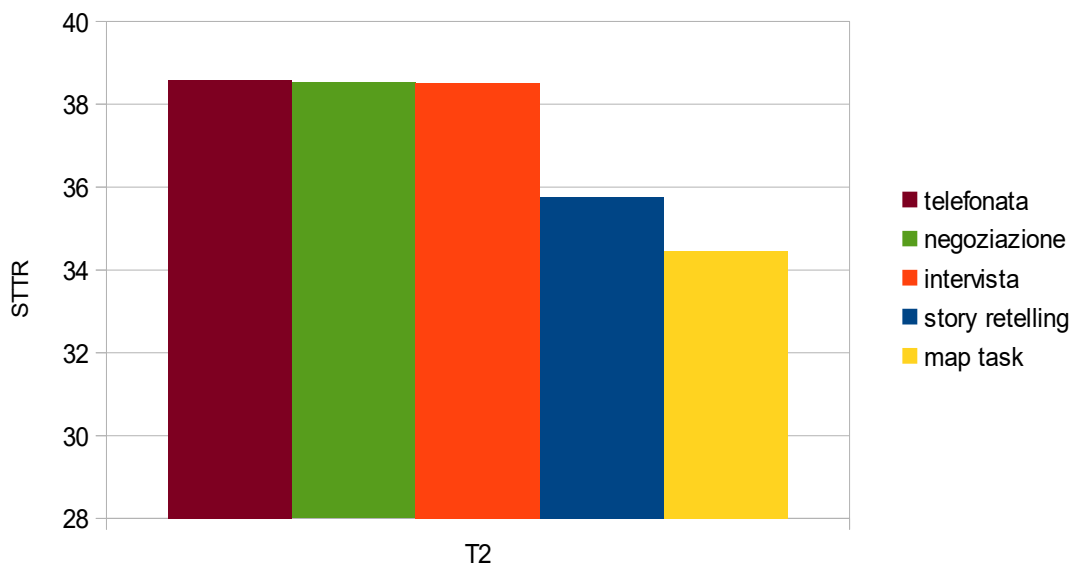
Come già riscontrato in studi precedenti, lo *story retelling* registra una crescita scarsa nel tempo, presumibilmente dovuta ad un calo della motivazione nella realizzazione del task, forse non adeguatamente impegnativo e stimolante per dei parlanti avanzati (Housen et al., 2008).

L'unico task la cui complessità decresce dal 2012 al 2015 è la *negoziazione*. Come abbiamo già detto più volte, in molti casi l'adeguatezza comunicativa viene anteposta alla complessità linguistica, anche dai parlanti nativi (Ortega, 2003; Pallotti, 2009). Nonostante il calo, questo compito registrava già dal primo anno i livelli più alti di diversità lessicale fra tutti i task. A questo proposito bisogna tenere conto del fatto che la *negoziazione* è il compito comunicativo più aderente alla quotidianità, in altre parole si tratta di un evento comunicativo con cui le studentesse hanno occasione di fare pratica tutti i giorni, dunque si potrebbe assumere che, in quanto attività più praticata, abbia consentito lo sfoggio di un repertorio lessicale più ampio rispetto ad esempio alla *telefonata*, dinamica meno presente nella vita di un gruppo di adolescenti non native.

Negli istogrammi alla pagina seguente è possibile osservare l'ordine dei task in base alla complessità lessicale generata nella produzione linguistica delle apprendenti. Dal primo al secondo anno la sequenza dei task cambia minimamente: la *telefonata* e l'*intervista*, task che sottoponevano le apprendenti a una fonte di stress (l'interlocutore sconosciuto e l'assenza del contatto visivo per la telefonata; l'interlocutore sconosciuto e l'espressione della personalità per l'intervista) raggiungono il livello di complessità della *negoziazione*.



Istogramma 3 – Ordine dei task per complessità lessicale nel 2012



Istogramma 4 – Ordine dei task per complessità lessicale nel 2015

Fra le domande di ricerca (paragrafo 2.5), ci si era chiesti se fosse possibile disporre i task in una scala che rispecchiasse la maggiore o minore complessità lessicale richiesta alle apprendenti per portarli a termine. In questo senso, i risultati analizzati nel presente capitolo hanno mostrato la presenza di tendenze molto omogenee tra le diverse apprendenti. Questa uniformità è molto rilevante in quanto ci serve ad appurare che la scala di task emersa dai dati dipende prevalentemente dalle caratteristiche strutturali dei task stessi (*task complexity* e *task conditions*), più che dalle variabili relazionate all'individuo che li svolge (*task difficulty*). Certamente abbiamo incontrato anche casi in cui le caratteristiche personali dell'apprendente hanno evidentemente influenzato l'esito linguistico del task, ma queste circostanze non hanno determinato un mutamento significativo nella tendenza complessiva. Di conseguenza, prendendo la scala di task che emerge dai risultati del 2015: *telefonata* > *negoziazione* > *intervista* > *story retelling* > *map task*, possiamo affermare che la *telefonata* elicitava una maggiore complessità lessicale rispetto allo *story retelling* a causa di caratteristiche specifiche di questo task. Ma quali sono, dunque, le caratteristiche intrinseche ai task responsabili di stimolare la complessità lessicale nella produzione linguistica delle parlanti non native?

Nel nostro studio ci siamo limitati a suddividere i task somministrati in “prevalentemente monologici” e “prevalentemente dialogici”. Sulla base di questa distinzione è possibile osservare quanto segue: i task altamente dialogici come la *negoziazione* e la *telefonata* elicitano sempre i livelli più elevati di complessità lessicale; così come l'*intervista*, che pur essendo un task prevalentemente monologico, presenta uno scambio continuo tra l'apprendente e l'intervistatrice che pone le domande; lo *story retelling*, task monologico per eccellenza, registra livelli bassi di complessità lessicale; lo stesso vale per il *map task*, in cui è la natura stessa del compito a porre dei limiti all'utilizzo di un lessico variato (per indicare un percorso su una mappa occorre ripetere frequentemente i nomi dei punti di riferimento, i verbi di moto, gli avverbi e le preposizioni di luogo). Nonostante questi limiti, il *map task* dal punto di vista dello sviluppo longitudinale mostra una crescita nettamente superiore a quella degli altri task, comparabile solo a quella della *telefonata*.

In sintesi, se escludiamo i casi in cui la struttura del task pone delle restrizioni



all'utilizzo del lessico, la tendenza che si osserva nei nostri dati è che i task prevalentemente dialogici elicitano una maggiore complessità lessicale rispetto ai task prevalentemente monologici. Inoltre, i task prevalentemente dialogici che presentano la variabile dell'assenza di contatto visivo tra gli interlocutori elicitano una maggiore crescita della complessità nel tempo.

Tali conclusioni ci offrono uno spunto di riflessione interessante sulle strategie didattiche volte a favorire l'apprendimento del lessico nell'italiano L2.

A questo proposito, l'ultimo capitolo del presente elaborato contiene alcune considerazioni e proposte indirizzate agli operatori.

## CAPITOLO 5

### *Considerazioni didattiche*

La procedura descritta e messa in pratica nel terzo e nel quarto capitolo del presente elaborato vuole essere un contributo alla discussione sugli approcci metodologici nella misurazione della diversità lessicale dei testi. Oltre a ciò, questo tipo di analisi presenta delle implicazioni pedagogiche: misurare la complessità lessicale della produzione L2 ci dà indicazioni sui diversi livelli di competenza lessicale dei singoli studenti o dei gruppi, permette di discriminare lo sviluppo lessicale nel tempo e di riconoscere quando un certo *plateau* è stato raggiunto e solo con un intervento mirato si può sbloccare il processo di apprendimento del lessico (Daller et al., 2013). L'indice di diversità lessicale qui proposto, applicabile facilmente a qualsiasi testo tramite l'utilizzo di un *tool* informatico, è un valido strumento che può essere utilizzato dai ricercatori nel campo dell'acquisizione di lingue seconde per l'analisi quantitativa dei dati, ma anche dagli stessi insegnanti per calibrare gli interventi didattici.

Inoltre, quando applicata a diversi tipi di task, come è il caso della nostra ricerca, la misurazione della diversità lessicale può fornire importanti indicazioni sulla relazione tra le caratteristiche del task e la performance, che consentono di comprendere quali sono i compiti comunicativi più adeguati all'esercizio o alla valutazione di certe abilità. Le informazioni così ricavate, oltre a guidare l'insegnante nella scelta delle attività più consone al raggiungimento di un dato obiettivo, servono come criterio per graduare i task, ovvero disporli in una sequenza che sia ottimale per l'apprendimento, in modo tale che agli studenti vengano proposte, in base al loro livello di competenza, delle attività che siano sufficientemente impegnative da promuovere l'apprendimento ma non così impegnative da divenire frustranti (Robinson, 2011).

Nel primo capitolo, al paragrafo 1.4, abbiamo parlato degli svariati modi in cui le caratteristiche del task possono influenzare la complessità, l'accuratezza e la fluenza dell'output linguistico. I risultati di alcune ricerche suggeriscono che la diversità

lessicale sia una delle dimensioni più suscettibili alla variabile della complessità del task (De Jong et al., 2012). A questo proposito, il nostro studio di caso ha evidenziato una relazione tra l'elemento  $\pm$ dialogico riferito ai task e la diversità lessicale calcolata nell'output delle apprendenti. Le caratteristiche interattive del task sembrano elicitarne la produzione di testi con una maggiore complessità lessicale. Questo risultato trova un'eco in altri studi empirici, ad esempio quello condotto da Michel (2011), che più generalmente registra nei dialoghi prodotti da non-nativi un linguaggio più fluente, accurato e lessicalmente variato rispetto a quello prodotto nei monologhi.

Si potrebbe dunque concludere che le attività che comportano un'interazione siano le più indicate per esercitare o testare la complessità lessicale nell'insegnamento dell'italiano L2, ma per capire meglio questa relazione occorre prima di tutto chiedersi cosa si intende per interazione e quali sono nello specifico le caratteristiche che fanno di un task un'attività interattiva.

## **5.1 Il ruolo dell'interazione nello sviluppo dell'interlingua**

Sono molti gli studiosi a sostenere che il dialogo sia un luogo privilegiato per l'acquisizione e lo sviluppo delle strutture interlinguistiche.

Hatch dichiara che: “si apprende a conversare, a interagire verbalmente, e da queste interazioni si sviluppano delle strutture sintattiche” (Hatch, 1978:404 in Pallotti, 1998). Questa visione si contrappone nettamente alla concezione “grammatico-centrica” che fino agli anni '70 ha dominato la scena didattica, secondo la quale prima si apprendono le regole grammaticali e le parole di una lingua, e solo in seguito le si “applica” alle conversazioni. Il ribaltamento di questa prospettiva ha visto la nascita dell'Approccio Comunicativo che, nelle sue svariate versioni, fa perno proprio sull'abilità di interazione orale.

Questa abilità è considerata anche la più complessa, in quanto integra in tempo reale le abilità di comprensione e di produzione orale, il tutto all'interno di un evento comunicativo con regole sociali, culturali e pragmatiche. Per dialogare bisogna prima di tutto saper definire il proprio ruolo all'interno della situazione sociale in cui si svolge la

conversazione ed adattare ad esso il registro linguistico; saper organizzare il discorso con una certa competenza strategica che ci permetta di raggiungere il fine pragmatico che ci siamo dati; saper interpretare le intenzioni dell'interlocutore e infine saper negoziare i significati quando questi non sono chiari (Balboni, 2015). Per fare ciò risulta molto utile conoscere dei “copioni situazionali”, cioè delle sequenze prevedibili di atti e mosse comunicative utilizzate solitamente in un determinato contesto (ad esempio durante una telefonata ad uno sconosciuto per richiedere informazioni). Inoltre, secondo i principi del già citato *Lexical Approach* (paragrafo 2.1), il lessico può essere il fulcro dello sviluppo di questa competenza, poiché l'apprendimento di combinazioni di parole molto frequenti in una data lingua è in grado di rendere la comunicazione orale più efficace e fluente.

Se l'interazione è un'abilità così complessa, come può favorire l'acquisizione di un'altra lingua? Chomsky ha ipotizzato l'esistenza di un “dispositivo di acquisizione linguistica” (il *Language Acquisition Device*), con cui si intende quell'insieme di meccanismi psico-biologici di cui ogni essere umano è dotato, che permettono di acquisire sia la lingua materna che le lingue seconde. Tuttavia, possedere questo “organo mentale” non è di per sé sufficiente ad imparare una lingua, poiché esso non si mette in moto se non in presenza della possibilità di interagire con altri esseri umani. Lo psicologo Jerome Bruner parla quindi di un “sistema di supporto per l'acquisizione linguistica” (*Language Acquisition Support System*) riferendosi a quell'insieme di interazioni grazie alle quali i bambini acquisiscono la loro lingua materna. Lo stesso fenomeno può essere esteso all'apprendimento della seconda lingua, per cui la possibilità di interagire con gli altri può giocare un ruolo fondamentale (Balboni, 2015).

Da alcuni studi risulta che l'input linguistico ottenuto tramite la partecipazione a scambi conversazionali facilita la comprensione dei significati, sia da parte di chi ha partecipato attivamente all'interazione (specialmente nelle prime fasi dell'apprendimento) sia di chi ne è stato solo uno spettatore (Pallotti, 1998). Secondo gli approcci comunicativi, la comprensione ha un ruolo fondamentale nell'acquisizione, perché è proprio concentrando l'attenzione sul significato dell'input (e non sulla sua forma) che l'apprendente acquisisce nuove strutture linguistiche<sup>28</sup>. Se è vero che

---

28 Si vedano a proposito le teorie di Krashen sull'importanza dell'input comprensibile (Balboni, 2015).

l'interazione favorisce la comprensione e che un input comprensibile favorisce l'acquisizione, allora l'utilizzo di un input fornito in modalità interattiva è certamente più efficace ai fini dell'acquisizione di uno fornito in modalità unilaterale.

Gli studi più recenti dimostrano, tuttavia, che l'input comprensibile è sì necessario ma non sufficiente ad apprendere correttamente tutte le regole di cui è composta una lingua. Per questo motivo gli approcci moderni suggeriscono di affiancare all'input comprensibile delle attività che dirigano l'attenzione degli apprendenti sulla forma linguistica (*form-focused activities*). È qui che l'interazione si dimostra nuovamente utile, poiché le attività dialogiche offrono la possibilità agli apprendenti di mettere alla prova le proprie ipotesi riguardo alle forme della lingua target, ricevendo subito un feedback su di esse.

Nella conversazione tra parlanti nativi il significato delle parole è un dato che non viene messo in discussione poiché generalmente si condividono le stesse conoscenze linguistiche; al contrario, il dialogo in presenza di parlanti non nativi è in gran parte composto da “negoziazioni del significato”, ovvero momenti in cui il significato delle parole e delle frasi deve essere discusso e stabilito volta per volta affinché il discorso possa procedere. La negoziazione avviene tipicamente tramite le richieste di chiarimento (*clarification requests*: “cosa hai detto?”) e le verifiche della comprensione (*comprehension checks*: “vai a destra! // a destra? // sì”), la formulazione di ipotesi e la loro conferma (Pallotti, 1998). Lo studio della negoziazione del significato ha conosciuto un rapida espansione, in particolare a proposito del suo contributo allo sviluppo linguistico della L2. Una delle motivazioni più diffuse in letteratura per l'impiego di task comunicativi nell'insegnamento delle lingue si basa proprio sull'ipotesi che l'acquisizione della seconda lingua venga stimolata dalla negoziazione del significato. In effetti, la definizione classica di “task comunicativo” enfatizza l'importanza del *focus on meaning*, cioè della preminenza dell'attenzione al significato sull'attenzione alla forma linguistica. In una visione meno categorica, però, il task può comunque essere considerato comunicativo anche nel caso in cui l'attenzione degli apprendenti si rivolga alla forma; spesso, infatti, sono proprio le attività che contengono un dialogo collaborativo a favorire il *focus on form*, poiché è nello sforzo di negoziare il significato che gli apprendenti vengono spinti a cercare attivamente di

formulare delle versioni grammaticalmente più corrette dei loro enunciati (Swain & Lapkin, 2001). Nei sistemi interlinguistici spesso coesistono più forme per esprimere lo stesso significato: può accadere quindi che l'apprendente selezioni la forma che gli sembra più adatta, ma notando che l'interlocutore manifesta dei problemi di comprensione, sposti la sua decisione su una forma alternativa, che nove volte su dieci si dimostra essere una forma più vicina alle regole della lingua target (Pallotti, 1998). Questo meccanismo di auto-correzione si attiva nell'apprendente sia quando l'interlocutore è un parlante nativo che quando è un non-nativo, anche se con alcune differenze. Quando ha problemi di incomprensione, l'interlocutore non-nativo si limita a segnalarlo e attende una riformulazione più chiara, mentre l'interlocutore nativo cerca di fornire delle possibili soluzioni. Questo fa della conversazione tra non-nativi un'interazione più povera e per questa ragione anche più utile alla focalizzazione sull'output, perché spinge gli apprendenti a compiere uno sforzo autonomo per modificare la propria produzione in direzione della lingua d'arrivo. Infatti, se l'interlocutore nativo offre il vantaggio di un confronto diretto con le forme linguistiche *target-like* e quindi un focus maggiore sull'input, allo stesso tempo fornisce delle soluzioni preconfezionate che l'apprendente deve semplicemente ratificare con un cenno di assenso; inoltre, la presenza di un nativo non sembra influenzare positivamente l'output in quanto alcuni studi dimostrano che gli apprendenti hanno l'abilità di riconoscere cosa è corretto e cosa è scorretto anche in assenza di un parlante nativo e anche quando le forme che essi stessi producono non sono conformi alle norme della lingua target (Pallotti, 1998).

L'interazione, dunque, agevola l'acquisizione della lingua seconda grazie alla sua duplice funzione di rendere l'input più comprensibile e di favorire il *focus on form* durante le operazioni di negoziazione del significato.

Tuttavia, essendo un'attività particolarmente complessa, la presenza dell'interazione all'interno di un task porta ad un aumento del livello di complessità intrinseca del task stesso (*task complexity*). Nonostante non siano ancora giunti ad un accordo soddisfacente, i ricercatori mirano a definire gli effetti della *task complexity* sulla produzione L2, per essere in grado un giorno di prevedere con precisione l'azione di determinati task sullo sviluppo interlinguistico. Gli studi più recenti in questo campo

(raccolti in: Skehan, 2001) offrono alcuni risultati generalizzati, fra i quali ci interessa in particolare il dato che vede l'accuratezza e la complessità della produzione L2 beneficiare entrambe della presenza dell'interazione nel task (Skehan, 2009a).

Se dal nostro studio di caso emerge un andamento simile, non è ancora chiaro quali siano gli aspetti dell'interazione direttamente responsabili per l'aumento della complessità dei task e di conseguenza anche di quella dell'output linguistico.

## 5.2 La difficoltà interazionale dei task

Nella ricerca SLA sulla task performance, l'interazione svolge tipicamente il ruolo di variabile dipendente: ad esempio si osserva come i processi cognitivi che caratterizzano un task abbiano effetti sistematici sugli aspetti interattivi come il numero di sequenze di negoziazione. I pochi studi che hanno assunto l'interazione in qualità di variabile indipendente, invece, l'hanno operazionalizzata in termini dicotomici, cioè come presenza o assenza dell'elemento interattivo nel task (nel nostro studio di caso:  $\pm$  dialogico). Tuttavia, è presumibile che diversi tipi di scambio dialogico – come un'intervista, un map task o una telefonata – diano vita a dinamiche interattive diverse tra loro, che apportano un contributo diverso alla complessità generale del task. Spesso in letteratura la *task complexity* è stata caratterizzata sommariamente, tramite nozioni vaghe (quali  $\pm$  *few participants*,  $\pm$  *few elements*, in cui “*few*” sta ad indicare una quantità indefinita) o giudizi soggettivi (Long, 2015), tuttavia per comprendere meglio l'intricata relazione tra le caratteristiche del compito comunicativo e la performance linguistica è necessario poter misurare con più precisione possibile la complessità intrinseca di un task.

Recentemente Pallotti e Rosi (2015) hanno proposto un metodo per misurare quella che hanno chiamato “*interactional difficulty*”, che offre una scala di complessità empiricamente fondata e più raffinata rispetto alla dicotomia  $\pm$ dialogico.

Da un punto di vista teorico, i due ricercatori tentano di identificare gli aspetti specificamente “interazionali” per differenziare il costrutto da altri tipi di competenza: ad esempio, la capacità di produrre una varietà di atti linguistici (chiedere, rispondere,

proporre, descrivere, spiegare ecc...) o un registro cordiale e sostenuto, sono aspetti che vengono spesso associati alla competenza interazionale, ma riguardano una più generale competenza pragmatica, che può manifestarsi non solo nelle conversazioni ma anche nei monologhi o nei testi scritti. Il numero di partecipanti, pur essendo un aspetto più strettamente legato all'interazione, ha un effetto contraddittorio sulla complessità del task, in quanto un maggior numero di interlocutori richiede l'abilità di gestire conversazioni parallele, ma implica anche una minore attività dei singoli nel dialogo, per questo motivo il parametro non è stato incluso nella definizione del costrutto (Pallotti, 2017). Solo tre fattori si possono considerare chiaramente indicativi della difficoltà interazionale di un task e allo stesso tempo sono operativizzabili: il numero di cambi di turno, il numero di turni di iniziativa e la presenza/assenza del contatto visivo tra i partecipanti (Pallotti e Rosi, 2015).

– *Numero di cambi di turno*

Il susseguirsi dei turni in una conversazione richiede ai parlanti coordinazione e tempismo nel prendere parola. La “telegraficità”, cioè lo scambio rapido di turni brevi, è un'abilità che viene acquisita molto lentamente dagli apprendenti di una lingua seconda; infatti la produzione dei parlanti non-nativi si caratterizza per la costruzione di turni lunghi e densi di informazioni, la cui prolissità si può interpretare come una risposta alla mancanza di sicurezza nella capacità di farsi capire (Nuzzo, 2009).

– *Numero di turni di iniziativa*

Nelle prime fasi di acquisizione di una lingua seconda, gli apprendenti tendono a rimanere passivi e intervengono nella conversazione solo se esplicitamente sollecitati da un interlocutore più competente. Lo sviluppo della competenza interazionale implica quindi l'assunzione da parte dell'apprendente di un ruolo più attivo nel far progredire il dialogo, tramite l'uso di mosse conversazionali che creano l'aspettativa di una continuazione (come ad esempio le domande, le proposte, gli inviti, i saluti e i ringraziamenti).

– *Contatto visivo*

La comunicazione multimodale, fluendo attraverso diversi canali, garantisce un alto grado di ridondanza e permette di compensare le eventuali mancanze



linguistiche tramite la comunicazione non-verbale. I parlanti non-nativi, infatti, capiscono meglio i messaggi quando sono accompagnati da gesti ed espressioni facciali, inoltre loro stessi si servono di questi strumenti nella produzione orale per colmare le proprie lacune lessicali. Lo sguardo, poi viene utilizzato da nativi e non-nativi come segnale di gestione dell'alternanza dei turni di parola. Per queste ragioni l'assenza del contatto visivo viene considerata un fattore che accresce la difficoltà interazionale del task.

Una volta individuati i comportamenti interazionali ritenuti più difficili perché acquisiti più tardi dagli apprendenti di una lingua, occorre valutare quali tipi di task richiedono questi comportamenti per essere svolti e in quale misura. Per farlo, Pallotti e Rosi (2015) hanno scelto di esaminare in primo luogo la performance dei parlanti nativi in diverse tipologie di task, poiché osservando solo la performance dei non-nativi sarebbe stato impossibile stabilire se l'assenza di determinati comportamenti fosse dovuta al fatto che non erano richiesti dal task o alle limitazioni linguistiche degli apprendenti. Senza dimenticare che i parlanti nativi non formano una comunità perfettamente omogenea, questo provvedimento ha permesso di affermare con più sicurezza che l'uso di determinate strutture osservato nei task dipende dal grado di difficoltà interazionale dei task stessi, più che dalle capacità dei partecipanti (Pallotti, 2017).

Le partecipanti e i task coinvolti nell'indagine sono stati selezionati (anche se in numero maggiore) dallo stesso corpus VIP utilizzato per il nostro studio di caso; questo, oltre a dimostrare la versatilità del corpus che si presta a innumerevoli analisi, ci permette di formulare ulteriori considerazioni sulla nostra ricerca incrociando i risultati.

Dopo una serie di misurazioni corrispondenti ai tre parametri sopra descritti, viene attribuito ad ogni task un valore indicante la sua difficoltà interazionale. La scala dei task che ne emerge risulta essere la stessa sia per i parlanti nativi che per i non-nativi: *telefonata* > *negoiazione* > *intervista* > *film retelling*.

Il *film retelling* è un compito sostanzialmente simile allo *story retelling*, dunque essenzialmente monologico, perciò è risultato in assoluto il task meno complesso dal punto di vista dell'interazione. L'*intervista* è caratterizzata da cambi di turno abbastanza

frequenti ma l'intervistato non prende quasi mai l'iniziativa. La *negoziazione* e la *telefonata* hanno riportato un livello di difficoltà interazionale simile e maggiore rispetto agli altri task, anche se per motivi diversi: nella prima sono molto frequenti i cambi di turno, mentre nella seconda i partecipanti svolgono un ruolo molto attivo nella conversazione (conducono i tempi dell'interazione, chiedono consigli o informazioni, ringraziano e congedano) a cui si aggiunge l'assenza di contatto visivo con l'interlocutore (Pallotti e Rosi, 2015).

Il fatto che questa scala di difficoltà interazionale dei task rispecchi nel dettaglio la scala di complessità lessicale emersa dal nostro studio (istogramma 4) è un ulteriore argomento che avvalorata la nostra tesi sulla correlazione esistente tra l'aspetto interattivo del task e la diversità lessicale generata nell'output linguistico.

Il costrutto della *interactional difficulty* allarga la nostra prospettiva e ci permette di individuare nello specifico alcune fra le caratteristiche dei task che potrebbero essere responsabili per l'aumento della complessità lessicale.

Sebbene le limitazioni intrinseche a questi studi non consentano di fare inferenze valide per tutti gli apprendenti di una lingua seconda, abbiamo tentato di offrire uno spunto di riflessione su alcune tematiche che sono di particolare rilevanza per la didattica, in quanto gli insegnanti necessitano di conoscere più a fondo il tipo e il livello di complessità dei diversi task e gli effetti prevedibili sulla produzione linguistica degli studenti, in una prospettiva di *testing* o di *task-based language teaching*.

### **5.3 L'uso dei task nell'insegnamento delle lingue**

La ricerca SLA è da sempre interessata a comprendere il modo in cui i parlanti di una lingua seconda adattano il loro parlato quando si trovano di fronte a diversi problemi di comunicazione.

Stando allo studio fin qui condotto, sembra che un aumento nella complessità interazionale del task inneschi dei meccanismi che portano all'aumento della diversità lessicale nell'output degli apprendenti. Nonostante gli studi a questo riguardo siano ancora scarsi (Michel, 2011), le considerazioni emerse sono in linea con la *Cognition*

*Hypothesis* di Robinson (2011). L'autore sostiene che l'aumento del carico cognitivo di un task stimoli l'apprendente a modificare il suo output per andare incontro a queste necessità, causando una rianalisi e una ristrutturazione del proprio sistema interlinguistico. La causa diretta di queste modificazioni viene attribuita a due fattori principali: le *communicative consequences* e i *functional requirements*. In primo luogo, un task più complesso genera più problemi di comunicazione, che i partecipanti tentano di risolvere tramite un maggior numero di *clarification requests* e *comprehension checks*; questi comportamenti favoriscono sia un feedback linguistico esterno da parte dell'interlocutore, che un feedback interno all'apprendente, dato dall'auto-monitoraggio e dalla continua comparazione dei propri enunciati con i modelli dell'input. In secondo luogo, alcuni studi hanno dimostrato che la complessità strutturale tende ad essere accompagnata dalla complessità funzionale del discorso, dunque è plausibile aspettarsi che una maggiore complessità della produzione orale si sviluppi in risposta a queste esigenze funzionali che l'apprendente si trova a dover fronteggiare nella comunicazione (Robinson, 2007c).

Un approccio meno cognitivo e più socioculturale, invece, avanza l'ipotesi che i task possano essere interpretati diversamente da chi li svolge in base ai bisogni e agli interessi personali; l'apprendente sarebbe quindi in grado di modellare il task per i propri fini, costruendo in modo collaborativo dei significati che sono personali e imprevedibili. Chiaramente questo punto di vista mette in dubbio la possibilità di analizzare le qualità del task in modo statico e utilizzarle per prevedere la performance degli apprendenti (Skehan, 2003).

Quello su cui sembrano concordare i sostenitori di un insegnamento *task-based* (TBLT), primo fra tutti Michael Long (2015) e il suo approccio psicolinguistico, è l'efficacia dell'interazione nell'agire in modo personalizzato sull'apprendimento: come sappiamo le incomprensioni che sorgono durante una conversazione sollevano il bisogno di negoziare i significati, a quel punto il feedback interviene apportando l'informazione utile esattamente nell'area del sistema linguistico in cui gli apprendenti hanno dimostrato di avere delle difficoltà. Per questa ragione i task vengono implementati nell'insegnamento come uno strumento tramite il quale l'apprendente può portare avanti la propria agenda, ricevendo da ogni interazione il *focus on form* più

rilevante per ristrutturare la propria interlingua (Skehan, 2003).

In quest'ottica, progettare un task con l'obiettivo di predisporre l'uso di determinate strutture linguistiche che si vogliono insegnare in un dato momento (ad esempio scegliere un task perché offre la possibilità di fare pratica con le frasi negative) significherebbe fraintendere lo scopo originale dei task nell'insegnamento e il loro approccio olistico all'uso della lingua. Secondo Long (2015) è quello che accade in molti libri di testo commercializzati dagli ultimi anni '90, promotori di un TBLT “non genuino”, in cui dei task contraffatti, o come li definirebbe Skehan “*structure-trapping tasks*”, vengono utilizzati per praticare strutture e funzioni in un syllabus tradizionale incentrato sulla grammatica e il *focus on form*.

Ma ci sono anche ricercatori, come Ellis (2003), che sostengono l'utilità dei “*focused tasks*”, in virtù del fatto che la loro chiarezza di struttura favorisce l'apprendimento per induzione. In base al loro focus sulle forme linguistiche, i task sono classificabili in “naturali”, “efficienti” o “necessari”: i primi consentono ma non predispongono l'uso di una particolare struttura linguistica; i secondi sono svolti più efficacemente se quella data struttura viene utilizzata; gli ultimi forzano l'uso di tale struttura (Skehan, 2003).

Qualsiasi sia la tipologia di task prescelta per l'insegnamento, è importante che le attività vengano proposte agli apprendenti in un ordine crescente di complessità. Secondo Robinson (2001) incrementare gradualmente il carico concettuale e comunicativo di un task determina l'impiego di maggiori risorse cognitive (principalmente l'attenzione e l'incorporazione dell'input nella memoria di lavoro) che innescano dei meccanismi di apprendimento, i quali progressivamente portano a una più profonda analisi, modificazione e riorganizzazione dell'interlingua, con conseguenti effetti sulla performance. Non rispettare le necessità e i limiti cognitivi degli apprendenti può avere conseguenze dannose sull'acquisizione linguistica: se il task proposto è troppo difficile, il rischio è che lo studente concentri tutta l'attenzione sul significato, facendo affidamento per la produzione su strategie (contesto, ellissi, linguaggio lessicalizzato) che riducono il valore pedagogico dell'attività; se, al contrario, il task è troppo facile e non presenta una sfida, l'apprendente può annoiarsi e non prendere sul serio le richieste, con un risultato nullo in termini di evoluzione

interlinguistica (Skehan, 1996). Le ragioni per adottare un syllabus costruito su una sequenza graduale di complessità dei task non sono diverse da quelle che sottendono le scelte educative in altri ambiti, come la matematica, che vede lo sviluppo di abilità complesse attraverso la pratica su compiti inizialmente semplici e progressivamente più impegnativi. In questo modo, poi, risulta più equilibrato il carico di lavoro percepito dagli apprendenti, con conseguenze positive sulle variabili affettive come ansia e motivazione (Robinson, 2007c).

In passato, il criterio utilizzato per graduare gli esercizi si basava essenzialmente sul contenuto linguistico, perciò un'attività era considerata più o meno complessa a seconda degli elementi linguistici in essa coinvolti. Secondo l'approccio *task-based* all'insegnamento linguistico, quello che gli apprendenti devono fare con la lingua è altrettanto importante del contenuto stesso della lingua. Per essere disposti in una sequenza come quella descritta sopra, i task pedagogici devono quindi essere analizzati nei loro elementi strutturali costitutivi.

Secondo Robinson è sufficiente modificare il carico cognitivo tramite le variabili che egli chiama *resource-directing* (vedi tabella 1 al paragrafo 1.4.3) per ottenere versioni più o meno complesse di uno stesso task: ad esempio produrre una narrazione su eventi che si svolgono nel passato e in un altro luogo (*– here and now*), quindi senza il supporto del contesto e con il ricorso alla memoria, richiede un maggior impiego di risorse cognitive oltre ad un migliore controllo sulla morfologia (tempi passati) e sulla sintassi (avverbi di tempo e luogo e connettori) rispetto alla descrizione di fatti che stanno accadendo ora, nel contesto circostante e davanti ai nostri occhi (Robinson, 2007c).

Un filone di ricerca si occupa invece di determinare gli effetti, sulla complessità del task e sulla produzione linguistica, di alcune attività che possono essere svolte prima o dopo il task. Per esempio, la complessità cognitiva di un task può essere ridotta da alcune *pre-task activities* mirate a richiamare negli apprendenti conoscenze schematiche che potrebbero essere rilevanti per lo svolgimento del compito e che lo rendono cognitivamente più familiare; le attività di pianificazione sia del linguaggio da utilizzare che dei significati da esprimere hanno dimostrato di avere un chiaro impatto sulla produzione linguistica in termini di maggiore complessità (lessicale o sintattica)

accuratezza e fluenza, poiché riducono il carico di elaborazione che gli apprendenti devono sostenere durante il task, consentendo un rilascio maggiore di attenzione verso la forma linguistica. L'ipotesi che sottende le *post-task activities* è che la consapevolezza di quello che verrà dopo possa influenzare il modo in cui gli apprendenti indirizzano la propria attenzione durante il task: se ad esempio gli studenti sanno di dover riportare pubblicamente il contenuto di un task, durante lo svolgimento in privato sono portati a dedicare significativamente più attenzione all'accuratezza (Skehan, 1996).

#### **5.4 Verso una maggiore complessità linguistica**

Lo sviluppo della complessità è uno degli obiettivi centrali dell'apprendimento linguistico poiché riguarda il processo di ristrutturazione continua tramite cui il sistema interlinguistico incorpora nuovi elementi e ne abbandona altri, divenendo più elaborato nella struttura e più coerente con la lingua target. L'aumento della complessità linguistica nella produzione degli apprendenti è dunque il segnale più chiaro dell'avvenuta acquisizione.

Questo procedimento potrebbe incontrare alcuni ostacoli, come la mancanza di interesse dell'apprendente nel diventare più *native-like* o semplicemente nel compiere lo sforzo di cambiare e riorganizzare la propria interlingua, o ancora l'insufficienza dell'input che fornisce le condizioni necessarie perché questa evoluzione avvenga. Inoltre, alcuni apprendenti potrebbero essere restii al cambiamento e preferire al rischio un sistema interlinguistico meno elaborato ma adattato alle proprie esigenze: tale strategia conservativa ha l'effetto di privilegiare l'accuratezza alle spese della complessità.

In sintesi, per agevolare la complessificazione dell'interlingua sono necessari una serie di fattori quali: una generale motivazione al cambiamento, la volontà di avvicinarsi alla performance dei parlanti nativi, un input adeguato, delle opportunità di interazione e infine un supporto nella ristrutturazione, che si può concretizzare tramite la proposta delle attività didattiche in una sequenza di difficoltà crescente e tramite lo svolgimento

di attività di preparazione e di riflessione prima e dopo il task (Skehan, 1996).

Nel nostro studio abbiamo messo l'accento sull'importanza dell'aspetto interattivo dei task nello sviluppo della complessità lessicale. I task pedagogici che comportano un'interazione hanno un effetto che Skehan (1996:48) definisce “stretching influence on interlanguage”, in quanto il focus è sulla realizzazione di un'attività concreta ma la complessità linguistica è indispensabile al suo completamento. Così, non solo la produzione delle apprendenti risulta effettivamente più complessa in questo tipo di task, ma aumentano anche le opportunità di ristrutturazione dell'interlingua, ossia le probabilità di acquisizione linguistica.

Abbiamo visto che ci sono aspetti dell'interazione più rilevanti di altri nel determinare la *difficoltà interazionale* di un task (numero di cambi di turno, numero di turni di iniziativa, assenza del contatto visivo) e che questi fattori sembrano essere relazionati alla produzione di un output più complesso, perciò devono essere tenuti in considerazione nella messa a punto di attività didattiche efficaci, così come nella valutazione di tali attività per la costruzione di una sequenza di difficoltà crescente.

Con lo stesso scopo, è possibile modellare le interazioni in modo da produrre un diverso quantitativo di negoziazione del significato – elemento fondamentale per l'acquisizione di nuove forme – giocando con le variabili strutturali del task a nostra disposizione. Ad esempio, per quanto riguarda le *participation variables* individuate da Robinson (2011:6) (vedi paragrafo 1.4.3), l'elemento *+convergent solution*, ovvero un task in cui i partecipanti devono concordare sulla risposta da dare, richiede una maggiore negoziazione rispetto ad un task divergente, in cui la discussione è completamente libera e non deve raggiungere un accordo (Long, 2015). Tra le *participant variables*, invece, abbiamo visto come il fattore *+same gender* potesse rendere più sicure le adolescenti nel task della telefonata ad uno sconosciuto, che quindi implicava già la difficoltà *-familiar*. Inoltre, alcuni studi riguardo alla variabile *±equal status and role* hanno mostrato che l'interazione tra gruppi o coppie di apprendenti garantisce uno scambio più animato e mosse conversazionali più varie rispetto all'interazione tra l'apprendente e l'esaminatore che appare invece più controllata (Pallotti, 2017). Se l'obiettivo è quello della complessità linguistica, dunque, è preferibile instaurare un dialogo tra due apprendenti, tuttavia non è sempre facile creare

delle coppie di lavoro adeguate: l'asimmetria di competenza linguistica fra gli interlocutori (*–same proficiency*), in certi casi può attenuare i benefici dell'interazione e avere degli effetti negativi sulla performance degli apprendenti (Skehan, 2003). Ne abbiamo avuto prova durante la nostra ricerca, nel caso della coppia Randa-Katerina: il livello di padronanza dell'italiano della prima studentessa si collocava sotto la media del resto del gruppo e questo ha influenzato l'interazione con la compagna nel task di coppia *organizzare una gita*; le due apprendenti si bloccavano a vicenda, Katerina dominava la negoziazione surclassando Randa, che a sua volta offriva un feedback minimale non sufficiente a sorreggere il dialogo.

In conclusione, tramite la scelta del tipo di task<sup>29</sup> e la sua manipolazione in base alle esigenze del target, l'insegnante può prendere una serie di provvedimenti per indirizzare la produzione degli apprendenti verso una maggiore complessità linguistica. Durante lo svolgimento del task, invece, il ruolo dell'insegnante deve essere defilato per non compromettere la naturalezza dell'interazione, il cui pregio è quello di andare ad agire proprio sui punti critici del sistema interlinguistico. Il docente deve quindi cogliere quando il suo intervento può essere decisivo per consentire la proficua prosecuzione dell'attività e quando invece le richieste di aiuto vengono usate dagli studenti per appianare qualsiasi difficoltà, rendendo inefficace il processo di negoziazione del significato. Inoltre, gli apprendenti possono portare il dialogo in qualsiasi direzione loro scelgano e l'insegnante deve essere pronto a fornire un aiuto totalmente imprevedibile: questo implica una preparazione molto più vasta rispetto al confortante ordine del giorno prestabilito. Probabilmente, la ragione per cui il *Task-based language teaching* non è più diffusamente impiegato nella didattica delle lingue è da ricercare proprio nel ruolo difficile e delicato che comporta per l'insegnante (Skehan, 2003).

---

29 Oltre ai task presentati in questa ricerca, per un elenco di attività interattive da utilizzare nell'insegnamento dell'italiano L2/LS si consulti Balboni (2015:182-186) o Frencich e Torresan (2005:69-81).



## Conclusioni

Nel presente lavoro si è cercato di investigare la diversità lessicale di un corpus di dati tramite l'uso dello *Standardized Type/Token Ratio*. Tra le tante proposte di indicizzazione della complessità lessicale, questo metodo ha dimostrato il vantaggio di coniugare diverse esigenze: in primo luogo la lemmatizzazione preventiva del testo, che garantisce l'affidabilità della misura anche nel caso di una lingua morfologicamente ricca come l'italiano e rende più accurata la comparazione di questi studi con quelli condotti su lingue tipologicamente distanti come l'inglese; in secondo luogo, l'utilizzo di un sistema di calcolo basato sui campionamenti casuali, che risolve il problema della sensibilità alla lunghezza del testo, permettendo la comparazione fra testi di lunghezza diversa; e infine la fruibilità della misura tramite una piattaforma informatica facilmente consultabile.

L'analisi dei dati, condotta con l'aiuto di questo strumento su un totale di 60 testi, corrispondenti a circa 19 ore di registrazioni audio, ha rilevato una crescita generalizzata della complessità lessicale nella produzione linguistica delle apprendenti dal 2012 al 2015.

Le differenze interindividuali si assottigliano nel corso degli anni, perciò le apprendenti affrontano una crescita diversa a seconda del loro punto di partenza ma tendono a raggiungere un livello simile, forse a causa del *ceiling effect*, fenomeno abbastanza comune nelle fasi avanzate di apprendimento, o semplicemente per il progressivo uniformarsi della loro performance a quella del gruppo di controllo, costituito da due studentesse parlanti native di italiano. Si distacca dalle altre la produzione linguistica della studentessa Randa, in cui i fattori affettivi si sono aggiunti alle difficoltà derivate dalla distanza tipologica della L1 (Tamil) rispetto alla lingua target.

Le differenze strutturali tra i compiti comunicativi proposti nella sperimentazione, e in particolare l'elemento  $\pm$ *dialogico*, sembrano influenzare, in modo coerente fra i diversi soggetti, la complessità lessicale utilizzata per portare a termine i

task. Ad una maggiore “difficoltà interazionale” del task è associata una maggiore complessità linguistica dell'output, ad eccezione del *map task* che per sua stessa natura pone dei forti limiti all'utilizzo di un lessico variato.

È opportuno ricordare che le misurazioni messe in pratica nel presente studio descrivono una sola delle molteplici dimensioni che compongono la complessità linguistica nei termini teorici in cui l'abbiamo presentata. Inoltre, i limiti del corpus analizzato non consentono inferenze valide per tutti gli adolescenti di sesso femminile apprendenti di italiano L2, né tantomeno per una più vasta comunità di parlanti di diversa età, sesso e background.

Tuttavia, i risultati emersi da questo studio hanno fornito uno spunto di riflessione sul ruolo dell'interazione nel processo di continua ristrutturazione al quale sono sottoposte le interlingue, che implica la loro graduale complessificazione. Durante il dialogo, gli apprendenti vengono necessariamente coinvolti in una negoziazione del significato, che ha il potere di agevolare la comprensione dell'input linguistico e favorire il *focus on form*, entrambi passaggi essenziali per l'acquisizione di una lingua seconda. Questo fa del task dialogico un eccellente strumento didattico, in cui il contenuto linguistico non deve essere prestabilito poiché l'interazione va ad agire naturalmente sulle zone critiche del sistema interlinguistico, mentre le variabili dell'interazione possono essere modellate dall'insegnante con lo scopo di stimolare lo sviluppo della complessità linguistica.

## Appendice

In questa sezione presentiamo le convenzioni utilizzate per la trascrizione dei dati, insieme ad un breve esempio di trascrizione (frammento iniziale) per ciascuno dei task considerati nell'indagine.

### Convenzioni di trascrizione

parole pronunciate ridendo	£parola£
risata vocalizzata	hehehe, hahaha, hihih
risata non vocalizzata	£hhh£
innalzamento o abbassamento improvviso dell'intonazione	/ ascendente, \ discendente
intonazione finale discendente	.
intonazione finale sospesa	,
intonazione finale ascendente	?
inspirazione (lunghezza proporzionale al numero di acca)	.hhh
espirazione (lunghezza proporzionale al numero di acca)	hhh.
basso volume	°parola°
alto volume	PAROLA
prolungamento di suono	paro:::la
inudibile (una x per fonema)	xxxxxx
trascrizione dubbia	(parola)
parlato veloce	f*parola*
parlato lento	s*parola*s
troncamento con colpo di glottide	paro-
pause di meno di 0.2 sec	#
pause più lunghe (il trattino basso separa secondi e decimi)	#0_5, #0_6, #3_5, #1, #2
enfasi su parola o frase	!parola! !tutta la frase!
Inizio di sovrapposizione	[
Fine sovrapposizione	]

## INTERVISTA

@Begin  
@Languages: it  
@Participant:S09 Student 09 Lali, INV Investigator Giovanna, CAM Camera  
Operator Fabiana  
@Task: intervista  
@Date: 2015  
@Languages of S09: L1 punjabi, L2 italian

INV: va bene? # sei pronta?  
S09: sì sì fhhh dai che ce la facciamo  
INV: perfetto # allora guarda ti chiedo di dirmi il tuo nome/ # un po' così quanti anni hai da dove vieni che scuola fai  
S09: allora io mi chiamo Lali e ho diciannove anni # vengo dall'India però sono in Italia da:: otto anni e:: poi: cioè insomma prima abitavo a Roma poi: ci ho vissuto un anno e poi ci siamo trasferiti qua: a Reggio Emilia e sono qua # cioè frequento lo Scaruffi fe: mi faccia altre domande  
INV: fcome no? # senti # la tua lingua che parlavi quando eri in India nel tuo paese qual'è?  
S09: eh punjabi parlo là # sì una: regione del nord dell'India Punjab che si parla punjabi là  
INV: ho capito # e conosci anche altre lingue oltre al punjabi e all'italiano?  
S09: sì:: certo conosco: beh parlo la mia lingua poi altre lingue dell'India che sa no? sono tantissime lingue in India # ogni regione ha la sua: lingua cultura tutto diverso # e poi parlo italiano inglese tedesco francese e un pochino l'avevamo studiato alle medie però adesso: [sinceramente non me lo ricordo  
INV: [francese alle medie # e tedesco invece lo studi alle superiori  
S09: sì sì tedesco lo facciamo da: cinque anni ormai sì questo è il quinto  
INV: e le altre lingue indiane invece che parli le hai imparate a scuola o: tramite:, [parlando con persone  
S09: [sì: a scuola la lingua quella: nazionale è hindi # però nella mia cioè regione si parla la proprio dialetto diciamo e: quella nazionale ovvio cioè: sappiamo [[sì sì si studia # e poi: cioè è un po' anche simile e poi anche per la televisione comunque # cioè è molto conosciuto diciamo  
INV: [[a scuola studiavi, parlavate in hindi # mh mh  
INV: certo # ho capito  
S09: e: poi italiano qua: ormai otto anni fcioè # sì: ce l'ho fatta anchef # e poi a scuola invece: inglese e tedesco e poi francese alle medie che # °l'avevo imparato là°  
INV: senti ricordi:, adesso sono passati otto anni ma ricordi un po' il tuo viaggio? il viaggio che hai fatto per venire qui la prima volta?  
S09: sì: cioè non nei particolari però in generale ovvio non è indimenticabile direi # ehm sì siamo: mh all'inizio non sapevo che dovevo venire però poi visto che c'era mio papà qua e io mio fratello e la madre eravamo: ancora in India # lui è venuto prima e poi ha fatto: preparato i documenti tutto e ci ha chiamato qua # allora mia madre mi disse che dovevamo: comunque cioè lasciare la scuola e partire per poi andare dal papà # diciamo che c'era sempre quella parte che diceva che bello no? andiamo dal papà invece l'altra che dovevo lasciare gli zii: e gli amici: i nonni: era un po' # triste diciamo quella parte # però: ero anche abbastanza contenta perché: comunque una nuova realtà: cioè otto anni fa India comunque non era proprio quella di oggi era un po' diversa # e: tutto sommato: era bello dai hihhi  
INV: c'era una parte anche bella  
S09: sì sì cioè quel nuovo un po': cioè mondo nuovo nuove: realtà conoscenze: era un po' emozionante da quello che ricordo # e: infatti poi: quando siamo venuti qua in Italia # all'inizio era un po' # cioè triste diciamo perché comunque non conoscevo nessuno qua e mi mancavano: i miei amici i nonni e

tutto # perché là comunque abitavamo insieme ai nonni/ # era un po': diciamo: difficile all'inizio però poi: cominciato andare a scuola e poi anche: dei vicini che erano molto gentili: cioè c'era una signora che ci trattava come i suoi nipoti quindi bellissimo # e: pian piano: ho cominciato a abituarci e ora mi trovo molto bene diciamo

INV: ho capito # e dei tuoi amici, mi dicevi prima che ti mancavano ti ricordi quando gli hai detto che saresti partita # com'è stato?

S09: sì sì c'era: mh # tutto tutto non mi ricordo però comunque c'erano degli amici coi:: magari ero più legata e avevo detto a loro sì che partivo # comunque piangevamo tutti hihhihi fsiif e io avevo: undici dodici anni là quindi: non è che: sapevo benissimo di cosa mi sarei trovata oppure che: per un tot di tempo non potevo più tornare # però poi: cioè insomma mamma che diceva no dai torniamo presto vedrai che sarà bello f'c'è papà che aspetta da moltof e: # cioè è stato # penso che anche loro sarà stato difficile infatti quando dopo li sentivo per telefono dicevano comunque che mancavi: che quando torni e:: insomma solito # alla partenza c'è sempre quella gioia di vedere nuove cose e quella invece la parte triste di lasciare: °il tuo paese°

INV: certo # sei tornata poi al tuo paese da quando sei qui?

S09: sì sì sono tornata:: tre volte # prima volta dopo: tre anni e mezzo tipo quattro quasi # la seconda invece: dopo un anno e poi: l'inverno scorso # sì

INV: e com'è quando torni che effetto ti fa trovare:, [ritrovare, troppo bello? Ah sì?

S09: [ftroppo bello sì bellissimo sì:f perché comunque c'ho dei legami ancora molto forti con i nonni che ti accolgono a braccia aperte come si dice e è bello trovarsi della gente: che: si vede sulla faccia la gioia di vedere una persona e: è bello veramente

INV: ti piacerebbe tornare a viverci in India?

S09: adess fhfhf cioè mi piacerebbe molto tornare tipo in vacanza diciamo per un mese due mesi # così sì però # vivere: non sarei molto sicura perché .hh cioè comunque ho vissuto tutto il tempo qua diciamo: la maggior parte della vita # là ero comunque bambina non è che, invece qua: ho # cioè ho accumulato delle diverse conoscenze comunque u:n un pensiero diverso da quello del tipico indiano: lì e sarebbe un po' difficile per me vivere là anche il modo di fare le cose di vivere abitudini # non cioè mh penso che sia difficile adesso mi sono abituata qua: e comunque sarebbe una cosa: quasi nuova per me andar là a vivere e: # cioè poi: bisogna vedere quando si trova nella situazione poi che si, escono fuori tutte le cose però: # cioè io prevedo una cosa così che mh se non fossi costretta io abiterei benissimo qua cioè andrei sicuramente a trovare ogni anno o due anni no? i parenti però: # cioè viverci no fhfhf preferisco Italia oppure un cioè un paese occidentale qua # o comunque cioè diverso per me India è ancora, diciamo che si sta sviluppando adesso molto cioè tra i paesi in via di sviluppo # però ci sono ancora delle cose che: magari: preferisco di non, cioè fstar lontana diciamof

INV: mh mh ha capito # senti quando sei arrivata in Italia qual'è stata la cosa che ti ha colpito di più com'è stato l'impatto? proprio all'arrivo

S09: ehm mi ricordo che pioveva quando siamo arrivati # e: c'era mio papà all'aeroporto quindi vedere lui già era: un'emozione cioè da tanto che non lo vedevo erano:: due anni penso # però comunque cioè due anni non vedere il papà è:: fun po'f # perciò cioè è stato bello vedere lui e poi anche era di notte e: tutte cioè non ricordo benissimo cosa notai per prima cosa però # cioè era la figura del papà penso che veder lui cioè fhfhf era: veramente bello # poi anche che non avevo mai viaggiato cioè prima di venire qua in Italia no? in aereo # quindi anche quell'effetto là che avevo ancora che ho volato hehe fil senso è quellof mi ricordo un po': # °quello°

INV: potresti dire alcune cose positive e cose negative del tuo primo periodo qua # dei primi mesi in Italia

S09: ah # sicuramente la cosa difficile è stata: cioè abituarsi a vivere un po': # più da soli perché comunque in casa mio papà andava a lavorare quindi eravamo io mio fratello e la mamma la maggior parte del tempo # invece lì in India: ero cresciuta là quindi i nonni zii vicini tutto # quindi quello: cioè per me è stato molto difficile poi anche il fatto di non conoscere la lingua quindi non riuscire: a parlare # però poi i fatti positivi/ invece che ho incontrato questa signora che era: la mamma del pad cioè del capo di mio papà

# quindi cioè la sua: noi la chiamavamo nonna quindi ci trattava:, e fino a ora siamo in contatto cioè lei mi chiama io: la vado a trovare d'estate a Roma # e conoscere lei è stato veramente bello da lì: ho cominciato un po' più a legarmi cioè all'Italia e pian piano poi: # oserei dire dimenticare un po' le cose dell'India haha e poi anche gli amici a scuola che: i primi # i primi mesi è stato un po' difficile non capire:: cioè comunque lì andavo a scuola quindi: diciamo che ero anche bravina fe qua non capire nulla era un po' triste # infatti fpiangevo anche a voltef # però poi le maestre erano molto carine # mi portavano fuori a: # cioè proprio all'inizio magari l'alfabeto parole: così # facili facili e poi dopo: cioè ho cominciato pian piano poi: essendo piccola comunque ho appreso molto presto # la lingua e quindi poi mi sono un po' diciamo un attimo orientata e poi: # poi è stato facile dai # cioè grossi problemi non credo di averli avuti comunque: c'era papà che: ormai da tanto tempo che abitava qua quindi: lui: cioè conosceva poi ci portava in giro così # quando trovava il tempo: cioè era bello anche stare con papà hehehe

INV: siete in quattro in famiglia mi dicevi?

S09: sì in quattro sì sì

INV: e ancora abiti con loro tutti e quattro insieme

S09: abitiamo insieme sì

INV: come sono i tuoi: famigliari # il papà la mamma tuo fratello giusto?

S09: sì fratello più picc [cioè,

INV: [prova: a dirmi un po' che tipo di persone sono

S09: hhh diciamo che mia madre la vedo più come un'amica diciamo siamo molto legate io e mia mamma # invece con papà: un po' meno: cioè essendo anche forse che da piccola cioè ho abitato sempre con la mamma forse anche per quello # però con papà, # cioè ho sempre buoni rapporti con tutti e due # grazie a dio # però: mi sento più legata cioè poi la figura femminile magari cioè penso che sia un po' per tutti di essere legati alla mamma che più a papà # però anche con papà cioè ci parlo tranquillamente:: # diciamo che forse: per il fatto che lui # cioè è molto più attaccato alla tradizione magari # e io: magari mi sento un po' più cioè di qua adesso e: magari a volte cioè voglio fare delle cose che # qua è normale invece lì in India forse un po' no quindi lui un po' cioè dice, magari

## STORY RETELLING

@Begin  
 @Languages: it  
 @Participants: S17 Student 17 Katerina, INV Investigator Giovanna, CAM Camera Operator Fabiana  
 @Task: story retelling  
 @Date: 2012  
 @Languages of S17: L1 Moldovan, L2 Italian

S17: allora #5 prima. # e::: c'era un ragazzino con il suo: cagnolino # che probabilmente ha ricevuto in regalo:: una rana # che l'ha messa dentro a un barattolo # di vetro .hhh allora e::: stanno insieme ad ammirarlo # ammirarla # la rana # e::: è notte # visto che fuori c'è la luna # dopodiché decide di andare a letto # insieme al cane # c-cane °dio santo # sono: stanca° .hhh e::: la rana nel frattempo: scappa da: dentro al barattolo # che: evidentemente non aveva il tappo # e al mattino quando il ragazzo si sveglia:: non la trova #3 e: la cerca::: dappertutto # dentro casa/ però: non riesce a trovarla # allora apre la finestra e::: comincia a chiamarla # e il cane: sale sul davanzale però sta per cadere # effettivamente cade/ e rompe il barattolo che: aveva in testa #3 il ragazzo si: arrabbia leggermente #3 e vanno:: a cercare la rana: in un bosco # vicino a casa #4 il cane e::: abbaia: a un alveare/ # invece il ragazzo: la sta cercando: dentro a u:n # fosso # °no:? boh # più o meno° # e::: infatti si: si avvicina # e secondo me sarebbe un #5 boh # un animale [fhhh

INV: [hehehe

S17: che:: secondo me le morde il naso #1 perché # sì .hhh ehm e invece il:

il cane fa: cadere l'alveare e: le api:: arrabbiate # evidentemente .hh e: cominciano a rincorrerlo # invece nel frattempo il rega il ragazzo ehm è salito su un albero # e sta guardando dentro # da dove sbuca un gufo che lo spaventa e lo fa cadere #3 allora: e: infastidito # sale: su una roccia e si appoggia # secondo lui a dei rami # che erano invece # corna di u:n corvo ((l'animale nell'immagine è un cervo)) # il quale alza la testa e: prende:: carica il ragazzo e comincia: a correre perché si è spaventato #3 arrivato:: a un diruppo # si ferma bruscamente e il ragazzo cade in acqua insieme al cane che stava nel frattempo ehm rincorrendo il # il corv # cervo °scusate° e:: il ragazzo e il cane cadono in acqua # e: dopodiché: lui sente il suono: delle: delle rane #2 e quindi va a guardare #2 infatti vede una famiglia di: di rane con la mamma il papà e i figlioletti #4 e:hm tornando a casa se ne prende uno dei piccolini #2 invece le altre rane lo stanno a guardare  
INV: finisce bene?  
S17: sì fhhf

### MAP TASK

@Begin  
@Languages: it  
@Participants: S18 Student 18 Randa, INV Investigator Giovanna, CAM Camera Operator Fabiana  
@Task: map task  
@Date: 2015  
@Languages of S18: L1 Tamil, L2 Italian

INV: quando vuoi # puoi iniziare a darmi qualche indicazione # e io al limite ti faccio delle domande  
S18: sì # e:: # bisogna partire dal stagno delle #1 libe # lule  
INV: sì  
S18: però sopra # sopra a destra  
INV: sopra / # a d  
S18: a destra # cioè si parte # e poi devi andare dritto # a sinistra # di poco #1 poi scendere un po' giù  
INV: a sinis # vado dritto a sinistra fino a dove? # io ho una barca qui  
S18: e:hm: # no devi # cioè dopo che sei andata un po' giù # devi anda # hai giardino delle visite?  
INV: no  
S18: gelataio?  
INV: sì  
S18: e:h okay # allora devi # devi: # fare # un cerchio grande # attorno al gelataio  
INV: ma passo # in mezzo alla barca?  
S18: io non ho la barca  
INV: ah tu non hai la barca # quindi dallo stagno delle libellule / # tu: # vai dritta dritta fino al gelataio  
S18: !no! # devi andare un po' dritta di poco  
INV: mh mh  
S18: poi scendere giù  
INV: sì  
S18: andare dritto a sinistra  
INV: sì  
S18: poi # andare dritto in su  
INV: mh mh  
S18: e poi # fare tipo # un tipo di cerchio # una curva sopra il gelataio  
INV: m:h  
S18: poi e:h # tipo da destra scendi giù

## NEGOZIAZIONE

@Begin  
@Languages: it  
@Participant:S09 Student 09 Lali, S10 Student 10 Monika, INV Investigator Laura, CAM Camera Operator Fabiana, SH1 Bed and Breakfast Il Sole d'Arechi, SH2 Bed and Breakfast La cittadella, SH3 Bed and Breakfast Sul corso, SH4 Ristorante Quo Vadis, SH5 Pizzeria Pizza&pizza, SH6 Ente provinciale del turismo, EA1 Expert Adult Manuela, EA2 Expert Adult 2 Serena, EA3 Expert Adult Luigi, EA4 Expert Adult Clorinda, EA5 Expert Adult 5 Giuliano, EP1 Expert Peer 1 Annachiara, EP2 Expert Peer 2 Mariasole  
@Task: organizzare una gita  
@Date: 2015  
@Languages of S09: L1 Punjabi, L2 Italian  
@Languages of S10: L1 Ukranian, L2 Italian

S09: allora andiamo a Salerno quindi #1 [[in mezzo  
S10: [[la prima cosa che dobbiamo decidere è  
S09: a che ora partire #1\_2 o no?  
S10: sì dai # mh mh  
S09: °se non parti°,  
INV: sono tre giorni abbiamo detto [eh #0\_5 per ventiquattro ventiquattro [[venticinque\  
S10: [mh mh  
S09: [mh mh  
S10: [[un secondo # scriviamo #1\_2 [ventiquattro, ventiquattro ventisei luglio  
S09: [[al mattino quindi/  
S10: hhhh allora quindi #3 cosa potremmo prendere il treno/ #3 che a Salerno/ #1 non so se  
S09: sì da Reggio Emilia #0\_5 dalla [stazione  
S10: [se conviene #1 okey #2 proviamo quindi #1\_2 e qua però non ci sono # i: i numeri dei delle ferrovie #1\_2 °facciamo che prendiamo la macchina [[hehehe°  
INV: [[fhhh  
S09: [[oh sì ora che ho la patente: # [[andiamo in macchina  
S10: [[esa- oh:: bravissima  
perfetto quindi questo l'abbiamo risolto non ci interessa hehehe  
INV: quindi tu guidi #1 Lali?  
S09: sì sì ho la- la patente [sì #1\_2 finalmente  
INV: [oh # sì  
INV: fatevi # cercate allora di magari chiedere qualche informazione comunque sul !costo! del viaggio [in macchina # perché:: # se conviene prendere l'autostrada: da # oppure no quanto può [[costare  
S10: [mh mh #1 mh che poi è l'autostrada  
S09: [[eh:: c'è il bigliettino da pagare  
INV: mh mh c'è il bigliettino da [pagare fhh  
S09: [beh la benzina fai un pieno # dovrete arrivare #2 [[abbiamo fatto da:  
INV: [[magari se capiamo più o meno quanti chilometri sono #1 riusciamo a capire quanti chilometri fa  
S09: da Reggio a Salerno secondo me saranno,  
S10: boh io non ne ho idea fa dire la verità .hhh hhhh  
S09: più di: #0\_8 no dai non spariamo  
S10: hehehe  
S09: [no va be da # da Roma fino a Reggio sono tipo cinquecento abbiamo fatto in macchina/  
S10: [magari poi chiediamo f\*lo\*f a qualcuno  
S10: quindi, mille? #1 no  
S09: be un po' di [meno sì:  
S10: [settecento dai # [[seicento



INV: [[ se vi fate un'ipotesi potete anche chiedere conferma a qualcu: no [#1\_2 poi decidete vo:i quale  
S09: [eh # secondo me saranno settecento e qualcosa  
INV: domanda fare ai ragazzi giovani e quali invece a una persona maga:ri non so [di quarant'anni trentase:tte # anche un settantenne può avere viaggiato tanto  
S09: [adulta  
S10: mh mh .hhh [allora  
S09: [speriamo che risponda il settantenne hehehe  
S10: heheh  
INV: hehehe # no sono tanti numeri quindi se non risponde uno poi risponde l'altro non vi preoccupate  
S09: allora Moni prendiam- prima chiamiamo: uno dei [ristoranti o:: # sì gli alberghi  
S10: [gli alberghi  
S10: secondo me uhm probabilmente proviamo con gli alberghi  
S09: mh.  
S10: così vediam:mo:: il miglior costo  
S09: mh mh  
S10: il miglior prezzo  
S09: chiamiamo due uno lo fai tu uno lo faccio io #2\_3 f\*go baby go\*f  
S10: fguarda #0\_8 avanti le più piccole  
S09: no no # vai dai dai dai su  
S10: bah hhhh  
S09: a chi vuoi chiamare però # così lo segno  
S10: non te lo dico #1 fè un segreto # così imparif

### TELEFONATA

%com 29:50 S10 telefona a SH5 Pizzeria Pizza&pizza, telefonata 5

SH5: pronto?  
S10: eh:: buongiorno #1 le volevo chiedere delle informazioni su: eh:m sugli orari che fate #3 e: siete aperti tutti i giorni o- op- mi- e:  
SH5: e: siamo aperti tutti i giorni  
S10: mh mh # [tutti i giorni sia a pranzo che a cena?  
SH5: eh sì  
S10: ah ah e: e poi le volevo chiedere # dato che ho trovato il numero nella rubrica non so dove vi trovate,  
SH5: eh?  
S10: e: non so dove vi trovate di preciso,  
SH5: e: s- sul lungoma:re:: Marconi signo'  
S10: ah ah # quindi dato che non sono del posto è in centro?  
SH5: eh: sì # è vicino al bar Ti amo  
S10: ah ah okey # e poi un'ultima informazio:ne e: servite solo la pizza # giusto/  
SH5: no pure primi pia:tti [secondi pia:tti #0\_5 contorni,  
S10: [ah  
S10: ah ah quindi anche qualcosa di tipico.  
SH5: eh sì pure fritta # eh tipico eh sì  
S10: [ah ah ehm: fun'ultima domanda e poi la lasciof # eh: quanto viene a costare bene o male/ #0\_5 i prezzi come sono.  
S09: [°è tipica°  
SH5: eh::: #1 venite che ve le faccio spiegare il prezzo  
S10: fah okey grazief hehehe  
SH5: fhhh  
S10: grazie mille # è stato gentilissimo arrivederci

%com 31:11 fine telefonata 5



## Bibliografia

- Baerman, M., Brown, D., Corbett, G. G., 2015, *Understanding and measuring morphological complexity*, Oxford University Press, New York.
- Berman, R. A., & Verhoeven, L., 2002, "Cross-linguistic perspectives on the development of text-production abilities: Speech and writing", *Written Language & Literacy*, 5(1), pp. 1-43.
- Balboni, P. E., 2015, *Le sfide di Babele. Insegnare le lingue nelle società complesse*, UTET Università, Torino.
- Beers, S. F., & Nagy, W. E., 2009, "Syntactic complexity as a predictor of adolescent writing quality: which measures? Which genre?", *Reading and writing*, 22(2), pp. 185-200.
- Berruto, G., 2006, *Corso elementare di linguistica generale*, UTET, Torino.
- Bettoni, C., 2001, *Imparare un'altra lingua. Lezioni di linguistica applicata*, Laterza, Bari.
- Bisang, W., 2009, "On the evolution of complexity: sometimes less is more in East and mainland Southeast Asia", in G. Sampson, D. Gil, & P. Trudgill (Eds.), *Language complexity as an evolving variable*, Oxford University Press, Oxford, pp. 34-49.
- Bisang, W., 2014, "Overt and hidden complexity. Two types of complexity and their implications", *Poznań studies in contemporary linguistics*, 50(2), pp. 127-143.
- Brezina, V., Pallotti, G., 2015, Morphological complexity tool, disponibile su [http://corpora.lancs.ac.uk/vocab/analyse\\_morph.php](http://corpora.lancs.ac.uk/vocab/analyse_morph.php)
- Brezina, V., & Pallotti, G., 2016, "Morphological complexity in written L2 texts", *Second Language Research*, pp. 1-21.

- Bulté, B., Housen, A., 2012, "Defining and operationalising L2 complexity", in A. Housen, F. Kuinen & I. Vedder (Eds.), *Dimensions of L2 Performance and Proficiency. Complexity, Accuracy and Fluency in SLA*, John Benjamins, Amsterdam/Philadelphia, pp. 21-46.
- Bulté, B., Housen, A., 2015, "Evaluating short-term changes in L2 complexity development", *CLAC*, 63, 42-76.
- Bygate, M., Skehan, P., & Swain, M. (eds.), 2001, *Researching pedagogic task. Second Language Learning, Teaching and Testing*, Pearson Education, Harlow.
- Campbell, D. J., 1988, "Task complexity: A review and analysis", *Academy of Management Review*, 13, pp. 40-52.
- Consiglio d'Europa, 2001, Quadro comune europeo di riferimento per le lingue: apprendimento, insegnamento, valutazione, disponibile su <http://www.memorbalia.it/descrittori/dalframeworkeuropeo.pdf> (08/10/2018).
- Dahl, Ö., 2004, *The growth and maintenance of linguistic complexity*, John Benjamins, Amsterdam/Philadelphia.
- Dahl, Ö., 2009, "Testing the assumption of complexity invariance: the case of Elfdalian and Swedish", in G. Sampson, D. Gill & P. Trudgill (Eds.), *Language complexity as an evolving variable*, Oxford University Press, Oxford, pp. 50-63.
- Daller, M., Turlik, J., & Weir, I., 2013, "Vocabulary acquisition and the learning curve", in S. Jarvis & M. Daller (Eds.), *Vocabulary knowledge: Human ratings and automated measures*, John Benjamins, Amsterdam/Philadelphia, pp. 185-217.
- De Bot, K., Lowie, W., & Verspoor, M., 2007, "A Dynamic Systems Theory approach to second language acquisition", *Bilingualism: Language and Cognition*, 10(1), Cambridge University Press, pp. 7-21
- De Clercq, B., Housen, A., 2015a, "A Cross-linguistic Perspective on Syntactic Complexity in L2 Development: Structural Complexity and Diversity", *The Modern Language Journal*.

- De Clercq, B., 2015b, "The development of lexical complexity in second language acquisition: a cross-linguistic study of L2 French and English", *EuroSLA Yearbook*, 15, pp. 69-94.
- De Clercq, B., Housen, A., 2015c, "The development of morphological complexity: a cross-linguistic study of L2 French and English", *Second Language Research*.
- De Jong, N. H., Steinel, M. P., Florijn, A., Schoonen, R. & Hulstijn, J. H., 2012, "The effect of task complexity on functional adequacy, fluency and lexical diversity in speaking performances of native and non-native speakers", in A. Housen, F. Kuiken & I. Vedder (Eds.), *Dimensions of L2 Performance and Proficiency. Complexity, Accuracy and Fluency in SLA*, John Benjamins, Amsterdam/Philadelphia.
- Devoto, G., Oli, C. C., 2000, *Dizionario della lingua italiana*, Le Monnier, Firenze.
- Dillard, J. P., & Pfau, M., 2002, *The persuasion handbook: Developments in theory and practice*, CA: Sage, Thousand Oaks.
- Ellis, R., 2003, *Task-based Language Learning and Teaching*, Oxford University Press, Oxford.
- Ellis, R., & Yuan F., 2005, "The effects of careful within-task planning on oral and written task performance", in R. Ellis (Eds.), *Planning and Task Performance in a Second Language*, John Benjamins, Amsterdam/Philadelphia, pp.167-192.
- Fergadiotis, G. et al., 2013, "Measuring Lexical Diversity in Narrative Discourse of People With Aphasia", *Am J Speech Lang Pathol*, 22(2), pp. 1-22.
- Ferrari, S., 2012, "A longitudinal study of complexity, accuracy and fluency variation in second language development", in A. Housen, F. Kuiken & I. Vedder (Eds.), *Dimensions of L2 Performance and Proficiency. Complexity, Accuracy and Fluency in SLA*, John Benjamins, Amsterdam/Philadelphia, pp. 277-98.
- Frencich, R., Torresan, P., 2005, *Giochi senza frontiere. Attività ludiche per l'insegnamento dell'italiano*, Alma, Firenze.
- Gill, T. G., Hicks, R. C., 2006, "Task complexity and informing science: A synthesis",

*Informing Science*, 9, pp. 1-30.

Graffi, G., & Scalise, S., 2003, *Le lingue e il linguaggio, introduzione alla linguistica*, Il Mulino, Bologna.

Guiraud, F., 1954, *Les Caracteres Statistiques du Vocabulaire. Essai de methodologie*, Presses Universitaires de France, Parigi.

Hassanali, K. N., Liu, Y., Iglesias, A., Solorio, T., Dollaghan, C., 2014, "Automatic generation of the index of productive syntax for child language transcripts", *Behaviour research methods*, 46(1), pp. 254-262.

Housen, A., Bulté, B., Pierrard, M., Van Daele, S., 2008, "Investigating lexical proficiency development over time. The case of Dutch-speaking learners of French in Brussels", *Journal of French Language Studies*, 18(3), pp. 277-298.

Housen, A., Kuiken, F., 2009, "Complexity, accuracy, and fluency in second language acquisition", *Applied Linguistics*, 30(4), pp. 461-473.

Housen, A., Kuiken, F., Vedder, I., (Eds.), 2012a, *Dimensions of L2 Performance and Proficiency. Complexity, Accuracy and Fluency in SLA*, John Benjamins, Amsterdam/Philadelphia.

Housen, A., Kuiken, F., Vedder, I., 2012b, "Complexity, accuracy and fluency. Definitions, measurement and research", in A. Housen, F. Kuiken & I. Vedder (Eds.), *Dimensions of L2 Performance and Proficiency. Complexity, Accuracy and Fluency in SLA*, John Benjamins, Amsterdam/Philadelphia, pp. 1-20.

Jarvis, S., 2013a, "Capturing the diversity in lexical diversity", *Language Learning*, 63, pp. 87-106.

Jarvis, S., 2013b, "Defining and measuring lexical diversity", in Jarvis, S., Daller, M. (Eds.), *Vocabulary knowledge: human ratings and automated measures*, John Benjamins, Amsterdam/Philadelphia, pp. 13-44.

Jefferson, G., 2004, "Glossary of transcript symbols with an Introduction", in G. H. Lerner (Eds.), *Conversation Analysis: Studies from the first generation*, John Benjamins, Amsterdam/Philadelphia, pp. 13-23.

- Jefferson, G., Drew, P., Heritage, J., Lerner, G. H., Pomerantz, A. (Eds.), 2015, *Talking about troubles in conversation*, Oxford University Press, New York.
- Johansson, V., 2008, "Lexical diversity and lexical density in speech and writing: a developmental perspective", *Lund Working Papers in Linguistics*, 53, pp. 61-79.
- Kusters, W., 2003, *Linguistic complexity. The influence of social change on verbal inflection*, LOT, Utrecht.
- Lambert, C., Kormos, J., 2014, "Complexity, accuracy, and fluency in task-based L2 research: Toward more developmentally based measures of second language acquisition", *Applied Linguistics*, 35(5), pp. 1-9.
- Larsen-Freeman, D., 2006, "The emergence of complexity, fluency, and accuracy in the oral and written production of five Chinese learners of English", *Applied Linguistics*, 27(4), pp. 590-619.
- Laufer, B., Nation, P., 1995, "Vocabulary size and use: Lexical richness in L2 written production", *Applied Linguistics*, 16(3), pp. 307-322.
- Lewis, M., 1993, *The Lexical Approach*, Language Teaching Publications, Hove.
- Lindqvist, C., Gudmundson, A., Bardel, C., 2013, "A new approach to measuring lexical sophistication in L2 oral production", in C. Bardel, C. Lindqvist & B. Laufer (Eds.), *L2 vocabulary acquisition, knowledge and use: new perspectives on assessment and corpus analysis*, EuroSLA, pp. 109-126.
- Long, M., 1985, "A role for instruction in second language acquisition: Task-based language teaching", in K. Hyltenstam & M. Pienemann, *Modelling and Assessing Second Language Acquisition*, Multilingual Matters Ltd, Clevedon.
- Long, M., 2015, *Second language acquisition and task-based language teaching*, John Wiley and Sons, Chichester.
- Malvern, D., Richards, B. J., Chipere, N., Duran, P., 2004, *Lexical diversity and language development: Quantification and assessment*, Palgrave Macmillan, Basingstoke.

- Mayer, M., 1969, *Frog, where are you?*, Penguins Young Readers, New York.
- McCarthy, P. M. & Jarvis, S., 2010, "MTLD, vocd-D, and HD-D: a validation study of sophisticated approaches to lexical diversity assessment", *Behavior Research Methods*, 42(2), pp. 381-392.
- McWhorter, J. H., 2001, "The world's simplest grammars are creole grammars", *Linguistic typology*, 5, pp. 125-166.
- McWhorter, J. H., 2011, *Linguistic simplicity and complexity: Why do languages undress?*, De Gruyter, Berlin/Boston.
- Mehnert, U., 1998, "The effects of different lengths of time for planning on second language performance", *Studies in second language acquisition*, 20(1), pp. 83-108.
- Michel, M. C., 2011, "Effects of task complexity and interaction on L2 performance", in P. Robinson (Eds.), *Second Language Task Complexity. Researching the Cognition Hypothesis of Language Learning and Performance*, John Benjamins, Amsterdam/Philadelphia.
- Miestamo, M., Sinnemaki, K., & Karlsson, F. (Eds.), 2008, *Language complexity: Typology, contact, change*, John Benjamins, Amsterdam/Philadelphia.
- Miestamo, M., 2009, "Implicational hierarchies and grammatical complexity", in G. Sampson, D. Gil & P. Trudgill (Eds.), *Language complexity as an evolving variable*, Oxford University Press, Oxford, pp. 80-97.
- Mitchell, M., 2009, *Complexity: A guided tour*, Oxford University Press, Oxford.
- Nichols, J., 2009, "Linguistic complexity: a comprehensive definition and survey", in G. Sampson, D. Gil & P. Trudgill, *Language complexity as an evolving variable*, Oxford University Press, Oxford, pp. 110-125
- Norris, J. M., & Ortega, L., 2009, "Towards an organic approach to investigating CAF in instructed SLA: The case of complexity", *Applied Linguistics*, 30(4), pp. 555-578.



- Nuzzo, E., 2009, "Buongiorno, ho bisogno dell'informazione per andare a Barcellona: Uno studio longitudinale sulle richieste di informazioni e suggerimenti in italiano L2", *Linguistica e Filologia*, 28, pp. 83-109.
- Ortega, L., 2003, "Syntactic complexity measures and their relationship to L2 proficiency: A research synthesis of college-level L2 writing", *Applied Linguistics*, 24(4), pp. 492-518.
- Ortega, L., Byrnes, H., 2008, "Introduction", in L. Ortega, H. Byrnes, *The Longitudinal Study of Advanced L2 Capacities*, Routledge, New York.
- Ortega, L., 2012, "Interlanguage complexity: A construct in search of theoretical renewal", in B. Kortmann & B. Szmrecsanyi (Eds.), *Linguistic complexity: Second language acquisition, indigenization, contact*, De Gruyter, Berlin/Boston, pp. 127-155.
- Pallotti, G., 1998, *La seconda lingua*, Bompiani, Milano.
- Pallotti, G., 1999, "I metodi della ricerca", in R. Galatolo, G. Pallotti (a cura di), *La conversazione. Un'introduzione allo studio dell'interazione verbale*, Raffaello Cortina, Milano.
- Pallotti, G., 2009, "CAF: Defining, refining and differentiating constructs", *Applied Linguistics*, 30(4), pp. 1-19.
- Pallotti, G., Ferrari, S., Nuzzo, E., Bettoni, C., 2010, "Una procedura sistematica per osservare la variabilità nell'interlingua", *Studi Italiani di Linguistica Teorica e Applicata*, 39, pp. 215-41.
- Pallotti, G., 2015a, "A simple view of linguistic complexity", *Second Language Research*, 31(1), pp. 117-134.
- Pallotti, G., 2015b, "Una nuova misura della complessità linguistica: l'indice di complessità morfologica (ICM)", *Rassegna italiana di linguistica applicata*, 2 (3), pp. 195- 215.

- Pallotti, G., Rosi, F., 2015, "La difficoltà interazionale nei task: definizione operativa e risultati di ricerca", in M. Chini (a cura di), *Il parlato in [italiano] L2: aspetti pragmatici e prosodici, [Italian] L2 Spoken Discourse: Pragmatic and Prosodic Aspects*, Franco Angeli, Milano, pp. 243-61.
- Pallotti, G., 2017, "Assessing Tasks: The Case of Interactional Difficulty", *Applied Linguistics*, pp. 1-23.
- Porcelli, G., 2003, "Segmenti lessicali e mappe semantiche", *ITALS Didattica e linguistica dell'italiano a stranieri*, I-1, pp. 65-74.
- Ramat, P., 1987, *Linguistic typology*, De Gruyter, Berlino.
- Read, J., 2000, *Assessing vocabulary*, Cambridge University Press, Cambridge.
- Rescher, N., 1998, *Complexity: A philosophical overview*, Transaction, London.
- Robinson, P., 2001, "Task complexity, task difficulty, and task production: exploring interactions in a componential framework", *Applied linguistics*, 22 (1), pp. 27-57.
- Robinson, P. (Eds.), 2007a, *Cognition and second language instruction*, Cambridge University Press, Cambridge.
- Robinson, P., 2007b, "Task complexity, theory of mind, and intentional reasoning: Effects on L2 speech production, interaction, uptake and perceptions of task difficulty", *International Review of Applied Linguistics in Language Teaching*, 45(3), pp. 193-213.
- Robinson, P., 2007c, "Task complexity, cognitive resources, and syllabus design: a triadic framework for examining task influences on SLA", in P. Robinson (Eds.), *Cognition and second language instruction*, Cambridge University Press, Cambridge.
- Robinson, P., (Eds.), 2011, *Second language task complexity: Researching the Cognition Hypothesis of language learning and performance*, John Benjamins, Amsterdam/Philadelphia.

- Sampson, G., 2009, "A linguistic axiom challenged", in G. Sampson, D. Gil & P. Trudgill (Eds.), *Language complexity as an evolving variable*, Oxford University Press, Oxford, pp. 1-18.
- Sampson, G., Gil, D. & Trudgill, P., (Eds.), 2009, *Language complexity as an evolving variable*, Oxford University Press, Oxford.
- Schmid, H., 1994, TreeTagger, TC project at the Institute for Computational Linguistics of the University of Stuttgart.
- Selinker, L., 1992, *Rediscovering interlanguage*, Longman, London.
- Shosted, R., 2006, "Correlating complexity: a typological approach", *Linguistic Typology*, 10, pp. 1-40.
- Skehan, P., & Foster, P., 2001, "Cognition and tasks", in P. Robinson (Eds.), *Cognition and Second Language Learning*, Cambridge University Press, Cambridge, pp. 183-205.
- Skehan, P., 1989, *Individual differences in Second-Language Learning*, Routledge, New York.
- Skehan, P., 1996, "A framework for the implementation of task based instruction", *Applied Linguistics*, 17(1), pp. 38-62.
- Skehan, P., 1998, *A cognitive approach to language learning*, Oxford University Press, Oxford.
- Skehan, P., 2001, "Tasks and language performance", in M. Bygate, P. Skehan, & M. Swain (Eds.), *Research pedagogic tasks: Second language learning, Teaching and testing*, Pearson Education, Harlow, pp. 167-185.
- Skehan, P., 2003, "Task-based instruction", *Language Teaching*, 36, pp. 1-14.
- Skehan, P., 2009a, "Modelling second language performance: Integrating complexity, accuracy, fluency and lexis", *Applied Linguistics*, 30(4), pp. 510-532.

- Skehan, P., 2009b, "Lexical performance by native and non-native speakers on language-learning tasks", in B. J. Richards, M. Daller, D. Malvern, P. Meara, J. Milton & J. Treffers-Daller (Eds.), *Vocabulary studies in first and second language acquisition: The interface between theory and application*, Palgrave Macmillian, London, pp. 107-124.
- Spoelman, M., & Verspoor, M., 2010, "Dynamic patterns in development of accuracy and complexity: A longitudinal case study in the acquisition of Finnish", *Applied Linguistics*, 31(4), pp. 532-553.
- Swain, M. & Lapkin, S., 2001, "Focus on form through collaborative dialogue: Exploring task effects", in M. Bygate, P. Skehan & M. Swain (Eds.), *Researching pedagogic task. Second Language Learning, Teaching and Testing*, Pearson Education, Harlow.
- Szmrecsanyi, B., Kortmann, B., 2012, "Introduction", in B. Kortmann & B. Szmrecsanyi (Eds.), *Linguistic complexity: Second language acquisition, indigenization, contact*, De Gruyter, Berlin/Boston, pp. 6-34.
- Towell, R., 2012, "Complexity, accuracy and fluency from the perspective of psycholinguistic second language acquisition research", in A. Housen, F. Kuiken & I. Vedder (Eds.), *Dimensions of L2 performance and proficiency: Complexity, accuracy and fluency in SLA*, John Benjamins, Amsterdam/Philadelphia.
- Verspoor, M., Lowie, W., & Van Dijk, M., 2008, "Variability in second language development from a dynamic system perspective", *The Modern Language Journal*, 92(ii), pp. 214-231.
- Wolfe-Quintero, K., Inagaki, S., & Kim, H.-Y., 1998, *Second language development in writing: Measures of fluency, accuracy, and complexity*, National Foreign Language Resource Center, Honolulu.
- Xanthos, A., Gillis, S., 2010, "Quantifying the Development of Inflectional Diversity", *First Language*, 30(2), pp. 175-198.