

LINGUAGGIO E APPRENDIMENTO LINGUISTICO

Metodi e strumenti tecnologici

a cura di

Francesca Bianchi - Paola Leone

studi AltLA **4**

AItLA

studi AltLA 4

LINGUAGGIO
E APPRENDIMENTO
LINGUISTICO

Metodi e strumenti tecnologici

a cura di

FRANCESCA BIANCHI - PAOLA LEONE

Milano 2016

L'AItLA pubblica una collana di monografie e di collettanee sui diversi temi della linguistica applicata. I manoscritti vengono valutati con i consueti processi di revisione di pari per assicurarne la conformità ai migliori standard qualitativi del settore. I volumi sono pubblicati nel sito dell'associazione con accesso libero a tutti gli interessati.

Comitato scientifico

Giuliano Bernini, Camilla Bettoni, Cristina Bosisio, Simone Ciccolone, Anna De Meo, Laura Gavioli, Elena Nuzzo, Lorenzo Spreafico.

Volume pubblicato con il contributo di BANCA MONTE DEI PASCHI DI SIENA,
erogato tramite il Dipartimento di Studi Umanistici dell'Università del Salento



© 2016 AItLA - Associazione Italiana di Linguistica Applicata
Via Cartoleria, 5
40100 Bologna - Italy
email: info@aitla.it
sito: www.aitla.it



Edizione realizzata da
Officinaventuno
Via Doberdò, 21
20126 Milano - Italy
email: info@officinaventuno.com
sito: www.officinaventuno.com

ISBN edizione cartacea: 978-88-97657-12-5
ISBN edizione digitale (pdf): 978-88-97657-13-2

Indice

PAOLA LEONE	
Introduzione	5

PARTE I

Tecnologie per la ricerca

ANNA DE MEO - MARILISA VITALE - ELISA PELLEGRINO	
Tecnologia della voce e miglioramento della pronuncia in una L2: imitazione e autoimitazione a confronto.	
Uno studio su sinofoni apprendenti di italiano L2	13
PATRIZIA SORIANELLO	
Indici predittivi probabilistici: il caso della geminazione consonantica in italiano L2	27

PARTE II

Tecnologie per comunicare

ALESSANDRO BITONTI	
L'e-mail nella comunicazione accademica fra pragmatica e coesione testuale	51
SABRINA FUSARI - ANTONELLA LUPORINI	
La comunicazione tra studenti e docenti via forum e e-mail: strategie di cortesia	67
STEFANIA SPINA	
Le conversazioni scritte dei <i>social media</i> : un'analisi multidimensionale	83
FRANCESCA CHIUSAROLI	
Scritture brevi e tendenze della scrittura nella comunicazione di Twitter	103

PARTE III

Tecnologie per apprendere

AMELIA BANDINI	
Wiki-eTandem: un progetto di apprendimento collaborativo a distanza	121
CLAUDIO NOBILI - CHIARA MELUZZI	
Riformulazioni attraverso Twitter da parte di studenti italiano L1 e LS: proposta di una tassonomia testuale	139

PAOLO FRASSI - OPHÉLIE TREMBLAY	
Il <i>Réseau Lexical du Français</i> : una banca dati per l'apprendimento del lessico francese	155
VALERIA CARUSO	
Dizionari elettronici e apprendimento delle espressioni idiomatiche: monitoraggio dei bisogni e prospettive future	173
ELISA CORINO	
“Scelto A perché frase B secondo me usa tanti verbi”. Attività riflessive <i>corpus-based</i> per elicitare competenze (meta)linguistiche	191

PAOLA LEONE

Introduzione

Negli ultimi decenni, il legame tra la linguistica applicata e le scienze tecnologiche dell'informazione e della comunicazione (TIC) è divenuto più forte e articolato. Entrambe le discipline si intersecano contribuendo l'una al potenziamento dell'altra, in un rapporto reciproco e fecondo. Lo studio dei reali usi comunicativi, delle modalità, dei tempi e dei processi di apprendimento linguistico sollecita la ricerca tecnologica. Allo stesso modo, le strumentazioni sempre più numerose e sofisticate permettono di creare nuovi contesti d'uso linguistico e di apprendimento, nonché di affinare metodi di analisi linguistica. Come attestano i contributi in questo volume, la stretta connessione tra le TIC e la linguistica applicata si realizza nella ricerca, nelle pratiche comunicative e nella didattica.

Il progresso tecnologico procede a un ritmo accelerato. A differenza di molti studi scientifici in cui l'uso delle strumentazioni fa da sfondo, il presente volume intende invece metterle in primo piano. Lo scopo è tracciare, attraverso studi empirici, l'impatto che esse hanno sulle caratteristiche del discorso mediato dal computer, sull'apprendimento linguistico e metalinguistico e sugli studi relativi al potenziamento di specifiche competenze in L2, come le capacità fonetica e fonologica, ambito in cui l'utilizzo delle strumentazioni per la ricerca sembra essenziale.

I lavori che presentiamo mirano a rispondere a quesiti di sicuro interesse scientifico. Di seguito alcuni: Quali sono i processi di apprendimento della pronuncia in L2? Qual è la struttura del discorso mediato dal computer? Quali sono le regolarità pragmatiche della comunicazione nei messaggi di posta elettronica? Quali sono le caratteristiche delle nuove forme di interazione nelle reti sociali virtuali? Come devono essere strutturati i dizionari online per apprendenti di L2? Come si può costruire un ambiente di apprendimento collaborativo tramite l'impiego di *social network*?

Tecnologie informatiche e della comunicazione nel volume

Nei contributi di De Meo e Sorianello le strumentazioni tecnologiche sono diverse. Esse sono rispettivamente hardware (*Tascam DR-07*) e software (*Sony Sound Forge 7.0*) per registrare; per l'analisi acustica del parlato il software Praat. Nella sezione dedicata alla ricerca si osserva anche l'impiego di metodi per la manipolazione delle caratteristiche acustiche, per la sintesi della voce e per l'analisi delle lingue naturali e dei processi di apprendimento linguistico. Nel lavoro di Sorianello le strumentazioni tecnologiche hanno permesso di analizzare le stringhe, computando indici percentuali di probabilità di realizzazione di determinati item. Gli strumenti hanno

consentito quindi di fare predizioni, di operare in breve calcoli statistici lunghi e complessi. In queste tipologie di studi, la “macchina” interviene quindi sul processo di ricerca stesso, ottimizzando i tempi e restituendo risultati scientificamente affidabili.

Servono invece per comunicare le strumentazioni tecnologiche trattate nei contributi interni alla seconda sezione di questo volume, i cui focus sono la posta elettronica (Bitonti, Fusari e Luporini), il forum in contesti didattici (Fusari e Luporini) e le reti sociali virtuali come Facebook e Twitter (Spina, Chiusaroli). Tali studi analizzano il testo e il processo comunicativo asincrono delle varietà scritte trasmesse. Le prospettive di analisi variano dal sociolinguistico, al pragmatico, al linguistico, ad un paradigma analitico multidimensionale. In questi studi, la tecnologia è integrata nel mezzo di comunicazione e il suo impiego genera nuove forme di interazione da studiare in comparazione con modalità comunicative meno recenti.

Nella sezione dedicata all'apprendimento, la tecnologia crea nuovi contesti comunicativi e di apprendimento collaborativo e partecipato (Bandini, Nobili e Meluzzi) e assiste anche l'apprendente nel suo percorso di approfondimento di conoscenze linguistiche e metalinguistiche attraverso la creazione di risorse quali dizionari e banche dati lessicali (Frassi e Tremblay, Caruso, Corino). I paradigmi per l'analisi a cui si fa riferimento in questi studi sono per lo più di carattere qualitativo. In essi, la tecnologia serve a realizzare esperienze d'uso linguistico tramite nuove modalità comunicative e a creare nuove fonti per accedere al sapere, motivando così nuove ricerche sul suo stesso utilizzo in ambito didattico.

Linguistica applicata e tecnologie? Verso la normalizzazione?

Nella sezione del volume dedicata all'uso delle macchine per la ricerca in Linguistica Applicata, non è inaspettata la presenza di contributi relativi all'apprendimento della fonetica in una lingua non-nativa. In tale ambito di ricerca l'uso delle tecnologie è fondamentale, visto che lo studio del suono è possibile solo catturandone la produzione. Il passaggio dall'ascolto alla registrazione, e all'analisi con le più recenti strumentazioni informatiche, ha permesso di migliorare la qualità scientifica della ricerca fonetica, creando una stretta connessione tra le esigenze della sperimentazione e l'evoluzione delle nuove tecnologie. Il rapporto tra i due ambiti è naturalmente e attualmente imprescindibile ed è, per così dire, normalizzato, ovvero considerato per nulla eccezionale.

Per quanto riguarda le tecnologie per la comunicazione, i contributi interni al volume approfondiscono aspetti del processo comunicativo nella posta elettronica, nei forum e nei *social network*. Si fa ormai un uso esteso di tali modalità comunicative asincrone, come dimostrano alcuni rilevamenti ISTAT. Per quanto riguarda la posta elettronica, i dati relativi al 2014 indicano che 8 internauti su 10 l'hanno utilizzata almeno una volta nei 3 mesi che precedono la rilevazione statistica di dicembre. Più della metà degli utilizzatori di Internet (57% circa) ha inviato messaggi su Facebook e Twitter. Un confronto con i dati del 2013 mostra un decremento

dell'impiego della posta elettronica a fronte di un aumento nell'impiego di altre forme di comunicazione mediate dal computer quali messaggi di reti sociali virtuali (nel 2013: 53,2%). Cresce inoltre l'utilizzo delle telefonate via web anche in video-conferenza. In generale, si assiste ad un aumento delle forme di scambio comunicativo mediato dalle tecnologie che permeano sempre più la nostra vita quotidiana. Interessante osservare come i *social network* non servano solo per mantenere relazioni amicali ma anche per prendere parte (o sentirsi parte?) alla vita pubblica e politica del paese. I dati (e anche il nostro stile di vita) confermano che l'uso delle tecnologie per comunicare tende alla normalizzazione.

Merita particolare attenzione l'uso delle tecnologie nell'insegnamento linguistico, ambito per il quale è necessaria una distinzione tra l'impiego delle TIC nella ricerca e nella pratica didattica. Il primo ha avuto notevoli sviluppi ben rappresentati dalle diverse riviste specialistiche del settore (es. CALICO, ReCALL), alcune delle quali ad accesso libero (es. Language Learning and Technology, Alsic), dai numerosi convegni specialistici, dai vari volumi pubblicati. Gli approfondimenti a carattere scientifico riguardano tanto le tecnologie che assistono nell'apprendimento (chiamate in inglese con l'acronimo CALL, ovvero Computer Assisted Language Learning), tanto quelle che servono a mediare la comunicazione tra esseri umani (chiamate in inglese con l'acronimo CMC, Computer Mediated Communication). Se gli studi sono in aumento, lo stesso non si può dire per l'uso didattico delle tecnologie in contesti formali. Nelle scuole e nelle università, il computer rimane uno strumento poco utilizzato. A titolo esemplificativo, in una ricerca OECD (Organization for Economic Co-operation and Development) relativa all'accesso ai computer nelle scuole di ragazzi età 15-16 anni, pubblicata a ottobre 2015, si evidenzia che nel 2012, in Italia, era disponibile un computer ogni quattro ragazzi. Solo due su tre quindicenni indicavano di utilizzare il computer a scuola. Sebbene si registri una percentuale in crescita rispetto al rilevamento del 2009, si può comunque affermare che l'impiego delle TIC nelle scuole, almeno per quella fascia di età, è distante dalla normalizzazione. Ne è prova il fatto che in molte scuole e in molti atenei i computer sono nei laboratori e vengono utilizzati in modo saltuario se non sporadico. Ma è proprio necessario cercare di procedere verso tale normalizzazione? Sono proprio chiari i vantaggi delle nuove tecnologie nella glottodidattica?

Come evidenziano gli studi nel volume, le tecnologie didattiche hanno un potenziale notevole. L'impiego dei computer a scuola o all'università non è di per sé innovativo se ad esso non si associa anche una reale capacità dell'insegnante di gestire i processi formativi e pedagogici che, grazie al web 2.0 e al web 3.0, possono divenire realmente partecipati e motivanti per tutti. Le tecnologie per la comunicazione promuovono, ad esempio, collaborazione tra studenti e tra docenti anche quando questi lavorano in paesi e contesti culturali differenti; sono quindi un utile mezzo per mettere in contatto parlanti esperti di lingue diverse. Non è secondario, inoltre, il ruolo che le tecnologie hanno nel favorire l'accesso a risorse e conoscenze secondo modalità flessibili e, quindi, adattabili anche ai molteplici bisogni (speciali e non) di apprendimento individuali.

I contributi in breve

Il volume comprende tre sezioni riguardanti le tecnologie per la ricerca, le tecnologie per comunicare e le tecnologie per apprendere. Nella prima sezione, il contributo di De Meo, Vitale e Pellegrino approfondisce l'acquisizione dell'italiano da parte di parlanti di cinese L1. Per valutare l'efficacia dell'imitazione e dell'autoimitazione nel training per il potenziamento delle competenze fonologiche in L2, si è fatto ricorso a un test percettivo nel quale parlanti nativi italofofoni giudicavano la funzione comunicativa (asserzione, ordine e richiesta) e il grado di accento straniero degli enunciati prodotti dagli apprendenti cinesi prima e dopo il training.

Sorianello analizza, invece, l'acquisizione delle geminate da parte di apprendenti di italiano L2, utilizzando un approccio chiamato *Analogical Modeling*. Lo studio evidenzia che gli apprendenti non-nativi si dimostrano più sensibili alla classe e alla sonorità di un fonema piuttosto che alla posizione dell'accento o alla frequenza della parola in esame.

Nella sezione dedicata alle tecnologie per comunicare, Bitonti analizza messaggi di posta elettronica tra studenti e docenti universitari, per mettere in evidenza le funzioni sintattiche e pragmatiche dei connettivi che in essi vengono adoperati. La ricerca mostra l'alta interattività dei testi e il significato sociale dei connettivi, correlato soprattutto con la necessità di attenuare l'asimmetria sociale per preservare la faccia positiva degli interlocutori.

Fusari e Luporini discutono il ruolo della cortesia in messaggi di posta elettronica e di forum tra studenti di inglese come lingua straniera e docenti di un corso universitario. Il contributo mette in luce *pattern* ricorrenti in relazione agli atti linguistici, alla modalità, alle metafore grammaticali. I dati mostrano che le strategie di cortesia degli apprendenti si traducono sul piano formale, ad esempio, nell'uso di verbi modali, di modificatori esterni e di metafore grammaticali interpersonali.

Spina, invece, mira a identificare le caratteristiche distintive di conversazioni sui *social network*, applicando i principi dell'analisi multidimensionale. La studiosa individua 41 tratti distintivi, ad es. segnali discorsivi, dislocazioni a destra e a sinistra, mitigazione e verbi di opinione. L'analisi fattoriale permette, inoltre, di estrarre 4 distinte dimensioni di variazione. Lo studio mostra come le varie forme prodotte e le diverse funzioni messe in atto su Facebook e Twitter riflettano la complessa rete di relazioni sociali che tali forme di comunicazione asincrona realizzano e mantengono.

Si focalizza esclusivamente sulla "scrittura breve" di Twitter il contributo di Chiusaroli che ne analizza le caratteristiche e i possibili orientamenti. Per rispondere ad esigenze di sintesi ed economia si fa ricorso a diverse strategie alcune delle quali sono convenzionali. Nel contributo si discute, in particolare, l'uso e l'effetto dello *slang* – quale vocabolario, abbreviazioni, acronimi, segni e simboli paragrafematici (punteggiatura e faccine, *hashtag*) – sulla egrammatica e sulla sintassi, per definirne nuove funzioni rispetto alle descrizioni più tradizionali di altre forme di comunicazione scritta.

Nell'ultima sezione, dedicata alle tecnologie per apprendere, Bandini sottolinea l'importanza delle tecnologie quale mezzo per costruire percorsi di apprendimento

collaborativi di una lingua straniera all'università. La studiosa discute un contesto "tandem" che si realizza grazie alla programmazione congiunta delle attività da parte di un ateneo italiano (Federico II di Napoli) e di uno tedesco (Università di Bochum), i cui studenti apprendono rispettivamente le lingue native dei loro partner (italiano e tedesco). La scrittura breve su Twitter diventa così un'opportunità per compensare la carenza di momenti di confronto autentico tra pari nativi della lingua di specializzazione. Si sottolinea anche l'opportunità di focalizzare l'attenzione su aspetti formali della LS durante lo svolgimenti di tali attività.

Nobili e Meluzzi discutono la creazione di un corpus di *tweet* scritti da studenti italofoeni nativi e studenti di lingua nativa slovacca che studiano italiano LS. Per la raccolta dati, gli autori fanno svolgere ai partecipanti un compito di riassunto di due generi testuali: una poesia e un articolo di giornale. L'analisi del corpus di *tweet* mette in evidenza tre parametri testuali la cui differente combinazione determina otto classi di testi prodotti dai partecipanti alla sperimentazione.

Frassi e Tremblay presentano una rete lessicale del francese, chiamata *Réseau Lexical du Français* e ne discutono alcune applicazioni per l'apprendimento e l'insegnamento linguistico. Gli autori partono dal presupposto che nella didattica linguistica è necessario potenziare anche l'uso consapevole di categorie metalinguistiche relative al lessico della lingua di studio. Costituiscono il fondamento teorico dello studio i principi di *Théorie Sens-Texte* e di *Lexicologie Explicative et Combinatoire*. Si propongono e si discutono anche alcune attività didattiche.

Caruso mira ad analizzare le possibili difficoltà che apprendenti di italiano L2 possono incontrare nello studio delle espressioni idiomatiche tramite la consultazione del dizionario Treccani online. La ricerca mostra che il dizionario, pensato per parlanti nativi, richiederebbe alcune modifiche, indicate dall'autrice, per poter guidare parlanti di italiano L2 nel potenziamento di tale competenza linguistica e lessicale. Nello studio si indicano tali caratteristiche, alcune delle quali sono in linea con *lexicographical function theory* e con l'OWID- *Sprichwörterbuch*.

Infine, lo studio di Corino approfondisce alcuni aspetti del potenziamento della competenza lessicale. Partendo dalla convinzione che nell'insegnamento linguistico occorre prestare attenzione alla forma linguistica e al potenziamento della consapevolezza metalinguistica, l'autrice analizza il comportamento di apprendenti di italiano L2 nello svolgimento di alcuni esercizi a scelta multipla. Si fa uso per questo scopo del corpus VALICO (Varietà di Apprendimento Lingua Italiana Corpus Online).

Il volume è di interesse per la ricerca e per la didattica linguistica. Il ricercatore troverà nelle pagine del testo una pluralità di prospettive di analisi che si ascrivono agli ambiti di ricerca di linguistica applicata. Il docente di lingua troverà spunti per innovare l'insegnamento linguistico assegnando alle tecnologie un posto di rilievo al passo con i tempi e con gli interessi dei giovani studenti.

PARTE I

TECNOLOGIE PER LA RICERCA

Tecnologia della voce e miglioramento della pronuncia in una L2: imitazione e autoimitazione a confronto. Uno studio su sinofoni apprendenti di italiano L2

This study investigates the effectiveness of imitation and self-imitation techniques for the acquisition of L2 prosody. For this purpose 19 elementary Chinese learners of L2 Italian were involved in individual sessions of prosody training, devoted to the exercise of the Italian melodic pattern of 2 assertions, 2 orders and 2 requests. The items used for the self-imitation training were obtained by transferring suprasegmental features from 2 native speakers' voices, used as control group, to the L2 ones. The influence of both techniques was assessed by comparing pre- and post-training performances through a perceptual test, in which 46 native Italian listeners evaluated the accentedness and the communicative function of the L2 productions. Both teaching strategies promoted a general improvement in learners' performances, but results vary depending on the communicative function considered.

1. Introduzione

Migliorare la pronuncia della L2 è uno degli obiettivi più difficili da conseguire per un apprendente adulto, il quale potrebbe aver sviluppato nella lingua target un livello di competenza morfo-sintattica, lessicale e testuale tale da consentirgli di comunicare alla stregua di un nativo (Munro - Derwing, 1995), conservando, tuttavia, nel proprio eloquio un mancato o parziale controllo delle caratteristiche segmentali e suprasegmentali della L2, manifestazione del cosiddetto accento straniero.

I lavori condotti nel settore della tecnologia della voce applicata all'apprendimento delle lingue si sono occupati in maniera considerevole di tale problematica (Chun, 2013; Eskenazi, 2009; Levis, 2007; Martin, 2012). Un ambito di studio con interessanti potenzialità e utili ricadute applicative, in particolare per l'acquisizione della competenza prosodico-intonativa in una L2, è quello che sottolinea l'importanza del rapporto tra grado di miglioramento della pronuncia e somiglianza tra la voce dello studente e quella dell'insegnante (Bissiri *et al.*, 2006; Jilka - Möhler, 1998; Nagano - Ozawa, 1990; Peabody - Seneff, 2006; Sundström, 1998; Tang *et al.*, 2001). In un lavoro di Probst *et al.* del 2002, condotto su apprendenti di inglese L2, era emerso che l'apprendimento della pronuncia veniva favorito quando la voce da imitare era simile a quella dell'apprendente in termini di velocità di articolazione e di curva intonativa, ammettendo così l'esistenza di uno *user-dependent golden speaker*. Secondo Felps *et al.* (2009), per imparare a gestire le caratteristiche segmen-

¹ Università degli studi di Napoli L'Orientale.

tali e soprasedimentali della L2 i risultati migliori si otterrebbero attraverso l'autoimitazione, vale a dire se l'apprendente avesse la possibilità di imitare enunciati in L2 prodotti dalla propria voce sinteticamente modificata sulla base del modello nativo della lingua target. La maggior parte di tali studi è stata condotta sull'inglese L2/LS; scarsa attenzione, invece, è stata posta sull'efficacia della tecnica dell'autoimitazione per l'apprendimento della prosodia in altre lingue.

2. Imitazione e autoimitazione nell'apprendimento della pronuncia in italiano L2

Il primo lavoro condotto sull'italiano L2, volto a indagare sperimentalmente l'effetto dell'autoimitazione per rapporto a quello dell'imitazione nell'apprendimento della pronuncia, è stato De Meo *et al.* (2013). Lo studio, attraverso l'utilizzo di una metodologia di apprendimento assistito dal computer e la manipolazione del segnale acustico delle produzioni non native, ha comparato gli effetti prodotti da due modalità di esercitazione, l'imitazione e l'autoimitazione, sulla produzione del livello soprasedimentale dell'italiano L2 in un gruppo di apprendenti cinesi di livello intermedio-avanzato, corrispondente al B2 del Quadro Comune Europeo di Riferimento (Council of Europe, 2001). Ai fini della ricerca, De Meo e colleghi hanno analizzato e sottoposto a valutazione percettiva le produzioni, pre- e post-esercitazione, di due enunciati con quattro diverse funzioni comunicative (richiesta, comando, concessione e minaccia), caratterizzate da un diverso livello di familiarità per apprendenti non nativi.

I risultati dello studio hanno mostrato una generale validità di entrambe le tecniche di esercitazione sia relativamente all'efficacia comunicativa degli enunciati prodotti, sia in termini di riduzione dell'accento straniero percepito. Tuttavia l'autoimitazione ha prodotto risultati mediamente più soddisfacenti, in particolare per la richiesta.

3. Lo studio

3.1 Obiettivo

Il presente studio si pone l'obiettivo di verificare la validità delle tecniche di imitazione e autoimitazione anche su apprendenti di italiano con un livello di competenza linguistico-comunicativa inferiore rispetto a quello testato in De Meo *et al.* (2013). Il lavoro si è articolato in due fasi progressive e interdipendenti: quella del pre-test, servita a selezionare gli enunciati da sottoporre a esercitazione, e quella sperimentale, dalla quale sono state ricavate le considerazioni relative all'efficacia delle tecniche in esame. Nella presentazione del lavoro verranno quindi ripercorse le tappe di sviluppo della ricerca.

3.2 Metodologia

L'autoimitazione, ovvero l'imitazione da parte degli apprendenti della loro stessa voce con prosodia italiana nativa, è stata resa possibile dall'utilizzo della tecnica del trapianto ritmico-intonativo (Yoon, 2007). Tale tecnica, infatti, attraverso la manipolazione del segnale acustico, permette di trasferire su una voce non nativa uno o più parametri acustici (tono, intensità, velocità di articolazione, frequenza e durata delle pause silenti) prelevati da una voce nativa, senza tuttavia alterare né la struttura fonica a livello segmentale né l'identità della voce manipolata. La procedura di manipolazione, che sfrutta l'algoritmo PSOLA - *Pitch - Synchronous Overlap and Add* (Charpentier - Moulines, 1989) implementato nel software Praat (Boersma, 2001), prevede una sequenza fissa di cinque fasi operative: 1. trattamento delle anomalie, 2. segmentazione dell'enunciato, 3. trapianto delle durate, 4. trapianto dell'intensità, 5. sovrapposizione del contorno intonativo (per ulteriori dettagli si veda Pettorino - Vitale, 2012). Al termine della procedura, la voce non nativa si presenta inalterata sul piano segmentale, ma conforme al modello nativo utilizzato dal punto di vista soprasegmentale.

3.3 Il pre-test

Nella fase di pre-test sono stati coinvolti 19 apprendenti cinesi di età compresa tra i 19 e i 22 anni, di cui 17 femmine e 2 maschi, tutti con un livello di competenza della lingua italiana corrispondente all'A2 del Quadro Comune Europeo di Riferimento. Al momento dello svolgimento della ricerca, gli studenti risiedevano a Napoli da circa tre mesi e frequentavano i corsi dell'Università di Napoli 'L'Orientale'. Tutti avevano studiato l'italiano in Cina per circa due anni. In qualità di parlanti modello sono stati coinvolti nello studio anche due soggetti di madrelingua italiana, un maschio e una femmina, di circa 27 anni, entrambi di origine campana, utilizzati come *donatori* dei tratti ritmico-prosodici. Gli apprendenti cinesi sono stati considerati *riceventi* nel corso della procedura di trapianto prosodico.

Per esigenze di comparabilità degli enunciati prodotti dai parlanti nativi con quelli dei non nativi e ai fini della procedura di manipolazione, i soggetti sono stati coinvolti in un'attività di parlato letto. Per la lettura sono stati scelti due enunciati di diversa lunghezza, ciascuno da produrre con tre diverse funzioni comunicative: asserzione (A), comando (C) e richiesta (R). Dal punto di vista della forma espressiva, le asserzioni sono state rese con frasi dichiarative alla terza persona singolare, i comandi con verbi al modo imperativo e le richieste come domande sì/no alla seconda persona singolare.

Trattandosi di studenti con un livello elementare di competenza linguistica, la funzione comunicativa da veicolare è stata indicata sia con la punteggiatura (punto fermo per gli enunciati assertivi, punto esclamativo per i comandi, punto interrogativo per le richieste), sia mediante indicazione esplicita della funzione linguistica corrispondente (asserzione, comando, richiesta). Accanto a ciascun blocco di enunciati è stata inserita un'immagine che, mediante una rappresen-

tazione iconica dell'azione espressa nell'enunciato, facilitasse la decodifica del contenuto. Tutta la fase di presentazione degli enunciati e dell'attività di lettura è stata supportata dalla presenza di un docente cinese di italiano L2, che ha fornito adeguata traduzione dei termini metalinguistici.

Di seguito l'elenco completo delle frasi utilizzate per l'attività di lettura:

Enunciati brevi

- (1) Prende le posate dal cassetto.
- (2) Prendi le posate dal cassetto!
- (3) Prendi le posate dal cassetto?

Enunciati lunghi

- (1) Finisce di prendere la scheda prima di consegnarla.
- (2) Finisci di prendere la scheda prima di consegnarla!
- (3) Finisci di prendere la scheda prima di consegnarla?

Dopo una breve fase di lettura silenziosa degli enunciati, gli studenti hanno registrato il corpus costituito dai sei enunciati in camera anecoica, in sessioni singole, mediante il software Sony Sound Forge 7.0, a una frequenza di campionamento di 44.100 Hz. La stessa procedura di registrazione è stata attuata con i parlanti di madrelingua italiana. Il corpus finale di parlato letto è risultato complessivamente costituito da 114 enunciati in italiano L2 (19 parlanti * 2 enunciati * 3 intenzioni comunicative) e da 12 enunciati in italiano L1 (2 parlanti * 2 enunciati * 3 intenzioni comunicative).

Nella fase di pre-test, gli enunciati prodotti dagli apprendenti cinesi sono stati sottoposti a una valutazione percettiva da parte di sei ascoltatori nativi italiani, di età compresa fra i 30 e i 33 anni, tutti docenti esperti di italiano L2, residenti in Campania, privi di competenze linguistiche in cinese. Tutti gli ascoltatori hanno dichiarato di essere privi di deficit uditivi. Il test è stato somministrato in maniera individuale, attraverso l'ascolto in cuffia degli enunciati, proposti in sequenza randomizzata.

Gli enunciati sono stati valutati sia per il grado di accento straniero (scala da 0 a 5 punti: 0 = accento nativo; 5 = accento straniero molto forte), sia per l'intenzione comunicativa veicolata (cinque alternative: asserzione A, comando C, richiesta R e due distrattori, minaccia M e altro Altro).

Dai risultati del pre-test percettivo si evince che l'eloquio degli apprendenti cinesi è stato giudicato da tutti gli ascoltatori come connotato da forte accento straniero (media = 4,1). Relativamente alla funzione comunicativa, come mostrato nella matrice di confusione in tab. 1, gli enunciati più chiari sono risultati le richieste, che hanno ottenuto una percentuale di esatto riconoscimento pari al 53,6%. Di più complessa interpretazione sono state le asserzioni e i comandi, identificati correttamente da meno della metà degli ascoltatori (46,3% asserzione; 28,7% comando). A ciò va aggiunto che le due funzioni comunicative vengono confuse reciprocamente in circa il 20% dei casi (asserzione > comando: 20,8%; comando > asserzione: 26,9%).

Tabella 1 - *Matrice di confusione tra le intenzioni comunicative attese e quelle percepite*

		<i>Intenzioni percepite</i>				
		A	C	R	M	Altro
<i>Intenzioni attese</i>	A	46,3%	20,8%	18,3%	2,8%	11,8%
	C	26,9%	28,7%	26,4%	8,8%	9,3%
	R	15,3%	16,3%	53,6%	4,1%	10,7%

In base ai risultati del pre-test, sono stati selezionati gli enunciati da sottoporre a esercitazione nella fase sperimentale del lavoro, eliminando per ogni soggetto le produzioni correttamente riconosciute da almeno quattro ascoltatori nativi su sei, e ridefinendo in tal modo il corpus di enunciati attraverso cui testare l'efficacia delle tecniche di imitazione e autoimitazione. Dei 114 enunciati di partenza, prodotti dai 19 soggetti coinvolti nel pre-test, ne sono stati quindi sottoposti a esercitazione solo 51, prodotti da 16 locutori (tab. 2).

Tabella 2 - *Confronto fra il corpus iniziale e quello sottoposto a esercitazione*

	<i>Pre-test</i>	<i>Esercitazione</i>
Parlanti cinesi	19	16
Enunciati	114	51

Per la metà degli enunciati (25/51) gli apprendenti si sono esercitati con un'attività di imitazione, per la restante parte (26/51) hanno invece imitato la propria voce prosodicamente rimodellata su quella nativa. Laddove il corpus lo rendeva possibile, le due tecniche sono state bilanciate anche per intenzione comunicativa, sebbene il corpus definitivo sia risultato costituito prevalentemente da comandi (42,3%) e asserzioni (40,4%) e solo in minima percentuale da richieste (17,3%), poiché queste ultime sono risultate più facilmente riconosciute nel pre-test. Rispetto alla variabile lunghezza della frase, il 47% del corpus è costituito da enunciati brevi (24), il 53% da quelli lunghi (27). Questi ultimi hanno ottenuto nel pre-test un numero minore di riconoscimenti corretti dell'intenzione comunicativa, a causa della maggiore difficoltà incontrata dagli apprendenti nella gestione della curva intonativa. L'allineamento delle variazioni tonali con le diverse unità linguistiche costitutive di un enunciato (foni, sillabe, parole, sintagmi) nel continuum di una curva melodica avente maggiore estensione temporale richiede una elaborazione dei tratti prosodici più complessa, che per i parlanti non nativi coinvolti è risultata inevitabilmente più problematica.

3.4 La fase sperimentale

3.4.1 Le attività di imitazione e autoimitazione

Come sintetizzato in tab. 3, per motivi tecnici l'attività di autoimitazione prevede alcune fasi di lavoro aggiuntive e preliminari rispetto all'attività di imitazione.

Tabella 3 - *Fasi attuative previste dalle attività di imitazione e autoimitazione*

<i>Fase</i>	<i>Imitazione</i>	<i>Autoimitazione</i>
<i>Trattamento del segnale acustico</i>		1. Segmentazione manuale degli enunciati target in italiano L1 e L2 in porzioni vocaliche e consonantiche 2. Trattamento delle anomalie e allineamento dei segmenti tra gli enunciati in L1 e in L2 3. Trapianto prosodico intonativo
<i>Esercitazione</i>	1. Ascolto singolo, in cuffia, degli enunciati prodotti dal parlante nativo 2. Esercitazione individuale 3. Nuova registrazione imitando il modello nativo	4. Ascolto singolo, in cuffia, degli enunciati con la prosodia del modello nativo trapiantata sulla voce dell'ascoltatore 5. Esercitazione individuale 6. Nuova registrazione autoimitando

Per la somministrazione dell'attività di autoimitazione è stato necessario modificare artificialmente i parametri ritmico-intonativi di ciascuna produzione non nativa, secondo la procedura descritta in §3.2, per dare a ciascun partecipante coinvolto nella sperimentazione la possibilità di ascoltare e imitare la propria voce con prosodia nativa italiana. L'imitazione, invece, non ha implicato la manipolazione del segnale acustico, in quanto ai parlanti non nativi è stato richiesto di imitare la voce naturale di un nativo italiano.

Ai fini dello studio, gli apprendenti cinesi sono stati suddivisi in due gruppi, entrambi sottoposti a una attività di esercitazione della prosodia attraverso apprendimento assistito dal computer, ma differenziati per tipologia di compito: imitazione vs. autoimitazione. La fase di esercitazione individuale ha previsto, quindi, l'ascolto in cuffia di enunciati modello naturali (imitazione) o manipolati (autoimitazione) e la ripetizione ad alta voce degli stessi, con l'obiettivo di avvicinarsi il più possibile al modello intonativo proposto. Raggiunto un grado di avvicinamento giudicato soddisfacente dagli stessi apprendenti, ciascuno studente ha registrato nuovamente gli enunciati esercitati.

3.4.2 Il test percettivo

Il corpus utilizzato per la valutazione percettiva è costituito complessivamente da 110 enunciati in italiano L2, di cui 51 prodotti prima di effettuare l'esercitazione (24 brevi + 27 lunghi) e 51 registrati dopo l'esercitazione basata sull'ascolto della

voce modello di riferimento (24 brevi + 27 lunghi). A questi enunciati ne sono stati aggiunti altri 8 (4 brevi e 4 lunghi), con funzione di distrattori, della stessa tipologia di quelli impiegati per la sperimentazione, ma prodotti da apprendenti non coinvolti nello studio. Per evitare che l'ascolto di un numero eccessivo di enunciati potesse incidere sull'esito del test percettivo, i 52 enunciati brevi sono stati separati dai 58 lunghi e i due sottogruppi di file audio sono stati somministrati a due diversi gruppi di ascoltatori nativi (test 1 e 2), mediante il software online Survey Gizmo. I file audio risultano così distribuiti:

test 1 (enunciati brevi)

- 24 pre-esercitazione
- 24 post-esercitazione (12 imitazione + 12 autoimitazione)
- 4 distrattori;

test 2 (enunciati lunghi)

- 27 pre-esercitazione
- 27 post-esercitazione (13 imitazione + 14 autoimitazione)
- 4 distrattori.

Ciascuno dei due gruppi di ascoltatori coinvolti nella valutazione percettiva era costituito da 20 soggetti, simili per profilo sociolinguistico al gruppo di ascoltatori impegnati nella fase di pre-test, tutti docenti di italiano L2 in formazione, residenti in Campania, di età media 35 anni, privi di competenze linguistiche in cinese. Tutti gli ascoltatori hanno dichiarato di essere privi di deficit uditivi. Come nella fase di pre-test, le valutazioni percettive, eseguite in maniera individuale attraverso l'ascolto in cuffia degli enunciati in sequenza randomizzata, hanno riguardato il grado di accento straniero del parlante e la funzione comunicativa degli enunciati.

3.4.3 Analisi dei risultati

Prima di procedere al commento dei risultati del test percettivo, è opportuno chiarire i criteri adottati per l'analisi dei dati.

Partendo da considerazioni generali sul riconoscimento della funzione comunicativa veicolata dagli enunciati in L2, si è proseguito con valutazioni specifiche sull'adeguatezza delle due tecniche didattiche di imitazione e autoimitazione ai fini del miglioramento della competenza comunicativa. È stata successivamente verificata l'efficacia di imitazione e autoimitazione per rapporto alla lunghezza degli enunciati. Infine è stata considerata la validità dell'autoimitazione e dell'imitazione per l'attenuazione dell'accento straniero.

Il rapporto tra le intenzioni comunicative attese e quelle percepite, rappresentato dalla matrice di confusione in tab. 4, ha confermato la tendenza già emersa nel pre-test: la richiesta risulta l'atto linguistico più facilmente riconoscibile da parte degli ascoltatori italiani (74,5%); meno netti sono stati i giudizi forniti sugli enunciati prodotti con funzione di comando (50,8%) e di asserzione (44,2%). Rispetto

alla fase di pre-test risulta incrementata la percentuale di ascoltatori che confonde l'asserzione con il comando (A > C, pre-test: 20,8%; test: 31,1%).

Tabella 4 - *Matrice di confusione tra intenzioni attese e quelle percepite nella fase successiva all'esercitazione*

		Intenzioni percepite				
		A	C	R	M	Altro
Intenzioni attese	A	44,2%	31,1%	14,3%	4,2%	6,3%
	C	20,5%	50,8%	17,2%	6,9%	4,6%
	R	6,5%	15,8%	74,5%	1,1%	2,2%

Relativamente alle valutazioni sulla validità delle due tecniche didattiche ai fini del potenziamento dell'efficacia comunicativa è bene evidenziare il miglioramento generale delle prestazioni degli studenti cinesi a seguito dell'esercitazione. Come mostrato in tab. 5, le percentuali di corretta identificazione degli enunciati nella fase successiva all'esercitazione (Cin_Post) superano quelle ottenute nella fase di pre-test (Cin_Pre), benché in misura diversa tra le tre funzioni comunicative considerate e per rapporto alla lunghezza dell'enunciato.

Tabella 5 - *Confronto delle percentuali di corretta identificazione delle intenzioni comunicative tra la fase di pre-test (Cin_Pre) e quella successiva all'esercitazione (Cin_Post) per tecnica (imitazione e autoimitazione) e test (1 e 2)*

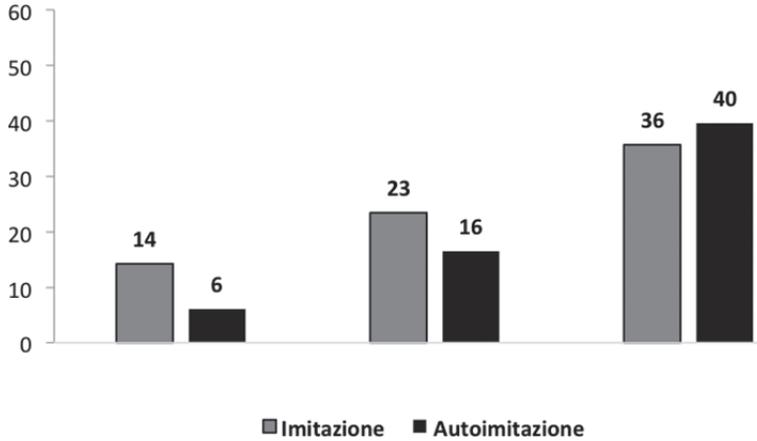
		TEST 1			TEST 2		
		A	C	R	A	C	R
Imitazione	Cin_Pre	41,3	17,5	55	25	25	25
	Cin_Post	40	37,5	75	54,8	51,7	76,2
	Scarto	-1,3	20	20	29,8	26,7	51,2
Autoimitazione	Cin_Pre	45	28,8	31,7	29,2	38,1	33,3
	Cin_Post	41	42,5	80	45,2	57,1	64,3
	Scarto	-4	13,8	48,3	16,1	19	31

Come evidenziato in fig. 1, l'asserzione è la funzione che migliora con più difficoltà, indipendentemente dalla tecnica utilizzata (imitazione o autoimitazione): lo scarto percentuale tra la fase di pre-test e quella successiva all'esercitazione si aggira intorno al 15% nel caso degli enunciati sottoposti a imitazione e al 5% per quelli esercitati con autoimitazione.

Sul versante opposto si ritrova la richiesta, la funzione comunicativa per la quale le tecniche didattiche impiegate hanno prodotto i risultati migliori. Per gli enunciati imitati lo scarto tra la percentuale di corretta identificazione dell'intenzione comunicativa ottenuta nella fase di pre-test e quella successiva all'esercitazione raggiunge circa il 35%; per quelli autoimitati la differenza raggiunge il 40%.

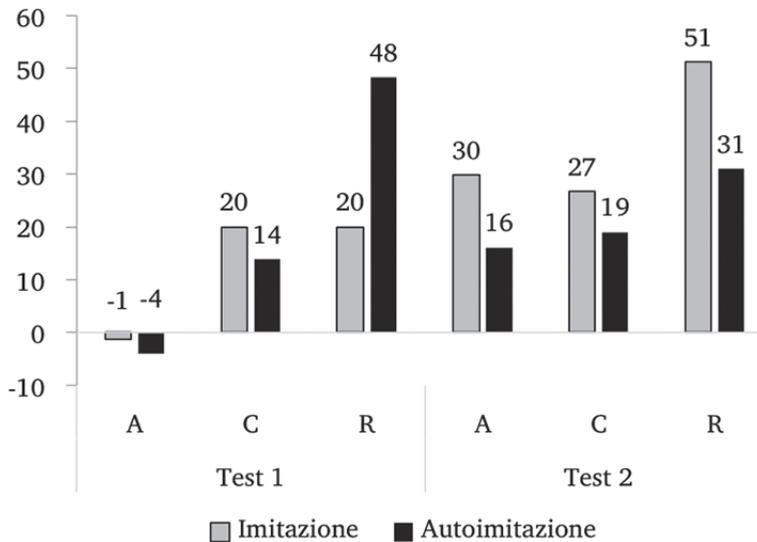
A metà strada si colloca il comando, per il quale lo scarto tra la fase di pre-test e quella successiva all'esercitazione varia dal 23,5% nel caso degli enunciati esercitati attraverso l'imitazione, al 16,4% per quelli esercitati mediante autoimitazione.

Figura 1 - Percentuali di scarto tra la fase di pre-test e quella successiva all'esercitazione per intenzione comunicativa e tecnica



Scorpendo ulteriormente i dati per lunghezza degli enunciati (test 1 e 2) (fig. 2) è possibile valutare con maggiore accuratezza il ruolo delle variabili considerate nella produzione di profili intonativi compatibili alle attese del nativo.

Figura 2 - Percentuali di scarto tra la fase di pre-test e quella successiva all'esercitazione per intenzione comunicativa, tecnica e lunghezza degli enunciati



Osservando i dati di fig. 2, si può affermare che nel caso delle richieste e dei comandi, entrambe le tecniche hanno contribuito al miglioramento delle produzioni degli studenti, indipendentemente dalla lunghezza degli enunciati. Per le frasi brevi, lo scarto medio tra la fase di pre-test e quella successiva all'esercitazione si aggira intorno al 17% per i comandi (imitazione: 20%; autoimitazione: 13,8%) e al 34% per le richieste (imitazione: 20%; autoimitazione: 48,3%). Per gli enunciati lunghi la soglia sale al 23% per i comandi (imitazione: 26,7; autoimitazione: 19%) e al 41% per le richieste (imitazione: 51,2%; autoimitazione: 31%). Nella realizzazione dell'asserzione, le due tecniche hanno avuto un effetto positivo solo sugli enunciati lunghi (imitazione: 30%; autoimitazione: 16,1%).

Analizzando, invece, l'efficacia della singola tecnica per funzione comunicativa e lunghezza dell'enunciato, è possibile osservare che:

- per la richiesta:
 - l'imitazione sortisce l'effetto più significativo nel caso degli enunciati lunghi (test 2), raggiungendo un scarto del 51,2% rispetto alla fase di pre-test; l'autoimitazione sembra essere più efficace nella riproduzione degli enunciati brevi, con uno scarto del 48,3% rispetto alla fase di pre-test;
- per il comando:
 - entrambe le tecniche sono risultate più efficaci nel caso degli enunciati lunghi (scarto test 1: imitazione = 20%; autoimitazione = 13,8%; scarto test 2: imitazione = 26,7%; autoimitazione = 19%);
- per le asserzioni:
 - l'esercitazione sembra aver migliorato la capacità degli apprendenti di riprodurre andamenti intonativi conformi alle attese degli ascoltatori nativi solo nel caso di enunciati lunghi (test 2). A seguito dell'attività di imitazione la percentuale di corretto riconoscimento aumenta quasi del 30%, mentre nel caso dell'autoimitazione si osserva un incremento del 16,1%. Per gli enunciati brevi (test 1), invece, non si osservano variazioni positive nell'identificazione delle asserzioni in fase successiva all'esercitazione: la percentuale di esatto riconoscimento subisce un lieve calo.

Per quanto riguarda le valutazioni relative al grado di accento straniero, la possibilità di riprodurre asserzioni, comandi e richieste imitando il modello nativo o autoimitando la propria voce ha prodotto una riduzione del giudizio di forestierismo solo nel caso degli enunciati brevi del test 1 (tab. 6). Di contro, non sono state rilevate variazioni significative tra la fase di pre-test e quella successiva all'esercitazione nel caso degli enunciati lunghi del test 2.

Tabella 6 - *Confronto delle percentuali di giudizio di accento straniero fornito per gli enunciati tra la fase di pre-test (Cin_Pre) e quella successiva all'esercitazione (Cin_Post) per tecnica (imitazione e autoimitazione) e test (1 e 2)*

		TEST 1			TEST 2		
		A	C	R	A	C	R
<i>Imitazione</i>	Cin_Pre	4,1	3,9	3,5	4,1	4,1	3,9
	Cin_Post	3	3,7	2,4	4,4	4	3,8
	Scarto	1,1	0,2	1,1	-0,3	0,1	0,1
<i>Autoimitazione</i>	Cin_Pre	4,1	3,6	3,6	4	4	4
	Cin_Post	3,7	2,3	2,9	3,9	4,1	4,5
	Scarto	0,4	1,3	0,7	0,1	-0,1	-0,5

Sebbene nessuno degli apprendenti cinesi sia sceso al di sotto del livello medio della scala proposta per il giudizio di accento straniero, a seguito dell'attività di esercitazione si osserva comunque un lieve miglioramento, indipendente dal tipo di enunciato (A, C, R) e dal tipo di esercitazione svolta (imitazione vs. autoimitazione).

4. Conclusioni

L'obiettivo di questo lavoro è stato quello di verificare la validità delle tecniche di imitazione e di autoimitazione ai fini del miglioramento della pronuncia e dell'efficacia comunicativa in italiano L2. Lo studio, condotto su apprendenti cinesi di livello A2, ha messo in evidenza che le due tecniche favoriscono nei partecipanti la memorizzazione e la riproduzione di profili prosodico-intonativi idonei a soddisfare il sistema di aspettative che consente all'ascoltatore nativo di comprendere gli intenti comunicativi dell'interlocutore.

Nel compito di riconoscimento della funzione comunicativa degli enunciati, la percentuale di coincidenza fra l'intenzione attesa e quella percepita è superiore nella fase successiva all'esercitazione rispetto a quella di pre-test, fatta eccezione per gli enunciati assertivi brevi (fig. 2). La richiesta, come già in De Meo *et al.* (2013), si conferma l'enunciato più semplice da riprodurre, imitare e/o auto-imitare e per il quale entrambe le tecniche hanno inciso sulla qualità delle produzioni degli apprendenti cinesi; tuttavia, a parità di incremento della percentuale di riconoscimento nella fase successiva all'esercitazione (50%), l'autoimitazione è stata più efficace per gli enunciati brevi, l'imitazione per quelli lunghi. Più complessi sono risultati l'asserzione e il comando, per i quali lo scarto tra la fase di pre-test e quella successiva all'esercitazione risulta più limitata (fig. 1).

Diversamente da quanto precedentemente emerso nello studio di De Meo *et al.* (2013), condotto su studenti cinesi con un livello di competenza linguistica più alta (B2), i dati di questo studio, condotto su apprendenti di livello A2, non forniscono indicazioni univoche sulla maggiore adeguatezza di una delle due tecniche, imitazione o autoimitazione, per il miglioramento della pronuncia a livello prosodico.

Con riferimento agli enunciati brevi, l'autoimitazione risulta essere più efficace per il miglioramento delle richieste, mentre appare meno incisiva per gli altri due tipi di enunciati. La tecnica dell'imitazione sembra aver giocato un ruolo più rilevante per gli enunciati lunghi.

Relativamente al grado di accento straniero percepito, le due tecniche didattiche, imitazione o autoimitazione, hanno influito in maniera positiva solo nel caso di enunciati brevi. A differenza di quanto era emerso nel lavoro condotto su parlanti cinesi di livello intermedio-avanzato (De Meo *et al.*, 2013), in questo studio nessuno degli apprendenti cinesi è sceso al di sotto del livello medio della scala proposta per il giudizio di accento straniero. Plausibilmente, con apprendenti di livello elementare, caratterizzati da un eloquio ancora ricco di deviazioni segmentali rispetto al modello nativo, l'esercitazione favorisce il miglioramento dell'efficacia comunicativa, ma contribuisce solo parzialmente ad attenuare l'intensità del giudizio di forestierismo. Al contrario, con parlanti di livello più avanzato, con un piano segmentale maggiormente accurato, è stato possibile apprezzare l'effetto dell'esercitazione sia in termini di efficacia comunicativa, sia di riduzione dell'accento straniero.

Prossimi sviluppi della ricerca riguarderanno l'estensione delle indagini sperimentali sull'efficacia delle tecniche di imitazione e autoimitazione ad apprendenti di italiano con altre lingue materne e con livelli di competenza linguistica differenziati.

Bibliografia

- BISSIRI M.P. - PFITZINGER H.R. - TILLMANN H.G. (2006), Lexical stress training of German compounds for Italian speakers by means of resynthesis and emphasis, in WARREN P. - WATSON C.I. (eds.), *Proceedings of the 11th Australian International Conference on Speech Science & Technology*, University of Auckland, New Zealand: 24-29.
- BOERSMA P. - WEENINK P. (2007), *Praat: doing phonetics by computer* (Version 5.0), University of Amsterdam, The Netherlands, Disponibile in: <http://www.praat.org> (data ultimo accesso 03 settembre 2015).
- CHARPENTIER F. - MOULINES E. (1989), Pitch-synchronous waveform processing techniques for text-to-speech synthesis using diphones, in *Proceedings of the First European Conference on Speech Communication and Technology – Eurospeech*, European Speech Communication Association, Paris: 2013-2019.
- CHUN D.M. (2013), Computer-assisted pronunciation teaching, in CHAPELLE C.A. (ed.), *The Encyclopedia of Applied Linguistics*, Wiley-Blackwell, Oxford: 823-834.
- COUNCIL OF EUROPE (2001), *Common European Framework of Reference for Languages: Learning, Teaching, Assessment*, Cambridge University Press, Cambridge.
- DE MEO A. - VITALE M. - PETTORINO M. - CUTUGNO F. - ORIGLIA A. (2013), Imitation/self-imitation in computer-assisted prosody training for Chinese learners of L2 Italian, in LEVIS J. - LE VELLE K. (eds.), *Proceedings of the 4th Pronunciation in Second Language Learning and Teaching Conference*, August 2012, Ames, IA, Iowa State University: 90-100.

- ESKENAZI M. (2009), An overview of spoken language technology for education, in *Speech Communication* 51(10): 832-844.
- FELPS D. - BORTFELD H. - GUTIERREZ-OSUNA R. (2009), Foreign accent conversion in computer assisted pronunciation training, in *Speech Communication* 51(10): 920-932.
- JILKA M. - MÖHLER G. (1998), Intonational foreign accent: speech technology and foreign language teaching, in *Proceedings of ESCA Workshop on Speech Technology in Language Learning*: 115-118.
- LEVIS J. (2007), Computer technology in teaching and researching pronunciation, in *Annual Review of Applied Linguistics* 27: 184-202.
- MARTIN P. (2012), Automatic prosodic comparison between model and imitation sentences in a second language teaching computerized environment, in DE MEO A. - PETTORINO M. (eds.), *Prosodic and rhythmic aspects of L2 acquisition. The case of Italian*, Cambridge Scholars Publishing, Newcastle upon Tyne: 263-276.
- MUNRO M. - DERWING T. (1995), Processing time, accent, and comprehensibility in the perception of native and foreign-accented speech, in *Language and Speech*, 38: 289-306.
- NAGANO K. - OZAWA K. (1990), English speech training using voice conversion, in *Proceedings of the 1st International Conference on Spoken Language Processing (ICSLP 90)*, Kobe, Japan: 1169-1172.
- PEABODY M. - SENEFF S. (2006), Towards automatic tone correction in non-native mandarin, in HUO Q. - MA B. - CHANG E-S. - LI H. (eds.), *Proceedings of the 5th international conference on Chinese spoken language processing*, Kent Ridge, Singapore: 602-603.
- PETTORINO M. - VITALE M. (2012), Transplanting native prosody into second language speech, in BUSÀ M.G. - STELLA A. (eds.), *Methodological Perspectives on Second Language Prosody. Papers from ML2P*, CLEUP, Padova: 11-16.
- PROBST K. - KE Y. - ESKENAZI M. (2002), Enhancing foreign language tutors - In search of the golden speaker, in *Speech Communication* 37(3-4): 161-173.
- SUNDSTRÖM A. (1998), Automatic prosody modification as a means for foreign language pronunciation training, in *Proceedings of ISCA Workshop on Speech Technology in Language Learning (STILL 98)*, Marholmen, Sweden: 49-52.
- TANG M. - WANG C. - SENEFF S. (2001), Voice transformations: From speech synthesis to mammalian vocalizations, in *Proceedings of Eurospeech 2001*, Aalborg, Denmark: 357-360.
- YOON K. (2007), Imposing native speakers' prosody on non-native speakers' utterances: The technique of cloning prosody, in *Journal of the Modern British & American Language & Literature* 25(4): 197-215.

PATRIZIA SORIANELLO¹

Indici predittivi probabilistici: il caso della geminazione consonantica in italiano L2

The acquisition of geminate consonants by second-language learners of Italian has been investigated using an exemplar-based approach called Analogical Modeling (AM). To this purpose, a database of 2000 lexical items containing a geminate phoneme has been collected. Each long phoneme has been specified in terms of articulatory, prosodic and lexical variables. The AM paradigm allows us to capture regularities in an apparently irregular process. AM generates all possible combinations, providing ranked outcomes. The analogical simulations demonstrate that the variables considered have not the same predictive force in inducing germination.

1. *La geminazione in italiano L2*

L'acquisizione del contrasto fonologico tra suoni lunghi e brevi è motivo di grande difficoltà per gli apprendenti stranieri che non possiedono tale tratto nella loro lingua materna. Nell'interlingua le consonanti degeminate sono pertanto numerose, e sono spesso causa di errori grammaticali o lessicali (Enomoto, 1996; Han, 1992; Harada, 2006; Kabak *et al.*, 2011; Kunnari *et al.*, 2001; MacAllister *et al.*, 2002; Mah - Archibald, 2003). Con riferimento all'italiano L2 (da ora it-L2), l'acquisizione delle geminate manifesta un elevato grado di variabilità, ma non sembra configurarsi come un processo causale. Dalle ricerche svolte, si desume che in it-L2 la geminazione ha inizio dai bisillabi più frequenti, preposizioni, forme verbali e modificatori, prediligendo inoltre la posizione postonica (Giannini - Costamagna, 1997, 1998; Giannini, 2003). A ciò si aggiunge l'azione della struttura ritmico-prosodica, evidenti fenomeni di *transfer* sono stati rilevati in apprendenti cinesi (Costamagna *et al.*, 2014). La realizzazione delle geminate sembra inoltre sensibile alla natura articolatoria del fonema, essendo meglio rappresentata tra le ostruenti sorde che tra le sonoranti (Soriano, 2009; 2014). Decisamente più complesso è però comprendere la dinamica dei fattori che predispongono il contesto 'ottimale' di questo processo, così come assegnare il giusto peso alle variabili in gioco. Per fare luce sulla questione, nella presente ricerca l'acquisizione della geminazione sarà indagata adottando una prospettiva di tipo probabilistico, mediante l'applicazione del modello *Analogical Modeling* (AM).

¹ Università degli Studi di Bari.

2. Un modello computazionale: *l'Analogical modeling*

L'*Analogical Modeling* (AM) rientra tra i modelli *exemplar-based*; si tratta di approcci, ad impostazione funzionalista, aventi come sfondo teorico la psicologia cognitiva. Essi prevedono che le esperienze compiute, gli stimoli o le informazioni nuove, siano di volta in volta comparati con gli elementi più simili già esperiti in passato per essere poi memorizzati, previo un processo di indicizzazione categoriale. Nel corso degli anni, questi modelli, tra cui ricordiamo almeno la teoria degli elementi, sono stati applicati, sotto forma di procedura computazionale, anche in ambito linguistico, soprattutto a livello fonetico (Johnson, 1997; Pierrehumbert, 2003, per una rassegna sul funzionamento di tali modelli e delle loro criticità rinviamo a Vietti, 2013). AM riesce a fare predizioni sul comportamento di un determinato esito linguistico, senza definire regole. Il modello, prima denominato *Analogical Modeling of Language* (AML), è stato rivisto più volte. Si tratta di un paradigma che associa i presupposti teorici elaborati da Royal Jon Skousen (1989; 1992) con un'implementazione computazionale (Skousen, 1989; Skousen *et al.*, 2002; Eddington - Lonsdale, 2007; per una rassegna si veda anche il sito web <http://humanities.byu.edu/am/> contenente anche le istruzioni sul *download* e sul funzionamento del programma). AM parte dal presupposto che ogni esperienza linguistica passata venga immagazzinata nel lessico mentale dell'uomo; questa traccia fungerà da *input* per i comportamenti linguistici futuri. Sul versante computazionale, AM è un sistema di calcolo elaborato in linguaggio Perl (<http://www.cpan.org>, per un approfondimento si veda Hammond, 2003) che simula il comportamento linguistico assegnando indici di realizzazione probabilistica. Ciò è possibile grazie al principio di analogia, un indice che computa il grado di somiglianza/dissomiglianza tra forme. La condizione necessaria affinché ciò possa avvenire è che sia disponibile un archivio di entrate linguistiche in cui sia rappresentato il processo che si intende indagare. Esso simula l'archivio mentale che ogni soggetto ha della propria lingua. L'archivio va costruito seguendo determinati schemi, i valori di ogni variabile (es. genere, numero, classe lessicale ecc.) devono essere codificati attraverso l'assegnazione di un indice numerico o alfabetico, ovvero un'etichetta categoriale. Ogni occorrenza linguistica è infatti specificata da un set di variabili, ridotte ad una stringa alfanumerica, e da un *outcome* già assegnato, ovvero l'elenco dei valori associabili ad una variabile, ad es. in italiano la categoria del Numero genera solo due valori: singolare e plurale. L'obiettivo di AM è quello di predire l'esito linguistico di un elemento per un dato contesto, sulla base delle ricorrenze ad esso più somiglianti trovate nell'archivio. Il contesto selezionato (*given context*), cioè l'insieme delle variabili che rappresentano l'informazione linguistica dell'*item* in esame, è quindi fondamentale.

La probabilità che una data occorrenza possa essere forma analogica di un'altra dipende da più fattori: 1) la somiglianza della forma al contesto dato (*proximity*), 2) la presenza, vicino al contesto dato, di altre forme simili, ma dal comportamento differente, in questo caso la forma non sarà scelta come candidato analogico (*heterogeneity*), 3) la presenza nell'archivio di molte forme linguistiche che si comportano in modo simile a quella da interrogare; tali forme possono influenzare la predizione,

aumentando la possibilità di essere selezionate (*gang effect*). Ogni volta che l'archivio è interrogato, le entrate sono suddivise in sottocontesti, cioè gruppi i cui *items* condividono il maggior numero di tratti con l'*input* selezionato.

Nel corso degli anni AM, grazie alle sue potenzialità esplicative, è stato variamente applicato per verificare il comportamento di svariati processi, non solo di ordine linguistico. Tra gli studi effettuati ricordiamo almeno quelli sulla formazione del *past-tense* in inglese (Derwing - Skousen, 1994) o dello *imperfekti* in finnico (Skousen, 1989), sul posizionamento dell'accento lessicale in spagnolo (Eddington, 2000), sulla selezione dell'articolo indeterminativo (Skousen, 1989) o del genere (Eddington, 2002)². Più recentemente, AM è stato impiegato anche in contesto acquisizionale (Lonsdale - Matsushita, 2013).

3. *La ricerca*

In questo studio l'acquisizione dei fonemi geminati da parte di apprendenti di it-L2 è stata analizzata nell'ambito del modello AM. Ciò ha permesso di isolare i fattori che incidono di più sull'acquisizione del fonema lungo e di cogliere delle regolarità all'interno di un processo apparentemente irregolare. Gli indici probabilistici ottenuti tramite l'interrogazione dell'algoritmo sono stati, quando possibile, confrontati con i dati reali estratti dal nostro *database*.

3.1 I soggetti

Sono stati considerati quattro diversi gruppi di apprendenti, accomunati dal fatto di non possedere nella loro lingua materna fonemi geminati. Si tratta per la precisione di sei spagnoli, tre tedeschi, cinque inglesi e cinque ecuadoriani. Gli apprendenti sono studenti Erasmus di livello A2/B2 e di età compresa tra 22 e 25 anni, ad eccezione degli ecuadoriani che rientrano in un programma di scambio internazionale.

3.2. I materiali

Gli apprendenti hanno partecipato a una o più sessioni di registrazione; queste ultime sono state svolte in un ambiente silenzioso ed acquisite mediante un *Tascam DR-07* (44 KHz, 32 bit, .wav file). I soggetti sono stati informati di essere audio-registrati, ma non delle finalità della ricerca. Ogni apprendente ha compilato una scheda sociolinguistica, rilasciando inoltre consenso informato sul trattamento dei propri dati. Il materiale linguistico consta di conversazioni dialogiche bidirezionali su esperienze di vita comune. Un campione di 1966 consonanti geminate è stato analizzato uditivamente ed acusticamente.

² Con riferimento alla lingua italiana, il modello AM è stato adottato per lo studio della gorgia toscana (Soriano *et al.*, 2005).

3.3. L'archivio

Il corpus è formato da fonemi geminati primari. I lessemi contenenti le entrate consonantiche sono stati inseriti in un foglio elettronico Excel e manualmente classificati per: 1) identificazione del parlante, b) identificazione del corpus, c) identificazione del fonema, d) contesto prosodico, e) identificazione del lessema.

I primi indici mirano all'identificazione sociolinguistica dell'apprendente (provenienza, età ecc.). Nel secondo gruppo sono confluite le informazioni relative al corpus da cui è stata estratta la parola *target* e la sigla del turno dialogico. Il terzo gruppo di variabili classifica il fonema lungo per: 1) Modo di articolazione (MA), 2) Luogo di articolazione (LA), 3) Sonorità (SO), 4) Classe di suono (CS). Queste variabili comprendono a loro volta i seguenti indici: MA (O: occlusiva, F: fricativa, A: affricata, N: nasale, L: laterale, V: vibrante); LA (B: bilabiale, L: labiodentale, D: dentale, P: palatale, V: velare), SO (U: sorda, V: sonora), CS (O: ostruenti, S: sonoranti).

Con riferimento al contesto prosodico, la geminata è stata indicizzata rispetto all'accento lessicale della parola³. Questa codifica non può tuttavia prescindere dalla sillaba (SIL). I suoni geminati hanno un intrinseco statuto eterosillabico: la prima parte del fonema costituisce una Coda sillabica, mentre la seconda parte l'Attacco della sillaba successiva, es. *affan.no*. Riprendendo la metodologia adottata in Payne (2005), abbiamo previsto tre diverse codifiche, precisamente ATONA (A): la geminata è preceduta e seguita da una vocale atona (es. *av.vo.cà.to*), PRETONICA (R): la geminata ricorre immediatamente prima dell'accento lessicale, la seconda parte del fonema rappresenta quindi l'onset della sillaba tonica (es. *ot.tò.bre*), POSTONICA (P): il fonema ricorre immediatamente dopo l'accento lessicale, in questo caso la prima parte della geminata rappresenta la coda della sillaba tonica, (es. *a.dès.so, pic.co.lo*).

Per avere un quadro più esaustivo, abbiamo calcolato anche la distanza che intercorreva tra la geminata e l'accento lessicale (Distanza Accento Lessicale, DAL), questo indice, espresso in sillabe, oscilla tra 0 (es. *settembre*) e 3 (*possibilità*).

L'ultimo gruppo di variabili mira a classificare il lessema. In questa direzione abbiamo considerato la sua durata in sillabe (Lunghezza di Parola, LUP), gli indici oscillano tra 2 (es. *bello*) e 6 (*assolutamente*), ricoprendo tutte le possibilità intermedie, la classe di parola (CL) e il suo indice di frequenza (FRE).

Per la codifica della natura morfologica della parola contenente la geminata, abbiamo assunto le linee guida adottate in CoLFIS (Bertinetto *et al.*, 2005). Le classi lessicali utilizzate sono riportate qui di seguito, tra parentesi l'etichetta apposta nel *database*:

Articoli (D), in italiano i fonemi lunghi non ricorrono all'interno di questa categoria, con la sola eccezione dell'articolo partitivo (es. vorrei *delle* mele) che pertanto è l'unica ricorrenza indicizzata.

Aggettivi (G): comprende gli aggettivi qualificativi (anche di grado superlativo o comparativo), i participi passati usati con valore aggettivale e gli aggettivi determinativi (es. dimostrativi, indefiniti ecc.).

³ Le lingue materne degli apprendenti sono tutte ad accento libero.

Numerali (U): nel corpus compaiono solo numeri cardinali (es. *otto*), i numeri ordinali contenenti una geminata (es. *ottantesimo*) sono assenti. Non si è reso necessario quindi distinguere tra uso aggettivale e sostantivale dei lemmi.

Nomi (N): comprende i nomi comuni di persona, animali, oggetti e concetti.

Pronomi (R): questa classe è più ridotta rispetto a quanto previsto dalla grammatica tradizionale; le uniche categorie presenti sono i dimostrativi (es. *quello*) e gli indefiniti (es. *nessuno*), assenti sono gli altri sottotipi, le cui forme in italiano non presentano un fonema lungo⁴.

Nomi Propri (O): include i nomi propri, di strade, toponimi ecc. Nel corpus non compaiono cognomi, né nomi di prodotti o ditte.

Verbi (V): include le forme verbali espresse nei modi finiti o indefiniti.

Verbi ausiliari (X): include le ricorrenze dei verbi ausiliari, *avere* ed *essere* aventi valore grammaticale e non predicativo.

Preposizioni (P): include solo le preposizioni articolate, essendo le semplici prive di un fonema lungo, e alcune ricorrenze di preposizioni improprie (es. *sotto*).

Avverbi (A): include avverbi e locuzioni avverbiali. Nei casi di polifunzionalità di un lessema, la funzione grammaticale è stata assegnata previa considerazione del contesto.

Congiunzioni (C): valgono i criteri citati per la classificazione degli Avverbi. Tra le congiunzioni abbiamo inserito anche diversi segnali discorsivi, spesso semanticamente vuoti, come *insomma* o *allora*.

Nei nostri corpora non è rappresentata la classe delle Interiezioni; espressioni contenenti un fonema lungo, come ad es. *accidenti*, *mannaggia*, *evviva* sono assenti.

Il parametro (FRE) è stato inserito per verificare se le parole statisticamente più frequenti contenenti un fonema lungo costituissero in it-L2 un bersaglio preferenziale per la produzione di geminate, come formulato in una serie di studi sulla diffusione delle nuove forme lessicali (Bybee, 2002; Giannini, 2003; Phillips, 1984). L'assegnazione di un indice di frequenza ha sollevato alcune riflessioni di ordine metodologico ed ha imposto delle scelte. I dizionari di frequenza disponibili per l'italiano nativo sono diversi, per numero e per impostazione metodologica; tra questi citiamo almeno il *Vocabolario di Base* (VdB, De Mauro, 1980), il *Lessico dell'Italiano parlato* (LIP, De Mauro et al., 1993), il *Corpus e Lessico di Frequenza dell'Italiano Scritto* (CoLFIS, Bertinetto et al., 2005). Fra gli strumenti disponibili, gli indici di frequenza sono stati ricavati dalla consultazione del LIP, un lessico interamente incentrato sull'analisi di campioni orali. Nel LIP è fornito un rango d'uso per 490.000 lemmi, di questi 7213 sono presenti con almeno 3 occorrenze. Per il perseguimento delle nostre finalità, l'utilizzo di un numero così elevato di indici di frequenza rischiava di compromettere la significatività del dato. Per questo motivo, i ranghi d'uso del LIP sono stati ricondotti a sole 6 classi, come riportato qui di seguito.

⁴ Costituiscono eccezione i pronomi personali *esso*, *ella* i quali però, essendo fortemente marcati in diafasia, non ricorrono nel parlato degli apprendenti.

Classe 1: lemmi con rango d'uso compreso tra 1 e 500, si tratta delle parole dell'italiano che hanno la più alta incidenza statistica; Classe 2: lemmi con frequenza da 501 a 1000, classe 3: lemmi con frequenza da 1001 a 2000, classe 4: lemmi con frequenza da 2001 a 4000, classe 5: lemmi con frequenza 4001 a 7213, Classe 6: lemmi la cui frequenza è superiore a 7213.

3.4 Le variabili fonetiche

La codifica fonetica assegnata a ciascun fonema geminato costituisce un anello fondamentale della ricerca poiché rappresenta l'*outcome* che AM assegnerà in modo probabilistico alle nuove forme interrogate sotto forma di punteggio percentuale. I valori dell'*outcome*, o esito linguistico, sono assegnati dal ricercatore, a seconda della finalità della ricerca. Nel nostro caso il ventaglio degli esiti associato all'*outcome* include solo tre codifiche: suono geminato, suono degeminato e altre realizzazioni (*infra*), essendo queste varianti le sole riscontrate nell'analisi fonetica del campione. La valutazione fonetica dei fonemi bersaglio è stata condotta su base uditiva ed acustica. L'accertamento uditivo ha previsto l'ascolto delle parole contenenti il fonema lungo, le parole non sono state valutate in isolamento, ma all'interno del loro contesto frasale. L'accertamento spettro-acustico è stato effettuato mediante il *software Praat* (Boersma - Weenink 2007, versione 5.0.16) ed ha previsto il rilievo della durata della consonante geminata (C) e quella della vocale (V) precedente. Al fine di poter stabilire delle utili comparazioni riguardo alla natura (lunga o degeminata) dei segmenti, abbiamo utilizzato due sottocampioni formati rispettivamente da fonemi geminati e scempi estratti da conversazioni di it-L1, e da fonemi scempi estratti dai corpora di it-L2. I dati acustici relativi a questi parametri sono discussi in Sorianello (2014), ad eccezione di quelli relativi agli ecuadoriani.

I fonemi lunghi sono stati classificati nelle seguenti categorie, tra parentesi l'indicazione della codifica assegnata:

1. segmenti geminati (G): si tratta di consonanti la cui realizzazione fonetica non diverge dal *target* fonologico. Questi suoni sono approssimabili per resa uditiva e per durata ai fonemi lunghi prodotti dai nativi;
2. segmenti degeminati (D): l'identità fonetica non coincide con quella del fonema bersaglio. La durata della C è praticamente dimezzata ed è comparabile a quella dei corrispettivi fonemi scempi prodotti in it-L2 dagli stessi apprendenti o prodotti in it-L1 dai soggetti nativi;
3. altre varianti (A): questo raggruppamento può considerarsi un sottotipo della variante degeminata. Non è insolito che la pronuncia di una geminata subisca, oltre ad una riduzione temporale, anche uno spostamento di Modo o di Luogo di articolazione, come la pronuncia approssimante o fricativa di occlusive ed affricate, es. *dobbiamo* > [de'vja:mo], *leggere* > ['le:jere].

4. *La predizione analogica*

AM possiede diversi vantaggi. Innanzitutto, la sua spiccata flessibilità rende possibile combinare le variabili in modo diverso, a seconda dell'azione che si intende di volta in volta valutare. Tutte le variabili assegnate sono inoltre potenzialmente uguali, non essendo specificate in anticipo per grado di salienza. Ciò consente di indagare anche quelle combinazioni contestuali (fonologiche, lessicali o altro) non ricorrenti nell'archivio, ma potenzialmente possibili nella lingua in esame. L'archivio è stato interrogato in modo mirato, incrociando variabili di vario tipo. Durante l'interrogazione, AM ha vagliato tutte le stringhe aventi in comune almeno un elemento della sequenza considerata, computando indici percentuali di probabilità. Sostanzialmente possiamo distinguere tre raggruppamenti, a seconda della prevalenza dei tratti implicati; le combinazioni testate sono le seguenti:

- indici articolatori: 1) MA-LA-SO-CS, 2) MA-LA-CS, 3) LA-SO (§ 4.1);
- indici lessicali: 1) CL-FRE, 2) CS- CL-FRE, 3) CS-SIL-FRE (§ 4.2);
- indici prosodico-accentuali: 1) LUP-SIL-DAL, 2) LUP-SIL-DAL-FRE 3) CS-LUP-SIL-DAL (§ 4.3).

Nella tab. 1 riportiamo a scopo riassuntivo l'elenco di tutte le variabili considerate e dei loro valori.

Tabella 1 - *Elenco riassuntivo delle variabili considerate, delle relative sigle e dei valori ad esse associati*

<i>Variabile</i>	<i>Sigla</i>	<i>Valori</i>
Modo di Articolazione	MA	O, F, A, N, L, V,
Luogo di Articolazione	LA	B, L, D, P, V
Sonorità	SO	U, V
Classe Suono	CS	O, S
Sillaba	SIL	A, R, P
Distanza Accento Lessicale (sill)	DAL	0, 1, 2, 3,
Lunghezza di Parola (sill)	LUP	2, 3, 4, 5, 6
Classe Lessicale	CL	A, G, U, N, R, O, V, X, A, C
Indice di Frequenza parola	FRE	1, 2, 3, 4, 5, 6

4.1 Le predizioni articolatorie

Dall'insieme degli esiti un dato appare prevalente: la possibilità che un fono sia realizzato geminato dagli apprendenti è fortemente condizionata dal Modo di articolazione (MA) e dalla sonorità del fonema (SO).

Interrogando l'archivio si evince come la massima resa geminata venga assegnata ad FDUO (83%, /ss/), la più bassa a FLVO (9%, /vv/). Complessivamente, i contesti articolatori ottimali risultano quelli in cui è presente un'ostruente sorda, ad es.

APUO (76%, /ttf/)⁵, ODUO, FLUO, (69%, /ff/), (66%, /tt/). I valori scendono considerevolmente se la stessa stringa definisce la corrispettiva consonante sonora. Nel nostro campione non vi sono ricorrenze di /gg/ (OVVO), è tuttavia possibile fornire una simulazione anche per questo sottocontesto, l'algoritmo infatti, procedendo per analogia, estrae dall'intero archivio le stringhe più simili a /gg/, quelle ad es. contenenti v (sonoro) e o (occlusiva), e computando un punteggio di geminazione pari al 49%.

Al fine di isolare l'azione predittiva delle singole codifiche, abbiamo limitato la stringa da interrogare al solo sottocontesto MA-LA-CS, eliminando così l'informazione relativa alla sonorità. Le predizioni analogiche risultano allineate alle precedenti; il contesto più robusto per presenza di geminazione è la combinazione FDO, Fricativa-dentale-Ostruente (83%) seguito da ODO (occlusiva-dentale-ostruente, 65%) e da APO (affricata-palatale-ostruente, 63%); gli indici di geminazione più bassi si rilevano per LDS e VDS, rispettivamente laterale (22%) o vibrante-dentale (25%). L'assenza della codifica relativa alla sonorità produce un sensibile effetto, almeno per alcuni sottocontesti, come ODO (/tt/, /dd/), FLO (/ff/, /vv/), OBO (/pp/, /bb/) i quali, privati di tale informazione, vedono abbassarsi il proprio potenziale di corretta pronuncia geminata, riproducendo una sorta di media tra valore della consonante sorda e quello della sonora.

Diverso il quadro che emerge interrogando solo le codifiche LA-SO. In questo set, AM lavorando su una stringa privata sia della variabile MA che di quella CS, rimette in gioco il valore predittivo del tratto di sonorità. La geminazione è favorita nei sottocontesti analogici DU e PU (75%), rispettivamente fonema dentale e palatale sordo, ma fortemente preclusa nei set LV (8%), BV (21%) e DV (24%), ovvero fonema labiodentale, bilabiale e dentale sonoro.

Le simulazioni analogiche ottenute sono in linea con i dati, questa volta effettivi e non simulati, che si ottengono dal nostro *database*. Lo scenario che si delinea è una decisa polarizzazione degli esiti: da un lato si pongono le ostruenti sorde (G: 71%), dall'altro le sonoranti (G: 25%), che al contrario ricevono il più basso tasso di allungamento temporale, precisamente in ordine decrescente: /mm/ (35%), /nn/ (29%), /rr/ (25%), /ll/ (22%). Un commento a parte va fatto per le ostruenti sonore il cui tasso di geminazione è pari al 29%. Questi fonemi sembrano particolarmente inclini alla degeminazione, precisamente /dd/ (58%), /bb/ (49%) e /vv/ (74%), pur trattandosi di ostruenti, ma sono anche i segmenti che realizzano il più alto numero di ALTRE varianti (rispettivamente 8% 36% e 17%), essendo frequentemente approssimanti (già Sorianello 2009, 2014).

⁵ Ovvero AFFRICATA-PALATALE-SORDA-OSTRUENTE, per la lettura delle stringhe si faccia riferimento a quanto riportato nei §§ precedenti e nella tab. 1.

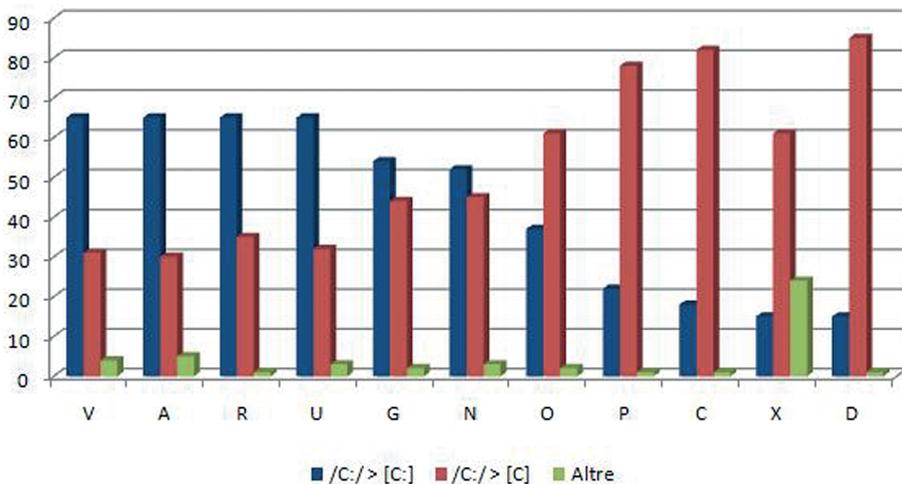
4.2 Le predizioni lessicali

Il rapporto tra geminazione e classe lessicale della parola fornisce un'ulteriore chiave di lettura degli esiti. In questa discussione, consideriamo prima i dati reali estratti dal campione, successivamente i risultati delle predizioni computate da AM.

4.2.1 Le predizioni lessicali: dati reali

Nell'archivio raccolto, il fonema geminato non raggiunge le medesime percentuali all'interno delle varie categorie lessicali. Come è evidente dall'istogramma riprodotto nella fig. 1, in it-L2 i valori più alti di geminazione si rilevano in Verbi, (65%, es. *posso, vanno*) Pronomi (65%, es. *tutto*), Numerali (65%, es. *otto*) e Avverbi (65%, es. *spesso*), per converso lo scempiamento del fonema lungo prevale nelle classi relative ad Articoli (85%), Congiunzioni (82%), e Preposizioni (78%).

Figura 1 - Valori percentuali reali di realizzazione geminata /C:/ > [C:], degeminata /C:/ > [C] e Altra, per classi lessicali. D: articoli, G: aggettivi, U: numerali, N: nomi, O: nomi propri, R: pronomi, V: verbi, X: ausiliari, P: preposizioni, A: avverbi, C: congiunzioni



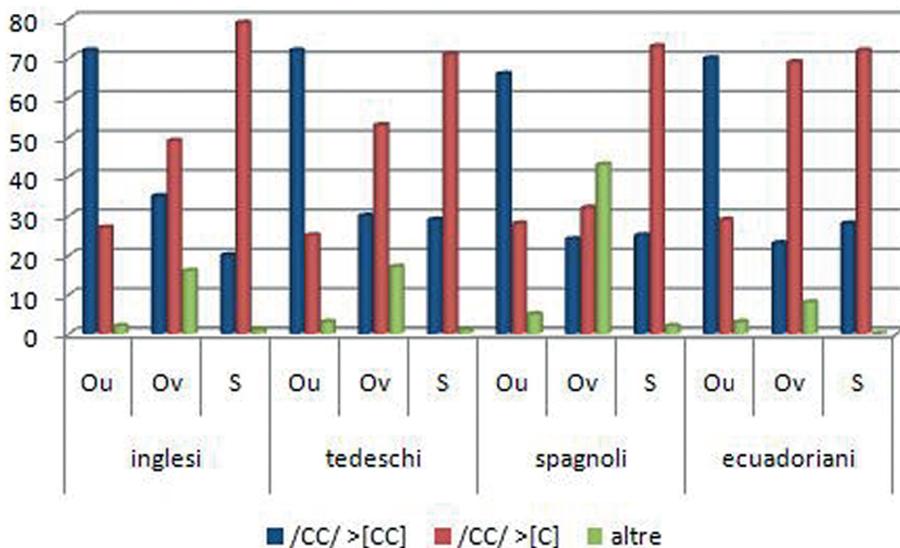
La geminazione è più rappresentata nelle parti del discorso dotate di alto carico semantico, prodotte anche per questo con un maggiore controllo articolatorio. Per contro, Preposizioni, Ausiliari, Articoli e Congiunzioni sono più inclini a subire processi di erosione articolatoria o, come in questo caso, di riduzione temporale. Le categorie grammaticali chiuse sono notoriamente caratterizzate da elevata fragilità morfologica; ad un ridotto materiale fonico si associa quasi sempre anche uno scarso contenuto semantico e una debole autonomia prosodica (già Simone, 1988).

I risultati mettono in luce la presenza di fattori interagenti. Scorporando i dati, si scopre che i valori più alti di degeminazione sono imputabili in modo sistematico ad alcune consonanti, tipicamente le sonoranti. Occorre tuttavia notare che i fonemi non sono ugualmente rappresentati nelle diverse parti del discorso. Negli Articoli e nelle Preposizioni il fonema implicato è sempre /ll/, negli Ausiliari tro-

viamo fondamentalmente solo /nn/ e /bb/ (*hanno, abbiamo, sarebbe*), mentre tra le Congiunzioni, oltre alla forma *allora* che è dominante, troviamo *infatti* e *siccome*.

In tutti i gruppi di apprendenti, le ostruenti sorde concorrono, in modo maggioritario, ad innalzare i valori della variante geminata. Per fare un esempio, la percentuale di degeminazione che si riscontra nel Nome è pari al 45%, in oltre la metà di questi casi però il fonema implicato è una sonorante. A prescindere dalla classe lessicale, il 74% della resa degeminata si deve allo scempiamento di una sonorante, il 50% allo scempiamento di una ostruente sonora e solo il 28% alla riduzione temporale di un'ostruente sorda. Questa tendenza è costante, dai dati non emerge infatti un significativo condizionamento imputabile alla L1 degli apprendenti, si veda quanto riportato nella fig. 2.

Figura 2 - Valori percentuali reali di geminazione, degeminazione e altre varianti, ripartite per Classe di Suono (Ov: Ostruente sorda, Os: Ostruente sonora, S: Sonorante) e per gruppo di apprendenti



4.2.2 Le predizioni lessicali: le simulazioni

La classe lessicale (CL) è stata oggetto di un'interrogazione specifica, in combinazione con la variabile frequenza (CL-FRE), con una variabile articolatoria (CS-FRE-CL) e con una accentuale (CS-SIL-FRE).

Un dato inatteso riguarda l'estrazione degli indici probabilistici relativi alla frequenza (stringa CL-FRE). L'ipotesi di partenza prevedeva percentuali maggiori di geminazione tra i lessemi statisticamente più frequenti. Le predizioni analogiche indicano invece come FRE sia un parametro pressoché ininfluenza, in quanto, da solo, non riesce a fornire predizioni differenziate. La probabilità che un fonema lungo sia ben realizzato dagli apprendenti è infatti sostanzialmente uguale, tanto nelle parole ad alta frequenza, quanto in quelle meno frequenti. Il risultato non subisce sostanziali variazioni se l'indice FRE è integrato con quello relativo alla CL. Assumiamo

come esempio quanto avviene all'interno delle categorie più rappresentate, come Nome e Verbo, rispettivamente 38,4% e 15,6% sul totale delle CL. Nel nostro archivio, N ha ricevuto tutte le codifiche di frequenza contemplate, da 1 a 6. Se il ruolo della frequenza fosse stato determinante ai fini della corretta produzione di consonanti lunghe, allora N con indice frequenziale 1, cioè massima frequenza lessicale, avrebbe dovuto mostrare, nel parlato degli apprendenti, tassi più elevati di geminazione rispetto a N con codifiche di frequenza, ad es. 5 o 6. Gli indici probabilistici sono controversi, si veda quanto riportato nella tab. 2.

Tabella 2 - *Indici predittivi (in %) CL-FRE relativi al Nome*

<i>Codifica</i>	<i>Geminata</i>	<i>Degeminata</i>	<i>Altre</i>
N1	53	42	5
N2	50	47	3
N3	51	45	4
N4	52	45	3
N5	53	44	3
N6	54	43	3

Le probabilità che all'interno del Nome un fonema dell'it-L1 venga realizzato geminato in it-L2, stando alle predizioni analogiche fornite da AM, sono pressoché le stesse, a prescindere dalla fascia di frequenza a cui appartiene. Lievemente diverso il quadro che risulta analizzando la categoria del Verbo (tab. 3). Paradossalmente, la codifica v4 produce una predizione di geminazione più alta, mentre la più bassa si ha per i Verbi più frequenti, quasi come se l'apprendente prestasse più attenzione alla pronuncia di parole poco comuni, ma beninteso non rare, visto che anche la fascia FRE6 include parole del lessico ad alta disponibilità (De Mauro *et al.*, 1993).

Tabella 3 - *Indici predittivi (in %) CL-FRE relativi al Verbo*

<i>Codifica</i>	<i>Geminata</i>	<i>Degeminata</i>	<i>Altre</i>
v1	48	49	3
v2	63	33	4
v3	63	34	3
v4	64	32	3
v5	63	33	4
v6	61	35	4

Né va ignorato il fatto che il campione degli *items* è sbilanciato; il sottogruppo di parole con codifica FRE1 è decisamente più numeroso rispetto agli altri (61,5% sul totale). L'indice FRE1 costituisce quindi una variabile dominante all'interno dell'archivio che finisce per influenzare la predizione, creando un forte effetto analogico.

Una parziale riprova di quanto detto proviene dai dati reali presenti nel *database*, i verbi più frequenti (V1) mostrano infatti la corretta geminazione del fonema nel 68% dei casi, i valori tuttavia non decrescono in modo lineare, essendo V5 pari a 40% e V6 a 50%. Alla luce dei dati, è ragionevole pensare che né FRE né CL siano predittori forti di geminazione; CL può solo aumentare o decrescere la probabilità che una parola possa essere o no realizzata con un fonema lungo, ma non determinarla

La stringa è stata allora integrata con la codifica CS. Questa sequenza (CS-CL-FRE) ha permesso di verificare in che modo tre variabili diverse, classe suono, frequenza e classe lessicale, agissero nel favorire la comparsa del suono geminato in it-L2. Abbiamo quindi comparato gli indici probabilistici per sequenze identiche, ad eccezione della codifica per la quale si intendeva valutarne l'effetto. A tal fine, saranno discusse le classi lessicali statisticamente più presenti nel campione, ma anche più eterogenee per fonemi lunghi rappresentati, ovvero, Nomi, Verbi e Aggettivi (tab. 4-5-6).

Il parametro FRE incide poco sulla lunghezza del fono *target*, le percentuali non sono infatti statisticamente significative, ad es. la stringa ON2, ovvero Ostruente-Nome con indice di Frequenza 2 (es. *gatto, caffè*) nell'eloquio di un apprendente di it-L2 ha la stessa possibilità di ricorrere come geminata rispetto alla stringa ON6, ovvero Ostruente-Nome con indice di Frequenza 6 (es. *mattoni, apprendistato*), rispettivamente 67% e 64%.

Tabella 4 - Indici probabilistici (in %) CS-CL-FRE relativi al Nome

Codifica	Geminata	Degeminata	Altre	Codifica	Geminata	Degeminata	Altre
ON1	64	35	1	SN1	27	73	–
ON2	67	25	8	SN2	35	65	–
ON3	68	32	–	SN3	30	70	–
ON4	*63	37	–	SN4	42	58	–
ON5	*64	29	7	SN5	35	65	–
ON6	*64	29	7	SN6	23	77	–

Tabella 5 - Indici probabilistici (in %) CS-CL-FRE relativi al Verbo

Codifica	Geminata	Degeminata	Altre	Codifica	Geminata	Degeminata	Altre
OV1	77	19	4	SV1	19	76	5
OV2	65	30	5	SV2	19	76	5
OV3	*80	20	–	SV3	50	50	–
OV4	*67	33	–	SV4	19	76	5
OV5	*40	60	–	SV5	25	74	1
OV6	*100	–	–	SV6*	100	–	–

Tabella 6 - *Indici probabilistici (in %) CS-CL-FRE relativi all'Aggettivo*

<i>Codifica</i>	<i>Geminata</i>	<i>Degeminata</i>	<i>Altre</i>	<i>Codifica</i>	<i>Geminata</i>	<i>Degeminata</i>	<i>Altre</i>
OG1	71	27	2	SG1	10	90	–
OG2	65	30	5	SG2	33	60	7
OG3	71	26	3	SG3	33	60	7
OG4	67	33	–	SG4	50	50	–
OG5	80	13	7	SG5	37	63	–
OG6	65	30	5	SG6	100	–	–

Con riferimento alla CL, se la parola è molto frequente (FRE1) la percentuale di geminazione più elevata si riscontra per Verbo (77%) e Avverbio (78%) ma solo se il fonema è un'Ostruente, diametralmente opposti i valori probabilistici ottenuti se il fonema coincide con un suono Sonorante, Avverbio 20%, Verbo 19%, Aggettivo 10%. Questa tendenza appare sostanzialmente confermata se la parola è poco frequente (FRE6), ad es. O6N (64%), S6N (23%), S6V (23%), S6O (33%).

L'indicazione della CS prevale sugli altri indici, la CL della parola non è determinante ai fini della corretta produzione del fonema lungo, ma lo è invece molto di più la CS a cui appartiene, come già rinvenuto anche nei dati reali.

Se la stringa è integrata con l'informazione sull'accento (CS-SIL-FRE), i valori più alti di geminazione sono ricavati se la stringa contiene O (ostruente), soprattutto se associata al valore accentuale P, (sillaba postonica). AM predice che il mancato raggiungimento del *target* articolatorio sia più frequente in OA-FRE rispetto a OR-FRE e OP-FRE. L'indice di frequenza del lessema appare trascurabile, infatti lo *score* di geminazione in OP1 (ovvero ostruente lunga, in posizione postonica in lessema con indice di frequenza 1) essendo pari a 79% è di poco superiore a quello computato per OP4 (73%), oppure per OP6 (75%).

4.3 Le predizioni prosodico-accentuali: dati reali

Nei dati effettivi la geminazione prevale quando la seconda parte del fonema lungo costituisce l'attacco della sillaba postonica (57%, es. *fréd.do, ca.pèl.li*), mentre è minore negli altri contesti accentuali, precisamente 43% se PRETONICA e 40% se ATONA. Ciò può essere spiegato con il fatto che le sillabe atone sono di norma brevi, per questo più facilmente candidate a subire processi di riduzione articolatoria. Accento lessicale e lunghezza di parola sono correlati: le parole che compongono il nostro corpus sono prevalentemente bisillabi (41%) e trisillabi (37%), scarsamente rappresentati solo i lessemi con un numero di sillabe superiore a cinque (3%); le ricorrenze di una geminata in contesto atono sono, di conseguenza, numericamente inferiori. Complessivamente, il contesto che sembra favorire la realizzazione geminata è quello postonico, bisillabi e trisillabi mostrano a tale riguardo alte percentuali di geminazione, rispettivamente

51% e 67%. Al di là della lunghezza del lessema, la presenza della resa geminata è maggiore quando il fonema coincide con un'ostruente sorda, postonica: 59%, pretonica: 29%, atona: 11%. Le percentuali decrescono notevolmente quando, a parità di accento, la geminata è sonorante, precisamente postonica: 27%, pretonica: 3%, atona: 2%.

4.3.1 Le predizioni prosodico-accidentali: dati simulati

In questo paragrafo prenderemo in esame le simulazioni generate per la stringa LUP-SIL-DAL. Tra queste variabili esistono dei vincoli, la codifica R seleziona sempre DAL=0, la codifica P seleziona sempre DAL=1, mentre la codifica U ha come preconditione che la parola contenga almeno 4 sillabe, in questo caso DAL oscilla tra 1 e 3, a seconda del posizionamento dell'accento.

In rapporto all'accento, come nei dati reali, i valori di predizione calcolati da AM sono più alti se la geminata è postonica (P), medi se tonica (R), bassi se posta tra atone (A); i maggiori indici predittivi (67%) sono infatti assegnati quando il fonema geminato riceve la codifica P, a prescindere dai parametri LUP e DAL. Nella tab. 7 riportiamo gli indici probabilistici estratti per le stringhe testate.

Tabella 7 - *Indici probabilistici (in %) LUP-SIL-DAL*

Codifica	Geminata	Degeminata	Altra	Esempio
2P1	51	48	1	quello
3P1	67	30	3	adesso
4P1	57	42	1	tantissimo
2R0	50	50	-	città
3R0	41	49	10	abbiamo
4R0	45	51	4	grammatica
5R0	55	43	2	interessante
4A1	45	50	5	piccolini
4A2	45	51	4	difficoltà
5A1	50	50	-	professoressa
5A2	13	80	7	piccolissima
5A3	28	68	4	possibilità

C'è tuttavia una correlazione inattesa, se aumenta LUP, l'indice probabilistico di geminazione decresce, tale effetto risulta ancora più cospicuo se aumenta anche DAL (*infra*). La stringa è stata a questo punto integrata con l'informazione inerente alla frequenza (LUP-SIL-DAL-FRE) per verificare se, a parità di schema accentuale, gli *scores* di geminazione subissero un mutamento al variare dell'indice di frequenza. Per sondare tale aspetto abbiamo esplorato il comportamento delle consonanti lunghe codificate in stringhe differenziate solo per indice di

frequenza (da 1 a 6). Nella tab. 8 riportiamo solo un estratto dei numerosi dati probabilistici ottenuti per lessemi trisillabi.

Tabella 8 - *Indici probabilistici (%) LUP-SIL-DAL-FRE, le stringhe precedute da * sono fittizie*

<i>Codifica</i>	<i>Geminata</i>	<i>Degeminata</i>	<i>Altre</i>
3P11	52	48	1
3P12	67	29	4
3P13	66	31	3
3P14	67	30	3
3P15	65	32	3
3P16	64	33	3
3R01	38	46	16
3R02	41	49	10
3R03	67	33	–
3R04	41	47	12
*3R05	42	47	10
*3R06	40	49	11

Le predizioni non hanno un andamento lineare, nel contesto postonico, se si esclude l'indice FRE1, tutti i valori sono simili. Anche nel contesto pretonico, le percentuali sono comparabili, ad eccezione di SR03, ma aumentano invece le probabilità che il fonema sia sostituito con ALTRE varianti.

Per finire abbiamo controllato in che modo le variabili accentuali interagissero con la classe del suono (cs-LUP-SIL-DAL). L'aggiunta di quest'ultima informazione ridisegna completamente il quadro delle predizioni fornite da AM. Per meglio cogliere le differenze, abbiamo interrogato coppie di stringhe identiche, ad eccezione della classe di suono. Le divergenze si apprezzano per tutte le sequenze considerate, sebbene per alcuni contesti siano più cospicue, si veda la tab. 9.

Nuovamente, gli indici di geminazione crollano, quando il fonema, indifferentemente ostruente o sonorante, si trova in contesto atono, una condizione quest'ultima vincolata alla presenza di una parola mediamente lunga. L'abbassamento degli indici probabilistici (es. in O5A2 e S5A2) non è però dovuto alla lunghezza della parola, come prova l'interrogazione di alcune stringhe fittizie, cioè non ricorrenti nel nostro campione di it-L2; i tassi di geminazione infatti si innalzano solo quando la geminata, pur inserita in un polisillabo, è adiacente ad un accento lessicale (es. O7P1).

Tabella 9 - *Indici probabilistici (%) CS-LUP-SIL-DAL, le stringhe precedute da * sono fittizie*

<i>Codifica</i>	<i>Geminata</i>	<i>Degeminata</i>	<i>Altra</i>	<i>Esempio</i>
O2P1	75	23	2	tutto
S2P1	25	75	-	bello
O3P1	76	20	4	adesso
S3P1	40	60	-	capelli
O4P1	89	11	-	bicicletta
S4P1	18	79	3	capodanno
O2R0	51	49	-	città
S2R0	-	-	100	vorrei
O3R0	45	43	12	abbiamo
S3R0	24	72	4	colleghi
S4R0	65	30	5	possibile
S4R0	3	97	-	ombrellone
O5R0				interessante
*S5R0				giocherellare
O6R0	50	41	9	
*S6R0	16	82	2	
O5A2	9	82	9	appartamento
S5A2	25	75	-	arrivederci
*O5P1	71	29		gentilissimo
*S5P1	16	79	5	tergicristallo
*O7P1	76	22	2	formidabilissimo
*O7A2	11	72	17	repubblicanesimo

5. Conclusioni

La ricerca condotta ha prodotto una fitta serie di dati che si presta a diverse letture interpretative, evidenziando anche alcune criticità. Partiamo da queste ultime: le predizioni analogiche risentono della struttura dei corpora analizzati. Nell'archivio, alcuni fonemi geminati sono infrequenti, (/vv/ o /dd/), altri assenti (/gg/). Scarsamente rappresentati sono pure le Congiunzioni o i polisillabi contenenti un fonema lungo. È ragionevole pensare che in questo caso l'algoritmo abbia costruito un set analogico troppo ampio, includendo anche quelle stringhe che condividevano con l'*item* interrogato un solo valore, perdendo così parte della propria azione predittiva. Questo è anche il motivo per cui dati effettivi e dati simulati non sempre combaciano.

Nel complesso, l'efficacia predittiva di AM è comunque apprezzabile: ciò è particolarmente evidente durante l'interrogazione di stringhe fittizie, una condizione molto informativa, poiché ha permesso di guardare oltre i confini dell'archivio rac-

colto e di prevedere dei comportamenti linguistici, anche per una serie di contesti, altrimenti non osservabili direttamente. Naturalmente, in questi casi il raffronto tra dati simulati e dati effettivi, di norma molto positivo, non è stato possibile.

Con riferimento alla geminazione in it-L2, l'azione di alcuni fattori appare confermata, mentre il ruolo di altre variabili risulta ridimensionato. Dagli esiti estratti da AM rileviamo che tutti i fonemi ricevono un valore probabilistico di geminazione. Parimenti, anche nei dati non simulati nessun fonema lungo è realizzato sempre geminato o sempre degeminato. Ciò vuol dire che negli apprendenti, nonostante il loro modesto livello di competenza dell'italiano, la formazione della categoria C [+Lungo] è già avviata. L'azione delle variabili testate non è però categorica, ma suscettibile di gerarchizzazione. Nel processo che porterà all'acquisizione dei fonemi geminati è possibile individuare l'azione di predittori forti, Classe di suono (CS) e Sonorità (SO), di predittori medi, Modo di Articolazione (MA), Sillaba (SIL), Distanza dall'Accento Lessicale (DAL) e Classe lessicale (CL), e di predittori deboli, Luogo di Articolazione (LA), Lunghezza di Parola (LUP) e Frequenza (FRE).

Il valore di CS è indubbio, ma altrettanto influente è il tratto di sonorità, gli indici predittivi, in linea con i dati effettivi, sono nettamente più alti per le ostruenti sorde rispetto a quelli estratti per sonoranti e ostruenti sonore. Per quest'ultima classe anzi riscontriamo valori marginali, è il caso di /vv/ o di /bb/ in cui la geminazione è prevista in modo sporadico, confermando il dato reale. È possibile che per questi contesti ci sia stata una forma di alterazione analogica, il numero delle ostruenti sonore presenti nell'archivio è infatti inferiore, solo il 9% sul totale dei fonemi, rispetto a quello delle ostruenti sorde (55%) o delle sonoranti (36%). Nell'archivio, l'associazione OSTRUENTE-GEMINATA-SONORA diventa così una combinazione contestuale inusuale, un'associazione poco produttiva, in realtà, come vedremo, anche a livello tipologico.

Un secondo gruppo di codifiche (MA, SIL, DAL, CL) mostra una forza predittiva secondaria, ad es. la posizione postonica, un contesto ottimale di geminazione (Thurgood, 1993), non innalza significativamente le predizioni se il fonema non equivale ad un'ostruente sorda. Parimenti, l'azione della Classe del suono (CL) sul processo non è autonoma, ma subordinata alla classe del suono coinvolto. Le rimanenti variabili svolgono un'azione indiretta, debolmente incisiva, ad es. LUP favorisce la variante degeminata solo quando è associata alla codifica ATONA, laddove FRE non sembra mai avere un ruolo centrale nel processo indagato.

Nell'insieme, le predizioni probabilistiche sembrano suffragare alcuni universali implicazionali formulati proprio in merito alla geminazione. Alcuni studi, sulla base di ampie evidenze di natura tipologica, hanno postulato come la geminazione sia fortemente correlata al grado di forza consonantica del fonema: la presenza di occlusive geminate sonore implicherebbe quella delle corrispettive sorde (Blevins, 2005; Thurgood, 1993), la ricorrenza di queste ultime sarebbe parimenti implicata dalle fricative geminate (Podesva, 2002). I riscontri tipologici non sono tuttavia univoci e richiamano una tematica piuttosto controversa. Da un lato, si assume che le geminate più frequenti tra le lingue del mondo siano le ostruenti sorde, ed *in*

primis le occlusive, poiché suoni ottimali (Podesva, 2002), le geminate sonore e le sonoranti sarebbero invece evitate per via della loro somiglianza, soprattutto in termini di intensità, con le vocali (Podesva, 2002; Taylor, 1985)⁶. Dall'altro, si osserva pure che l'esigua presenza di ostruenti sonore geminate potrebbe essere il riflesso dell'altrettanta scarsa incidenza delle scempie, ad es. le fricative sonore scempie sono poco frequenti sul piano tipologico (Dmitrieva, 2012). In controtendenza, però si pongono le nasali geminate, le quali, a dispetto della loro sonoranza, sono tutt'altro che rare, anzi secondo Jaeger (1978) sarebbero la classe di geminate in assoluto più frequente tra le lingue⁷.

In generale, le predizioni di AM provano che nell'interlingua il processo di geminazione, pur nella sua natura progrediente, inizia dalle consonanti articolatoriamente più stabili. Le ostruenti sorde sono infatti unità ottimali sul fronte percettivo ed acustico, segmenti fonologicamente forti e non marcati, e quindi anche per questo più inclini a realizzare un'opposizione di lunghezza consonantica. È d'altra parte noto che i principi di naturalezza e di marcatezza tipologica svolgano un ruolo fondamentale nell'interlingua, prefigurandosi come indici predittivi del grado di difficoltà di un processo acquisizionale.

Bibliografia

BERTINETTO P.M. - BURANI C. - LAUDANNA A. - MARCONI L. - RATTI D. - ROLANDO C. - THORNTON A. (2005), *Corpus e Lessico di Frequenza dell'Italiano Scritto (CoLFIS)*. Disponibile in: <http://linguistica.sns.it/CoLFIS/Home.htm> (data ultimo accesso 03 settembre 2015).

BLEVINS J. (2005), The typology of geminate inventories: Historical explanations for recurrent sound patterns, in *Seoul Linguistics Forum 2005*, Seoul National University, Seoul: 121-137.

BOERSMA P. - WEENINK P. (2007), *Praat: doing phonetics by computer* (Version 5.0), University of Amsterdam, The Netherlands, Disponibile in: <http://www.praat.org> (data ultimo accesso 03 settembre 2015).

BYBEE J. (2002), Word frequency and context of use in the lexical diffusion and phonetically conditioned sound change, in *Language Variation and Change* 14: 261-290.

COSTAMAGNA L. - MONTILLI C. - RICCI I. (2014), The role of prosodic features in the acquisition of consonant gemination by Chinese learners, in COSTAMAGNA L. - CELATA C. (eds.), *Consonant gemination in first and second language acquisition*, Pacini Editore, Pisa: 47-82.

DE MAURO T. (1980), *Guida all'uso delle parole*, Editori Riuniti, Roma.

⁶ I principi a cui si fa normalmente riferimento per argomentare la scarsa presenza di sonoranti geminate sono tre: principio di aerodinamicità, principio dello sforzo articolatorio, principio della percepibilità.

⁷ Podesva (2002) e Maddieson (2008) rilevano invece che le nasali geminate, si pongono, per incidenza tipologica, subito dopo le occlusive sorde geminate. Seguono in ordine decrescente liquide, fricative, approssimanti e affricate.

DE MAURO T. - MANCINI F. - VEDOVELLI M. - VOGHERA M. (1993), *Lessico di frequenza dell'italiano parlato*, Etas Libri, Milano.

DERWING B. - SKOUSEN R. (1994), Productivity and the English past-tense: testing Skousen's Analogical Model, in LIMA S.D. - CORRIGAN R.L. - IVERSON G.K. (eds.), *The reality of linguistic rules*, John Benjamins, Amsterdam: 193-218.

DMITRIEVA O. (2012), *Geminate typology and the perception of consonant duration*, Dissertation submitted to the Department of Linguistics and the Committee in Graduate Studies of Stanford University.

EDDINGTON D. (2000), Spanish stress assignment within the analogical modeling of language, in *Language* 76: 92-109.

EDDINGTON D. (2002), Spanish gender assignment in an analogical framework, in *Journal of quantitative Linguistics* 9: 49-75.

EDDINGTON D. - LONSDALE D. (2007), Analogical Modeling: An Update, Online Lecture Notes of the 12th ESSLL, *I Workshop Exemplar-based Models of Language, Acquisition and Use*; Dublin, Ireland; August 2007. Disponibile in: <http://humanities.byu.edu/am/amupdate.pdf> (data ultimo accesso 03 settembre 2015).

ENOMOTO K. (1996), Interlanguage phonology: the perceptual development of durational contrasts by English-speaking learners of Japanese, in *Edinburgh Working Papers in Applied Linguistics* 3: 25-35.

GIANNINI S. (2003), Typological comparison and interlanguage phonology: maps or gaps between typology and language learning of sound systems?, in GIACALONE A. (ed.), *Typology and second language acquisition*, Mouton de Gruyter, Berlin: 403-440.

GIANNINI S. - COSTAMAGNA L. (1997), Language learning strategies in interlanguage phonology of Italian L2: a case history: the acquisition of consonant length, in LEATHER J. - JAMES A. (eds.), *New sounds 97*, Proceedings of the 3rd International Symposium on the acquisition of second language speech, Klagenfurt Universitätverlag, Klagenfurt: 96-103.

GIANNINI S. - COSTAMAGNA L. (1998), Acquisizione di categorie fonologiche e diffusione lessicale del mutamento linguistico: affinità strutturali, in *Archivio Glottologico Italiano* 85: 150-187.

HAMMOND M. (2003), *Programming for linguists: Perl for language researches*, Blackwell Publishing, Oxford.

HAN M. (1992), The timing control of geminate and single stop consonants in Japanese: a challenge for nonnative speakers, in *Phonetica* 49: 102-127.

HARADA T. (2006), The acquisition of single and geminate stops by English speakers in a Japanese immersion program, in *Studies in Second Language Acquisition* 28: 601-632.

JAEGER J. (1978), Speech aerodynamics and phonological universals, in *Proceedings of the Annual Meeting of the Berkeley Linguistics Society*, 4: 311-329.

JOHNSON K. (1997), Speech perception without speaker normalization. An exemplar model, in: JOHNSON K. - MULLENNIX J.W. (eds.), *Talker variability in speech processing*, San Diego, Academic Press: 145-165.

KABAK B. - RECKZIEGEL T. - BRAUN B. (2011), Timing of second language singletons and geminates, in WAI-SUM L. - ZEE E. (eds.), *Proceedings of the 17th International Congress*

of *Phonetic Sciences*, (China, Hong Kong, 17-21 August 2011), City University of Hong Kong: 994-997.

KUNNARI S. - NAKAI S. - VIHMAN M.M. (2001), Cross-linguistic evidence for the acquisition of geminates, in *Psychology of Language and Communication* 5: 13-24.

LONSDALE D. - MATSUSHITA H. (2013), Modeling speech errors by analogy, in WEST R.L. - STEWART T.C. (eds.), *Proceedings of the 12th International Conference on Cognitive Modelling, (ICCM 2013)*, Ottawa, Canada: 17-22.

MACALLISTER R. - FLEGE J.E. - PISKE T. (2002), The influence of L1 on the acquisition of Swedish quantity by native speakers of Spanish, English and Estonian, in *Journal of Phonetics* 30: 229-258.

MADDISON I. (2008), Glides and gemination, in *Lingua* 118: 1926-1936.

MAH J. - ARCHIBALD J. (2003), Acquisition of L2 length contrast, in LICERAS J.M. - ZOBL H. - GOODLUCK H. (eds.), *Proceedings of the 6th generative approaches to second language acquisition conference (GASLA 2002)*, Cascadilla, Somerville (Ma): 208-212.

PAYNE E.M. (2005), Phonetic variation in Italian consonant gemination, in *Journal of the International Phonetic Association* 35(2): 153-189.

PIERREHUMBERT J. (2003), Probabilistic phonology: discrimination and robustness, in BOD R. - HAY J. - JANNEDY S. (eds.), *Probabilistic Linguistics*. Cambridge MA, MIT Press: 177-228.

PHILLIPS B. (1984), Word frequency and the actuation of sound change, in *Language* 60: 320-342.

PODESVA R.J. (2002), Segmental constraints on geminates and their implication for typology, Paper presented at the 76th Annual Meeting of the Linguistic Society of America, San Francisco.

SIMONE R. (1988), Fragilità della morfologia e contesti turbati, in GIACALONE RAMAT A. (a cura di), *L'italiano tra le altre lingue. Strategie di acquisizione*, Il Mulino, Bologna: 91-98.

SKOUSEN R. (1989), *Analogical modeling of language*, Kluwer, Dordrecht.

SKOUSEN R. (1992), *Analogy and structure*, Kluwer, Dordrecht.

SKOUSEN R. - LONSDALE D. - PARKINSON D.B. (eds.), (2002), *Analogical modeling: an exemplar-based approach to language*, John Benjamins, Amsterdam.

SORIANELLO P. (2009), L'acquisizione del tratto di lunghezza consonantica in italiano L2, in ROMITO L. - GALATÀ V. - LIO R. (a cura di), *La fonetica sperimentale. Metodo e applicazioni*, Atti del IV Convegno Nazionale dell' AISV, (Rende, 3-5 Dicembre 2007), CD-Rom, EDK, Torriana: 40-61.

SORIANELLO P. (2014), Italian geminate consonants in L2 acquisition, in COSTAMAGNA L. - CELATA C. (eds.), *Consonant gemination in first and second language acquisition*, Pacini Editore, Pisa: 25-46.

SORIANELLO P. - BERTINETTO P.M. - RICCI I. (2005), Alle sorgenti della variabilità della 'gorgia' fiorentina: un approccio analogico, in COSÌ P. (a cura di), *Misura dei Parametri. Aspetti tecnologici ed implicazioni nei modelli linguistici*, Atti del I Convegno Nazionale AISV, Associazione Italiana di Scienze della Voce, Padova, 2-4 dicembre 2004, EDK Editore, Brescia, CD-Rom: 327-362.

TAYLOR M. (1985), Some patterns of geminate consonants, *University of Chicago Working Papers in Linguistics* 1: 120-129.

THURGOOD G. (1993), Geminate: a cross-linguistic examination, in NEVIS J.A. - MCMENAMIN G. - THURGOOD G. (eds.), *Papers in honor of Frederick H. Bregelman on the occasion of 25th anniversary of the Department of Linguistics*, California State University, Fresno: 129-139.

VIETTI A. (2013), Alcune riflessioni sulla teoria degli esemplari e la variazione linguistica, in *Rivista Italiana di Dialettologia* 36, 7-22.

PARTE II

TECNOLOGIE PER COMUNICARE

ALESSANDRO BITONTI¹

L'e-mail nella comunicazione accademica fra pragmatica e coesione testuale

This paper investigates the syntactic and pragmatic functions of connectives in e-mail communication. The study analyses a corpus of students' e-mails addressed to university professors and shows that connectives, with their communicative functions, produce a strongly interactive text type. In the data under analysis, connectives are strongly relevant in the distribution of the information in the discourse and also convey social meanings connected to the positive face of the interlocutors. More specifically, despite the diamesic, diaphasic and diastratic connotation of the communicative context, connectives seem to be used to attenuate the participants' social asymmetry and to prevent possible conflicts arising from the content of the text itself.

1. *E-mail e testualità*

L'analisi che si propone in questo studio riguarda la testualità della comunicazione tramite e-mail, mezzo che, benché sembri destinato a scomparire negli scambi privati e soprattutto giovanili, gode ancora di ampio utilizzo nel settore pubblico e nel mondo del lavoro (Lorenz, 2007). Come sostengono Skovholt - Svennevig (2013: 589), "e-mail has become the central communication tool for most office workers. It is the medium in which people carry out their daily professional activities and in which workers and business partners build and maintain professional and interpersonal relations".

I messaggi di posta elettronica rappresentano, a oggi, uno strumento di comunicazione che arriva sempre più di frequente a sostituire, insieme ai social network e alle diverse applicazioni internet, le conversazioni telefoniche tradizionali (Jansen, 2002: 156). Dürscheid - Frehner (2013: 35) sostengono anche che:

people send each other e-mails in situations in which they would not have addressed each other earlier on, and so it has become much easier to approach another person when needing assistance. Teachers, for instance, frequently receive e-mails from their students who ask questions about homework assignments, upcoming exams, or personal matters, inquiring about problems for which they would have found a solution without the teacher's help in earlier times.

Anche in ambito accademico la posta elettronica funge da canale di interazione privilegiato; in particolare gli studenti universitari utilizzano questa nuova forma

¹ Masarykova univerzita.

epistolare per interagire con i docenti, per chiedere loro indicazioni e, spesso, per negoziare informazioni.

Baron (2000: 235) ritiene inoltre che “since we construct e-mail in social isolation, and since we see the medium as ephemeral, we don't feel particularly constrained by the social conventions that govern face-to-face exchange or written communication”. Pertanto, l'e-mail presenta limiti diversi da quelli imposti dalla conversazione faccia-a-faccia e non ha ancora del tutto sviluppato convenzioni linguistiche proprie. Benché si tratti di un testo scritto (che prevede una struttura più complessa e curata rispetto ai testi orali), le sue caratteristiche primarie sono, come molti autori sostengono, la scarsa pianificazione e l'esigua accuratezza morfosintattica. Per via di questa loro particolarità, i messaggi di posta elettronica possono contenere errori e strutture frasali poco elaborate. Tali elementi, secondo Milroy - Milroy (1985), vengono in genere tollerati dagli interlocutori, rendendo meno formale e meno formalizzato questo strumento di comunicazione asincrona.

Gli studi sul tema, come prova Bazzanella (2003), hanno ormai superato la dualità “scritto/parlato” di questo genere testuale e hanno spostato l'attenzione sull'asse della variazione diafasica e diastratica rispetto all'uso, ad esempio, di registri che variano in dipendenza dell'argomento, degli attanti e delle intenzioni comunicative (per alcune indicazioni relative alla scrittura delle e-mail si veda Gualdo *et al.*, 2014). Tuttavia, nel caso della Comunicazione Mediata dal Computer (d'ora in poi CMC), e nel caso specifico di questo studio, non è il grado di formalità del testo che diventa rilevante, ma la cura e il controllo dello scrivente sul piano delle forme linguistiche.

Nella presente ricerca ci si occuperà di alcuni meccanismi di coesione (par. 3) presenti in un corpus di 68 e-mail di studenti universitari, dei corsi di laurea triennale e magistrale in lingue straniere dell'Università del Salento, che interagiscono con i docenti per richiedere informazioni o chiarimenti rispetto alle attività didattiche. Si tratta perciò di comunicazioni prodotte e trasmesse da parlanti giovani di istruzione superiore che, grazie alle loro competenze linguistiche ed informatiche, manipolano i sistemi e si fanno promotori, secondo la nostra ipotesi, di innovazione linguistica.

Sul piano delle strutture, la coesione testuale sembra incidere sulla funzione comunicativa del messaggio producendo, a seconda delle tipologie di atti linguistici presenti, testi, oltre che scarsamente pianificati, di lunghezza e concisione variabile. La testualità che ne deriva si traduce attraverso elementi di coesione che riguardano tutti i livelli della lingua, come ad esempio l'uso di espressioni anaforiche e deittiche e di connettori che segnalano le relazioni logico-sintattiche degli enunciati.

Nel caso specifico di questo studio si evidenzierà come i connettivi, oltre a essere essenziali per la distribuzione delle informazioni, producono un alto livello di interattività e puntano sulla funzionalità, o meglio ancora sull'efficienza del messaggio (Dressler - De Beaugrande, 1994).

2. *L'e-mail fra diamesia e diafasia*

Oggi più che mai le strumentazioni tecnologiche rappresentano un elemento costante e rilevante nella quotidianità dei parlanti e sono in grado di determinare variazioni linguistiche significative le cui prospettive di analisi devono mettere a fuoco non soltanto il cambiamento delle strutture, ma anche le dinamiche sociolinguistiche e pragmatiche che intercorrono fra i fruitori delle tecnologie della comunicazione.

Nell'ultimo decennio si è assistito a un'evoluzione dei canali della CMC e si è passati da prodotti multimediali statici (web 1.0) a prodotti dinamici che consentono un'interazione maggiore con i contenuti presenti nel web. Siamo quindi nell'epoca del web 2.0 (O'Reilly, 2005) in cui i parlanti gestiscono e, soprattutto, costruiscono relazioni fra i contenuti utilizzando strumenti e prodotti che appartengono sia alla prima quanto alla seconda fase dello sviluppo tecnologico. Proprio Fiorentino (2010: 200), ad esempio, sostiene che "gli strumenti e i prodotti del web 2.0, vecchi o nuovi che siano, risultano tutti caratterizzati da un alto grado di mescolanza e di interattività"; così ai blog, agli spazi personali, alle risorse wiki, al podcasting, a youtube, alle applicazioni di messaggiera istantanea si affiancano strumenti per così dire obsoleti come chat, e-mail, newsgroup, mailing list o forum.

Considerata la struttura dinamica e ipertestuale del web di oggi, i tipi di testualità sono influenzati non soltanto dagli elementi linguistici a disposizione dei parlanti, ma anche e in maniera prevalente dalle caratteristiche del mezzo e dalla sua dimensione socio-pragmatica; si pensi alle modalità di trasmissione sincrona e asincrona, alla simmetria/asimmetria dei ruoli dei partecipanti, alle caratteristiche e alla dimensione del gruppo di condivisione, alle modalità di accesso alla pagina (personale o pubblica), all'argomento trattato (si veda a tal proposito il volume curato da Held - Schwarze, 2011).

La lingua trasmessa attraverso i nuovi media si caratterizza per una forte dialogicità e per una continua e frequente interazione fra gli utenti. Le comunicazioni risultano pertanto fortemente contestualizzate e i ruoli sociali vengono stabiliti e coordinati secondo le modalità tipiche dell'oralità e degli scambi interazionali in presenza (cfr. Berruto, 2005).

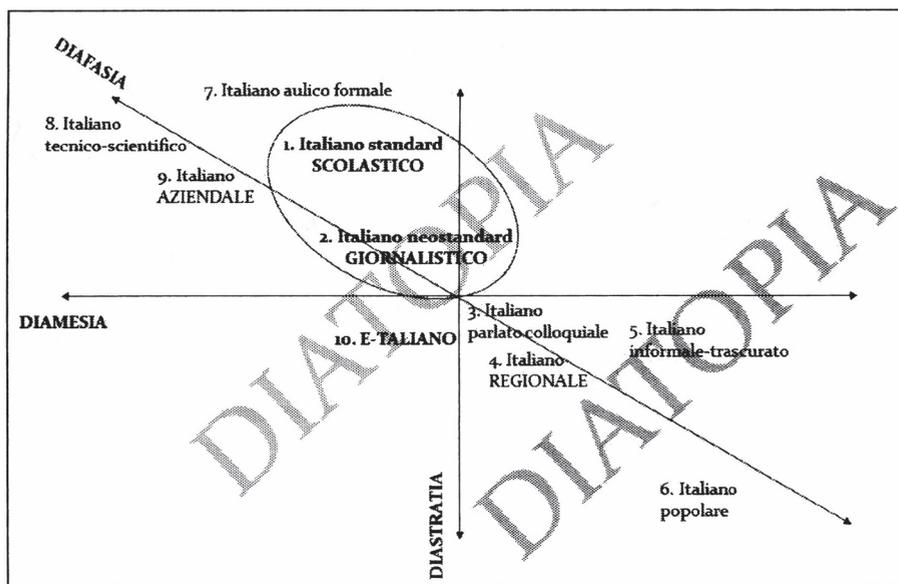
Se si osserva il repertorio linguistico italiano (fig. 1), la varietà trasmessa ("e-taliano" per Antonelli, 2014) non riguarda unicamente la dimensione diamesica, ma anche quegli ambiti della variazione relativi alla diafasia e alla diastratia; si tratta in sostanza, secondo Antonelli (2014: 539), di "una varietà scritta diafasicamente informale e diastraticamente anche bassa".

Lo stesso Berruto (2005) sottolinea come nella CMC il mezzo grafico si orienti verso un livello di interattività molto elevato, simulando le pratiche orali della lingua: frammentarietà, velocità e socialità diventano così i caratteri prevalenti.

Ancora, in merito al dualismo scritto/parlato, si è senza dubbio di fronte a uno sviluppo ulteriore della lingua: la scrittura sta rivivendo un momento di slancio grazie alle nuove tecnologie che, in quanto strumenti di massa, determinano un uso scritto dell'italiano, piuttosto fluido e dai contorni imprecisi e sovrapposti, tenden-

te verso la zona bassa del repertorio. Tale orientamento si può individuare ad esempio sul piano dell'organizzazione testuale (lo scritto dei nuovi media risulta spesso non sufficientemente pianificato), sul piano semantico (in quanto il contenuto del messaggio è predominante rispetto alle forme linguistiche utilizzate per veicolarlo) e sul livello socio-pragmatico (il contesto, in particolare i rapporti fra i partecipanti allo scambio e gli scopi comunicativi, determina gli usi linguistici) (cfr. Berruto, 2004; Tempesta, 2006; Sabatini, 2012).

Figura 1 - *Il repertorio dell'italiano (da Antonelli 2014: 539, con adattamenti)*



Come nel parlato spontaneo, anche i testi digitali sono marcati in diastratia e la loro tendenza verso l'alto o verso il basso dipende dai parlanti che li producono. Data l'ampia democraticità dei prodotti multimediali, tutti i gruppi sociali sviluppano contenuti e linguaggi variegati, facendo emergere gli usi e le competenze linguistiche degli italiani. Il web rappresenta infatti un contesto nel quale diverse classi di parlanti interagiscono fra di loro più di quanto non accada in contesti non-virtuali: al dibattito o alla semplice fruizione di un messaggio partecipano professionisti ma anche cittadini più o meno istruiti e di diverse fasce generazionali.

In particolare, come sostiene Tempesta (2006: 39) “sembra difficile pronosticare fino a che punto l'influenza del *carpe diem* linguistico giovanile possa incidere sul sistema dell'italiano e sull'organizzazione del repertorio”, tuttavia le strumentazioni tecnologiche sembrano contribuire a “rendere giovane, meno paludato, per alcuni versi irruente, l'italiano”. Proprio Tempesta (2010), per esempio, sottolinea come nel parlato di registro informale dei giovani prevalga l'uso di marcatori pragmatici di tipo interattivo con funzione focalizzante e cooperativa, evidenziando l'approccio collaborativo e il gioco di faccia positivo nei confronti dell'interlocutore.

3. *Coesione e connettivi*

La coesione è notoriamente una caratteristica del discorso che utilizza dei dispositivi di connessione (di tipo morfologico, lessicale, sintattico e interpuntivo) assicurando la continuità e i rapporti grammaticali fra le strutture linguistiche e garantendo l'accesso ai significati del testo e, quindi, allo sviluppo della coerenza testuale. Questi dispositivi coesivi servono a giustificare, a segnalare o semplicemente a marcare le relazioni logiche e sintattiche fra diversi nuclei tematici e a definire, di conseguenza, spostamenti o cambiamenti dei significati. Bazzanella (2007: 81) conferma inoltre che:

i meccanismi di coesione possono variare nella quantità e nella qualità in base al tipo di testo, se non mancare quasi totalmente, come nella lingua infantile, o nel linguaggio quotidiano caratterizzati da forte ellitticità, indeterminatezza, giustapposizione paratattica invece che subordinazione, rimando al contesto.

Riprendendo ancora Bazzanella (2007: 78-79), tra i vari dispositivi necessari per la coesione testuale si utilizzano:

pronomi, referenti testuali, anafore pronominali e lessicali, proforme, tempi verbali, congiunzioni, connettivi [...], correlazioni lessicali [...], articoli determinativi [...], accordo morfologico, struttura dell'informazione, ripetizione parziale o totale [...], parafrasi, ellissi, punteggiatura, deissi testuale in genere.

Semplificando, Serianni (2007: 98) ritiene che le “sfumature discorsive o argomentative” di un testo siano determinate, oltre che da altri elementi (es. la deissi, le espressioni anaforiche e cataforiche o l'ellissi)², dai connettivi testuali, ovvero da classi di parole come le preposizioni, le congiunzioni o determinate locuzioni complesse, come sintagmi verbali o intere clausole, che possono svolgere diverse funzioni pragmatiche a seconda del contesto nel quale si collocano (Colombo, 2012). Di queste stesse classi di parole ci occuperemo nel paragrafo 5.

Il dibattito sui connettivi è ancora molto vivo: essendo alimentati da più categorie morfologiche e assumendo diverse funzioni comunicative, i connettivi costituiscono una classe aperta e risulta anche difficile redigerne un elenco completo. Possono appartenere a questa classe, pertanto, “tutti gli elementi che collegano tra loro segmenti più o meno ampi di testo” (Palermo, 2013: 189).

Nella stessa direzione sembrano muoversi Sabatini *et al.* (2011: 611) i quali introducono il concetto di “legamento sintattico”, intendendo con questa dizione “le parole e le espressioni (avverbi, congiunzioni, locuzioni congiuntive ecc.) che collegano i vari punti del discorso volto nel testo”. Tali legamenti sintattici vengono distinti in base alla funzione (di tempo, di spazio, di causa ed effetto, di numerazione, di corrispondenza, di contrapposizione, di dimostrazione, di valutazione, di esposizione).

² Si veda anche Dressler - De Beaugrande (1994). Significativo, ai fini della coesione, è anche l'uso della punteggiatura e di alcuni segnali discorsivi che confermano, ancora una volta, il legame con le modalità orali della lingua, attraverso la conseguente composizione di frasi di tipo prevalentemente paratattico.

In questa sede si analizzeranno le relazioni di collegamento affidate ai connettivi (cfr. par. 5) che, come sostiene Palermo (2013: 190), non sono “elementi inerentemente coesivi, ma indirettamente coesivi, grazie al loro specifico significato e alla relazione che configurano tra ciò che precede e ciò che segue”. La funzione pertanto dei connettivi è quella di definire la strutturazione del testo e delle sue parti al fine di garantire la coerenza *stricto sensu* e la continuità comunicativa fra i parlanti.

Si accoglie per questo studio, dunque, la formulazione di Palermo (2013: 210), in continuità con Bazzanella (1985), che distingue tra connettivi semantici e connettivi pragmatici: “mentre i connettivi semantici esprimono una relazione tra fatti, i connettivi pragmatici segnalano il punto di vista dei parlanti sull’enunciato o sull’atto di enunciazione”. Questi ultimi sono dei veri e propri segnali discorsivi, appartengono a diverse categorie grammaticali e la loro caratteristica primaria è la polifunzionalità: la loro funzione rispecchia il punto di vista del locutore e varia col modificarsi degli elementi con i quali cooccorrono.

4. *Materiali e metodi*

Il corpus esaminato è costituito da 68 e-mail che sono state analizzate attraverso il programma di interrogazione di corpora TextSTAT. Attraverso il software è stato possibile selezionare i connettivi sulla base delle forme e, grazie alla wordlist ottenuta, determinarne le occorrenze. Successivamente, tenendo conto dei contesti frasali (ovvero il cotesto precedente e successivo alla parola chiave) ed extra-frasali, sono state identificate le funzioni pragmatiche (demarcative e interazionali). Tali funzioni sfumano spesso l’una nell’altra (par. 3) e, anche all’interno dello stesso contesto, non hanno necessariamente valore univoco.

Sulla base della letteratura considerata (par. 3) e delle frequenze prevalenti, l’analisi si è concentrata su 200 congiunzioni, 119 forme avverbiali e 62 locuzioni di diverso tipo, tra cui perifrasi, sintagmi verbali e intere proposizioni (Serianni, 2007).

5. *Risultati*

Sul piano delle funzioni, i connettivi che caratterizzano il corpus sono orientati principalmente verso la funzione metatestuale di tipo demarcativo (50%), con valore di apertura, proseguimento e chiusura, e verso la funzione interattiva di modulazione (43%). Il restante 7% riguarda casi numericamente non rilevanti in cui le espressioni giuntive hanno funzione di parafrasi, di correzione o di esemplificazione. Veniamo ora alla discussione dei dati.

Già Bazzanella (2010) aveva rilevato che le funzioni metatestuali sono quelle più frequenti nello scritto e anche nel nostro caso, benché l’ambito di analisi sia quello della lingua trasmessa, la maggior parte degli operatori di coordinazione serve a esprimere i rapporti fra gli atti linguistici che compongono il testo.

Tra gli operatori di coordinazione, frequente è l’utilizzo di *ma* in qualità di congiunzione avversativa (Scorretti, 2001: 245). In alcuni casi, come in (1) e (2), questa

esprime opposizione, insieme con una negazione esplicita, al sintagma o alla proposizione con cui è messo in relazione³:

- (1) Sono già due ore che ci provo con tre pc diversi e dal telefono ma non c'è niente da fare!
- (2) Per sicurezza ho controllato sul forum, ma non ho trovato una data confermata.

In altri casi, più frequenti, la costruzione avversativa serve a segnalare il passaggio a un altro argomento o, meglio ancora, a introdurre e a indirizzare l'interlocutore verso il punto focale dell'affermazione. È il caso di (3)-(7):

- (3) Sapevo che la scadenza fosse oggi, *ma* [il sistema informatico]⁴ mi dice che il codice che inserisco è errato.
- (4) Mi sono iscritta al corso di grammatica italiana, *ma* per motivi di lavoro non sono riuscita a frequentarlo finora.
- (5) Il prof. xxx non sarà in commissione *ma* ci ha detto che non serve neanche la sua relazione sull'elaborato.
- (6) Ho appreso che l'elaborato va bene *ma* devo impostare meglio una griglia di osservazione.
- (7) Scusi lei *ma* avevo bisogno di saperlo subito perché fra stasera e domattina devo procedere alle stampe.

Interessanti sono le occorrenze, più rare, in cui *ma* compare all'inizio di frasi ed è utilizzato con valore non coordinativo, come in (8) e (9). Ci si trova di fronte a un elemento che serve a indicare, oltre che immediatezza espressiva, un completamento sintattico e semantico con un contenuto già noto agli interlocutori o presente nel testo della mail.

- (8) Vabene la rongrazio ci vediamo venerdì allora. *Ma* quando finisce il tutorato?
- (9) Gentile professore, *Ma* per caso l'ultima lezione c'è già stata?

Ancora, si sottolinea come questo connettore pragmatico sia utilizzato per introdurre delle semplici frasi dichiarative, eccetto che per (8) e (9) dove l'atto linguistico formulato è di tipo interrogativo. In tutti i casi esposti, dunque, l'operatore di coordinazione si trova in posizione comunicativamente rilevante, utile a segnalare il focus della frase, palesando la sua funzione prevalentemente demarcativa nell'introduzione di un nuovo tema.

³ Gli esempi forniti contengono spesso errori, imprecisioni linguistiche e refusi tipici della scrittura mediata dal computer che sono stati conservati e non corretti. Per motivi di spazio si riporta solo il segmento interessato.

⁴ Tra parentesi quadre si riproducono elementi extratestuali oppure l'omissione di parti di testo non rilevanti ai fini dell'esemplificazione. Come accade nell'oralità, è facile trovare in queste comunicazioni delle informazioni implicite, note agli interlocutori, che sono esito della concisione discorsiva.

I connettori riportati negli esempi (10) e (11) hanno una funzione metatestuale, come segnali demarcativi di apertura.

- (10) *innanzitutto* grazie per i suggerimenti, sistemerò quella parte che lei mi ha “corretto”. è stata una svista!!!... *in merito alla task invece*, l’idea di partenza, cioè la reading e l’esercizio iniziale con le immagini sono prese dal testo ma la strutturazione, la sequenza e tutte le altre attività ideate da me.
- (11) *Visto che* ancora sono l’unica partecipante del corso sulla piattaforma non mi è tanto chiaro ciò che dobbiamo consegnare.

Sia l’avverbio *innanzitutto* che le locuzioni *in merito a* e *visto che* determinano l’attacco del discorso. La locuzione *in merito a* si pone in catena con l’avverbio *invece* collegando, in una relazione di continuità, il nuovo topic a un elemento extratestuale (contenuto nel testo input del docente). Questi marcatori segnalano l’inizio di un nuovo argomento e lo focalizzano all’interno dell’intero testo prodotto dallo studente.

Ancora funzione demarcativa hanno i connettori negli esempi (12)-(14). La congiunzione *infatti* segnala il passaggio a un atto dichiarativo; analogamente la congiunzione *eppure* e la locuzione avverbiale *in più* articolano il passaggio (in direzione semplicemente aggiuntiva) da un’informazione a un’altra.

- (12) Nel frattempo, ho cercato di rimettermi al passo con arabo. *Infatti*, mi mancano esclusivamente 5 esami (di arabo), rispettivamente uno del primo anno, due del secondo e due del terzo anno.
- (13) Non sto riuscendo ad accedere alla piattaforma on line. [...]! *Eppure* fino a questa mattina non ho avuto problemi, *infatti* ho svolto l’analisi grammaticale da lei proposta.
- (14) io ho già fatto 3 assenze purtroppo, [...] *in più* volevo dirle che io ci terrei molto a seguire l’altro suo corso, però purtroppo non potrò venire ogni mercoledì.

Valore demarcativo ha anche il connettore *comunque*. Questo si trova soprattutto come indicatore di chiusura, in (15)-(19), utilizzato per segnalare la conclusione di un blocco argomentativo.

- (15) per problemi vari non ho potuto eseguire molti esercizi in piattaforma... è possibile avere *comunque* i 2 CFU?
- (16) Potrebbe *comunque* conseguire i 2 CFU se riesce a inviarle il file del ProjectWork entro Venerdì?
- (17) Spero di essere *comunque* inserita tra i partecipanti, nonostante i pochi giorni di ritardo.
- (18) *Comunque*, il problema principale che riscontro, [...], è la mancanza di appelli che mi agevolino a provare/riprovare a sostenere gli esami.
- (19) Tutti gli esercizi li sto svolgendo *comunque* da casa e spero lei possa capire.

L'avverbio, anche nei casi in cui è posto in apertura di frase, contrassegna la conclusione dell'asserzione precedente, collegando il periodo, secondo le modalità dell'ellipsis anaforica e cataforica, a quanto dichiarato in precedenza o successivamente.

Questo avverbio conserva in ogni caso il proprio valore avversativo, di demarcazione testuale, mettendo in relazione delle frasi dichiarative ma anche collegando degli atti di domanda con altre sezioni del periodo, come in (15) e (16). In (17) la funzione avversativa è rafforzata dalla presenza del sintagma preposizionale introdotto da *nonostante*.

I connettivi con valore di modulazione si presentano spesso sotto la forma di avverbi derivati: *gentilmente*, *possibilmente*, *effettivamente*, *inutilmente*, *nuovamente* servono ad attenuare il contenuto proposizionale. Mentre nei casi (20) e (21) l'avverbio è utilizzato per marcare ancora di più l'atto linguistico nel tipo ottativo, nell'esempio (22) questo mitiga l'imperativo e la finalità iussiva dell'enunciato.

(20) vorrei *gentilmente* sapere qual'è il termine ultimo di consegna della tesi.

(21) le allego la bozza perché vorrei *possibilmente* un suo parere.

(22) dia *gentilmente* un'occhiata anche lei.

Nei casi di (23)-(26) il punto di vista adottato è quello dell'interlocutore. Lo studente che scrive si pone in una condizione di empatia con il testo ricevuto dal docente, in particolare dimostrando accordo in (23), con l'avverbio posto in posizione iniziale (in questo caso ha anche valore demarcativo con funzione di apertura), ed esplicitando dei segnali di cortesia in (24)-(26):

(23) Gent.mo prof X, *effettivamente* ho dato il mio ultimo esame a luglio scorso,

(24) non volendo farle perdere *inutilmente* tempo,

(25) la ringrazio *nuovamente*.

(26) le chiedo scusa se la disturbo *nuovamente*, *ma* man mano che si lavora sorgono i dubbi.

In quest'ultimo esempio l'avverbio si pone in catena con la congiunzione *ma* che conferisce certamente valore avversativo alla frase, ma la cui funzione primaria è quella di introdurre un nuovo argomento.

Avverbi come *assolutamente*, *completamente* e *necessariamente* hanno la funzione di rafforzamento degli enunciati. Si tratta per (27) e (28) di forme avverbiali che intensificano la forza illocutoria dei due atti dichiarativi, mentre per (29) il tipo interrogativo è rafforzato oltre che dall'avverbio anche dal modale con valore deontico.

(27) perché in realtà una tesi non deve essere *assolutamente* un copia e incolla da internet.

(28) ho perso in quest'ultima sessione di esami *completamente* la motivazione e la volontà di proseguire i miei studi.

(29) ci sono dei requisiti particolari per il layout della tesina finale che devono essere *necessariamente* seguiti?

Come evidente negli esempi (30)-(33), l'avverbio frasale *purtroppo* ricorre all'interno di atti dichiarativi sia all'inizio che alla fine della proposizione.

- (30) io ho già fatto 3 giorni di assenze *purtroppo*, perchè tra lavoro e influenza non sono potuta venire.
- (31) io *purtroppo* non riesco ad inviarle le attività sulla piattaforma.
- (32) volevo scusarmi della mia assenza *ma purtroppo* ho avuto e sto avendo dei problemi di salute.
- (33) io ci terrei molto a seguire l'altro suo corso di "morfosintassi", *però purtroppo* non potrò venire ogni mercoledì.

La funzione attribuita a questo elemento è di tipo interazionale. L'avverbio valutativo non fa che esprimere lo stato d'animo del parlante che modula, rafforzandolo, il valore semantico della proposizione. Anche la catena con l'avverbio avversativo in (32) e (33) intensifica il coinvolgimento emotivo dell'emittente oltre alla forza illocutoria della frase. Inoltre, mentre in (30) e (31) la posizione del connettore è nella frase reggente, in (32) e (33) questo appare nella coordinata.

Valore attenuativo sembrano avere alcune formule di cortesia del tipo *scusi se*, come in (34)-(36), dove si palesa lo stato d'animo dello scrivente e si sottolineano i ruoli asimmetrici dei due parlanti oltre all'autorità del destinatario.

- (34) mi *scusi se* le scrivo di domenica, però sono tornata oggi dalla Spagna
- (35) *scusi se* la disturbo oggi che è festa
- (36) mi *perdoni se* la disturbo

Frequenti sono anche le occorrenze di locuzioni con valore attenuativo del tipo *se possibile*. La formula, strutturalmente assimilabile alla protasi dei costrutti condizionali, rimanderebbe a una locuzione avverbiale e copre il ruolo di connettivo pragmatico.

- (37) Vorrei aggiungermi sulla lista dei frequentanti, *se possibile*.
- (38) Proverò a farla e poi *se vi è possibile* verrò a trovarla lunedì nel Vostro studio.
- (39) Dia gentilmente un'occhiata anche Lei, *se Le è possibile*, e mi faccia sapere se nella forma (o altro) va bene.

Tali espressioni giuntive si collocano, a mo' di inciso, all'interno del periodo tra due segni di interpunzione (eccetto nell'esempio (38) dove lo studente, per motivi probabilmente legati alla velocità di scrittura, ha ommesso le due virgole) segnalando la volontà di mitigare gli eventuali effetti conflittuali dell'affermazione contenuta nella frase reggente. Tale funzione è rafforzata in alcune catene come in (40)-(42).

- (40) Vorrei iscrivermi al suo corso, *se è ancora possibile*.
- (41) *Appena mi sarà possibile*, cercherò di venire presso il Suo studio durante l'orario di ricevimento.

(42) *Se è possibile però*, vorrei lo stesso fissare un appuntamento per un colloquio.

L'avverbio *ancora* contenuto in (40) suggerisce un ancoraggio deittico a un elemento esterno al periodo, mentre la congiunzione *appena* in (41) aggiunge valore probabilistico alla locuzione epistemica. La catena contenuta in (42) è caratterizzata dalla congiunzione avversativa *però* che, rafforzata dalla locuzione avverbiale *lo stesso*, ha valore di aggiunta di informazione, mettendo in relazione l'asserzione precedente con la proposizione successiva.

Nel corpus in esame sono presenti diverse locuzioni complesse (cfr. parr. 3 e 4) che tendono ad assumere funzioni molteplici: in generale queste sembrano avere valore di modulazione ma risultano comunque funzionali per la strutturazione del messaggio.

Alcune di queste espressioni fungono da vere e proprie formule di attenuazione dell'enunciato. Si tratta molto spesso di atti di domanda espliciti costituiti da frasi interrogative indirette, come negli esempi seguenti:

(43) *Mi chiedevo se* mi sto rivolgendo alla persona giusta.

(44) *Volevo chiederle se* il periodo di svolgimento del corso sarà successivo al giorno Giovedì 24 aprile.

(45) *Vorrei sapere se* è confermato l'inizio del corso di morfosintassi per domani dalle otto e trenta alle dieci e trenta.

(46) *Avrei bisogno di sapere se* la tesina finale va consegnata in segreteria qualche giorno prima del 30.

Il valore interrogativo è stabilito dalla frase principale costruita con un verbo illocutorio di richiesta di informazione alla prima persona e dall'introduttore di subordinazione *se*. Si è di fronte a indicatori formulaici che, da un punto di vista pragmatico, servono anche da focalizzatori e da demarcatori, in quanto da un lato puntano l'attenzione sul contenuto della frase subordinata e, dall'altro, segnalano, posti in apertura di enunciato, la pianificazione del discorso.

Sulla funzione metatestuale di tipo focalizzante si collocano perifrasi costruite insieme all'imperfetto indicativo del verbo servile *volere*, come in (47)-(49):

(47) *Volevo chiederle* informazioni circa la verbalizzazione dei CFU.

(48) *Volevo inoltre sapere* come si svolgeranno le lezioni online.

(49) *La volevo avvisare* che io venerdì potrò venire a verbalizzare dopo le 9 e 30.

Negli esempi riportati sopra, ci troviamo di fronte a un imperfetto attenuativo o, meglio ancora, di cortesia: "il parlante, per una sorta di pudore o (a seconda dei casi) di deferenza verso l'interlocutore si comporta come se tale intenzione non valesse più al momento dell'enunciazione" (Bertinetto, 2001: 82). Questa modalità di interazione è utilizzata dagli attanti per mantenere costante, data l'asimmetria dei rapporti, l'equilibrio del goffmaniano gioco di faccia (Goffman, 1988).

6. *Alcune riflessioni*

Le possibilità offerte dalle nuove tecnologie e dalle nuove applicazioni per le comunicazioni⁵ stanno generando una proliferazione di testi *trasmessi* prodotti da tutte le classi di parlanti. Si potrebbe dire che i nuovi media consentono una comunicazione democratica per la quale chiunque può elaborare testi e divulgarli nella rete.

Antonelli (2014: 549-550) sostiene l'idea di un "mutamento varietistico" sottolineando che questa nuova testualità può essere considerata non come un'innovazione del sistema, bensì come un movimento più o meno distante dalla norma linguistica: "come già per l'italiano dell'uso medio, anche in questo caso siamo di fronte a usi non nuovi in sé - né esclusivi di questa varietà -, ma presenti con una diversa concentrazione e soprattutto in diverse condizioni di co-occorrenza".

Per questa ragione, la dimensione variazionale, quella sociolinguistica e quella pragmatica, sembrano le chiavi di lettura più appropriate per l'analisi di questo ambito della comunicazione che, data la concisione, la fluidità e la frammentazione, sembra guidare i processi di innovazione linguistica.

A oggi, le e-mail rappresentano un canale di trasmissione, non ancora obsoleto, con cui una buona parte dei fruitori di internet ha familiarità e dimestichezza. La funzione di questo canale è molteplice: può essere utilizzato per trasmettere file e link, per scambiare informazioni, per fare inviti, per salutare o anche per pubblicizzare prodotti. Questa funzionalità composita determina generi testuali diversi. In particolare, il genere dialogico sembra essere quello prevalente in quanto, sia pure in modalità asincrona, i parlanti richiedono e danno informazioni producendo atti linguistici funzionali ai propri obiettivi comunicativi.

I dispositivi utilizzati nella produzione di questi testi sono anch'essi plurimi. In questo studio, che rappresenta un percorso di ricerca nella sua fase iniziale ma dagli sviluppi certamente significativi, ci si è occupati in particolare delle relazioni di connessione a livello frasale segnalate dai connettori pragmatici (come in Palermo, 2013).

Come si è avuto modo di vedere, nelle e-mail la situazione, l'intenzione comunicativa e l'intertesto sono elementi fortemente connessi. A mettere in relazione questi fattori della testualità contribuiscono alcuni connettivi pragmatici che, sebbene non siano accompagnati da quegli elementi paralinguistici tipici degli usi parlati, si ritrovano anche nelle forme testuali della CMC e, come nel caso in esame, anche nei messaggi di posta elettronica.

Si è osservato dunque che, data la natura grafica del testo, prevale certamente l'attenzione alla distribuzione delle informazioni legata ai connettori con valore demarcativo (es. *ma, comunque, infatti, innanzitutto*) che segnalano l'introduzione, la prosecuzione e il termine delle diverse unità tematiche. Tuttavia, pragmaticamente e linguisticamente rilevante è l'utilizzo di espressioni giuntive con funzione di modu-

⁵ Si vedano a tal proposito le rilevazioni EURISPES (2012), CENSIS (2013), ISTAT (2014) già ampiamente discusse in Antonelli (2014). Le fonti concordano nel sostenere che per i cosiddetti "nativi digitali" i canali di comunicazione privilegiati sono i telefoni cellulari e gli smartphone; l'interazione passa attraverso le applicazioni online e i social network, con un conseguente uso limitato delle e-mail.

lazione. Si pensi a connettori come *gentilmente, possibilmente, se possibile, purtroppo, scusi se*, i quali, in direzione attenuativa (a volte anche focalizzante), contribuiscono a identificare ed evidenziare l'asimmetria e il potere sociale fra gli interlocutori. Risulta evidente la volontà di stabilire un clima positivo per poter garantire la circolarità delle informazioni espresse. La funzione che i connettivi stabiliscono è, allora, fortemente interattiva e il loro uso fa sì che gli atti prodotti siano di tipo implicito, regolando lo sviluppo degli enunciati attraverso formulazioni di cortesia basate sulla faccia positiva e sulla prevenzione di possibili conflitti.

Va dunque evidenziato che la coordinazione è certamente una categoria sintattica ma, come dimostrato, è in grado di veicolare funzioni e usi pragmatici utili, per un verso, a pianificare la concatenazione tematica degli enunciati (segnalando focus argomentativi o dichiarativi, relazioni oppositive tra le frasi e ancora passaggi, sospensioni o aperture ad altro argomento) e, per l'altro, a mantenere buone relazioni fra i partecipanti, assicurandone la cooperazione e l'empatia reciproca e rispettandone i ruoli.

Le scelte linguistiche operate dai giovani parlanti risultano altamente funzionali: in un contesto profondamente connotato sul piano diamesico, tematico e delle relazioni sociali, solo considerando le variabili pragmatiche e quelle più strettamente linguistiche (mettendo cioè in relazione le funzioni comunicative e le strutture morfosintattiche) è stato possibile individuare alcuni caratteri ed esigenze di un genere testuale ancora in allestimento, ma che potrebbe sviluppare una certa influenza e tendenze innovative anche su altre aree dell'italiano.

Bibliografia

- ANTONELLI G. (2014), L'e-taliano: una nuova realtà tra le varietà linguistiche italiane?, in GARAVELLI E. - SUOMELA-HÄRMÄ E. (a cura di), *Dal manoscritto al web. Canali e modalità di trasmissione dell'italiano. Tecniche, materiali e usi nella storia della lingua*, vol. II, Franco Cesati Editore, Firenze: 537-556.
- BAZZANELLA C. (1985), L'uso dei connettivi nel parlato. Alcune proposte, in FRANCHI DE BELLIS A. - SAVOIA M.L. (a cura di), *Sintassi e morfologia della lingua italiana d'uso*, Bulzoni, Roma: 83-93.
- BAZZANELLA C. (2003), Nuove forme di comunicazione a distanza, restrizioni contestuali e segnali discorsivi, in MARASCHIO N. - POGGI SALANI T. (a cura di), *Italia linguistica anno mille. Italia linguistica anno duemila*, Bulzoni, Roma: 403-415.
- BAZZANELLA C. (2007), *Linguistica e pragmatica del linguaggio*, Laterza, Bari.
- BAZZANELLA C. (2010), I segnali discorsivi, in SALVI G. - RENZI L. (a cura di), *Grammatica dell'italiano antico*, vol. II, Il Mulino, Bologna: 1339-1357.
- BARON N. (2000), *Alphabet to email. How written English evolved and where it's heading*, Routledge, New York/London.
- BERRUTO G. (2004), *Prima lezione di sociolinguistica*, Laterza, Bari.
- BERRUTO G. (2005), Italiano parlato e comunicazione mediata dal computer, in HÖLKER K. - MAASS C. (a cura di), *Aspetti dell'Italiano parlato*, Lit Verlag, Münster: 137-156.

- BERTINETTO P.M. (2001), Il verbo, in RENZI L. - SALVI G. - CARDINALETTI A. (a cura di), *Grande grammatica italiana di consultazione*, vol. II, Il Mulino, Bologna: 13-161.
- CENSIS (2013), *Gli under 30 protagonisti dell'evoluzione digitale della specie* (sintesi per la stampa dell'*Undicesimo rapporto Censis/Ucsi sulla comunicazione. L'evoluzione digitale della specie*). Disponibile in: http://www.censis.it/7?shadow_comunicato_stampa=120930 (data ultimo accesso 24 agosto 2015).
- COLOMBO A. (2012), *La coordinazione*, Carocci, Roma.
- DRESSLER W.U. - DE BEAUGRANDE R.A. (1994), *Introduzione alla linguistica testuale*, Il Mulino, Bologna (ediz. orig. 1981).
- DÜRSCHIED C. - FREHNER C. (2013), Email communication, in HERRING C.S. - STEIN D. - VIRTANEN T. (eds.), *Pragmatics of computer-mediated communication*, De Gruyter, Berlin: 35-54.
- EURISPES (2012), *Indagine conoscitiva sulla condizione dell'infanzia e dell'adolescenza in Italia 2012*. Disponibile in: http://www.azzurro.it/sites/default/files/Sintesi%20Indagine%20conoscitiva%20Infanzia%20Adolescenza%202012_2.pdf (data ultimo accesso 24 agosto 2015).
- FIorentino G. (2010), Forme di scrittura in rete: dal web 1.0 al web 2.0, in APRILE M. (a cura di), *Lingua e linguaggio dei media*, Aracne, Roma: 193-206.
- GOFFMAN E. (1988), *Il rituale dell'interazione*, Il Mulino, Bologna.
- GUALDO R. - RAFFAELLI L. - TELVE S. (2014), *Scrivere all'università. Pianificare e realizzare testi efficaci*, Carocci, Roma.
- HELD G. - SCHWARZE S. (a cura di), (2011), *Testi brevi. Teoria e pratica della testualità nell'era multimediale*, Peter Lang, Frankfurt am Main.
- ISTAT (2014), *Cittadini e nuove tecnologie*. Disponibile in: <http://www.istat.it/it/archivio/143073> (data ultimo accesso 24 agosto 2015).
- JANSEN E. (2002), *NetLingo. The internet dictionary*, Netlingo Inc, Ojai.
- LORENZ C. (2007), The death of e-mail. Teenagers are abandoning their Yahoo! and Hotmail accounts. Do the rest of us have to?, in *Slate Magazine*, November 14th. Disponibile in: <http://www.slate.com/id/2177969> (data ultimo accesso 24 agosto 2015).
- MILROY J. - MILROY L. (1985), *Authority in language. Investigating language prescription and standardization*, Routledge, New York/London.
- O'REILLY T. (2005), *What is web 2.0. Design patterns and business models for the next generation of software*. Disponibile in: <http://www.oreilly.com/pub/a/web2/archive/what-is-web-20.html> (data ultimo accesso 24 agosto 2015).
- PALERMO M. (2013), *Linguistica testuale dell'italiano*, Il Mulino, Bologna.
- SABATINI F. - CAMODECA C. - DE SANTIS C. (2011), *Sistema e testo. Dalla grammatica valenziale all'esperienza dei testi*, Loescher editore, Torino.
- SABATINI F. (2012), L'italiano dell'uso medio: una realtà tra le varietà linguistiche italiane, in COLETTI V. - COLUCCIA R. - D'ACHILLE P. - DE BLASI N. - PROIETTI D. (a cura di), *L'italiano nel mondo moderno. Saggi scelti dal 1968 al 2009*, vol. II, Liguori, Napoli: 3-36.
- SERIANNI L. (2007), *Prima lezione di grammatica*, Laterza, Bari.

SCORRETTI M. (2001), Le strutture coordinate, in RENZI L. - SALVI G. - CARDINALETTI A. (a cura di), *Grande grammatica italiana di consultazione*, vol. I, Il Mulino, Bologna: 241-284.

SKOVHOLT K. - SVENNEVIG J. (2013), Responses and non-responses in workplace emails, in HERRING C.S. - STEIN D. - VIRTANEN T. (eds.), *Pragmatics of computer-mediated communication*, De Gruyter, Berlin: 589-611.

TEMPESTA I. (2006), Linguaggio dei giovani o lingua giovane? Quale rapporto fra l'italiano dei giovani e il repertorio, in MARCATO G. (a cura di), *Giovani, lingue e dialetti*, Unipress, Padova: 33-42.

TEMPESTA I. (2010), Dialetto e italiano in azione. I segnali discorsivi nel repertorio (meridionale) contemporaneo, in RUFFINO G. - D'AGOSTINO M. (a cura di), *Storia della lingua italiana e dialettologia*, Centro di Studi Filologici e Linguistici Siciliani, Palermo: 712-728.

La comunicazione tra studenti e docenti via forum e e-mail: strategie di cortesia

This study discusses the role of politeness in computer mediated communication between students and teachers in an English Linguistics course held in 2012-2014 within a BA in Foreign Languages and Literatures at the University of Bologna. E-mails and forum posts on an e-learning platform are assembled in an electronic corpus and compared, in order to identify recurrent patterns of linguistic behaviour, with respect to speech acts, modality, grammatical metaphors, and especially politeness. As students belong to 16 different nationalities, patterns of intercultural variation are also considered. The data show that students deployed politeness strategies competently and were aware of the specificity of these two electronic media: this emerges from their use of modal verbs, external modification, interpersonal grammatical metaphors, terms of address and degree of directness.

1. *Introduzione: strategie di cortesia nella comunicazione docente-studenti*

L'utilizzo delle tecnologie informatiche applicate all'apprendimento linguistico è un fenomeno sempre più diffuso e oggetto di studi che, già da diversi anni e da diverse prospettive, prendono in considerazione l'effetto degli strumenti informatici sul rapporto docente-discente (Accietto - Zorzi, 1998; Thorne, 2008; Eslami *et al.*, 2015). Gli studi più recenti si concentrano sulla comunicazione informatica, in particolare sulla posta elettronica, proponendo un connubio tra analisi del discorso e pragmatica interculturale per esaminare il ruolo di questo mezzo di interazione in contesti accademici (Biesenbach-Lucas, 2007; Economidou-Kogetsidis, 2011). Le ricerche in questo campo si estendono anche ad altre modalità di comunicazione, specialmente asincrona, quali i forum di discussione all'interno delle piattaforme di *e-learning* più diffuse in ambito universitario (Miyazoe - Anderson, 2010), nonché i *social network* (Aydin, 2014). Tra gli scopi principali di queste indagini si colloca l'analisi dell'efficacia di tali strumenti per la promozione dell'apprendimento collaborativo e di una conoscenza partecipata (Popescu, 2012). Grande rilevanza assume anche l'analisi delle strategie di aggiustamento e compensazione attuate per far fronte alla mancanza di tratti paralinguistici nel mezzo informatico: aspetto, questo, centrale nella recente area di ricerca della ciberpragmatica (Yus, 2011). Va comunque rilevato come gran parte dei lavori sulla comunicazione informatica tra studenti e docenti a livello universitario riguardi attualmente contesti socio-culturali diversi da quello italiano. Il presente studio si colloca in tale panorama di ricerche, propo-

¹ Università di Bologna.

nendosi di approfondire le strategie di cortesia attuate dagli studenti universitari italiani nel corso di scambi mediati dal computer quali e-mail e forum.

Si presentano i risultati dell'analisi di un corpus contenente e-mail e messaggi asincroni sul forum integrato nella piattaforma di *e-learning* Moodle (Dougiamas - Taylor, 2003): i testi risalgono agli anni accademici 2012-2013 e 2013-2014 e sono stati prodotti da studenti e docenti di un corso universitario di Lingua e Linguistica inglese a livello triennale. Il corpus si articola in due sottocorpora, l'uno contenente le e-mail, l'altro dedicato ai messaggi su piattaforma *e-learning*. È doveroso precisare che, allo stato attuale, il corpus ha dimensioni ridotte, essendo ancora in costruzione: è in programma, infatti, l'inserimento graduale di testi degli anni accademici successivi. Tuttavia, nonostante le limitazioni dal punto di vista quantitativo, l'analisi evidenzia alcune caratteristiche dell'uso delle strategie di cortesia, nei diversi mezzi informatici presi in esame, che risultano meritevoli di approfondimento.

Lo studio si divide in quattro sezioni. Dapprima, si descrivono il corpus e i partecipanti, per nazionalità, genere, corso di laurea, livello di competenza linguistica in inglese e grado di "natività digitale" (Bennett *et al.*, 2008; Bennett - Maton, 2010). La seconda e terza sezione affrontano rispettivamente il sottocorpus contenente le e-mail e quello costituito dai messaggi sul forum di Moodle, con particolare attenzione alle strategie di cortesia in relazione alla natura asimmetrica della comunicazione (Brown - Levinson, 1987), quali si evincono ad esempio da formule allocutive, espressione della modalità e metafore grammaticali di tipo interpersonale (Halliday - Matthiessen, 2004; Thompson, 2014). Infine, si espongono le conclusioni dello studio, anche in merito alle somiglianze e differenze tra e-mail e messaggi sul forum di un *learning management system* (Moodle), correlate almeno in parte alla loro diversa natura in ordine allo scopo e contesto di utilizzo, nonché alla competenza linguistica e pragmatica degli studenti².

2. Corpus e partecipanti

Gli studenti sono iscritti al primo anno dei corsi di laurea triennale (Lingue e Letterature Straniere e Lingue, Mercati e Culture dell'Asia) attivati presso la Scuola di Lingue, Letterature, Traduzione e Interpretazione dell'Università di Bologna. Tra i requisiti di ammissione ad ambedue i corsi figura l'accertamento del livello B1 del Quadro Comune Europeo delle Lingue, tramite l'*Oxford Online Placement Test* (OOPT), a cui tutti gli studenti devono sottoporsi se intendono inserire la lingua inglese nel proprio piano di studi. Dai dati raccolti nell'anno accademico 2013-2014, risulta che la maggioranza degli studenti (71%) ha un livello di competenza superiore, con una preponderanza di B2 (47,6% del totale). Il corso di Lingua e Linguistica inglese 1, cui fa riferimento questo studio, è attivato sia in modalità tradizionale, con 30 ore di didattica frontale in presenza, sia in modalità di *blended*

² Le sezioni 1 e 4 si devono a Luporini; le sezioni 2 e 3 sono opera di Fusari; le conclusioni (sezione 5) sono frutto del lavoro congiunto delle autrici.

e-learning, con tre incontri in presenza e circa 24 ore di didattica a distanza sulla piattaforma di *e-learning* Moodle. Ciascuno studente è libero di scegliere con quale modalità frequentare il corso. Malgrado la giovane età degli studenti (tra 19 e 23 anni, con una sola studentessa “matura”, ultracinquantenne), che li qualificerebbe come nativi digitali, ben il 96% sceglie di seguire il corso in modalità tradizionale, per ragioni in gran parte logistiche, dal momento che il corso *e-learning* è disponibile solo al secondo semestre e pertanto il relativo esame va posticipato fino alla sessione estiva (Fusari, in corso di stampa). A eccezione di due studentesse britanniche, l'inglese è una lingua straniera per i partecipanti, che appartengono alle seguenti nazionalità esposte in ordine alfabetico: bulgara, cinese, estone, georgiana, giordana, italiana, lituana, marocchina, polacca, portoghese, romena, russa, spagnola, tedesca e ucraina. Nel sottocorpus contenente le e-mail, sono rappresentate quasi tutte queste nazionalità, mentre in quello relativo alla piattaforma di *e-learning* Moodle, come vedremo alla sezione 4, prevalgono gli studenti italiani. Gli studenti non italiani rappresentati all'interno del sottocorpus di e-mail sono per il 71,4% Erasmus, per il 17,9% residenti in Italia e interni all'Ateneo bolognese, per il 7,1% Erasmus Mundus e per il 3,6% appartengono al cosiddetto “contingente extracomunitari”. È di genere femminile l'82% del campione complessivo, in linea con il profilo demografico generale dei due corsi di laurea³.

Le insegnanti sono rispettivamente una docente a contratto e una docente di seconda fascia, ambedue addottorate nel Settore Scientifico Disciplinare di Lingua e Traduzione Inglese (L-LIN/12), di madrelingua italiana e dotate di una competenza linguistica in inglese a livello C2. Benché entrambe abbiano al proprio attivo una formazione specifica in materia di *e-learning*, solo la prima può considerarsi nativa digitale, in quanto anagraficamente più giovane (Bennett - Maton, 2010). Il corpus, codificato in .txt e analizzato con Antconc 3.2.4, è privo di annotazioni grammaticali, che si prevede di aggiungere in una fase successiva del progetto, non solo per le parti del discorso, ma anche per ulteriori metadati, avvalendosi del programma UAM Corpus Tool 3 (O'Donnell, 2009).

3. *E-mail*

Il sottocorpus contenente le e-mail si riferisce al corso in presenza e contiene tutti i messaggi di posta elettronica scritti in inglese dagli studenti alla docente, con le relative risposte. Il numero dei messaggi e degli studenti, suddivisi per nazionalità, è illustrato in tab. 1, dove si nota anche un'eccezione, contrassegnata dal simbolo Ø: si tratta dell'unico caso nel corpus in cui è la docente a contattare una studentessa, che non risponde.

³ I dati demografici relativi alla popolazione studentesca dei due corsi di studio, attivi presso l'Ateneo di Bologna, sono reperibili nei rispettivi Rapporti di *Quality Assurance*: <http://www.unibo.it/QualityAssurance/Reports2014/Report-0979-2014.pdf> e <http://www.unibo.it/QualityAssurance/Reports2014/Report-0980-2014.pdf>.

Tabella 1 - Numero degli studenti e delle e-mail da loro inviate alle docenti

<i>Nazionalità</i>	<i>n. studenti</i>	<i>n. e-mail</i>
Britannica	2	5
Bulgara	1	1
Cinese	1	1
Estone	1	9
Georgiana	1	4
Giordana	1	2
Italiana	3	5
Lituana	3	5
Marocchina	1	∅
Polacca	2	9
Portoghese	6	21
Romena	3	5
Spagnola	1	2
Tedesca	1	1
Ucraina	1	2
TOTALE	28	72

Complessivamente, il sottocorpus contiene 5248 parole scritte dagli studenti sotto forma di e-mail, la cui lunghezza media ammonta a 72,9 parole. La più frequente è la prima persona soggetto, *I*, che introduce solitamente l'esposizione di un problema individuale per risolvere il quale si richiede l'aiuto della docente: la fraseologia più frequente contenente tale pronome è in effetti *I want/wanted/would (like)*, che prelude alla formulazione di un atto direttivo. Dal canto suo, la docente ha scritto e-mail della lunghezza media di 110 parole, per un totale di 6817 vocaboli all'interno del sottocorpus in esame, il più frequente dei quali è *the*, come tipico della maggioranza dei corpora in inglese. La maggiore lunghezza delle e-mail della docente si trova altresì in parziale contraddizione con Bloch (2002), Chen (2006) e Shim (2013), secondo cui la lunghezza dei messaggi di posta elettronica è inversamente proporzionale al potere sociale esercitato da chi li scrive. In questo sottocorpus, avviene esattamente il contrario: la docente, infatti, fornisce risposte dettagliatissime e spesso addirittura esplicita la natura dell'atto linguistico (rappresentativo o direttivo) che va compiendo, affermando che una determinata azione è *my advice/information* oppure *obligatory for you/all students to do*. Va notato però che la maggiore lunghezza dei messaggi della docente, sia per e-mail, sia su Moodle, come vedremo alla sezione 4, non è unicamente imputabile a un divario sociale con gli studenti, ma risente almeno in parte anche della sua maggiore competenza linguistica in L2.

3.1 Gli atti linguistici

Nel presente lavoro (in linea con altri di analogo argomento, ad esempio Merrison *et al.*, 2012), gli studenti compiono in gran parte atti direttivi, attivando le funzioni

discorsive (secondo il modello di Halliday - Matthiessen, 2004: 107-111) “domanda” e “comando” per chiedere un appuntamento, ottenere chiarimenti su esami e problematiche amministrative, farsi inoltrare materiali di studio/esercizio ecc. Il contenuto generale delle e-mail degli studenti è riassunto in tab. 2.

Tabella 2 - *Contenuto delle e-mail degli studenti*

<i>Contenuto e-mail</i>	<i>n. e-mail</i>
Richiesta informazioni esami	32
Richiesta informazioni corso	19
Ringraziamento	10
“Crisi da bocciatura”	6
Richiesta materiali	2
Richiesta appuntamento	2
Richiesta referenze	1

La “crisi da bocciatura” cui si fa riferimento in tab. 2 rappresenta un evento purtroppo piuttosto diffuso nella carriera degli studenti Erasmus. Per alcuni, infatti, non è del tutto chiaro il fatto che, nel sistema universitario italiano, è necessario superare l’esame finale di ciascun insegnamento per cui si intendono acquisire crediti formativi in base al proprio *Learning Agreement*: costoro, di fronte a un esito negativo di tale esame, si mostrano convinti di poter acquisire un pacchetto di crediti, magari inferiore al totale previsto, in virtù delle lezioni frequentate o di eventuali *portfolio* di esercizi o altre prove di competenza, svolti durante l’anno accademico, che documentino, se non il profitto, almeno l’impegno profuso. Scoprendo che ciò non è possibile, nei sei casi di cui alla tab. 2, gli studenti coinvolti chiedono alla docente di trovare un modo per aggirare le regole, narrando con grande emotività e con dovizia di particolari le nefaste conseguenze di un loro rientro in patria con meno crediti del dovuto. Tale richiesta viola palesemente svariate massime conversazionali, *in primis* quelle che Leech (1983: 104-139) definisce “massima del tatto” e “massima della simpatia”, arrivando talvolta a un conflitto vero e proprio, dove la docente è esplicitamente accusata di non aver letto o compreso la gravità delle conseguenze che lo studente sconterebbe rientrando dall’Erasmus senza i crediti previsti (es. “I wanted to know if you took that in account when you suggested me those lessons in the previous email?”). Paradossalmente, nel tentativo di suscitare simpatia (nel senso etimologico del termine) nella docente, affinché gli venga incontro attribuendogli un voto superiore a quello conseguito, lo studente le chiede di contravvenire alle regole sancite dalla propria istituzione, in base a cui non è possibile ritoccare i voti al rialzo, se non sostenendo nuovamente il relativo esame. Tale conflittualità non si può però liquidare come esempio di scorrettezza individuale, né interpretare come indizio di una maggiore propensione da parte di alcune culture ad aspettarsi dai docenti la volontà e possibilità di eludere le norme istituzionali: generalizzazioni di questo tipo sono infatti del tutto inadeguate a spiegare il comportamento prag-

matico degli apprendenti (House, 1999: 84; McKay, 2002: 75-77). Le motivazioni sono da ricercare piuttosto nel complesso dei fattori contestuali che definiscono la situazione comunicativa e che comprendono le convenzioni linguistiche e culturali dei paesi di origine, ma non si esauriscono in esse (Poppi, 2012: 179). Questo tipo di evento rappresenta, in effetti, un incidente pragmatico, legato alle aspettative che gli studenti si formano in merito all'acquisizione dei crediti formativi universitari, in base sia a quanto accade nel proprio paese di origine, sia a quanto viene loro illustrato (o a come comprendono quanto loro illustrato) durante le riunioni informative organizzate sia negli Atenei di origine, sia nelle sedi ospitanti. È pertanto un problema più vasto, che spetta a tutti gli enti coinvolti nell'Erasmus risolvere, con una maggiore attenzione al rispetto delle massime della qualità e della quantità nell'esecuzione degli atti linguistici rappresentativi volti a informare gli studenti delle modalità di conseguimento di crediti universitari nei diversi paesi membri dell'UE. Sarebbe auspicabile, ad esempio, che gli opuscoli informativi e gli incontri di orientamento destinati agli studenti Erasmus, sia in entrata, sia in uscita, si soffermassero in modo specifico sulle regole che disciplinano l'acquisizione dei crediti nei diversi paesi e anche nei diversi Atenei interessati dagli scambi. I dati del presente studio dimostrano, infatti, che non è sufficiente demandare questo compito informativo ai singoli docenti, poiché la puntigliosa descrizione delle modalità di esame da parte della docente, nelle e-mail antecedenti la fase di valutazione, ha comunque lasciato spazio a ben sei casi di fallimento pragmatico.

Si è già notato che la maggioranza degli atti è di natura direttiva, e l'analisi delle e-mail permette di rilevare che gli studenti ricorrono a svariate strategie per mitigare la minaccia della faccia intrinseca in questi atti (Economidou-Kogetsidis, 2011: 3206). La modificazione esterna (Blum-Kulka *et al.*, 1989) è attuata in special modo attraverso le scuse, talvolta espresse al termine di lunghi preamboli (es. "As you know, I joined your course after 2 weeks as lectures were started. I do not know what I have to do"), mentre la modificazione interna si avvale soprattutto di soggettivizzatori (es. "I wonder if, I understand if, I think it would be interesting to"), di aggiunte modali di commento (es. *unfortunately, hopefully*), di mitigatori (es. *it seems that, rather, quite, or something*), nonché dell'uso di *please* quale complemento all'imperativo, volto a invocare la cooperazione dell'interlocutore (Blum-Kulka *et al.*, 1989: 283). Tale uso di *please*, tuttavia, risulta meno frequente rispetto alle attese, con sole 11 occorrenze: gli studenti, almeno implicitamente, sono pertanto consapevoli che, pur lasciando un certo spazio di libertà all'interlocutore, *please* non è di per sé sufficiente a mitigare la forza impositiva dell'imperativo (Hartford - Bardovi Harlig, 1996: 59)⁴.

⁴ Sul rischio di fallimento pragmatico legato a un uso indiscriminato di *please* quale mitigatore dell'imperativo nella comunicazione studente-docente, si vedano Biesenbach-Lucas (2007), Economidou-Kogetsidis (2011) e Shim (2013).

3.2 I verbi modali

Anche l'uso dei verbi modali riveste un ruolo importante per l'esecuzione degli atti direttivi da parte degli studenti. In particolare, i dati in tab. 3 mostrano che gli studenti sanno utilizzare in modo abbastanza efficace la metafora grammaticale interpersonale (Thompson, 2014: 246-251), ad esempio formulando un direttivo sotto forma di domanda, allo scopo di realizzare un atto linguistico indiretto.

Tabella 3 - *Verbi modali nelle e-mail degli studenti*

<i>Modale</i>	<i>Frequenza</i>	<i>Esempi</i>
can	33	Please, can you help me?/Which advices can you give me?/Can we discuss my issue?
would	29	I would like to know/I would be extremely grateful if you
could	21	Could you let me know?/Could you add us to the list?/I was wondering if we could change the date
will	21	I hope that will fix all I need/I will greatly appreciate your help
have/need to	20	I have/need to register/submit/check
should	13	I do not know which exam/group should I choose
be able to	3	We want to ask whether we are able to/We are not able to/Will you be able to?
may	3	How this question may be solved for me?/May I sign up for verbalizzazione?/I wonder if you may tell me
must	2	Truly I must say that this system is a complete maze/I attach you the blank which must be sent directly at the e-mail address

Malgrado alcuni errori grammaticali (es. il plurale *advices*, la mancata inversione dell'ausiliare nel modo interrogativo e l'uso deontico di *to be able to*), la competenza nella scelta dei modali più adatti al contesto si rivela generalmente elevata. Ne è una prova la rarità di *must*, in linea con Leech *et al.* (2009), che ne attestano il drastico declino nell'inglese contemporaneo: in questo sottocorpus, le uniche due occorrenze di *must* sono riferite, in senso deontico (in questo sottocorpus, *must* non è mai utilizzato in senso epistemico), ad adempimenti burocratici. La docente stessa usa *must* solo quattro volte, preferendogli altre espressioni deontiche, in particolare quelle perifrastiche *it is essential*, *it is obligatory*, che codificano il direttivo come un fatto, dando luogo a un altro tipo di metafora grammaticale interpersonale (Halliday - Matthiessen, 2004: 616), nella quale la modalità è espressa attraverso una proiezione ipotattica.

3.3 I ringraziamenti e gli allocutivi

Tornando alla tipologia di atti linguistici, va notato che gli unici atti apparentemente non direttivi compiuti dagli studenti, ossia i ringraziamenti, celano a loro volta, nella maggioranza dei casi, un direttivo. L'espressione *thank/s*, infatti, è quasi sempre usata a sostegno dello scopo fondamentale dell'e-mail, che consiste in una richiesta⁵. I ringraziamenti veri e propri, oltre che rari, sono anche brevissimi, benché la maggioranza (80%) sia corredata di intensificatori (es. *very, so much, ever so much, a lot*) e alcuni ricorrono a strategie di cortesia positiva, basate sulla solidarietà, in cui lo studente si pone non come un soggetto subalterno, bisognoso dell'aiuto della docente, ma come un interlocutore che desidera abbattere le gerarchie per mostrare la propria consapevolezza delle esigenze e dei diritti della docente. Ciò si evidenzia in modo particolare in una e-mail, scritta dall'unica studentessa "matura" rappresentata in questo studio, che correda il suo ringraziamento con un augurio alla docente per la pronta guarigione di un membro della sua famiglia, per assistere il quale la docente si era assentata dal lavoro per un giorno. Questa strategia, definita *sweetener* da Blum-Kulka *et al.* (1989), è tipicamente evitata nella comunicazione asimmetrica (Chejnová, 2014: 185), ma in questo contesto il rischio pragmatico corso dalla studentessa si motiva con la sua età e con il desiderio, già manifestato precedentemente dalla docente, di creare un ambiente didattico partecipato, pur nel rispetto delle differenze sociali e di ruolo (Merrison *et al.*, 2012: 1094) che trova conferma nelle formule allocutive in tab. 4, nonché in italiano nell'uso reciproco del pronome allocutivo di cortesia "lei". Il grado di formalità della comunicazione è, del resto, solitamente inferiore nelle e-mail rispetto ad altre forme di interazione, sia sincrona sia asincrona, come ebbe a notare Baron (1998), molto prima che la comunicazione on-line conoscesse la capillare diffusione odierna.

Tabella 4 - *Formule allocutive nelle e-mail degli studenti*

<i>Allocutivo</i>	<i>Frequenza</i>
Dear Dr/Prof. X	17
Good morning/afternoon/evening	13
Dear Ms/Mrs X	11
Dear Nome/Hello	5
Dear Nome Cognome	5
Dear teacher	3
∅ (solo in scambi che si compongono di più e-mail)	18

A conclusione di questa sezione, va notato che alcuni scambi di e-mail sono risultati pragmaticamente infelici, fino a raggiungere il fallimento comunicativo in tre casi,

⁵ Lo stesso avviene in Chejnová (2014), secondo cui tale ringraziamento non incide positivamente sul livello di cortesia, ma rappresenta piuttosto una pressione, dando per scontato che la docente ottempererà alla richiesta.

in cui la docente o gli studenti non hanno risposto all'e-mail loro rivolta. Nel primo caso, già menzionato in tab. 1, la docente comunica a una studentessa Erasmus Mundus che non ha superato l'esame: la studentessa, anziché evocare lo scenario comunicativo testé definito "crisi da bocciatura", sparisce letteralmente nel nulla, non facendosi più vedere né sentire, malgrado i numerosi tentativi di reperirla da parte della docente. In questo caso, è probabile che la studentessa abbia interpretato la comunicazione della docente, invero piuttosto breve e diretta, come lesiva della massima del tatto (Leech, 1983: 104). Negli altri due casi, è la docente a interrompere il flusso della comunicazione: in un caso, le era stata richiesta una lettera di referenze, nell'altro l'invio di materiali affrontati in classe. Entrambe le e-mail erano brevissime, prive di firma e dotate di formule di cortesia puramente convenzionali (*could you, thank you*): ciò ha indotto la docente a ritenerle troppo dirette, nonché indicative di un lavoro pragmatico insufficiente in relazione alla natura dell'atto linguistico.

4. Messaggi sul forum di Moodle

Prendiamo ora in esame il sottocorpus contenente i messaggi scambiati tra studenti e docente, ma anche tra studenti e studenti, all'interno del forum presente su Moodle: si tratta di 10.858 parole, suddivise in quattro sezioni, come illustrato in tab. 5.

Tabella 5 - Numero dei messaggi e delle parole sul forum di Moodle

Sezione forum	n. messaggi	n. parole
Studenti 2012-2013	25	1073
Docente 2012-2013	15	1733
Studenti 2013-2014	70	3023
Docente 2013-2014	48	5029

Nell'anno accademico 2012-2013, gli studenti partecipanti al corso in modalità *blended e-learning* erano 14, tutti di nazionalità italiana, con una proporzione femmine/maschi di 13:1. In questa sezione del sottocorpus figurano 25 messaggi scritti dagli studenti, per un totale di 1073 parole e una lunghezza media per messaggio di 42,9 parole. Nell'anno accademico successivo, gli *e-learner* erano 13, di cui dieci italiani, due studentesse straniere interne all'Università di Bologna (di nazionalità rispettivamente polacca e russa) e una studentessa Erasmus portoghese; predomina anche in questo caso la componente femminile, con un rapporto di 11:2. In questa sezione del sottocorpus è però notevolmente maggiore il numero dei messaggi scritti dagli studenti, per un totale di 70, equivalenti a 3023 parole, con una lunghezza media di 43,2 parole. La differenza è sicuramente riconducibile al fatto che la docente non è la stessa: nel 2013-2014, infatti, è subentrata nella titolarità dell'insegnamento in *e-learning* una docente nativa digitale, il cui stile comunicativo *on-line*

si differenzia da quello della collega che l'ha preceduta nel 2012-2013, soprattutto per quanto riguarda la propensione a spronare esplicitamente gli studenti a interagire, non solo sottolineando l'obbligatorietà della scrittura di messaggi sul forum, ma anche valorizzandone la funzione euristica per la compartecipazione della classe virtuale alla soluzione di dubbi e problemi. Nei due casi, all'interno della sezione 2013-2014, in cui la partecipazione attiva degli studenti inizia a diminuire, è la docente a stimolarne la ripresa aprendo una nuova discussione ("Dear students, is there anything unclear about Unit 1? Or anything you would like to expand on?"), cosa che non fa mai la docente del 2012-2013: quest'ultima, 'immigrata digitale', secondo la classificazione di Bennett *et al.* (2008), tende piuttosto a valorizzare l'autonomia degli studenti, lasciandoli soli responsabili del forum, e dunque non intervenendo per riaccendere la discussione quando langue. Malgrado questa apprezzabile differenza, ambedue le docenti detengono un ruolo dominante nel forum, come si evince anche dal maggior numero di parole da loro scritte (15 messaggi nel 2012-2013, di una lunghezza media di 115,5 parole e un totale di 1733 parole; 48 messaggi nel 2013-2014, di una lunghezza media di 104,8 parole e un totale di 5029 parole) e dal loro ruolo di moderatrici, quindi, titolate a gestire il turno di parola e a orientare la discussione in una particolare direzione, evitando ad esempio il cosiddetto *off topic* (fuori tema). Concorre a motivare la lunghezza dei messaggi della docente, come già visto, anche il suo status di parlante esperta.

4.1 Il contenuto dei messaggi su Moodle

Diversamente da quanto accade per le e-mail (tab. 2), pare qui più opportuno analizzare il contenuto dei messaggi degli studenti sul forum attraverso una lista lemmatizzata delle parole più frequenti (*wordlist*) anziché individuando l'atto linguistico che viene compiuto: infatti, sul forum, non sempre è possibile istituire una corrispondenza univoca tra messaggio e atto linguistico, perché spesso avviene che più messaggi concorrano alla realizzazione del medesimo atto. Sotto questo aspetto, è senz'altro possibile ipotizzare che il sottocorpus relativo al forum di Moodle evidenzia una maggiore mimesi del parlato rispetto a quello contenente le e-mail. La *wordlist* in tab. 6 è stata generata con Antconc 3.2.4, applicando una lista di esclusioni (*stoplist*) per escludere le parole grammaticali più frequenti, non particolarmente rilevanti per l'oggetto di studio.

L'elevato livello di interattività e la dimensione dialogica caratteristica del forum si manifestano, in entrambi gli anni accademici, nella notevole frequenza non solo dei pronomi personali, ma anche delle espressioni di ringraziamento: una concordanza di *thank/s* ne evidenzia infatti l'utilizzo, ancora una volta, sia come supporto a un atto direttivo, sia come espressione di gratitudine non unicamente convenzionale, ma anche denotativa della reciproca soddisfazione rispetto allo scambio comunicativo (Aston, 1995). Le parole elencate in tab. 6 comprovano altresì il fatto che i forum *on-line*, anche in ambito didattico, vedono una riduzione della formalità e della distanza sociale non solo rispetto alle omologhe comunicazioni in presenza, ma anche rispetto alle e-mail, come si evince dall'uso di *hello*, *hi* e soprattutto degli *emo-*

icon, le cosiddette “faccine” che esprimono lo stato d’animo di chi scrive. Antconce le considera come parole, permettendone l’ingresso nella lista di frequenza, un fatto molto utile ai fini dell’analisi: si nota immediatamente, ad esempio, una profonda differenza nell’uso degli *emoticon* da parte degli studenti nei due anni accademici di riferimento. Nel 2012-2013, infatti, compaiono 26 *emoticon*, mentre ne figura solo uno nella sezione del sottocorpus relativa al forum del 2013-2014. Ancora una volta, a motivare questa difformità è probabilmente il diverso stile comunicativo delle docenti: la presenza meno assidua della docente sul forum del 2012-2013 ha forse indotto gli studenti a utilizzare uno stile meno controllato, con il conseguente ricorso a caratteristiche tipiche dei *social media* e degli sms (Zappavigna, 2012; Tagg, 2012). Gli *emoticon*, inoltre, sono assenti nel sottocorpus contenente le e-mail, il che lascia supporre che gli studenti non li considerino appropriati per comunicare con la docente in questa forma.

Tabella 6 - *Lista di frequenza lemmatizzata.*
Prime dieci parole usate dagli studenti sul forum di Moodle

2012-2013			2013-2014		
	<i>Frequenza</i>	<i>Lemma</i>		<i>Frequenza</i>	<i>Lemma</i>
1.	57	be	1.	153	I
2.	45	I	2.	149	be
3.	27	it	3.	67	you
4.	26	smiley	4.	64	have
5.	19	you	5.	60	it
6.	15	have	6.	40	thank
7.	13	thanks	7.	39	can
8.	12	do/theme	8.	34	do/we
9.	9	hello/topical	9.	26	like/would
10.	8	some	10.	23	hi

4.2 Gli allocutivi

Anche le formule allocutive usate dagli studenti si prestano a un paragone con quanto visto riguardo le e-mail.

In tab. 7, si nota infatti la preferenza per saluti di tipo informale, quali *bello* e *hi*; in due casi, compaiono le formule tipiche del parlato spontaneo *good afternoon/good morning* (Marques, 2008), abbastanza frequenti anche nelle e-mail, come già visto. Le regole contestuali della *netiquette* (le “buone maniere” da utilizzarsi *online*) sono comunque rispettate, come dimostra l’estrema rarità delle conversazioni che si aprono senza un saluto (due nel 2012-2013 e una nel 2013-2014). Nel corso dei due anni accademici oggetto di analisi, è altresì possibile individuare una forma di convergenza sociolinguistica sul tipo di allocutivi utilizzati: data la natura aperta e collaborativa del forum, gli studenti sono infatti portati a convergere progressi-

vamente sulle medesime forme di saluto, selezionando quale allocuzione preferita rispettivamente *bello* nel 2012-2013 e *hi* nel 2013-2014 (è da notare il fatto che, in quest'ultimo anno accademico, *bello* non viene mai usato). Anche il maggiore utilizzo dell'allocutivo formale *dear* nel 2013-2014 si può spiegare in termini di convergenza, poiché la docente nativa digitale ne fa un impiego più ampio rispetto alla collega dell'anno precedente: nel 2013-2014, su 48 messaggi della docente, 28 (58,3%) cominciano con *dear*, mentre nel 2012-2013 il dato è assai più contenuto (33,3%). È dunque probabile che gli studenti si siano adeguati, almeno in parte, agli allocutivi utilizzati dalla loro rispettiva docente. Infine, non si può sorvolare sul fatto che, nel 2012-2013, ben il 75% dei saluti di apertura *bello* e *hi* è seguito da un *emoticon* sorridente: si tratta, con ogni probabilità, di una strategia di cortesia positiva dalla duplice funzione di ridurre la distanza sociale e di preparare un atto direttivo, mitigando la minaccia della faccia (es. "Hello people! :-) I didn't understand very well the topic ... Can anyone help me?").

Tabella 7 - *Formule allocutive nei messaggi degli studenti sul forum di Moodle*

2012-2013		2013-2014	
<i>Allocutivo</i>	<i>Frequenza</i>	<i>Allocutivo</i>	<i>Frequenza</i>
Hello	9	Hi	23
Hi	3	Dear Professor	5
Dear all/Good morning	1	Good afternoon	1
∅	11 (di cui 2 a inizio conversazione)	∅	41 (di cui 1 a inizio conversazione)

4.3 I verbi modali

A conclusione di questa sezione, la tab. 8 illustra il tipo e la frequenza dei verbi modali nei messaggi degli studenti sul forum di Moodle.

Benché non sia possibile, in questa sede, una disamina esaustiva, sono comunque ravvisabili alcuni comportamenti ben definiti da parte degli studenti. Pur con alcune differenze quantitative, il dato già rilevato sulle e-mail relativo alla padronanza di una vasta gamma di verbi modali trova conferma nel sottocorpus di messaggi sul forum di Moodle: *ought to* e *shall* (anch'esso, come *must*, in forte declino, Leech *et al.*, 2009) non compaiono; *must* è assente nei messaggi degli studenti nel 2012-2013, compare una volta, in senso epistemico, nei messaggi scritti dalla docente in quell'anno, mentre figura esclusivamente come oggetto di considerazioni metalinguistiche nel 2013-2014. A riscuotere la maggiore frequenza relativa sono *can*, *would* e *will*, al punto che i primi due sono attestati anche nella *wordlist* in tab. 6. Anche in questo sottocorpus, tali verbi sono spesso usati per realizzare metafore grammaticali interpersonali, grazie alle quali si ottiene un atto linguistico indiretto formulando un imperativo tramite una (apparentemente) più negoziabile domanda.

Tabella 8 - Verbi modali nei messaggi degli studenti sul forum di Moodle

Modale	Frequenza 2012-2013	Frequenza 2013-2014	Esempi
can	4	39	Can anyone help me?/So, can we also say that the defining sentences are always embedded?
would	4	26	it'd be good to know if I made some kind of mistake/I would have explanations about the concept of social semiotic
have to	2	8	I want to know ... if we have to revise the theoretical part/Do we have to consider also "can" as part of the VG?
must	0	5	"Must" is usually used for internal obligation
could	1	4	Could you give me some help?/sorry for my bad English; I wish it could improve!
may	0	4	I may obviously be wrong!
will	6	4	this will be definitely helpful pages smiley/I hope prof. X will help and correct me
should	0	2	I should pay more attention next time?
be able to	1	1	I'm not able to find it/For some reason I wasn't able to figure out some examples

Da questo corpus nel suo complesso, considerando quindi sia le e-mail, sia i messaggi sul forum di Moodle, si evince pertanto che la metafora grammaticale rappresenta un potente mitigatore, volto a salvare la faccia (nel senso goffmaniano di "immagine pubblica", Goffman, 1955) degli interlocutori, alleviando il peso sociale legato a una richiesta di beni o servizi (Halliday - Matthiessen, 2004: 106-111). Questa strategia è chiaramente visibile in tab. 8, laddove gli studenti chiedono "Can anyone help me?" (anziché ingiungere "Help me"), si avvalgono di perifrasi indirette come "I hope Prof. X will help and correct me" (anziché chiedere direttamente un aiuto, "Prof. X, help and correct me") o talvolta addirittura utilizzano i verbi modali in modo grammaticalmente erraneo (es. "I *would have* explanations about the concept of social semiotic" anziché "I *would like* explanations ecc."), ma pragmaticamente più efficace rispetto al corrispondente atto diretto (es. "Provide explanations"). Questo utilizzo della metafora grammaticale dimostra non solo una competenza pragmatica relativamente elevata (e talvolta più elevata di quella *strettamente* grammaticale), ma anche la capacità di confrontarsi adeguatamente con la piattaforma di *e-learning*, malgrado la sua *netiquette* ibrida, che prevede una maggiore informalità rispetto alle e-mail, ma sempre nell'osservanza delle gerarchie sociali legate al contesto didattico e alla presenza (anche discreta) della docente.

5. Conclusioni

Il corpus esaminato denota la capacità degli studenti, la maggioranza dei quali ha un livello di competenza linguistica in inglese pari a B2, di destreggiarsi con una vasta gamma di strategie di cortesia interagendo on-line con la docente: il ricorso alla modificazione esterna (Blum-Kulka *et al.*, 1989) e alla metafora grammaticale (Halliday - Matthiessen, 2004; Thompson, 2014), nonché la varietà dei verbi modali, tutto sommato conforme al comportamento linguistico dei loro coetanei di madrelingua inglese (Leech *et al.*, 2009) dimostrano che gli studenti hanno una competenza pragmatica piuttosto elevata. Ne è un chiaro indizio anche la varietà delle strategie di mitigazione, che vanno ben oltre l'uso convenzionale, generalmente attribuito ai parlanti non nativi, e nondimeno irritante per i parlanti nativi, di *please+imperativo* (Economidou-Kogetsidis, 2011: 3206). Quanto alla scelta degli allocutivi, se nel sottocorpus di Moodle si deve in gran parte a strategie di convergenza e adeguamento all'interno della classe on-line, via e-mail è probabilmente correlabile con la lingua e cultura di origine degli studenti: ne è un esempio il fatto che sono solo le parlanti native a rivolgersi alla docente per nome, pur temperando l'informalità dell'allocutivo con un più distaccato *dear* (Formentelli, 2009). I dati su cui si fonda il presente studio non permettono, tuttavia, di istituire una corrispondenza diretta tra lingua madre degli studenti e tipo di allocuzione. Il quadro generale riflette piuttosto l'esistenza di una sorta di "classe europea" (composta di studenti sia interni all'Ateneo, sia coinvolti in progetti di scambio internazionale) che, pur a fronte di un discreto grado di variabilità interculturale, condivide la maggioranza dei problemi e delle necessità da manifestare alla docente e li affronta con strategie pragma-linguistiche in gran parte comuni.

I dati tratti dalla piattaforma *e-learning* denotano invece una maggiore interattività, nonché il rispetto di una *netiquette* ibrida, legata sia al mezzo stesso, sia al contesto, con un gradiente di formalità che dipende anche dal ruolo, più o meno partecipe, assunto dalla docente nei due anni accademici osservati. In generale, il forum *on-line*, come mezzo comunicativo, vede una riduzione della distanza sociale sia rispetto alle corrispettive situazioni dialogiche in presenza, sia rispetto alle e-mail e, sotto questo aspetto, il forum di Moodle qui analizzato non fa eccezione. Rispetto al sottocorpus contenente le e-mail, vi sono alcune somiglianze (es. l'uso di *thanks/thank you* a supporto di un atto direttivo e la metafora grammaticale come mitigatore), ma anche importanti differenze, *in primis* l'utilizzo degli *emoticon* per attuare strategie di cortesia positiva. L'elemento forse più importante che permette di differenziare la comunicazione docente-discente via e-mail da quella tramite forum, in base ai dati analizzati in questo studio, è che, mentre l'e-mail vede la presenza di soli due partecipanti (studente e docente), il forum di Moodle contempla una platea più ampia, con comunicazioni sia tra studenti e docente, sia tra studenti e studenti, consentendo quindi l'instaurarsi di strategie di adeguamento e convergenza, nonché la formazione di una vera e propria classe virtuale.

Bibliografia

- ACCIETTO T. - ZORZI D. (a cura di) (1998), *Nuove tecnologie e didattica delle lingue*, CLUEB, Bologna.
- ASTON G. (1995), Say “thank you”: some pragmatic constraints in conversational closings, in *Applied Linguistics* 16(1): 57-86.
- AYDIN S. (2014), Foreign language learners’ interactions with their teachers on Facebook, in *System* 42: 155-163.
- BARON N. (1998), Letters by phone or speech by other means: the linguistics of email, in *Language and Communication* 18(2): 133-170.
- BENNETT S. - MATON K. - KERVIN L. (2008), The digital natives debate: a critical review of the evidence, in *British Journal of Educational Technology* 39(5): 775-786.
- BENNETT S. - MATON K. (2010), Beyond the digital natives’ debate: towards a more nuanced understanding of students’ technology experiences, in *Journal of Computer Assisted Learning* 26(5): 321-331.
- BIESENBACH-LUCAS S. (2007), Students writing e-mails to faculty: an examination of e-politeness among native and non-native speakers of English, in *Language Learning and Technology* 11(2): 59-81.
- BLOCH J. (2002), Student/teacher interaction via e-mail: the social context of Internet discourse, in *Journal of Second Language Writing* 11(2): 117-134.
- BLUM-KULKA S. - HOUSE J. - KASPER G. (1989), *Cross-cultural pragmatics: requests and apologies*, Ablex, Norwood.
- BROWN P. - LEVINSON S. (1987), *Politeness: some universals in language use*, Cambridge University Press, Cambridge.
- CHEN C.E. (2006), The development of e-mail literacy: from writing to peers to writing to authority figures, in *Language Learning & Technology* 10(2): 35-55.
- CHEJNOVÁ P. (2014), Expressing politeness in the institutional e-mail communications of university students in the Czech Republic, in *Journal of Pragmatics* 60(1): 175-192.
- DOUGIAMAS M. - TAYLOR P. (2003), Moodle: using learning communities to create an open source Course Management System, in LASSNER D. - McNAUGHT C. (eds.), *Proceedings of the world conference on educational multimedia, hypermedia and telecommunications 2003*, AACE, Chesapeake, VA: 171-178.
- ECONOMIDOU-KOGETSIDIS M. (2011), “Please answer me as soon as possible”: pragmatic failure in non-native speakers’ e-mail requests to faculty, in *Journal of Pragmatics* 43(13): 3193-3215.
- ESLAMI Z.R. - AZIZULLAH M. - DINI S. (2015), The role of asynchronous computer mediated communication in the instruction and development of EFL learners’ pragmatic competence, in *System* 48: 99-111.
- FORMENTELLI M. (2009), Address strategies in a British academic setting, in *Pragmatics* 19(2): 179-196.
- FUSARI S. (in corso di stampa), A blended e-learning course of Systemic Functional Linguistics, in CERVINI C. (a cura di), *Interdisciplinarietà e apprendimento linguistico nei nuovi contesti formativi. L'apprendente di lingue tra tradizione e innovazione*, Quaderni del CeSLiC, Bologna.

- GOFFMAN E. (1955), On face-work: an analysis of ritual elements in social interaction, in *Psychiatry: Journal of Interpersonal Relations* 18(3): 213-231.
- HALLIDAY M.A.K. - MATTHIESSEN C.M.I.M. (2004), *An introduction to functional grammar*, 3rd edition, Arnold, London.
- HARTFORD B.S. - BARDOVI HARLIG K. (1996), "At your earliest convenience": a study of written student requests to faculty, in BOUTON L.F. (ed.), *Pragmatics and language learning, monograph series*, Vol. 7, University of Illinois Division of English as an International Language (DEIL), Urbana Champaign, IL: 55-69.
- HOUSE J. (1999), Misunderstandings in intercultural communication: interactions in English as a lingua franca and the myth of mutual intelligibility, in GNUTZMANN, C. (ed.), *Teaching and learning English as a global language*, Stauffenburg Verlag, Tübingen: 73-89.
- LEECH G. (1983), *Principles of pragmatics*, Longman, London.
- LEECH G. - HUNDT M. - MAIR C. - SMITH N. (2009), *Change in contemporary English. A grammatical study*, Cambridge University Press, Cambridge.
- MARQUES G. (2008), Establishing a continuum in spoken and written language in student emails, in NØRGAARD N. (ed.), *Systemic Functional Linguistics in use. Odense working papers in language and communication*, Vol. 29, University of Southern Denmark, Odense: 564-576.
- MCKAY L.S. (2002), *Teaching English as an international language*, Oxford University Press, Oxford.
- MERRISON A.J. - WILSON J.J. - BETHAN L.D. - HAUGH M. (2012), Getting stuff done: comparing email requests from students in higher education in Britain and Australia, in *Journal of Pragmatics* 44(9): 1077-1098.
- MIYAZOE T. - ANDERSON T. (2010), Learning outcomes and students' perceptions of online writing: simultaneous implementation of a forum, blog, and wiki in an EFL blended learning setting, in *System* 38(2): 185-199.
- O'DONNELL M. (2009), The UAM Corpus Tool: software for corpus annotation and exploration, in BRETONES CALLEJAS C.M. - FERNÁNDEZ SÁNCHEZ J.F. - IBÁÑEZ IBÁÑEZ J.R. - GARCÍA SÁNCHEZ M.E. - CORTES DE LOS RÍOS M.E. - SALABERRI RAMIRO S. - CRUZ MARTÍNEZ M.S. - PERDÚ HONEYMAN N. - CANTIZANO MÁRQUEZ B. (eds.), *Applied linguistics now: understanding language and mind/La lingüística aplicada hoy: comprendiendo el lenguaje y la mente*, Universidad de Almería, Almería: 1433-1447.
- POPESCU E. (2012), Providing collaborative learning support with social media in an integrated environment, in *World Wide Web* 17(2): 199-212.
- POPPI F. (2012), *Global interactions in English as a lingua franca*, Peter Lang, Bern.
- SHIM Y.S. (2013), International faculty perceptions of requestive emails by Korean university students, in *Multimedia-Assisted Language Learning* 16(4): 111-131.
- TAGG C. (2012), *Discourse of text messaging*, Continuum, London.
- THOMPSON G. (2014), *Introducing functional grammar*, 3rd edition, Routledge, London.
- THORNE S.L. (2008), Computer-mediated communication, in HORNBERGER N.H. (ed.), *Encyclopedia of language and education*, 2nd edition, Vol. 4: *Second and foreign language education*, Springer, Berlin: 325-336.
- YUS F. (2011), *Cyberpragmatics*, John Benjamins, Amsterdam/Philadelphia.
- ZAPPAVIGNA M. (2012), *The discourse of Twitter and social media*, Continuum, London.

STEFANIA SPINA¹

Le conversazioni scritte dei *social media*: un'analisi multidimensionale

The present study aimed to identify the distinctive features of conversations on social media by means of multidimensional analysis. Through the investigation of the systematic co-occurrence of 41 selected linguistic features, such as discourse markers, left and right dislocations, mitigation and opinion verbs, factor analysis was used to extract four distinct dimensions of variation from eleven types of conversation. Multidimensional analysis favoured functional interpretation of the main features of Facebook and Twitter conversational interactions and led to conclude that these features are strictly connected to the complex relationships established between participants and their audience. Multidimensional analysis also suggested that the different ways these relationships are established and maintained affect the discourse strategies adopted in different social media.

1. *La conversazione come genere testuale*

Secondo una visione tradizionale, la conversazione è un'attività socialmente organizzata che ha luogo attraverso lo scambio verbale tra due o più partecipanti, prevalentemente di carattere informale e basato sulla lingua parlata (Bazzanella, 2010). Nella sua forma tipica (Berretta, 1994), o prototipica (Bazzanella, 2002), quella della conversazione faccia-a-faccia, essa è caratterizzata dalla sincronia, dalla compresenza di un numero delimitato di interlocutori e dalla possibilità di feedback immediato.

Con la diffusione della comunicazione mediata dal computer (CMC) e la nascita e il rapido sviluppo di generi testuali nuovi, si creano nuove modalità di interazione dialogica, che impongono una ridefinizione del genere "conversazione", e, insieme, l'adozione di metodologie di analisi adeguate a coglierne le differenze. Lo scopo di questo lavoro è la presentazione e la descrizione di una metodologia di analisi delle conversazioni e di come questa possa essere applicata per identificare i tratti caratterizzanti e distintivi della comunicazione sui *social media*.

1.1 La conversazione nella comunicazione mediata dal computer

La diffusione della CMC ha tra le altre conseguenze lo sviluppo di una serie di nuove forme di conversazione, basate sul codice scritto digitale: oltre alle interazioni nelle chat (Pistolesi, 2004), esse annoverano le più recenti interazioni nei *social media*. Fin dalla seconda metà degli anni 80 del secolo scorso, diversi studi ne hanno messo in evidenza via via aspetti diversi: il fatto di essere 'speech writ down' (Horowitz - Samuels, 1987:

¹ Università per Stranieri di Perugia.

27), o di costituire forme di conversazione i cui turni sono regolati dalla mediazione di un computer (Murray, 1989); il fatto di essere percepite attraverso la vista e non attraverso l'udito (Colomb - Simutis, 1996), e di essere persistenti, e non più volatili come la conversazione orale (Erickson - Herring, 1999; Fiorentino, 2011); il fatto, infine, di rappresentare forme di conversazione su vasta scala, con un pubblico potenzialmente molto ampio (Sack, 2000).

L'emergere di una nuova tipologia di interazione, la "conversazione mediata dal computer" (Herring, 2011), ha inoltre portato ad una riformulazione della definizione tradizionale di conversazione, legata esclusivamente alla sua modalità orale e faccia-a-faccia: conversazione è dunque "any exchange of messages between two or more participants, where the messages that follow bear at least minimal relevance to those that preceded or are otherwise intended as responses" (Herring, 2011).

Tale definizione è in grado di inglobare anche le forme più recenti di conversazione mediata dal computer, quelle che hanno luogo sui *social media*, sviluppatasi a partire dal 2003².

Una riflessione preliminare su questo tipo di interazioni è in grado di metterne in luce alcune caratteristiche generali. Esse si presentano come conversazioni:

- basate sullo scritto digitale;
- asincrone;
- senza compresenza fisica dei partecipanti;
- con destinatari non necessariamente noti, potenzialmente molto numerosi, che possono variare nel corso delle interazioni, attraverso meccanismi di selezione (come ad esempio la menzione, una strategia molto diffusa in particolare su Twitter);
- persistenti: non sono interazioni volatili, ma si possono salvare, ricercare, recuperare, riproporre, e modificare, e sono per questo "searchable talk", secondo la definizione di Zappavigna (2012: 95);
- dal punto di vista linguistico, tendenti all'oralità, ma con una propria, peculiare organizzazione testuale e discorsiva e con lo sviluppo di regole conversazionali proprie (Spina, 2012).

La conversazione rappresenta di conseguenza una modalità comunicativa mista (Bazzanella, 2010), le cui diverse realizzazioni si collocano lungo un continuum tra lingua parlata e lingua scritta. La diffusione di forme di conversazione mediata dal computer accentua dunque la crisi dei modelli di interpretazione e di analisi basati esclusivamente sull'opposizione scritto-parlato; tale crisi, già evidenziata da alcuni studi fondamentali, come Biber (1988), diventa ancora più evidente da quando la conversazione mediata dal computer ha "ampliato i confini del dialogo e piegato la parola scritta al dominio dell'immediatezza" (Pistolessi, 2004: 34).

² Il 2003 è l'anno di fondazione di Myspace, seguito, l'anno successivo, da Facebook, attualmente il network più frequentato al mondo. Non sono considerati in questa sede, perché non più attivi, alcuni precursori dei *social media*, nati alla fine degli anni '90, come ad esempio SixDegrees o Friendster.

Per analizzare in profondità le caratteristiche linguistiche delle nuove forme di conversazione realizzate sui *social media* sono dunque indispensabili metodologie che tengano conto della loro natura multidimensionale.

2. Metodologie multidimensionali: l'analisi fattoriale

L'approccio multidimensionale utilizzato in questo lavoro è basato su una metodologia che applica tecniche statistiche (nel nostro caso, l'analisi fattoriale) allo studio di generi testuali con un ampio spettro di variazione: molto raramente, infatti, l'insorgere di un tratto linguistico è funzione di un singolo elemento che ne provoca la selezione da parte di un parlante; più spesso, una particolare scelta linguistica è frutto di un insieme di fattori che, in modi e misure diversi, agiscono su tale scelta (Gries, 2013: 247).

Originariamente, l'analisi fattoriale è stata usata in linguistica per analizzare la variazione tra generi testuali scritti e parlati in inglese; al lavoro fondamentale di Biber (1988) hanno fatto seguito diversi altri studi, curati dallo stesso autore (Biber, 1992; 1995; 2003; 2004). In ambito italiano, Santini (2001) ha applicato l'analisi fattoriale a un corpus di testi giornalistici di tipologie differenti.

Tabella 1 - *Tipologia, caratteristiche e dimensione del corpus degli undici generi di conversazione analizzati*

<i>conversazione</i>	<i>tokens</i>	<i>orale</i>	<i>sincrona</i>	<i>compresenza</i>	<i>pianificata</i>	<i>tra pari</i>	<i>vincoli lungh.</i>
faccia a faccia	187.454	+	+	+	-	+	-
telefonica	284.567	+	+	-	-	+	-
istituzionale	144.124	+	+	+	-	-	-
intervista	402.420	+	+	+/-	-/+	-	-
Facebook	492.508	-	-	-	+	+	-
Twitter	110.744	-	-	-	+	+	+
chat	119.280	-	+	-	-	+	-
talkshow tv	221.008	+	+	+	-	-	+
spettacolo tv	80.238	+	+	+	+	-	+
fiction tv	127.026	-	-	-	+	+/-	+
film	626.487	-	-	-	+	+/-	+

In questo studio, l'analisi fattoriale è stata applicata a un insieme di undici generi testuali che condividono il fatto di essere conversazioni, con l'obiettivo di analizzare le caratteristiche delle interazioni nei *social media* in rapporto ad altri generi di conversazione. La tab. 1 riporta gli undici generi testuali considerati e, per ognuno, i dati utilizzati, che

sono tutti tratti dal *Perugia corpus* (Spina, 2014), un corpus di riferimento dell'italiano contemporaneo, scritto e parlato.

La matrice riportata nella tab. 1 segnala la presenza/assenza di alcuni elementi caratterizzanti, come l'oralità, la sincronia, la compresenza fisica, la pianificazione, la parità di diritti-doveri tra gli interlocutori e i vincoli nella lunghezza dei testi prodotti. La matrice evidenzia l'eterogeneità delle conversazioni prese in considerazione; tra di esse, le due tipologie meno distanti tra loro sono proprio le interazioni in Facebook e in Twitter, che si differenziano esclusivamente per l'ultimo tratto, quello relativo ai vincoli nella lunghezza dei testi: Twitter, a differenza di Facebook, ha infatti un limite massimo di 140 caratteri per ogni testo inviato.

In presenza di uno spettro di variazione così ampio, l'analisi fattoriale può risultare efficace anche perché è una metodologia di tipo esplorativo, che ha come obiettivo principale quello di generare ipotesi, di suggerire piste di ricerca che possono essere approfondite attraverso altre metodologie. Essa è inoltre in grado di ridurre un numero elevato di variabili linguistiche a pochi elementi più semplici, la cui variazione è quindi più facile da interpretare (Biber, 1988). L'assunto che costituisce la base teorica nello studio della variazione di generi testuali attraverso l'analisi fattoriale è che la co-occorrenza sistematica di tratti linguistici riflette una funzione condivisa. Tale assunto non è peraltro esclusivo di un approccio quantitativo allo studio del linguaggio: già Hymes (1984) identificava gli "speech styles" come varietà definite da un insieme condiviso di tratti linguistici che co-occorrono; e Halliday (1988: 162) scriveva che un genere testuale è "a cluster of associated features having a greater-than-random tendency to co-occur".

Date queste premesse, l'approccio multidimensionale adottato in questo lavoro ha previsto le seguenti cinque fasi:

1. selezione dei tratti linguistici da analizzare;
2. conteggio della frequenza dei tratti selezionati negli undici generi testuali oggetto di analisi;
3. analisi fattoriale: individuazione delle dimensioni di variazione, caratterizzate dalla co-occorrenza sistematica di gruppi di tratti linguistici;
4. interpretazione di ciascuna dimensione di variazione;
5. posizionamento di ciascun genere testuale all'interno del continuum delle dimensioni di variazione.

2.1 Tratti linguistici selezionati

La selezione dei tratti linguistici da analizzare attraverso l'analisi fattoriale è stata effettuata sulla base di due criteri fondamentali: in primo luogo, la rilevanza della loro presenza/assenza all'interno del genere testuale "conversazione", così come è descritta in letteratura, in base alla loro funzione comunicativa. I segnali discorsivi, ad esempio, sono elementi linguistici dal valore pragmatico-testuale molto utilizzati nel parlato conversazionale e con una grande varietà di funzioni (Bazzanella, 1994; 1995), e sono stati di conseguenza selezionati come tratti linguistici rilevanti ai fini di questa analisi.

Tabella 2 - I 56 tratti linguistici selezionati

<i>n.</i>	<i>tratto</i>	<i>n.</i>	<i>tratto</i>
	AGGETTIVI	27	III pers. s.
1	qualificativi	28	I pers. pl.
2	superlativi	29	II pers. pl.
3	accrescitivi - enfatici (<i>pazzesco, mitico...</i>)	30	III pers. pl.
	AVVERBI	31	indefiniti
4	di tempo	32	dimostrativi
5	di luogo	33	rafforzativi dei dimostrativi (<i>questo qui, ...</i>)
6	di modo	34	sintagmi preposizionali
	NOMI	35	intensificatori (<i>proprio, davvero... + agg.</i>)
7	tutti i nomi	36	prefissi intensificativi (<i>macro-, maxi-, mega-...</i>)
8	nominalizzazioni (<i>-mento, -ità, -zione</i>)	37	mitigazione
9	nomi in <i>-ata</i>	38	lessico delle emozioni
	VERBI	39	domande dirette
10	tutti i verbi	40	segnali discorsivi
11	I pers. s.	41	frasi esclamative
12	II pers. s.	42	interiezioni (<i>Eh! Ciao! Accidenti! Oddio! ...</i>)
13	III pers. s. e pl.	43	dislocazioni (a destra e a sinistra)
14	I pers. pl.	44	frase scissa
15	II pers. pl.	45	<i>c'è</i> presentativo
16	imperfetto indicativo	46	espressioni di opinione
17	passato prossimo	47	periodo ipotetico imperfetto indicativo
18	futuro	48	periodo ipotetico imperfetto coniuntivo
19	condizionale	49	<i>stare</i> + gerundio
20	imperativo (+ clitico)	50	<i>stare</i> + <i>a</i> + infinito
21	modali	51	coordinazione frasale (<i>E</i> a inizio frase)
22	causativi		SUBORDINAZIONE
23	passivi (+ <i>da</i>)	52	coniunzioni subordinanti
24	pronominali (<i>farcela, avercela, andarsene...</i>)	53	<i>che</i> subordinante
	PRONOMI PERSONALI E POSSESSIVI	54	pronomi relativi \neq <i>che</i>
25	I pers. s.	55	subordinate introdotte da preposizione
26	II pers. s.	56	gerundio

In secondo luogo, la loro reperibilità nel corpus utilizzato, attraverso interrogazioni automatiche. Il *Perugia corpus* è infatti annotato per categoria grammaticale, e consente

l'estrazione di informazioni di tipo grammaticale, lessicale, e in parte sintattico; l'esigenza di analizzare tratti linguistici specifici ha comunque dovuto essere temperata con la possibilità di reperirli ed estrarli in modo automatico.

Il processo di applicazione di questi due criteri ha portato alla selezione finale di 56 tratti linguistici, riportati nella tab. 2. I tratti sono di tipo grammaticale, lessicale (ad esempio il lessico delle emozioni, individuato attraverso un insieme di avverbi, verbi, nomi e aggettivi³), sintattico (dislocazioni a destra e a sinistra, frasi scisse, subordinate ecc.), o pragmatico (come la mitigazione, ottenuta attraverso mezzi lessicali o frasali; cfr. Caffi, 2001).

Di questi 56 tratti originariamente selezionati, solo 41 sono stati conservati per l'esecuzione dell'analisi fattoriale: 15 tratti non hanno concorso all'individuazione delle dimensioni di variazione descritte più avanti, perché avevano un peso troppo esiguo, e dunque non statisticamente significativo ai fini della misura della co-occorrenza. I 15 tratti non utilizzati sono i superlativi, gli avverbi di modo, i nomi, i nomi in *-ata*, i pronomi di I, II e III persona plurale, il futuro, i verbi di III persona singolare e plurale, i verbi causativi, i prefissi intensificativi, i periodi ipotetici con l'imperfetto indicativo e congiuntivo, e le costruzioni con *stare*+gerundio e *stare+a*+infinito. I restanti 41 tratti, invece, hanno tutti avuto un peso significativo nella definizione delle dimensioni di variazione.

3. *Analisi fattoriale*

Gli undici sotto-corpora utilizzati come campioni rappresentativi degli undici generi di conversazione sono stati suddivisi, in modo casuale, ciascuno in cinque parti distinte: il computo dell'occorrenza dei 56 tratti inizialmente selezionati è stato dunque effettuato per ciascuno dei 55 testi ottenuti (5 x 11), e successivamente normalizzato ad un valore di 1.000, per risolvere il problema della non omogeneità delle dimensioni dei sotto-corpora (Santini, 2001), evidenziato dalla tab. 1.

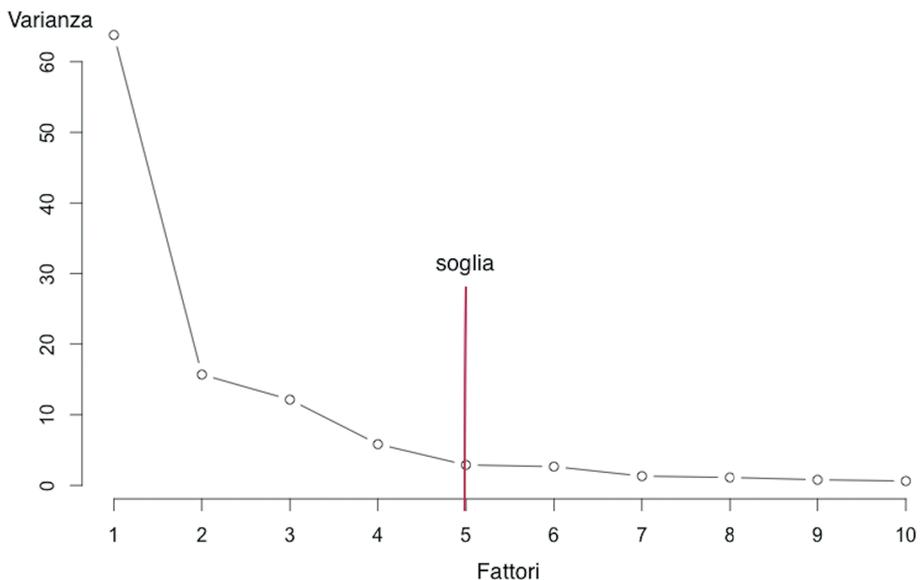
L'operazione successiva è stata il calcolo della matrice delle correlazioni, che costituisce la base di partenza dell'analisi: misurando quanto le variabili analizzate siano correlate tra loro, essa indica infatti se i dati considerati sono adatti ad essere sottoposti ad un'analisi fattoriale. In particolare, se nessuna correlazione supera il coefficiente di 0,3, i risultati dell'analisi non possono essere considerati affidabili (Santini, 2001). Nel nostro caso, la matrice mostra invece che molte delle correlazioni tra le variabili supera la soglia di 0,3; questo dato consente di procedere con l'analisi fattoriale.

Il passaggio successivo è costituito dall'analisi fattoriale vera e propria, con l'estrazione dei fattori e la determinazione di quanti di essi possono essere selezionati

³ Si tratta di un insieme di 103 avverbi (ad es. *dolorosamente*), 181 verbi (ad es. *detestare*, *adorare*), 158 aggettivi (*meraviglioso*, *terribile*), 297 nomi (*affetto*, *rabbia*) che esprimono una delle quattro coppie di emozioni di sorpresa/attesa, felicità/tristezza, rabbia/paura, disgusto/accettazione, o di emozioni da esse derivate, individuate da Plutchik (2001); gli elementi lessicali sono stati adattati da *WordNet-Affect* (Strappafava - Valitutti, 2004), un'estensione del database lessicale *Wordnet* (Fellbaum, 1998).

come dimensioni di variazione. La scelta della soglia di varianza al di sotto della quale è consigliabile eliminare i fattori individuati, perché non rilevanti, è un passaggio fondamentale: uno dei criteri normalmente utilizzati è quello di eliminare i fattori con varianza inferiore a 1 (Biber, 1988). Nel nostro caso, i fattori da mantenere per l'analisi ammonterebbero ad otto. Un altro possibile criterio è quello di escludere i fattori al di sotto del punto in cui c'è un ampio intervallo tra i valori della varianza. La fig. 1 mostra che questo punto si colloca in particolare dopo il fattore 4; la scelta è stata quella di includere nell'analisi solo i primi quattro fattori⁴, e di individuare dunque quattro dimensioni di variazione.

Figura 1 - I primi 10 fattori individuati dall'analisi fattoriale e la loro varianza



Infine, i quattro fattori selezionati per l'analisi sono stati ruotati⁵, come è prassi nell'analisi fattoriale (Biber, 1988). La rotazione consiste nel presentare i dati ottenuti in una forma matematicamente equivalente all'originale non ruotato, ma che rappresenta i fattori in un modo più facilmente interpretabile (Santini, 2001).

È importante sottolineare come l'estrazione e la determinazione dei fattori, cioè delle dimensioni di variazione di un insieme di generi testuali, sia un procedimento strettamente linguistico (perché basato sulla co-occorrenza di tratti linguistici), empirico e quantitativo. Attraverso questo procedimento, l'analisi fattoriale ha portato all'identificazione di quattro dimensioni di variazione. La descrizione e l'interpretazione delle quattro dimensioni individuate, che sono invece operazioni di tipo qualitativo, saranno espone nei prossimi quattro paragrafi.

⁴ La varianza cumulata dei primi quattro fattori, scelti per l'analisi, ammonta all'88,2% della varianza totale.

⁵ In particolare, ai quattro fattori è stata applicata una rotazione di tipo Promax (Biber, 1988).

3.1 Dimensione 1: interattività/informatività

Il primo fattore individuato dall'analisi rappresenta la dimensione di variazione, definita dai due poli dell'interattività e dell'informatività, che copre la maggior parte della variabilità dei tratti linguistici selezionati, ed è dunque il più forte e il più chiaramente delineato.

Tabella 3 - *I tratti linguistici della dimensione di variazione interattività/informatività (1)*

<i>tratto</i>	<i>loading</i>
pron_IIs	0,826
verbi_Is	0,804
verbi_IIs	0,764
pron_Is	0,731
interiezioni	0,731
segnali discorsivi	0,643
imperativo	0,551
domande dirette	0,549
coord_frase	0,469
v. pronominali	0,437
avv_luogo	0,434
esclamative	0,376
nominalizzazioni	-0,842
nomi	-0,822
sintagmi_prep	-0,792
aggettivi	-0,727
pre_sub	-0,628
che_sub	-0,618
relativi	-0,602
passivi	-0,586
gerundio	-0,450

La tab. 3 mostra che i tratti linguistici che hanno un peso maggiore⁶ in questa dimensione sono i verbi e i pronomi di prima e seconda persona singolare, le interiezioni, i segnali discorsivi, l'imperativo, le domande dirette e le frasi esclamative. Si tratta di risorse linguistiche prevalentemente verbali, con un focus deciso sull'interazione, che introducono un riferimento diretto ai destinatari, o che, come i segnali discorsivi, hanno una natura prevalentemente pragmatica e sono mirate soprattutto

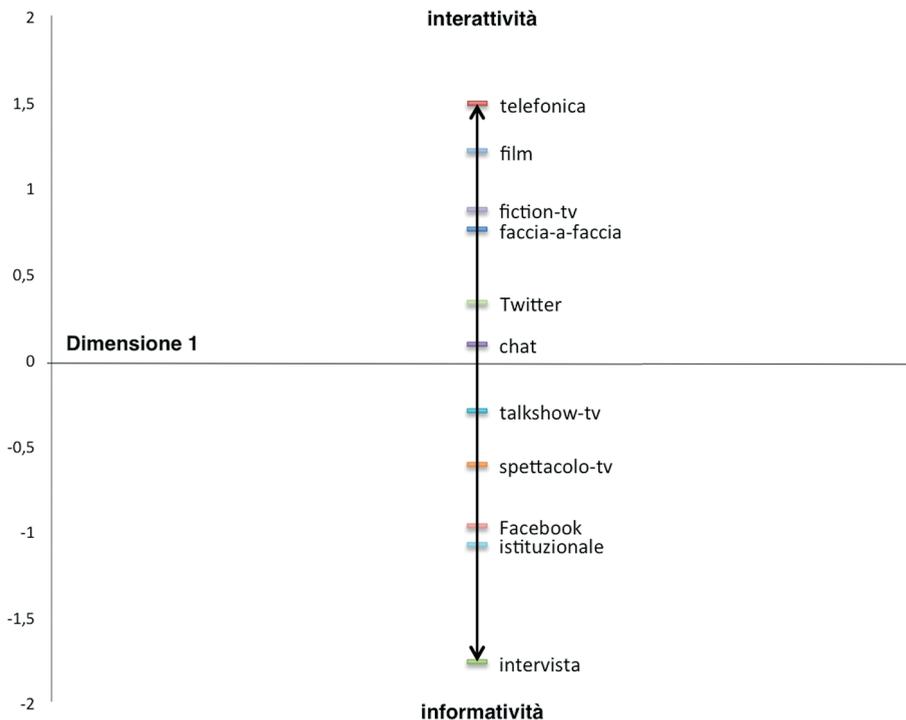
⁶ Il coefficiente associato a ciascun tratto linguistico può variare da 0 a 1, ed indica la misura in cui un tratto è correlato con la dimensione di variazione di cui fa parte. Più il coefficiente è vicino ad 1, più questa correlazione è forte. In generale, si considerano parte integrante di una dimensione di variazione i tratti con peso > 0,3 (Santini, 2001). I tratti con coefficiente negativo (da -0,3 a -1) sono in distribuzione complementare con quelli con coefficiente positivo.

ad ancorare la conversazione agli interlocutori e al contesto (Bazzanella, 1995). La presenza, fra i tratti rilevanti, di coordinazione frasale, verbi pronominali e avverbi di luogo, caratterizza ancora più chiaramente la dimensione come interattiva, perché indica una forma di comunicazione che procede in modo additivo più che gerarchico (Voghera, 1992), che è basata sulla deissi spaziale oltre che personale (Bazzanella, 1994), e che si presenta come tendenzialmente informale (come suggerito dalla presenza dei verbi pronominali *volerci*, *farcela* o *andarsene*, ecc.).

I tratti linguistici di segno negativo associati a questa dimensione, per contro, sono prevalentemente di tipo nominale (nominalizzazioni, nomi, aggettivi), funzionali ad una presentazione gerarchica delle informazioni (subordinate introdotte da preposizione e congiunzione, pronomi relativi, gerundio) e ad una generale impressione di distacco (passivi); il loro focus è dunque sulla densità informativa (Voghera, 1992).

La dimensione di variazione dal peso più rilevante è dunque quella racchiusa tra i due poli di interattività ed informatività; tale dimensione è presente in tutti gli altri studi sulla variazione di generi testuali basati sull'analisi multidimensionale (Biber, 1988; 1995; Santini, 2001). Per questo motivo Biber (2004) arriva ad ipotizzare che questa dimensione possa costituire un universale nella variazione tra generi.

Figura 2 - *Gli 11 generi di conversazione all'interno della dimensione interattività/informatività (1)*



La fig. 2 mostra che, per quanto riguarda i due generi di conversazione oggetto di attenzione nel presente studio, vale a dire quelle sui *social media*, Twitter è uno dei cinque generi più interattivi, mentre Facebook è posizionato nella parte inferiore del grafico, perché manifesta spiccati caratteri di informatività.

I testi brevi inviati in Twitter riflettono dunque un atteggiamento prevalentemente orientato all'interazione e alla relazione con i propri interlocutori, come mostrano gli esempi 1-3; la presenza di segnali discorsivi, interiezioni, *emoticon*, domande dirette e verbi e pronomi di prima e seconda persona sono la spia di questo atteggiamento orientato all'interlocutore più che all'esclusiva trasmissione di contenuti informativi.

- (1) Beh, vado (: Buonanotte #off
- (2) ah ah ah buona domenica teso'!!!
- (3) Domani andrei all'IKEA etc. vuoi venire anche tu? Altri baci, altro ciao.

Le conversazioni in Facebook (cfr. es. 4-5), prive di vincoli di lunghezza, sono spesso di carattere più informativo, e presentano tratti di densità ed astrattezza che raramente si ritrovano in Twitter.

- (4) Il giocatore è stato quindi ricoverato in ospedale per alcuni accertamenti.
- (5) così ho chiamato il numero verde, il quale nn ha saputo darmi le informazioni che volevo riguardanti la mia chiamata

3.2 Dimensione 2: elaborazione sintattica/espressione non strutturata di emozioni

Anche il secondo fattore emerso dall'analisi statistica presenta un insieme abbastanza vasto di tratti con una forte correlazione reciproca (cfr. tab. 4). La dimensione è caratterizzata in particolare dalla presenza di subordinate, introdotte da *che* e da altre congiunzioni, dall'uso di meccanismi di focalizzazione (il *c'è* presentativo, le dislocazioni a destra e a sinistra e le frasi scisse), e dal riferimento diretto al tempo, al luogo e ai partecipanti (gli avverbi di tempo e i dimostrativi, anche in forma rafforzata). I tratti con peso di segno negativo, in distribuzione complementare con i precedenti, individuano invece una tendenza alla manifestazione esplicita di emozioni, con un riferimento particolare a destinatari indicati alla seconda persona plurale, e una mancanza di strutturazione in subordinate (uso dell'imperativo, di frasi esclamative e di domande dirette).

La dimensione è dunque caratterizzata da una variazione interna collocata tra i due estremi dell'elaborazione sintattica, mirata in particolare alla focalizzazione e alla deissi, da un lato, e dell'espressione non strutturata di emozioni, dall'altro. La presenza di subordinate di primo grado è un tratto tipico del parlato dialogico (Berretta, 1994; Voghera, 1992), come anche l'uso frequente di meccanismi di focalizzazione, che rispondono "a un'esigenza di semplicità della struttura informativa" (Berretta, 1994: 257). La fig. 3 mostra tuttavia che en-

trambe le tipologie di conversazione sui *social media* si situano sul versante del continuum legato all'espressione non strutturata di emozioni; in Twitter, e soprattutto in Facebook, si tende a comunicare al proprio pubblico di *follower* o di amici i propri stati d'animo, spesso emotivi, con un'articolazione della frase scarsamente strutturata, come mostrano gli esempi 6-9.

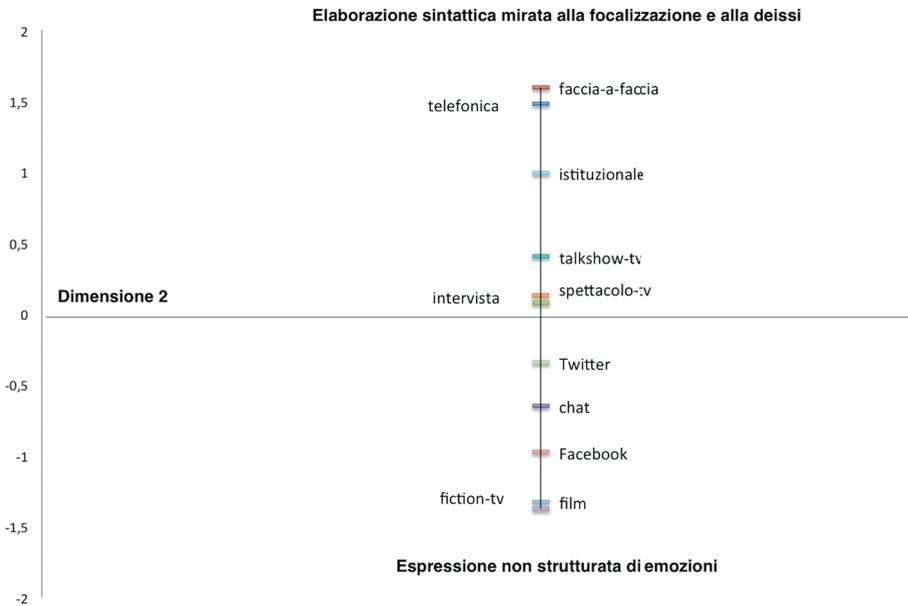
Tabella 4 - *I tratti linguistici della dimensione di variazione elaborazione sintattica/ espressione non strutturata di emozioni (2)*

<i>tratto</i>	<i>loading</i>
cong_sub	0,756
presentativo	0,640
rafforz. demo	0,592
demo	0,532
che_sub	0,530
avv_tempo	0,509
dislocazioni	0,473
scissa	0,403
esclamative	-0,793
domande dirette	-0,666
lessico_emo	-0,664
verbi_Ipl	-0,456
imperativo	-0,356

- (6) emozionante qlla marea di persone! Milano vuole crederci ... e io anke!
- (7) Sì è vero... l'ho letto nel cinema della mia città ho visto RED e credetemi è un film pessimo!
- (8) :(Magari!!! Siamo tutti con voi! Fategliela vedere alla Moratti e al Berlusca!

L'espressione di emozioni prevale dunque, specie in Facebook, sull'uso di meccanismi di strutturazione della frase utilizzati diffusamente in altre forme di conversazione. Il dato è meno evidente in Twitter, dove peraltro la focalizzazione è spesso ottenuta non attraverso un ordine marcato delle parole, ma per mezzo dell'*hashtag*, una risorsa ad hoc per mettere in evidenza specifici elementi lessicali (Zappavigna, 2015).

Figura 3 - *Gli 11 generi di conversazione all'interno della dimensione elaborazione sintattica/ espressione non strutturata di emozioni (2)*



3.3 Dimensione 3: narrazione/coinvolgimento

Il terzo fattore individuato dall'analisi statistica è contrassegnato, come mostra la tab. 5, dalla co-occorrenza di riferimenti a persone esterne alla situazione comunicativa (verbi e pronomi di terza persona), tempi verbali del passato, deissi (avverbi di luogo, dimostrativi) e strutturazione delle frasi in subordinate (introdotte soprattutto da *che*); tali tratti sono tradizionalmente associati a testi di tipo narrativo (Biber, 1988). I tratti in distribuzione complementare sono invece essenzialmente strategie mirate all'intensificazione e al coinvolgimento degli interlocutori.

La variazione della dimensione 3 è dunque compresa tra i due poli della narrazione e del coinvolgimento. Sebbene una dimensione di variazione relativa al parametro della narrazione fosse presente in precedenti analisi multidimensionali (Biber, 1988; 2004), negli studi precedenti il polo "testi narrativi" era opposto genericamente a "testi non narrativi", caratterizzati da tratti linguistici come l'uso del presente indicativo. Nel nostro caso, invece, narrazione si oppone a coinvolgimento; i tratti linguistici che sono usati per raccontare eventi o situazioni sono in distribuzione complementare con quelli che mirano al coinvolgimento degli interlocutori.

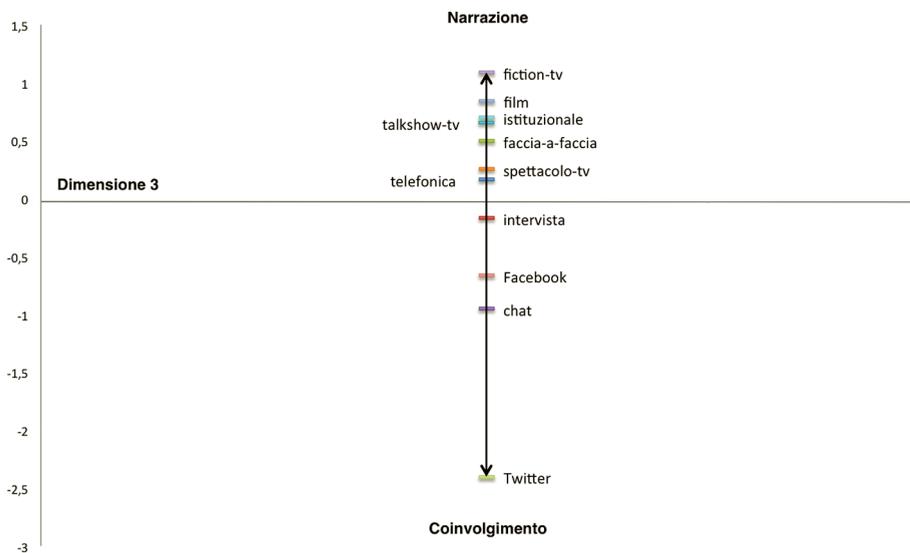
La fig. 4 mostra il modo in cui le conversazioni analizzate si dispongono nel continuum di variazione della dimensione. Se i due generi esplicitamente finalizzati a narrare storie (la fiction televisiva e i film) occupano, come era da attendersi, la posizione che tende di più verso il polo della narrazione, le due conversazioni sui *social media* si situano, ancora una volta, in posizioni differenti: le interazioni in Twitter sono completamente orientate a coinvolgere gli interlocutori, e non a narrare storie,

almeno secondo il canone tradizionale di narrazione; le interazioni in Facebook si situano invece in una posizione più intermedia, in cui, sebbene le strategie di intensificazione e di coinvolgimento svolgano un ruolo preponderante, non sono assenti modalità discorsive di tipo narrativo.

Tabella 5 - *I tratti linguistici della dimensione di variazione narrazione/coinvolgimento (3)*

<i>tratto</i>	<i>loading</i>
pron_III _s	0,675
imperfetto	0,646
verbi_III _{s_pl}	0,603
dimostrativi	0,559
pass_pross	0,435
modali	0,433
avv_luogo	0,367
che_sub	0,345
intensificatori	-0,579
verbi_1 _s	-0,445
verbi_1 _{pl}	-0,443

Figura 4 - *Gli 11 generi di conversazione all'interno della dimensione narrazione/coinvolgimento (3)*



Le due forme di conversazione che hanno luogo sui *social media* si rapportano dunque in modo diverso con la narrazione di eventi. In Twitter (es. 9-10) prevale in modo netto l'uso di tratti non narrativi (la prima persona plurale dei verbi, gli intensificatori), che mirano ad enfatizzare e a coinvolgere. In Facebook, invece, i due poli

si bilanciano in misura maggiore, e trovano spazio anche tratti tipicamente narrativi come negli es. 11-12.

- (9) Se facciamo a noi stessi una domanda difficile e ci diamo la risposta che speravamo, beh è semplicemente l'inizio della felicità.
- (10) @utente anche solo la parte vegetariana è davvero gustosa, te lo consiglio! :)
- (11) Oggi pomeriggio mentre ero davanti al meccanico una signora ha parcheggiato la macchina di 20 cm sulle strisce dicendo che sarebbe tornata in 2 minuti.
- (12) Questa mattina appena mi sono svegliata mi sembrava che erano già iniziate le vacanze.

3.4 Dimensione 4: espressione di un'opinione soggettiva

Il quarto ed ultimo fattore è caratterizzato, come mostra la tab. 6, da un tratto linguistico dal peso molto rilevante (le espressioni di opinione), che dà l'impronta a tutta la dimensione; insieme, concorrono all'individuazione del fattore una serie di tratti mirati ad attenuare l'opinione espressa (il condizionale, una serie di risorse lessicali e frasali finalizzate anch'esse alla mitigazione, come *alquanto, per lo più, in parte, tendenzialmente, più o meno, su per giù, in un certo senso, non è detto che, è possibile che* ecc., gli indefiniti e i verbi modali). I pronomi di prima persona singolare, che pure hanno un peso leggermente inferiore allo 0,3 stabilito come soglia minima, sono comunque funzionalmente coerenti con gli altri tratti.

Tabella 6 - I tratti linguistici della dimensione di variazione espressione di un'opinione soggettiva (4)

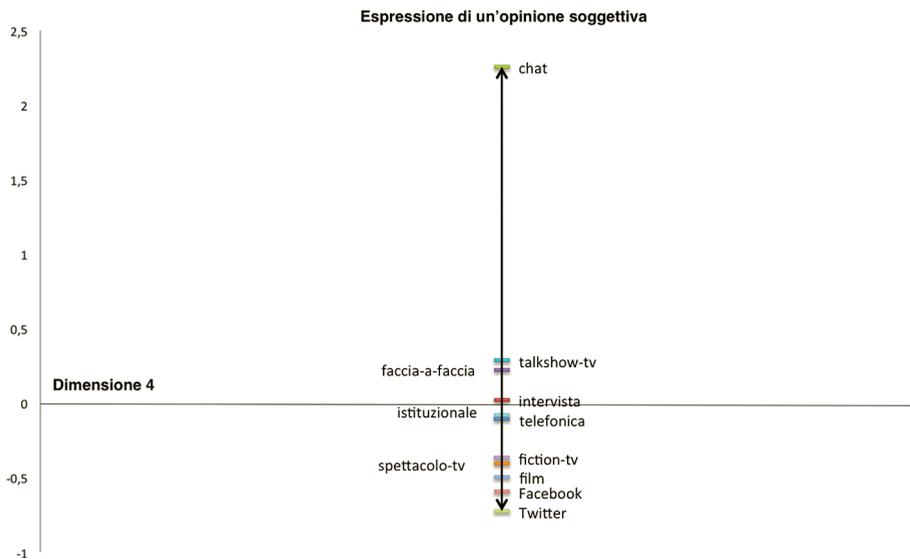
<i>tratto</i>	<i>loading</i>
espr_opinione	0,987
condizionale	0,694
mitigazione	0,570
indefiniti	0,361
modali	0,347
pron_Is	0,296

La dimensione in questione, che non presenta tratti linguistici di segno negativo, è dunque caratterizzata dall'espressione, potenzialmente attenuata, di un punto di vista personale; può essere dunque definita come la dimensione dell'espressione di un'opinione soggettiva.

La fig. 5 mostra la posizione degli undici generi di conversazione all'interno di questo continuum di variazione: dieci di essi sono piuttosto vicini, e si dispongono tutti intorno ad un valore medio. Anche i due generi relativi ai *social media* hanno un comportamento simile nei confronti dell'espressione di un'opinione; nessuno

dei due, infatti, è usato in genere per esprimere un punto di vista soggettivo in modo argomentato, servendosi di risorse linguistico-discorsive ad hoc, come i verbi parentetici o le espressioni come *per me, secondo me, a mio giudizio*, ecc. Più che argomentare un'opinione, sui *social media* si esprimono emozioni, come ha peraltro dimostrato la descrizione della dimensione 3.

Figura 5 - Gli 11 generi di conversazione all'interno della dimensione espressione di un'opinione soggettiva (4)



4. Discussione

L'analisi fattoriale ha messo in luce alcune delle strategie discorsive che caratterizzano le conversazioni sui *social media* rispetto ad altre forme di interazione. In primo luogo, tali conversazioni sono contraddistinte dal tentativo costante di manifestare la propria presenza connessa in rete (Baron, 2008): la dimensione di variazione (3) indica in modo chiaro che la comunicazione in Twitter è principalmente mirata al coinvolgimento dei propri naturali destinatari – i *follower* o gli amici – ai quali si vuole costantemente mostrare la propria disponibilità alla connessione (Miller, 2008). Si tratta di un coinvolgimento ricercato e al tempo spesso offerto da chi partecipa alle interazioni in rete e finalizzato principalmente alla definizione di un contesto comune di connessione; il messaggio che si vuole inviare, come mostra l'esempio 13, è in sostanza: “Sono qui, e sono disponibile alla conversazione: voi ci siete?”.

(13) buongiorno, mi sono appena svegliata e ho un sonno assurdo .-.

Sui *social media*, inoltre, non si narrano storie con gli strumenti linguistici e discorsivi utilizzati tradizionalmente per raccontare: in Twitter e Facebook si narra coinvolgendo direttamente gli interlocutori, trasportandoli all'interno della storia e dell'e-

mozione che essa è in grado di suscitare. Un'altra strategia che emerge dall'analisi è infatti quella mirata all'aggregazione emotiva del proprio "pubblico personale" (Schmidt, 2014) attorno a temi o valori condivisi. Il ricorso costante all'emotività, individuato chiaramente dalla dimensione di variazione (2), è anche finalizzato a suscitare una reazione nella comunità dei partecipanti. La focalizzazione su un tema o su una posizione è ottenuta più attraverso l'espressione forte ed esplicita delle emozioni che il tema è in grado di evocare che attraverso le risorse linguistiche tipicamente usate nell'interazione parlata, come ad esempio il mutamento dell'ordine non marcato delle parole. Lo stesso uso dell'*hashtag*, su cui tanto fanno leva le interazioni in Twitter, ingloba la doppia funzione di focalizzazione su un tema e di aggregazione degli interlocutori attorno a quel tema. L'*hashtag* funge dunque da centro di coordinamento dell'intero *tweet*; la scelta dell'elemento lessicale a cui associare questa funzione, spesso negoziata con altri interlocutori, è dettata anche dal grado di emotività o di ironia che esso riesce a suscitare (Zappavigna, 2015).

Il pubblico dei destinatari è dunque nella comunicazione sui *social media* il perno attorno al quale ruota l'organizzazione delle conversazioni; tutte le strategie discorsive adottate in tali interazioni mirano alla creazione e al mantenimento delle peculiari relazioni sociali che si instaurano con tale pubblico (Zappavigna, 2012); un pubblico personale, come si è detto, formato da un numero variabile ma potenzialmente molto elevato di interlocutori, che viene adattato, modulato e ridimensionato nelle singole interazioni, ad esempio attraverso l'uso della menzione (Herring - Honeycutt, 2009).

In quest'ultimo aspetto risiede una delle innovazioni più rilevanti delle conversazioni nei *social media*: tradizionalmente, è il parlato ad essere collegato a questa funzione sociale di gestione della relazione tra gli interlocutori (Berretta, 1994: 243); nella comunicazione mediata dal computer questa funzione è invece svolta dalla scrittura. Le conversazioni scritte consentono ai partecipanti di rimanere in contatto a distanza, rassicurandosi reciprocamente e costantemente sulla propria presenza connessa e negoziando continuamente le proprie relazioni interpersonali.

In questo senso, le interazioni sui *social media* sono forme di "conversazione distribuita" (Schmidt, 2014: 9), perché "dispongono a un dialogo apparentemente costante, sempre attivo nella mente degli attori. [...] Ogni messaggio è un turno che attende una replica, una conferma, un rinvio" (Pistoiesi, 2014: 375).

In aggiunta, l'analisi fattoriale ha evidenziato che le conversazioni sui *social media* hanno già sviluppato modalità comunicative differenziate: ciascuna delle due tipologie esaminate presenta infatti caratteristiche peculiari. Solo la dimensione di variazione (4) ha individuato uno spazio di variazione in cui esse occupano posizioni contigue: la dimensione di variazione (1) ci mostra che le conversazioni in Twitter, molto più di quelle in Facebook, tendono all'interattività piuttosto che alla densità informativa; le nanostorie (Wasik, 2009) di Twitter sono inoltre raccontate coinvolgendo gli interlocutori piuttosto che attraverso strategie narrative tradizionali, come invece più spesso avviene in Facebook.

Le conversazioni sui *social media* si specializzano dunque a seconda dei diversi spazi comunicativi in cui hanno luogo; tali spazi non differiscono per il canale fisico (lo scritto digitale di rete è lo stesso per i due ambienti), ma per almeno due elementi caratterizzanti: in primo luogo, per le restrizioni nella lunghezza dei testi. La tab. 1 mostra che questo elemento è l'unico tra quelli considerati che differenzia le interazioni in Twitter e in Facebook. La frammentarietà che caratterizza i micro-contenuti di Twitter non è il frutto, come avviene per la conversazione faccia-a-faccia, di vincoli temporali e della pressione della sincronia sulle scelte linguistiche (Voghera, 1992), ma di vincoli spaziali. Ogni *tweet* è il frammento di una conversazione distribuita che trova il suo senso solo nell'associazione con tutti gli altri frammenti. In secondo luogo, per il diverso modo di aggregare il proprio pubblico personale. L'uso dell'*hashtag* è una caratteristica distintiva delle conversazioni in Twitter: il suo ruolo di coordinamento conferisce un carattere di specificità nel modo di aggregare e coinvolgere i propri destinatari attorno a temi o valori condivisi.

5. Conclusioni

Questo lavoro si è posto l'obiettivo di applicare una metodologia di tipo multidimensionale all'analisi delle conversazioni, per identificare le caratteristiche peculiari e distintive di quelle che hanno luogo sui *social media*.

L'analisi fattoriale si è rivelata una metodologia efficace, perché in grado di descrivere la variazione tra generi testuali in termini di spazio multidimensionale; la dicotomia scritto/parlato non è infatti sufficiente per descrivere la variazione tra le diverse forme di conversazione, che si differenziano per una costellazione di elementi che agiscono in modo simultaneo.

Lo spazio di variazione tra gli undici generi di conversazione analizzati ha fatto emergere quattro fattori, distinti attraverso i tratti linguistici che li caratterizzano. L'interpretazione della co-occorrenza sistematica di tali tratti ha portato alla definizione di quattro dimensioni, caratterizzate da spazi di variazione così definiti: interattività/informatività, elaborazione sintattica/espressione non strutturata di emozioni, narrazione/coinvolgimento, espressione di un'opinione soggettiva. Tale interpretazione ha messo in luce alcune caratteristiche peculiari delle conversazioni nei *social media*, legate spesso al modo in cui vengono selezionati gli interlocutori e al tipo di relazione interpersonale che con essi ciascun partecipante è in grado di instaurare e di negoziare.

L'analisi multidimensionale ha inoltre evidenziato differenze consistenti tra i due generi di conversazione specificamente analizzati, conseguenza dei diversi modi in cui essi aggregano i vari tipi di pubblico e delle restrizioni che sono loro poste nella lunghezza dei testi.

In prospettiva futura, le linee di ricerca suggerite da questo lavoro indicano la necessità di un maggiore raffinamento nella selezione dei tratti linguistici, per esplorare possibili dimensioni di variazione sfuggite a quest'analisi, oltre che l'opportunità

di un approfondimento qualitativo dei risultati ottenuti, con strumenti differenti rispetto all'analisi fattoriale.

Bibliografia

- BARON N.S. (2008), *Always on: Language in an Online and Mobile World*, Oxford University Press, Oxford.
- BAZZANELLA C. (1994), *Le facce del parlare. Un approccio pragmatico all'italiano parlato*, La Nuova Italia, Firenze.
- BAZZANELLA C. (1995), I segnali discorsivi, in RENZI L. - SALVI G. - CARDINALETTI A. (a cura di), *Grande grammatica italiana di consultazione*, Il Mulino, Bologna: 225-257.
- BAZZANELLA C. (2002), *Prototipo, dialogo e configurazione complessiva*, in BAZZANELLA C. (a cura di), *Sul dialogo. Contesti e forme di interazione verbale*, Guerini Studio, Milano: 19-34.
- BAZZANELLA C. (2010), Conversazione, in *Enciclopedia dell'italiano*, Treccani, Roma. Disponibile in: [http://www.treccani.it/enciclopedia/conversazione_\(Enciclopedia_dell'Italiano\)/](http://www.treccani.it/enciclopedia/conversazione_(Enciclopedia_dell'Italiano)/) (data ultimo accesso 24 agosto 2015).
- BERRETTA M. (1994), Il parlato italiano contemporaneo, in SERIANNI L. - TRIFONE P. (a cura di), *Storia della lingua italiana. Scritto e parlato*, vol. II, Einaudi, Torino: 239-270.
- BIBER D. (1988), *Variation across speech and writing*, Cambridge University Press, Cambridge.
- BIBER D. (1992), On the complexity of discourse complexity: A multidimensional analysis, in *Discourse Processes* 15: 133-163.
- BIBER D. (1995), *Dimensions of register variation: A cross-linguistic comparison*, Cambridge University Press, Cambridge.
- BIBER D. (2003), Variation among university spoken and written registers: A new multi-dimensional analysis, in MEYER C.F. - LEISTYNA P. (eds.), *Corpus analysis: Language structure and language use*, Rodopi, New York: 47-70.
- BIBER D. (2004), Conversation text types: A multi-dimensional analysis, in G. PURNELLE C. FAIRON - A. DISTER (eds.), *Le poids des mots: Proceedings of the 7th International Conference on the Statistical Analysis of Textual Data*, Presses universitaires de Louvain, Louvain: 15-34. Disponibile in: http://lexicometrica.univ-paris3.fr/jadt/jadt2004/pdf/JADT_000.pdf (data ultimo accesso 24 agosto 2015).
- CAFFI C. (2001), *La mitigazione. Un approccio pragmatico alla comunicazione nei contesti terapeutici*, LIT Verlag, Münster.
- COLOMB G.G. - SIMUTIS J.A. (1996), Visible conversation and academic inquiry: CMC in a culturally diverse classroom, in HERRING S.C. (ed.), *Computer-mediated communication: Linguistic, social and cross-cultural perspectives*, John Benjamins, Amsterdam/Philadelphia: 203-222.
- ERICKSON T. - HERRING S.C. (1999), *Persistent Conversation: Home page*. Disponibile in: http://www.toмери.org/HICSS_PC.html (data ultimo accesso 24 agosto 2015).
- FELLBAUM C. (1998), *WordNet: An Electronic Lexical Database*, MIT Press, Cambridge.

- FIorentino G. (2011), Scrittura liquida e grammatica essenziale, in CARDINALE U. (a cura di), *A scuola d'italiano a 150 anni dall'Unità*, Il Mulino, Bologna: 219-241.
- GRIES S.TH. (2013), *Statistics for linguistics with R*, De Gruyter Mouton, Berlin/New York.
- HALLIDAY M.A.K. (1988), On the language of physical science, in GHADDESSY M. (ed.), *Registers of written English: Situational factors and linguistic features*, Pinter, London/New York: 162-178.
- HERRING S.C. (2011), Computer-mediated conversation: Introduction and overview, in *Language@Internet* 8. Disponibile in: <http://www.languageatinternet.org/articles/2011/Herring> (data ultimo accesso 24 agosto 2015).
- HERRING S.C. - HONEYCUTT C. (2009), Beyond Microblogging: Conversation and Collaboration via Twitter, in *Proceedings of the Forty-Second Hawai'i International Conference on System Sciences (HICSS-42)*, IEEE Press, Los Alamitos: 1-10. Disponibile in: <http://ella.slis.indiana.edu/~herring/honeycutt.herring.2009.pdf> (data ultimo accesso 24 agosto 2015).
- HOROWITZ R. - SAMUELS J.S. (1987), Comprehending oral and written language: Critical contrasts for literacy and schooling, in HOROWITZ R. - SAMUELS J.S. (eds.), *Comprehending oral and written language*, Academic Press, New York: 1-52.
- HYMES D. (1984), Sociolinguistics: Stability and Consolidation, in *International Journal of the Sociology of Language*, 45: 39-46.
- MILLER V. (2008), New Media, Networking and Phatic Culture, in *Convergence: The International Journal of Research into New Media Technologies* 14(4): 387-400.
- MURRAY D.E. (1989), When the medium determines turns: turn-taking in computer Conversation, in COLEMAN H. (ed.), *Working with Language. A Multidisciplinary Consideration of Language Use in Work Contexts*, De Gruyter Mouton, Berlin/New York: 319-337.
- PISTOLESI E. (2004), *Il parlar spedito. L'italiano di chat, e-mail e sms*, Esedra, Padova.
- PISTOLESI E. (2014), Scritture digitali, in ANTONELLI G. - MOTOLESE M. - TOMASIN L. (a cura di), *Storia dell'italiano scritto. Italiano dell'uso*, vol. III, Carocci, Roma: 349-375.
- PLUTCHIK R. (2001), The Nature of Emotions, in *American Scientist* 89(4): 344-350.
- SACK W. (2000), Conversation Map: An interface for very large-scale conversations, in *Journal of Management Information Systems* 17(3): 73-92.
- SANTINI M. (2001), Text typology and statistics. Explorations in Italian press subgenres, in *Italian Journal of Linguistics* 13(2): 339-374.
- SCHMIDT J. (2014), Twitter and the Rise of Personal Publics, in WELLER K. - BRUNS A. - BURGESS J. - MAHRT M. - PUSCHMANN C. (eds.), *Twitter and Society*, Peter Lang, New York: 3-14.
- SPINA S. (2012), *Openpolitica. Il discorso dei politici italiani nell'era di Twitter*, FrancoAngeli, Milano.
- SPINA S. (2014), Il Perugia Corpus: una risorsa di riferimento per l'italiano. Composizione, annotazione e valutazione, in BASILI R. - LENCI A. - MAGNINI B. (a cura di), *Proceedings of the First Italian Conference on Computational Linguistics CLiC-it 2014*, Pisa University Press, Pisa: 354-359. Disponibile in: <http://clic.humnet.unipi.it/proceedings/vol1/CLICIT2014168.pdf> (data ultimo accesso 24 agosto 2015).

STRAPPÀFAVA C. - VALITUTTI A. (2004), WordNet-Affect: an affective extension of WordNet, in *Proceedings of 4th International Conference on Language Resources and Evaluation (LREC 2004)*, LREC, Lisbon: 1083-1086.

VOGHERA M. (1992), *Sintassi e intonazione nell'italiano parlato*, Il Mulino, Bologna.

WASIK B. (2009), *And Then There's This: How Stories Live and Die in Viral Culture*, Viking, New York.

ZAPPAVIGNA M. (2012), *The Discourse of Twitter and Social Media. How We Use Language to Create Affiliation on the Web*, Continuum, London.

ZAPPAVIGNA M. (2015), Searchable talk: The linguistic functions of hashtags, in *Social Semiotics* 25(3): 274-291.

FRANCESCA CHIUSAROLI¹

Scritture brevi e tendenze della scrittura nella comunicazione di Twitter

This paper analyses some tendencies in writing on Twitter that are connected to the specific need for brevity and economy, together with peculiar and specific linguistic features adopted as conventional internal rules. The elements of the Twitter slang – special vocabulary; abbreviations and acronyms; spelling; and paragraphematic signs and symbols, such as punctuation and emojis, hashtag and at-sign – are discussed with reference to their effects on grammar and syntax, and for the re-definition of new functions against traditional uses. This study moves on from the definition of “scritture brevi” (“short writings”) as developed in Chiusaroli - Zanzotto (<https://sites.google.com/site/scritturebrevi/>) and now at www.scritturebrevi.it. It also addresses the “language of Twitter”, especially for the Italian case, within the wider area of writing in Computer Mediated Communication.

1. *La scrittura di Twitter nel diasistema delle scritture della rete*

È noto il funzionamento di Twitter, il *social network* di Jack Dorsey modellato sulla dimensione “quantitativa” ispirata al modello dell’SMS: 140 caratteri compongono il testo/*tweet*, configurando un contesto di possibile interazione che attualmente risulta particolarmente adeguato alla comunicazione pubblica da uno a uno, o meglio, nel caso di condizioni e figure istituzionali (politiche, dello spettacolo, e così via), da uno a molti, secondo la relazione *following-follower* (<https://twitter.com/>)².

Peculiari condizioni della comunicazione di Twitter sono: lo statuto “gergale” riconosciuto al linguaggio, il mantenimento dell’interazione all’interno della comunità coesa, la particolare espressione del “parlato” in forma scritta (“digitato”), che connota le pratiche dell’*instant messaging* in rete³ ed è pienamente esplicitato nello statuto del *tweet* (nelle intenzioni assimilato all’immagine del “cinguettio”). A partire da tali premesse si osserva l’affermazione di pratiche grafiche distintive e

¹ Università di Macerata.

² Per un’utile sintesi informativa si legga l’intervista a Dorsey sul sito del *Los Angeles Times*, 18 febbraio 2009: <http://latimesblogs.latimes.com/technology/2009/02/twitter-creator.html>

³ Per la definizione di tali fenomenologie della scrittura in rete, per cui proponiamo di applicare la categoria concettuale di “scritture brevi”, procediamo dalla ricca riflessione critica sui caratteri dell’italiano del web, svolta nell’arco cronologico che muove dagli anni Settanta del Novecento. Qui citiamo, utili per la prospettiva indagata, gli interventi di Nencioni (1976), Sabatini (1982), Dardano (1994), Schwarze (2003), Maraschio-De Martino (2010), Orletti (2004), Pistolesi (2004; 2014), Arcangeli (2005), Bazzanella (2005), Berruto (2005), Lorenzetti - Schirru (2006), Fiorentino (2007), Antonelli (2007; 2014), Poli (2012), Garavelli - Suomela-Härmä (2013), Tavosanis (2011).

peculiari del mezzo, insieme alla predisposizione di un ricco repertorio di simboli e di abbreviazioni, alcuni già noti alla scrittura comune ed altri specifici (Chiusaroli, 2012; 2014c; 2014d), tali da consentire di riconoscere uno stile crittografico responsabile spesso della reticenza di molti utenti all'accesso. I testi di Twitter forniscono pertanto all'analisi linguistica una somma di materiali utili alla verifica dei principi di codifica e di funzionamento della comunicazione digitale in contesti speciali, soprattutto nel confronto con l'obbligato limite spazio-temporale in cui si determina l'applicazione delle regole. Lo speciale ambiente tecnologico (Ong, 1982) fa rilevare infatti un processo di rifunzionalizzazione di forme della scrittura riguardante tutti i livelli dell'analisi linguistica, ovvero:

1. il piano ortografico, per le strategie di confronto con la fonetica della lingua;
2. il piano morfosintattico, per la valutazione del peso informativo degli elementi linguistici nei procedimenti di selezione e combinazione, e per la dominanza della dimensione sintagmatica (il contesto) contro l'ambiguità;
3. il piano lessicale, per l'adozione di un vocabolario scritto speciale, in particolare riguardante alcune sigle, abbreviazioni e forme neologiche comuni, anche ricavate da meccanismi di interferenza;
4. il piano testuale, pragmatico e della riproduzione del parlato, per l'inserimento di elementi grafici riguardanti specificamente la resa delle componenti prosodiche fondamentali;
5. il piano più generale dell'atto linguistico, per la tipica presenza, nel *tweet*, dei classici elementi del modello postale della comunicazione (emittente, ricevente o destinatario, referente o argomento), rappresentati "fisicamente" nello scritto.

Particolare attenzione, nell'economia della comunicazione di Twitter, è rivolta al ruolo dell'*hashtag* – # – e alla funzione della chiocciola – @ – potendosi riconoscere a questi segni speciali la natura di "segnalatori di funzione" e insieme anche un precipuo ruolo nella ristrutturazione del piano sintattico del messaggio.

La presente trattazione ha lo scopo di esaminare tali caratteristiche proprie della scrittura di Twitter, collocando forme e funzioni all'interno dell'orizzonte teorico definito dal concetto di "scritture brevi" elaborato in Chiusaroli - Zanzotto (<https://sites.google.com/site/scritturebrevi/>) e ora in www.scritturebrevi.it, a partire dalla definizione:

L'etichetta "scritture brevi" è proposta come categoria concettuale e metalinguistica per la classificazione di elementi grafici come abbreviazioni, acronimi, sigle, punteggiatura, segni, icone, indici e simboli, elementi figurativi, espressioni testuali e codici visivi per i quali risulti dirimente il principio della "brevità" connesso al criterio dell'"economia". In particolare sono comprese nella categoria "scritture brevi" tutte le manifestazioni della scrittura che, nella dimensione sintagmatica, si sottraggono al principio della linearità del significante, alterano le regole morfosintattiche convenzionali della lingua scritta, intervengono nella costruzione del messaggio nei termini di "riduzione, contenimento, sintesi" indotti dai supporti e dai contesti. La categoria ha applicazione nella sincronia e nella diacronia linguistica, nei sistemi standard e non standard, negli ambiti generali e specialistici.

2. Esiste una lingua di Twitter?

Nella presente analisi proponiamo alcune considerazioni linguistiche nell'ambito di un'ipotesi di definizione della "lingua" o "linguaggio di Twitter", con esempi soprattutto in italiano e con riguardo verso la comunicazione scritta plurilingue agevolata da Internet.

A tale proposito, va detto in premessa che l'identificazione della scrittura di Twitter con la dimensione quantitativa non può, da sola, rendere ragione della specificità del mezzo. La specificità, se esiste, deve infatti certamente poter includere tutte le possibili strategie della brevità (Schwarze, 2011; Held - Schwarze, 2011 e prima Dardano *et al.*, 2008), ma soprattutto individuare le peculiarità identificative del mezzo, ovvero quegli elementi che rendono un *tweet* "indecifrabile" fuori dal contesto, e i suoi caratteri estranei alle regole ordinarie della leggibilità. Se "M'illumino d'immenso" può essere considerato testo breve o scrittura breve, l'occorrenza del verso ungarettiano su Twitter non ne decreta l'inserimento tra le espressioni nella "lingua/linguaggio di Twitter"⁴, restando, per tale caso, conservata l'identità di "citazione (letteraria)", riproducibile nel *tweet* come in altri contesti. Le frequenti pratiche di citazione, traduzione e riscrittura⁵ attivano meccanismi di sintesi interessanti per l'analisi testuale (su cui cfr. Trifone, in stampa), ma cospicuo per il piano generale dell'indagine linguistica è, a nostro avviso, soprattutto il problema della codifica dei testi di Twitter quando caratterizzati dalla tipica scrittura gergale (fig. 1), fatta di elementi grafici, morfologici e sintattici che, fuori dall'ambiente di origine, comportano le ben note difficoltà di ricezione e comprensione per i non fruitori, o per la lettura tramite processore automatico (Pennacchiotti - Popescu, 2011).

Figura 1 - Esempio di *tweet* "gergale"



⁴ Il popolare dizionario collettivo *Urban Dictionary* propone, per il termine "twitterese", la definizione generica: "The language or diction used in posting Twitter messages" (<http://it.urbandictionary.com/>; s.v.); cfr., con riferimento anche all'italiano, la nostra analisi "Ma tu lo scrivi il twitterese?": <http://www.scritturebrevi.it/2014/06/05/twitterese>.

⁵ Si vedano, a titolo di esempi, i progetti promossi da Scritture Brevi raccolti nella piattaforma *Storify* (<https://storify.com/search?q=%23scritturebrevi>) e naturalmente in www.scritturebrevi.it.

Per l'acquisizione dei dati utili sulla "lingua di Twitter" risultano inoltre non particolarmente rilevanti le abbreviazioni identificative della scrittura della generazione cosiddetta "digitale" (Stefinlongo, 2002), rare su Twitter, almeno in italiano, ed anzi perlopiù connotate negativamente dagli stessi utenti⁶, con scarsa incidenza sulla reale portata dei fenomeni di brevità indotti dal mezzo (Chiusaroli - Zanzotto, 2012a; Zanzotto - Pennacchiotti, 2012).

La presente analisi riguarda dunque soltanto secondariamente l'elenco delle cosiddette "abbreviazioni dell'SMS", o "abbreviazioni di Internet", abbondantemente riportate in rete⁷, dalle qualità che naturalmente rinviano alla dimensione comunicativa "globale" (Crystal, 2001; 2003; 2004; 2008; Poli, 2012). L'identità generazionale, solitamente "adulta", degli utenti di Twitter in Italia, come pure l'ambiente professionale e tendenzialmente colto (per certi versi "elitario"), determinano lo scarso uso delle classiche forme del *texting* – fonetiche o fonotattiche, come <k> al posto di <ch>, o <nn> al posto di <non> – comunque circoscritte a poche discontinue contingenze, eventualmente accolte nell'uso, ma non esclusive né distintive del mezzo.

3. *Lessico di Twitter*

Nella categoria "abbreviazioni di Twitter" si ritiene opportuno collocare a giusto titolo le sigle specifiche del *social network*, che non a caso risultano incomprensibili al neofita e oscure al non utente, tali da indurre la predisposizione in rete di glossari dedicati⁸. Tali elementi, che soprattutto occorrono come prestiti dall'inglese, consentono innanzi tutto di riconoscere alla "lingua di Twitter" un imprescindibile fondamento nella scrittura, con conseguenze incidenti sulle regole d'uso.

Per la rilevanza sulla grammatica e sulla terminologia, citeremo qui i casi delle sigle più popolari – ottenute tramite procedimento acronimico da iniziali di parole o di morfemi: *MT* = *Modified Tweet*; *DM* = *Direct Message*; *FF* = *Follow Friday*; *TT* = *Trending Topic*; *RT* = *Retweet* (Chiusaroli, 2014c).

Negli impieghi di tali sigle si notano l'invariabilità formale rispetto al valore grammaticale e il tipico trattamento dei fenomeni di riduzione (Thornton, 2004; Grossmann - Rainer, 2004; Di Pace - Pannain, 2012), nonché i meccanismi di identificazione deduttiva delle categorie morfo-lessicali (sostantivi, verbi e loro varianti flesse) sulla base del contesto. La disambiguazione viene regolarmente consentita dalle condizioni sintattiche e dalla catena sintagmatica (Chiusaroli, 2014a; 2014b), presupposti che governano anche i processi della selezione/sottrazione degli ele-

⁶ Si veda, come esempio, il caso del sondaggio svolto dal quotidiano *Repubblica* su Twitter con *hashtag* #laparolacheodio, su cui cfr. <http://www.scrittorebrevi.it/2013/06/29/laparolacheodio>.

⁷ A partire da Wikipedia:

http://it.wikipedia.org/wiki/Gergo_di_Internet#Il_gergo_comune_di_internet

⁸ *In primis* quello predisposto dallo stesso *Centro assistenza* di Twitter, con l'eloquente premessa "Il nostro gergo non ti è chiaro?": <https://support.twitter.com/groups/50-welcome-to-twitter/topics/204-the-basics/articles/464314-il-glossario-di-twitter>.

menti funzionali come articoli e preposizioni, rispetto ai quali si definiscono le diverse letture:

- (1) *DM (Direct Message)*: N: Scrivimi in DM; V: DM la tua mail
- (2) *RT (Retweet)*: N: Il mio RT/ i miei RT; V: I primi 10 che RT li seguò! / Caro, RT il mio tweet / Cari, RT il mio tweet
- (3) *FF (Follow Friday)*: Oggi è #FF! / #FF per @XY; V: #FF @XY

Accanto a fenomeni classici di interferenza, che evidentemente decretano la primazia del modello inglese (come, sopra, *DM* per *Messaggio Diretto*), non mancano problemi di trasferimento interlinguistico nel caso di sigle coalescenti. Ad esempio, *HT* in italiano è ormai convenzionalmente riconosciuta e usata come abbreviazione di “hashtag”, mentre in inglese la stessa sigla rende la forma idiomatica *Hat Tip*, per indicare invece l’attribuzione di paternità di un testo o fatto citato⁹.

Il carattere gergale degli esempi conferma la specializzazione della scrittura di Twitter, lì dove le forme indicate restano consuetamente ignorate fuori dell’ambiente del *social network*, con ulteriori possibili casi di collisioni omografiche. L’interpretazione di TT/tt come “tutto/i”, nota e codificata nel *texting*, appare ad esempio ufficialmente soppiantata, in Twitter, dalla “corretta” corrispondenza *Trending Topic* (per la graduatoria delle “tendenze”), rilevante per numero di occorrenze e affermata dalla convenzione interna al mezzo:

- (4) Se inserisci #scritturebrevi arriviamo in TT (tt)
- (5) #scritturebrevi per TT (tt)

Oltre al repertorio delle sigle, andrà aggiunta all’analisi la considerazione di un vocabolario specialistico dedicato, composto di forme ed espressioni lessicali che indicano le attività dell’utente su Twitter. Insieme a prestiti dall’inglese che il dizionario dell’italiano ha recentemente acquisito¹⁰ – come *follower* e soprattutto *tweet* – a questo punto divenute ufficialmente parole della nostra lingua, si registra l’occorrenza di forme non normate di prestiti integrati, apprezzabili per la brevità, che si realizzano ancora sul piano della scrittura – es. le grafie fonetiche *twit/tuit* al posto di *tweet*, o *ritweet/retwit/ritwit/rituit* e *retweetare/ritweetare/ritwittare/rituittare* al posto di *retweet* e *retwittare* – oppure calchi di impiego gergale come l’irregolare *stellinare* per indicare l’atto di “aggiungere un *tweet* ai preferiti” (poco pratica parafrasi da *to prefer*) ispirato alla figura della “stellina” predisposta dal sistema per la collegata operazione¹¹, o l’altrettanto atipico cuorare, relativo all’apposizione del nuovo simbolo del cuore (da novembre 2015).

⁹ Cfr. *Urban Dictionary*, s.v.: <http://www.urbandictionary.com/define.php?term=hat+tip>. Si rimanda anche alla trattazione in “HT, what’s that?”: <http://www.scritturebrevi.it/2013/08/22/ht-whats-that-che-carattere>.

¹⁰ Per l’inglese cfr. il nostro “L’OED su Twitter”: <http://www.scritturebrevi.it/2013/06/13/loed-su-twitter-che-carattere/>; per l’italiano Chiusaroli, 2014c.

¹¹ “Attenti a quei due. Neologismi per Scritture Brevi”: <http://www.scritturebrevi.it/2013/06/15/attenti-a-quei-due-neologismi-per-scritturebrevi>.

4. *La scrittura dell'hashtag*

La questione del vocabolario introduce l'argomento del ruolo fondamentale dell'*hashtag* – #¹² – in Twitter (Chiusaroli, 2014d), il simbolo più caratteristico e innovativo della scrittura in rete, per i potenti effetti sulle dinamiche sociali della comunicazione (Niola, 2014), e altresì – ciò che qui maggiormente interessa – per le conseguenze non meno rilevanti sul fronte della scrittura, in quanto testo e ortografia.

La peculiarità più interessante dell'*hashtag* è nella capacità ipertestuale, una qualità di elemento “aggregatore” che tecnicamente è assolta con l'abbandono delle regole ordinarie della punteggiatura e soprattutto con la pratica della *scriptio continua* elevata a regola. Tipica manifestazione è la composizione di stringhe polimorfiche complesse (#scritturebrevis, #fatespazio, #adottaunsegno), dimostratesi particolarmente idonee alle esigenze della comunicazione pubblica odierna: #labuonascuola, #lavoltabuona, #passodopopasso, #arrivoarrivo (Spina, 2012. In particolare sul “caso Renzi” Grandi - Tallei, 2014).

Per la considerazione della struttura dell'*hashtag*, l'inscindibilità risulta dunque indispensabile, poiché il simbolo del cancelletto può mantenere insieme le parti a patto che queste non contengano elementi “scissori”. Simile presupposto comporta, per conseguenza, la tipica attivazione di processi di accorpamento/agglutinazione allo scopo di evitare l'impiego di elementi grafici inibitori dell'unità formale, e naturalmente, ed innanzi tutto, al fine di eliminare l'inadeguato spazio bianco separatore. È significativa la perdita di funzione dei simboli di unione più tradizionali “&” oppure il classico trattino medio “-”, inutilizzabili per l'*hashtag* composto in quanto inidonei a preservarne l'unità formale. Per le funzionalità collegate ai linguaggi dell'informatica resta invece efficace il trattino basso (*underscore*), benché statisticamente di limitata occorrenza rispetto alla prevalente *scriptio continua*, mentre, come vedremo, la maiuscola si colloca al rango di variante, soprattutto utile per la funzione delimitativa e di stile: #scritturebrevis; #scritture_brevi; *scritture-brevi; *scritture&brevi; *#scritture brevis

La decadenza funzionale dei segni di punteggiatura e diacritici intacca così varie regole fondamentali della scrittura tradizionale, alterando lo standard ereditato (Mortara Garavelli, 2007; 2008). Ad esempio, avendo l'apostrofo il limite di rompere la coesione formale, potremo avere più preferibilmente in sua vece il trattino basso, che visualizza e rappresenta concretamente lo spazio costituito dall'apostrofo stesso, e in alternativa ancora la scrittura continua, più ricorrente, nonostante l'accresciuta difficoltà di decifrazione: #letàlirica / #l_etàlirica / #l_età_lirica

Impedisce la costituzione del link il numerale, inefficace allorché isolatamente legato all'*hashtag* (*#2015), mentre la medesima categoria appare funzionale se seguita o preceduta da caratteri alfabetici (#capodanno2015 #2015capodanno). Rispetto alla necessità di citare una cifra numerica, idonea soluzione alternativa

¹² Definizione e istruzioni riportate dal *Centro assistenza* di Twitter: <https://support.twitter.com/articles/253564>.

sarà la scrittura per lettera, sempre priva di spaziature: #duemilaquindici, #benvenuto2015.

Ancora dalle convenzioni dell'informatica deriva la perdita della rilevanza della maiuscola, ovvero l'assenza di un reale *discrimen* tra maiuscola e minuscola nei procedimenti della scrittura o selezione automatica dei testi, rinviando, il diverso corpo, ai medesimi risultati del motore di ricerca, con conseguente effetto sull'uso non sorvegliato dell'opposizione rispetto alle regole ortografiche dello standard: #ScrittureBrevi = #scritturebrevi = #Scritturebrevi.

Sempre con effetti di deterioramento delle regole ortografiche si rileva la perdita della funzionalità dell'accento rispetto alla dimensione del collegamento ipertestuale, così che appare sistematicamente affermata la soppressione dei caratteri accentati, ridotti a varianti, anche in contesti "ufficiali", come è stato il caso di #maturita2014, dato insieme e al posto di #maturità2014 anche nei *tweet* di *account* istituzionali¹³.

L'accento può essere utilmente scritto nei casi di possibile ambiguità, ma, come avviene per la maiuscola, esso resterà privo di efficacia formale, dando origine a nuove inusitate omografie, come per #festadelpapà coincidente con #festadelpapa¹⁴, una uniformazione automatica importante per la ricaduta sul senso.

Non essendo contemplate dal sistema, sono del tutto assenti varianti stilistiche come differenze tra font o corsivi e grassetto, e dunque le funzioni usualmente annessi a tali notazioni nella scrittura saranno eventualmente rappresentate attraverso l'introduzione di elementi grafici o logonimici espliciti (es. le virgolette per le citazioni oppure l'inserimento dell'abbreviazione "cit.").

5. La lettura dell' *hashtag*

La perdita di rilievo per i tradizionali strumenti ortografici, e la tendenza comune a formare *hashtag* nelle produzioni (*tweet*) individuali (livello saussuriano della *parole*), sono condizioni capaci di complicare i procedimenti di identificazione del senso, per le già viste collisioni delle forme omografiche, per le difficoltà di lettura della *scriptio continua* (es. il contestatissimo #tutoglioincludo di una recente campagna sindacale¹⁵), o per plurima segmentazione e l'ambiguità (dedicato alla morte di Margaret Thatcher, l'*hashtag* #nowthatchersdead era anche leggibile come "now that Cher's dead"¹⁶). Tra le soluzioni ai problemi di riconoscimento vi è, ad esempio, il recupero dell'opposizione maiuscola/minuscola, ovvero la sua rifunzionalizzazione come espediente per definire, almeno a livello di "struttura superficiale", la distintività percettiva/ricettiva nella lettura e decodifica dell'*hashtag* composto.

¹³ "A proposito di Maturità": <http://www.scritturebrevi.it/2014/06/18/a-proposito-di-maturita-e-dellaccento-che-carattere>.

¹⁴ "Un regalo per papà": <http://www.scritturebrevi.it/2015/03/19/un-regalo-per-papa>.

¹⁵ http://www.repubblica.it/politica/2014/10/21/news/_tutoglioincludo_storia_e_parentele_di_un_incomprensibile_hashtag_cgil-98649029.

¹⁶ <http://www.dailymail.co.uk/news/article-2306137/Margaret-Thatcher-dead-Confused-CHER-fans-thought-shed-died-nowthatchersdead-Twitter-hashtag.html>.

Ad esempio, le forme #narrArte #narrarTe identificano, attraverso le maiuscole, due differenti universi del discorso (il primo dedicato all'arte, il secondo di marca letteraria).

6. L'elemento prosodico nella scrittura: emoticon, emoji

Speciale rilevanza assumono elementi con funzione paragrafematica per la disambiguazione del tono, in particolare i cosiddetti *emoticon* o *smiley* (Dresner - Herring, 2010). Di tali segni, aventi qualità logo- o pittografica, andrà notata l'origine dalla scrittura da combinazioni di punteggiatura, diacritici e lettere – es. il classico “duepunti trattino-parentesi” per il “sorriso” :-) – cui si associa la graduale, ormai pressoché definitiva, perdita di consistenza dell'originaria percezione dell'ordine sequenziale in favore della visualizzazione “globale” :-) > ☺ (tendenza già comune nella lettura di grafie alfabetiche, su cui, ad es. Cardona, 1981).

Non soltanto nel caso di Twitter, il repertorio dei pittogrammi si è presto arricchito a comprendere, oltre alle “faccine”, immagini di referenti della realtà quotidiana, per i quali i dispositivi digitali hanno allestito supporti dedicati, le cosiddette “tastiere emoji”, che consentono la digitazione senza il passaggio intermedio delle sequenze dei segni di punteggiatura. Agli emoji è deputato il ruolo più articolato della rappresentazione dell'emozione ed insieme della pratica espressione dei contenuti, nella sostanza assumendo, nella comunicazione, la funzione referenziale, denotativa e connotativa. La fortuna della scrittura per immagini in rete risulta confermata dalla verifica dell'altissima frequenza d'uso (registrata su www.emojitracker.com), o recentemente dall'occasione del riconoscimento (seppure per impropria o errata attribuzione) all'icona del cuore del ruolo di “parola” più cercata in rete nell'anno 2014¹⁷, con conseguente necessaria riflessione su un possibile nuovo statuto di “parola”. I limiti e le potenzialità (Chiusaroli, 2015 e in stampa) si riscontrano in vari esperimenti di scrittura e traduzione di testi complessi, come *Emojibama* (traduzione in emoji dei discorsi di Obama)¹⁸, o la tentata apertura di un *social network* come *Emoji* (<http://www.emoji.li/>). Si tratta di proposte fondate sull'intento della sostituzione totale delle scritture alfabetiche e, dunque, di luoghi idonei per valutare i gradi della trasponibilità, linguistica, interlinguistica e intersemiotica, delle immagini, della pluralità categoriale nella dimensione grammaticale, ovviamente della capacità sintetica, della “riduzione”, in senso etimologico, della lingua a repertorio lessicale (Chiusaroli, 1998; 2001; 2007) in prospettiva universalista (Chomsky, 1966; Eco, 1993; Burkhardt, 1987).

¹⁷ Come rileva l'analisi in <http://www.dailymail.co.uk/news/article-2889949/Heart-emoji-widely-used-character-social-media-blogs-global-news-outlets-2014.html>.

¹⁸ <http://www.theguardian.com/us-news/ng-interactive/2015/jan/20/-sp-state-of-the-union-2015-address-obama-emoji>.

7. *La struttura sintattica*

La proprietà dell'*hashtag* di costituirsi quale elemento formalmente isolato comporta conseguenze interessanti rispetto alla costruzione della stringa sintattica, vuoi per la possibilità di entrare nel messaggio come parte testuale pertinente ed “interna” o “integrata”, vuoi per la natura di formalizzatore del *topic* e/o del *comment*, in tal senso con trattamento come struttura separata dal corpo del messaggio.

Le possibili applicazioni mostrano l’attribuzione delle più varie categorie grammaticali alle parole con *hashtag* con conseguenze sulla dinamica testuale, dal momento che l'*hashtag* può valere come forma unica e unitaria, sintagmatica o sintematica, oppure nell’economia del *tweet* le sue componenti possono essere recuperate e trattate per il loro valore linguistico lessicale. Le diverse circostanze d’uso e, *in primis*, la necessità dell’adesione al *topic*, determinano così la costante alterazione della morfologia ordinaria della lingua (6), di pari passo con gli impieghi normati (7), con conseguenze significative per l’annotazione (Nakagawa - Mori, 2003):

(6) Ho scritto una #scritturebrevi bellissima

(7) Le tue #scritturebrevi sono bellissime

Tra i possibili esiti dello speciale trattamento morfosintattico è la frequente attribuzione all'*hashtag* della funzione di “soggetto” della frase con categoria di nome proprio, e la personificazione nel ruolo di “parlante/emittente”:

(8) #scritturebrevi è felice

Ugualmente in deroga alla regola grammaticale appare l’opzione per la forma invariabile nel caso di alternative flesse, preferenza necessitata dal sistema della condivisione, con conseguente riduzione delle forme flesse e la corrispondente tendenza verso le strutture isolanti, a salvaguardia del canale comunicativo:

(9) #sonounlettore appassionato di romanzi gialli

(10) #sonounlettore appassionata di romanzi gialli

Per contro, l’adesione per adeguamento alla morfologia ordinaria determina inevitabilmente la separazione dalla filiera comune e la formazione di percorsi di interazione disgiunti ed autonomi. Diversamente dagli esempi (9) e (10), nei casi (11) e (12) il rispetto della notazione di genere ha come effetto la creazione di un duplice canale di discussione (maschile vs. femminile), a svantaggio della condivisione collettiva consentita dall'*hashtag* comune:

(11) #sonounlettore di romanzi gialli

(12) #sonounalettore di romanzi gialli

Tra gli *hashtag* possono comparire forme composte a base verbale. Anche in tal caso la struttura “immodificabile” comporta la creazione di strutture incorporanti e la conseguente sostantivizzazione della stringa con cancelletto o, viceversa, il mante-

nimento della proprietà di verbo, ma necessariamente indeclinabile, con eventuali limitazioni espressive:

- (13) #fatespazio alla virgola! / Questo è un #fatespazio;
 (14) C'è chi #adottaunsegno all'ora, anch'io adottato un segno allora! / Questo è un #adottaunsegno per #scritturebreve

Parimenti conseguente appare l'effetto operato sulla norma sintattica dalla costruzione lineare privata degli elementi funzionali e di accordo nei molti casi in cui l'*hashtag* appare liberamente collocato rispetto al testo, frequentemente senza segni di separazione:

- (15) Forma interessante per #scritturebreve = Forma interessante. #scritturebreve =
 Forma interessante #scritturebreve

8. *La scrittura dell'account*

La deriva sintattica osservabile negli esempi è ulteriormente accentuata dall'occorrenza dell'*account*, elemento preceduto dalla chiocciola (per la storia del segno Arcangeli, 2015) che corrisponde all'indirizzo del nome dell'utente iscritto alla piattaforma *social* di Twitter. La speciale rilevanza riconosciuta agli attanti della comunicazione (mittente e ricevente) si trova nella regola dell'inclusione dell'*account* destinatario nel corpo del *tweet*, come nella scelta originaria di fissare a 140 caratteri il limite massimo di testo, rispetto ai 160 dell'SMS, sottraendo idealmente 20 caratteri per il nome *account* dello scrivente¹⁹.

Negli esempi che seguono (16, 17, 18) osserveremo in particolare il ruolo della sintassi per la definizione funzionale dell'*account*, ovvero, a seconda della collocazione nel *tweet*, la differenziazione della dimensione comunicativa tra i due estremi del massimo indirizzamento e della massima generalizzazione.

Mentre l'infinito contesto del *social medium* determina la naturale dispersione delle interazioni, la fortuna del testo prodotto può contare essenzialmente sulla ricerca di spie testuali significative, rispondenti a regole (metalinguistiche) e a rituali che fanno emergere la comunicazione dalla trasmissione neutra indirizzandola al giusto pubblico ed alla appropriata dimensione dialogica (Zanzotto *et al.*, 2011).

Agisce in questa direzione proprio il posizionamento dell'*account* all'interno del *tweet*, che determina la destinazione del testo vuoi ai singoli, vuoi alla più grande comunità degli utenti (Chiusaroli, 2014c). Se collocato in apertura, l'*account* è "destinatario" in senso esclusivo. Si configurano con l'*account* iniziale la risposta al *tweet* ed anche ogni messaggio "indirizzato": a tale scopo il sistema ci presenta utilmente il *tweet* con l'*account* già posizionato. Collocazioni diverse rispetto alla posizione iniziale determinano automaticamente la visualizzazione a tutti i *follower*, quali "riceventi", accanto allo specifico destinatario. Di qui la pratica convenzionale di iniziare il *tweet* con un "punto" che precede la chiocciola, elemento grafico minimo

¹⁹ Cfr. intervista a Dorsey cit. alla nota 1.

sufficiente a spostare il messaggio dalla bacheca personale a quella comune, simbolo e indicazione dell'interesse pubblico del *tweet*, oltre il singolo destinatario:

(16) @XY la Consolazione della filosofia uscirà martedì in edicola. #Boezio

(17) La Consolazione della filosofia uscirà martedì in edicola #Boezio @XY

(18) .@XY la Consolazione della filosofia uscirà martedì in edicola. #Boezio

Nei tre esempi appena citati, il primo testo sarà dunque “recapitato” all'utente XY, il secondo e il terzo si collocheranno nella bacheca comune, ma nell'esempio (18) la presenza del punto, segno minimo differenziatore, rende la transazione la più concretamente efficace per la immediata visibilità nel continuo svolgersi dello scambio. In un testo che si gioca entro i pochi caratteri consentiti, il punto iniziale è stragemma grafico minimo della massima portata comunicativa ed ha l'effetto di un potentissimo richiamo all'attenzione, capace di garantire orientamento, indirizzamento ed emanazione.

Privato della sua funzione originaria, ortografica, come segno della “pausa”, il punto appare così ri-assunto (o ri-abilitato), non solo nella percezione personale, ma anche nella dimensione tecnica, con un ruolo proprio e riconoscibile solo all'interno di Twitter, una scrittura convenzionale e gergale impossibile da decodificare fuori dal contesto.

9. Conclusioni: la grammatica di Twitter per una grammatica delle “scritture brevi”

Il caso del punto è forse il più significativo e sintomatico della prospettiva qui indagata, che riguarda il generale processo di erosione e di ridefinizione, in senso intrinsecamente funzionale, oltre che di stile, delle pratiche e delle regole della lingua scritta rispetto agli elementi dello speciale atto linguistico e comunicativo di Twitter.

Lo spazio del *tweet* manifesta dunque un particolare contesto di espressione utile per la verifica dell'efficacia di procedimenti compositivi sottoposti alle necessità della sintesi e della velocità quali consistenti condizioni regolate dalle tecnologie e dall'uso: per noi ancora una occasione per la costruzione della grammatica delle “scritture brevi” tra convenzione e sistema.

Bibliografia

- ANTONELLI G. (2007), *L'italiano nella società della comunicazione*, Il Mulino, Bologna.
- ANTONELLI G. (2014), *L'e-taliano: una nuova realtà tra le varietà linguistiche italiane?*, in GARAVELLI E. - SUOMELA-HÄRMÄ E. (a cura di), *Dal manoscritto al web: canali e modalità di trasmissione dell'italiano*, Atti del XII Congresso SILFI Società Internazionale di Linguistica e Filologia Italiana (Helsinki, 18-20 giugno 2012), Cesati, Firenze: II, 537-556.
- ARCANGELI M. (2005), *Lingua e società nell'era globale*, Meltemi, Roma.

- ARCANGELI M. (2015), *Biografia di una chiocciola. Storia confidenziale di @*, Castelvechi, Roma.
- BAZZANELLA C. (2005), Tratti prototipici del parlato e nuove tecnologie, in Burr E. (a cura di), *Tradizione ed innovazione, Atti del Convegno della Società internazionale di linguistica e filologia italiana (SILFI) (Duisburg 28 giugno-2 luglio 2000)*, Cesati, Firenze: 427-441.
- BERRUTO G. (2005), Italiano parlato e comunicazione mediata dal computer, in HÖLKER K. - MAASS C. (a cura di), *Aspetti dell'italiano parlato*, Lit Verlag, Münster/Hamburg/Berlin/Wien/London: 137-156.
- BURKHARDT H. (1987), The Leibnizian "characteristica universalis" as link between grammar and logic, in BUZZETTI D. - FERRIANI M. (eds.), *Speculative grammar, universal grammar, and philosophical analysis of language*, John Benjamins, Amsterdam/Philadelphia: 43-63.
- CARDONA G.R. (1981), *Antropologia della scrittura*, Loescher, Torino; nuova ed. 2009, UTET, Torino.
- CHIUSAROLI F. (1998), *Categorie di pensiero e categorie di lingua. L'idioma filosofico di John Wilkins*, Il Calamo, Roma.
- CHIUSAROLI F. (2001), Una trafila secentesca di *reductio*, in ORIOLES V. (a cura di), *Dal 'paradigma' alla parola. Riflessioni sul metalinguaggio della linguistica. Atti del Convegno, Università degli Studi di Udine - Gorizia, 10-11 febbraio 1999*, Il Calamo, Roma: 33-51.
- CHIUSAROLI F. (2007), 'Naturale' vs. 'Artificiale' nella linguistica secentesca, in Poli D. (a cura di), *Lessicologia e metalinguaggio, Atti del Convegno, Dipartimento di Ricerca Linguistica, Letteraria e Filologica, Università di Macerata, 19-21 dicembre 2005*, Il Calamo, Roma: II, 821-833.
- CHIUSAROLI F. (2012), Scritture brevi oggi. Tra convenzione e sistema, in CHIUSAROLI F. - ZANZOTTO F.M. (a cura di), *Scritture brevi di oggi, Quaderni di Linguistica Zero 1*, Napoli, Università degli studi di Napoli "L'Orientale", Napoli: 4-44. Disponibile in: http://www.unior.it/index2.php?content_id=8011&content_id_start=1 (data ultimo accesso 31 agosto 2015).
- CHIUSAROLI F. (2014a), Procedimenti onomaturgici e scritture brevi della rete, in *AION Linguistica 3*: 57-93.
- CHIUSAROLI F. (2014b), Relazioni paradigmatiche e sintagmatiche nella costruzione del segno e l'elemento creativo nella lingua, in *Studii de Stiinta si Cultura 10*, 2(37): 49-57. Disponibile in: http://www.revista-studii-uvvg.ro/index.php?option=com_content&view=category&id=74:volumul-x-numarul-2-37-2014&layout=default (data ultimo accesso 31 agosto 2015).
- CHIUSAROLI F. (2014c), Scritture brevi di Twitter: note di grammatica e di terminologia, in ORIOLES V. - BOMBI R. - BRAZZO M. (a cura di), *Metalinguaggio. Storia e statuto dei costrutti della linguistica*, Il Calamo, Roma: I, 435-448.
- CHIUSAROLI F. (2014d), Sintassi e semantica dell'hashtag: studio preliminare di una forma di scritture brevi, in BASILI R. - LENCI A. - MAGNINI B. (eds.), *The First Italian Conference on Computational Linguistics, CLiC-it 2014 - Proceedings, 9-10 December 2014*, Pisa University Press, Pisa: I, 117-121. Disponibile in: <http://clic.humnet.unipi.it/proceedings/Proceedings-CLiC-it-2014.pdf> (data ultimo accesso 31 agosto 2015).

- CHIUSAROLI F. (2015), La scrittura in emoji tra dizionario e traduzione, in BOSCO C. - TONELLI S. - ZANZOTTO F.M. (eds.), *The Second Italian Conference on Computational Linguistics, CLiC-it 2015 - Proceedings, 3-4 December 2015*, Trento, Fondazione Bruno Kessler, Accademia University Press, Torino., 2015, 88-93. Disponibile in: <http://www.aacademia.it/elenco-libri?acat=linguistica> (data ultimo accesso 31 dicembre 2015).
- CHIUSAROLI F. (in stampa), *Scritture brevi in emoji, dalla scrittura alla lettura*, in CHIUSAROLI F. - CICCARINI M. (a cura di), *Brevitas. Letture e scritture a confronto, Atti del Workshop, Università di Roma "Tor Vergata", 25-26 febbraio 2015*.
- CHIUSAROLI F. - ZANZOTTO F.M. (2012), Informatività e scritture brevi del web, in CHIUSAROLI F. - ZANZOTTO F.M. (2012), (a cura di), *Scritture brevi nelle lingue moderne, Quaderni di Linguistica Zero 2*, Napoli, Università degli studi di Napoli "L'Orientale", Napoli: 3-20. Disponibile in: http://www.unior.it/index2.php?content_id=8011&content_id_start=1 (data ultimo accesso 31 agosto 2015).
- CHOMSKY N. (1966), *Cartesian linguistics: a chapter in the history of rationalist thought*, Harper & Row, New York; 2009, 3rd ed. with a new introduction by MCGILVRAY J., Cambridge University Press, Cambridge.
- CRYSTAL D. (2001), *Language and the Internet*, Cambridge University Press, Cambridge.
- CRYSTAL D. (2003), *English as a global language*, II ed., Cambridge University Press, Cambridge.
- CRYSTAL D. (2004), *A glossary of netspeak and textspeak*, Edinburgh University Press, Edinburgh.
- CRYSTAL D. (2008), *Txtng. The Gr8 Db8*, Oxford University Press, Oxford.
- DARDANO M. (1994), *Come parlano gli italiani*, La Nuova Italia, Firenze.
- DARDANO M. - FRENGUELLI G. - DE ROBERTO E. (2008) (a cura di), *Testi brevi. Atti del Convegno internazionale di studi, Università di Roma Tre, 8-10 giugno 2006*, Aracne, Roma.
- DI PACE L. - PANNAIN R. (2012), Sigle, acronimi: dimensione grafica e statuto lessicale, in CHIUSAROLI F. - ZANZOTTO F.M. (2012b): 208-228. Disponibile in: http://www.unior.it/index2.php?content_id=8011&content_id_start=1 (data ultimo accesso 31 agosto 2015).
- DRESNER E. - HERRING S.C. (2010), Functions of the non-verbal in CMC: emoticons and illocutionary force, in *Communication Theory* 20: 249-268.
- ECO U. (1993), *La ricerca della lingua perfetta nella cultura europea*, Laterza, Roma/Bari.
- FIorentino G. (2007), Nuove scritture e media: le metamorfosi della scrittura, in FIorentino G. (a cura di), *Scrittura e società. Storia, cultura, professioni*, Aracne, Roma: 175-207.
- GARAVELLI E. - SUOMELA-HÄRMÄ E. (2013) (a cura di), *Dal manoscritto al web: canali e modalità di trasmissione dell'italiano, Atti del XII Congresso SILFI Società Internazionale di Linguistica e Filologia Italiana (Helsinki, 18-20 giugno 2012)*, Cesati, Firenze.
- GRANDI M. - TALLEI R. (2014), *#Arrivo.Arrivo. La corsa di @matteorenzi da Twitter a Palazzo Chigi*, Fazi, Roma.
- GROSSMANN M. - RAINER F. (2004) (a cura di), *La formazione delle parole in italiano*, Niemeyer, Tübingen.

- HELD G. - SCHWARZE S. (2011) (a cura di), *Testi brevi. Teoria e pratica della testualità nell'era multimediale*, Peter Lang, Frankfurt.
- LORENZETTI L. - SCHIRRU G. (2006), La lingua italiana nei nuovi mezzi di comunicazione: SMS, posta elettronica e Internet, in GENSINI S. (a cura di), *Fare comunicazione. Teoria ed esercizi*, Carocci, Roma: 71-98.
- MARASCHIO N. - DE MARTINO D. (2010) (a cura di), *Se telefonando... ti scrivo. L'italiano al telefono, dal parlato al digitato. I giovani e la lingua. Atti dei Convegni, Firenze, 11 maggio-26 novembre 2007*, Accademia della Crusca, Firenze.
- MORTARA GARAVELLI B. (2007), *Prontuario di punteggiatura*, Laterza, Roma/Bari.
- MORTARA GARAVELLI B. (2008), *Storia della punteggiatura in Europa*, Laterza, Roma/Bari.
- NAKAGAWA H. - MORI T. (2003), Automatic term recognition based on statistics of compound nouns and their components, in *Terminology* 9(2): 201-219.
- NENCIONI G. (1976), Parlato-parlato, parlato-scritto, parlato-recitato, in *Strumenti critici* 10: 1-56; rist. in NENCIONI G. (1983), *Di scritto e parlato, Discorsi linguistici*, Zanichelli, Bologna: 126-179.
- NIOLA M. (2014), *Hashtag. Cronache di un paese connesso*, Bompiani, Milano.
- ONG W.J. (1982), *Orality and literacy. The technologizing of the word*, Methuen, London/New York; trad. it. (1986), *Oralità e scrittura. Le tecnologie della parola*, Il Mulino, Bologna.
- ORLETTI F. (2004) (a cura di), *Scrittura e nuovi media. Dalle conversazioni in rete alla Web usability*, Carocci, Roma.
- PENNACCHIOTTI M. - POPESCU A.M. (2011), A machine learning approach to Twitter user classification, in ADAMIC L.A. - BAEZA YATES R.A. - COUNTS S. (eds.), *Proceedings of the Fifth International Conference on Weblogs and Social Media, Catalonia, Spain, July 17-21*, The AAAI (Association for the Advancement of Artificial Intelligence) Press, Menlo Park, California: 281-288.
- PISTOLESI E. (2004), *Il parlar spedito. L'italiano di chat, e-mail e SMS*, Esedra, Padova.
- PISTOLESI E. (2014), Scritture digitali, in ANTONELLI, G. - MOTOLESE, M. - TOMASIN, L. (a cura di), *Storia dell'italiano scritto. Italiano dell'uso*, vol. III, Carocci, Roma: 349-375.
- POLI D. (2012), Pensiero elaborato e modalità esperienziali: una convergenza per una nuova gestione del sapere linguistico collettivo in Italia, in ORIOLES V. (a cura di), *Per Roberto Gusmani. Linguistica storica e teorica*, Forum, Udine: II, 337-367.
- SABATINI F. (1982), La comunicazione orale, scritta e trasmessa: la diversità del mezzo, della lingua e delle funzioni, in BOCCAFURNI A.M. - SERROMANI S. (a cura di), *Educazione linguistica nella scuola superiore. Sei argomenti per un curriculum*, Consiglio Nazionale delle Ricerche, Roma: 105-127.
- SCHWARZE S. (2003), La metamorfosi della lettera: epistola vs. e-pistola, in RAINER F. - STEIN A. (a cura di), *I nuovi media come strumenti per la ricerca linguistica*, Peter Lang, Frankfurt: 141-155.
- SCHWARZE S. (2011), Brevità e “testo breve” – stimoli interpretativi in chiave della linguistica testuale, in HELD G. - SCHWARZE S. (2011) (a cura di), *Testi brevi. Teoria e pratica della testualità nell'era multimediale*, Peter Lang, Frankfurt: 15-30.

SPINA S. (2012), *Openpolitica. Il discorso dei politici italiani nell'era di Twitter*, Franco Angeli, Milano.

STEFINLONGO A. (2002), *I giovani e la scrittura: attitudini, bisogni, competenze di scrittura delle nuove generazioni*, Aracne, Roma.

TAVOSANIS M. (2011), *L'italiano del web*, Carocci, Roma.

THORNTON A.M. (2004), Riduzione, in GROSSMANN M. - RAINER F. (2004): 555-566.

TRIFONE P. (in stampa), L'italiano 'twitterario', in DOVETTO F.M. - LEONARDI S. - MANCO A. - MORLICCHIO E. (a cura di), *Scritture Brevi e variabilità diatecnica, Convegno interannuale PRIN SCRIBE, Università di Napoli L'Orientale, Scuola Alta Formazione Procida, 26-28 giugno 2014*.

ZANZOTTO F.M. - PENNACCHIOTTI M. (2012), Language evolution in social media, in CHIUSAROLI F. - ZANZOTTO F.M. (2012) (a cura di), *Scritture brevi nelle lingue moderne, Quaderni di Linguistica Zero 2*, Napoli, Università degli studi di Napoli "L'Orientale", Napoli: 209-230. Disponibile in: http://www.unior.it/index2.php?content_id=8011&content_id_start=1 (data ultimo accesso 31 agosto 2015).

ZANZOTTO F.M. - PENNACCHIOTTI M. - TSIOUTSIOLIKLIS K. (2011), Linguistic redundancy in Twitter, in *Proceedings of the 2011 Conference on Empirical Methods in Natural Language Processing, Edinburgh, Scotland, UK, July 27-31*, John McIntyre Conference Centre, A meeting of SIGDAT, Association for Computational Linguistics, Edinburgh, UK: 659-669. Disponibile in: http://www.researchgate.net/publication/221012842_Linguistic_Redundancy_in_Twitter (data ultimo accesso 31 agosto 2015).

PARTE III

TECNOLOGIE PER APPRENDERE

AMELIA BANDINI¹

Wiki-eTandem: un progetto di apprendimento collaborativo a distanza

In formal language learning contexts, “in-tandem learning” activities can help language learners to overcome the lack of opportunities of direct contacts with target language speakers and to focus on specific linguistic forms and thus to become aware of their own language competences. The integration of e-tandem activities into an institutional course can therefore be a solution for combining formal and informal language learning. This paper outlines an e-Tandem experience carried on between the Universities of Naples and Bochum (DE) and focuses particularly on the analysis of written productions. In fact, although e-Tandems are generally planned to improve oral skills, Wiki-eTandem aimed to promote activities involving the four language skills, through reading and writing tasks and face-to-face “e-interaction”. By observing students’ production and particularly by considering their writings and the feedbacks that they gave to each other as expressions of personal learning styles and goals, the teachers were able to redefine learning activities with a view to increasing learner-centered teaching.

1. *Presupposti teorici*

I corsi istituzionali di lingua straniera (LS) possono essere stimolati ed arricchiti, come è noto, da occasioni di contatto e collaborazione con madrelingua, finalizzate a promuovere la fase di produzione linguistica. Questa, concepita secondo il modello di Krashen (1981) come esito del processo di acquisizione, è stata invece proposta in altri filoni di ricerca come ulteriore momento di apprendimento (Swain, 1985). In fase di *output* si innescano processi cognitivi di autovalutazione in quanto gli apprendenti testano le proprie competenze ed individuano le eventuali lacune da colmare, sviluppando in tal modo meccanismi di autonomia. Tali procedimenti si attiverebbero, in base alla *Output Hypothesis* di Swain, soprattutto nella fase di scrittura, durante la quale si ha il tempo di riflettere sul proprio prodotto, sia dal punto di vista delle scelte lessicali, che della correttezza formale e dell’organizzazione testuale.

Ulteriori studi si sono concentrati anche su altre forme di produzione, evidenziando gli effetti positivi dell’interazione come stimolo alla riflessione (meta) linguistica. Gass e Varonis (1985b) confrontano sequenze interazionali fra copie di apprendenti formate da parlante nativo/non nativo e da parlante esperto/non esperto, ambedue non-nativi, evidenziando una più approfondita riflessione

¹ Università degli Studi di Napoli Federico II.

metalinguistica nel caso delle coppie non-native. Long (1996) sottolinea come l'interazione faciliti l'apprendimento della L2 perché connette in modo produttivo l'input, le capacità di comprensione dell'apprendente e l'output. In questo senso si può dire che l'interazione stimola lo sviluppo dell'autonomia, infatti "our autonomy develops not in isolation but through interaction with other people, and consequently autonomous language learning schemas will fail unless they are firmly embedded in social-interactive processes" (Little, 2002: 18). In sintesi la "conversational interaction in a second language forms the basis for the development of language rather than being only a form for practice of specific language features" (Gass, 2005: 234).

L'incremento di attività interazionali nel processo di apprendimento di una LS va anche nella direzione suggerita dai più recenti approcci acquisizionali al docente di lingua, che deve cercare di offrire un input cospicuo, comprensibile, interessante e motivante, "ma anche interattivo e coinvolgente, stimolando il discente a intervenire e a produrre output" (Chini, 2011: 11) e a confrontare le sue produzioni con l'*input* ricevuto.

In contesti istituzionali, però, è proprio la fase di produzione linguistica a presentare le maggiori problematiche. I corsi universitari, ad esempio, sono tradizionalmente organizzati in modo formale, seguono un approccio istruttivo esplicito e propongono percorsi guidati, focalizzati sulla forma linguistica. Studi teorici e applicativi (Della Putta - Visigalli, 2010; Matthiae, 2012) suggeriscono infatti che l'approccio cognitivo, che stimola la riflessione metalinguistica, influenza positivamente il processo formativo (Long, 1988), in quanto il lavoro esplicito su forma e regole grammaticali della L2 contribuisce all'acquisizione di essa, particolarmente nel caso di apprendenti adulti (Chini, 2005; 2011). In contesti di apprendimento formali la fase di *output* viene penalizzata in quanto per un verso, la presenza del docente innalza il filtro affettivo ed inibisce l'attività produttiva, e per un altro, l'interazione all'interno della stessa classe genera l'insorgere di forme di comunicazione non autentica ed inaccurata (Cziko, 2004; Lightbown - Spada, 2006). Il lavoro in-tandem con un madrelingua può costituire in questi casi una proficua opportunità per creare un contesto di interazione autentico ed informale nel quale l'apprendente, abbassando il filtro affettivo grazie alla collaborazione *peer-to-peer*, attiva processi di negoziazione spontanei e naturali che favoriscono il suo processo di apprendimento (Brammerts, 1996). Ulteriori vantaggi per gli apprendenti sono costituiti sia dalla possibilità di esperire attività di apprendimento collaborativo (Telles - Vassallo, 2006), che di lavorare autonomamente. Essi infatti si aiutano vicendevolmente ad acquisire consapevolezza delle proprie competenze grazie al feedback reciproco e, allo stesso tempo, decidono e negoziano fra di loro quali attività svolgere, con quali tempi ed in quale momento, assumendosi in tal modo la responsabilità del proprio processo di apprendimento (Holec, 1979; Little, 1991).

La tecnologia del Web 2.0 offre la possibilità di realizzare molteplici forme virtuali di incontri tandem, come è testimoniato da numerose esperienze, per esem-

pio dal noto *Teletandem Brasil*, pensato per implementare l'interazione orale e, sfruttando le potenzialità della chat, la riflessione metalinguistica, nonché dal più recente *Seagull-Tandem*, creato da un *network* di università europee, che stimola i partecipanti ad interagire seguendo le attività di produzione orale e scritta proposte in 15 *topic sheets*².

Sulla base di queste considerazioni ed esperienze è nato il Wiki-eTandem, un progetto finalizzato ad implementare attività collaborative a distanza fra apprendenti l'uno la lingua madre dell'altro (tedesco/italiano). Il progetto è stato messo a punto insieme al collega Jörg Meuter della Ruhr Universität di Bochum. Le attività di *in-tandem learning* sono state proposte quale integrazione facoltativa a quelle presentate durante i corsi istituzionali. L'idea di inserire un eTandem in un corso strutturato è stata concepita principalmente per indurre gli apprendenti all'interazione e stimolare quindi la produzione linguistica. In questo modo si è tentato di creare un contesto di apprendimento integrato nel quale alla fase di lavoro consapevole, strutturato ed organizzato dal docente in sede di lezione frontale, si sono affiancati momenti di contatto linguistico diretto e autentico in un contesto sociale "naturale". Il presente studio ha lo scopo di dimostrare come le attività svolte in-tandem dai partecipanti abbiano favorito il potenziamento delle competenze in LS stimolando sia la riflessione metalinguistica, grazie al feedback del partner, che l'insorgere di un processo acquisizionale, scaturito naturalmente dal contatto con il/la madrelingua. A tal fine verranno presentate le produzioni scritte di due coppie di apprendenti ed i rispettivi feedback correttivi. Verrà inoltre mostrato come l'analisi dei feedback dei partecipanti ha offerto un utile spunto ai docenti per riformulare la propria prassi didattica in chiave *learner centred*.

2. Il Wiki-eTandem

Il lavoro in-tandem può riguardare attività sia di scrittura che di conversazione. L'interazione scritta, ad eccezione di quella via chat, è asincrona e dà il tempo di calibrare le reazioni agli input ricevuti, cioè di riflettere sulla propria ed altrui produzione, di concentrarsi sia sui contenuti che sulla forma del testo, di consultare gli opportuni sussidi. Nello scambio *face-to-face* (F2F), invece, l'interazione è immediata, nel senso che la conversazione può essere interrotta in ogni momento, è difficile concentrarsi sia sui contenuti che sulla sua forma, ma è data l'opportunità di esperire anche gli eventi paratestuali che accompagnano la comunicazione e semplificano la comprensione. Occorre però tenere presente che nell'interazione F2F via computer, la condivisione del contesto comunicativo è solo parziale, limitata cioè a quanto è possibile esperire attraverso lo schermo (Leone, 2009).

Nel Wiki-eTandem l'interazione scritta si svolge su un sito appositamente creato, realizzato con un software wiki. La decisione di usare questo tipo di software, che stimola particolarmente la collaborazione fra i partner, si riallaccia alle più

² <http://seagull-tandem.eu>.

recenti tendenze nell'ambito dei tandem linguistici, per quanto la metodologia sia già stata sperimentata in altri campi disciplinari (Judd *et al.*, 2010). I wiki infatti sono uno strumento ampiamente utilizzato nella didattica on-line in quanto sono organizzati in modo flessibile e divisibile in varie categorie e consentono di poter essere modificati e aggiornati continuamente – conservando però traccia di quanto fatto –: per questa ragione possono costituire un feedback costantemente consultabile da parte di studenti e insegnanti ed un *dossier* linguistico telematico progressivamente aggiornabile. Essi possono essere considerati risorse utili allo sviluppo di attività di *Computer Supported Collaborative Learning* (Stahl *et al.* 2006), in quanto creano situazioni inclusive che consentono l'accrescimento delle capacità di ognuno all'interno di un gruppo senza ingenerare tensione emotiva o alzare il filtro affettivo (Bedini - Balò, 2008).

La collaborazione con il *Zentrum für Fremdsprachenausbildung* si reitera ormai da cinque anni. Ogni anno viene creato un nuovo sito al quale sono invitati ad accedere solo i partecipanti di uno specifico corso. È stato scelto di limitare l'accesso al sito perché in questo modo si è voluto realizzare una *web community* costituita da coloro che condividono il progetto formativo: a nostro avviso ciò può contribuire ulteriormente all'abbassamento del filtro affettivo grazie alla creazione di un sentimento di condivisione fra coloro che partecipano al progetto.

La descrizione delle attività e gli esempi che vengono riportati in questa sede si riferiscono all'a.a. 2014-15. Come nella maggioranza dei Tandem, le coppie sono state create in base ai dati dei questionari compilati dai partecipanti. Oltre alle informazioni biografiche, si è cercato di far emergere anche il profilo linguistico degli apprendenti, non solo per quanto concerne le lingue obiettivo (Bosisio, 2012) del tandem, ma anche per quanto riguarda le altre LS possedute: a questo fine è stato proposto un *self-assessment* basato sui livelli di competenza del Quadro Comune Europeo di Riferimento (QCER). Essere a conoscenza delle competenze linguistiche dei partecipanti è stato di aiuto ai docenti soprattutto nella fase di analisi degli errori presenti nelle produzioni scritte, che è stata condotta con gli strumenti dell'analisi contrastiva, come si evidenzierà in seguito.

Tradizionalmente le attività di apprendimento in-tandem sono libere, esistono però anche esempi, come è il caso del *Seagull-Tandem* precedentemente citato, in cui le attività di apprendimento sono guidate. Considerando che il Wiki-eTandem è inserito nei nostri rispettivi corsi, ma anche che gli studenti universitari sono abituati ad essere guidati nel loro percorso di formazione, si è pensato ad un eTandem fortemente strutturato. Abbiamo proposto schede contenenti argomenti³ ed attività in riferimento alle situazioni comunicative

³ I titoli delle nove schede di lavoro presentate sul sito erano: Presento una persona famosa del mio paese, Vacanze, Tempo libero, Studiare all'estero, La mia giornata, Shopping, Cosa c'è da mangiare? Come vivo/come mi piacerebbe vivere, Vogliamo festeggiare?

previste ai livelli A1 e A2 del QCER poiché si intendeva coinvolgere non solo apprendenti con competenze linguistiche intermedie, ma anche chi possedeva un livello linguistico iniziale nella lingua obiettivo. I partecipanti, che erano tenuti a completare il lavoro di almeno quattro schede per il riconoscimento di 3 CFU, hanno svolto il lavoro in maniera assolutamente autonoma per quanto riguarda la scelta degli argomenti, i tempi e le modalità del lavoro. Una volta alla settimana il docente era disponibile nel laboratorio linguistico per quattro ore durante le quali forniva chiarimenti o assistenza a chi ne faceva richiesta.

Una volta entrati nel sito, i partecipanti trovavano le istruzioni per il lavoro ed una cartella contenente le schede, ognuna incentrata su un tema e suddivisa in tre fasi distinte, che proponevano attività collegate alle 4 abilità linguistiche. Ogni coppia creava una propria cartella all'interno del sito nella quale inseriva di volta in volta le proprie elaborazioni.

Nella prima fase, dedicata allo sviluppo della competenza di lettura e comprensione di un testo scritto, il primo compito assegnato consisteva nello scrivere un brano nella propria lingua madre sull'argomento scelto. In tal modo ognuno ha avuto a disposizione un testo nella propria lingua target e si è esercitato nella sua decodifica, venendo a contatto con il lessico e la morfologia specifici pertinenti alla tematica trattata. Il secondo compito era quello di produrre una mappa concettuale nella propria LS contenente il lessico inerente all'argomento scelto. La mappa è stata prospettata non solo come rappresentazione grafica delle competenze lessicali di ogni partecipante su un determinato *topic* – ampliate dal lavoro sul testo ricevuto dal partner – ma anche come strumento per iniziare la vera e propria collaborazione con il/la proprio/a collega, che aveva il compito di intervenire per correggere le inesattezze ed eventualmente completare la mappa. Le *mind-map* hanno inoltre promosso nel Wiki-eTandem uno scambio interculturale poiché hanno innescato curiosità verso espressioni lessicali sconosciute: è il caso, ad esempio, delle mappe sull'argomento del Natale, attraverso le quali i partecipanti italiani hanno scoperto, grazie alle integrazioni dei partner tedeschi (*der Adventkranz, der Adventskalender...*), le tradizioni tipiche del periodo dell'Avvento, non particolarmente presenti nella nostra cultura (Fig.1).

Il lavoro proposto nella prima fase delle schede è stato concepito anche come preparazione all'interazione orale – seconda fase di ciascuna scheda – nella quale i partner si sono incontrati F2F via *Skype* per parlare dell'argomento scelto; come in tutti i tandem metà del tempo doveva essere interamente devoluta ad una lingua e l'altra metà all'altra.

Infine, nella terza fase, dedicata allo sviluppo della competenza di scrittura, ognuno ha prodotto un testo scritto nella propria lingua target ed emendato quello scritto dal partner.

In tutte e tre le fasi di ogni scheda di lavoro i partecipanti sono stati invitati a dare un feedback attento e comprensibile ai propri partner.

Figura 1 - *Mappa cognitiva di una partecipante italoфона sul Natale*

Mind Map:



Al termine di ogni scheda si è proposta una fase di riflessione esplicita: ad ogni partecipante è stato infatti chiesto di riflettere autonomamente e valutare tutte e tre le fasi del lavoro svolto in termini di *input* ricevuto, *output* prodotto, soddisfazione per i risultati ottenuti, formulazione di obiettivi di apprendimento immediati, da perseguire nella successiva attività. A questo scopo è stata approntata una apposita scheda (Fig. 2) in riferimento a quanto previsto nel Portfolio Europeo delle Lingue (PEL), strumento per riflettere, (auto)valutare e progettare il proprio apprendimento (Little, 2012). Ogni portfolio elaborato e l'insieme di essi doveva infatti rappresentare, "in primis" per l'apprendente stesso, la storia linguistica dell'esperienza del Wiki-eTandem, vista attraverso la lente dello sviluppo delle competenze linguistiche e interculturali. È stato inoltre chiesto agli apprendenti di inviare di volta in volta i portfolio compilati ai docenti, che li hanno utilizzati come strumento di

orientamento per l'attività didattica in quanto in alcuni casi, come si mostrerà in seguito, hanno suggerito la trattazione e/o l'approfondimento di alcuni argomenti.

Figura 2 - Scheda di riflessione e autovalutazione del lavoro svolto



PORTFOLIO

Cognome	
Nome	
Indirizzo e-mail	
Il mio/la mia partner	

Rifletta partendo dai seguenti suggerimenti, sul lavoro in-tandem.

Usi una nuova scheda per ogni scheda di lavoro elaborata.

Titolo della scheda di lavoro:	
1° fase	
“Mi sono preparato all'incontro Skype svolgendo le seguenti attività”	
„ Il testo scritto dal partner/dalla partner è stato utile?”	
„Cosa ho imparato di nuovo?”	
2° fase	
“Quali difficoltà ho avuto durante l'incontro Skype?”	
Cosa vorrei migliorare o fare diversamente nel prossimo incontro skype?”	
„ L'aiuto datomi dal partner/dalla partner è stato come avrei voluto?”	
„Cosa ho imparato di nuovo?”	
„Cosa mi ha corretto?”	
3° fase	
“Dopo l'incontro Skype ho svolto le seguenti attività ...”	
„ L'aiuto datomi dal partner/dalla partner è stato come avrei voluto?”	
„Cosa ho imparato di nuovo?”	
„Cosa mi ha corretto?”	

Alla fine dell'intero percorso è stato chiesto ai partecipanti di redigere un proprio bilancio finale su tutto il lavoro svolto per valutare "ex-post" l'effettivo raggiungimento degli obiettivi formativi iniziali.

Ogni coppia di utenti ha dunque creato le proprie pagine all'interno del sito, ognuna dedicata alle attività su uno specifico argomento: in questo modo il sito si è arricchito progressivamente di materiali ai quali tutti i partecipanti avevano accesso. Ciò ha portato notevoli vantaggi soprattutto per quanto riguarda la fruizione di testi autentici redatti da madrelingua (*input* scritto) che non si è limitata a quanto prodotto dal rispettivo partner, ma si è ampliata ai prodotti di tutti i partecipanti.

Il software *wiki* lascia aperta a tutti gli utenti la possibilità non solo di commentare i testi nella apposita sezione "*comments*", ma anche di "editarli" modificandoli, procedendo per aggiunte, cambiando e/o cancellando ciò che è stato scritto e registra ogni modifica in una cronologia che permette di ritornare alla versione precedente del testo e osservare gli interventi che sono stati operati. Questa funzione permette agli utenti del sito, quindi anche ai docenti, di seguire diacronicamente l'evoluzione di ogni testo e quindi delle competenze linguistiche dei partecipanti, e di perseguire in tal modo un duplice fine. Da una parte le produzioni scritte ed i feedback ricevuti dai partner sono immediatamente utili sia agli apprendenti per ripensare all'input ricevuto, sia ai docenti per valutare, di volta in volta, il lavoro dei partecipanti; da un'altra esse costituiscono, sempre in riferimento al PEL, il *Dossier*, cioè l'insieme dei documenti che illustrano le esperienze e i risultati descritti nella biografia linguistica.

3. *Le produzioni scritte: opportunità per migliorare la didattica frontale*

Nel Wiki-eTandem è stato dunque dato molto spazio alla parte scritta, chiedendo ai partecipanti non solo di redigere testi nella propria lingua madre e nella lingua obiettivo, ma anche di emendare quelli dei propri partner e di compilare i portfolio.

Le produzioni scritte hanno offerto ai docenti l'opportunità di osservare ed analizzare l'interlingua (IL) degli apprendenti, fornendo spunti sia per la valutazione dei progressi dei discenti in termini di trasformazione dell'*input* in *intake*, sia per la messa a fuoco di criticità da trattare in sede frontale. L'interlingua, come è noto, costituisce infatti un sistema linguistico in continuo movimento e divenire che rappresenta in ogni sua fase la norma che il parlante stesso crea e ipotizza per la lingua straniera (Mezzadri, 2002, Chini, 2011). Infatti "l'analisi delle IL degli apprendenti può essere utile per meglio conoscerli e per adattare l'input di insegnamento al processo di acquisizione in corso. Un'analisi delle IL non fine a sé stessa, ma funzionale a tracciare il profilo dell'apprendente" (Bosisio, 2012: 118-119).

L'analisi critica delle produzioni degli apprendenti è stata condotta seguendo un duplice approccio: quello indicato da Bosisio (2012) come analisi contrastiva debole – che intende spiegare a posteriori gli errori effettuati facendo riferimento alla lingua madre ed alle altre lingue straniere possedute – ed il criterio pedagogico, secondo il quale "la classificazione degli errori riflette la distinzione chomskiana tra

competenza ed esecuzione e comprende dunque gli errori transitori (*mistakes*) e quelli sistematici (*errors*)” (Bosisio, 2012: 45). A nostro avviso, posto che nell’interlingua – almeno ad uno stato iniziale dell’apprendimento – sono presenti strutture linguistiche tipiche della lingua madre e delle altre lingue possedute, gli *errors*, scaturiti da mancata conoscenza, richiedevano interventi strutturali, mentre i *mistakes*, momentanei e non sistematici, potevano essere (auto)corretti attraverso un idoneo percorso guidato, suggerito dal docente (Mezzadri, 2002) o dal partner. Le interlingue sono state infine prese in considerazione anche diacronicamente, analizzando la trasformazione degli *input* in *intake* come testimonianza dell’evolversi delle competenze linguistiche degli apprendenti, poiché la sola analisi degli errori sembrava non potesse condurre ad una valutazione completa del processo di apprendimento.

Nell’ottica di una didattica incentrata sull’apprendente, inoltre, si è rivelata utile anche l’analisi dei *feedback* correttivi. La modalità con la quale essi venivano proposti dai discenti ai rispettivi partner ritraeva infatti, a nostro avviso, gli stili e gli obiettivi di apprendimento dei singoli partecipanti e, conseguentemente, costituiva un’utile indicazione per ri-orientare e adeguare la didassi quotidiana alle loro esigenze. Nel paragrafo che segue verranno presentati i *feedback* ed i portfolio di due coppie di apprendenti. I primi sono rappresentativi di due diverse modalità di approccio all’apprendimento, mentre le riflessioni proposte nei portfolio costituiscono una testimonianza della capacità di autovalutazione e definizione esplicita e puntuale degli obiettivi formativi da parte dei partecipanti.

4. *Presentazione ed analisi di alcune attività*

In questo paragrafo si propongono le attività svolte nella edizione 2014-15 del Wiki-eTandem da due coppie di apprendenti. Della prima si presentano stralci di alcuni testi, esempi dei *feedback* correttivi e dei portfolio. Della seconda, invece si mostreranno i soli *feedback* correttivi, allo scopo di evidenziarne le differenze tipologiche con quelli forniti dalla prima coppia, come segnale di stili ed obiettivi di apprendimento diversi.

Nel primo caso la partnership è formata da due apprendenti, IS germanofona ed AN italoфона, con competenze in diverse lingue straniere. Ambedue dichiarano di ambire ad un miglioramento delle proprie capacità espressive nella rispettiva LS: AN intende imparare nuovi vocaboli ed espressioni da poter usare eventualmente nel caso di un soggiorno in Germania, mentre IS desidera, in particolare, migliorare le proprie competenze inerenti al linguaggio di uso quotidiano e conoscere la cultura e le tradizioni italiane.

Nella prima produzione in italiano dell’apprendente germanofona IS si rilevano problematiche connesse all’uso di:

- verbi pronominali con clitico incorporato, es. (4) “rilassare”;
- modali, dei quali viene fatto un uso estensivo come nelle lingue anglosassoni e per i quali risulta difficoltosa la combinazione con il clitico, es. (1) “può fare”, (2) “può andare”, (3) “puoi andare”;

- ortografia di connettivi, es. (6) “In oltre”;
- espressioni temporali, es. (3) “Alla sera” (8) “quando Natale è presto” (9) “troppo spesso”.

Sono presenti inoltre casi di errata costruzione delle frasi complete (5) “Mi piace di passare” (7) “Amo di sperimentare”, nonché inesattezze morfologiche e lessicali, in parte derivanti da transfer dal tedesco (LM) (9) “vicino a mio fratello” e dall’inglese (LS) (10) “m’ingaggio”:

- (1) Questo può fare vicino a un lago. In estate potrebbe nuotare nella piscina all’aperto
- (2) Per esempio può andare nel ‘Ruhr’
- (3) Alla sera puoi andare in un quartiere speciale
- (4) Abbiamo un grande giardino che è perfetto per rilassare e leggere al sole
- (5) Mi piace di passare il tempo con loro
- (6) In oltre cucinare e informare mi affascino
- (7) Amo di sperimentare nuove ricette e fare le torte
- (8) E adesso quando Natale è presto mi piace informare i biscotti
- (9) Troppo spesso vado vicino a mio fratello
- (10) M’ingaggio molto nella mia parrocchia

Per quanto riguarda la classificazione delle espressioni non conformi alla norma in *errors* e *mistakes*, si nota come nel caso della costruzione delle complete, ad esempio, siano presenti nello stesso testo anche espressioni come (11) “Mi diverto ad ascoltare, mi piace nuotare” che fanno propendere per una classificazione delle frasi (5) e (7) come *mistake*, nel senso che la scrivente non ignora la costruzione delle complete, ma probabilmente ha bisogno di una azione di rinforzo:

- (11) Mi diverto ad ascoltare musica e allo stesso tempo fare jogging o solo risposare fuori. In oltre mi piace nuotare, andare in bicicletta

La partner di IS, AN, emenda il testo modificandolo grazie alla funzione *edit* presente nel sito. Il *feedback* è a correzione diretta, secondo la tassonomia di Sheen (2007), con evidenziazione grafica (Nuzzo, 2013), nel senso che le parti del testo che si discostano dalla norma italiana vengono eliminate e riformulate utilizzando un colore diverso, senza fornire approfondimenti metalinguistici. In tal modo AN cerca di innescare un meccanismo di riflessione autonoma e di indurre la partner ad un’autocorrezione. Alcune espressioni lessicali non perfettamente conformi alla norma italiana vengono inoltre tralasciate, forse perché non disturbano la comunicazione: (8) “informare biscotti”. Comunque il *feedback* fornito appare accurato ed attento alla forma linguistica. Anche IS offre ad AN la stessa tipologia di *feedback*, che si mantiene costante in tutte le schede di lavoro elaborate.

Osservando quanto dichiara nel portfolio, si enuclea però che IS apprezza soprattutto l'opportunità di leggere un vero ed autentico testo italiano (12) e di espandere in tal modo il proprio vocabolario (13), mentre non ci sono riferimenti agli errori/sbagli di tipo formale, che pure erano stati scrupolosamente emendati da AN: l'acquisizione di nuovo lessico (14) costituisce evidentemente per l'apprendente la premessa per il raggiungimento del suo obiettivo a lungo termine, cioè l'ampliamento delle proprie capacità di comunicazione nella lingua obiettivo.

(12) "Il testo scritto dal partner/dalla partner è stato utile?"
 – *Ja! Um einen „echte/richtigen“ Italienischen Text zu lesen*

(13) "Cosa ho imparato di nuovo?"
 – *themenspezifische Vokabeln*

(14) "Cosa vorrei migliorare o fare diversamente nel prossimo incontro skype?"
 – *Vokabeln aus Annies Text nachschlagen (cercare)*

Dal portfolio emerge inoltre che le partecipanti sono riuscite a risolvere autonomamente alcune difficoltà di comprensione insorte durante l'interazione orale scrivendo le espressioni più problematiche nello spazio di *Skype* riservato alla *chat*⁴: sono state quindi capaci di attivare autonomamente strategie di problem solving dimostrando in tal modo la propria "autonomia" cioè la capacità di mettere in atto comportamenti adatti al contesto e funzionali al raggiungimento del proprio scopo.

Il secondo lavoro di IS rivela che l'*input* formale di AN si è comunque trasformato in *intake* in almeno due casi: quello dell'espressione errata del verbo pronominale con clitico incorporato e quella riguardante il transfer da LM (si confronti l'esempio (15) con gli esempi (4) e (9)).

(15) A casa mi piace rilassarmi o andare a trovare mio fratello

Si notano inoltre problemi con l'ordine degli elementi nelle frasi (16) e (17):

(16) [...] solo ho l'università la sera

(17) [...] solo parlo con i miei amici

e permangono difficoltà con le espressioni temporali in (18), (18) e (20)

(18) Ogni secondo martedì, m'incontro con gli altri che sono nella 'Eine-Welt-Gruppe'

(19) Il mercoledì vado all'università non prima di 10 ore

(20) Il venerdì i miei corsi finiscono per le 12

⁴ "Quali difficoltà ho avuto durante l'incontro Skype?" -am Anfang ein Paar Tonprobleme (Lag an der Technik) - teilweise einzelne sprachliche Probleme, mit Händen und Füßen und Umschreibungen hat alles trotzdem geklappt und wir haben uns gut verstanden "All'inizio problemi con l'audio (problemi tecnici) - a volte abbiamo incontrato difficoltà linguistiche, ma con le unghie e con i denti e alcune trascrizioni tutto ha funzionato e ci siamo comprese bene".

Il *feedback* correttivo di AN segue lo stesso approccio tipologico del precedente, anche se presenta una maggiore scrupolosità, in quanto, piuttosto che sostituire le espressioni inaccurate o errate, le evidenzia in rosso e suggerisce la riparazione in azzurro (21).

(21) IS: “Il lunedì comincia **rilassato** (**tranquillamente**) e lentamente **perché solo ho l’università la sera** (**perché ho università solo la sera**). Prima di andare all’università ho tempo per studiare. La mattina mi piace fare colazione ed allo stesso tempo leggere il giornale.

Il martedì passo molto tempo all’università. Frattanto ho anche **il tempo libero** (**del tempo libero**), **in cui** (**durante il quale**) studio o solo parlo con i miei amici. Ogni secondo martedì m’incontro con gli altri che sono nella ‘Eine-Welt-Gruppe’.

Il mercoledì vado all’università **non prima di 10 ore** (**non prima delle 10**). Per questo il bus non è così zeppo di studenti, perché **la maggioranza venga** (**la maggior parte viene**) all’università per le otto. In seguito a ciò vado a casa. La sera ho un altro corso.

Il giovedì è uguale. Ho i corsi la mattina, poi tempo **libro** (**libero**) e un corso la sera. A casa mi piace rilassarmi o andare a trovare mio fratello.

Il venerdì i miei corsi finiscono per le 12. **Provo di studio** (**provo a studiare**) molto il venerdì, cosicché non devo fare molto **nella fine della settimana** (**nel fine settimana**).

Nella fine della settimana faccio molto con la mia famiglia (**nel fine settimana passo molto tempo con la mia famiglia**). Mi piace **andare di fuori** (**uscire**) o fare qualcosa con i miei amici. Spesso mio fratello e la sua amica vengono **per mangiare** (**vengono a mangiare/pranzare/cenare**) con noi.

Le nostre settimane sono molto differenti, perché Annie fa molto sport e abbiamo attività diverse. Ma tutti **i due** (**tutte e due/entrambe**) passiamo molto tempo all’università”.

Nel portfolio IS segnala di avere problemi riguardo la capacità di comprendere ed esprimere fluentemente le parti del giorno ed in generale le espressioni numeriche e temporali, come riportato nell’esempio (22): in questo modo, a nostro avviso, il portfolio è diventato uno strumento attraverso il quale la discente sollecita un approfondimento specifico.

(22) “Quali difficoltà ho avuto durante l’incontro Skype?”

– Zahlen/Uhrzeiten auf Anhieb auszudrücken, Zeiträume erklären
‘essere fluente nelle espressioni numeriche/temporali, esprimere gli intervalli di tempo’

“L’aiuto datomi dal partner/dalla partner è stato come avrei voluto?”

– Mir mit den Zahlen geholfen

‘Mi ha aiutato con i numeri’

Anche nelle produzioni successive le due partner hanno lavorato molto sul lessico rivedendo ed ampliando le rispettive mappe concettuali, hanno fornito un *feedback* esplicito con evidenziazioni grafiche, hanno continuato a scegliere di “editare” i testi non facendo mai ricorso, come invece hanno fatto altre coppie, alla sezione *comments*. La tipologia di *feedback* scelta ha reso palese, a nostro avviso, uno stile di apprendimento che predilige interventi correttivi espliciti, ma privi di commenti metalinguistici, che lascino cioè la possibilità di ricorrere a strumenti autonomi di riflessione e approfondimento. L’attenzione delle apprendenti è sembrata inoltre maggiormente focalizzata sugli aspetti comunicativi, piuttosto che su quelli formali della lingua, che

hanno costituito un problema solo nei casi in cui impedivano la trasmissione dei contenuti. L'ampliamento del lessico era dunque l'obiettivo formativo intermedio della coppia, lo strumento grazie al quale le apprendenti ritenevano di poter acquisire una maggiore possibilità di comunicazione. Queste indicazioni sono state utili nella programmazione di attività didattiche frontali orientate specificamente a questa e alle altre coppie che hanno fornito questa tipologia di feedback. Esse da una parte sono state improntate ad aiutare il processo di ampliamento del vocabolario proponendo letture e approfondimenti morfologici, ad esempio, sulle modalità di formazione delle parole in tedesco, mentre da un'altra sono state incentrate su attività di riflessione induttiva per stimolare l'autocorrezione, come quelle nelle quali gli errori vengono classificati per tipologia usando simboli convenzionali e lasciando all'apprendente il compito di riflettere ed autocorreggersi (Mezzadri, 2002). L'osservazione diacronica del lavoro della coppia ha dimostrato un miglioramento nella produzione scritta, sia per quanto riguarda la ricchezza del repertorio lessicale che la forma linguistica.

Le modalità con le quali è stato fornito il *feedback* correttivo, come precedentemente accennato, non sono state uguali per tutte le coppie. Nella maggioranza dei casi i partecipanti hanno scelto di editare i testi intervenendo direttamente su quanto scritto dai partner, ma alcuni di loro hanno invece fatto uso della sezione *comments*. Dall'analisi della tipologia dei *feedback* è risultato evidente che partecipanti con stili di apprendimento diversi e con diversi obiettivi hanno dato *feedback* di diversa tipologia. Allo scopo di evidenziare ciò, si propone il lavoro di una seconda coppia di apprendenti, formata da MA germanofona e GI italofona. Già dal questionario si evidenziava che MA individuava il raggiungimento di una competenza approfondita, anche metalinguistica, della lingua obiettivo, come lo strumento indispensabile per ampliare la propria capacità di esprimersi, dichiarava infatti di finalizzare l'attività eTandem all'implementazione di un rapporto più personale con la lingua in quanto desiderava sentirla più vicina a sé.

Le due apprendenti non hanno "editato" le produzioni scritte della partner, ma le hanno "commentate" nell'apposita sezione (Fig. 3).

Il loro *feedback* correttivo, seguendo anche in questo caso la tassonomia di Sheen, è stato a correzione diretta metalinguistica con evidenziazione grafica, poiché le due discenti hanno commentato quanto scritto dalla partner dal punto di vista della forma linguistica, cercando di esplicitare le "regole linguistiche" della propria lingua e/o hanno proposto soluzioni alternative ad espressioni inaccurate o erranee. Nei *feedback* si nota inoltre un'alternanza d'uso fra lingua madre e lingua obiettivo; la lingua madre - per MA il tedesco - è stata usata nelle parti discorsive, mentre la lingua obiettivo in quelle dedicate alla riflessione metalinguistica: in questo modo le apprendenti hanno cercato di rendere i propri commenti quanto più chiari possibile per la partner. Differentemente dall'esempio precedente, quindi, la correttezza formale assumeva in questo caso un ruolo di primo piano nell'evento comunicativo: questo ha suggerito ai docenti di diversificare il proprio comportamento in sede di lezione frontale rispetto a quello tenuto nei confronti della coppia AN/IS. In questo caso ed in tutti i casi analoghi, è stata infatti stimolata la riflessione metalinguistica, segnalando ed ap-

profondendo esplicitamente gli errori, si sono invitate i/le discenti a produrre tabelle che schematizzassero alcuni argomenti di morfosintassi, da presentare poi all'intera classe, si sono infine proposti esercizi di riformulazione di testi soprattutto per quanto riguarda gli aspetti morfosintattici.

Figura 3 - *Il feedback commentato della coppia GI/MA*

Maïke Prause said

at 7:04 pm on Nov 16, 2014

[Reply](#) [Delete](#)

Ciao Givonna,

lavoro molto al presente, e difficile per me di andare al tandem-corso martedì sera.

E possibile per te di fare la tandem insieme per un altro tempo?

Mi compiacero di la tua risposta. Maïke

Maïke Prause said

at 7:38 pm on Nov 16, 2014

[Reply](#) [Delete](#)

Hallo Giovanna! Gerne korrigiere ich auch deinen Text:

Neaples si chiamo in tedesco: Neapel

si dice: Mir bereiten Kunstfilme, Literatur, Kunst und Musik große Freude //oder// mir gefallen Kunstfilme, Literatur, Kunst und Musik.

non si dice: Filme "machen" Spaß. tu potresti dire: die Filme von Woody Allen gefallen mir sehr gut// bereiten mir große Freude.

Si tu dici qualcosa che TU ami o che TU piaci, tu potresti dire: MIR macht es sehr viel Spaß zu kochen, MIR macht es sehr viel Spaß die Filme von Woody Allen zu schauen etc...

Si dice: einEN Film, DEN ich sehr MAG (presente!) (tu ami questo film ancora, non?) ancora meglio tu dici: EIN Film, DER mir sehr gut gefällt, ist.....

Noi facciamo una grande differenza fra "mögen" e "lieben". (Ich habe drei Bücher sehr geliebt). Ich mag drei Bücher sehr: / Mir gefielen 3 Bücher (,die ich gelesen habe) sehr gut.

(tu votresti dire che tu le ami ancora o in passato?)

Ich habe keine Facebook Seite aber ich habe seit einer Woche einen Skype Account. Ich muss jetzt nur mein "alias" erinnern xD (che combinazione, ho dimenticato mon password per il Tandem fa una settimana :-)

Ich koche gern, besonders Kuchen. (si tu facci Kuchen, backst du. (Infinitiv: Backen) Uno cucinare la cena ma uno cuocere la torta.

Bcken kann man: Kuchen, Brot, Brötchen, Plätzchen (für Weihnachten), Kekse, Torten....

Scrivo una prenetazione corta in Italiano anche io a la mia pagina.

Si tu votresti, mi scrivi una e-mail si tu ha coretto il mio text. Non guardo sempre per il tandem:

5. Conclusioni

Il Wiki-eTandem ha perseguito obiettivi formativi specifici sia per gli apprendenti che per i docenti. Per quanto riguarda i primi, la principale finalità delle attività proposte era quella di stimolare la produzione linguistica. Il Wiki-eTandem è riuscito, a nostro avviso, nel suo intento, in quanto si è creato un contesto sociale informale che ha favorito, grazie alla interazione fra pari, l'*output* e conseguentemente ha stimolato l'acquisizione della capacità di comunicare in lingua straniera. La fluidità e la correttezza linguistica in fase di produzione sia orale che scritta è infatti migliorata nei partecipanti alle attività eTandem in misura sensibilmente maggiore rispetto ai colleghi che non vi hanno preso parte, come è risultato evidente sia dall'analisi delle produzioni scritte, che in sede di colloquio orale. Anche i partecipanti hanno valutato positivamente l'esperienza, come si evidenzia nel bilancio finale di AN (23):

(23) AN: "Ho trovato il Tandem un'esperienza totalmente positiva [...] La capacità di comprendere il testo è di molto migliorata infatti mi risulta molto più semplice leggere e capire soprattutto le frasi secondarie con "dass" che prima del Tandem utilizzavo sporadicamente ed in maniera incerta. Sicuramente il mio più grande problema durante gli incontri Skype è stato capire cosa la mia partner dicesse in quanto parlava molto velocemente e qualche volta abbiamo dovuto utilizzare la tastiera durante la videochiamata. Prima degli incontri ho sempre preparato una lista di vocaboli riguardanti l'argomento scelto e di verbi che sarebbero potuti essere utili. Abbiamo letto le nostre composizioni, ognuno leggeva il testo dell'altra in maniera tale da migliorare la pronuncia nella nostra lingua target. Mi è stato utile partecipare a questo lavoro poiché ho imparato moltissimi nuovi vocaboli e soprattutto perché ho conosciuto le tradizioni di Isabel. Weihnachten è stato infatti l'argomento che ho preferito tra [...]. Non ci siamo scritte solo sulla pagine wiki e si Skype, abbiamo anche chattato su whatsapp."

I discenti hanno inoltre sviluppato la capacità di lavorare autonomamente (Holec, 1979; Little, 1991) in quanto hanno risolto situazioni di difficoltà e hanno gestito bene i tempi andando anche oltre quanto era stato richiesto: 4 schede di lavoro in due mesi. Dall'analisi dei portfolio emerge tuttavia che occorre stimolare ulteriormente gli apprendenti verso un'analisi critica metalinguistica dei *feedback* ricevuti, al fine di focalizzare le proprie lacune e riflettere sugli strumenti per colmarle: sono mancate infatti riflessioni critiche in questo senso. Per raggiungere questo obiettivo ipotizziamo in una prossima esperienza in primo luogo di presentare il PEL illustrandone le finalità ed esplicitando in tal modo lo scopo con il quale è stato pensato il portfolio proposto, ed in secondo luogo di trattare in sede di lezione frontale il tema "imparare ad imparare". Esso rappresenta fra l'altro una delle otto Competenze Chiave per l'Apprendimento Permanente individuate nel 2006 dal Parlamento Europeo e dal Consiglio come "necessarie per la realizzazione e lo sviluppo personali, la cittadinanza attiva, l'inclusione sociale e l'occupazione"⁵.

⁵ Una sintesi del testo è consultabile sul sito EUR-LEX <http://eur-lex.europa.eu/legal-content/IT/TXT/?uri=uriserv:c11090> (data ultimo accesso 30 luglio 2015).

Per quanto riguarda invece i docenti, lo scopo del Wiki-eTandem era quello di fornire indicazioni utili ad un'ottimizzazione della didattica in presenza, in un'ottica *learner centered*. Come si è cercato di evidenziare attraverso gli esempi portati, l'analisi delle produzioni degli apprendenti, della loro interlingua, dei *feedback* correttivi e dei portfolio è stata senz'altro di aiuto a mettere a fuoco gli elementi morfosintattici più problematici e ha fornito suggerimenti utili per adeguare gli interventi didattici allo stile degli apprendenti. Purtroppo, però, l'organizzazione dei corsi a Napoli e a Bochum, che prevede un solo semestre per gli insegnamenti linguistici, ha impedito di organizzare sistematicamente delle unità didattiche di approfondimento da proporre in sede di lezione frontale, che tenessero conto degli stili di apprendimento e delle esigenze formative dei singoli apprendenti, e ci ha costretti spesso a interventi personalizzati al di fuori dell'orario di lezione.

A nostro avviso, per cogliere appieno le potenzialità del Wiki-eTandem occorrerebbero due semestri: nel primo si potrebbero sviluppare le attività di *output* in collaborazione, mentre nel secondo ci si potrebbe dedicare ad analizzare approfonditamente le produzioni, i *feedback* ricevuti ed i portfolio insieme agli apprendenti, anche in sede di lezione frontale. Si potrebbe in tal modo lavorare più proficuamente per attivare processi di autovalutazione finalizzati a facilitare l'acquisizione della competenza metalinguistica e sviluppare maggiormente meccanismi di autonomia negli apprendenti.

Bibliografia

- BEDINI S. - BALÒ R. (2008), Gli usi didattici del wiki per l'apprendimento collaborativo e cooperativo nell'Italiano L2/LS: l'esempio di ADGWIKI, in *Bollettino Itals* (25), disponibile in: <http://www.ital5.it/editoriale/bollettino-ital5-giugno-2008> (data ultimo accesso 30 luglio 2015).
- BOSISIO C. (2012), *Interlingua e profilo d'apprendente. Uno sguardo diacronico fra linguistica acquisizionale e glottodidattica*, EDUCatt, Milano.
- BRAMMERTS H. (1996), Language learning in tandem using the Internet, in WARSCHAUER M. (ed.), *Telecollaboration in foreign language learning. Proceedings of the Hawai'i Symposium*, University of Hawai'i Press, Honolulu: 121-130.
- CZIKO G.A. (2004), Electronic Tandem Language Learning (eTandem): A Third Approach to Second Language Learning for the 21st Century, in *Calico Journal* 22(1): 25-39.
- CHINI M. (2005), *Che cos'è la linguistica acquisizionale*, Carocci, Roma.
- CHINI M. (2011), Qualche riflessione sulla didattica di L2 ispirata alla recente ricerca acquisizionale, in *Italiano LingueDue* 3(2): 1-22.
- DELLA PUTTA P. - VISIGALLI M. (2010), Come aiutare l'acquisizione della morfosintassi del sintagma nominale italiano: confronto fra tre percorsi glottodidattici rivolti a studenti anglofoni, in *Italiano LingueDue* (2): 23-52
- GASS S.M. (2005), Input and Interaction, in DOUGHTY C.J. - LONG M.H. (eds), *The Handbook of Second Language Acquisition*, Blackwell, Malden USA: 224-255.

- HOLEC H. (1979), *Autonomie et apprentissage des langues étrangères*, Editions du Conseil de l'Europe, Strasbourg.
- JUDD T. - KENNEDY G. - CROPPER S. (2010), Using wikis for collaborative learning: Assessing collaboration through contribution, in *Australasian Journal of Educational Technology* 26(3): 341-354.
- LEONE P. (2009), Comunicazione mediata dal computer e apprendimento linguistico: gli incontri Teletandem, in *Studi di Glottodidattica* 1: 90-106, disponibile in: <http://ojs.cime-doc.uniba.it/index.php/glottodidattica/article/viewFile/184/55> (data ultimo accesso 30 luglio 2015).
- LIGHTBOWN P.M. - SPADA N. (2006), *How languages are learned*, Oxford University Press, New York.
- LITTLE D. (1991), *Learner autonomy 1: Definitions, issues and problems*, Authentik, Dublin.
- LITTLE D. (2002), Learner autonomy and autonomous language learning: some theoretical perspectives and their practical implications, in EVANGELISTI P. - ARGONDIZZO C. (a cura di), *L'apprendimento autonomo delle lingue straniere*, Rubettino, Catanzaro: 17-31.
- LITTLE D. (2012), *L'integrazione linguistica degli immigrati adulti ed il Portfolio europeo delle lingue: Introduzione*, Consiglio d'Europa, Strasburgo, http://www.coe.int/t/dg4/linguistic/liam/Source/Tools/LIAM-ELP_Introduction_IT.pdf (data ultimo accesso 30 luglio 2015).
- LONG M.H. (1988), Instructed interlanguage development, in BEEBE L. (ed.), *Issues in second language acquisition: Multiple perspectives*, Newbury House, Rowley, MA: 115-141.
- LONG M. (1996), The Role of the Linguistic Environment in Second Language Acquisition, in RITCHIE W.C. - BATHIA, T.J. (eds), *Handbook of Second Language Acquisition*, New York Press, New York: 413-468.
- MATTHIAE C. (2012), La riflessione metalinguistica nei manuali d'italiano L2: case study, in *Italiano LinguaDue* 1: 343-351.
- MEZZADRI M. (2002), La correzione degli errori in *In.it*, anno 3. (1 2): 4-9.
- KRASHEN S.D. (1981), *Principles and Practice in Second Language Acquisition*, English Language Teaching series, Prentice-Hall International, London.
- NUZZO E. (2013), Il feedback correttivo fra pari nell'apprendimento dell'italiano in rete: osservazioni a partire da un corpo di interazioni asincrone, in *Italiano LinguaDue* 1: 15-28.
- SHEEN Y. (2007), The Effect of Focused Written Corrective Feedback and Language Aptitude on ESL Learners' Acquisition of Articles, in *TESOL Quarterly* 41(2): 255-283.
- STAHL G. - KOSCHMANN T. - SUTHERS D. (2006), Computer-supported collaborative learning: An historical perspective, in SAWYER K.R. (ed.), *Cambridge handbook of the learning sciences*, Cambridge University Press, Cambridge, UK: 409-426, http://gerrystahl.net/csc/CSCL_English.pdf (data ultimo accesso 30 luglio 2015).
- SWAIN M. (1985), Communicative Competence: Some roles of Comprehensible Input and Comprehensible Output in its Development, in GASS S. - MADDEN C.G. (eds.), *Input in Second Language Acquisition*, Newbury House, Rowley, MA: 235-253.
- TELLES J.A. - VASSALLO M.L. (2006), Foreign language learning in- tandem: Theoretical principles and research perspectives, in *The ESPecialist* 25(1): 1-37.

TELLES J.A. - VASSALLO M.L. (2006), Foreign language learning in- tandem: Telerandem as an alternative proposal in CALLT, in *The ESspecialist* 27(2): 189-212.

VARONIS E.M. - GASS S. (1985b), Non-native/ non-native conversations: A model for negotiation of meaning, in *Applied Linguistics* 6(1): 71-90.

Riformulazioni attraverso Twitter da parte di studenti italiano L1 e LS: proposta di una tassonomia testuale³

The current paper describes the creation of a corpus of tweets in Italian written by Italian and Slovakian students. The tweets were elicited by asking the students to summarise in tweets (target texts) two source texts in Italian, i.e. a poem and a newspaper article. Qualitative analysis of the target texts highlighted three textual parameters whose combination resulted in a taxonomy of eight textual classes. The analysis of the distribution of the textual classes showed that it varied according to genre of the source texts and student group, which confirms the validity of the proposed taxonomy.

1. *Introduzione*

Il presente contributo si inserisce nella macro-area di studi sul processo di comunicazione in rete mediato dal computer (CMC), che dai pioneristici studi sugli usi linguistici nelle e-mail, nei forum di discussione e nelle chat room di circa venti anni fa, è oggi estesa al Web 2.0, nozione che include i *social network* come Twitter. La dinamicità del campo di studi sulla CMC è evidente a livello di definizione della materia. Negli anni '90 del secolo scorso, la CMC cominciò a suscitare attenzione da parte della comunità scientifica (Herring, 1996: 1-2). Un punto di svolta negli studi è certamente segnato dall'attenzione verso l'uso delle tecnologie della rete come mezzo di scambio e di condivisione attiva e partecipata di informazioni: l'uso del web come piattaforma sociale comporta oggi, nell'ambito della CMC, nuovi contenuti da analizzare, nuovi contesti, nuovi modelli d'uso e, conseguentemente, nuove prospettive di ricerca su nuovi domini come i *social network*, e su lingue altre rispetto all'inglese (Herring *et al.*, 2013: 13).

Il presente contributo si colloca in tale panorama di ricerche, e si propone di presentare e validare una tassonomia delle riformulazioni di testi in *tweet*. Tramite un approccio *data-driven* si mira inoltre a valutare se e in che misura la tipologia di tali riformulazioni possa correlare con due macro-variabili: genere del testo riformulato e contesti di italiano L1 e LS. Rispetto a tali macro-variabili, il lavoro si è basato sulla raccolta di un corpus di cento *tweet* durante un esperimento di tipo controllato.

¹ Università Matej Bel di Banská Bystrica.

² Libera Università di Bolzano.

³ Il lavoro è il frutto di una stretta collaborazione tra i due autori. I singoli paragrafi sono stati pensati e discussi congiuntamente; tuttavia, per i soli fini accademici, Claudio Nobili è responsabile della stesura dei paragrafi 1 e 2, Chiara Meluzzi di quella dei paragrafi 3 e 4.

Nel contributo si illustreranno le caratteristiche di Twitter e le potenzialità di (ri)scrittura breve offerte da tale mezzo (§ 1.1). Successivamente sarà definito il principio di riformulabilità di un testo, applicato alle nuove tecnologie (§ 1.2). Sarà inoltre presentata la metodologia per la raccolta e per l'analisi dei dati (§ 2 e § 3). Infine, vengono presentate nuove prospettive di ricerca sull'utilità e spendibilità della tassonomia elaborata, in particolare in contesti non controllati, e in relazione ad altri processi di scrittura breve non legati alle tecnologie della rete (§ 4).

1.1 Twitter nell'ambito delle scritture brevi oggi

La brevità è tratto distintivo di Twitter. Esso, infatti, consente di condividere con interlocutori *following* (coloro che un utente, un *tweeter*, sceglie di 'seguire') e *followers* (coloro che 'seguono un utente') messaggi (*tweet*) di una lunghezza massima di 140 caratteri (un *tweet* di 140 caratteri si chiama *twoosh*), inclusa un'eventuale foto. A livello linguistico, il *social network* mostra indubbe similitudini con altre scritture digitali (es. SMS, chat), caratterizzate, per esempio, da abbreviazioni, errori di battitura legati alla velocità di produzione e all'informalità del mezzo, sigle, acronimi, segni non verbali (Pistolesi, 2014: 364-371). Tra questi ultimi la chiocciola (@) e l'*hashtag* (#), peculiari di Twitter, si stanno progressivamente estendendo nell'uso ad altri *social network* (es. Facebook).

Il simbolo della chiocciola (@) (Arcangeli, 2015), seguito dal nome di un utente senza spazio bianco separatore, serve per creare un nome utente, e per menzionare un altro utente durante la produzione di un *tweet*. Il simbolo cancelletto (#), seguito da una o più unità lessicali continue, ovvero senza spazio separatore al suo interno, rende il nesso un cosiddetto *hashtag*, la cui funzione linguistica è stata recentemente indagata in vari studi (es. Zappavigna, 2015). A livello definitorio, l'*hashtag* può essere inteso come:

elemento capace di generare link, ciò che ne determina lo statuto di "aggregatore" del mezzo, indispensabile elemento per la costituzione del dibattito interno e della composizione della relativa comunità social. Proprio tale specifico attributo si pone alla base delle componenti formali dell'*hashtag*, da cui originano le difficoltà di decifrazione (Chiusaroli, 2014: 118).

Più in generale, problemi di comprensione legati alla CMC sono stati già evidenziati da De Mauro (2002: 133), il quale ha rilevato che "rispetto a ogni altra tecnologia, a tutt'oggi è ancora il rapporto con la pagina scritta quello che ci consente le condizioni migliori per quella elaborazione interiore, silenziosa, per quei controlli, per quell'andare su e giù per i dati che le videate non riescono a eguagliare". Da ciò può derivare un'eventuale compromissione del processo di comprensione, che in Twitter può assumere un valore maggiore dipendente dalle caratteristiche di scrittura distintive del mezzo appena illustrate (Chiusaroli, 2014: 117).

Le problematicità dei processi di scrittura e comprensione attraverso Twitter potrebbero essere alla base della scarsa frequenza d'uso del mezzo. Secondo l'edizione 2014 dell'indagine *European Digital Landscape*, pubblicata da *We are social*, che ha coinvolto 40 paesi europei, in Italia il 41% degli utenti attivi sui *social network* (a

loro volta il 42% della popolazione) possiede e usa un account Twitter. L'indagine non riporta purtroppo un dato analogo per la Slovacchia, secondo termine di paragone nel presente studio. Dalla relazione di Kollar - Czwickowics (2013: 35) emerge soltanto che: "Twitter [...] does not appear to have gained great popularity. There are no prominent subscribers and it is rarely used for marketing".

Lo scarso uso di Twitter non esclude, bensì incoraggia un tentativo di esplorazione sistematica delle potenzialità di scrittura creativa attraverso questo mezzo, nell'ottica di quella che Chiusaroli in Pierucci (2015) ha definito "educazione digitale".

In riferimento al processo di scrittura attraverso Twitter in italiano, Trifone (2014) ha delineato alcuni aspetti dell'"italiano twitterario" o 'twitterese', in senso funzionale, ovvero "l'idoneità della lingua a produrre un testo letterario anche senza oltrepassare la soglia dei 140 caratteri", una 'micronarrativa twitteraria'. Senza quindi porre attenzione verso la correttezza grammaticale dei *tweet* in relazione all'italiano standard, ciò che Trifone (2014) evidenzia è la possibilità del mezzo di produrre testi brevi, anche come riformulazione di testi altri. Ne sono alcuni esempi significativi la riscrittura del *Decameron* di Boccaccio in cento *tweet/twoosh* per cento giorni, promossa dalla Società Dante Alighieri all'interno di un progetto curato da Massimo Arcangeli tra agosto e novembre 2013, e la riscrittura, tra giugno e luglio del 2014, in forma di *twoosh*, di un brano delle *Operette morali* o di un pensiero dello *Zibaldone* di Leopardi lanciata da *M'è dolce il cinguettar... #Leopardeide*, progetto ideato sempre da Arcangeli in collaborazione con il blog *Scritture brevi*⁴.

La possibilità di riformulazione in forma breve di testi altri attraverso Twitter è alla base del primo stadio della presente ricerca, che si propone di indagare il prodotto (*tweet/twoosh*) del processo di riformulazione testuale da parte di studenti di facoltà umanistiche d'italiano L1 e LS (L1 slovacca). Prima di entrare nel merito dello studio qui proposto, pare necessaria una premessa teorica sul principio di riformulabilità di un testo, e sulla sua possibile applicazione alle nuove tecnologie.

1.2 Riformulazione: dal segno al testo

La nozione teorica di sostegno al presente lavoro è il principio di riformulabilità di un segno definito nella sua interezza e complessità in un orizzonte semiotico:

Dato un qualunque segno, dai principi costitutivi di ogni semiotica ricaviamo che esso è sempre passibile di sostituzione con un segno diverso, se c'è e se adempie all'ufficio di veicolare i sensi. *In questo senso* possiamo dire che non vi è limite alla riformulazione di qualunque segno. (De Mauro, 2008: 107).

Il principio di riformulabilità può essere esteso dal segno al testo, il quale "si ha quando differenti segni di un calcolo o di una lingua sono connessi tra loro per circoscrivere grazie ai loro significati un significato nuovo, un significato d'insieme, che li trascende e li accomuna" (*ibidem*: 109).

⁴ <http://www.scritturebrevi.it>.

Alla base di qualunque processo di riformulazione testuale, non mera riduzione dei contenuti essenziali del testo di partenza (Serianni, 2003), vi è un primo passaggio cognitivo di comprensione e selezione di tali contenuti, come sostengono Fornasiero - Tamiozzo Goldmann (2013: 90). La fase di comprensione è dunque alla base di qualsiasi processo di riformulazione testuale.

Se si fa riferimento al solo codice linguistico e a un solo codice linguistico, nonché ai soli testi scritti, possiamo parlare di riformulazione testuale scritta intralinguistica come “qualsiasi processo che trasformi un prototesto in un metatesto” nella stessa lingua (Osimo, 2010: 10; Popovič, 2006)⁵. Applicando produttivamente questa nozione alle nuove tecnologie, è possibile approdare alla nozione di riformulazione testuale scritta intralinguistica digitale, nel senso di processo di riscrittura breve, attraverso il mezzo Twitter, di un prototesto (primo testo o testo originario) in un metatesto (testo ulteriore) nella stessa lingua.

In questo studio, i prototesti, occorrenza comunicativa che soddisfa le condizioni di testualità individuate da De Beaugrande - Dressler (1994: 18-27), corrispondono a due testi in italiano variabili per genere, letterario e giornalistico.

2. Metodologia per la costruzione di un corpus sperimentale in chiave comparativa

Il protocollo di ricerca per il presente studio si basa su tre fasi principali: selezione dei prototesti, creazione di un account Twitter, selezione del campione di riferimento.

In primo luogo sono stati selezionati i prototesti in virtù della variabile del genere testuale e della semplicità complessiva e lessicale. Il primo testo, la poesia di Alda Merini *Io non ho bisogno di denaro*, di 72 parole, incluso titolo e nome dell'autrice, è costruito sulla dicotomia materialismo-spiritualismo; il secondo, un articolo giornalistico sul web tratto dal sito della rivista Focus⁶, *Calcio...che emozione!*, di 349 parole, riporta uno studio inglese secondo il quale gli uomini esprimerebbero le loro emozioni durante o dopo una partita di calcio visto o giocato. I prototesti sono stati selezionati perché complessivamente e lessicalmente semplici, soggetti, pertanto, a univocità d'interpretazione e decodifica. Tale assunto è stato verificato successivamente tramite un'apposita scheda di valutazione somministrata al campione della ricerca (§ 2.1).

Una volta selezionati i prototesti, è stato creato un account Twitter unico e specifico per la ricerca, privo di *following* e *followers*, e intenzionalmente oscurato in questa prima fase. In tal modo, durante la successiva fase di raccolta dati, gli studenti avrebbero dovuto produrre i propri *tweet* come riassunti dei prototesti (due *tweet* da parte di ciascun studente) recandosi a turno a un computer presente in aula, sen-

⁵ In questa sede, utilizzeremo le etichette di “prototesto” e “metatesto” in chiave puramente operativa, senza alcuna implicazione teorica o metodologica.

⁶ <http://www.focus.it/cultura/calcio-che-emozione>.

za la possibilità, una volta completato il *tweet* e lasciata la postazione, di apportare ulteriori modifiche al testo, non essendo a conoscenza delle credenziali di accesso all'account, note soltanto ai ricercatori. La ricerca pertanto è stata concepita come esperimento controllato, novità metodologica rispetto alle precedenti esperienze di riscrittura di testi, letterari in particolare, attraverso Twitter, realizzate in contesti autentici (§ 1.1).

Infine, è stato individuato un campione di riferimento costituito da studenti di madrelingua italiana di tre atenei (Università degli Studi del Piemonte Orientale, Università degli Studi di Pavia, Università degli Studi di Milano-Bicocca), e di studenti di italiano L2 dell'Università slovacca Matej Bel di Banská Bystrica. Più precisamente, il campione di riferimento è articolato in 25 studenti di italiano L1, dei quali 19 iscritti a un Corso di Laurea triennale, e 6 a uno di Laurea magistrale, e 25 studenti di italiano L2, dei quali 12 iscritti a un Corso di Laurea triennale, e 13 a uno di Laurea magistrale. In questa prima fase della ricerca, non si è tenuto in considerazione delle variabili 'livello di laurea' e, nel contesto L2, 'anni di studio dell'italiano', poiché l'esiguità del campione compromette fortemente la rappresentatività dei diversi sottocampioni. Dato tuttavia l'interesse di queste variabili rispetto alle macro-variabili linguistico-testuali (gruppo degli studenti e tipo di prototesto), se ne terrà certamente conto nelle fasi successive della ricerca in previsione di un già progettato allargamento del campione.

I dati sono stati raccolti a margine di lezioni universitarie che vedevano già presenti in classe gli studenti. Nel contesto di italiano L1 si trattava di lezioni dei corsi di sociolinguistica e di etnolinguistica dei rispettivi atenei, mentre nel contesto d'italiano L2 l'elicitazione è avvenuta a margine delle lezioni di sintassi italiana, esercizi di comunicazione in lingua italiana, e aspetti dell'italiano contemporaneo. Questo giustifica la naturale propensione degli studenti slovacchi a produrre testi in italiano e non in un'altra lingua, inclusa la propria L1. Tuttavia, è bene specificare che questo esperimento non ha rappresentato alcuna attività didattica, e di questo gli studenti, sia italiano L1 sia L2, erano consapevoli, accettando quindi spontaneamente di partecipare alla raccolta dati.

La raccolta dati si è svolta in cinque fasi consecutive:

1. somministrazione e compilazione da parte degli studenti di un questionario, al fine di elicitare le principali informazioni anagrafiche e, tra l'altro, indagare l'eventuale frequenza d'uso di Twitter;
2. somministrazione del prototesto letterario ed esplicitazione da parte del ricercatore delle istruzioni relative all'attività successiva ("Avete 5 minuti di tempo massimo per leggere il testo per poi riassumerlo in un *tweet*");
3. lettura interiore;
4. scrittura del *tweet*, articolata a sua volta in due sotto-fasi. In primo luogo, il ricercatore, dopo aver effettuato l'accesso all'account unico per la ricerca, ha premesso a ogni *tweet* due *hashtag*: #itinbreve, dal momento che, come già è stato evidenziato, oggetto della ricerca è un processo di scrittura breve in italiano, e un secondo *hashtag* complesso composto dal prefisso ITA o SLO, a seconda dell'i-

- italiano o dello slovacco come lingua madre, da un numero a due cifre identificativo di ciascun studente, e dalle maiuscole L o G in base al genere, letterario o giornalistico, del prototesto (es. #ITA01L). A questo punto, gli studenti hanno completato il *tweet*, con la possibilità di visualizzare il prototesto, ma senza possibilità di consultazione reciproca, recandosi uno alla volta al computer in classe;
5. ripetizione dei punti precedenti per il prototesto giornalistico;
 6. compilazione, da parte degli studenti, di una scheda di valutazione e comprensione dei prototesti.

Per quanto detto rispetto al legame tra comprensione e riformulazione testuale (§ 1.2), la necessità di sottoporre agli studenti una scheda di valutazione e comprensione dei prototesti è stata dettata da due ragioni fondamentali: verificare l'ipotesi iniziale di semplicità generale e lessicale dei prototesti; verificare l'effettiva comprensione del contenuto dei prototesti, al fine di escludere che la particolarità di alcuni tipi di metatesti fosse imputabile a una mancata comprensione dei prototesti.

La scheda di valutazione è strutturata in quattro domande per ciascun prototesto. Le prime due domande riguardano il grado di difficoltà del prototesto; sono articolate su cinque valori, da "molto facile" a "molto difficile", e richiedono inoltre l'esplicitazione di al massimo tre parole o espressioni non comprese. Le altre due domande, a risposta multipla, sono incentrate sulle principali unità informative dei prototesti, prevedendo pertanto una sola risposta corretta delle tre fornite.

Allo stato attuale, il corpus consta di cento *tweet/twoosh* in italiano raccolti secondo la metodologia fin qui descritta, ed equamente divisi in base alle due macrovariabili del genere del prototesto (50 *tweet* per il testo letterario e 50 per il testo giornalistico) e del gruppo linguistico degli studenti (50 *tweet* in italiano L1 e 50 in italiano LS).

2.1 Osservazioni preliminari

Dal questionario emerge che la maggioranza del campione (e la quasi totalità degli studenti slovacchi) non possiede un account Twitter, e se posseduto è usato raramente, dati in linea con le ricerche sociologiche precedentemente menzionate (§ 1.1).

I prototesti sono stati valutati facili, e ne sono state comprese tutte le parole. Tuttavia, un solo studente italiano L1 ha indicato "abitino" per il prototesto letterario ("di sogni che abitino gli alberi", v. 7), e "terapeutico" (il calcio come "qualcosa di quasi terapeutico") per quello giornalistico come parole non comprese, mentre circa la metà degli studenti di italiano LS ha dichiarato di non aver compreso "mormorino" ("di stelle che mormorino agli orecchi degli amanti", vv. 9-10) per il prototesto letterario, e "infervorati" ("tanti infervorati discorsi su questa o quella squadra") per quello giornalistico. È interessante osservare che le parole non comprese dal prototesto giornalistico occorrono nello stesso paragrafo iniziale.

Relativamente alla generale comprensione dei prototesti, in contesto italiano L1 vi è un solo caso di risposta errata relativa al prototesto letterario, mentre in contesto

italiano LS vi sono tre risposte errate sempre relative al prototesto letterario, due delle quali, come nel caso italiano L1, attribuibili forse a una non attenta lettura della domanda. La terza e ultima risposta errata è dovuta a un mancato riconoscimento del contenuto generale del prototesto letterario, nonostante una delle informazioni principali sia stata comunque correttamente selezionata e riproposta successivamente nel *tweet*.

Prima di procedere con l'analisi dei dati della ricerca, possiamo sintetizzare almeno tre conclusioni parziali. La prima riguarda la variabile diatecnica che, intesa come alfabetizzazione e competenza al mezzo tecnologico, non può essere applicata al presente studio per il non uso di Twitter dichiarato dal campione della ricerca. La seconda tiene conto dei prototesti: l'assunto iniziale sulla loro semplicità complessiva e lessicale è stato confermato, ed essi sono stati compresi eccetto un caso nel contesto italiano LS. È questo un dato fondamentale per l'analisi successiva delle produzioni di *tweet* a qualsiasi livello. La terza conclusione, infine, di natura qualitativa, considera il diverso comportamento dei due sottogruppi del campione durante le fasi di raccolta dati: gli studenti slovacchi hanno mostrato sia una maggiore disponibilità ad effettuare l'esperimento sia una maggiore attenzione durante lo svolgimento dello stesso; al contrario, gli studenti italiani si sono rivelati in generale poco disponibili alla ricerca, e distratti durante lo svolgimento dell'esperimento. Questo diverso comportamento parrebbe aver avuto delle ricadute sui dati, come evidenziato in § 3.3.

3. *Analisi dei dati a livello testuale*

In prima istanza l'analisi mira a evidenziare le correlazioni tra l'uso di determinate strategie di riformulazione e il tipo di prototesto, nonché il gruppo linguistico, italiano L1 o LS, degli autori dei singoli metatesti. Per rispondere a questi interrogativi di natura variazionista è stato però necessario porsi criticamente nei confronti dei metatesti stessi, cercando di definire i possibili parametri caratterizzanti la riformulazione scritta intralinguistica digitale su Twitter.

3.1 Classificazione delle tipologie dei metatesti

L'individuazione dei possibili parametri è avvenuta tramite una prima analisi qualitativa dei singoli *tweet*, cercando di evidenziarne elementi comuni e divergenti. Gli esempi riportati in (1) e (2), infatti, sono metatesti derivati dal prototesto letterario, che in questa sede indicheremo con l'etichetta operativa di "metatesti letterari" (e, in parallelo, si avranno dunque i "metatesti giornalistici"). In entrambi gli esempi è stata selezionata e riprodotta una informazione presente nel prototesto, ossia l'opposizione tra poesia e denaro; nel secondo esempio, però, il richiamo al prototesto è reso esplicito tramite il duplice riferimento al "testo" e alla "autrice".

- (1) #itinbreve #SLO20L il valore della poesia è più grande del valore dei soldi

- (2) #itinbreve #ITA17L il testo sottolinea la necessita dell'autrice di staccarsi le necessita materiali in favore di sentimenti e pensieri

Molto più complesso risulta invece il terzo esempio, in cui vi è un rimando diretto al prototesto, ottenuto tramite la citazione dell'intero titolo della poesia e del nome dell'autrice, ma al quale non segue, come in (2), un riassunto del prototesto, ma una riflessione altra, che rappresenta un posizionamento dello scrivente rispetto al testo fornito.

- (3) #itinbreve #ITA04L Il testo Io non ho bisogno di denaro di Alda Merini è una voce fuori dal tempo che si staglia contro le abitudini odierne

Esempi di inserimento di riflessioni personali, che vanno oltre il prototesto, si ritrovano anche in altri *tweet* appartenenti sia alla tipologia letteraria sia a quella giornalistica. Vi sono tuttavia casi in cui tutti e tre i fenomeni appena evidenziati co-occorrono, andando a creare una struttura testuale che potremmo definire geometricamente perfetta, come nell'esempio (4) in cui viene prima riassunto il prototesto, poi si inserisce un rimando diretto al prototesto stesso, e infine viene aggiunto un posizionamento personale dello scrivente, diretto ed esplicito, in un segmento che si presenta separato anche graficamente rispetto al resto del testo, come si può notare dall'inserimento dei puntini di sospensione che rimarcano dunque questo carattere di aggiunta di una informazione ulteriore.

- (4) #itinbreve #SLO07L non abbiamo bisogno di soldi ma di sentimenti, così ha scritto Alda Merini in sua poesia... e ha ragione

Da questi e da altri esempi pare dunque di poter individuare tre parametri principali all'interno dei metatesti:

1. ancoraggio al prototesto tramite un rimando linguistico diretto;
2. selezione di informazioni (contenute nel prototesto);
3. aggiunta di informazioni (non presenti nel prototesto).

Tabella 1 - *Le tipologie testuali teoriche dei metatesti analizzati secondo la combinazione dei tre parametri individuati*

	<i>Ancoraggio al prototesto</i>	<i>Selezione di informazioni</i>	<i>Aggiunta di informazioni</i>
1. Non produzione	-	-	-
2. Ancoraggio puro	+	-	-
3. Riassunto puro	-	+	-
4. Commento puro	-	-	+
5. Riassunto Ancorato	+	+	-
6. Commento ancorato	+	-	+
7. Riassunto e Commento Non Ancorato	-	+	+
8. Riassunto e Commento Ancorato	+	+	+

Poiché questi tre parametri possono presentarsi singolarmente oppure co-occorrere in vario modo all'interno dei diversi metatesti, essi sono stati considerati come tratti

distintivi e dicotomici, disposti in una matrice di presenza o assenza secondo un classico modello di analisi componenziale: la combinazione di questi parametri restituisce otto tipologie testuali teoriche, classificate nella tab. 1.

La nostra proposta di tassonomia contempla anche la presenza di una non produzione, teoricamente possibile in quanto lo studente poteva limitarsi a pubblicare il *tweet* con i soli *hashtag* inseriti dal ricercatore. Data però la natura sperimentale del compito, non si danno casi di non produzioni all'interno del nostro corpus, così come risultano assenti i casi di ancoraggio puro. Tutte le altre tipologie testuali trovano invece una effettiva realizzazione all'interno del corpus: gli esempi (1) e (2) sono quindi definiti, all'interno della tassonomia, rispettivamente come riassunto puro e riassunto ancorato, mentre l'esempio (3) rappresenta un caso di commento ancorato. Il testo riportato in (4), come già evidenziato, rappresenta il caso più complesso, in cui cioè tutti e tre i parametri co-occorrono all'interno del *tweet*, rappresentando quindi un caso della tipologia 8, ossia un riassunto e commento ancorato. Si noti che l'ancoraggio può essere esplicitato in maniera diversa, magari anche sfruttando le potenzialità del mezzo Twitter attraverso l'uso dell'*hashtag* come nell'esempio in (5), anch'esso appartenente alla tipologia 8 della nostra classificazione.

- (5) #itinbreve #ITA19L #aldamerini comprende l'importanza e la funzione taumaturgica della poesia, magia che risveglia l'emozione e la vita

Di non immediata trasparenza risulta invece la tipologia 4 che abbiamo definito commento puro: essa si esplica qualora il *tweet* non contenga alcun rimando al prototesto, né come ancoraggio né attraverso la selezione di una o più informazioni. Un esempio di *tweet* appartenente a questa tipologia si trova in (6) in cui non vi è alcun legame con il prototesto giornalistico, dato che anche il riferimento al calcio non è pertinente rispetto alle informazioni fornite.

- (6) #itinbreve #SLO09G Non devi giocare calcio però fai un esercizio di regalare il tuo sorriso, abbraccio e poi vedrai cosa succede.

3.2 Distribuzione delle tipologie dei metatesti nel corpus

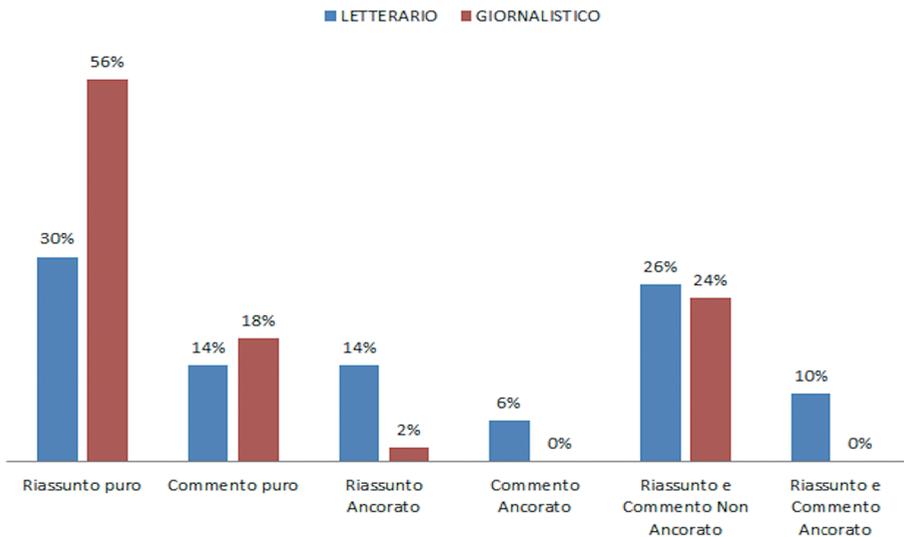
I cento metatesti del nostro corpus sono stati classificati secondo la tassonomia delineata in § 3.1 ed è stata osservata la distribuzione di ciascuna tipologia testuale all'interno del corpus in relazione alle due macro-variabili del gruppo linguistico degli scriventi e del genere del prototesto.

Prima di illustrare i risultati dell'analisi, è importante evidenziare come tutte le correlazioni presentate siano risultate anche statisticamente significative ($p < 0,005$), ad eccezione della distribuzione per tipologia del metatesto e genere del prototesto nel sottogruppo di italiano LS. Trattandosi però di un corpus allo stato attuale ancora numericamente contenuto, che risulta ancora più ridotto nel caso di divisione in sotto-campioni, si è preferito non caricare di eccessiva importanza i test statistici, interpretando i risultati in chiave prettamente qualitativa. La riscontrata significatività statistica conforta però l'analisi intrapresa circa le variabili che si è de-

ciso di utilizzare, rimarcando come per il sottogruppo di italiano LS possano esserci in gioco variabili diverse da quelle qui considerate.

Allo stato attuale della ricerca, si è in primo luogo osservato come la tipologia testuale più frequente sia quella del riassunto puro, riscontrata in 43 *tweet* su cento: ciò non stupisce date le istruzioni fornite ai partecipanti durante la fase di raccolta dati (§ 2). La seconda tipologia più frequente (25 *tweet* su cento) è quella relativamente complessa del riassunto e commento non ancorato, seguita però dal commento puro (ben 16 *tweet* su cento). Con frequenze più basse ricorrono anche le altre tipologie previste dalla nostra tassonomia: il riassunto ancorato si ritrova infatti in otto *tweet*, il riassunto e commento ancorato in cinque *tweet* e, infine, si hanno tre *tweet* di commento ancorato. Questi dati ci mostrano come il parametro dell'ancoraggio sia in generale scarsamente utilizzato, dato che è presente solo in 16 *tweet* di varie tipologie testuali, nonostante, come si è detto, gli studenti avessero la consegna specifica di produrre un riassunto da ciascun prototesto, e avessero sempre davanti i prototesti stessi durante la scrittura dei *tweet*. All'opposto, invece, l'aggiunta di informazioni altre rispetto al prototesto è molto frequente, ricorrendo nella forma di commento, sia puro sia in correlazione con altri parametri, per un totale di 49 *tweet*, dato che sorprende non solo per la già citata natura del compito richiesto, ma anche per la necessaria brevità imposta dal mezzo Twitter.

Figura 1 - Distribuzione delle tipologie dei metatesti rispetto al genere del prototesto

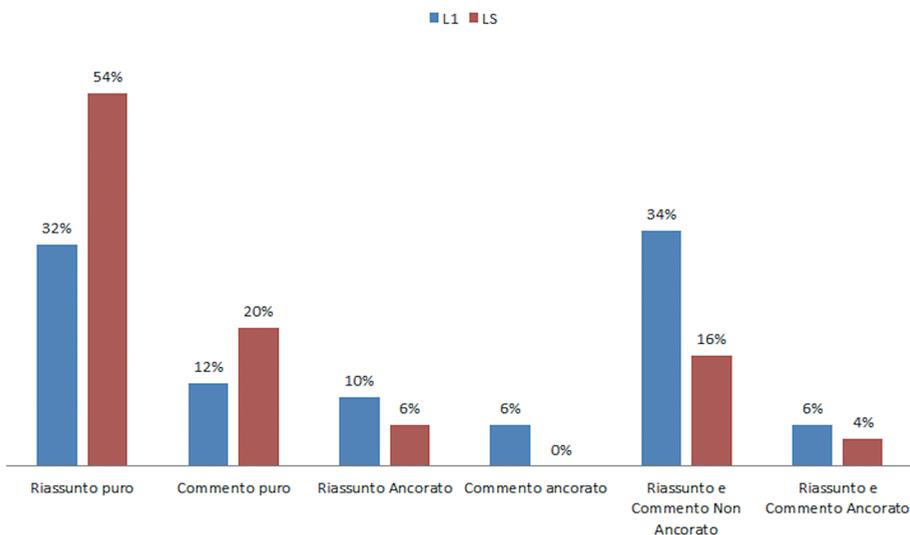


Per cercare di spiegare meglio questi dati, se ne è osservata dunque la variazione rispetto alle due macro-variabili precedentemente illustrate. I risultati riportati nella fig. 1 mostrano delle differenze abbastanza evidenti rispetto al genere del prototesto. In particolare, si nota come la strategia del riassunto puro è di gran lunga preferita non solo in generale, ma soprattutto se legata al prototesto giornalistico, dato che il 56% dei metatesti giornalistici appartiene a questa tipologia, contro il 30%, valore

comunque massimo, dei metatesti letterari. Per questi ultimi si osserva però come sia molto frequente che alle informazioni selezionate venga aggiunto anche un posizionamento personale in diverse forme di commento, ma soprattutto nella tipologia di riassunto e commento non ancorato (ben 26% delle occorrenze). Inoltre l'ancoraggio al prototesto, parametro in generale poco utilizzato, risulta fortemente legato al prototesto letterario, presentandosi sia nella forma di riassunto o commento ancorati (rispettivamente, 14% e 6% delle occorrenze), sia nella tipologia testuale più complessa di riassunto e commento ancorato (10% delle occorrenze).

Rispetto al gruppo linguistico di appartenenza degli studenti, si osserva una netta differenza tra il gruppo italiano L1 e quello italiano LS, come mostrato dalla fig. 2. Si può infatti notare come nel gruppo di italiano LS prevalga fortemente la tipologia del riassunto (27 *tweet* su 50, corrispondenti al 54%), mentre sono gli studenti di madrelingua italiana a introdurre in misura percentualmente rilevante un posizionamento personale rispetto ai prototesti, come si può notare dall'elevata presenza di riassunti e commenti non ancorati (34%), nonché dal fatto che gli unici casi di commenti ancorati siano prodotti da madrelingua italiana (tre casi, corrispondenti al 6%). Un dato in questo senso apparentemente discordante riguarda la distribuzione della tipologia testuale che abbiamo definito di commento puro, che risulta lievemente maggioritaria nel gruppo slovacco (dieci casi, 54%) rispetto al gruppo italiano (sei casi, 32%); è possibile che questa distribuzione sia legata alle dimensioni ridotte del corpus, ma è anche ipotizzabile che esistano differenze interne, di natura semantico-testuale, tra i diversi *tweet* classificati come commenti, che sarà interessante indagare in seguito.

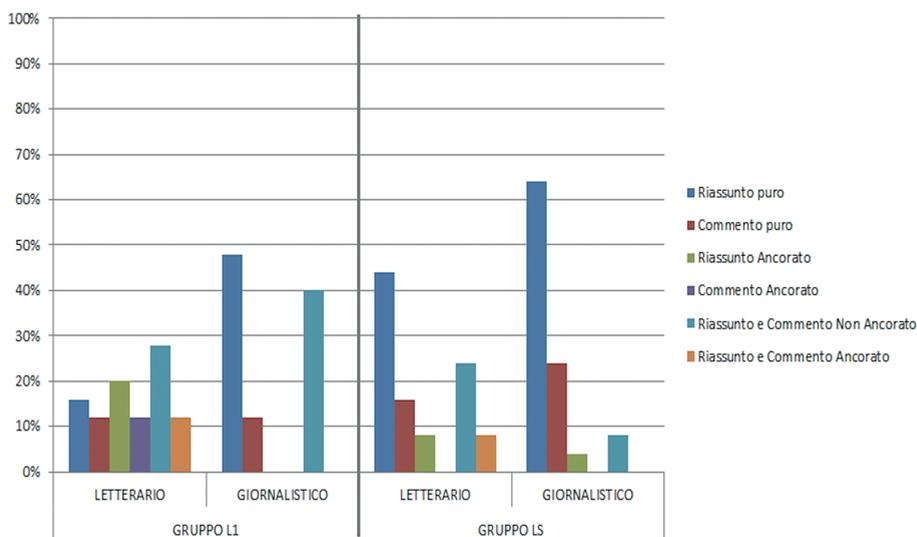
Figura 2 - Distribuzione delle tipologie dei metatesti rispetto al gruppo linguistico degli studenti



Le due macro-variabili sono state poi analizzate congiuntamente. Date le più volte menzionate dimensioni ridotte del corpus (solo 50 *tweet* per ogni sottogruppo) è opportuno che i risultati di questa analisi siano interpretati con cautela e considerati come indicazioni orientative per gli sviluppi successivi del lavoro.

L'analisi ha evidenziato diversità nella distribuzione delle tipologie dei metatesti nel gruppo italiano L1 e italiano LS rispetto alla tipologia dei prototesti (fig. 3).

Figura 3 - *Distribuzione delle tipologie dei metatesti rispetto al genere dei prototesti nei due sottogruppi italiano L1 e italiano LS*



Per il gruppo italiano L1 emerge una prevalenza della tipologia del riassunto puro (12 casi su 25, 48%) legata al prototesto giornalistico, seguita però da un quasi uguale numero di casi (ben dieci su 25, 40%) di riassunto e commento non ancorato, confermando la tendenza degli italiani ad andare oltre i prototesti forniti, aggiungendo quindi qualcosa di personale. Nel gruppo italiano LS, come si è osservato in precedenza, la tipologia del riassunto puro è nettamente prevalente, ricorrendo in 16 casi su 25 (64%) per i metatesti giornalistici, e in 11 casi su 25 (44%) per quelli letterari, confermando che questa è la tipologia testuale più utilizzata dal gruppo italiano LS, a prescindere dal livello di conoscenza della LS, qui non considerato (§ 2). Una differenza rispetto ai due prototesti emerge invece in relazione al parametro dell'ancoraggio, il quale risulta praticamente assente nei metatesti giornalistici (un caso, 4%), mentre è più utilizzato in quelli letterari (quattro casi, 16%). Anche la tipologia di riassunto e commento non ancorato risulta più frequente nei metatesti letterari (sei casi, 24%) rispetto a quelli giornalistici (due casi, 8%).

3.3 Discussione dei risultati

L'analisi di questo primo corpus di cento *tweet* ha permesso di giungere ad alcune interessanti considerazioni sulle riformulazioni testuali scritte intralinguistiche di-

gitali. In primo luogo si è potuto osservare come, nonostante il compito richiesto agli studenti fosse quello di riassumere i prototesti in forma breve, i metatesti mostrino la presenza anche di elementi diversi (ancoraggio e aggiunta di informazioni) rispetto all'attesa selezione e riformulazione di informazioni.

L'analisi della distribuzione delle otto tipologie testuali teoriche all'interno del corpus ha permesso di evidenziare come la tipologia del riassunto puro sia di gran lunga preferita, ma come in realtà essa ricorra con più frequenza all'interno delle produzioni degli studenti italiano LS nonché nei metatesti giornalistici. Si è potuto notare come gli studenti italiano L1 tendano molto più frequentemente al commento, soprattutto in aggiunta ad altre informazioni estratte dal prototesto, evidenziando quindi una preferenza per strategie testuali complesse. Questo non accade per il gruppo di italiano LS, in cui prevale la tipologia del riassunto puro, ma in cui il parametro di aggiunta di informazioni altre rispetto al prototesto non è assente né infrequente, specialmente per quanto riguarda i metatesti letterari. Questa distribuzione potrebbe ovviamente essere dettata dalla diversa competenza in italiano, specialmente in fase di produzione, ma a nostro giudizio essa potrebbe essere legata alla maggiore attenzione che il gruppo LS ha prestato alle istruzioni fornite dal ricercatore (§ 2). Data la metodologia di elicitazione dei dati, dunque, si può inferire che gli studenti slovacchi si siano attenuti maggiormente alle indicazioni fornite rispetto a quelli italiani, per cui la diversa distribuzione delle tipologie dei metatesti potrebbe riflettere una differenza più di natura pragmatico-culturale che non semplicemente una minore competenza linguistica. Riteniamo che le indicazioni del task non costituiscano una variabile in sé, in quanto sono state mantenute costanti per i due gruppi di studenti L1 e LS. La possibile nuova variabile da tenere in considerazione è dunque l'atteggiamento rispetto al task stesso, non la sua formulazione; tale nuova variabile, poiché di tipo pragmatico-culturale, si può considerare inclusa nella macro-variabile gruppo linguistico degli studenti, ma potrebbe rivelarsi interessante una ulteriore indagine su questo aspetto specifico anche in ottica pragmatica o, eventualmente, glottodidattica, dato il contesto di italiano LS degli studenti slovacchi selezionati. L'analisi conferma comunque l'esistenza di una variazione nelle tipologie testuali dei metatesti in riferimento a tale macro-variabile anche da questo punto di osservazione.

4. Conclusioni e prospettive di ricerca

Utilizzando la nozione di riformulazione scritta intralinguistica e applicandola al mezzo digitale Twitter, proprio per le sue intrinseche caratteristiche di brevità, è stato possibile giungere ad elaborare una tassonomia delle riformulazioni secondo tre parametri (ancoraggio al prototesto, selezione di informazioni, aggiunta di informazioni). Questa tassonomia definisce otto tipologie testuali teoriche, le quali mostrano una distribuzione variabile all'interno del nostro corpus sia rispetto al genere testuale del prototesto (letterario, giornalistico) sia rispetto al gruppo linguistico degli scriventi (italiano L1, italiano LS) come discusso in § 3.3.

Tale tassonomia risulta interessante in quanto propone una categorizzazione per tipologie testuali dei due principali passaggi linguistici e cognitivi richiesti dal riassunto, ossia la comprensione e selezione delle informazioni di un prototesto, e la loro successiva riformulazione e ridistribuzione all'interno di un metatesto (Fornasiero - Tamiozzo Goldmann, 2013). Al di là delle specificità richieste dal mezzo Twitter, non si esclude una possibile applicazione di tale tassonomia ad altri processi di riformulazione di prototesti anche al di fuori della CMC. Tale applicazione costituisce senz'altro un interessante sviluppo futuro della ricerca, anche in ottica contrastiva tra digitale e non digitale, tenendo comunque fisso il limite del numero di caratteri, verificando quindi empiricamente la posizione di De Mauro (2002) in proposito, e ricordata anche in apertura del presente contributo (§ 1).

Come più volte ricordato, questo studio rappresenta un primo stadio verso una analisi più estesa delle riformulazioni brevi intralinguistiche digitali. Una analisi qualitativa su diversi livelli (es. morfosintattico, pragmatico-discorsivo) rappresenta sicuramente un successivo passo necessario al fine di osservare possibili variazioni anche all'interno delle stesse tipologie testuali qui delineate⁷. Avendo inoltre constatato l'effettiva validità della metodologia di indagine predisposta per questa prima raccolta dati, si procederà a una estensione del campione dei metatesti sia per il gruppo di italiano L1 sia per quello di italiano L2; per quest'ultimo si prevede inoltre la raccolta di metatesti su Twitter anche nella L1 degli studenti, ossia lo slovacco.

Sarebbe inoltre interessante considerare, come sviluppo ulteriore della ricerca, anche nuovi generi dei prototesti, dal momento che è stata osservata una variazione significativa rispetto a questa macro-variabile. Si sta inoltre valutando la possibilità di considerare altre tipologie di riformulazioni lungo un continuum di brevità da una lunghezza minima, data da Twitter, a una lunghezza massima, altrimenti definita a priori, per osservare se e come ciò influisca sulle diverse tipologie testuali, analizzando quindi più puntualmente il legame tra tipologia testuale e brevità.

Infine, non sono da sottovalutare le possibili implicazioni didattiche di ricerche di questo tipo, specialmente in relazione a quella che Chiusaroli ha definito "educazione digitale", già richiamata in apertura del presente contributo.

Bibliografia

ARCANGELI M. (2015), *Biografia di una chiocciola. Storia confidenziale di @*, Castelvecchi, Roma.

CHIUSAROLI F. (2014), Sintassi e semantica dell'hashtag: studio preliminare di una forma di Scritture Brevi, in BASILI R. - LENCI A. - MAGNINI B. (eds.), *The First Italian Conference on Computational Linguistics CLiC-it 2014 Proceedings*, Pisa University Press, Pisa: 117-121.

⁷ Questo sviluppo della ricerca costituisce l'argomento di un contributo a cura degli autori al workshop "Formazione delle parole e trattamento morfosintattico di scritture brevi della comunicazione digitale" del XLIX Congresso Internazionale di Studi della SLI (Malta, 24-26 settembre 2015).

- DE BEAUGRANDE R. - DRESSLER W. (1994), *Introduzione alla linguistica testuale*, Il Mulino, Bologna.
- DE MAURO T. (2002), *Capire le parole*, Laterza, Roma/Bari.
- DE MAURO T. (2008), *Lezioni di linguistica teorica*, Laterza, Roma/Bari.
- FORNASIERO S. - TAMIOZZO GOLDMANN S. (2013), *Scrivere l'italiano. Galateo della comunicazione scritta*, Il Mulino, Bologna.
- HERRING S. (1996), Introduction, in HERRING S. (ed.), *Computer-Mediated Communication: Linguistic, Social and Cross-Cultural Perspectives*, John Benjamins, Amsterdam: 1-10.
- HERRING S. - STEIN D. - VIRTANEN T. (2013), Introduction to the pragmatics of computer-mediated communication, in HERRING S. - STEIN D. - VIRTANEN T. (eds.), *Pragmatics of Computer-Mediated Communication*, De Gruyter Mouton, Berlin: 3-32.
- KOLLAR M. - CZWITOKOWICS T. (2013), *Mapping digital Media: Slovakia. A report by the Open Society Foundation*, Open Society Media Program, London.
- OSIMO B. (2010), *Propedeutica della traduzione*, Hoepli, Milano.
- PIERUCCI M.L. (2015), *Un successo il progetto universitario di Chiusaroli che spopola sul web*. Disponibile in: http://www.corriereadriatico.it/SPETTACOLI/amp_8203_progetto_universitario_macerata_francesca_chiusaroli_web_scritture_brevi/notizie/11_40858.shtml (data ultimo accesso 20 luglio 2015).
- PISTOLESI E. (2014), Scritture digitali, in ANTONELLI G. - MOTOLESE M. - TOMASIN L. (a cura di), *Storia dell'italiano scritto. Italiano dell'uso*, vol. III, Carocci, Roma: 349-375.
- POPOVIČ A. (2006), *La scienza della traduzione. Aspetti metodologici. La comunicazione traduttiva*, Hoepli, Milano.
- SERIANNI L. (2003), *Italiani scritti*, Il Mulino, Bologna.
- TRIFONE P. (2014), *L'italiano twitterario*, relazione al Convegno "Scritture brevi" e variabilità diatecnica, Università di Napoli L'Orientale, Procida, 26-28 giugno 2014.
- WE ARE SOCIAL (ed.) (2014), *European Digital Landscape 2014*. Disponibile in: http://147.102.16.219/demo1/attachments/124_european%20digital%20landscape%202014.pdf/ (data ultimo accesso 20 luglio 2015).
- ZAPPAVIGNA M. (2015), Searchable talk: the linguistic functions of hashtags, *Social Semiotics*, 25(3): 274-291.

Il *Réseau Lexical du Français*: una banca dati per l'apprendimento del lessico francese

In this paper, we explore the lexicon in terms of its importance for second language (L2) teaching. More specifically, we present The *Réseau Lexical du Français* (Polguère, 2014), in order to illustrate its possible applications for L2 learning and teaching. The *Réseau Lexical du Français* (RLF) is a network of lexical units connected to each others by lexical links which encode semantic or combinatorial lexical relations. In order to understand and use this original lexical resource, some metalinguistic notions must be mastered. The first part of the paper will introduce the RLF and the importance of dictionaries and lexical databases in L2 learning and teaching. We will then present metalinguistic notions, within our theoretical framework, the *Meaning-Text Theory* and its lexical branch, the *Explanatory and Combinatorial Lexicology*. In the last part of the article, we propose pedagogical activities using adjectival lexical entries from the RLF and the related lexical networks in which they appear.

1. *Introduzione*

Come hanno dimostrato diversi studi (August *et al.*, 2005; Schmitt, 2000), il lessico assume un ruolo chiave nell'apprendimento di una lingua. Le ricerche quantitative sul lessico (Kieffer - Lesaux, 2012; Verhallen - Schoonen, 1993; Schmitt, 2010) hanno consentito di identificare il numero di parole che dovrebbero essere apprese da un discente. Altri aspetti non squisitamente quantitativi sono da considerarsi imprescindibili ai fini dell'acquisizione di una buona competenza lessicale: questi riguardano il lato 'qualitativo' del lessico, ovvero la conoscenza, da parte dell'apprendente, di informazioni inerenti a tali parole. Il nostro contributo si colloca all'interno di questa prospettiva e mira a dimostrare il ruolo dei dizionari e di altre risorse lessicografiche nello sviluppo delle competenze lessicali di un discente.

Il nostro studio verterà sul *Réseau Lexical du Français* (d'ora in poi RLF), una risorsa lessicografica che non può essere annoverata tra i dizionari nel senso 'tradizionale' del termine: in effetti, i dati lessicali sono strutturati sotto forma di rete e tale strutturazione rispecchierebbe il modello di organizzazione del nostro lessico mentale come presentato da Aitchison (2003) e da Wolter (2006).

Prima di addentrarci nell'analisi di tale risorsa, tratteremo in modo più generale dell'utilità del dizionario nella didattica delle lingue, precisando altresì che faremo es-

¹ Università di Verona.

² Université du Québec à Montréal.

senzialmente riferimento alle risorse lessicografiche della tradizione francese e alla loro possibile applicazione nell'ambito dell'insegnamento della lingua francese.

Seguiranno la presentazione della *Théorie Sens-Texte* e della *Lexicologie Explicative et Combinatoire* e l'illustrazione delle nozioni di *vocabole*, *lexie*, *definizione*, *relazioni semantiche* (*polisemia*, *sinonimia*, *antonimia*, *iperonimia*) e *collocazione* così come sono state elaborate nell'ambito dello stesso quadro teorico. Tali nozioni corrispondono ad altrettante competenze preliminari di ordine metalinguistico necessarie all'apprendimento del lessico tramite la banca dati RLF.

Nell'ultima parte, un insieme ridotto di *lexies* aggettivali ci consentirà di soffermarci su tre aspetti esemplificativi che ci permetteranno di caratterizzare il significato e l'utilizzo di tali *lexies* rispetto all'insegnamento e all'apprendimento di una lingua: la parafrasi; la polisemia; le relazioni sintagmatiche e paradigmatiche.

Ci concentreremo, in particolare, sui seguenti aggettivi: RAPIDE e CRIMINEL³ per la parafrasi; NERVEUX per la polisemia; FERROVIAIRE per le relazioni sintagmatiche e paradigmatiche.

La scelta è ricaduta su queste *lexies* per diverse ragioni: innanzitutto, esse appartengono al lessico 'fondamentale' della lingua francese; inoltre, le loro caratteristiche semantiche intrinseche ci consentono di illustrare meglio l'utilizzo didattico della banca dati RLF secondo le tre prospettive che abbiamo scelto; infine, le rispettive reti lessicali sono particolarmente ricche, presentando così un interesse anche dal punto di vista dello sviluppo del lessico di un apprendente.

Ciò che presenteremo è l'utilizzo della banca dati RLF in quanto insieme di potenzialità ai fini didattici. Tali potenzialità verranno messe in luce sulla base di un procedimento onomasiologico: il punto di partenza, quindi, sarà costituito dal significato – che verrà ricostruito in diversi modi – a partire dal quale si avrà accesso alla *lexie*. Più particolarmente, per ogni singolo caso, simuleremo le domande che un apprendente potrà porsi, sulla base dell'entourage lessicale di ogni *lexie* – l'insieme di nodi visualizzabili simultaneamente – all'interno di ogni singola rete lessicale, per proporre possibili parafrasi finalizzate all'accesso alle unità lessicali scelte.

2. Lessico, dizionari e risorse lessicali

2.1 Lessico, dizionari e apprendimento delle lingue

Diversi studi (Giasson, 2011; Scott *et al.*, 2008) hanno dimostrato che il dizionario (e le altre risorse lessicali ad esso simili, come ad esempio la banca dati RLF) svolge un ruolo importante nell'insegnamento/apprendimento di una lingua, favorendo, in special modo, lo sviluppo delle conoscenze lessicali, le quali costituiscono di fatto un indicatore rilevante del successo scolastico (Biemiller, 2011). Inoltre, tale risorsa consentirebbe di consolidare le competenze relative alla comprensione scritta e alla produzione scritta

³ Conformemente all'uso della LEC, utilizzeremo il maiuscoletto per indicare le unità lessicali di tipo *vocabole* e *lexie* e il grassetto per indicare il nome assegnato alle funzioni lessicali.

(Beech, 2004; Bishop, 2000; Hartmann, 1999; Carstens, 1995), tanto negli apprendenti di una L1, quanto in quelli di una L2.

Oggi gli apprendenti hanno accesso ad una grande varietà di dizionari elettronici che facilitano l'accesso all'informazione lessicale (*Trésor de la langue française informatisé*, *Petit Robert électronique*, *Antidote*, *Usito*, ecc.). Da qualche decennio, vengono messe a punto banche dati lessicali che hanno legami di parentela coi dizionari, ma che si concentrano su un tipo d'informazione lessicale particolare, come ad esempio le relazioni di sinonimia (Wordnet), le collocazioni di un ambito particolare (DAFLES), ecc. Grazie agli sviluppi dell'informatica e alle possibilità offerte dal mondo del digitale, sono state inoltre concepite e realizzate risorse lessicali di nuova generazione: la banca dati RLF ne è un esempio.

La disponibilità di risorse lessicali contenenti una grande varietà d'informazioni non presuppone necessariamente che gli utenti siano in grado di fruirne pienamente. In effetti, la consultazione e l'utilizzo di un dizionario, cartaceo o elettronico, oppure di una banca dati lessicale, richiede un alto grado di competenza e una varietà di conoscenze e abilità. Come dimostrato da Cowie (1999), gli apprendenti non dispongono di conoscenze e competenze sufficienti all'utilizzo efficace di tale tipologia di strumenti. Per quanto riguarda gli insegnanti, invece, essi hanno spesso una visione ingenua del dizionario (Neubach - Cohen, 1988; Anctil, 2011) e non sembrano avere una conoscenza approfondita della ricchezza delle informazioni in esso contenute. Così, poiché non sono a conoscenza del ventaglio di risorse a loro disposizione, né tantomeno delle particolarità che caratterizzano ogni singolo strumento lessicografico, incentrano raramente le loro lezioni sulle modalità più efficaci con cui effettuare ricerche nei dizionari (Dolezal - McCreary, 1999; Atkins, 1998; Nesi, 1999).

Lew (2013) riassume il problema ricordando che esistono due ingredienti che possono garantire un uso appropriato ed efficace dei dizionari (e, aggiungiamo noi, delle risorse lessicali): 1) dizionari (o risorse lessicali) di grande qualità che siano al tempo stesso 'user-friendly' e 2) utilizzatori che siano consapevoli di ciò che fanno quando consultano un dizionario (o altre risorse lessicali).

2.2 Il Réseau Lexical du Français

Lanciato nel 2011 presso il Laboratoire ATILF-CNRS e diretto da Alain Polguère, il progetto RLF si propone di descrivere la rete lessicale della lingua francese "selon les principes de la Lexicologie Explicative et Combinatoire, tels qu'appliqués dans le *Dictionnaire explicatif et combinatoire du français contemporain* [...] et dans le projet *DiCo / LAF*" (Barque - Polguère, 2013: 231).

La banca dati RLF non è un dizionario nel senso proprio del termine: appartiene ai progetti di tipo 'lexical networks' come WordNet (Fellbaum, 1998) e FrameNet (Baker *et al.*, 2003; Ruppenhofer *et al.*, 2010); così, tale banca dati "is similar to a lexical system [...]: a huge network of lexical units connected by a broad range of lexical links encoding semantic or combinatorial lexical relations" (Lux-Pogodalla - Polguère, 2011: 55). In quanto rete lessicale, si differenzia dai dizionari tradizionali poiché le unità lessicali trattate non vengono descritte secondo un'organizzazione lineare – classificazione in

ordine alfabetico e descrizione testuale delle proprietà semantiche e grammaticali. Le unità lessicali vengono invece organizzate secondo un sistema di grafi lessicali⁴: ogni nodo corrisponde a un'unità lessicale e le relazioni che tessono la rete corrispondono a relazioni sintagmatiche o paradigmatiche. La descrizione specifica di tutte le proprietà delle unità lessicali – proprietà grammaticali, definizione analitica, aspetti combinatori, legami derivazionali, esempi, fraseologia – si trova alla base della configurazione tramite il sistema di grafi lessicali (ovvero la macrostruttura) e ne costituisce la microstruttura. Per una presentazione dettagliata degli aspetti della microstruttura e della macrostruttura della banca dati rimandiamo a Polguère (2014). La nostra attenzione si concentra, in questo contributo, sulla macrostruttura in quanto riserva a cui attingere il materiale per l'elaborazione di attività pedagogiche incentrate sul lessico. Teniamo a precisare che, ad oggi, la banca dati si trova ancora ad uno stadio embrionale e non è accessibile al grande pubblico: per questo, i grafi generati non riflettono le relazioni definitive fra unità lessicali.

3. Il quadro teorico e le nozioni ad esso connesse

3.1 La *Théorie Sens-Texte* e la *Lexicologie Explicative et Combinatoire*

La *Lexicologie Explicative et Combinatoire* (d'ora in poi LEC) (Mel'čuk *et al.*, 1995) è il ramo lessicale della *Théorie Sens-Texte* (d'ora in poi TST) (Mel'čuk, 1973; Polguère, 2011).

La TST si propone di rendere conto dell'associazione che un locutore di una lingua naturale è in grado di fare fra un *sens* e l'insieme delle entità lessicali tramite le quali tale *sens* si può manifestare. Per fare ciò, la TST considera le lingue come sistemi di regole grammaticali che, “*appliquées de façon séquentielle du niveau sémantique jusqu'au niveau phonétique et vice-versa, établissent une correspondance bidirectionnelle Sens – Texte*” (Polguère, 2011: 83).

La LEC è il ramo lessicale della TST e si occupa di fornire un apparato logico-formale adeguato allo studio del lessico. Particolare importanza viene data alla nozione di parafrasi e alle risorse linguistiche – le *lexies* – atte ad esprimere tale (o tali) parafrasi: la TST e la LEC concepiscono il lessico di una lingua come strutturato sulla base di un insieme di relazioni paradigmatiche e sintagmatiche che vengono codificate mediante il sistema delle *funzioni lessicali* (Mel'čuk, 1996: 38).

Queste relazioni vengono formalizzate mediante la formula tradizionale della funzione nel senso matematico del termine – $f(x) = y$ – in cui le due variabili, trasposte in ambito lessicologico, vengono associate a due *lexies* di una lingua, chiamate per convenzione L_1 e L_2 .

La funzione lessicale **Anti** è un esempio di funzione lessicale paradigmatica che consente di formalizzare la relazione di antonimia fra due *lexies*, come nell'esempio che segue:

⁴ La visualizzazione tramite grafi lessicali è stata elaborata da Kodex Lab.

Anti (*grand*) = *petit*.

Per la TST e la LEC il legame fra lessicologia e lessicografia è indissociabile: dalle riflessioni teoriche e dalle formalizzazioni proposte tramite lo studio del lessico delle lingue naturali – la lessicologia – scaturisce la descrizione sperimentale del lessico, ovvero la pratica lessicografica.

3.2 Nozioni metalinguistiche per l'apprendimento del lessico

Numerosi studi nell'ambito della lessicologia di tradizione francofona hanno identificato le nozioni metalinguistiche di base propedeutiche all'insegnamento/apprendimento del lessico e, collateralmente, alla consultazione efficace di un dizionario (Polguère - Tremblay, 2003; Tremblay, 2009): *vocabole*, *lexie* (o *unità lessicale*), *definizione*, *collocazione* e diversi tipi di relazioni semantiche (*polisemia*, *sinonimia*, *antonimia*, *iperonimia*). Tali nozioni – alcune delle quali sono state elaborate nell'ambito della LEC⁵ – sono indispensabili per la comprensione del tipo di approccio che proponiamo nel presente articolo⁶. Poiché il nostro studio riguarda l'insegnamento/apprendimento della lingua francese, le unità lessicali scelte e le rispettive definizioni verranno mantenute in francese.

La prima nozione lessicale fondamentale è quella di *polisemia*. Essa designa la proprietà di una parola, o più precisamente di un *vocabole*, di avere diverse accezioni che condividono almeno una componente semantica. Ad esempio, il *vocabole* NERVEUX è polisemico poiché possiede almeno quattro accezioni:

- NERVEUX1: 'relatif aux nerfs': *le système nerveux*.
- NERVEUX2: 'dont les muscles et les tendons sont très apparents': *des mains nerveuses*.
- NERVEUX3: 'qui résulte d'un état d'excitation psychologique, d'une émotivité intense': *rire nerveux*.
- NERVEUX4: 'Se dit d'une personne émotive, agitée, qui peut difficilement rester calme': *Elle est toujours très nerveuse avant de prononcer une conférence*.

Ognuno di questi significati corrisponde a una parola della lingua, nozione che designeremo di seguito con il termine *lexie* o *unità lessicale*. Ogni dizionario codifica le diverse *lexies* di un *vocabole* polisemico in maniera diversa: per esempio, nel *Petit Robert* ogni definizione è preceduta da un numero che consente di distinguere le unità lessicali fra di loro. Le unità lessicali di uno stesso *vocabole* sono collegate da una o più relazioni di significato: nel nostro caso, le *lexies* del *vocabole* NERVEUX contengono il significato di NERF.

Un'ulteriore nozione sulla quale vale la pena soffermarci è quella di *definizione*. In lessicografia, la redazione di una definizione dipende da un lavoro rigoroso di analisi del significato e segue una procedura formale. Innanzitutto, è necessario identificare la classe di parole del *definiendum*, poi scegliere il termine generico (o l'espressione equivalen-

⁵ Ci riferiamo, in particolare, alle nozioni di *vocabole* e di *lexie*. Le altre nozioni sono comuni ai diversi approcci, ma assumono una particolare specificità in relazione alla nozione di *fonction lexicale* così come è stata brevemente presentata nel par. 2.1. Per maggiori dettagli si veda Mel'čuk *et al.* (1995).

⁶ Forniremo le traduzioni in italiano solo qualora queste siano già in uso; in caso contrario, manterremo le denominazioni francesi.

te) che comparirà nella prima parte della definizione. Per esempio, il sostantivo *TORSE* (*L'homme bombait fièrement le torse*) sarà definito tramite l'etichetta semantica generica 'partie du corps'. La seconda parte della definizione include ciò che viene chiamato *différence spécifique*; si tratta di unità di significato che permettono di distinguere un'unità lessicale da altre unità lessicali che possiedono lo stesso senso generico. Per esempio, i sostantivi *POITRINE* e *TORSE* saranno definiti tramite la medesima etichetta semantica generica 'partie du corps', ma verranno distinti tramite unità di significato diverse ('comprenant les épaules et la poitrine' per *TORSE* e 'qui s'étend du cou à l'abdomen' per *POITRINE*).

La caratterizzazione di una parola passa altresì per l'identificazione di diversi tipi di relazioni semantiche (sinonimia, antonimia, iperonimia) che essa intrattiene con altre parole: è sulla base di tali relazioni che si struttura il lessico. Per questa ragione, capire l'organizzazione del lessico può aiutarci a cogliere meglio il significato delle parole. Ogni unità lessicale può, in effetti, caratterizzarsi sulla base delle relazioni semantiche che la legano ad altre unità lessicali all'interno dell'insieme del lessico. L'analisi delle relazioni semantiche proprie di ciascuna unità lessicale consente di capire meglio, al contempo, ciò che i rispettivi significati hanno in comune e ciò che, invece, li differenzia.

Questo insieme di nozioni costituisce la base per poter capire il contenuto di un articolo di dizionario e la strutturazione delle informazioni in una risorsa lessicografica come la banca dati RLF.

4. *Analisi dei dati tratti dalla banca dati RLF: il caso degli aggettivi*

4.1 La parafrasi

Nella pratica lessicografica della banca dati RLF, l'analisi strutturata del significato delle *lexies* avviene mediante una scomposizione del significato in due parti. La prima descrive la struttura attanziale dell'unità lessicale, mentre la seconda contiene una parafrasi analitica (etichettata mediante un linguaggio XML) con una componente centrale e una o più componenti periferiche inquadrate da un'etichetta semantica, come nell'esempio che segue (Barque - Polguère, 2013: 238):

```
<DEFINITION_RLF>
  <DEFINIENDUM>X admire Y pour Z</DEFINIENDUM>
  <DEFINIENS label = "apprécier">
    <PP>Z étant un acte ou une caractéristique de Y</PP>
    <CC>L'individu X apprécie Y</CC>
    <PC role="intensité">beaucoup</PC>
    <PC role="cause">du fait des qualités exceptionnelles de Z</PC>
  </DEFINIENS>
</DEFINITION_RLF>
```

Abbiamo scelto, a titolo esemplificativo, gli aggettivi *RAPIDE* e *CRIMINEL* per le loro caratteristiche semantiche intrinseche, e per l'interesse che presenta la scomposizione dei rispettivi significati tramite la definizione lessicografica.

È opportuno sottolineare che la definizione lessicografica di aggettivi (così come la definizione di avverbi) si distingue dalla definizione di *lexies* appartenenti ad altre parti del discorso (nome e verbo) per l'assenza di unità lessicali iperonimiche: a causa di ciò, l'etichetta semantica generica corrisponde alla parafrasi dell'unità lessicale iperonimica assente (Frassi, 2012; 2014). Tale parafrasi contiene un sostantivo che può intrattenere un legame morfologico con l'aggettivo da definire. Per certi aggettivi – nella fattispecie gli aggettivi primari – l'etichetta semantica e la componente centrale si appoggiano su *lexies* nominali che non intrattengono alcun legame morfologico con il *definiendum* (es.: RAPIDE = 'qui a une certaine vitesse'), mentre per altri – e in particolare i derivati morfologici, chiamati anche aggettivi di relazione o pseudo-aggettivi se sono derivati denominali – l'etichetta semantica e la componente centrale si appoggiano su *lexies* nominali che costituiscono, in termini morfologici, la radice del derivato (es.: NERVEUX = 'relatif aux nerfs').

Tale peculiarità delle *lexies* aggettivali pone un problema non secondario nell'ambito dell'insegnamento-apprendimento di una lingua (nel nostro caso il francese), dal punto di vista dell'accesso all'unità lessicale da parte dell'apprendente a partire dalla parafrasi, secondo quindi un procedimento onomasiologico.

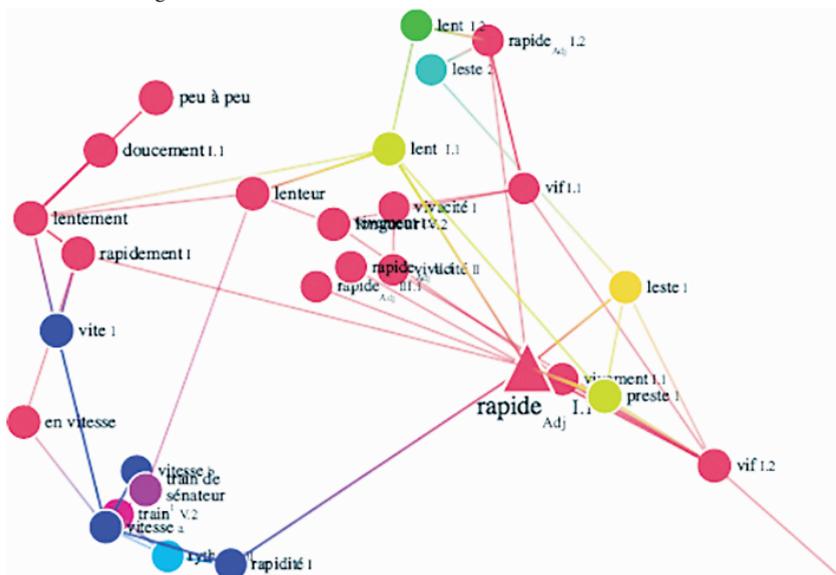
Le due *lexies* che abbiamo scelto ci consentiranno di dimostrare le virtù potenziali di una banca dati di tipo 'rete lessicale' seguendo questo procedimento, tenuto conto delle differenze sostanziali che esistono fra un aggettivo primario (RAPIDE) e un aggettivo di relazione (CRIMINEL).

In quanto aggettivo primario, RAPIDE non è, dal punto di vista morfologico, un derivato da una radice nominale o verbale. La parafrasi minima che corrisponde all'etichetta semantica (e quindi alla componente centrale) deve appoggiarsi su una *lexie* nominale che non intrattiene alcun legame morfologico con la *lexie* RAPIDE. In questo caso si tratta di VITESSE: RAPIDE = 'qui a une certaine vitesse'.

Nel caso di CRIMINEL, invece, la parafrasi deve tener conto del legame con la radice nominale CRIME. A seconda dell'accezione di CRIMINEL – si tratta, in effetti, di un *vocable* polisemico – avremo CRIMINEL1 = 'relatif au crime' (*brigade criminelle*) e CRIMINEL2 = 'qui est un crime' (*acte criminel*). Per evitare di utilizzare, nella definizione lessicografica, una *lexie* che abbia il medesimo significato del *definiendum* – che non aiuterebbe l'apprendente a capire il significato di CRIMINEL – è verosimilmente improbabile che un apprendente che non conosce il significato di CRIMINEL possa conoscere il significato di CRIME – è possibile appoggiarsi su una *lexie* nominale che costituisce l'iperonimo di CRIME, ovvero MÉFAIT⁷: avremo così 'relatif au méfait' per CRIMINEL1 e 'qui est un méfait' per CRIMINEL2.

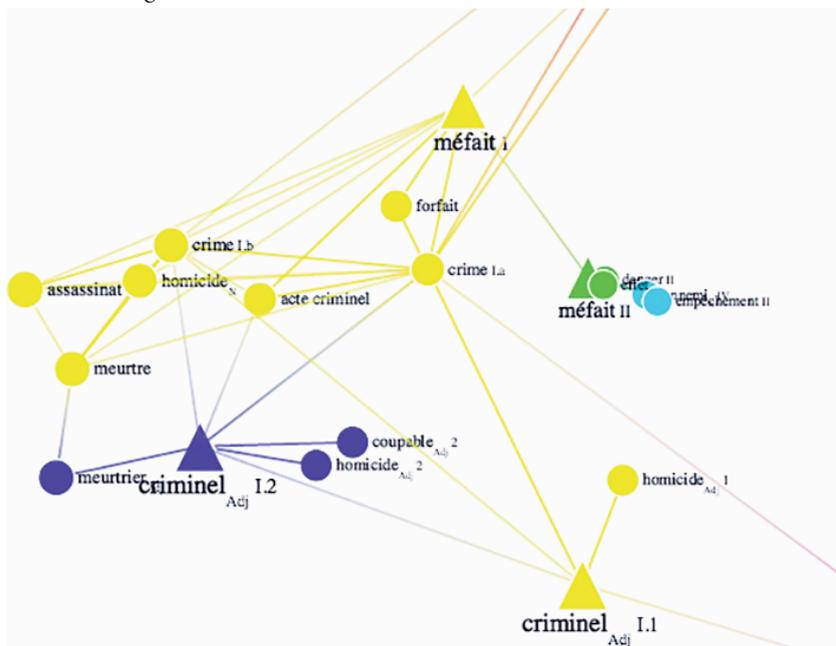
Le rispettive reti lessicali ci consentono di accedere in modo intuitivo ai significati di ognuna delle *lexies* scelte attraverso la parafrasi.

⁷ Riconosciamo che l'accesso a MÉFAIT è relativamente difficile per un apprendente principiante di francese L2; l'esercizio si adatterebbe meglio a livelli più elevati e nella direzione contraria, ovvero partendo da CRIME per accedere alla *lexie* più generica. Ricordiamo che la banca dati RLF si presta anche ad attività didattiche per il francese L1; l'accesso a CRIME tramite MÉFAIT sarebbe più adatto a questo tipo di apprendenti.

Figura 1⁸ - La rete lessicale di RAPIDE nella banca dati RLF

RAPIDE I.1: 'qui a une certaine vitesse'

Figura 2 - La rete lessicale di CRIMINEL nella banca dati RLF



CRIMINEL I.1: 'qui est relatif au méfait';

CRIMINEL I.2: 'qui est un méfait'

⁸ Nella versione originale l'immagine è tridimensionale, per questa ragione, in questa figura e nelle successive, alcune unità lessicali risultano sovrapposte.

Entrambi i casi mostrano chiaramente come la rete lessicale includa non solo le *lexies* che, in un modo o in un altro (tramite relazioni sintagmatiche o paradigmatiche, attanti sintattici profondi ecc.), sono collegate ad una *lexie* di partenza all'interno di una rete lessicale, ma altresì quelle di cui ci si può servire per ricostituire il significato della *lexie* stessa.

Per quanto riguarda RAPIDE, ritroviamo il sostantivo VITESSE il quale costituisce l'appoggio nominale nella parafrasi che corrisponde all'etichetta semantica; ritroviamo, inoltre, RAPIDITÉ, derivato morfologico di RAPIDE. Tuttavia, analogamente al caso di CRIMINEL/CRIME, è piuttosto inverosimile che un apprendente che non conosce il significato di RAPIDE possa conoscere il significato di RAPIDITÉ. A questo punto, nella fase di accesso all'unità lessicale tramite la parafrasi, la domanda che l'apprendente si porrà per trovare la *lexie* di cui necessita conterrà probabilmente la *lexie* VITESSE: *Quale unità lessicale esprime il significato 'qui a une certaine vitesse'?* Invece, se la *lexie* RAPIDITÉ è già conosciuta dall'apprendente, la domanda che si porrà – purché possieda un certo numero di nozioni metalinguistiche – è: *Qual è il derivato semantico aggettivale del sostantivo RAPIDITÉ?*

Il caso di CRIMINEL è leggermente diverso: l'aggettivo, infatti, è un derivato morfologico denominale a partire dalla radice CRIME. Basandosi esclusivamente sulla sua rete lessicale, l'apprendente potrebbe partire dalla *lexie* nominale iperonimica di CRIME, cioè MÉFAIT, che compare, in effetti, nella parafrasi che corrisponde all'etichetta semantica di questa stessa *lexie*:

- Qual è la *lexie* aggettivale che esprime il significato 'qui est relatif à un méfait' (o 'qui est un méfait')?

È tuttavia possibile che la *lexie* CRIME faccia già parte del bagaglio lessicale dell'apprendente; dandosi il caso, potrebbe porsi due tipi di domande purché abbia, a livello metalinguistico, delle conoscenze minime riguardanti le funzioni lessicali:

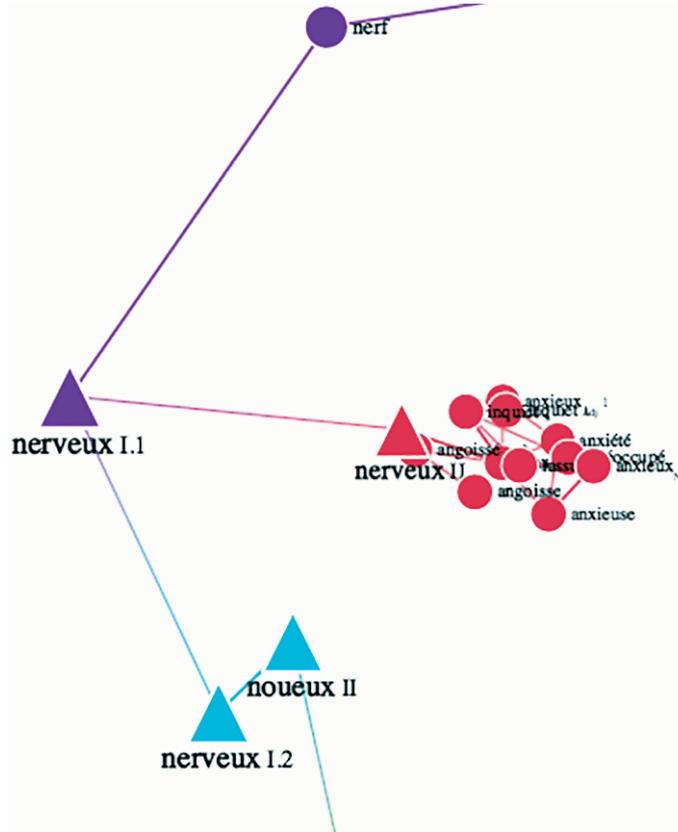
- Qual è l'aggettivo che ha lo stesso significato di CRIME?
- Qual è l'aggettivo che ha lo stesso significato di CRIME?

4.2 La polisemia

Il secondo aspetto di cui trattiamo è la polisemia. Come nel caso precedente, si tratterà di accedere a un'unità lessicale a partire dal significato. Ciò che rende specifico il caso della polisemia rispetto alla parafrasi è la necessità di considerare che un *vocable* possa essere suddiviso in diverse *lexies* che, come abbiamo visto nel § 3.2, condividono una o più componenti di significato. In questo caso verranno sfruttati sia l'approccio onomasiologico che l'approccio semasiologico.

Proponiamo, come esempio, la *lexie* aggettivale NERVEUX:

Figura 3 - La rete lessicale di NERVEUX nella banca dati RLF



NERVEUX I.1, che significa ‘qui est relatif aux nerfs’ (*système nerveux*) è collegato direttamente – come nel caso CRIMINEL/CRIME – alla radice di cui è un derivato morfologico denominale. L’assenza, nella rete lessicale, di un’unità lessicale iperonimica rispetto a NERF dipende dal fatto che ‘corps humain’, che potremmo ritrovare nell’etichetta semantica della parafrasi, non è stato trattato come entrata lessicale nella banca dati RLF. Il legame con NERF appare tuttavia piuttosto evidente nella rete, e il suo significato sufficientemente comprensibile per un apprendente che consulti questa rete lessicale specifica.

A partire da NERVEUX I.1 osserviamo una ramificazione che riguarda due legami semantici che presentano un’intersezione di significato: il primo muove in direzione di NERVEUX I.2 e l’altro verso NERVEUX II.

Il significato di NERVEUX I.2, ‘qui a des nervures’ (*des mains nerveuses; des bras nerveux*) può essere accomunato all’accezione I.1 tramite il legame di quasi sinonimia NERF/NERVURE, dato che entrambe le *lexies* significano ‘tendons, ligaments’.

Il significato di NERVEUX II, ‘qui manifeste un certain état’ (*Je me sentais nerveux*) è una trasposizione figurata legata all’immagine della tensione dei nervi come

manifestazione di una data emozione che potremmo collocare fra l'irritazione e il nervosismo.

È possibile didattizzare la polisemia a partire da una rete lessicale come quella di NERVEUX tramite due approcci diversi.

Il primo considera la polisemia un'intersezione di senso e procede in maniera onomasiologica: a partire dalle etichette semantiche, le riflessioni dell'apprendente potrebbero riguardare le componenti comuni di significato e le differenze semantiche sensibili fra una *lexie* e l'altra dello stesso *vocable*. L'apprendente si porrebbe le seguenti domande:

- Quale *lexie* del *vocable* NERVEUX corrisponde al significato 'qui est relatif aux nerfs'?
- Quale *lexie* del *vocable* NERVEUX corrisponde al significato 'qui a des nervures'?
- Quale *lexie* del *vocable* NERVEUX corrisponde al significato 'qui manifeste un certain état (lié aux nerfs)'?

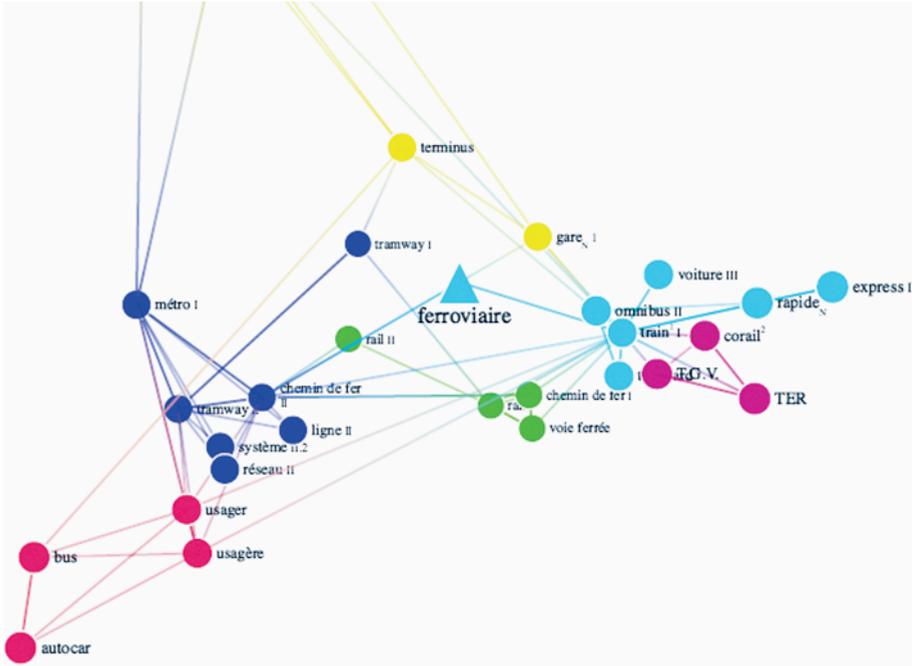
Le attività didattiche potrebbero così includere l'esplorazione dei legami e le divergenze fra unità minime di significato. Per esempio, si potrebbe far notare la lontananza o la prossimità delle diverse *lexies* dal supporto nominale del significato I.1, cioè NERF, oppure a partire dalla *lexie* semanticamente più lontana (NERVEUX II, poi NERVEUX I.2 per concludere con NERVEUX I.1), o, ancora, a partire dalla *lexie* semanticamente più vicina (nell'ordine, NERVEUX I.1, NERVEUX I.2, NERVEUX II).

Un'altra possibilità, più intuitiva, potrebbe riguardare la comprensione dei significati delle tre *lexies*, non tramite le rispettive etichette semantiche, ma esaminando le *lexies* che si trovano nelle immediate vicinanze. Questa attività consentirebbe di lavorare maggiormente con i sinonimi di NERVEUX I.2 e NERVEUX II (che sono collegati, rispettivamente, a NOUEUX II et a ANXIEUX) sfruttando anche la nozione di conversione pura per NERVEUX, a partire da NERF. Si potrebbe poi proseguire con la parte che abbiamo suggerito qui sopra, ovvero le intersezioni di significato con l'utilizzo (e le formulazioni) di etichette semantiche da collegare alle diverse *lexies* del *vocable*.

3.3 Le relazioni sintagmatiche e paradigmatiche

In quest'ultimo paragrafo presentiamo, in modo semplificato, i procedimenti che si trovano alla base dell'organizzazione mentale del lessico di un apprendente (Aitchison, 2003; Wolter, 2006) a partire dalla 'visualizzazione' delle relazioni che diverse *lexies* possono intrattenere. Queste relazioni partono da una *lexie* già nota all'apprendente; l'obiettivo consiste nel ritrovare la *lexie* di cui l'apprendente necessita per formulare un messaggio specifico. Forniremo alcuni esempi a partire dalla *lexie* aggettivale FERROVIAIRE, intorno alla quale si sviluppa una rete lessicale particolarmente densa riguardante i mezzi di trasporto su rotaia.

Figura 4 - La rete lessicale di FERROVIAIRE nella banca dati RLF



In questo caso non si tratta più di lavorare soltanto sul rapporto fra la *lexie* e il proprio significato, ma anche sui tipi di relazioni sintagmatiche e paradigmatiche controllate dalla *lexie* stessa. Tuttavia, per evidenziare il legame con le *lexies* più vicine a FERROVIAIRE, è opportuno prendere in considerazione l’etichetta semantica.

Come nei precedenti casi, la domanda che l’apprendente potrebbe porsi, relativamente al legame di prossimità fra FERROVIAIRE e CHEMIN DE FER, è la seguente:

- Qual è la *lexie* aggettivale che significa ‘qui est relatif aux chemins de fer’?

Oltre al legame con CHEMIN DE FER, notiamo, per la *lexie* FERROVIAIRE, due legami diretti con le seguenti *lexies*: GARE e TRAIN.

La relazione che intercorre fra GARE e FERROVIAIRE è di tipo sintagmatico: *gare ferroviaire* – che è distinta, ad esempio, da *gare routière* o da *gare maritime* – è una collocazione del francese. Sebbene la *lexie* più in uso in francese corrente per veicolare il significato ‘bâtiment établi à un point de la voie ferrée où s’arrêtent les trains pour embarquer et débarquer passagers et/ou marchandises’ sia GARE, l’esistenza e l’uso di *gare ferroviaire* ci impongono di presentare e sfruttare anche questa relazione sintagmatica.

La relazione fra TRAIN e FERROVIAIRE è di tipo paradigmatico: è possibile parafrasare l’aggettivo FERROVIAIRE – tramite una parafrasi estremamente semplificata – con ‘qui a affaire aux trains’.

La rete lessicale di FERROVIAIRE si presta inoltre ad essere utilizzata per le relazioni che altre *lexies* tessono fra di loro, anche non considerando il legame diret-

to con FERROVIAIRE. Per esempio, ci si potrebbe concentrare sulla relazione fra GARE e TRAIN, la cui spiegazione richiede l'esplicitazione della struttura attanziale di GARE tramite la definizione. Come abbiamo ricordato nel § 4.1, la LEC, nel confezionare le proprie definizioni lessicografiche, non si limita ad enumerare le unità minime di significato, ma prevede anche l'esplicitazione della struttura attanziale del *definiendum*. In base a questi presupposti, la definizione di GARE sarebbe: 'Bâtiment de l'agglomération (Y) pour les trains (X)'. La *gare ferroviaire* è effettivamente un luogo in cui transitano i treni; per questo, TRAIN è il sostantivo che indica, tipicamente, il primo attante semantico di GARE.

La relazione fra TRAIN e GARE può essere sfruttata in due direzioni.

Da un lato, l'apprendente può porsi la seguente domanda per trovare la *lexie* TRAIN:

– Qual è il nome del mezzo di trasporto che, tipicamente, si trova in una stazione?

Dall'altra, a partire dalla *lexie* TRAIN – quindi dal primo attante semantico di GARE –, la domanda potrebbe essere formulata nel modo seguente:

– Qual è il nome del luogo dove si possono trovare, tipicamente, dei treni?

Altre attività possono essere elaborate a partire dalle *lexies* che si trovano nella rete lessicale. Per esempio, a partire da TRAIN, sarebbe possibile lavorare sui tre principali tipi di treni che circolano in Francia (TGV, CORAIL, TER), anche se ciò richiederebbe un livello di competenza culturale più elevato; oppure, a partire da CHEMIN DE FER, sulle tipologie di mezzi di trasporto che possono circolare, come i treni, su rotaia: TRAMWAY et MÉTRO.

Relazioni di sinonimia, inoltre, collegano la *lexie* CHEMIN DE FER ad altre *lexies*: VOIE FERRÉE et RAIL.

I pochi esempi di utilizzo della rete lessicale che abbiamo illustrato (dal punto di vista delle relazioni sintagmatiche e paradigmatiche) non sono esaustivi né rappresentativi delle potenzialità pedagogiche della banca dati RLF. Teniamo a sottolineare un aspetto innovativo di questa configurazione: nonostante le reti lessicali che mostriamo siano ancora provvisorie, il diverso grado di prossimità di una *lexie* rispetto ad un'altra *lexie* che costituisce il punto di partenza dell'analisi della rete, consente, in maniera intuitiva, di dedurre le relazioni fra *lexies* nonché di esplicitarle. La visione globale dei gruppi di *lexies* più o meno prossimi rispetto ad una *lexie* di partenza, ben più evidente rispetto a quanto ci possa offrire la lettura di un dizionario tradizionale, costituisce uno dei grandi vantaggi di un progetto lessicografico come quello che abbiamo presentato, particolarmente adatto all'utilizzo nell'insegnamento delle lingue.

5. Conclusioni

A conclusione del nostro lavoro, vorremmo ritornare su alcuni vantaggi della banca dati RLF in quanto risorsa lessicale per l'insegnamento e l'apprendimento di una

lingua; ci pare opportuno presentare anche alcuni suggerimenti volti a renderla quanto più efficace possibile.

In primo luogo, la presentazione delle reti lessicali associate agli aggettivi RAPIDE, NERVEUX, CRIMINEL e FERROVIAIRE ci ha permesso di mettere in luce i vantaggi che una banca dati di tipo 'rete lessicale' presenta rispetto a un dizionario tradizionale. Sebbene la maggior parte dei dizionari espliciti alcune relazioni semantiche fra *lexies* – pensiamo, ad esempio, ai sinonimi e agli antonimi – e taluni abbiano fatto del loro carattere 'analogico' il loro punto di forza (ci riferiamo, nella fattispecie, al *Grand Robert* e al *Petit Robert*), il carattere lineare del dizionario costituisce spesso un ostacolo ad una visione globale di tutte le relazioni. Solo una rappresentazione come quella che abbiamo esposto può fornire un'idea della complessità delle relazioni che esistono fra *lexies*; inoltre, solo una rappresentazione di questo tipo permette un accesso immediato e intuitivo alle *lexies* ai fini dell'apprendimento del lessico.

Tale rappresentazione è alla portata di insegnanti ed apprendenti L1 o L2, per svariate attività, come ad esempio la preparazione di attività che riguardano, nello specifico, determinate unità lessicali (ad esempio il lessico dei trasporti, del corpo umano, degli animali ecc.), oppure per capire meglio l'organizzazione del lessico e strutturare in maniera ottimale, in prospettiva più ampia e articolata, l'apprendimento del lessico.

Inoltre, una conoscenza più approfondita dei fenomeni lessicali e delle nozioni che permettono di descriverli, così come un utilizzo migliore delle risorse di tipo 'rete lessicale', offrirebbero agli apprendenti la possibilità di diagnosticare meglio i problemi ai quali devono far fronte nelle attività di produzione scritta o di comprensione scritta.

Per quanto riguarda la necessità di formare gli utilizzatori in merito a nozioni lessicali fondamentali, ci preme sottolineare, nel caso della banca dati RLF, l'esigenza di conoscere un certo numero di 'convenzioni' lessicografiche. In questo senso, la banca dati RLF dovrebbe essere accompagnata da chiare istruzioni ad uso di coloro che intendano consultarla. Poiché, per il momento, la modalità di inserimento e strutturazione dei dati lessicografici è pensata per facilitare i lessicografi che lavorano all'implementazione della banca dati, in fase di messa in rete della banca dati RLF, e in vista della sua fruizione da parte del grande pubblico, la ricerca nel settore dovrà concentrarsi sull'accessibilità e sulla consultazione di tale strumento lessicografico in modo da renderlo quanto più 'user-friendly' possibile.

Ringraziamenti

Ringraziamo M.F. Bonadonna e M. Lisi per la loro paziente rilettura del presente articolo e F. Bianchi e P. Leone per i loro preziosi suggerimenti.

Bibliografia

- AITCHISON J. (2012), *Words in the Mind: An Introduction to the Mental Lexicon* (4th ed), Wiley-Blackwell, Chichester, West Sussex/Malden, MA.
- ANCTIL D. (2011), L'erreur lexicale au secondaire: analyse d'erreurs lexicales d'élèves de 3^e secondaire et description du rapport à l'erreur lexicale d'enseignants de français, Dissertation doctorale, Département de didactique de la Faculté des sciences de l'éducation, Université de Montréal.
- ATKINS B.T.S. (ed.) (1998), *Using Dictionaries: Studies of Dictionary Use by Language Learners and Translators*, Max Niemeyer Verlag, Tübingen.
- AUGUST D. - CARLO M. - DRESSLER C. - SNOW C. (2005), The Critical Role of Vocabulary Development for English Language Learners, in *Learning Disabilities Research and Practice* 20: 50-57.
- BARQUE L. - POLGUÈRE A. (2013), Enrichissement des définitions du Trésor de la langue française informatisé dans une perspective lexicographique, in *Lexique* 21: 221-244.
- BAKER C.F. - FILLMORE C.J. - CRONIN B. (2003), The Structure of the FrameNet Database, in *International Journal of Lexicography* 16(3): 281-296.
- BEECH J.R. (2004), Using a Dictionary: Its Influence on Children's Reading, Spelling, and Phonology, in *Reading Psychology* 25: 19-36.
- BIEMILLER A. (2011), Vocabulary Development and Implications for Reading Problems, in MCGILL-FRANZEN M. - ALLINGTON R.L. (eds.), *Handbook of Reading Disability Research*, Routledge, New York: 208-218.
- BISHOP G. (2000), Developing Learner Strategies in the Use of Dictionaries as a Productive Language Learning Tool, in *The Language Learning Journal* 22(1): 58-62.
- CARSTENS A. (1995), Language Teaching and Dictionary Use: An Overview, in *Lexikos* 5: 105-116.
- COWIE A.P. (1999), *English Dictionaries for Foreign Learners. A History*, Clarendon Press, Oxford.
- DOLEZAL F.T. - MCCREARY D.R. (1999), *Pedagogical Lexicography Today: A Critical Bibliography on Learners' Dictionaries with Special Emphasis on Language Learners and Dictionary Users*, Lexicographica Series Maior 96, Max Niemeyer Verlag, Tübingen.
- FELLBAUM C. (1998), *WordNet: An Electronic Lexical Database*, The MIT Press, Cambridge MA.
- FRASSI P. (2012), Des définitions pas comme les autres: les définitions d'adjectifs, in NEVEU F. - MONI TOKE V. - BLUMENTHAL P. - KLINGER T. - LIGAS P. - PRÉVOST S. - TESTON-BONNARD S. (éds.), *3^e Congrès Mondial de Linguistique Française*, EDP Sciences, Paris: 891-907.
- FRASSI P. (2014), La définition lexicographique des dérivés sémantiques adjectivaux, in NEVEU F. - BLUMENTHAL P. - HRIBA L. - GERSTENBERG A. - MEINSCHAEFER J. - PRÉVOST S. (éds.), *4^e Congrès Mondial de Linguistique Française*, EDP Sciences, Paris: 741-758.
- GIASSON J. (2011), *La lecture. Apprentissages et difficultés*, Gaëtan Morin Éditeur, Boucherville.

- KIEFFER M.J. - LESAUX N.K. (2012), Knowledge of Words, Knowledge about Words: Dimensions of Vocabulary in First and Second Language Learners in Sixth Grade, in *Reading and Writing* 25(2): 347-373.
- LEW R. (2013), Online Dictionary Skills, in IZTOK K. - KALLAS J. - GANTAR P. - KREK S. - LANGEMETS M. - TUULIK M. (éds.), *Electronic Lexicography in the 21st Century: Thinking Outside the Paper*, Trojina, Institute for Applied Slovene Studies/Eesti Keele Instituut: 16-31.
- LUX-POGODALLA V. - POLGUÈRE A. (2011), Construction of a French Lexical Network: Methodological Issues, in SAGOT B. (éd.), *Proceedings of the International Workshop on Lexical Resources*: 54-61. http://alpage.inria.fr/~sagot/pub/WoLeR_2011_proceedings.pdf
- MEL'ČUK I. (1973), Towards a Linguistic 'Meaning \Leftrightarrow Text' Model, in *Trend in Soviet Theoretical Linguistics* 18: 33-57.
- MEL'ČUK I. - CLAS A. - POLGUÈRE A. (1995), *Introduction à la lexicologie explicative et combinatoire*, Duculot, Paris.
- MEL'ČUK I. (1996), Lexical Functions: A Tool for the Description of Lexical Relations in a Lexicon, in WANNER L. (éd.), *Lexical Functions in Lexicography and Natural Language Processing*, Benjamins, Amsterdam/Philadelphia: 37-102.
- NESI H. (1999), The Specification of Dictionary Reference Skills in Higher Education, in HARTMANN R.R.K. (ed.), *Dictionaries in Language Learning. Recommendations, National Reports and Thematic Reports from the Thematic Network Project in the Area of Languages, Sub-Project 9: Dictionaries*, Freie Universität Berlin, Berlin: 53-67.
- NESI H. - HALL R. (2002), A Study of Dictionary Use by International Students at a British University, in *International Journal of Lexicography* 15(4): 277-306.
- NEUBACH A. - COHEN A.D. (1988), Processing Strategies and Problems Encountered in the Use of Dictionaries, in *Dictionaries: Journal of the Dictionary Society of North America* 10: 1-19.
- POLGUÈRE A. (2011), Perspective épistémologique sur l'approche linguistique Sens-Texte, in *Mémoires de la Société de Linguistique de Paris* XX: 79-114.
- POLGUÈRE A. (2014), From Writing Dictionaries to Weaving Lexical Networks, in *International Journal of Lexicography* 27(4): 396-418.
- POLGUÈRE A. - TREMBLAY O. (2003), Qu'y a-t-il à l'intérieur de NOIX? Ou comment décortiquer les unités lexicales, in *La Lettre de l'AIRDF* 33(3): 27-30.
- RUPPENHOFER J. - ELLSWORTH M. - PETRUCK, M.R.L. - JOHNSON C.R. - SCHEFFCZYK J. (2010), *FrameNet II: Extended Theory and Practice*, International Computer Science Institute, Berkeley, CA.
- SCHMITT N. (2000), *Vocabulary in Language Teaching*, Cambridge University Press, New York.
- SCHMITT N. (2010), *Researching Vocabulary: A Vocabulary Research Manual*, Palgrave Press, Basingstoke.
- SCOTT J. - NAGY B. - FLINSPACH S. (2008), More than Merely Words: Redefining Vocabulary Learning in a Culturally and Linguistically Diverse Society, in FARSTRUP A.

- SAMUELS J. (eds.), *What Research Has to Say About Vocabulary Instruction*, International Reading Association, Newark, DE: 182-210.

TREMBLAY O. (2009), Une ontologie des savoirs lexicologiques pour l'élaboration d'un module de cours en didactique du lexique. Thèse de doctorat non publiée, Université de Montréal, Montréal.

VERHALLEN M. - SCHOONEN R. (1993), Lexical Knowledge of Monolingual and Bilingual Children, in *Applied Linguistics* 14(4): 344-363.

WOLTER B. (2006). Lexical Network Structures and L2 Vocabulary Acquisition: The Role of L1 Lexical/Conceptual Knowledge, in *Applied Linguistics* 27(4): 741-747.

VALERIA CARUSO¹

Dizionari elettronici e apprendimento delle espressioni idiomatiche: monitoraggio dei bisogni e prospettive future

This study investigates the lexicographical needs of a group of L2 Italian learners performing a production task using idioms. In order to achieve this aim, the efficacy of an electronic monolingual dictionary frequently used by foreign students was put to the test. Results show that, on the whole, the given dictionary provides inadequate support; however some of its features offer important cues to the target users. These include the clear statement of topics and comments in the idiom definition, and the etymological explanation of how literal meanings evolved into figurative ones. The results of this study will be used for planning an electronic pedagogical dictionary that meets the requirements demanded by *lexicographical function theory*.

1. *Dizionari elettronici e apprendimento delle espressioni idiomatiche*

Questo studio si propone di avviare una riflessione sul trattamento lessicografico delle espressioni idiomatiche, con l'intento di pianificare un dizionario italiano dei modi di dire per apprendenti stranieri.

I modi di dire sono espressioni multiparola “bloccat[e], sia per quanto riguarda la sostituibilità dei membri [...], sia per quanto riguarda la possibilità di modificarli” (Ježek, 2005: 198) e rappresentano gli elementi più complessi del lessico di una lingua, dal momento che sono formati da più parole e sono soggetti a restrizioni morfosintattiche collegate con quelle semantiche e pragmatiche (Burger, 2010; Cowie, 1981; Dobrovol'skij, 2001; Fillmore *et al.*, 1988).

Come i proverbi, inoltre, le espressioni idiomatiche cristallizzano prassi e consuetudini sociali (Aprile, 2005: 45-47) e sono depositarie di un sapere popolare che viene spesso rielaborato in diverse forme di comunicazione pubblica, dalla pubblicità alle conversazioni tra pari, come sottolinea il *Common European Framework of Reference for Languages* (Council of Europe, 2007: 120). I modi di dire sono quindi una componente essenziale della competenza socioculturale dei parlanti di una lingua, e la comprensione delle principali espressioni idiomatiche nei testi scritti rientra tra le competenze attese già per il livello B2 del *Quadro* (Council of Europe, 2007: 69). Studi condotti in ambito psicolinguistico hanno dimostrato che i parlanti nativi acquisiscono una competenza specifica per queste unità lessicali solo nella tarda infanzia, di pari passo con lo sviluppo delle abilità metalinguistiche e della

¹ Università degli Studi di Napoli 'L'Orientale'.

capacità di ricavare inferenze dal contesto (Benneli *et al.*, 2006; Levorato - Cacciari, 2002).

2. *Inquadramento della ricerca e sue finalità*

Per identificare i tipi di difficoltà riscontrabili nell'autoapprendimento delle espressioni idiomatiche italiane, vengono presentati i risultati di uno studio pilota condotto su 18 apprendenti di italiano L2 di livello intermedio-avanzato (B2 e C1), parlanti nativi di una lingua isolante (cinese e vietnamita), finalizzato a monitorare l'efficacia informativa di un dizionario elettronico monolingue che i partecipanti hanno dichiarato di consultare abitualmente: il *Treccani* online [TREACCANI]. Di recente (Fuertes-Olivera - Tarp, 2014: 94) è stata infatti sottolineata l'importanza delle indagini sulle fasi di 'post-consultazione' dei dizionari, finalizzate a verificare se le informazioni in essi contenute consentono agli utenti di svolgere efficacemente un compito. Nel caso del test che viene qui presentato è stato possibile, ad esempio, identificare alcune caratteristiche lessicografiche importanti per supportare parlanti non-nativi nell'autoapprendimento delle espressioni idiomatiche. In futuro, le caratteristiche evidenziate dal presente studio serviranno ad integrare i campi informativi di un dizionario elettronico dei fraseologismi per renderlo più adeguato all'apprendimento.

In tal senso, va ricordato che la riflessione metalessicografica sui dizionari elettronici ha mutuato dall'informatica due parametri essenziali per stimare la qualità di un dizionario: l'efficacia e l'efficienza. Come spiega Heid (2011), un dizionario elettronico, o un software, è efficace se consente agli utenti di trovare i dati di loro interesse senza sovraccargarli con informazioni inutili; è efficiente, invece, se garantisce tempi rapidi per lo svolgimento della ricerca: "The quicker the user gets access to the data needed the more efficient is the software" (Heid, 2011: 289).

2.1 Il test e i partecipanti

Per verificare essenzialmente il primo di questi parametri, ovvero l'efficacia, lo studio che è stato condotto si è soffermato sulle capacità dei partecipanti di ricavare informazioni da un dizionario. La ricerca è stata suddivisa in due fasi, svoltesi a distanza di una settimana l'una dall'altra: nella prima, gli studenti dovevano estrarre una lista dei modi di dire relativi ai lemmi "mano" e "testa" contenuti nel dizionario *Treccani*, un compito non banale data la grande quantità di dati presentati in queste voci con un *layout* poco riadattato per la versione online del dizionario². Gli studenti dovevano inoltre spiegare brevemente, per iscritto, il significato di ciascun *idiom*,

² Al contrario, i dizionari elettronici più innovativi si distinguono anche per l'organizzazione strutturale delle loro entrate, le quali consentono una rapida identificazione dei dati in esse contenuti, usando un'accurata distribuzione del testo (*text blocks*) e degli appositi 'indicatori strutturali' (*structural indicators*, Wiegand, 1989). Si veda al riguardo l'attenta analisi di Gouws (2014).

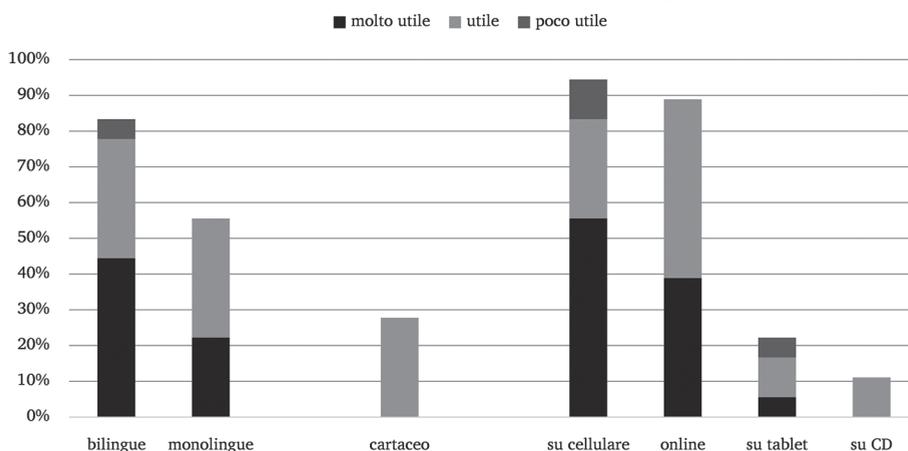
in modo da concentrarsi sulla sua semantica ed iniziare a familiarizzare con queste espressioni.

La seconda fase dello studio consisteva invece in un compito di riempimento basato su testi semi-autentici ricavati dalla stampa. Gli studenti dovevano scegliere il fraseologismo mancante all'interno di una selezione di 15 *idiom* tra quelli precedentemente estratti dal *Treccani*, che sono stati presentati loro sotto forma di un elenco corredato dalle glosse del dizionario. Durante lo svolgimento della prova, gli studenti potevano quindi leggere tutti i dati presenti nel *Treccani* pur non avendo accesso diretto al dizionario stesso.

Hanno svolto il test 18 parlanti cinesi e vietnamiti, in Italia da sei mesi come studenti di scambio presso l'Università degli Studi di Napoli 'L'Orientale'. La maggior parte (55%) studiava l'italiano da tre anni, un altro gruppo (28%) da due ed un terzo (17%) da quattro anni. Otto di loro avevano un livello di conoscenza dell'italiano pari al B2 del QCER (*Quadro Comune Europeo di Riferimento per le lingue*), i restanti dieci erano di livello C1.

Prima di svolgere il test, sono state rivolte ai partecipanti alcune domande sulle loro abitudini di consultazione dei dizionari (fig. 1). L'indagine ha rivelato un uso estremamente limitato dei dizionari cartacei, solo il 30% dei partecipanti ha dichiarato infatti di adoperarli; quasi tutti, invece, hanno detto di usare un dizionario elettronico sul loro cellulare (95%) oppure online (89%). Inoltre, per lo studio delle lingue straniere, i partecipanti hanno dichiarato di preferire le risorse bilingue (83%) rispetto a quelle monolingue (55%) e, tra queste, il dizionario più usato è il *Treccani* online (95%).

Figura 1 - *Abitudini di consultazione lessicografica dei partecipanti*



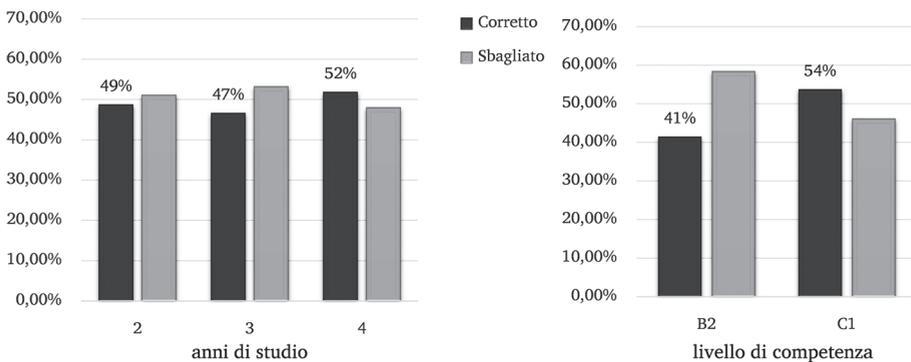
3. Risultati

Nel presentare gli esiti del test, dopo aver fornito i risultati generali, la discussione si soffermerà sulle tipologie di difficoltà riscontrate dai partecipanti, desumibili per

lo più da una stima dei tipi di errori che sono stati commessi. Successivamente verranno invece valutate le carenze della descrizione lessicografica, facendo riferimento all'analisi degli errori che sono stati commessi sui singoli fraseologismi e ai tipi di glosse offerte dal dizionario. Il test ha evidenziato infatti che alcune espressioni idiomatiche non hanno posto quasi nessuna difficoltà ai partecipanti, mentre altre sono state utilizzate quasi sempre scorrettamente. Per cercare evidenze di eventuali caratteristiche lessicografiche che potessero essere correlate con i successi e gli insuccessi al test proposto, è stata condotta un'analisi approfondita delle tipologie e modalità di definizione offerte dal *Treccani*.

I risultati generali del test di completamento somministrato evidenziano come il *Treccani* online non riesca a supportare adeguatamente gli apprendenti in un compito produttivo. Gli item svolti in maniera corretta ammontano infatti a poco più della metà del totale (56%) e gli anni di studio e il livello di competenza dei partecipanti non determinano un incremento statisticamente significativo delle risposte esatte fornite (figg. 2 e 3).

Figure 2-3 - Percentuali di risposte corrette per livello di competenza linguistica e anni di studio



Le differenze percentuali di risposte positive date dagli studenti di livello B2 (41% di risposte esatte in media) e C1 (54%) risulta infatti non significativa al t-test ($p > 0,05$)³. Analogamente, sono tutte negative al t-test ($p > 0,5$) le differenze percentuali di risposte corrette date, rispettivamente, da chi studia l'italiano da due anni (49% di risposte esatte in media) rispetto a chi lo studia da tre (47%) e da quattro (52%) anni; e da chi lo studia da tre rispetto a chi lo studia da quattro anni.

3.1 Problematiche dell'apprendimento delle espressioni idiomatiche italiane

Analizzando più nel dettaglio i risultati del test, va sottolineato come, per i partecipanti, comprendere i fraseologismi italiani presenti lo stesso grado di difficoltà del saperli usare, dal momento che gli errori commessi si suddividono quasi equamente

³ Si ricorda che il t-test consente di stabilire se il valore medio di una distribuzione tra due set di dati è significativo.

tra quelli lessicali e morfosintattici: i primi sono più legati a problemi di comprensione, i secondi più a quelli di produzione (tab. 1).

Tabella 1 - *Tipi di errori commessi nel test: lessicali e morfosintattici*

<i>Tipo di errore</i>	<i>Percentuale</i>
lessicale	56%
non coerente/parzialmente coerente	51%
modifica morfosintattica dei costituenti	5%
morfosintattico	43%
accordo	21%
modo	12%
tempo	4%
persona	4%
Altro (verbo reggente, ausiliare)	2%
pragmatico	1%

Per quantificare correttamente i tipi di difficoltà riscontrati dai partecipanti, è stato infatti necessario scorporare, tra gli errori lessicali, quelli che dipendono da variazioni morfosintattiche non ammesse dalla fissità strutturale dell'*idiom* e, pertanto, dovrebbero essere evidenziati con particolare enfasi in un dizionario pedagogico per apprendenti stranieri. Questo impedirebbe normalizzazioni che vengono di fatto bloccate dalla cristallizzazione del fraseologismo nel lessico della lingua (cfr. Fillmore *et al.*, 1988), come mostrano le variazioni nell'ordine dei costituenti tratte dalle produzioni degli studenti, *fa di *sua* testa o *facciamo *della nostra* testa, oppure i cambiamenti di numero relativi al 'termine chiave' (Dobrovól'skij - Pirainen, 2005), *hanno *la mano legata* o *a portata di *mani*.

Sono squisitamente lessicali, invece, gli errori che inficiano la coerenza testuale perché la semantica del fraseologismo non è adeguata al contesto, come nel caso seguente:

- (1) Quando noi chiamiamo la polizia di notte per segnalare rumori e situazioni pericolose ci rispondono che non possono intervenire. Perché non c'è abbastanza personale e non ce la fanno a far tutto. E i poliziotti stessi sono esasperati e ci dicono che **tagliano la testa al toro*.

Tra gli errori prettamente morfosintattici, invece, i più numerosi sono quelli di accordo (tab. 1), che si dividono a loro volta tra errori di accordo contestuale (71%) e di accordo tra i costituenti dell'*idiom* stesso (29%). Il punto 2 di seguito mostra un errore di accordo contestuale, mentre il punto 3 un errore di accordo tra i costituenti:

- (2) A Milano riaprono le scuole ma molti bambini non potranno andare in un asilo pubblico. Le scuole materne a Milano, infatti, non bastano e il Comune non ha i soldi per aprirne delle nuove. "È urgente che lo Stato ci **si mette le mani nei capelli*" spiega il vicesindaco. Milano è già la città italiana dove c'è il

maggior numero di scuole dell'infanzia comunali, ma queste non sono ancora sufficienti.

- (3) Ai giornalisti che chiedono se la riduzione delle tasse provocherà un aumento del debito pubblico, il Ministro dell'Economia risponde così: "Non dobbiamo *fasciarci la testa prima di rompersela, ma valutare cosa fare man mano che i problemi ci si presentano, queste sono le mie intenzioni".

In letteratura è stato peraltro evidenziato come la scarsa illustrazione, da parte dei dizionari, dei comportamenti morfosintattici dei modi di dire ne pregiudichi sia il corretto uso sia la comprensione (Burger, 2010; Dobrovolskij, 2009). Nel test somministrato, ad esempio, l'inserimento dell'espressione idiomatica *dare una mano* (cfr. punto 4 dell'elenco) in una subordinata al congiuntivo retta da un verbo impersonale, come *bisognare*, potrebbe aver compromesso la capacità dei partecipanti di selezionare l'espressione idiomatica corretta, causando i molti errori lessicali che sono stati commessi per questo specifico fraseologismo (cfr. fig. 4 più avanti):

- (4) NUOVE SCUOLE MATERNE IN ARRIVO. FIRMATO L'ACCORDO CON LO STATO
A Milano riaprono le scuole ma molti bambini non potranno andare in un asilo pubblico. Le scuole materne di Milano, infatti, non bastano e il Comune non ha i soldi per aprirne delle nuove: "Bisogna che lo Stato ci
– spiega il vicesindaco. Milano è già la città italiana dove c'è il maggior numero di scuole dell'infanzia comunali, ma queste non sono ancora sufficienti".

Sarebbe quindi importante che il dizionario offrisse delle note sulle maggiori criticità di accordo contestuale, con una particolare enfasi sulla flessione in frasi subordinate, sfruttando anche le soluzioni tecniche disponibili in ambito elettronico per rendere la loro presentazione efficace ed efficiente allo stesso tempo (si veda Heid, 2014).

3.2 Problematiche riconducibili alla descrizione lessicografica

Accanto ai dati generali sulle diverse tipologie di errori riscontrate, risulta particolarmente macroscopica la disparità dei risultati ottenuti per i singoli *idiom* testati (fig. 4): alcuni vengono sbagliati quasi sempre (come *dare una mano*); altri quasi mai (come *alzare le mani su qualcuno*).

Si è cercato quindi di capire se il tipo di spiegazione del dizionario o le modalità con cui essa viene fornita agli utenti influiscono sui risultati del test. A questo scopo è stato stilato un elenco delle caratteristiche rilevabili nelle glosse del *Treccani* (fig. 5), ovvero: la presenza di esempi, la descrizione della componente pragmatica dell'*idiom*, l'uso di un lessico controllato, il tipo di definizione fornita e le modalità con cui essa viene formulata. La presenza di ciascuna di queste componenti nelle definizioni analizzate può essere apprezzata riferendosi al grafico della figura 5, mentre verranno discussi nel dettaglio, in paragrafi ad essi dedicati, solo i parametri correlati con i risultati statisticamente significativi che i partecipanti hanno ottenuto al test.

Figura 4 - Percentuali di errori lessicali, morfosintattici e pragmatici per singoli idiom

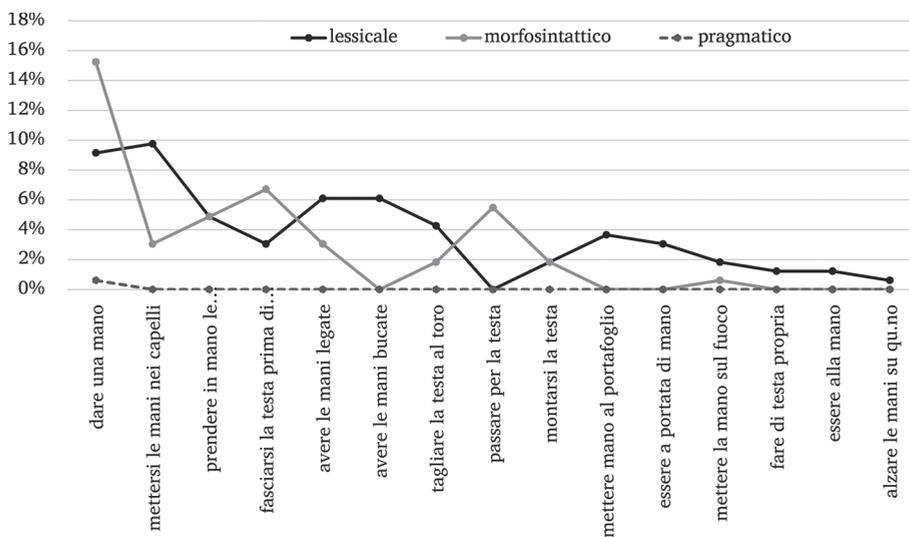
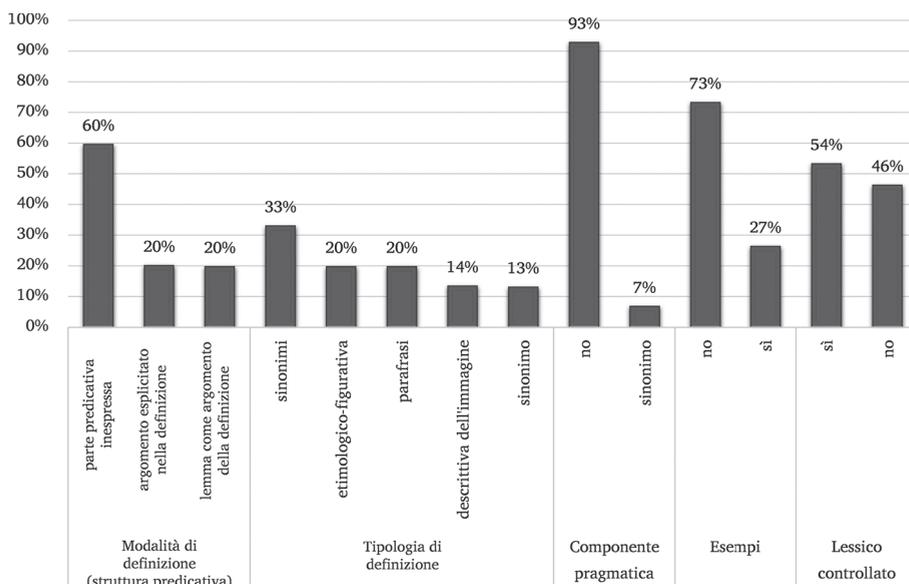


Figura 5 - Caratteristiche lessicografiche rilevate nelle definizioni del Treccani

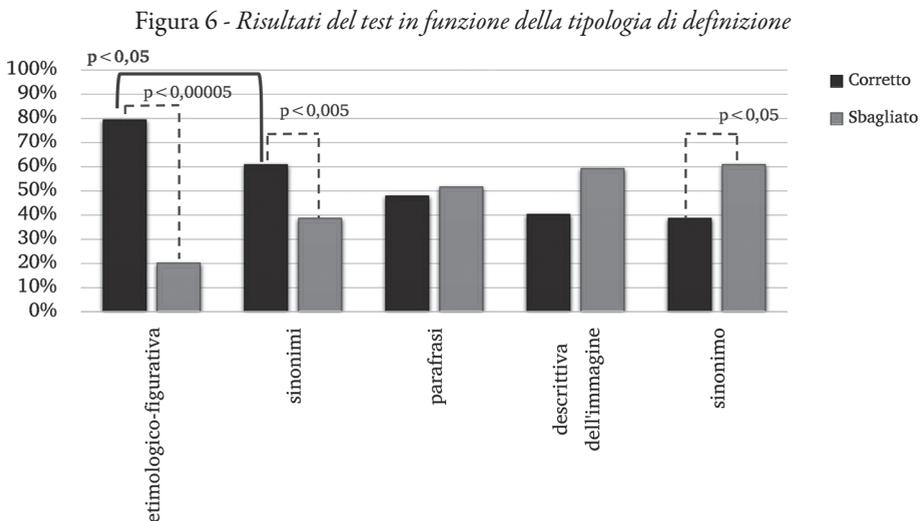


Si precisa inoltre che, per il corretto rilevamento delle caratteristiche delle definizioni, ci si è riferiti all'esaustiva tassonomia di Wiegand, che nel corso di circa trent'anni ha elaborato una dettagliata teoria delle componenti strutturali dei dizionari ("theory of dictionary form", per una sintesi si veda Wiegand, 1989; Wiegand - Smit, 2013). Dal punto di vista metodologico, invece, lo studio monitora le informazioni che un gruppo di partecipanti riesce a ricavare da un dizionario per risolvere un compito specifico. Tale ricerca rientra in un filone di studi sull'uso dei dizionari

(*users' research*) che Fuertes-Olivera e Tarp (2014, 92-96) hanno denominato “post-consultation studies”, per differenziarlo dalle indagini che riguardano sia la fase della consultazione (“intra-lexicographical consultation phase”), sia quella che la precede (“pre-consultation phase”). Quest’ultimo tipo d’indagini dovrebbero aiutare ad identificare i motivi per cui gli utenti decidono di rivolgersi a un dizionario, le altre, invece, monitorano il comportamento degli utenti mentre sono impegnati in una ricerca lessicografica.

3.2.1 Tipologie di definizioni

L’analisi statistica condotta con il t-test ha fatto emergere due strategie definitorie efficaci e una decisamente inadeguata a descrivere la semantica degli *idiom* a degli apprendenti stranieri (fig. 6).



Una descrizione di tipo etimologico-figurativo risulta la più idonea e registra l’80% di risposte corrette⁴. Si consideri ad esempio la glossa di *essere alla mano* fornita dal *Treccani*: “di cosa che è vicina, comoda, pronta (*con le prove alla m.*), e anche di persona affabile, cortese, facile a essere avvicinata”. In questa definizione viene descritta l’immagine contenuta nel modo di dire e, allo stesso tempo, viene suggerito come il significato idiomatico sia derivato da quello letterale, fornendo una sorta di breve etimologia. Meno efficaci, ma correlate comunque con risultati significativamente positivi al test (61%)⁵, sono le definizioni offerte attraverso l’uso di più sinonimi, come quella di *alzare le mani su qualcuno*: “disporsi a percuoterlo, a schiaffeggiarlo o sim. (anche, *picchiarlo senz’altro*)”. Al contrario, quando la glossa consiste in un

⁴ La differenza tra le percentuali di risposte positive e quelle negative riportate da tutti i partecipanti per questo tipo di definizioni ha peraltro un’altissima significatività statistica al t-test ($p < 0,00005$).

⁵ Anche in questo caso le differenze percentuali tra le risposte corrette e quelle sbagliate è risultata significativa al t-test ($p < 0,005$).

unico sinonimo del modo di dire, ad esempio *montarsi la testa*: “illudere, illudersi”, gli apprendenti non sembrano in grado di ricavarne informazioni utili e danno solo il 39% di risposte esatte⁶.

L’analisi statistica del t-test mostra invece che le differenze percentuali tra le risposte corrette e quelle sbagliate ottenute per le altre due modalità di definizione non sono significative, ovvero le definizioni basate su una parafrasi del significato idiomatico o, al contrario, quelle che descrivono esclusivamente l’immagine soggettiva non sono correlabili né con esiti positivi né negativi al test. Un esempio del primo tipo di definizione è quella relativa a *fasciarsi la testa prima di rompersela*, ovvero: “non è opportuno affliggersi d’un male, o cercarvi rimedio, prima ch’esso sia sopravvenuto”. Per l’altra tipologia può essere citato *passare per la testa*, glossato nel dizionario come “affacciarsi alla mente”.

Inoltre, comparando tra loro le percentuali di risposte corrette, il t-test rivela che le definizioni di tipo etimologico-figurativo ottengono risultati significativamente migliori rispetto alle definizioni che usano vari sinonimi ($p < 0,05$, fig. 6), confermando l’assoluta preferibilità di questo genere di commento semantico⁷. Questo dato è peraltro in linea con gli esiti di altri studi sull’apprendimento delle espressioni idiomatiche. Steinel, Hulstijn e Steinel (2007: 455) ad esempio hanno dimostrato, attraverso un esperimento di PAL (Paired Associate Learning⁸), che la spiegazione etimologica facilita la creazione di un’immagine mentale dell’*idiom* da parte di apprendenti stranieri, favorendo la memorizzazione a lungo termine del fraseologismo. Secondo gli autori, infatti, l’immagine mentale costituisce una seconda modalità di *encoding* dell’unità lessicale, garantendo una doppia tipologia di memorizzazione.

Più dati convergono quindi nell’indicare le caratteristiche che vanno specificate all’interno di una definizione per apprendenti stranieri: essa dovrebbe contenere una descrizione sia del significato letterale, sia di quello metaforico dell’*idiom* ed offrire allo stesso tempo una spiegazione etimologica di questo cambiamento semantico.

3.2.2 Modalità di formulazione della definizione

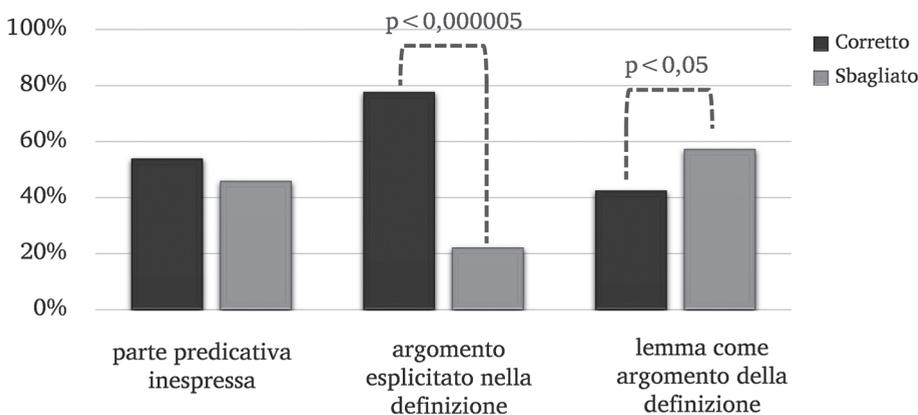
Altri risultati del test dimostrano quanto anche la modalità con cui viene formulata la definizione, con riferimento alla sua struttura predicativa, sia rilevante (fig. 7).

⁶ La differenza tra le risposte corrette e quelle sbagliate per questo tipo di definizione è risultata significativa al t-test ($p < 0,05$).

⁷ Si fa riferimento alla bipartizione strutturale introdotta da Wiegand (1989a) per gli articoli dei dizionari (*Wörterbuchartikel/dictionary article*), che si dividono in un commento sulla forma (*Formkommentar*) e in un commento semantico (*Sematischerkommentar*).

⁸ Particolarmente usati in campo medico, gli esperimenti di *Paired Associate Learning* consentono di stimare le capacità mnemoniche dei soggetti testati. Essendo basati su sessioni di apprendimento a cui segue un test di verifica, tra gli anni ’50 e ’60 sono stati usati nell’ambito della linguistica acquisizionale per valutare l’apprendimento del lessico di una lingua straniera. Più di recente, a partire dai primi anni del 2000, questa metodologia d’indagine empirica è stata impiegata per stimare gli effetti della direzione (dalla L1 alla L2 o viceversa) con cui si apprende il vocabolario di una L2 (si veda Steinel *et al.*, 2007 per una breve rassegna).

Figura 7 - Risultati del test in funzione della struttura predicativa della definizione



I partecipanti ottengono infatti risultati nettamente positivi nel task di completamento quando le definizioni si avvicinano ad uno stile enunciativo naturale (“argomento esplicitato nella definizione”, in fig. 7, con il 78% di risposte esatte), basato su una limitata rimozione delle “strutture e [delle] funzioni sintattiche” (Wiegand *et al.*, 2013: 320) che è invece tipica della condensazione testuale usata nei dizionari (Wiegand, 2009). Il recupero di una forma “naturale” del testo è stata peraltro invocata con forza da alcuni lessicografi britannici (Sinclair, 2004; Hanks, 1987), che l’hanno inserita come norma stilistica dei loro dizionari pedagogici [COBUILD]: tutte le definizioni in essi presenti sono infatti scritte come ‘enunciati completi’ (*full sentence definitions*, FSD), provvisti delle strutture argomentali e dei giuntivi testuali usati nel linguaggio naturale. Herbst (1996) sintetizza le caratteristiche di questa modalità definitoria mettendo in luce come essa assomigli alle formulazioni prodotte spontaneamente dai parlanti e, pertanto, come essa sia facilmente comprensibile. Le *full sentence definitions* (o ‘enunciati completi’) esemplificano inoltre le reggenze e i contesti d’uso dei lemmi poiché, per formulare una definizione che abbia una struttura enunciativa completa, è necessario esplicitare gli eventuali argomenti associati al termine, come accade nella definizione di ‘kick’ ripresa dal COBUILD online: “If you kick someone or something, you hit them forcefully with your foot”. Non vanno tuttavia dimenticate le riserve espresse da Rundell (2006: 328), che mette in risalto come questa tipologia di commento semantico sia preferibile solo nei casi in cui essa non impone la scrittura di commenti semantici troppo lunghi e complessi.

Tra le definizioni estratte dal *Treccani* per il test non ci sono tuttavia veri e propri enunciati completi; si possono riconoscere solo alcuni commenti semantici che, al contrario degli altri, esplicitano chiaramente l’argomento su cui il lemma da definire predica qualcosa. Ad esempio, *tagliare la testa al toro* viene spiegato come “risolvere decisamente, troncando definitivamente una questione: *questo è un argomento che taglia la testa al toro*”. Nella definizione si ritrova chiaramente espressa una parte predicativa (“risolvere decisamente, troncando definitivamente”) ed una argomentale

a cui il predicato si riferisce (“una questione”), una modalità enunciativa che preserva la completezza logica della spiegazione semantica.

Al contrario, in “*mettersi le mani nei capelli*, in segno di disperazione”, il sottolemma⁹ (“mettersi le mani nei capelli”) della voce ‘mano’ viene usato come argomento dell’enunciato e la definizione (“in segno di disperazione”) coincide con la parte predicativa.

Quando, invece, la spiegazione semantica viene fornita usando un unico sinonimo, il predicato risulta sostanzialmente indefinito, poiché non viene espresso. Ad esempio, in “*passare per la testa*, affacciarsi alla mente”, la parte predicativa che lega i due sinonimi, il verbo ‘significa’ (*passare per la testa* [significa] affacciarsi alla mente), viene completamente omessa.

Delle tre modalità di definizione solo la prima, quella con l’argomento chiaramente espresso nel commento semantico, risulta correlata con gli esiti positivi del test (78% di risposte corrette); quando invece è il lemma a fungere da argomento, i partecipanti appaiono in serie difficoltà (43% di risposte corrette), comprovate anche dal t-test, che evidenzia la significatività della differenza percentuale tra le risposte corrette e sbagliate per questa tipologia di definizione ($p < 0,05$). Al contrario, le definizioni con predicato inespresso risultano né vantaggiose né svantaggiose¹⁰.

Sembra quindi necessario che un dizionario per apprendenti stranieri abbia definizioni con una struttura predicativa completa, contenenti sia gli elementi che svolgono la funzione di argomenti, sia quelli con funzione di predicati. Sarebbe inoltre preferibile che questi commenti semantici si avvicinassero al modello degli enunciati completi in uso in molti dizionari pedagogici inglesi per poter introdurre sia le strutture argomentali sia le possibili variazioni flessive dell’espressione idiomatica.

4. *Dal monitoraggio dei bisogni alla compilazione di un dizionario elettronico*

Il tipo di indagine empirica fin qui presentata, analizzando la reale capacità degli utenti di estrarre informazioni dai dizionari, dovrebbe orientare la prassi lessicografica e consentire la compilazione di risorse efficaci ed efficienti. Si cercherà quindi d’illustrare brevemente la spendibilità dei risultati raggiunti con il test per la compilazione di una voce di un dizionario pedagogico elettronico di nuova concezione, detto “monofunzionale” (Bergenholtz - Bergenholtz, 2011). Questa tipologia di dizionario è stata elaborata all’interno di un filone di studi, detto *lexicographical*

⁹ Seguendo Hausmann e Wiegand (1989), la denominazione lessicograficamente più corretta per riferirsi a questa unità strutturale del dizionario dovrebbe essere quella di ‘indirizzamento non lemmatico’ (*non-lemmatic address*), dal momento che il *Treccani* non presenta un’organizzazione basata su ‘nicchie’ (*niche*). Ritengo tuttavia che in questo contesto la denominazione di ‘sottolemma’ risulti più intuitiva e comprensibile per riferirsi all’unità in oggetto. Per approfondimenti terminologici si rimanda anche a Gouws (2002).

¹⁰ La differenza percentuale tra le risposte corrette e quelle sbagliate per le definizioni con predicato inespresso non si rivela infatti significativa al t-test.

function theory, che ha messo in primo piano il ruolo delle nuove tecnologie per l'evoluzione della lessicografia e, in particolare, per la tipologia dei dizionari.

Per compilare il modello di voce lessicografica proposto, verranno ripresi i campi informativi di un dizionario elettronico della lingua tedesca, l'*OWID-Sprichwörterbuch* [Owid - Sprichwörterbuch]¹¹ (Steyer - Ďurčo, 2013), che sarà arricchito delle caratteristiche precedentemente discusse, ovvero definizioni scritte con una struttura enunciativa completa in cui risulti esplicito il legame tra il significato letterale e quello idiomatico associati al lemma. L'organizzazione "monofunzionale" della risorsa imporrà inoltre che i dati associati al lemma siano presentati in tre voci distinte, così come previsto dalle linee guida della *lexicographical function theory*.

4.1 Dizionari "monofunzionali"

Partendo da considerazioni generali su cosa siano realmente i dizionari, la *teoria delle funzioni lessicografiche* (*lexicographical function theory*, cfr. Tarp, 2008) ha messo in luce la loro natura di strumenti ("utility tools") finalizzati a colmare lacune informative circoscritte ("punctual"). Un dizionario, infatti, deve rispondere a diverse tipologie di esigenze, che variano a seconda del tipo di utente e del contesto in cui il dizionario viene usato, determinandone la funzione: cognitiva, per i dizionari che aiutano ad acquisire nuove conoscenze; comunicativa, per i problemi più strettamente connessi alla lingua; operativa, se il dizionario deve spiegare come svolgere un compito; simbolica, se è necessario decodificare dei simboli (come bandiere, segnali stradali e tutti i codici convenzionali non linguistici). A partire da queste linee guida, il fattore determinante nella strutturazione del dizionario non coincide più con la scelta sostanzialmente bipartita tra un'organizzazione onomasiologica o semasiologica, il lessicografo deve invece pianificare il dizionario in base alle funzioni che esso è chiamato a svolgere. Mentre i dizionari tradizionali, detti "polifunzionali", accolgono all'interno delle loro voci dati utili a svolgere diversi tipi di compiti, ad esempio scrivere e comprendere testi allo stesso tempo, i dizionari "monofunzionali", come gli *Accounting Dictionaries-Diccionarios de Contabilidad* [ACCOUNTING DICTIONARIES], distribuiscono le informazioni su uno stesso lemma in voci diverse, appartenenti ad altrettanti dizionari collegati tra di loro, che supportano l'utente in specifiche tipologie di compiti. La realizzazione tecnica di questo genere d'impostazione strutturale è affidata a dei filtri che distribuiscono i dati contenuti nel database lessicografico in diverse voci, sulla base delle funzioni che sono state preimpostate dal lessicografo (Heid, 2014). La parcellizzazione delle informazioni così ottenuta riduce il carico cognitivo richiesto agli utenti che intraprendono una ricerca, aumentando la loro capacità informativa e, conseguentemente, l'efficacia e l'efficienza dei dizionari stessi. In questo tipo di risorsa, infatti, l'utente deve speci-

¹¹ Questo dizionario è dedicato alla descrizione lessicografica dei proverbi che, come ricordano Dobrovol'skij e Pirainen (2005: 52), "devono essere considerati come una classe di fraseologismi strettamente legati ai modi di dire stessi".

ficare il tipo d'informazione di cui ha bisogno, ad esempio 'comunicativa', per visualizzare una delle voci associate al lemma che lo aiuti a svolgere il compito richiesto.

4.2 Per un dizionario elettronico monofunzionale dei modi di dire italiani

Viene presentato di seguito (figg. 8, 9, 10) il modello di tre entrate di dizionari "monofunzionali" contenenti le caratteristiche rilevanti per l'apprendimento emerse dal test. Queste voci illustrano l'*idiom* "mettersi le mani nei capelli" in rapporto a tre diverse funzioni: "capire un testo", per i compiti comunicativi di tipo ricettivo, "scrivere un testo", per quelli produttivi, "imparare nuove parole", per supportare gli utenti nell'apprendimento delle espressioni idiomatiche. I campi informativi di base sono stati ripresi dal citato *OWID-Sprichwörterbuch*.

Figura 8 - Voce di dizionario elettronico "monofunzionale" per 'capire un testo'

Mettersi le mani nei capelli (per la disperazione) componenti: mettere; mani; capelli

Capire un testo	
Significato	
- Idiomatico	Se qualcuno si mette le mani nei capelli , vuol dire che è disperato per qualcosa che è successo, o sta succedendo.
- Letterale	È come se si volesse strappare i capelli per sfogarsi.
atto/funzione linguistica:	lamentarsi; anche ironicamente
connotazione:	negativa; ironico e scherzoso
registro:	informale; colloquiale

Figura 9 - Dizionario elettronico "monofunzionale" per 'imparare nuove parole'

Imparare nuove parole	
Significato	
- Idiomatico	Se qualcuno si mette le mani nei capelli , vuol dire che è disperato per qualcosa che è successo, o sta succedendo.
- Letterale	È come se si volesse strappare i capelli per sfogarsi.
- Origine	Il gesto di mettersi le mani nei capelli veniva usato, in epoca antica, nei funerali dalle 'prefiche', donne pagate per piangere e disperarsi per la persona che era morta.
atto/funzione linguistica:	lamentarsi; anche ironicamente
connotazione:	negativa; ironico e scherzoso
registro:	informale; colloquiale
trasformazioni sintattiche:	<u>Gli ho fatto mettere</u> le mani nei capelli (per qualcosa) / <u>ho fatto mettere</u> le mani nei capelli a qualcuno
- tipi di frasi:	<p>Frase esclamative: <i>Ho visto l'acqua in giro e mi sono messa le mani nei capelli!</i> Anche introdotte da "Roba da": <i>'Roba da mettersi le mani nei capelli' dice l'industriale veneto</i></p> <p>Subordinate introdotte da "c'è da": <i>c'è da mettersi le mani nei capelli vedendo quello che hanno combinato.</i></p>

Si evidenzia come, nel rispetto dei risultati del test, nel dizionario "imparare nuove parole" sia presente una glossa etimologica, mentre sono molto ricche le informazioni sulle variazioni nella forma e sulle restrizioni d'uso fornite per la funzione produttiva ("scrivere un testo"). In quest'ultimo dizionario, inoltre, possono essere integrati link a corpora, come [CORIS/CODIS] o ['LA REPUBBLICA' CORPUS], che possono fornire esemplificazioni aggiuntive sia delle diverse "modalità d'uso nei testi", sia delle "variazioni" morfosintattiche ammesse.

Figura 10 - Dizionario elettronico “monofunzionale” per ‘scrivere un testo’

Scrivere un testo	
Dominio →	<i>esprimere i sentimenti - disperazione</i>
Flessione e variazioni:	Mi sono messo le mani nei capelli
Significato	
- Idiomatico	Se qualcuno si mette le mani nei capelli , vuol dire che è disperato, esasperato per qualcosa che è successo, o sta succedendo davanti ai suoi occhi.
atto/funzione linguistica:	lamentarsi; anche ironicamente
connotazione:	negativa; ironico e scherzoso
registro:	informale; colloquiale
Frequenza d'uso in contesti specifici	soprattutto in conversazioni tra amici
Frequenza d'uso in testi specifici	dialoghi; discorso riportato
Modalità d'uso nei testi	
- tipi di frasi:	Frasi esclamative: <i>Ho visto l'acqua in giro e mi sono messa le mani nei capelli!</i> Anche introdotte da “Roba da”: <i>'Roba da mettersi le mani nei capelli' dice l'industriale veneto</i> Subordinate introdotte da “c'è da”: <i>c'è da mettersi le mani nei capelli vedendo quello che hanno combinato.</i>
- connettori:	C'è da; roba da
forme negative:	raro: ? <i>Non ci mettiamo le mani nei capelli se succede questa cosa,</i> può essere sostituito con: <i>non ci mettiamo a piangere se succede questa cosa</i>
Variazioni	Sostituzione dei componenti: verbo + con le mani nei capelli [esempi → <i>ha gridato con le mani nei capelli; erano scappati via con le mani nei capelli; stanno pensando con le mani nei capelli che cosa fare</i>] Trasformazioni sintattiche tipiche: <i>Gli ho fatto mettere le mani nei capelli (per qualcosa) [esempi]/ ho fatto mettere le mani nei capelli a qualcuno</i>

5. Conclusioni

Il passaggio al formato elettronico offre ai dizionari l'occasione di una radicale trasformazione strutturale. Vocabolari sempre più ricchi di informazioni possono infatti preservare la capacità informativa degli utenti attraverso un'adeguata presentazione dei loro dati.

Adottando come metodologia investigativa il test sulle fasi di post-consulazione (“post-consultation study”), questo studio ha collezionato evidenze su come arricchire e, al contempo, rendere più adeguata la descrizione lessicografica per una specifica tipologia di dizionari. I risultati raggiunti attraverso la somministrazione di un apposito test provano infatti che, per essere efficace, un dizionario idiomatico pedagogico dovrebbe contenere anche note di tipo etimologico, con cui venga spiegato il passaggio dal significato letterale a quello figurato del fraseologismo, mentre la definizione deve essere formulata in maniera esplicita, sotto forma di una enunciazione completa che non comprometta l'intelligibilità della sua struttura logica.

Nella seconda parte di questo lavoro è stato invece fornito un esempio concreto di come l'arricchimento di dati e l'aumento di accuratezza richiesti dai risultati dell'indagine empirica possono essere organizzati all'interno di una risorsa di tipo elettronico di nuova concezione. Inglobando i dati risultati rilevanti al test lessicografico, il dizionario che è stato abbozzato distribuisce differentemente i suoi campi informativi all'interno di tre voci: selezionando funzioni diverse, le voci che vengono visualizzate contengono dati differenti. Ad esempio, conformemente ai risultati del test, la descrizione del significato è molto più accurata nella voce che supporta l'apprendimento (“imparare nuove parole”, fig. 9) rispetto a quella che viene visualizzata.

lizzata per scrivere un testo (fig. 10), quando plausibilmente l'utente ha bisogno di essere rassicurato sul valore semantico di un'espressione che conosce già. Questa parcellizzazione del commento semantico evita il tipo di sovraccarico cognitivo su cui anche Rundell (2006) ha invitato a riflettere, riferendosi tuttavia a definizioni meno ricche d'informazioni rispetto a quelle che sono state prospettate sin qui per le espressioni idiomatiche.

Naturalmente, in conformità con l'impostazione metodologica sin qui seguita, l'efficacia del tipo di dizionario proposto necessita di essere sottoposta al vaglio degli utenti. La ricerca verrà quindi completata mediante un'ulteriore indagine empirica, in cui un gruppo di apprendenti stranieri, parlanti nativi di una lingua isolante, saranno chiamati a svolgere lo stesso compito di riempimento consultando i tre dizionari monofunzionali appena presentati.

Bibliografia

APRILE M. (2005), *Dalle parole ai dizionari*, Il Mulino, Bologna.

BENNELI B. - BELACCHI C. - GINI G. - LUGANGELI D. (2006), 'To define means to say what you know about things': the development of definitional skills as metalinguistic acquisition, in *Journal of Child Language* 33: 71-97.

BERGENHOLTZ H. - BERGENHOLTZ I. (2011), A Dictionary Is a Tool, a Good Dictionary Is a Monofunctional Tool, in FUERTES-OLIVERA P.A. - BERGENHOLTZ H. (eds.) *e-Lexicography: The Internet, Digital Initiatives and Lexicography*, Continuum, London/New York: 187-207.

BOSCH S. - FAASS G. (2014), Towards an Integrated E-Dictionary Application. The Case of an English to Zulu Dictionary of Possessives, in ABEL A. - VETTORI C. - RALLI N. (eds.), *Proceedings of the XVI EURALEX International Congress. The User in Focus*, 15-19 July 2014, EURAC research, Bolzano/Bozen: 1087-1098.

BURGER H. (2010), *Phraseologie. Eine Einführung am Beispiel des Deutschen*, Erich Schmidt, Berlin.

CARR M. (1997), Internet Dictionaries and Lexicography, in *International Journal of Lexicography* 10(3): 209-230.

COUNCIL OF EUROPE (2007), *Common European Framework of Reference for Languages: Learning, Teaching, Assessment*, Council of Europe. Disponibile in: http://www.coe.int/t/dg4/linguistic/Cadre1_en.asp (data ultimo accesso 30/07/2015).

COWIE A.P. (1981), The Treatment of Collocations and Idioms in Learners' Dictionaries, in *Applied Linguistics* II(4): 223-235.

DOBROVOĽSKIJ D. (2001), Pragmatische Faktoren bei der syntaktischen Modifizierbarkeit von Idiomen, in LIEDTKE F. - HUNDSNURSCHER F. (Hrsg.), *Pragmatische Syntax*, Niemeyer, Tübingen: 271-308.

DOBROVOĽSKIJ D. (2009), Zur lexikografischen Repräsentation der Phraseme (mit Schwerpunkt auf zweisprachigen Wörterbüchern), in MELLADO BLANCO C. (Hrsg.), *Theorie und Praxis der idiomatischen Wörterbücher*, (Lexicographica Series Maior), de Gruyter, Berlin/New York: 149-168.

- DOBROVOĽSKIJ D. - PIIRAINEN, E. (2005), *Figurative language: cross-cultural and cross-linguistic perspectives*, Emerald, London.
- FILLMORE C.J. - KAY P. - O'CONNOR J. (1988), Regularity and idiomaticity in grammatical constructions: the case of 'let alone', in *Language* 64: 521-538.
- FUERTES-OLIVERA P.A. - TARP S. (2014), *Theory and Practice of Specialised Online Dictionaries*, (Lexicographica Series Maior), de Gruyter, Berlin/New York.
- GOUWS R.H. (2002), Nicheing as a Macrostructural Procedure, in *Lexikos*, 12: 133-158.
- GOUWS R.H. (2014), Article structures: Moving from Printed to e-Dictionaries, in *Lexikos*, 24: 155-157.
- HANKS P. (1987), Definitions and Explanations, in SINCLAIR J.M. (ed.), *Looking Up*, Collins, London: 113-136.
- HEID U. (2011), Electronic Dictionaries as Tools: Toward an Assessment of Usability, in FUERTES-OLIVERA P.A. - H. BERGENHOLTZ (eds.), *e-Lexicography: The Internet, Digital Initiatives and Lexicography*, Continuum, London/New York: 287-304.
- HEID U. (2014), Natural Language Processing Techniques for Improved User-friendliness of Electronic Dictionaries, in ABEL A. - VETTORI C. - RALLI N. (eds.), *Proceedings of the XVI EURALEX International Congress: The User in Focus*, 15-19 July 2014, EURAC Research, Bolzano/Bozen: 47-61.
- HERBST T. (1996), On the way to the perfect learners' dictionary: A first comparison of OALD5, LDOCE3, COBUILD2 and CIDE, in *International Journal of Lexicography* 9 (4): 321-357.
- JEŽEK E. (2011), *Lessico. Classi di parole, strutture, combinazioni*, 2° edizione, Il Mulino, Bologna.
- LEVORATO M.C. - CACCIARI C. (2002), The creation of new figurative expressions: psycholinguistic evidence in Italian children, adolescents and adults, in *Journal of Child Language* 29: 127-150.
- RUNDELL M. (2006), More than one way to skin a cat: why full-sentence definitions have not been universally adopted, in CORINO E. - MARELLO C. - ONESTI C. (eds.), *Proceedings of 12th EURALEX International Congress*, Edizioni Dell'Orso, Alessandria: 323-337.
- SINCLAIR J. (2004), In praise of the dictionary, in WILLIAMS G. - VESSIER S. (eds.), *Proceedings of the Eleventh EURALEX Congress*, Université de Bretagne-Sud, Lorient (France): 1-12.
- STEINEL M.P. - HULSTIJN J.H. - STEINEL W. (2007), Second-language idiom learning in a paired-associate paradigm: Effects of direction of learning, direction of testing, idiom imageability, and idiom transparency, in *Studies in Second Language Acquisition* 29: 449-484.
- STEYER K. - ĐURČO P. (2013), Ein korpusbasiertes Beschreibungsmodell für die elektronische Sprichwortlexikografie, in BENAYOUN J.M. - KÜBLER N. - ZOUOGBO J.P. (eds.), *Parémiologie. Proverbes et formes voisines*, Sainte Gemme, Band 3: 219-250.
- TARP S. (2008), *Lexicography in the Borderland between Knowledge and Non-knowledge. General Lexicographical Theory with Particular Focus on Learner's Lexicography*, (Lexicographica Series Maior), de Gruyter, Berlin/New York.
- WIEGAND H.E. (1989), Arten von Mikrostrukturen im allgemeinen einsprachigen Wörterbuch, in HAUSMANN F.J. - REICHMANN O. - WIEGAND H.E. - ZGUSTA L. (Hrsg.),

(1989), *Wörterbücher/Dictionnaires. An International Encyclopedia of Lexicography*, de Gruyter, Band, Berlin/New York 1: 462-501.

WIEGAND H.E. (2009), Nichtnatürlich über natürliche Sprache schreiben. Zu einigen formalen Aspekten von Wörterbuchartikeln, in *Heidelberger Jahrbücher [Themenabend „Sprache“ Hrsg. v. Ekkehard Felder]* 53: 247-285.

WIEGAND H.E. - FEINAUER I. - GOUWS R.H. (2013), Types of dictionary articles in printed dictionaries, in GOUWS R.H. - HEID U. - SCHWEICKARD W. - WIEGAND H.E. (Hrsg.), *Dictionaries. An International Encyclopedia of Lexicography: Supplementary volume. Recent Developments with Special Focus on Computational Lexicography*, Mouton de Gruyter, Berlin: 314-366.

WIEGAND H.E. - SMIT M. (2013), Microstructures in printed dictionaries, in GOUWS R.H. - HEID U. - SCHWEICKARD W. - WIEGAND H.E. (Hrsg.), *Dictionaries. An International Encyclopedia of Lexicography: Supplementary volume. Recent Developments with Special Focus on Computational Lexicography*, Mouton de Gruyter, Berlin: 149-214.

Dizionari e Corpora

[ACCOUNTING DICTIONARIES] (2015), *Accounting Dictionaries-Diccionarios de Contabilidad*, Hamburg, Lemma GmbH, <http://lemma.com/dictionary/special/center-for-leksikografi/accounting/?dict=Spanish+Accounting>.

[COBUILD] (2015) *Collins COBUILD English Dictionary*, London/Glasgow, HarperCollins, <http://www.collinsdictionary.com/dictionary/english-cobuild-learners>.

[CORIS/CODIS] ROSSINI FAVRETTI R. (ed.) (1998-2015), *CORIS/CODIS - CORpus di Riferimento dell'Italiano Scritto*, Università degli Studi di Bologna, http://corpora.ficlit.unibo.it/coris_eng.html.

['LA REPUBBLICA' CORPUS] BARONI M. - BERNARDINI S. - COMASTRI F. - PICCIONI L. - VOLPI A. - ASTON G. - MAZZOLENI M. (ed.) (2004), *'La Repubblica' Corpus*, <http://dev.sslmit.unibo.it/corpora/corpus.php?path=&name=Repubblica>.

[OWID-SPRICHWÖRTERBUCH] OWID – *Sprichwörterbuch. Online-Wortschatz-Informationssystem Deutsch*, Institut für Deutsche Sprache, Mannheim, <http://www.owid.de/wb/sprw/start.html>.

[SABATINI-COLETTI] SABATINI F. - COLETTI V. (eds.) (2011), *il Sabatini Coletti. Dizionario della Lingua Italiana*, Milano, RCS Libri.

[TRECCANI] *Vocabolario Treccani*, Istituto della Enciclopedia Italiana, <http://www.treccani.it/vocabolario>.

[WIKTIONARY] (2015), *Wiktionary*, Creative Commons Attribution-ShareAlike License, https://en.wiktionary.org/wiki/Wiktionary:Main_Page.

ELISA CORINO¹

“Scelto A perché frase B secondo me usa tanti verbi”. Attività riflessive *corpus-based* per elicitare competenze (meta)linguistiche

Current research in SLA is based on the assumption that some level of attention to form is needed for language acquisition to take place; it has also been extensively demonstrated that a higher level of awareness correlates positively with language development.

Starting from the evidences from VALICO (*Varietà di Apprendimento Lingua Italiana Corpus Online*), multiple choice exercises were prepared aiming at eliciting metalinguistic comments, thus enhancing the passage from *noticing* to *understanding* through a process of awareness.

A case study on the items containing the Italian pronoun *si* will be presented and recordings of class discussions will be analysed in order to answer the main research questions: what are the actual reasons behind the choice of a given distractor? And, does a good choice reflect good metalinguistic competence and language awareness?

1. Introduzione

La presente ricerca ha lo scopo di verificare la correlazione tra la competenza d'uso della lingua straniera (in questo caso l'italiano) e la capacità di descrivere la lingua stessa, ovvero la competenza metalinguistica degli studenti.

Sulla scorta delle teorie sull'apprendimento che rilevano l'importanza dell'*awareness* nei processi di *noticing* per favorire l'*understanding* (Schmidt, 1990; Robinson, 1997), si osserveranno i comportamenti degli studenti di fronte alla richiesta di esplicitare le ragioni che li hanno guidati nell'attribuzione di punteggi scalari ai distrattori di esercizi a scelta multipla.

L'analisi ha lo scopo di esplicitare le ragioni che sottendono l'attribuzione dei punteggi e di stabilire se una spiegazione metalinguistica approssimativa o errata può corrispondere ad un elevato grado di correttezza nella soluzione dell'esercizio, o se viceversa solo chi ha piena padronanza e sa verbalizzare in modo preciso le regole che governano il sistema lingua ha risultati positivi.

Verrà descritto il processo di costruzione degli esercizi che hanno elicitato le discussioni in classe, con particolare riferimento agli item che contengono il clitico *si*. A questo seguirà l'analisi dei risultati ottenuti da studenti anglofoni e lusofoni e delle registrazioni delle loro riflessioni. Infine, verranno avanzate ipotesi sulla reale competenza linguistica di questi studenti.

¹ Università di Torino.

1.1 Apprendimento tra *noticing* e *understanding*

La riflessione metalinguistica è uno strumento che fa emergere “gli apparati di riconoscimento delle strutture linguistiche e di controllo delle regole” (Vedovelli, 1990: 187), in quanto mette in luce la percezione da parte del parlante della propria competenza sulla lingua. Sono molti gli studi condotti nel campo dell’acquisizione delle lingue seconde (*Second Language Acquisition*, SLA) nell’ultimo decennio (Long, 1996; Grassi - Magiarini, 2010; Schmidt, 2010, tra i tanti) che sottolineano come il *focus on form* sia un presupposto necessario affinché il processo di acquisizione abbia luogo.

A partire dal modello di Schmidt (1990), e nonostante le molte critiche, l’*awareness* – la consapevolezza sull’oggetto di studio linguistico – è riconosciuta come componente centrale per la rappresentazione dei meccanismi interni al sistema linguistico della L2 degli apprendenti, che sono per lo più abituati ad eseguire compiti scritti mirati all’applicazione delle regole spiegate e a controllare la correttezza delle loro esecuzioni con la verifica dell’insegnante.

Robinson (1997) concorda con Schmidt nel sostenere l’idea che apprendimento e consapevolezza siano strettamente correlate e sottolinea come l’apprendente abbia la chiara necessità di notare esplicitamente una certa forma dell’input – *noticing* – prima che questa possa essere processata dalla sua interlingua e attraverso l’*understanding* – la concettualizzazione e sistematizzazione – diventare *intake*.

La correlazione positiva fra lo sviluppo della competenza metalinguistica e il livello di conoscenza della L2 negli apprendenti adulti è stata più volte dimostrata da studi che hanno rilevato come la capacità di individuare e verbalizzare le regole d’uso delle forme linguistiche segna la transizione dalla conoscenza implicita alla conoscenza esplicita della lingua (Tomlin - Villa, 1994; Alanen, 1995; Robinson, 1997).

I dati riportati nei più recenti Radwan (2005) e Roher (2007) e i dati osservati per l’inglese da Butler (2002), per l’italiano da Marellò (2009) e Rosi (2010), infine, confermano che un alto livello di consapevolezza linguistica e metalinguistica rispetto al sistema hanno influenze positive sulla progressione nell’apprendimento. In particolare la capacità degli apprendenti di correggere, descrivere e spiegare gli errori, la capacità di rilevare e analizzare strutture della LS sono strettamente correlate all’effettiva produzione linguistica degli apprendenti stessi.

1.2 Test a scelta multipla per sviluppare la competenza metalinguistica

L’ipotesi interazionista di Long (1996) suggerisce che l’apprendimento di una lingua straniera è facilitato dalla presenza di processi interazionali grazie al ruolo che la comunicazione svolge nell’interconnessione tra input, attenzione selettiva e produzione dell’output. Sulla scorta del *focus on form* auspicato da Long, Ellis (2001) suggerisce una didattica focalizzata sulla forma (*form-focused instruction*) che descrive come “any planned or incidental instructional activity that is intended to induce language learners to pay attention to linguistic form” (Ellis, 2001: 1-2).

Gli esercizi a scelta multipla utilizzati in questo studio di caso, creati a partire dal corpus di apprendenti VALICO (Varietà di Lingua Italiana Corpus Online)², sono stati concepiti proprio per stimolare la riflessione sulla forma e di indurre gli apprendenti a tradurre le loro considerazioni in forma dichiarativa, dialogando sulle scelte operate e sulle ragioni di tali scelte.

Nelle discussioni in classe, il feedback dell'insegnante e dei compagni acquista in questo senso un valore aggiunto, dato dall'interazione e dalla negoziazione dei significati (e dell'appropriatezza dell'una o dell'altra scelta), che contribuisce a focalizzare l'attenzione dell'apprendente sulle possibili discrepanze tra la lingua di arrivo e le possibili forme dell'interlingua che Schmidt - Frota (1986) definivano come *noticing the gap*.

La discussione costruita intorno alla riflessione metalinguistica si configura come feedback e come modalità per mettere chiaramente in luce la distinzione fra competenza implicita e conoscenza esplicita (Bardel - Falk, 2012; Paradis, 2009), rivelando importanti informazioni sulla capacità degli studenti di:

1. individuare le deviazioni dalla norma;
2. usare un lessico specialistico – quello della descrizione grammaticale – per riferire e descrivere fenomeni;
3. descrivere un fenomeno linguistico in modo corretto, distinguendo forme e funzioni e definendolo in modo preciso, non solo per ragioni legate alla percezione e dovute alla sua frequenza nell'input.

Sono numerosi i casi in cui la discussione può far emergere ragionamenti che un insegnante avrebbe difficoltà a ricostruire, in cui le ipotesi interpretative del ricercatore in merito alle scelte operate dagli apprendenti vengono smentite, gettando nuove luci sui processi che hanno governato una certa scelta (si veda l'interessante discussione sulla natura del pronome/articolo *gli* in Colombo, 2009).

Elicitare in modo esplicito le ragioni delle scelte, poi, rivela il pensiero che ha guidato la selezione da parte dei discenti e con esso gli eventuali *covert errors*, che altrimenti sarebbe stato difficile individuare come tali, e invalida il fattore relativo alla probabilità statistica di scegliere per caso la risposta giusta, permettendo inoltre di estendere le considerazioni sull'apprendimento di una lingua ai processi cognitivi che sottendono le produzioni degli apprendenti, così come essi stessi li esplicitano.

2. *Materiali e metodi*

Come verrà meglio dettagliato nei paragrafi 2.1 e 2.2, il presente studio di caso utilizza esercizi a scelta multipla creati a partire da estrazioni di fenomeni sintattici osservati in VALICO e costruiti citando occorrenze di alcuni fenomeni morfosintattici e lessicali presenti nel corpus (Marello, 2009; Corino - Marello, 2011). Ogni esercizio include l'opzione corretta e tre distrattori.

² Corpus di apprendenti di italiano LS/L2, liberamente consultabile all'indirizzo www.valico.org.

Gli esercizi sono poi stati sottoposti a un gruppo di studenti anglofoni in soggiorno di studio a Siena e Firenze, e lusofoni presso l'Università di San Paolo in Brasile. Il livello degli apprendenti è A2/B1.

Agli studenti è stato inoltre chiesto di commentare oralmente i distrattori dando una spiegazione circostanziata delle ragioni per cui una certa scelta può essere considerata giusta o sbagliata.

La discussione elicitata a partire dagli esercizi a scelta multipla si inserisce nel solco di quelle ricerche (Grassi - Mangiarini, 2010) che sottolineano i vantaggi del feedback esplicito nella forma dell'informazione metalinguistica. Tale ambito di ricerca si rifa, in particolare, al percorso di correzione indicato da Ellis *et al.* (2006). Ellis e colleghi prevedono che lo studente spieghi esplicitamente le scelte linguistiche da lui operate, articolando le proprie riflessioni in diverse fasi quali: focalizzazione dell'errore, suggerimento tecnico e *prompt*.

Le registrazioni delle discussioni in classe, basate su un ampio insieme di esercizi collegati a VALICO³, da una parte offrono agli studiosi di acquisizione risorse e materiali nuovi da analizzare in relazione ai processi cognitivi che regolano le scelte e la formazione del sistema-interlingua; dall'altra si presentano come un'opportunità per i docenti in formazione di entrare in contatto con le problematiche di uno specifico gruppo di apprendenti ancora prima di conoscere la propria classe: confrontarsi con le registrazioni costringe il docente ad interrogarsi sulle spiegazioni da dare, sulle modalità di formularle e sulla loro efficacia, oltre naturalmente ad aprire una finestra sul paradigma di spiegazioni linguistiche che potrebbero emergere durante la fase di confronto.

2.1 La costruzione degli esercizi

Gli esercizi creati a partire da VALICO sono utilizzabili a livello A1, A2 e B1 per tutti i tipi di apprendenti, indipendentemente dalla loro L1, per la sintassi dei tempi e dei modi fra principale e subordinate, ma anche per i pronomi clitici e per l'accordo del participio nei tempi composti o per i pronomi riflessivi, che sono più o meno difficili per tutti. Questi sono stati integrati con item specifici per apprendenti anglofoni, francofoni, ispanofoni, lusofoni, russofoni e sinofoni, che contengono distrattori "naturali" ispirati o letteralmente ripresi dalle frasi prodotte dagli apprendenti stessi nei testi raccolti nel corpus.

Per questo studio di caso è stato scelto il pronome *si*, che costituisce una fonte di difficoltà e incertezze a tutti i livelli (Corino - Marellò, in stampa): gli apprendenti mostrano una certa imprecisione nel distinguere l'uso impersonale da quello passivante e, sovrapponendoli, ne sovraestendono e ipergeneralizzano l'uso, oltre a mostrare difficoltà nell'accordo con il verbo, come mostrano i seguenti esempi⁴:

³ Gli esercizi sono liberamente scaricabili in formato testuale dal sito di VALICO ed è possibile svolgerli online grazie al software SoFIA (Software di Formazione Insegnanti ed Autovalutazione, Procopio 2013), www.valico.org/quizproc/finale.

⁴ Gli esempi riportati sono ripresi dal corpus nella loro forma originaria, compresi errori di ortografia.

- (1) Questi corsi si fanno dai professori universitari.
[L1 romeno, annualità 2]
- (2) Tuttavia , il suo nome si ha detto sul [loudspeaker] per il telefono
[L1 inglese, annualità?]
- (3) Ben presto gli altri personaggi della scena si verranno immischiati nel flusso travolgente dell'azione.
[L1 ceco, annualità +]
- (4) C'era una volta una principessa, si chiamava Roppunsel, si amava un cavaliere.
Ma ai suoi genitori non piacevano que il cavaliere
[L1 inglese, annualità?]

Il Profilo della lingua italiana (Spinelli - Parizzi, 2010) inserisce la competenza d'uso delle funzioni riflessive fin dall'A1 (molti sono i verbi riflessivi annoverati nella lista lessicale), integra l'impersonale al B1 e completa il quadro con il *si* passivante al B2. Corino - Marellò (in stampa) osservano come proprio la costruzione impersonale sia fonte di errore in particolare laddove il *si* impersonale usato alla terza persona dei verbi transitivi viene sovraesteso e si innesta su un intransitivo che ha anche uso impersonale, ad esempio *sembrare*:

- (5) Prima si sembrava che lei fosse contenta , però a dopo , si è messa ad arrabbiarsi con lui
- (6) La zia davvero si è innamorata e si sembrava che fosse più felice che mai prima.

Gli esempi (4) e (6) sono emblematici della frequente sovraestensione del *si* in presenza di proposizioni coordinate: la selezione della particella in riferimento ad un verbo determina infatti la sua sovraestensione anche all'altra azione. Si tratta di un procedimento indipendente dalla funzione del *si*, poiché ricorre sia in presenza della funzione riflessiva, sia nel caso di impersonali (come nella scelta multipla *molte cose si possono comprare...*).

Inoltre, nel corpus VALICO, non mancano casi di evidente transfer interlinguistico, sia positivo che negativo. Gli apprendenti germanofoni, influenzati dalla funzione del *man* tedesco, tendono a fare un uso copioso delle forme impersonali, come in (7) e (8), e ad accordarle erroneamente alla terza persona singolare:

- (7) Parigi mi sembra come molte piccole città insieme , non si sente di vivere in una grande città. Si conosce tutti i vicini , si ha il piccolo negozio?? do ve si compra gli alimenti.
- (8) Sicuramente, con una laurea in "economia aziendale internazionale" si ha tutte le possibilità di lavorare nell'azienda che si vuole e si può scegliere

Similmente anche nei testi dei francofoni e degli anglofoni si riconosce l'interferenza di *on* e *one*:

- (9) Di più uno si ricorda del nome della città scritta sulla freccia dove aveva fatto la mietitura.

Tali considerazioni hanno dunque costituito una base di partenza per la selezione delle occorrenze dalle quali trarre i distrattori per gli esercizi a scelta multipla intorno alla particella pronominale *si* e alle sue numerose funzioni, con particolare riferimento all'uso impersonale, passivante e riflessivo, e ai problemi di accordo che la sua presenza comporta a livello sintagmatico.

2.2 La consegna degli esercizi

Per elicitarne riflessioni su forme e funzioni e successivamente sollecitare riflessioni ad alta voce che non si focalizzassero solo sulla risposta corretta, ma che prendessero in esame anche le occorrenze sbagliate e ne dessero conto definendone le caratteristiche e i motivi della deviazione dalla norma, è stato chiesto agli studenti di attribuire punteggi decrescenti ai distrattori degli esercizi. La consegna recita:

Scegli la frase corretta e attribuisce a quella un punteggio 4; alle altre attribuisce punteggi decrescenti 3, 2, 1 in base al tuo giudizio di scorrettezza (3 quasi giusta, 2 meno giusta, 1 la più sbagliata)

La relazione tra il valore assegnato dagli studenti al distrattore e la capacità metalinguistica di individuare e verbalizzare le regole d'uso delle forme linguistiche sarà dunque significativa per definire il rapporto tra conoscenze implicite e conoscenze esplicite e quindi anche il livello di competenza degli studenti, questa volta anche in chiave produttiva e non solo meramente ricettiva.

Da tale relazione emergono quindi da una parte le sicurezze e le insicurezze degli apprendenti, la differenza tra la struttura acquisita e pienamente conosciuta e padroneggiata e la percezione legata a frequenza e familiarità con una certa forma, la produzione di motivazioni incoerenti con la selezione effettuata; dall'altra si sottolineano le autocorrezioni dovute al *noticing* che diventa *understanding* grazie al momento di confronto e discussione e alla conseguente presa di coscienza delle ragioni di sistema sottese ad una certa scelta.

3. Risultati

Per illustrare i risultati degli studenti particolarmente significativo l'esercizio seguente:

- A. Molte cose si possono comprare quando si hanno tanti soldi
- B. Molte cose si possono essere comprate quando si hanno tanti soldi
- C. Molte cose si può essere comprato quando si hanno tanti soldi
- D. Molte cose si può comprato quando si hanno tanti soldi

Dei sette studenti anglofoni che hanno partecipato alla discussione sull'attribuzione dei punteggi ben quattro hanno selezionato la B come risposta corretta e hanno attribuito ad A tre punti; due studenti dichiarano di aver cambiato idea poco prima

di consegnare e di essere passati da A a B e da A a C, tutti concordano nell'assegnare un solo punto a D.

Emergono quindi errori che riflettono incertezze rispetto all'uso della forma passiva in alternativa a quella pronominale, gli studenti dichiarano la propria indecisione e candidamente ammettono di aver fatto una scelta dettata dal caso, di aver "tirato a indovinare".

Se nel caso di C anche lo studente che l'ha identificata come corretta è in grado di identificare l'errore di accordo e spiegarlo:

- (10) io io ho cambiato <ahm> ma<aa> il ricordo perché: può ma: adesso <ahmmm> vedo molte cose è plurale e so<o><ahh> A giusto ma: C è sbagliato <ehh> io schelto io ho schelto C perché<ee>last minute change perché può ma non realisatto ho non realisatto molte cose è plurale

per quanto riguarda la differenza tra A e B, invece, nessuno studente, indipendentemente dal fatto che ha attribuito quattro punti ad A o a B, dimostra di riuscire motivare il perché B non è la scelta corretta, né spiegare perché A invece lo è.

- (11) non mi ricordo qual altro era ancora fra di due quasi tutti questi per me era fra due altre risposte [...] <ahmm> è un po' difficile m ho scelto B e per me ho scelto B perché <uhm> ho scelto B perché<eee> non mi ricordo qual altro era ancora fra di due quasi tutti questi per me era fra due altre risposte ma ho scelto B <ahh><ah> (.) devo trovare quale altro solo <uhm>. Tra B e A non sono sicuro non mi ricordo perché <mm> mi è voluto un po' di tempo da indovinare questo ma per me era un po' <o> non ero molto sicuro ma ho indovinato B [...] credo che era metà metà
- (12) per me ho scelto A<aa> e<ee> per me fra A e B <eh> ho scelto A perché<ee><uhm> non so perché [...] non so frase per me <uhh> difficoltà di capire

Il distrattore B viene riconosciuto da molti come corretto probabilmente in virtù della coerenza dell'accordo S-V e V-O, uno studente commenta:

- (13) molte cose si possono, *si possono* è riflessivo per *le cose*: e quindi<ii><ehh> deve essere: plurale, <ee>*essere comprato* è giusto, *essere comprato*, non solo *comprato*, possono comprato, perché non ha senso, quando si hanno soldi, non so, butthat is... soldi è giusto per <iI> le quattro. Quindi B per 4, e poi: C per 3 perché *essere comprato*. <uhm> poi<ii> poi A per<rr> due, perché è *si possono* è plurale, poi D è più sbagliato.

In questo caso chi parla ha notato l'accordo tra soggetto (*le cose*) e verbo (*si possono*), in opposizione a C che invece presenta il verbo alla terza persona singolare. Notiamo come lo studente non riconosca il *si* come passivante, ma lo riconduca alla funzione riflessiva, l'unica che probabilmente ha già analizzato in modo esplicito – ricordiamo che gli studenti sono di livello A2/B1 mentre il Profilo segnala il *si* passivante come una competenza B2. La mancata identificazione della funzione del pronome comporta un corrispondente errore nell'osservazione sulla ridondanza

data dalla successiva costruzione con il verbo *essere*, invece considerata pienamente accettabile proprio come espressione necessaria della diatesi passiva.

Un altro studente fornisce una terza via proponendo una soluzione alternativa e dichiarando di aver scelto il distrattore che più si avvicinava alla sua variante, ancora una volta il verbo *essere* si rivela il perno intorno al quale vengono operate le selezioni, il pronome sembra qui non avere alcun peso sintattico:

- (14) per me qualcosa più giusto sarebbe *molte cose possono essere comprate quando si ha hanno tanti soldi* ma non c'era allora ho scelto qualcosa che era sarebbe più vicino e ho scelto B

Uno solo individua la ridondanza di costruzioni passivanti, ma lo fa in modo implicito e non è in grado di verbalizzare consapevolmente la ragione per cui B non è la frase corretta:

- (15) ho scelto A perché <ee><uhm> non so perché frase B <uhm> secondo me is<uh> tanti verbi quando penso <uhh> ho pensato<o><uhh> non so frase per me <uhh> difficoltà di capire ma<a> ma<a><uhh> ci sono sì tro- troppi verbi

La difficoltà dell'esercizio è poi certamente complicata dalla presenza di due clitici con funzioni diverse: uno passivante, l'altro impersonale; la presenza del secondo, "fisso" perché non segnato da modificazioni nell'elaborazione dei distrattori, sembra giustificare l'uso del primo e indurre la scelta di B, preferendolo ad A. Nelle discussioni elicitate, tuttavia, gli studenti si concentrano sulla prima parte della frase, quasi ignorando la seconda; anche il commento dell'insegnante (16) non mette in luce le differenze tra i due: si illustra la differenza tra A e B e si descrive il *si* impersonale passivante come un *si* riflessivo, perdendo di vista l'effettiva funzione della particella pronominale nel caso in questione e fornendo una spiegazione (meta) grammaticale poco efficace.

- (16) perché *essere comprato* lo usi poi nel passivo. Se tu... se tu ad esempio dici *molte cose possono essere comprate quando si hanno tanti soldi* va bene, però qui c'è la presenza del (.) del *si*, della particella *si* (.) che cambia un po' il significato della frase, capito? Si possono comprare (.) è possibile comprarle (.) capito?
[...]
Se tu invece avessi tolto il *si* potevi benissimo dire *molte cose possono essere comprate quando si hanno tanti soldi*, una frase passiva (.) invece con la frase riflessiva quella giusta è la A

Osserviamo che la prima parte della spiegazione data è corretta, ma l'attribuzione di una funzione riflessiva al *si* è inappropriata, inoltre manca il focus sulla differenza tra il primo *si* [possono comprare] e il secondo *si* [hanno tanti soldi]: passivante il primo, impersonale il secondo.

Nella discussione gli studenti sembrano in effetti focalizzarsi più sulla questione ortografica che su quella morfosintattica

- (17) [...] perché la forma corretta del verbo è coltivare, non cultivare

- (18) <chh> anch'io non sono sicura penso che sia coltivare non cultivare

Rispetto alla riflessione metalinguistica sulla correttezza dei distrattori, in questo caso i punti attribuiti dagli studenti sono più distribuiti: non c'è una preponderanza nella scelta della frase da quattro punti, emergono anzi selezioni estremamente incongruenti.

Quattro punti sono stati ad esempio assegnati da più di due studenti sia ad A, che a B e a D. D'altra parte A ha è stata segnalata da più di uno studente come "la risposta più sbagliata" con punteggio 1. In particolare un'apprendente, che ha dato 4 alla risposta D, 3 a B e 2 a C, sostiene che accordare il verbo con il complemento diretto comporterebbe un'interpretazione riflessiva della frase, implicitamente si inferisce quindi che il *si* deve essere considerato in funzione impersonale, e quindi il verbo va coniugato alla terza persona singolare, come in D:

- (19) la A è la più sbagliata perché credo che in italiano non si dice si coltivano il grano e la vite (.) si dice si coltiva il grano e la vite perché <ee> mi sembra che altrimenti il grano e la vite si coltivarono a se stessi

La differenza tra B e C apre una finestra di discussione sulla dimensione diafasica e sull'accettabilità dell'accordo del verbo transitivo unicamente con il complemento diretto o, dando una lettura eminentemente impersonale del *si*, anche alla terza persona singolare, costruzione abbastanza diffusa nell'italiano neostandard, almeno orale. Dalle registrazioni non emergono molte riflessioni in questo senso, la discussione sull'accordo e sulla diatesi risulta anzi marginale.

Le ragioni di accordo che sottendono la scelta della studentessa in (19) potrebbero essere considerate valide anche per l'altra apprendente che ha attribuito quattro punti a D, eppure la registrazione mette in luce un'interpretazione altra, che difficilmente avrebbe potuto emergere in assenza della discussione elicitata dall'esercizio:

- (20) la frase più corretta è la D perché nell'azienda dei miei genitori si coltiva il grano e la vite <ehh> perché nell'azienda è singolare e coltiva è la scritta giusta del verbo e la frase più sbagliata è A Nell'azienda dei miei genitori si coltivano il grano e la vite perché coltivano è al plurale e la scritta con la C /ku/ U /u/ltivano

Si tratta di una sorta di *covert error*: l'insegnante potrebbe ipotizzare che la ragione della scelta sia, insieme all'ortografia del verbo, l'accordo del verbo con il clitico in funzione impersonale, invece l'accordo individuato dall'apprendente è tra *azienda* e il verbo.

Solo uno studente dà una spiegazione completa che denota piena consapevolezza delle regole di accordo, anche se non esplicitate

- (21) si dice coltivare (.) allora senz'altro la B è la giusta e per me <e> la persona – per esempio la A io ho messo 3 perché per me è più importante la persona fare la concordanza che <e> – perché la persona ha sbagliato una lettera, soltanto una lettera (.) ma comunque ha mantenuto la concordanza giusta

3.1 Anglofoni e lusofoni a confronto

Alcuni item sono stati sottoposti a gruppi di anglofoni e lusofoni brasiliani dello stesso livello di competenza in L2 (entrambi A2/B1). Si può ipotizzare che la vicinanza tipologica tra le lingue giochi un ruolo determinante nella definizione della competenza metalinguistica degli apprendenti e nella loro capacità di descrivere le scelte, tuttavia i dati a disposizione, pochi al momento per trarre generalizzazioni affidabili, non confermano pienamente le previsioni.

Emblematico è:

- A. Un uomo e una donna si sono fatti prendere una foto
- B. Un uomo e una donna si hanno fatto prendere una foto di loro
- C. Un uomo e una donna si sono fatti fare una foto
- D. Un uomo e una donna ha fatto prendere una foto di loro

L'esercizio presenta due fondamentali difficoltà: l'interferenza lessicale legata alla collocazione 'fare una foto' che sia in inglese (*to take a picture*) che in portoghese (*tirar uma foto*) si costruiscono con il verbo 'prendere', e l'uso del riflessivo unito al verbo causativo.

- (22) this is was difficult question with *prender* with <uhh> fare una foto (.) but then prendere is *to take* and there is also <o>*scattare* so(o) (.) <uh>
[EN]

Come in altri casi è la questione lessicale ad essere il discrimine scelto dagli apprendenti per definire il continuum tra giusto e sbagliato tra i distrattori presentati. Interessante però è notare come rispetto alla collocazione siano gli studenti lusofoni a commettere il maggior numero di errori: mentre gli anglofoni, che d'altra parte si trovano nella situazione privilegiata del bagno linguistico, riconoscono il falso amico e identificano l'unica occorrenza di 'fare una foto' come quella corretta.

Vale la pena riflettere sulla difficile arte di creare buoni distrattori: in questo caso l'aver lasciato una sola occorrenza di 'fare una foto' ha portato gli studenti ad escludere a priori la scelta, in virtù di un'abitudine a confrontarsi con una certa regolarità nella struttura dei distrattori, operando una scelta più dettata dal calcolo delle probabilità che dall'effettivo ragionamento sulla forma della lingua.

Di fronte alla scelta tra *fare* e *prender* una foto gli studenti hanno fatto ricorso alla più classica delle spiegazioni date dai parlanti ingenui e dai non linguisti: il 'suono' o la sensazione:

- (23) io ho scelto come la giusta la B e la più peggiore è la C. è quasi una cacofonia questa frase

Che ci siano delle ragioni dovute alla frequenza e all'esposizione all'input è ovvio, ma solo uno studente – un brasiliano che è stato in Italia – pare esserne consapevole:

- (24) questo è un problema per essere brasiliano perché non sapevo se *si sono fatti* che c'è o *si hanno fatto* <ehh> nelle frasi A, B, C se si era un'espressione idiomatica

dell'italiano. Perché ho pensato di *fare una foto* perché ho sentito questo molto in Italia. *Fare una foto* e no *prendere una foto*.

[PT]

La costruzione riflessiva è stata riconosciuta dalla maggior parte degli apprendenti, indipendentemente dalla L1 di provenienza; in linea con i contenuti tracciati dal Profilo, la funzione del clitico in questo esercizio pare porre meno problemi rispetto a quelli discussi in precedenza.

Sia i lusofoni che gli anglofoni riconoscono la natura del pronome e ne derivano la relazione con il verbo e la selezione dell'ausiliare, tanto che per i brasiliani la scelta di *essere* o *avere* è considerata discriminante nell'attribuzione di 3 o 4 punti, laddove la scelta è fra A e B:

(25) la prima è giusta (.) la B è la 3 perché sbaglia l'ausiliare dovrebbe essere si sono fatti

(26) STUDENTE1: secondo me la corretta *Un uomo e una donna si sono fatti prendere una foto* (.) la B punteggiatura 3, la C 2, e la D 1

STUDENTE2: <uh> la A 4, la B 3, la C 2, la D 1 perché <e> nella B l'ausiliare è sbagliato STUDENTE3: sì perché non si può utilizzare un verbo riflessivo con ausiliare *avere*

STUDENTE2: sì l'ausiliare *avere* nella B è sbagliato

(27) ah qui mi sono ricordato dell'inglese, di questi pronomi che sono/ che si svolgono alle stesse persone riflessivi. Ad esempio un uomo e una donna si sono fatti per loro stessi una foto, così sarebbe per- è la più corretta (.) *si hanno fatto* no, perché quando c'è uno riflessivo deve sempre fare il passato con il verbo *essere* come ausiliare

Il pronome tonico del distrattore B è stato idealmente inserito come possibile variante del riflessivo; non sembra però aver sortito l'effetto previsto, poiché molti apprendenti brasiliani, invece di considerarlo un'alternativa alla costruzione con il *si* (come in (30)), lo sentono come "confuso" (cfr. (28)), mentre gli anglofoni non lo notano nemmeno (solo in (35) vi si accenna indirettamente).

(28) STUDENTE1: nella D *ha fatto prendere una foto di loro* mi sembra un po' confuso

STUDENTE2: sì sì è la più confusa secondo me
[PT]

(29) ci sono un uomo e una donna <eh> così è un plurale e il verbo è coniugato nel singolare *ha fatto* <ehh> credo anche che non si dice *di loro*, io credo che devi avere un verbo <eh> flexivo / riflessivo

[PT]

L'accordo S-V (ausiliare) è pienamente padroneggiato ed esaurientemente spiegato da entrambi i gruppi di apprendenti, che sono in genere concordi nell'assegnare un solo punto alla frase D

- (30) il soggetto è plurale e composto, non so se si dice composto in italiano, <eh><ehm> il verbo è al singolare (.) per questo per me suona più sbagliato [PT]
- (31) e la D è la più sbagliata perché *un uomo e una donna ha fatto* sbaglia la concordanza [PT]
- (32) DOCENTE: quella più sbagliata quale credi che sia?
STUDENTE: <uh> per numero 3 <uhhh> (.) <uhhh> (.) credo<o> D (.) perché<e> D (.) un uomo e una donna *ha fatto* prendere una foto di loro [EN]

Solo un apprendente ipotizza la possibilità di una lettura potenzialmente corretta di D, giustificando l'interpretazione con l'identificazione del soggetto singolare

- (33) lettera D voto 3 *un uomo e una donna ha fatto prendere una foto di loro* (.) se pensare solo alla donna per concordare *ha fatto prendere* forse va bene [PT]

Se la forma dell'ausiliare non crea discussioni, l'accordo del participio passato rivela alcuni dubbi; in particolare in (34) la studentessa brasiliana sostiene che l'accordo debba essere fatto con l'oggetto diretto e non con il soggetto:

- (34) il problema mi sembra che è questo sono fatti, un plurale perché è una foto nella A si sono fatti perché non mi sembra giusto perché è solo una foto

Alcuni studenti, sia anglofoni che lusofoni, poi, hanno dichiarato di avere delle difficoltà legate alla presenza del verbo fattivo, alcuni ne sono consapevoli, altri non riescono a definire in modo chiaro ed esplicito la ragione dei propri dubbi. L'accumulo di verbi unito all'uso riflessivo porta ad una confusione legata alla difficoltà di distinguere il piano grammaticale-funzionale, in cui l'uomo e la donna sono soggetto, e il piano dei ruoli semantici, in cui l'agente è qualcun'altro e l'uomo e la donna sono pazienti.

- (35) but I guess the difficult is also the reflexive *si hanno prendere una foto di loro* (.) so I'm not really sure if a man or a woman took a picture of them or had a picture taken of themselves

A causare perplessità è soprattutto il fatto che la costruzione *far fare qualcosa*, implica la presenza di un agente, che non è esplicitato nelle frasi proposte

- (36) prima ho pensato che sarebbe la C, ma *un uomo e una donna si sono fatti fare una foto* ma per me manca che si sono fatti fare una foto da qualcuno (.) qualcuno ha fatto [PT]
- (37) Secondo me devi usare *fare* invece di *prendere* perché qualcun altro lo deve fare per loro invece ma prendere funziona secondo me con A [EN]

4. Conclusioni

Le registrazioni analizzate offrono una casistica rappresentativa del comportamento degli apprendenti di fronte all'esigenza di dare una spiegazione alle proprie scelte e della relazione che le riflessioni metalinguistiche hanno con l'appropriatezza delle scelte stesse.

I dati inoltre indicano che pochi studenti hanno assegnato un punteggio appropriato e hanno saputo spiegare perché (è il caso, ad esempio, della scelta dell'ausiliare con un verbo riflessivo): in molti danno motivazioni impressionistiche e parziali, ammettendo la propria confusione; in questi casi gli apprendenti sembrano essere ancora in fase di *noticing* e non di pieno possesso delle strutture. Per così dire, non sono ancora arrivati all'*understanding*.

Laddove i distrattori contenevano delle interferenze con la L1 degli apprendenti, sono poi emerse occasioni in cui gli studenti hanno attribuito i punteggi con estrema sicurezza e hanno sostanzialmente le loro scelte in modo convinto, ma le ragioni addotte si sono rivelate fallaci, dato che l'assunto di partenza era sbagliato.

Sul piano metodologico, VALICO si è dimostrato una interessante fonte per elaborare distrattori *corpus-based*, basati quindi sulle reali difficoltà degli apprendenti. Esso fornisce, quindi, materiale autentico facilmente manipolabile nella prospettiva didattica richiesta. Gli esempi riportati sono emblematici in questo senso, poiché mettono bene in luce come la particella *si* con le sue funzioni diverse costituisca in effetti un elemento di difficoltà per gli apprendenti. Se è vero che ai livelli A2/B1 viene richiesta la sola competenza d'uso della funzione riflessiva, è d'altra parte vero che quasi nessuno degli studenti coinvolti nella discussione ha identificato la forma passivante e tanto meno quella impersonale, che invece erano il perno di alcuni dei distrattori proposti; gli apprendenti, e anche un insegnante, hanno invece ricondotto tutti i *si* alla funzione riflessiva, non notando le differenze sintattiche e semantiche del clitico in contesti diversi.

Per concludere, questo primo studio di caso sul rapporto tra competenza linguistica e capacità di dare spiegazioni metalinguistiche ha messo in evidenza come quest'ultima sia strettamente legata all'interiorizzazione del sistema di una lingua. Posti di fronte a minime, ma significative, variazioni, gli studenti sono spesso disorientati e mettono in atto modalità selettive che non emergerebbero in assenza della discussione che segue l'attribuzione dei punteggi. Si pensi ad esempio alle strategie di esclusione stocastica dei distrattori o ai ragionamenti sulla costruzione e sulla formulazione delle scelte. D'altra parte anche gli insegnanti possono trarre benefici da questa attività, riflettendo essi stessi sulla natura delle deviazioni dalla norma, approfondendo così la loro formazione e preparazione di fronte ai problemi posti dai discenti. Stimolare processi interazionali ed elicitarne riflessioni metalinguistiche orientate al *noticing*, favorendo così lo sviluppo dell'*awareness* e dell'*undersanding* è una buona pratica che deve essere integrata nella didattica delle lingue. Farlo attraverso esercizi come quelli presentati permette di elicitarne discussioni che smascherano i *covert error* e verificano la reale corrispondenza del livello degli studenti ai

parametri segnati dai referenziali per le lingue, inoltre facilita il processo di *noticing*, permette agli apprendenti di verificare ipotesi e di consolidare principi già acquisiti.

Bibliografia

- ALANEN R. (1995), Input enhancement and rule presentation in second language acquisition, in SCHMIDT R. (Ed.), *Attention and Awareness in Foreign Language Learning*, University of Hawaii at Manoa, Honolulu: 259-302.
- BARDEL C. - FALK Y. (2012), The L2 status factor and the declarative/procedural distinction, in CABRELLI AMARO J. - FLYNN S. - ROTHMAN J. (eds.), *Third language Acquisition in Adulthood*, John Benjamins, Amsterdam: 61-78.
- BUTLER Y. (2002), Second Language learners' theories on the use of English articles, in *Studies in Second Language Acquisition* 24: 451-480.
- COLOMBO O. (2009), L'apprendimento dell'alterazione in italiano L2 in produzioni scritte di francofoni, in CORINO E. - MARELLO C. (a cura di), *VALICO. Studi di linguistica e Didattica*, Guerra, Perugia: 111-135.
- CORINO E. - MARELLO C. (2011), Der Fehler als Distraktor. Multiple-choice Verfahren und metalinguistische Überlegungen im DaF-Unterricht, in BOSCO COLETSOS S. - COSTA M. - EICHINGER L. (a cura di), *Atti del convegno Tedesco-Italiano: confronti linguistici/Deutsch Italienisch: Sprachvergleiche*, Winter, Heidelberg: 189-207.
- CORINO E. - MARELLO C. (in stampa), Verbi impersonali anzi usi impersonali dei verbi. Microstrutture di dizionari bilingui e monolingui italiani ed esigenze degli apprendenti, in MINERVA N. - SAN VICENTE F. - STURIALE M. (eds), *Norm and Usage in Bilingual Lexicography: 16th-21st Century*, Nodus, Münster.
- ELLIS R. (2001), Introduction: Investigating form-focused instruction, in *Language Learning* 51(1): 1-46.
- ELLIS R. - S. LOEWEN - ERLAM R. (2006), Implicit and Explicit Corrective Feedback and the Acquisition of L2 Grammar, in *SSLA*, 28: 339-368.
- GRASSI R. - MANGIARINI C. (2010), Feedback "implicito" e feedback "esplicito" a confronto: uno studio sperimentale, in GRASSI R. - PIANTONI M. - GHEZZI C. (a cura di), *Interazione didattica e apprendimento linguistico*, Guerra, Perugia: 129-161.
- LONG M.H. (1996), The role of linguistic environment in second language acquisition, in RITCHIE W.C., - BHATIA T.K. (eds.), *Handbook of Language Acquisition: vol. 2. Second Language Acquisition*, Academic Press, New York: 413-468.
- MARELLO C. (2009) Distrattori tratti da corpora di apprendenti di italiano LS/L2, in CORINO E. - MARELLO C. (a cura di), *VALICO. Studi di linguistica e Didattica*, Guerra, Perugia: 177-193.
- PARADIS M. (2009), *Declarative and Procedural Determinants of Second Languages*, John Benjamins, Amsterdam.
- RADWAN A.A. (2005), The effectiveness of explicit attention to form in language learning, in *System* 33: 69-87.

ROBINSON P. (1997), Generalizability and automaticity of second language learning under implicit incidental, enhanced, and instructed conditions, in *Studies in Second Language Acquisition* 19: 223-247.

ROHERK. (2007), Metalinguistic Knowledge and Language Ability in University-Level L2 Learners, in *Applied Linguistics* 29(2): 173-199.

ROSI F. (2010), Il ruolo dell'insegnamento nell'emergere della competenza metalinguistica, in GHEZZI C. - GRASSI R. - PIANTONI M. (a cura di), *Interazione didattica e apprendimento linguistico*, Guerra, Perugia: 79-100.

SCHMIDT R. (1990), The role of consciousness in second language learning, in *Applied Linguistics* 11: 17-46.

SCHMIDT R. (2010), Attention, awareness, and individual differences in language learning, in CHAN W.M. - CHI S. - CIN K.N. - Istanto J. - NAGAMI M. - SEW J.W. - SUTHIWAN T. - WALKER I., *Proceedings of CLaSIC 2010*, Singapore, National University of Singapore, Centre for Language Studies: 721-737.

SCHMIDT R. - FROTA S. (1986), Developing basic conversational ability in a second language: A case study of an adult learner of Portuguese, in DAY R. (ed.), *Talking to learn: Conversation in second language acquisition*, Newbury House, Rowley, MA: 237-326.

SPINELLI B. - PARIZZI F. (2010), *Profilo della lingua italiana. Livelli di riferimento del QCER A1, A2, B1 e B2*, La Nuova Italia, Firenze.

TOMLIN R. - VILLA V. (1994), Attention in cognitive science and second language acquisition, in *Studies in Second Language Acquisition* 16: 183-203.

VEDOVELLI M. (1990), La percezione della standardizzazione nell'apprendimento naturale dell'italiano L2, in BANFI E. - CORDIN P. (a cura di), *Storia dell'italiano e forme dell'italianizzazione*, Atti del XXIII Congresso S.L.I., Bulzoni, Roma: 141-156.

Negli ultimi decenni, il legame tra la linguistica applicata e le scienze tecnologiche dell'informazione e della comunicazione (TIC) è divenuto più forte e articolato. Entrambe le discipline si intersecano contribuendo l'una al potenziamento dell'altra, in un rapporto reciproco e fecondo. La stretta connessione tra le TIC e la linguistica applicata si realizza nella ricerca, nelle pratiche comunicative e nella didattica.

I lavori presentati in questo volume mirano a rispondere a quesiti di sicuro interesse scientifico, ivi inclusi: Quali sono i processi di apprendimento della pronuncia in L2? Qual è la struttura del discorso mediato dal computer? Quali sono le regolarità pragmatiche della comunicazione nei messaggi di posta elettronica? Quali sono le caratteristiche delle nuove forme di interazione nelle reti sociali virtuali? Come devono essere strutturati i dizionari online per apprendenti di L2? Come si può costruire un ambiente di apprendimento collaborativo tramite l'impiego di *social network*? Il volume è di sicuro interesse per la ricerca e per la didattica linguistica. Il ricercatore vi troverà una pluralità di prospettive di analisi che si ascrivono agli ambiti di ricerca di linguistica applicata. Il docente di lingua troverà spunti per innovare l'insegnamento linguistico assegnando alle tecnologie un posto di rilievo al passo con i tempi e con gli interessi dei giovani studenti.

Francesca Bianchi (*biobluob*) è ricercatrice e docente di Lingua inglese presso l'Università del Salento. Interessata da sempre all'uso delle tecnologie nella didattica delle lingue e nella ricerca linguistica, ha collaborato con il CNR di Pisa con L3S Research Center di Hannover alla creazione di ambienti di apprendimento online. Inoltre ha partecipato a progetti di ricerca nazionali e ha pubblicato articoli e curatele che applicano le tecnologie (Web e corpora) alla didattica e alla ricerca e viceversa.

Paola Leone è ricercatrice e docente di Didattica delle Lingue Moderne presso l'Università del Salento. Dal 1990 partecipa a progetti di ricerca nazionali e internazionali relativi all'apprendimento linguistico e alla formazione insegnanti. I suoi interessi di ricerca sono nel campo della comunicazione mediata dal computer e l'apprendimento delle lingue. Ha pubblicato articoli su riviste nazionali e internazionali.



**MONTE
DEI PASCHI
DI SIENA**
BANCA DAL 1472

Volume pubblicato
con il contributo
di Banca Monte dei Paschi
di Siena, erogato tramite
il Dipartimento
di Studi Umanistici
dell'Università del Salento

studi AitLA vol. 4

LINGUAGGIO E APPRENDIMENTO LINGUISTICO: METODI E STRUMENTI TECNOLOGICI

a cura di
Francesca Bianchi - Paola Leone

AitLA - Associazione Italiana di Linguistica Applicata
via Cartoleria, 5 - 40100 Bologna - Italy
email: info@aitla.it | sito: www.aitla.it

Edizione realizzata da
Officinaventuno
via Doberdò, 21 - 20126 Milano - Italy
email: info@officinaventuno.com | sito: www.officinaventuno.com

ISBN: 978-88-97657-13-2