

Scelta del topic e codificazione anaforica nelle composizioni argomentative in ceco L2

Anna Maria Perissutti

1. *Premessa*

Questa ricerca si situa nell'ambito degli studi sull'acquisizione delle abilità testuali e presenta i risultati delle indagini preliminari del progetto Wrilab2, Laboratorio online di lettura e scrittura in ceco, italiano, sloveno e tedesco come L2¹. Essa si concentra su uno degli aspetti ritenuti più complessi nella scrittura in L2², l'organizzazione dell'informazione, e intende verificare eventuali divergenze nella scelta del topic di frase³ e dei mezzi per codificarne la ripresa anaforica in composizioni argomentative scritte in ceco L1 e in ceco L2.

Esiste un discreto numero di studi che si concentra sull'organizzazione dell'informazione in L2 e in particolare sulla parte topicale di essa (Chini 2008, Carroll, von Stutterheim 2003 e altri). Pur mettendo in evidenza tendenze interessanti su cui torneremo nel corso della nostra analisi, questi studi si concentrano in modo pressoché esclusivo sul genere narrativo: basati sul modello della *Quaestio*, essi confrontano la descrizione orale di scene di film raccontate da apprendenti di L2 con quelle prodotte da soggetti L1, mettendone in evidenza le differenze in termini di scelta del topic e di espressione linguistica dello stesso. Gli studi esistenti sull'organizzazione dell'informazione in testi argomentativi scritti in L2 sono dedicati pressoché esclusiva-

¹ Si tratta di un progetto europeo cofinanziato dalla Commissione Europea nell'ambito del Programma per l'Apprendimento Permanente (Lifelong Learning Programme), per il triennio 2014-2016. Il progetto, a cui partecipano nove partner di cinque Paesi europei (Italia, Austria, Germania, Repubblica Ceca e Slovenia), si propone di dar vita a un laboratorio on-line basato su Moodle, dove testi argomentativi in ceco, italiano, sloveno e tedesco vengono smontati, analizzati e ricomposti e gli studenti sono guidati nelle fasi del processo di scrittura, dalla progettazione del testo, attraverso la redazione della bozza, fino alla revisione finale. Tra gli obiettivi di questo progetto c'è l'individuazione delle caratteristiche linguistiche dei testi argomentativi scritti in ceco, italiano, sloveno e tedesco da apprendenti L2 rispetto ai testi dello stesso genere scritti da nativi. Per maggiori informazioni si rimanda al sito <www.wrilab2.eu>.

² A proposito delle aree ritenute maggiormente difficili in L2, si vedano gli studi di Flowerdew 2001, Granger 2004, Mauranen 1993, Carroll *et al.* 2000.

³ Per topic intendiamo, in prima approssimazione, l'argomento su cui verte la frase. Si veda il punto 3 per un trattamento più dettagliato del termine.

mente all'inglese (Ådel 2006, Berry 1995, Crismore 1989, Crismore *et al.* 1993 e molti altri). Non esistono studi sulla composizione di testi argomentativi in ceco come L2.

La nostra indagine prenderà avvio dalla presentazione dei due corpora analizzati (punto 2), per poi passare alla definizione del concetto di topic e ai criteri utilizzati per la sua identificazione (punto 3). Al punto 4 del nostro contributo illustreremo il fenomeno dell'anafora, confrontando la distribuzione dei mezzi linguistici utilizzati per marcare l'anafora. Al punto 5 presenteremo infine le conclusioni della ricerca.

2. Composizioni analizzate

Vista la complessità della ricerca, che richiede un'analisi dettagliata delle scelte topicali in testi argomentativi di apprendenti e nativi, il numero delle composizioni prese in esame è stato piuttosto limitato: si tratta di un corpus di 15 composizioni scritte in ceco L2 da apprendenti di diverse madrelingue, estratte dal corpus CzeSl - *Czech as a Second Language with Spelling, Grammar and Tags*, facente parte del Czech National Korpus⁴. Abbiamo selezionato le composizioni di livello B1 dedicate a un unico argomento, i mass media, e rispondenti ad un unico topic discorsivo, espresso nel titolo *Řídí náš život média?*, 'I mass media condizionano la nostra vita?'. Presentiamo di seguito le caratteristiche degli informanti (cf. TABELLA 1).

Il numero complessivo di parole delle composizioni di ceco L2 da noi esaminate è di 3111. Per poter descrivere e valutare l'interlingua utilizzata dagli scriventi di ceco L2⁵, abbiamo preferito evitare riferimenti astratti alla norma testuale ceca e comparare le composizioni scritte dagli apprendenti con quelle scritte da loro coetanei in ceco L1, estratte dal portale *Český jazyk Literatura aneb studentský underground*, <www.cesky-jazyk.cz>⁶. Le composizioni che abbiamo estratto da questo corpus appartengono al genere *esej* 'saggio breve' e sono dedicate a un tema simile a quello delle composizioni di L2; si tratta in particolare dei seguenti 6 saggi brevi:

⁴ Si tratta di un corpus di ceco come lingua seconda, lemmatizzato, contenente circa un milione di parole. Esso mette a disposizione composizioni originali scritte da studenti di ceco L2 di 54 diverse L1 e con diversi livelli di competenza, da A1 a C2.

⁵ La scelta del gruppo di controllo appare assai più complicata negli studi sull'acquisizione rispetto a quella che si presenta in altri studi basati sui corpora. Qual è infatti la "norma", lo standard con cui comparare l'interlingua dei testi degli apprendenti L2? In letteratura esistono diverse opinioni in merito (per una panoramica sull'argomento si confronti Ådel 2006).

⁶ Si tratta di un portale ricco di materiale sulla lingua e la letteratura ceca, tra cui un elenco di composizioni di diversi generi testuali, scritte in ceco L1 da studenti delle superiori.

L1	Età	Sesso	Codice alfanumerico che identifica le composizioni all'interno del corpus CzeSI
Cinese	21	Femminile	AA_CJ_003
Cinese	21	Femminile	AA_CK_003
Cinese	21	Femminile	AA_CH_003
Cinese	21	Maschile	AA_CZ_003
Cinese	17	Maschile	AA_CV_003
Cinese	17	Femminile	AA_CL_003
Giapponese	26	Femminile	SVOBJ_KO_009
Ungherese	21	Femminile	AA_EK_003
Russo	25	Maschile	KAR_VP_007
Russo	18	Maschile	KAR_AD_001
Russo	17	Femminile	KAR_AK_007
Polacco	21	Femminile	AA_DS_004
Polacco	21	Femminile	AA_DS_003
Inglese	22	Femminile	AA_KB_002
Tedesco	21	Femminile	AA_JB_003

Tabella 1.

Caratteristiche degli autori delle composizioni L2 analizzate, tratte dal corpus CzeSI

- *Sdělovací prostředky (1)* ‘Mezzi di comunicazione di massa (1)’;
- *Sdělovací prostředky (2)*, ‘Mezzi di comunicazione di massa (2)’;
- *Televize – kamarád nebo zloduch?* ‘La televisione, entità amica o maligna?’;
- *Reklama a její vliv na nás*, ‘La pubblicità e la sua influenza su di noi’;
- *Lže nám televize?* ‘La televisione ci inganna?’;
- *Konzumní společnost*. ‘La società dei consumi’.

Il numero complessivo di parole del corpus di ceco L1 è di 3106, dunque paragonabile per dimensioni a quello di L2 analizzato.

La ricerca basata sui due corpora appena presentati si propone di individuare i mezzi utilizzati dal gruppo di madrelingua cechi e da quello di ceco L2 per codificare linguisticamente la ripresa anaforica dei topic. Prima di procedere

con l'analisi, ci sembra essenziale fornire la definizione di topic da cui siamo partiti e chiarire le modalità che abbiamo utilizzato per identificare i topic di frase nelle composizioni dei due gruppi presi in esame.

3. *Definizione del concetto di topic e fissazione dei criteri utilizzati per la sua identificazione*

La nozione di topic è notoriamente una delle più dibattute in letteratura⁷. Da alcuni questa nozione è intesa come una categoria pragmatica universale (Lambrech 1994), da altri in chiave semantica-cognitiva (Maslova, Bernini 2006) o come una testa funzionale alla periferia sinistra della frase (Rizzi 1997). Diverse sono le posizioni in merito ai criteri per la definizione di questa nozione anche in campo funzionalista (Dittmar 1992). Nel nostro contributo ci riferiamo al topic di frase nell'accezione di Lambrecht (1994: 147), distinguendo innanzitutto il topic in quanto referente, cioè l'elemento su cui si forniscono informazioni con la frase, dall'espressione del topic, costituente che codifica il referente topicale. Per topic intendiamo dunque il referente su cui verte l'enunciato e per il quale quanto vi si afferma è rilevante. A livello cognitivo, le caratteristiche del topic sono l'identificabilità, lo stato di attivazione alto nella memoria, l'accessibilità (si vedano Chafe 1976, Givón 1983); a livello informativo invece il topic si caratterizza come dotato di scarso dinamismo comunicativo⁸.

Nella nostra indagine abbiamo analizzato le scelte topicali a livello locale, di clausola, le unità sintattiche e informative di base del testo. Riteniamo infatti con Chini 2008 che il livello di clausola consenta di cogliere possibili elementi fini, concettuali e formali, del tessuto e della logica testuale. I criteri da noi adottati per identificare i topic di frase sono i seguenti:

1. il topic di una clausola è il referente su cui il parlante intende fornire informazioni per merito della clausola stessa. Ferrari (2009: 107) parla a questo proposito di relazione di "aboutness" e cita l'esaustiva definizione di Lambrecht (1994: 131), a cui si deve l'introduzione del termine di "aboutness":

In a given situation the proposition is construed as being about this referent, i.e. as expressing information which is relevant to and which increases the addressee's knowledge of this referent.

⁷ Diversi e non sempre equivalenti i termini utilizzati in letteratura per esprimere il concetto: si parla di "soggetto psicologico", "tema", "topic". Per una sintesi si veda Berretta 1995.

⁸ Pur essendo consapevoli della possibile valenza topicale di altri elementi, quali i circostanziali di tempo e di luogo che inquadrano l'evento e le entità di riferimento, il topic per noi sarà rappresentato unicamente dai partecipanti primari degli eventi.

La relazione di “aboutness” che definisce il topic non va intesa in senso generico, ma si applica ad uno o più referenti comunicativamente speciali, cioè di cui il contesto richiede, per pertinenza comunicativa, un arricchimento semantico.

2. il referente topicale è accessibile, identificabile, attivo, di solito espresso da costituenti non salienti. Non possono essere topic referenti nuovi, anche aventi il ruolo di soggetto. È il caso delle frasi identificative (1) e di quelle presentative (2)⁹:

(1) *Kennedy umřel!*
 Kennedy^{NOM} morire^{PST.PFV.3SG}
 ‘Kennedy è morto!’.

(2) *Dvě nehody s cyklisty.*
 Due incidenti^{NOM.PL.F} con ciclisti^{INS.PL.M}
 ‘Due incidenti con ciclisti’.

Essendo il contenuto delle frasi (1) e (2) tutto nuovo, si tratta di frasi senza topic;

3. laddove siano plausibili più candidati al ruolo di topic, viene privilegiato il topic con il ruolo di soggetto, come nella frase (3) dove il topic principale è *Pavel* mentre *Hana* è topic secondario:

(3) ***Pavel** si vzal Hanu, i když ji*
 Pavel^{NOM} si^{REFL} prendere^{PST.PFV.3SG} Hana^{ACC} anche se lei^{ACC}
 nemiloval.
 non amare^{PST.IPFV.3SF}
 ‘**Pavel** sposò Hana anche se non l’amava’.

In casi come:

- (4) “X passeggia” = X è topic se già introdotto in precedenza
 (5) “Y vede X” = X è topic, se Y è nuovo, non accessibile
 (6) “Y chiama X” = Y è topic perché è soggetto accessibile e attivo, mentre X è topic secondario
 (7) Nella struttura passiva “X è chiamato da Y” = X è topic

⁹ Come spiega Andorno (1999: 206), mentre le strutture identificative individuano il referente a cui il predicato si riferisce, le strutture presentative rispondono a una domanda generica come *che cosa succede/è successo?*.

Proponiamo di seguito una sequenza testuale tratta dalla composizione SVOBJ_KO_009 del corpus di ceco L2, *CzeSL*; la sequenza presentata all'esempio (8) evidenzia la segmentazione del testo in clausole, segmentazione da noi utilizzata per l'analisi delle composizioni. Il topic di frase è evidenziato in grassetto nell'originale ceco e sottolineato nella traduzione italiana:

(8)	<i>Zaprvé</i>	<i>určuji</i>	<i>média</i>	<i>náš</i>	<i>životní styl.</i>	
	Innanzitutto ^{AVV}	definire ^{PRS.3PL}	mass media ^{NOM.3PL}	nostro	di vita	stile ^{ACC}
	<i>Každý den</i>	(my)	<i>dostáváme</i>	<i>různé</i>	<i>reklamy</i>	
	Ogni giorno	(SBJ Ø noi)	ricevere ^{PRS.1PFV.1PL}	diverse	pubblicità ^{ACC}	
	<i>Přes</i>	<i>televizi</i>	<i>nebo</i>	<i>časopisy.</i>		
	attraverso	televisione ^{ACC}	o	riviste ^{ACC}		
	<i>Média</i>	<i>nám</i>	<i>tak</i>	<i>ukážou,</i>	<i>co</i>	<i>je</i>
	Massmedia ^{NOM.PL}	noi ^{DAT}	così	dimostrare ^{PRS.PFV.3PL}	cosa	essere ^{PRS.3SG}
	<i>průměrní</i>	<i>životní</i>	<i>styl.</i>			
	medio	di vita	stile ^{NOM}			

(SVOBJ_KO_009)

'Ogni giorno **(SBJ Ø noi)** riceviamo diverse pubblicità attraverso la televisione o le riviste. **I massmedia** così ci mostreranno qual è lo stile di vita medio'

Dopo aver chiarito la definizione di topic e i criteri che abbiamo utilizzato per identificare i topic di frase nelle composizioni di ceco L1 e L2, passiamo ad analizzare la distribuzione dei mezzi linguistici utilizzati dai due gruppi per codificare la ripresa anaforica del topic. La sezione seguente sarà appunto dedicata a questo argomento.

4. Codificazione dell'anafora in ceco L1 e L2

In questa sezione intendiamo illustrare i mezzi linguistici con cui i due gruppi analizzati marcano la ripresa anaforica, cioè la ripresa di un referente già introdotto nel discorso con un'espressione antecedente.

L'anafora è stata riconosciuta fin dagli inizi della linguistica testuale come uno degli elementi costitutivi del testo: per costruire un testo coerente e coeso, lo scrivente deve infatti fornire informazioni sufficientemente chiare al lettore sullo sviluppo dei contenuti testuali e sul movimento referenziale, utilizzando espressioni adeguate al contesto, alle conoscenze dell'interlocutore e a quelle condivise, oltre che conformi alla grammatica della lingua utilizzata. Come hanno dimostrato numerosi studi di linguistica testuale a partire dagli anni '80, (Givón 1983, Fox 1987, Tomlin 1987, Berretta 1990), vi è un'interazione tra la strutturazione dei testi e il materiale delle espressioni anaforiche. Uno dei pri-

mi ad interessarsi del materiale anaforico in prospettiva testuale è stato Givón (1983), che ha esposto l'importante principio della iconicità, inerente la correlazione tra l'accessibilità del topic e la codificazione dell'anafora. Secondo questo principio, più i referenti testuali sono accessibili, continui e prevedibili, più leggeri sono i mezzi linguistici con cui ci si riferisce ad essi (mezzi pronominali, ellissi); più discontinui, nuovi, imprevedibili sono i referenti testuali a cui ci si riferisce, più pesanti sono i mezzi utilizzati (SN pieni, costruzioni focalizzate)¹⁰. Determinante per l'accessibilità topicale e quindi per il materiale anaforico utilizzato sono, secondo Givón 1983, soprattutto la distanza lineare dall'antecedente, cioè il numero di confini di frase tra antecedente e anafora, e l'interferenza semantica, cioè la presenza di altri elementi testuali che rendono necessaria l'esplicitazione lessicale dell'anafora. Come ha dimostrato Berretta (1990) per l'italiano, sulla scelta del mezzo linguistico per codificare l'anafora incide anche la strutturazione del testo in unità o sequenze testuali: in posizione iniziale di una sequenza testuale vengono tendenzialmente usate anafore marcate (SN pieni, pronomi tonici), cioè materiale lessicalmente indipendente, mentre in posizione non iniziale di sequenza testuale vengono tendenzialmente usate anafore non marcate (segno Ø, accordo verbale, pronomi clitici), semanticamente dipendenti ed esprimenti una maggiore coesione con il contesto precedente.

L'anafora e la sua codificazione linguistica rappresentano dunque fenomeni molto complessi, realizzati in maniera specifica da ogni lingua e le cui tendenze dipendono dal registro utilizzato e dalla tipologia testuale. È ovvio allora che l'acquisizione di questi fenomeni in L2 richieda notevoli competenze linguistiche, stilistiche e pragmatiche, competenze la cui acquisizione nel caso del ceco L2 non viene favorita dai manuali e dalle grammatiche di ceco per stranieri che non offrono spiegazioni in merito. Ci è sembrato perciò interessante proporre un sondaggio dei mezzi linguistici utilizzati dal gruppo di ceco L1 e da quello di ceco L2 nei testi analizzati.

In ambito boemistico si sono occupati della questione Mathesius (1947), Zimová (1994), Esvan (1996), Trovesi (2004) ed altri.

In ceco l'utilizzo delle forme per il riferimento a persone e entità nei testi risente delle caratteristiche tipologiche e strutturali della lingua, a soggetto pronominale non obbligatorio o pro-drop, con ordine di base SVO, ma sensibile alla struttura dell'informazione, all'articolazione topic-comment: per introdurre nel discorso un referente nuovo, si ricorre a SN pieni, indefiniti (a meno che non si tratti di referenti unici nel contesto), collocati in posizione postverbale di focus; non avendo articoli, marche morfologiche specializzate nell'indicazione della definitezza/indefinitezza del referente, la posizione del referente introdotto da segno Ø nella frase svolge un ruolo essenziale.

Per il riferimento a un referente già introdotto si ricorre invece alla posizione preverbale. A seconda del ruolo sintattico della ripresa, della distanza dalla

¹⁰ Si parla di anafore non marcate per le riprese leggere (pronomi clitici, accordo grammaticale, soggetto Ø di verbi finiti) e di anafore marcate per quelle pesanti (SN pieni, costruzioni focalizzate).

menzione precedente dello stesso referente e delle scelte stilistiche operate dagli scriventi, la ripresa anaforica di un referente Dato può essere realizzata dai seguenti mezzi coreferenziali:

- mezzi vuoti, cioè l'accordo soggetto-verbo e l'anafora zero;
- mezzi pronominali, clitici e tonici;
- mezzi lessicali.

Nella TABELLA 2 esponiamo i risultati del confronto tra i mezzi utilizzati per marcare la ripresa anaforica di un referente Dato nelle composizioni di ceco L1 e in quelle di ceco L2 da noi analizzate. Per la terminologia utilizzata rinviamo allo studio di Chini *et al.* (2003).

Confrontiamo le scelte linguistiche del gruppo degli studenti di ceco L2 con quelle del gruppo dei nativi.

4.1. I mezzi coreferenziali vuoti

Per quanto riguarda i mezzi coreferenziali vuoti (accordo soggetto-verbo, anafora zero), notiamo che essi sono presenti sia in ceco L1 che in ceco L2. Gli apprendenti esaminati dunque mostrano di aver acquisito l'uso coreferenziale dell'accordo soggetto-verbo, dove il soggetto pronominale è inespresso (pronomi zero), strategia anaforica tipica del ceco. Tuttavia, nelle composizioni degli apprendenti di madrelingua cinese, si nota un uso errato dei mezzi coreferenziali vuoti, in contesti dove la lingua ceca impone l'uso di un mezzo anaforico marcato di natura pronominale, come nell'esempio (9):

(9)	Reklama	je	všude	na naši	světě.
	Pubblicità ^{NOM}	essere ^{PRS 3SG}	ovunque	in nostro	mondo ^{LOC}
	<i>Nejenom</i>	Ø	<i>vidíme</i>	<i>z televize,</i>	<i>z Internetu,</i>
	Non solo	Ø	vedere ^{PRS.1PL}	da televisione ^{GEN} ,	da Internet ^{GEN}
	<i>ale také</i>	<i>na ulici,</i>	<i>v knize</i>	<i>i na dveře</i>	<i>metra.</i>
	ma anche	in strada ^{LOC}	in libro ^{LOC}	e su porta ^{ACC}	metro ^{GEN}

(AA_CK_003)

‘La pubblicità è ovunque nel nostro mondo. Non solo [la] vediamo in televisione, in Internet, in un libro ma perfino sulle porte del metrò.’

Nell'esempio (9) l'uso del segno Ø è errato perchè esso sostituisce un elemento pronominale in posizione sintattica di oggetto mentre il sostantivo antecedente con cui esso coriferisce è in posizione di soggetto.

Avendo notato questa tendenza all'uso errato del segno Ø in funzione anaforica solo nelle composizioni dei madrelingua cinesi, supponiamo che essa possa essere dovuta all'influenza del sistema testuale della L1.

TIPO	SOTTOTIPO	CECO L1	CECO L2- livello B1
MEZZI VUOTI	SEGNO Ø	40	53
MEZZI PRONOMINALI CLITICI	PRONOMI PERSONALI CLITICI	12	7
MEZZI PRONOMINALI TONICI	TEN	10	4
	TENTO	4	0
	TENHLE	2	0
	ON	4	0
	KTERÝ	8	20
	JENŽ	8	0
	COŽ	7	
MEZZI LESSICALI	STESSO SN	20	93
	SN SINONIMO, IPERONIMO, IPO- NIMO	46	20
	TEN + STESSO SN	0	4
	TEN+ SN SINONIMO, IPERONI- MO, IPONIMO	7	3
	TENTO + STESSO SN	3	
	TENTO + SN SINONIMO, IPERO- NIMO, IPONIMO, PARAFRASANTE	22	6
	VÝŠE ZMNIŇENÝ, UVEDENÝ, ZPRVE POLOŽENÝ	9	0

Tabella 2.

Numero di occorrenze dei mezzi coreferenziali utilizzati per la ripresa anaforica di un referente in ceco L1 e L2

4.2. I mezzi pronominali clitici

Questi mezzi, che hanno un grande potere coesivo e sono particolarmente adatti alle riprese anaforiche, sono usati più frequentemente nei testi dei nativi cechi. Nel corpus da noi analizzato si tratta in particolare del pronome personale clitico di terza persona singolare (*ho*^{ACC.M}, *je*^{ACC.M/N}, *ji*^{ACC.F}) e plurale (*je*^{ACC.PL}), all'accusativo e ai casi obliqui (*mu*^{DAT.M}, *jí*^{DAT.F}, *jím*^{DAT.PL}, *ho*^{GEN.M}, *jí*^{GEN.F}, *jich*^{GEN.PL}, *něm*^{LOC.M}, *ní*^{LOC.F}, *nich*^{LOC.PL}, *jím*^{STR.M}, *jí*^{STR.F}, *jími*^{DAT.PL}). In (10) forniamo un esempio di questo uso:

- (10) *Zaměřím* -li se na **noviny**, mohu **je**
 concentrare^{PRS.PFV.1SG} -se REFL su **giornali**^{ACC} potere^{1SG} **essi**^{ACC}
rozdělit do dvou skupin: bulvár
 dividere^{INF.PFV} in due gruppi^{GEN.PL}: stampa scandalistica
a seriózní tisk.
 e seria stampa^{NOM.SG}

(Sdělovací prostředky 1)

‘Volgendo l’attenzione ai **giornali**, posso dividerli in due gruppi: le testate scandalistiche e quelle serie’.

La frequenza d’uso dei pronomi clitici nelle composizioni degli apprendenti è inferiore rispetto a quella delle composizioni dei nativi (7 occorrenze contro le 12 presenti nei testi dei nativi). Questo elemento è piuttosto importante ed è una spia della minore coesione dei testi degli apprendenti.

4.3. I mezzi pronominali tonici

Circa l’uso dei mezzi pronominali tonici nei due gruppi analizzati, mentre nei testi dei madrelingua cechi abbiamo riscontrato 43 occorrenze, in quelli degli apprendenti solo 24. Oltre alla maggiore frequenza di questi mezzi nei testi dei madrelingua, spicca anche la varietà di mezzi pronominali usati da essi rispetto alla monotonicità di quelli utilizzati dagli apprendenti. Limitandoci ai pronomi dimostrativi, notiamo in ceco L1 la concorrenza del pronome personale *on* ‘egli’ (*ona* ‘ella’, *ono* ‘esso’) con il dimostrativo *ten* ‘quel’, (*ta* ‘quella’, *to* ‘quello’). Tale concorrenza è regolata da principi testuali piuttosto sottili, di cui si occupa in modo molto dettagliato Zimová (1994), ma che non vengono trattati dalle grammatiche per stranieri. Secondo l’autrice, mentre il pronome *ten* viene tendenzialmente utilizzato per la ripresa di un referente vicino, il pronome personale *on* è utilizzato per la ripresa di un referente distante. Questo fenomeno è chiarito dall’esempio seguente (11), tratto dalla composizione del corpus L1 *Konzumní společnost*:

- (11) *Jak již bylo řečeno, v základu pyramidy leží*
 Come già AUX.PST.3SG dire^{PASS.PST} in base^{LOC} piramide^{GEN} giacere^{PRS.3PL}
potřeby **fiziologické.** **O těch** *již bylo*
 bisogni^{NOM.PL} fisiologici^{NOM.PL} Di questi^{LOC.PL} già AUX.PST.3SG
prohlášeno, že je konzumní společnost
 affermare^{PASS.PST} che essi^{ACC.PL} consumistica^{NOM} società^{NOM}
uspokojuje více než dostatečně [...]
 soddisfare^{PRS.IPFV.3SG} più che sufficientemente

‘Come è già stato detto, alla base della piramide vi sono i bisogni fisiologici. **Di questi** è già stato affermato che **essi** vengono soddisfatti in maniera più che soddisfacente dalle società dei consumi [...]’

Nell’esempio (11) notiamo appunto che mentre il pronome *ten* (*o těch*) realizza la ripresa dell’espressione antecedente ravvicinata (*potřeby fyziologické* ‘i bisogni fisiologici’), il pronome *on* (declinato nell’esempio 11 all’accusativo *je*), realizza la ripresa a distanza dello stesso referente (*potřeby fyziologické* ‘i bisogni fisiologici’). Si tratta tuttavia, come specifica Zimová (1994), di una semplice tendenza.

Per quanto riguarda i pronomi relativi, mentre gli apprendenti usano solo il pronome *který* ‘il quale’ i nativi usano anche *jenž* ‘il quale’ e il pronome relativo *což* ‘che’, che coriferisce a interi enunciati precedenti. Osserviamo l’esempio seguente, tratto dalla composizione del corpus L1 *Lže nám televize?*:

- (12) [...] *křivka* *znázorňující* *počet* *vražd* *ve Spojených státech* *a*
 curva^{NOM} indicante numero omicidi^{GEN} in Stati Uniti^{LOC} e
Kanadě *svým* *strmým* *nárůstem* *sleduje* *křivku*
 Canada^{LOC} propria^{INS} decisa^{INS} crescita^{INS} seguire^{PRS.IPFV.3SG} curva^{ACC}
růstu *počtu* *televizorů,* ***což*** *by* *mělo* *být*
 aumento^{GEN} numero^{GEN} televisioni^{GEN PL} ***che***^{COND} dovere^{PST.3SG} essere
bráno *jako* *velice* *alarmující* *zjištění [...]*
 considerare^{PASS.PST.3SG} come molto allarmante dato

‘[...] la curva che indica il numero degli omicidi negli Stati Uniti e in Canada segue con la propria brusca crescita la curva dell’aumento del numero dei televisori, **cosa che** dovrebbe essere considerata come un dato allarmante [...]’

Notiamo che il pronome relativo *což* ‘cosa che’ riprende l’intero contesto della frase precedente.

4.4. I mezzi lessicali

Le divergenze maggiori nell’uso dei mezzi anaforici nelle composizioni dei due gruppi analizzati emergono nell’ambito dei mezzi lessicali. Per gli apprendenti appare problematica la reintroduzione di referenti Dati non accessibili, perché la menzione precedente è distante o perché ci sono altri referenti in competizione: in questo caso essi non utilizzano i determinatori o gli aggettivi necessari per richiamare un referente già noto (*vyše zmíněný* ‘succitato’, *onen* ‘quel’, *nahoře uvedený* ‘di cui sopra’) ma si limitano a riprendere il referente con lo stesso SN pieno usato nell’antecedente. Abbiamo riscontrato 93 occorrenze di questo tipo nel gruppo degli apprendenti contro 20 del gruppo dei nativi.

Il settore paradossalmente più problematico in ceco L2 risulta però essere quello delle forme anaforiche per referenti massimamente accessibili, in quanto presenti nella clausola precedente, talora addirittura con lo stesso ruolo sintattico di soggetto o di oggetto. Mentre i nativi tendono a non utilizzare lo stesso SN pieno per riprendere un referente antecedente, accessibile, e ricorrono all'uso di sinonimi, iperonimi, iponimi o di sostantivi che parafrasano il significato dell'antecedente, gli apprendenti invece tendono a riprendere il referente antecedente utilizzando lo stesso SN pieno espresso nell'antecedente stesso. Per permettere di cogliere il fenomeno, presentiamo la sequenza di un testo L1 (esempio 13), tratto dalla composizione *Sdělováci prostředky 2*, e la sequenza di un testo di ceco L2, esempio (14):

- (13) ***Sdělovací prostředky*** vznikly už před miliony let
 Mass media^{NOM.PL} sorgere^{PST.PFV.3PL} già prima milioni^{INS} anni^{GEN.PL}
v pravěku, kde se začala rozvíjet řeč.
 in antichità^{LOC} dove^{REFL} iniziare^{PST.PFV.3SG} sviluppare^{INF.IPFV} linguaggio^{NOM.}
Tento způsob sdělování informací se
Questo modo comunicazione^{GEN} informazioni^{GEN.PL} REFL
vyvinul z potřeby pravěkého člověka se
 sviluppare^{PST.PFV.3SG} da^{REFL} bisogno^{GEN} antico^{GEN} uomo^{GEN} REFL
dorozumět [...]
 comunicare^{INF.IPFV}

‘I mezzi di comunicazione di massa sorsero milioni di anni fa nel mondo antico quando iniziò a svilupparsi il linguaggio. Questo sistema di trasmissione dell’informazione si sviluppò dal bisogno dell’uomo antico di comunicare [...]’

Notiamo in (13) che il referente introdotto dall’antecedente ed espresso dal SN *sdělovací prostředky* ‘mezzi di comunicazione di massa’, viene ripreso nel testo successivo dal SN pieno *tento způsob sdělování informací* ‘questo sistema di comunicazione dell’informazione’. Osserviamo invece una tipica sequenza testuale delle composizioni del corpus di ceco L2:

- (14) ***Média*** pro nás jsou absolutně
 Mass media^{NOM.PL} per noi^{ACC} essere^{PRS.3PL} assolutamente
nezbytné. To je hlavní důvod,
 indispensabili. Questo^{NOM.3SG} essere^{PRS.3SG} principale motivo
proč můžou média řídit náš
 perchè^{REFL} potere^{PRS.3PL} mass media^{NOM.PL} governare^{INF.IPFV} nostra^{ACC}
život.
 vita^{ACC}

I mass media per noi sono assolutamente indispensabili. Ciò rappresenta il motivo principale per cui i **mass media** possono condizionare la nostra vita’.

Notiamo in (14) che il referente introdotto dall’antecedente ed espresso dal SN *média* ‘mezzi di comunicazione di massa’, viene ripreso nel testo successivo dallo stesso SN. Questa tendenza, evidente soprattutto nelle progressioni tematiche a tema costante (le preferite dal gruppo di apprendenti), dà luogo a catene anaforiche ridondanti e monotone.

4.5. I determinatori definiti nella ripresa di un antecedente accessibile

Volgiamo ora lo sguardo alla ripresa di un referente vicino, realizzata da SN pieni preceduti da determinatori definiti *ten* ‘quello’, *tento* ‘questo’.

La distribuzione di questi determinatori nel riferimento anaforico appare regolata da sottili norme testuali che non vengono trattate dai manuali e dalle grammatiche di ceco per stranieri.

Come spiega Zimová (1994: 33), nella anafora (o riferimento endotestuale) il determinante *tento* ‘questo’ è portatore del segno “vicino al parlante” e *ten* ‘quello’ è portatore del segno “lontano dal parlante”. Tuttavia in certe condizioni contestuali si può giungere alla neutralizzazione dell’opposizione.

Nel nostro corpus L1 i nativi tendono ad utilizzare il determinante *tento* (si veda l’esempio 13), mentre nello stesso contesto gli apprendenti di ceco L2 tendono ad usare *ten* (esempio 15):

- (15) *Dnes máme reklamu všude a média*
 Oggi avere^{PRS.1PL} pubblicità^{ACC} ovunque. e **massmedia**^{NOM.3PL}
hraje větší roli než dříve. Jsme vždycky
 svolgere^{PRS.IPFV.3SG} maggiore ruolo^{ACC} di prima. Essere^{PRS.1PL} sempre
obklopováni s televizí, radiem nebo internetem.
 circondati^{PASS.IPFV.1PL} con televisione^{INS} radio^{INS} o internet^{INS}
Nemůžeme skoro žít bez těch moderních věcí [...]
 NEG-potere^{PRS.1PL} quasi vivere^{INF.IPFV} senza **quelle moderne cose**^{GEN.PL} [...]

‘Oggi abbiamo pubblicità ovunque e i mass media hanno un ruolo maggiore di prima. Siamo sempre circondati dalla televisione, dalla radio o da Internet. Non possiamo quasi vivere senza quelle moderne cose [...].’

Notiamo poi l’assenza nel corpus di ceco L2 di fenomeni di ripresa lessicale più complessi che si manifestano in L1, in cui il SN antecedente viene ripreso utilizzando:

1. SN aventi un significato generico (quali ad esempio *otázka* ‘questione’ *věc* ‘cosa’, *směr* ‘tendenza’, *skutečnost* ‘realtà’) che coriferiscono non solo verso il SN precedente ma anche verso intere porzioni di testo, come nell’esempio (16):

(16)	[...] <i>křivka</i>	<i>znázorňující</i>	<i>počet</i>	<i>vražd</i>	<i>ve Spojených státech</i>	
	curva ^{NOM}	indicante	numero	omicidi ^{GEN}	in Stati Uniti ^{LOC}	
	<i>a Kanadě</i>	<i>svým</i>	<i>strmým</i>	<i>nárůstem</i>	<i>sleduje</i>	<i>křivku</i>
	e Canada ^{LOC}	propria ^{INS}	decisa ^{INS}	crescita ^{INS}	seguire ^{PRS.IPFV.3SG}	curva ^{ACC}
	<i>růstu</i>	<i>počtu</i>	<i>televizorů,</i>	což	<i>by mělo</i>	
	aumento ^{GEN}	numero ^{GEN}	televisioni ^{GEN PL}	che	COND	dovere ^{PST.3SG}
	<i>být</i>	<i>bráno</i>	<i>jako</i>	<i>velice</i>	<i>alarmující</i>	<i>zjištění [...]</i>
	essere	considerare ^{PASS.PST.3SG}	come	molto	allarmante	dato
	<i>Smutnou</i>	<i>skutečností</i>	<i>dneška</i>	<i>se</i>	<i>staly</i>	<i>i</i>
	Triste^{INS}	realtà^{INS}	oggi	REFL	diventare ^{PST.PFV.3PLF}	anche
	<i>reklamní spoty</i>	<i>vysílané</i>		<i>den</i>	<i>za dnem</i>	
	le pubblicità ^{NOM}	trasmettere ^{PSS.PST.PL.F}		giorno	dopo	giorno

‘[...] la curva che indica il numero degli omicidi negli Stati Uniti e in Canada segue con la propria brusca crescita la curva dell’aumento del numero dei televisori, cosa che dovrebbe essere considerata come un dato allarmante [...]. **Una triste realtà** oggi è rappresentata anche dalle pubblicità trasmesse giorno dopo giorno...’.

2. SN definiti che, oltre a svolgere la funzione di ripresa, offrono nuove informazioni sul tema: si tratta del fenomeno del diatema, che richiede una grande padronanza lessicale e stilistica per essere realizzato:

(17)	<i>Autor</i>	<i>zde</i>	<i>publikuje</i>	<i>svou</i>	<i>novou</i>	<i>báseň.</i>
	Autore ^{NOM}	qui	pubblicare ^{PRS.IPFV.3SG}	propria ^{ACC}	nuova	poesia ^{ACC}
	<i>V</i>	<i>lyrickém</i>	<i>obrazě</i>	<i>podzimní</i>	<i>přírody</i>	<i>nechává</i>
	In	lirico ^{LOC}	immagine ^{LOC}	autunnale ^{GEN}	natura ^{GEN}	CAUS.PRS.IPFV.3SG
	<i>promlouvat...</i>					
	parlare ^{INF.IPFV}					

‘Qui l’autore pubblica la sua nuova poesia. **Nell’immagine lirica della natura autunnale** fa parlare [...]’

Il maggiore ricorso degli apprendenti di ceco L2 a strategie di tipo lessicale piuttosto che a quelle di tipo morfologico usate dai nativi potrebbe essere legato alla tendenza delle interlingue alla sovraesplicitzza, radicata nel desiderio di evitare fraintendimenti e di far passare in modo chiaro il messaggio. Questa eccessiva ridondanza potrebbe anche dipendere dal diverso tipo di pianificazione

testuale che negli apprendenti sembra essere più locale. Come spiegano Chini *et al.* 2003 infatti, gli apprendenti di L2, dovendo tenere sotto controllo problemi di diverso livello (lessicale e morfosintattico soprattutto) in seguito alla loro competenza carente, finiscono per segmentare il testo in unità sintattiche relativamente autonome e dunque più facilmente gestibili ma, in quanto autonome, più bisognose di forme di ripresa esplicite e pesanti.

5. Conclusioni

L'obiettivo del nostro contributo consisteva nell'identificare i mezzi linguistici utilizzati nel corpus dei madrelingua cechi e in quello degli apprendenti di ceco L2 per marcare la ripresa anaforica. La nostra analisi ha permesso di rilevare le seguenti differenze:

- circa l'uso dei mezzi coreferenziali vuoti (accordo soggetto-verbo, anafora zero), abbiamo notato che essi sono usati sia dai nativi che dagli apprendenti. Il gruppo di apprendenti cinesi (il sottogruppo più numeroso nel nostro corpus di ceco L2) estende l'anafora zero in contesti che in ceco richiedono un mezzo anaforico pronominale, perché il referente da riprendere è in posizione sintattica diversa. Questo dato fa pensare all'influsso della L1;
- circa i mezzi pronominali clitici, importanti per la coesione del testo, essi sono più frequenti nei testi dei madrelingua cechi (12 occorrenze contro le 7 degli apprendenti);
- circa i mezzi pronominali tonici, essi sono più frequenti nel corpus dei madrelingua (43 occorrenze contro le 20 degli apprendenti); oltre alla maggiore frequenza, spicca anche nei testi dei nativi la maggiore varietà di questi mezzi. Gli apprendenti usano esclusivamente il pronome *ten* 'quello' o il pronome relativo *který* 'il quale';
- per quanto riguarda invece i mezzi lessicali, mentre i nativi tendono a non utilizzare lo stesso SN pieno per riprendere un referente antecedente, accessibile e ricorrono all'uso di sinonimi, iperonimi, iponimi o di sostantivi che parafrasano il significato dell'antecedente, gli apprendenti invece tendono a riprendere il referente antecedente utilizzando lo stesso SN pieno espresso nell'antecedente stesso.

Bibliografia

- Ädel 2006: A. Ädel, *Metadiscourse in L1 and L2 English*, Amsterdam-Philadelphia 2006.
- Andorno 1999: C. Andorno, *Dalla grammatica alla linguistica. Basi per uno studio dell'italiano*, Torino 1999.
- Berretta 1990: M. Berretta, *Catene anaforiche in prospettiva funzionale: antecedenti difficili*, "Rivista di Linguistica", II, 1990, 1, pp. 91-120.
- Berretta 1995: M. Berretta, *Ordini marcati dei costituenti maggiori di frase: una rassegna*, "Linguistica e Filologia", 1995, 1, pp. 125-170.
- Berry 1995: M. Berry, *Thematic Options and Success in Writing*, in: M. Ghadessy (a cura di), *Thematic Development in English Texts*, London 1995, pp. 55-84.
- Carroll et al. 2000: M. Carroll, J. Murcia-Serra, M. Watorek, *The Relevance of Information Organization to Second Language Acquisition Studies*, "Studies in Second Language Acquisition", XXII, 2000, 3, pp. 441-446.
- Carroll, von Stutterheim 2003: M. Carroll, C. von Stutterheim, *Typology and Information Organization*, in: A. Giacalone Ramat (a cura di), *Typology and Second Language Acquisition*, Berlin 2003, pp. 365-402.
- Chafe 1976: W. Chafe, *Givenness, Contrastiveness, Definiteness, Subjects, Topics and Point of View*, in: Ch.N. Li (a cura di), *Subject and Topic*, New York 1976, pp. 25-56.
- Chini 2008: M. Chini, *Individuazione del topic in testi di apprendenti, fra teoria e empiria*, in: B. Ahrenholz, U. Bredel, W. Klein, M. Rost- Roth, R. Skiba (a cura di), *Empirische Forschung und Theoriebildung: Beiträge aus Soziolinguistik, Gesprochene-Sprache- und Zweitspracherwerbsforschung*, Frankfurt am Mein 2010, pp. 203-216.
- Chini et al. 2003: M. Chini, S. Ferraris, A. Valentini, B. Businaro, *Aspetti della testualità*, in: A. Giacalone Ramat (a cura di), *Verso l'italiano. Percorsi e strategie di acquisizione*, Roma 2003, pp. 179-219.
- Crismore 1989: A. Crismore, *Talking with Readers: Metadiscourse as Rhetorical Act*, Frankfurt am Mein 1989.
- Crismore et al. 1993: A. Crismore, R. Markkanen, M. Steffensen, *Metadiscourse in Persuasive Writing: A Study of Texts Written by American and Finnish University Students*, "Written Communication", X, 1993, 1, pp. 39-71.

- Dittmar 1992: N. Dittmar, *Introduction: Topic in Grammar and Discourse*, "Linguistics", XXX, 1992, pp. 1-3.
- Esvan 1996: F. Esvan, *L'uso anaforico del pronome sostantivale "ten" in ceco contemporaneo*, in: R. Benacchio, F. Fici, L. Gebert (a cura di), *Determinatezza e indeterminatezza nelle lingue slave*, Firenze 1996, pp. 61-75.
- Ferrari 2009: A. Ferrari, *La progressione tematica rivisitata*, "Vox Romanica", LXVIII, 2009, pp. 98-128.
- Flowerdew 2001: L. Flowerdew, *The Exploitation of Small Learner Corpora in EAP Material Design*, in: M. Ghadessy, R. Roseberry (a cura di), *Small Corpus Studies and ELT*, Amsterdam-Philadelphia 2001, pp. 363-379.
- Fox 1987: B. Fox, *Anaphora in Popular Written English Narratives*, in: R. Tomlin (a cura di), *Coherence and Grounding in Discourse*, Amsterdam-Philadelphia 1987.
- Givón 1983: T. Givón, *Topic Continuity in Discourse: An Introduction*, in: T. Givón (a cura di), *Topic Continuity in Discourse: A Quantitative Cross Language Study*, Amsterdam-Philadelphia 1983, pp. 1-41.
- Granger 2004: S. Granger, *Computer Learner Corpus Research: Current Status and Future Prospects*, in: U. Connore, T. Upton (a cura di), *Applied Corpus Linguistics: A Multi-dimensional Perspective*, Amsterdam-New York 2004, pp. 123-145.
- Lambrecht 1994: K. Lambrecht, *Information Structure and Sentence Form. Topic, Focus and the Mental Representations of Discourse Referents*, Cambridge 1994.
- Maslova, Bernini 2006: E. Maslova, G. Bernini, *Sentence Topics in the Languages of Europe and Beyond*, in: G. Bernini, M.L. Schwarz (a cura di), *Pragmatic Organization of Discourse in the Languages of Europe*, Berlin 2006, pp. 67-120.
- Mathesius 1947: V. Mathesius, *Privlastkové ten, ta, to v hovorové češtině*, in: V. Mathesius (a cura di), *Čestina a obecný jazykozpyt*, Praha 1947, pp. 185-189.
- Mauranen 1993: A. Mauranen, *Cultural Differences in Academic Rhetoric. A Textlinguistic Study*, Frankfurt am Mein 1993.
- Rizzi 1997: L. Rizzi, *The Fine Structure of the Left Periphery*, in: L. Haegeman (a cura di), *Elements of Grammar*, Dordrecht 1997, pp. 281-337.
- Tomlin 1987: R. Tomlin, *Linguistic Reflections of Cognitive Events*, in: R. Tomlin (a cura di), *Coherence and grounding in Discourse*, Amsterdam-Philadelphia 1987, pp. 455-479.

- Trovesi 2004: A. Trovesi, *La genesi degli articoli determinativi*, Milano 2004.
- Zimová 1994: L. Zimová, *Způsoby vyjadřování větných členů v textu. Konkurence pojmenování, pronominalizace a elize*, Brno 1994.

Siti consultati

<www.cesky-jazyk.cz> (ultimo accesso: 27.11.2015)

<<http://utkl.ff.cuni.cz/learncorp/>> (ultimo accesso: 26.10.2015)

<www.wrilib2.eu> (ultimo accesso: 21.11.2015)

Abstract

Anna Maria Perissutti

Topic Selection and Anaphoric Codification in Czech L2 Argumentative Texts

The research presented in this paper has been realized in the framework of Wrilib2, an European LLP project co-financed by the European Commission for the period 2014-2016 and aimed at creating an on-line reading and writing laboratory for Czech, German, Italian and Slovenian as L2.

One of the objectives of the project is to analyze the textual strategies of L2 writers, confronting them to those of L1 writers, with the aim of discovering the properties of L2 written interlanguage, while highlighting features and areas that need to be strengthened.

The research is based on two small corpora: a Czech L2 corpus composed of 15 argumentative compositions, devoted to the discourse topic of "Mass media" and taken from the Czech learner corpus CzeSl, and a Czech L1 corpus composed of 6 argumentative compositions (of the genre "esej"), taken from the portal *Český jazyk Literatura aneb studentský underground*, <www.cesky-jazyk.cz>.

After having exposed the criteria used to define topics in L1 and L2 texts, the research deals with anaphoric codification in Czech L1 and Czech L2 compositions and highlights the preference of Czech L2 writers for lexical means, in place of morphological means more often used by Czech L1 writers.