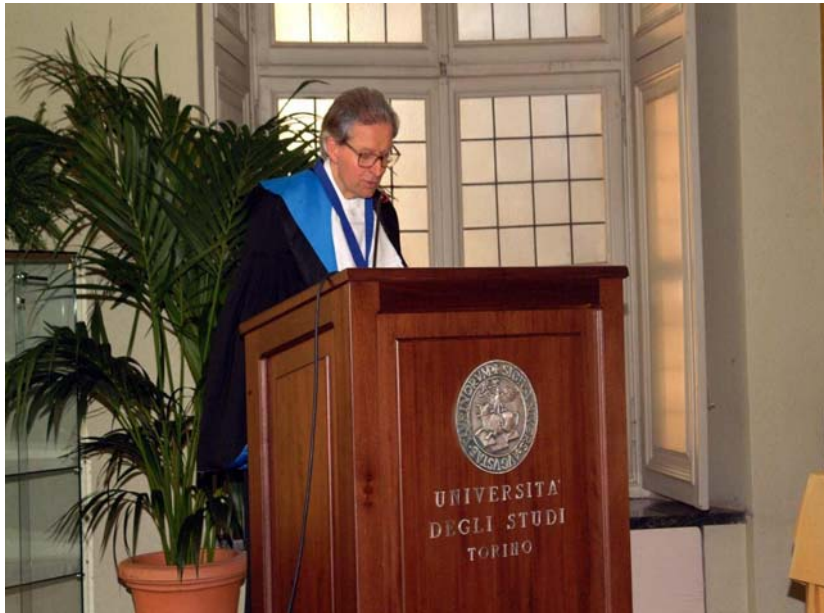


# TAVOLA ROTONDA CON PETŐFI JÁNOS SÁNDOR

Linguistica del testo, semiotica del testo, linguistica dei corpora:  
quali rapporti? (Torino, 27 febbraio 2004)

---

*a cura di Manuel BARBERA*



Petőfi János Sándor durante la Lectio magistralis

## **0. Premessa**

Petőfi János Sándor il pomeriggio del 27 febbraio 2004 ricevette dall'ateneo torinese una laurea honoris causa in Lingue e letterature straniere<sup>1</sup>. Nel mattino della medesima giornata, sempre all'Università di Torino, il locale Dottorato in linguistica, linguistica applicata e ingegneria linguistica organizzò una tavola rotonda con lui sul tema *Linguistica del testo, semiotica del testo, linguistica dei corpora: quali rapporti?* Vi presero parte, oltre a Petőfi attorno a cui l'evento fu costruito, Ugo Volli, Carla Marellò, Dario Corno, Maria Teresa Prat

---

<sup>1</sup> La sua lectio magistralis, Il conferimento del senso ai testi verbali. Aspetti dell'insegnamento delle lingue e delle letterature, è poi stata pubblicata in *Pensieri e parole del Novecento*, a cura di Elisabetta Soletti e Cristina Onesti, Alessandria, Edizioni dell'Orso, 2012, pp. 167-181.

Zagrebelsky, il sottoscritto ed alcuni dottorandi, tra cui i nostri Adriano Allora, Stefania Ferraris ed Alessandro Panunzi<sup>2</sup>, ed ancora altri colleghi e dottorandi di Torino e Macerata.

La pubblicazione di quella tavola rotonda è stata troppo a lungo ritardata, e molti eventi sono frattanto trascorsi, alcuni fausti (Alessandro Panunzi è ora associato di Linguistica generale a Firenze) ed altri molto meno (Petőfi, a più di ottant'anni, ci ha lasciati nel febbraio del 2013); adesso i principali interventi sono stati finalmente sbobinati. Gli spunti scritti che originariamente accompagnavano la lettera di Carla Marello sono qui riprodotti pressoché immutati, solo alleggeriti di alcuni grassetti e con pochissimi ritocchi; i testi sbobinati, poi, sono offerti praticamente come furono pronunciati: non vi sono state grosse revisioni d'autore ed il curatore si è perlopiù limitato all'interpunzione, alla paragrafematica ed alla formattazione; le risposte di Carla Marello e mie (probabilmente provocate a bella posta) non sono più replicate qui, come avrebbero dovuto, perché, richiedendo più spazio, sono state più diffusamente rielaborate per una comunicazione orale dal titolo *Corpus-based text linguistics* presentata al workshop internazionale *János S. Petőfi – In Memoriam (An International Workshop)*, tenutosi all' Universidad Complutense de Madrid, il 23 Aprile 2015, ora pubblicato in M. Borreguero Zuloaga & L. Vitacolonna (eds.): *The Legacy of János S. Petőfi: Text Linguistics, Literary Theory and Semiotics*, Newcastle-upon-Tyne, 2018 "Cambridge Scholars".

«L'incontro – diceva Carla Marello nella lettera di invito –, oltre a voler dare ai dottorandi la possibilità di esporre le loro idee a esperti esterni al collegio docenti, vorrebbe porre le basi per una discussione teorica da proseguire in sedi e modi da stabilire in seguito, anche sulla base di quanto verrà detto. Spunto della tavola rotonda è la constatazione che mentre i rapporti fra linguistica e semiotica del testo sono sempre stati stretti (e anzi indissolubili nell'opera di studiosi come appunto J.S. Petőfi), i legami fra linguistica del testo e linguistica dei corpora solo ad alcuni sembrano strettissimi».

### **1. Spunti per la riflessione (Carla Marello)**

La corpus linguistics è una text corpus linguistics? Finora è stata per lo più dominio di ricerche sulle collocazioni all'interno della frase. Quello che osservavano Biber e Finegan nel 1991 (in *English Corpus Linguistics. Studies in Honour of Jan Svartvik*, a cura di Karin Aijmer e Bengt Altenberg, London, Longman, 1991), «reflecting the state of linguistic studies generally, few studies have exploited the corpora to analyse characteristics of texts rather than characteristics of sentences», è ancora vero, ma molto sta cambiando.

La corpus linguistics è nata – a prescindere dalla preistoria filologica – come una branca dell'elaborazione computeristica del linguaggio naturale: non è stata (finora) molto creativa nel campo della progettazione di software di ricerca, ma piuttosto votata all'applicazione di software creati per indagare aspetti al massimo frasali, perché le ricerche corpus-based di ambito inglese non erano molto interessate ad aspetti testuali. E quelle corpus-driven non potranno essere serie a livello di testi finché i corpora per tipo di testo non saranno veramente grandi.

---

<sup>2</sup> Alessandro Panunzi aveva parlato della sua esperienza nel preparare il corpus C-ORAL-ROM, una raccolta di parlato spontaneo lemmatizzata ed etichettata per parti del discorso in un'ottica testuale; Stefania Ferraris aveva spiegato in che misura i suoi precedenti nel costruire corpora di apprendenti stranieri di italiano abbiano guidato gli studi che stava conducendo sul cloze test, uno dei test più testuali che esistano; Adriano Allora, infine, si era occupato dei problemi che pone costruire un insieme di corpora, congiuntamente interrogabili, in cui siano al contempo presenti testi prodotti per la fruizione in rete e testi prodotti originariamente per essere letti stampati.

Levinson nel 1983 (Pragmatics, 1983; trad. it. La pragmatica, Bologna, il Mulino, 1985) sosteneva che i modelli proposti dalla linguistica del testo erano “prematuri” in quanto troppo induttivi e basati, più che su veri e propri corpora, sulla intuizione del ricercatore: limiti superabili solo grazie alla metodologia dell’analisi della conversazione. Adesso i corpora per tipo di testo ci sono / possono esserci. Però pare che la dimensione testuale nell’elaborazione del linguaggio naturale stia più a cuore ai cognitivisti, ai cultori di frame semantics, che ai cultori di linguistica dei corpora.

Ora che perfino la scelta del tagset morfosintattico è collegata a riflessioni sul tipo di testo da annotare; che i browser intelligenti fanno uso di tecniche stocastiche corrette da caratteristiche testuali (in che parte del testo in rete sta quella parola e con quali si accompagna a che distanza); che la web semantics è un po’ l’incarnazione moderna web-oriented della ricerca che un tempo si chiamava Textsorten / Texttypologie; il connubio fra linguistica del testo e linguistica dei corpora, sic stantibus rebus, mi pare assodato. Innegabile almeno nei fatti, anche se non percepito così chiaramente a livello teorico come sarebbe invece auspicabile, sia da chi fa ricerche corpus-based, sia soprattutto da chi crede di farle corpus-driven.

Angela Ferrari e Emilio Manzotti concludevano il loro capitolo La linguistica del testo in La linguistica italiana alle soglie del 2000 a cura di Cristina Lavinio, Roma, Bulzoni, 2002, p. 438, dicendo: «e nella linguistica italiana, dove sta andando la LdT [linguistica del testo]? [...] Occorre distinguere tra concetti della LdT e LdT più in generale, perché se è vero che i primi sono entrati nella manualistica e nella lessicografia attuale e in tutti gli ambiti e in tutte le pratiche (didattica dell’italiano, insegnamento della scrittura, traduzione, analisi sociolinguistiche, tipologiche ecc.), riguardo alla LdT in generale il discorso è più diversificato. Certo abbiamo potuto registrare indagini approfondite di singoli fenomeni testuali (le riprese anaforiche, alcune relazioni logiche, alcuni tipi di testo ecc.) e applicazioni didattiche che non hanno a nostro avviso equivalente fuori d’Italia. Tuttavia non abbiamo trovato negli ultimi dieci anni<sup>3</sup>, tranne alcune eccezioni di cui abbiamo parlato, né riflessioni puntuali rivolte ad orizzonti più teorici né indagini più generali che facciano il punto sulla LdT, intrecciando altri ‘fili del discorso’ e ripensando altre ‘condizioni di coerenza’». Dalla ricognizione di Ferrari e Manzotti non si evince chiaramente se fuori d’Italia la riflessione teorica e le indagini più generali siano sufficienti e ce ne siano quindi da adottare come modello anche in Italia.

È sbagliato pensare che gli orizzonti teorici e le indagini più generali sono quelli che la text corpus linguistics obbliga ad affrontare quando chiede al ricercatore “che POS tagset vuoi adottare per annotare efficacemente il tuo corpus?” e soprattutto “che POS+X tagset devi adottare se vuoi fare ricerche oltre la frase?”, “che markup testuale e di tipo testuale vuoi?”

Semiotica del testo multimediale e corpus linguistics non hanno per ora i grandi rapporti che dovrebbero invece pianificare: eppure i corpora di testi multimediali stanno crescendo e andranno prima o poi studiati dal punto di vista della formazione

a – di un linguaggio di query multimediale (che parte da elementi iconici per cercare anche elementi in veste linguistica e viceversa; per cercare elementi misti e sovrapposti)

b – di un markup che tenga conto dell’intreccio dei codici, della nascita di tipi di testi e di tipi di utenti nuovi (ad es. (parti di) testi che solo la macchina può leggere sono in aumento).

Si potrebbe anche obiettare che questa è una fuga in avanti imposta dal mezzo; che per inseguire nuovi testi si tralasciano le necessarie riflessioni generali sul codice lingua naturale nei testi, riflessioni che ancora attendono sode sistemazioni text corpus-based.

<sup>3</sup> A colmare questa mancanza ha poi provveduto la Ferrari medesima [nota del curatore].

Va lasciato tutto nelle mani di chi si occupa di TEI, Unicode e, appunto, web semantics? È dominio di informatici e documentaristi? Linguisti e semiotici chioseranno poi? Non si corre il rischio di prendere i dati che sono stati dati (codificati) da altri con altre prospettive? Io resto dell'avviso che i dati sono sempre presi, mai dati, e vedrei con gran favore una linguistica dei corpora di testi (multimediali e non – esiste un testo elettronico non multimediale?) semioticamente consapevole a monte.

## **2. Intervento di Ugo Volli**

Mi sembra molto interessante e utile aprire qui una tappa del dialogo scientifico fra semiologi e linguisti: un dialogo che non si è mai interrotto anche perché la semiotica, almeno nella sua versione europea, nasce consapevolmente ed esplicitamente dalla linguistica. La semiotica strutturale di ispirazione francese si pensa dalla sua nascita, negli anni Sessanta, come erede, naturalmente parziale, della tradizione di pensiero che risale a Saussure, Hjelmslev e che si è espansa poi oltre la linguistica con l'influenza di Lévi-Strauss e ha subito forti influenze benvenistiane. Bisogna dunque partire da una affiliazione che è anche, ovviamente, un'influenza metodologica forte della linguistica, anche se bisogna ammettere che si tratta forse di una linguistica un po' vecchia, nel senso che il modello linguistico cui i semiologi si riferiscono è più storico che attuale.

Come si potrebbe svolgere l'interazione tra linguistica e semiotica? Non credo sia opportuno pensare alla semiotica solo come una specie di linguistica dei testi multimediali. Non credo che la multimedialità o la differenza di materia espressiva sia il punto determinante nella differenza fra le due discipline. È vero che Saussure quando parla della semiologia si riferisce a pratiche comunicative extralinguistiche come i segnali di cortesia, i segnali militari e altri significanti diversi da quelli del linguaggio. E tuttavia la semiotica ha lavorato molto anche su forme di significazione che hanno natura linguistica, basti pensare al lungo lavoro sulla letteratura o sulla narratività che ha impegnato molta linguistica anche italiana e torinese, da Cesare Segre a D'Arco Silvio Avalle.

Direi che fra le due discipline ci sono due grandi differenze.

Innanzitutto la semiotica lavora al di là della linguistica del testo, nel senso che si propone l'analisi di testi di grandi dimensioni e dei loro effetti di testo complessi; normalmente la semiotica lavora su segmenti di processo linguistico più lunghi anche di quelli studiati dalla linguistica testuale, per esempio di interi romanzi o numeri di giornale, per cui ci sarebbe una specie di successione di dimensioni e di complessità.

La seconda differenza fra semiotica e linguistica (e tutto sommato quella che ci interessa maggiormente) consiste nell'evidenziare come problema principale non il funzionamento sintattico del testo ma i suoi effetti di senso, dunque le dimensioni semantiche e pragmatiche. Noi diamo per conosciuto, il lavoro linguistico e cerchiamo di capire che cosa, nelle configurazioni dell'espressione linguistica e non linguistica, produca significative configurazioni sul piano del contenuto, cioè produca del senso.

Da questo punto di vista il principio del metodo variazionale sul piano espressivo è una sorta di punto di partenza per il tipo di ricerche semiotiche. In genere emerge (e questo è accettato dalla maggior parte di tutti i semiologi) che questi effetti raramente sono locali nella catena linguistica e raramente sono analizzabili semplicemente a livello di superficie. Diversi elementi, diverse configurazioni espressive collocate in luoghi diversi e a diversi livelli per esempio in un testo narrativo (dall'organizzazione sonora del significante alle azioni descritte fino al ricorso a conoscenze enciclopediche evocate da elementi lessicali) confluiscono per produrre effetti di senso importanti. Per capire questi effetti complessi bisogna fare ricorso a categorie profonde in senso generativo insieme a categorie di tipo sintattico (di una sintassi

estesa), per esempio alla semiotica greimasiana che ha generalizzato gli attanti sintattici di Tesnière. Per i semiologi capire una storia vuol dire fra l'altro identificare delle distribuzioni di ruoli attanziali, dei modi in cui questi ultimi rientrano in una variazione sintagmatica, ma anche attribuire loro ruoli tematici di natura semantica, il che suppone da parte nostra la capacità di distinguere le figure superficiali del testo dagli investimenti sintattici e semantici variabili a cui vengono sottoposti.

In un racconto un certo personaggio non ci interessa in quanto tale, ma lo analizziamo come il luogo in cui si proiettano da un lato certi ruoli attanziali che via via variano secondo certe strutture sintagmatiche e dall'altro certi ruoli tematici, che invece rilevano di una semantica. Tutto ciò presuppone un forte apparato metalinguistico, e tanto più forte quanto più procediamo in profondità, da proiettare sul testo per identificare delle strutture passionali (nel senso di posizioni emotive, non in senso sociologico) condivise, delle strutture ideologiche o assiologiche, delle posizioni enunciazionali connesse che contribuiscono al funzionamento del testo.

Tutto questo ovviamente si basa su un certo numero di ipotesi teoriche forti sulla semantica dei testi cui facciamo riferimento. In altri termini si tratta di provare a vedere se funzionano certe restrizioni sulle possibili forme della semantica dei testi.

Uno di tali presupposti, fra i più fondamentali, è che non ci siano fatti di senso isolati, ma solo relazioni che producono effetti di significazione, e quindi si tratta di identificare queste relazioni vedendo come il parlante o il lettore si basino su quel che già sanno, sia il sapere individuale o appartenga alla enciclopedia condivisa.

Quanto accade negli ipertesti o nelle chat, come nei romanzi e nelle poesie, fa pesantemente riferimento a del sapere sul mondo, sulle interazioni, sugli altri, che il lettore possiede; ed è chiarissimo che in ogni forma di comunicazione vige un principio di economia per cui chi produce il testo dice meno che può appoggiandosi il più possibile sul sapere degli altri, semplicemente evocandolo e attivando per mezzo del testo certe porzioni di conoscenza, di enciclopedia possedute dal suo lettore; per cui una stessa unità testuale di superficie può riferirsi in contesti diversi a sensi del tutto differenti, evocando dei blocchi di sapere attraverso la semantica contestuale o in maniera anaforica. Se si cerca di capire come funziona un giornale, per esempio come viene trattato da diversi giornali un caso politico di attualità, è abbastanza ovvio che chi scrive fa pesanti ipotesi implicite sul sapere e sulle passioni del suo ascoltatore, e ci lavora sopra.

C'è insomma una differenza di livelli e di strumenti analisi, ma credo stia maturando forse la possibilità di confrontarsi sull'interfaccia. Una volta le differenze tra un'analisi semiotica e un'analisi linguistica erano molto grandi, era difficile riempire la lacuna in mezzo anche se sapevamo che lavoravamo sulla stessa sostanza espressiva. Adesso, grazie a ciò che abbiamo imparato e grazie anche ai progressi dovuti alla tecnica, che dà ai linguisti la possibilità di utilizzare corpora complessi, forse c'è la possibilità di cominciare ad avvicinarsi ai due lati del tunnel, di confrontare le nostre discipline su oggetti concreti.

Tutto sommato la semiotica si trova in mezzo fra la linguistica e la sociologia perché presuppone delle competenze sociali. Però l'idea è di rielaborare queste competenze sociali in una forma che possa essere assimilata a quella della grammatica, cioè a dei saperi organizzati in norme implicite analoghe a quelle elaborate dalla linguistica. Da questo punto di vista, lavorando in direzione inversa rispetto a quella di alcune tendenze sociolinguistiche che tendono ad isolare i singoli tratti linguistici, a noi interessa cercare di capire se è possibile definire dei frammenti di langue e quindi formalizzare questi blocchi di conoscenza, fare delle ipotesi teoriche forti per capire come funzionano i testi. Io credo che non siano mancati dei successi su

questo fronte, così come ce ne sono stati sul fronte della comunicazione visiva, della comunicazione audiovisiva (relativa alle immagini in movimento in particolare), della musica.

Tutti questi aspetti della comunicazione sono unificati, secondo me, da un approccio antinaturalistico – che è un altro tema importante –, da un progetto scientifico che è il contrario del progetto epistemologico dominante oggi in filosofia, cioè quello di sottolineare gli aspetti che sono convenzionali, quindi in qualche modo storici, culturali, in qualche modo antropologici, perché quella è una grande frontiera, invece degli aspetti della effettiva realizzazione linguistica, o delle competenze semiotiche nel cervello che interessano di più la ricerca anglosassone; mentre l'altra grande frontiera riguarda la scoperta di modelli di comprensione che siano analoghi a ciò che sta sotto alle vostre tag, che è una analisi, un “pre-sapere” una “pre-conoscenza” su come potrebbe o dovrebbe funzionare la struttura sintattica del linguaggio.

Si tratta di comprendere se riusciamo a capire almeno per pezzetti, perché certo non è possibile avere un'immagine complessiva, se è possibile avere dei “pre-giudizi” su come funzionano dei pezzi di semantica del linguaggio nel senso ampio che dicevo prima o delle altre forme di discussione.

### **3. Intervento di Petőfi János Sándor**

Sono grato a Carla Marellò per aver reso possibile questo incontro. Le faccio gli auguri, per questo incontro, che lei considera un punto di partenza, cui seguiranno incontri più lunghi in cui tutti noi abbiamo il tempo sufficiente per spiegare ciò che pensiamo.

Prima di tutto, cercherò di riflettere su ciò che hanno presentato i dottorandi, poi sulla lettera o handout presentato da Carla Marellò e infine sulla relazione presentata dal collega semiologo.

I corpora: il primo problema è come definire e giustificare la scelta. I corpora dovrebbero soddisfare requisiti qualitativi e quantitativi, soprattutto se li si analizza con mezzi statistici. Se non si formulano questi requisiti qualitativi e quantitativi, non si possono emettere giudizi sui risultati che emergono.

Scusate, ma dal punto di vista terminologico sono molto pignolo. Quando studiavo per diplomarmi in Matematica il professore di Statistica iniziava la lezione dicendo: “dovete sapere bene che la statistica è una configurazione giusta di dati non giusti”.

Come giustificare la provenienza dei dati? Se si tratta di questo lavoro che voi avete iniziato, come creare questa pilot statistics, cioè tentare di identificare il problema e poi proseguire su un livello metalinguistico diverso e/o anche sul contenuto?

Il problema è che quando i dottorandi terminano i dottorati non hanno la possibilità o la voglia di continuare la ricerca. Invece, si dovrebbe trattare di una staffetta, e i dottorandi stessi dovrebbero passarsi il testimone. E questa era la prima osservazione.

Per quanto riguarda i corpora in internet, non so cosa si possa fare. Occorre distinguere tra corpora aperti e corpora chiusi in quanto gli ipertesti chiusi funzionano in modo diverso rispetto a quelli aperti. Inoltre gli ipertesti chiusi sono fatti quasi sempre con testi scritti. Dentro la cosiddetta lexia, che fa parte dell'ipertesto, le regole sono diverse rispetto agli ipertesti aperti.

In relazione a questi ipertesti bisogna tenere conto della multimedialità. Secondo me, non esistono testi comunicati unimediali o monomediali perché anche un testo normale ha un medium concettuale e un medium visivo o acustico. Questo è molto importante anche per quanto riguarda i corpora basati sul parlato o sullo scritto.

Io preferirei che l'analisi dei corpora venisse accompagnata da una videoregistrazione, perché molte volte non si può appurare l'esistenza o meno di un'ironia basandosi semplicemente sulle trascrizioni dei testi.

Per esempio, quali funzioni hanno i cosiddetti connettivi dummies? Cioè e può sostituire un ma, un però, un perché, insomma quasi tutti i tipi di connettivi? Per quanto riguarda l'analisi, questo non è sempre chiaro. Questa osservazione è collegata a ciò che posso dire qui di seguito.

Se si vuole collegare l'analisi dei corpora con gli aspetti testuali, occorre fornire informazioni riguardanti il cosiddetto co-testo, cioè in quale contesto verbale sono inserite le frasi o porzioni di testo che si analizzano. Naturalmente non ci si dovrebbe fermare soltanto sul co-testo. Per quanto riguarda la lexicon enciclopedia e così via, non si termina l'analisi con l'enciclopedia o con la conoscenza sintattica.

A lezione uso sempre un mini testo banale:

- (1) Pietro non è andato a scuola oggi. Il tempo era brutto.  
Cercate gli elementi che collegano linguisticamente queste due frasi.

Non esistono. Semplicemente il collegamento di questi due elementi viene dalla conoscenza che qualcuno possiede riguardo al mondo. La conoscenza non finisce con la conoscenza cosiddetta "linguistica"; come trattare anche questa conoscenza per spiegare meglio gli elementi presenti in un testo, cioè co-testo per quanto riguarda il contesto verbale e con-testo nel senso ampio? In questo caso io lavorerei non solo con pensieri induttivi, ma anche con pensieri abduttivi ("quando ho visto questo e quello, la situazione era quella", "adesso vedo questa costellazione e penso che anche in questo caso la situazione è quella").

E ora, alcune osservazioni terminologiche.

Oltre al problema creato dall'uso dei termini semiologia e semiotica, scuola di De Saussure e scuola di Peirce, che sono scuole diverse che conducono a risultati diversi, ci sono problemi relativi ai termini coesione e coerenza.

Per quanto riguarda la coesione, è più utile fare una distinzione tra coesione formale e coesione semantica, oppure tra connessità (per coesione formale) e coesione (per coesione semantica). La coerenza è rivolta al mondo, cioè alla connessità.

Nelle lingue con generi grammaticali l'uso di anafore e catafore è formale, si può decidere con sicurezza quali elementi sono connettivi. La coesione è una cosa meno precisa perché l'iperonimia e gli altri generi di rapporti lessicali fanno parte della coesione. Si può dire:

- (2) Ho scritto una lettera. Questa cosa disturbava X.

C'è un elemento coesivo dove non ci sono elementi precisi a livello superficiale che possano essere considerati come collegamenti.

Per quanto riguarda la coerenza, secondo me non è un aspetto dei testi scritti o parlati. È qualcosa legato con parte del mondo di cui si parla.

Non trovando termini migliori, proporrei costringenza, portando a tre i termini che definiscono il testo in relazione ai legami che stringe: connessità, coesione e costringenza.

La costringenza fa riferimento al rapporto tra porzioni del mondo. Se uso l'esempio di prima:

- (3) Pietro non è andato a scuola oggi. Il tempo era brutto.

C'è una coerenza per le persone che trovano un rapporto tra questi due stati di cose, altrimenti non c'è. Ma non si trovano elementi linguistici che dicano che si dovrebbe trattare questo frammento del mondo come un unico frammento.

Si dice che la statistica è un insieme giusto di dati non precisi.

Dipende come si crea una statistica: si tende a trovare elementi che rappresentino in modo oggettivo (o almeno in apparenza) cosa si fa. Per entrare nei dettagli, non so come trattare le cose nel modo giusto. Da circa trentacinque anni mi occupo di questa materia e mi rendo conto che in periodi di cinque-sei-sette anni si ricomincia di nuovo; molte volte anche sotto nomi diversi che possono essere necessari per trovare i fondi, per dire che non si fanno cose fatte in precedenza, ma si fanno cose nuove. Non ho niente in contrario, se si precisa cosa si fa una volta ottenuti i fondi.

Dico questo perché non so come si prendono e come si raccolgono i corpora. Come giustifico la mia analisi sui corpora? Molte volte non si ha tempo o non si ha la possibilità di fare un'analisi profonda dei lavori fatti per quanto riguarda un tema scelto.

Ho lavorato in contesti diversi per quasi diciassette anni in Germania e per quasi quattordici in Italia. Per quanto riguarda l'analisi di corpora ho visto (sono stato anche giudice in progetti ungheresi) che i temi ritornano con gli stessi risultati, ci si ferma, e si ricomincia di nuovo a parlare delle cose. Le cose possono essere diverse. Per quanto riguarda l'analisi dei corpora c'è un lavoro vastissimo da parte dei ricercatori tedeschi che hanno analizzato i corpora e dopo li hanno eliminati perché non sono stati registrati dati pertinenti che permettano un'analisi profonda.

Come si registrano i corpora per analizzarli? Si ha la registrazione verbale scritta, ma naturalmente l'analisi è diversa se si ha una registrazione audiovisiva. Naturalmente con quest'ultima vengono contraddetti i risultati che emergono rispetto a quando si analizza un corpus in mancanza di una registrazione audiovisiva. Come far fronte al problema? Non lo so. Saprei come fare, ma molte volte non c'è tempo, non ci sono fondi, o non c'è voglia sufficiente per complicare la vita. Si tenta di delineare un lavoro che porti risultati e in un certo senso porta sempre risultati. I problemi sono terminologici.

L'uso della terminologia può comportare problemi profondi. Parlare di sintattica dei corpora. Cos'è? Semantica dei corpora? In che senso? Semantica che si riferisce agli elementi verbali presenti, o alle cose di cui si parla?

Usare sintattica o semantica ha spiegazioni o basi completamente diverse. Soprattutto per quanto riguarda l'handout di Carla Marello, la semiotica del testo multimediale e la corpus linguistics non hanno per ora i rapporti che dovrebbero invece pianificare. Perché?

Sicuramente non li hanno perché è complicato trattare i problemi, si lascia perdere e si tenta di delineare una problematica che permetta risultati di lavoro che alla fine non sono risultati reali per quanto riguarda le cose. Farei questa distinzione in modo radicale. In Germania lavoravo in organi sovraccademici per valutare i progetti; molte volte si ritorna alle stesse problematiche, con gli stessi risultati, ci si ferma e non c'è una continuazione.

Per quanto riguarda la semiotica del testo, anche questo non è un concetto ben preciso. Si dovrebbe creare e introdurre una terminologia (e mi rivolgo in particolare ai colleghi giovani). Molto spesso si usa una terminologia che rimane vaga. Le parole sono uguali, ma i concetti sono diversi e i risultati non sono comparabili. Posso fare degli esempi: linguistica testuale è un'espressione sbagliata perché la linguistica non può trattare tutti gli elementi che sono rilevanti per quanto riguarda un testo; teoria testuale è un'espressione troppo impegnativa, non copre le cose.



Ho creato alla fine questa semiotica testuale perché pensavo che in questo modo si potesse parlare degli aspetti centrali del testo e da lì è nato il termine semiotica testologica. Il problema è che molte volte nei lavori non vengono spiegati i concetti che si usano. L'interpretazione è quasi libera e ognuno interpreta in base alla sua posizione. Forse è inevitabile, ma si dovrebbe tentare di essere più onesti. Io proporrei a Carla Marellò di fare un elenco dei problemi centrali e di organizzare incontri in cui potrebbero essere trattati, perché i giovani rischiano di iniziare un lavoro e di perdere l'entusiasmo e la possibilità di finire un lavoro ben definito; almeno dare un elenco dei problemi centrali che si potrebbero trattare.

#### **4. Intervento di Maria Teresa Prat Zagrebelsky**

Io parlo come ricercatrice e docente di Inglese come lingua straniera. Per me è stata a un certo punto una necessità occuparmi dei corpora, perché ne esistono in grandissima quantità per l'inglese; si possono comprare, sono delle imprese descritte, c'è già grande materiale pronto a disposizione: c'è stato un enorme dibattito, che è andato avanti tanto tempo e che io a un certo punto ho dovuto recuperare con grande fatica; quindi per me occuparmi di corpora è un po' una necessità e anche un'ansia. Dico subito qual è l'ansia.

L'ansia è (e su questo non rispondo ai tre dottorandi ma anzi pongo un quesito) che questo tipo di ricerca presuppone conoscenze linguistiche, sensibilità, ed una formazione ai problemi generali molto ampia, molto buona, proprio perché i dati che vengono dai corpora sono tanti e quindi bisogna saperli interpretare bene, saperli controllare. E questo già riempie la vita.

Ed oltre tutto bisogna anche avere una conoscenza dello strumento informatico. Tutti i grandi linguisti inglesi ormai da anni lavorano in équipe con degli esperti, spesso giovanissimi e bravissimi, di ambito computazionale. E poi c'è la statistica, che io per esempio non ho mai studiato. E tutte le volte che devo fare anche piccoli ragionamenti statistici, che pure sono importanti (è importante capirli, riprodurli), vengo presa dall'ansia.

La domanda che conclude il mio intervento è questa: quale sarà, dovrà essere, potrà essere la formazione dei futuri linguisti che secondo me devono avere non solo una base linguistica forte ma anche aver modo di agire negli altri due campi che ho menzionato? Cominceremo a lavorare formando dei gruppi in cui ci siano queste diverse competenze, come si fa in ambito scientifico dove in genere si mettono insieme diverse competenze, o dovremo, noi, voi, avere questa diversa formazione, almeno a un certo livello? E si dovranno modificare, arricchire i corsi di studio?

Io mi pongo questi problemi per la formazione dei futuri linguisti, non solo di chi fa il dottorato ma anche di chi studia nei nostri corsi (nelle triennali, nelle lauree di specializzazione, ecc.); e mi trovo in difficoltà a introdurre dei cambiamenti anche nel settore in cui opero.

Quindi più che dare delle risposte, ho fatto anche delle domande, e sarebbe interessante sapere dai tre dottorandi qui presenti come coniughino queste diverse competenze, che consentono di maneggiare con sicurezza e senza ansia gli strumenti nuovi.

#### **5. Repliche di Petőfi János Sándor**

A Prat. Lei faceva domande cui non sono state trovate risposte fino adesso. Per la verità non so come combinare le competenze.

Naturalmente bisognerebbe sapere come poter combinare le competenze in un gruppo, ma la vita accademica, forse in nessun paese, permette di creare gruppi che restano uniti per lungo tempo, il tempo sufficiente per elaborare le cose. Si tende ad elaborare competenze tante quante

sono possibili in sé. Naturalmente queste hanno i loro confini. Non si può andare oltre e non esiste una ricetta.

Io ho raccolto competenze diverse ma non con un progetto; la vita mi obbligava a raccoglierle. Combinare matematica da una parte e linguistica dall'altra credo sia abbastanza interessante, e adesso vengono anche le competenze tecniche cioè usare gli strumenti tecnici moderni.

I giovani dovrebbero essere invitati a studiare bene almeno tre lingue. E se possibile non lingue della stessa famiglia (per esempio non italiano, spagnolo e portoghese) ma possibilmente, e adesso parlo pro domo, una lingua non indoeuropea, perché dà una visione completamente diversa.

Che cosa si dovrebbe collegare? I giovani dovrebbero arrivare a un punto, e adesso ci sono diverse possibilità, e la mia esperienza per quanto riguarda l'università italiana è che l'inglese si colloca al primo posto, la lingua tedesca quasi non esiste, cioè noi abbiamo corsi in cui università diverse possono trattare le cose in modo diverso, ma si dovrebbe tentare di evitare di ripetere il lavoro fatto in precedenza. Non ha senso fare una terza, quarta versione delle cose già fatte, ma si potrebbero raccogliere i dati in contesti diversi.

Per quanto riguarda l'analisi dei corpora c'è un contesto inglese, angloamericano, svedese, norvegese, tedesco. Si dovrebbe poter analizzare e valutare i risultati e i metodi che loro hanno utilizzato per poter fare un lavoro che possa essere ritenuto una continuazione di un lavoro comune.

Ma si comincia sempre quasi da un punto zero e si arriva a un punto X e, o perché non ci sono più soldi, o perché non c'è più possibilità accademica, ci si ferma là e i prossimi ricominciano dall'inizio.

## **6. Intervento di Dario Corno**

Vorrei approfittare dell'occasione, senza usare le virtù della "mitigazione", per dirle con grande sincerità, del piacere di poterla incontrare qui a Torino nel momento in cui le viene offerta la laurea. Come ci ha ricordato Volli, ci sarebbe sempre uno sfondo "patemico" (e cioè ricco di emozioni) nei testi, e quello a cui vorrei accennare io rinvia a una storia che coinvolge Carla Marellò e in misura decisamente minore anche me.

Nel 1972-73 (eravamo leggermente più giovani, allora...) due stanze si fronteggiavano a Palazzo Nuovo, nella nostra Facoltà di Lettere. In una stanza c'era Petőfi: c'era lei – non lei ovviamente ma i suoi libri e la sua giovane studiosa –, c'erano assieme la sua passione terminologica, il bisogno di fare scienza anche nelle nostre discipline allora nascenti sotto la comune etichetta di "scienze umane". Nell'altra stanza, c'era – se così si può dire – una sorta di "parco giochi" di origine greimasiana, dove qualcuno di noi si alimentava alle idee di Greimas che penetravano allora nella nostra cultura, tutte incentrate più che sull'edificazione di una scienza certa, sul tentativo di capire che tipo di persona è la persona che usa una lingua quando la esaminiamo dal punto di vista della sola lingua e non da altri punti di vista. Se facciamo questo, come ha ricordato splendidamente Volli, ci occupiamo – oggi come ieri – del "senso".

Noi andavamo in queste stanze per trovare delle sicurezze e lì abbiamo imparato molte cose perché l'uno e l'altro, Petőfi e Greimas, hanno segnato una svolta decisiva nelle ricerche di semiotica del testo.

Quella svolta fu – per noi che eravamo stati formati all'idea che i testi fossero "oggetti" un po' come, secondo l'immagine di Juri Lotman, lo sono i cioccolatini dentro la carta stagnola – fondamentale. Eravamo infatti stati abituati dalla scuola a una stilistica e a una linguistica bloccate sull'idea della "staticità" del senso. Come dire: il senso è una cosa che prelevi dai testi; è in qualche misura già pre-formato.

Invece lei, mi scusi la rapidità e la perentorietà del giudizio, impresse assieme a Greimas una svolta alle nostre idee e alle idee di molti altri: le canalizzò verso una visione più dinamica e processuale degli oggetti linguistici. Fin da allora abbiamo iniziato a pensare che sarebbe stato interessante, magari proficuo, occuparci delle cose che capitano durante la costruzione del testo e non dopo. Oggi quello che lei aveva previsto è una sicurezza oggettiva: il testo è sempre più un oggetto dinamico, al punto che la libertà dell'interprete nei confronti della costruzione testuale sembra dimostrare vie decisamente più dinamiche di quanto fosse legittimo supporre in quegli anni lontani in cui noi studiavamo i suoi libri e i libri di Greimas. Oggi l'interprete fa molte più cose con i testi e (magari) fa, come lei ci ha fatto brillantemente capire, cose più sciocche. Tuttavia, ne fa molte di più. Qualcuno ha sostenuto che ci sarebbero nuove libertà in questo senso, fra tutte la libertà di pensare con la propria testa, oltre che con il testo. Lei ci ha indicato la strada.