

Bando definitivo

Scuola Invernale TRIPLE [SIT] 2010

Il Lessico *Classi di parole, metodi, computazione*

Roma, 25-29 gennaio 2010

Descrizione

Il laboratorio di ricerca TRIPLE annuncia la seconda edizione della Scuola Invernale TRIPLE (SIT2010). Anche in questa edizione saranno affrontati sia temi teorici, attraverso la discussione di modelli e metodi di analisi e di rappresentazione, sia temi applicativi, attraverso la presentazione degli strumenti e delle metodologie in uso nella linguistica computazionale e nella pratica lessicografica. Le lezioni saranno affiancate da laboratori dedicati ad approfondimenti ed esercitazioni pratiche.

Luogo e date

La SIT2010 si terrà a Roma, presso l'Università Roma Tre, dal 25 al 29 gennaio 2010.

Organizzazione

La SIT2010 dura 5 giorni e prevede 4 corsi (composti ciascuno da 3 lezioni), 5 sessioni di laboratorio e 2 conferenze, per un totale di 28 ore.

L'offerta didattica della SIT2010 è articolata come segue.

Corsi

- Nomi e nominalizzazioni
Antonietta Bisetto & Livio Gaeta
- La struttura argomentale dei verbi
Michela Cennamo & Elisabetta Jezek
- Linguistica computazionale
Alessandro Lenci
- Lessicografia
Vincenzo Lo Cascio

Laboratori

- Linguistica e lessicografia computazionale
Alessandro Lenci & Francesca Masini
- Lessicografia
Vincenzo Lo Cascio

Conferenze

- La forma (e la rappresentazione) della 'parola' cinese tra fonologia, semantica, questioni cognitive e antropologico-culturali
Emanuele Banfi
- Basi teoriche di un dizionario analogico
Raffaele Simone

Le sintesi di tutte le attività sono allegate in fondo a questo bando. La frequenza a tutte le lezioni assicura, per gli interessati, 6 CFU.

Il programma è descritto nella Tabella 1.

Tabella 1. Programma

	Lunedì 25	Martedì 26	Mercoledì 27	Giovedì 28	Venerdì 29
9.00-10.15	<i>Registrazione</i>	CORSO Cennamo/Jezek	CORSO Lenci	CORSO Bisetto/Gaeta	CORSO Lo Cascio
10.15-10.30	<i>Saluti</i>	<i>Pausa caffè</i>			
10.30-11.45	CONFERENZA Raffaele Simone	CORSO Lenci	CORSO Bisetto/Gaeta	CORSO Lo Cascio	CONFERENZA Emanuele Banfi
11.45-12.00	<i>Pausa caffè</i>				
12:00-13.15	CORSO Cennamo/Jezek	CORSO Cennamo/Jezek	CORSO Lenci	CORSO Bisetto/Gaeta	CORSO Lo Cascio
13:15-15:00	<i>Pranzo</i>				
15.00-17.00	LABORATORIO Computazionale	LABORATORIO Computazionale	LABORATORIO Computazionale	LABORATORIO Lessicografia	LABORATORIO Lessicografia

Destinatari

Laureandi e laureati magistrali, dottorandi, giovani ricercatori e operatori professionali della lessicografia e della linguistica computazionale.

Iscrizione

Gli aspiranti partecipanti dovranno compilare il modulo della domanda di partecipazione (scaricabile dal sito della SIT2010) e spedirlo all'indirizzo triple@uniroma3.it entro il **9 novembre 2009**. Dato il carattere altamente specifico della scuola, le domande verranno esaminate e selezionate. I risultati della selezione verranno comunicati entro il 20 novembre 2009. I partecipanti effettivi dovranno perfezionare l'iscrizione tramite il pagamento della quota di iscrizione entro il **4 dicembre 2009**. Le quote di iscrizione, suddivise per tipologia di partecipante, sono le seguenti:

Tabella 2. Quote di iscrizione

Tipologia di partecipante	Quota
Laureandi e neo-laureati magistrali	€ 100
Dottorandi	€ 200
Dottori di ricerca, post-doc, ricercatori	€ 250
Operatori professionali	€ 500

Borse di studio

La SIT 2010 metterà a disposizione due borse di studio a copertura delle spese di iscrizione. Le borse saranno assegnate sulla base del punteggio ricavato dalle informazioni fornite nel modulo di partecipazione.

Per ulteriori informazioni rivolgersi al

TRIPLE:

Dipartimento di Linguistica - Università Roma Tre

Via Ostiense 236 - 00146 Roma

tel: +39 06 57338343 - fax: +39 06 57338344

e-mail TRIPLE: triple@uniroma3.it

sito TRIPLE: <http://host.uniroma3.it/laboratori/triple>

URL SIT 2009: http://host.uniroma3.it/laboratori/triple/Scuola_School.html

ALLEGATO 1

Sintesi delle attività della SIT2010

(conferenze, corsi, laboratori)

La forma (e la rappresentazione) della ‘parola’ cinese tra fonologia, semantica, questioni cognitive e antropologico-culturali

[CONFERENZA]

Emanuele Banfi*

1. È nota la difficoltà teorica nel definire la nozione di ‘parola’ intesa quale categoria generale applicabile a lingue del mondo tipologicamente diverse (Anderson 1985; Ramat 1990 e Ramat 2005). La questione, di per sé già problematica nel caso di lingue agglutinanti e flessive, si fa ancor più complessa nel caso di lingue isolanti e, nello specifico, nel caso del cinese ove, in forza delle caratteristiche della struttura fonologica di tale lingua – e delle conseguenti ‘ricadute’ che tali caratteristiche hanno sugli altri livelli di analisi (morfologia, sintassi, lessico) –, la definizione di ‘parola’ è stata (ed è) oggetto di un ampio dibattito teorico (Duanmu 1998; Packard 2000; Yip 2000; Arcadia 2008): ciò che è stato (ed è) discusso è il rapporto tra le forme foniche (gli *cí* 词) sottese alla ‘parola’ cinese e le loro rappresentazioni (grafematico-)semantiche e cognitivamente salienti (gli *zì* 字: i logogrammi/i caratteri). Mentre il termine *zì* 字 ‘carattere’ ricorre in tutta la plurisecolare tradizione linguistica cinese, il termine *cí* 词 – indicante la forma fonica sottesa alla ‘parola’ e poi, più semplicemente, la ‘parola’ – compare invece per la prima volta solo agli albori del sec. XX (nel 1907, per la precisione) per divenire oggetto di reale discussione scientifica solo a partire dagli anni ’50 del secolo scorso quando, dibattendosi in Cina a livello ufficiale l’eventualità di adottare un sistema di scrittura alfabetico, si avviò contemporaneamente tra i linguisti cinesi (Dai 1993: 106-109; Duanmu 1998: 135) la discussione intorno alla ‘ammissibilità’/‘applicabilità’ della nozione stessa di ‘parola’ nelle descrizioni della loro plurimillennaria tradizione linguistica. Così, ad es., Y.R. Chao, il padre della linguistica cinese moderna, ritenendo la nozione di ‘parola’ qualcosa di (troppo) tipicamente ‘occidentale’, scriveva (Y.R. Chao (1968: 136),: “[...] It is [...] a matter of fiat and not a question of fact whether to apply the word ‘word’ to a type of subunit in the Chinese sentence”; e, ancora, qualche decennio più tardi, un altro autorevole linguista cinese, Shuxiang Lü (1981: 45) affermava categoricamente: “As a matter of fact, one does not need the notion ‘word’ in order to discuss Chinese grammar”.

2. Di fatto, al centro del problema, stava (e sta) il rapporto tra forme foniche (gli *cí* 词) e le loro rappresentazioni semantiche (gli *zì* 字) e ciò alla ricerca di una qualche forma di ‘equilibrio’ tra i due, necessariamente connessi, piani di analisi. Il nucleo del materiale fonologico del sistema lessicale cinese è costituito da 418 unità sillabiche, non tutte singolarmente ed egualmente distinte da uno stesso numero di opposizioni tonali e non tutte necessariamente combinabili in unità superiori (bisillabiche, per lo più). Ne consegue che – data la particolare natura dei suoi costituenti sillabici, a livello della oralità – la ‘forma’ della ‘parola’ (tipica) cinese (e ciò vale per tutte le varietà diatopiche dello spazio sinofono) appare caratterizzata da frequentissimi casi di omofonia; e, che, di conseguenza, l’etichettatura semantica di singole unità lessicali formate da monosillabi (o da bisillabi) è affidata unicamente al più generale contesto frasale.

3. La lingua cinese scritta ‘risolve’ però, di fatto, i casi di omofonia e le relative difficoltà di etichettatura semantica di singole unità lessicali. Vero e proprio sistema ‘a sé’, sovrapposto – quale potente elemento unificatore – alla frammentazione diatopica dello spazio linguistico cinese, la lingua cinese scritta prevede infatti che la ‘rappresentazione’ del significato delle ‘parole’ (degli *cí* 词, unità monosillabiche nel cinese antico; mono-/bisillabiche, prevalentemente, nel cinese moderno) sia immediatamente marcata da evidenze che rinviano al piano cognitivo e che sono ‘rappresentate’/‘inscritte’ entro la struttura degli *zì* 字.

4. Se gli *cí* 词 sono mere forme foniche (per lo più, tranne nel caso di voci onomatopiche) semanticamente immotivate (anche in prospettiva diacronica, e anche tenendo conto dei problemi di ‘erosione’ fonologica intercorsi, nella struttura della sillaba cinese, tra la fase arcaica della lingua e la sua forma attuale, come mostrato da Baxter/Sagart 1998), gli *zì* 字 invece, grazie alla loro natura (tendenzialmente) iconica, rinviano al sistema della secolare ‘codificazione’/‘classificazione’ del

materiale lessicale cinese: tale sistema, inaugurato dai redattori del primo dizionario cinese – lo *Ēryǎ* (爾雅, in caratteri tradizionali; o 尔雅, in caratteri semplificati) il *Pronto correttore*, risalente al 206 a.C. – è stato via via utilizzato nel corso dei secoli fino ad oggi, e ancora dai redattori dello *Hànyǔ Dà Cídiǎn* (漢語大詞典, in caratteri tradizionali; o 汉语大词典, in caratteri semplificati) il *Grande dizionario della lingua cinese*, senz'altro il maggiore repertorio lessicografico esistente della lingua cinese pubblicato a Shanghai tra il 1986 e il 2001 sotto la direzione di Luo Zhufeng.

5. Nucleo portante del processo di ‘codificazione’/‘classificazione’ degli *zì* 字 è il costante, programmatico richiamo alla dimensione cognitiva. Infatti, singolarmente, in ciascuno degli *zì* 字 ricorrono unità grafematiche rinvianti a puntuali macro-sfere semantiche: i cosiddetti ‘radicali’, i quali a loro volta sono connessi con dati esperienziali ‘elementari’ evocanti, tra le altre nozioni, elementi e fenomeni naturali (sole, luna; terra, montagna, fiume; acqua, fuoco; pioggia, ghiaccio, ecc.), vegetali (albero, erba, cereale, frutto, ecc.), animali (cane, serpente, uccello, insetto, pesce, ecc.), alcuni manufatti (armi, abiti, mezzi di trasporto, ecc.) ma anche l’essere umano (maschio, femmina) e le sue evidenze anatomiche (mano, gamba, occhio, orecchio, bocca, ecc.). In base a tali elementi e tenendo conto di parametri di classificazione che affondano le proprie motivazioni nella straordinaria ricchezza del pensiero cinese antico (Cheng 1997), i lessicografi cinesi hanno ‘organizzato’ la potenzialmente ‘disordinata’ serie degli *cí* 词 (unità immotivate del piano del significante) e li hanno posti entro ‘insiemi’ tendenzialmente coerenti dal punto di vista cognitivo, potenti ‘catalizzatori’ di processi di significazione e, parallelamente, di classificazione.

6. Nella relazione, tenendo conto del quadro teorico precedentemente indicato e di considerazioni generali d’ordine semantico (Chierchia 1997; Chierchia/Mc Collen-Ginet 2000²; Ježek 2005), verranno presentati alcuni esempi di organizzazione semantica propria di classi lessicali del cinese mandarino. In particolare saranno presentate e discusse le modalità di organizzazione del significato di ‘parole’ cinesi evocanti alcuni culturalmente importanti concetti ‘astratti’ (quali, ad es., *zhēnlǐ* 真理 ‘verità’, *zhèngyì* 正义 ‘giustizia’, *zìyóu* 自由 ‘libertà’). Nel far ciò si descriveranno i rapporti intercorrenti, nelle singole ‘parole’ cinesi oggetto di analisi, tra i rispettivi livelli fonologico, semantico/cognitivo e, non da ultimo, quello antropologico-culturale.

Riferimenti bibliografici essenziali

- Anderson, Stephen R. (1985), *Typological distinctions in word formation*, in T. Shopen (ed.), *Language Typology and Syntactic Description*. III: *Grammatical Categories and the Lexicon*, Cambridge, Cambridge University Press: 3-56.
- Arcadia, Giorgio Francesco (2008), *La derivazione lessicale in cinese mandarino*, Milano, FrancoAngeli.
- Baxter, William H. & Laurent Sagart (1998), *Word formation in Old Chinese*, in J.L. Packard (ed.), *New Approaches to Chinese Word Formation. Morphology, Phonology and the Lexicon in Modern and Ancient Chinese*, Berlin-New York, Mouton de Gruyter, 1993: 35-76.
- Chao, Yuen Ren (1968), *A Grammar of spoken Chinese*, Berkeley, University of California Press.
- Cheng, Anne (1997), *Histoire de la pensée chinoise*, Paris, Seuil.
- Chierchia, Gennaro (1997), *Semantica*, Bologna, il Mulino.
- Chierchia, Gennaro & Sally McCollen-Ginet (2000²), *Meaning and Grammar. An Introduction to Semantics*, Cambridge Mass., MIT Press.
- DeFrancis, John (1989), *Visible Speech : The Diverse Oneness of Writing Systems*, Honolulu, University of Hawai’i Press.
- Duanmu, San (1993), *Wordhood in Chinese*, in J.L. Packard (ed.), *New Approaches to Chinese Word Formation. Morphology, Phonology and the Lexicon in Modern and Ancient Chinese*, Berlin-New York, Mouton de Gruyter, 1993:135-196.
- Ježek, Elisabetta (2005), *Lessico. Classi di parole, strutture, combinazioni*, Bologna, il Mulino.
- Lü, Shuxiang (1981), *Talking about language*, Beijing, San Lian Shudian.
- Norman, Jerry (1988), *Chinese*, Cambridge, Cambridge University Press.
- Ooi, Vincent B.Y. et alii (eds.) (2005), *Words in Asian Cultural Contexts. Proceedings of the 4th Asialex Conference*, Singapore, National University of Singapore.
- Packard, Jerome L. (2000), *The Morphology of Chinese. A Linguistic and Cognitive Approach*, Cambridge, Cambridge University Press.
- Po-Ching, Yip (2000), *The Chinese Lexicon. A Comprehensive Survey*, London-New York, Routledge.
- Ramat, Paolo (1990), *Definizione di ‘parola’ e sua tipologia*, in M. Berretta, P. Molinelli, A. Valentini (a cura di/Hrsg.), *Parallela 4. Morfologia/Morphologie*, Atti del V Incontro Italo-Austriaco della SLI (Bergamo, 2-4 ottobre 1989), Tübingen, Narr: 3-15.
- Ramat, Paolo (2005), *Per una definizione di parola*, in Id. *Pagine linguistiche*, Roma-Bari, Laterza: 106-121.

Unger, Marshall J. (2004), *Ideograms: Chinese Characters and the Myth of Disembodied Meaning*, Honolulu, University of Hawai'i Press.
Xiang-Ling Dai, John (1993), *Syntactic, phonological and morphological words in Chinese*, in J.L. Packard (ed.), *New Approaches to Chinese Word Formation. Morphology, Phonology and the Lexicon in Modern and Ancient Chinese*, Berlin-New York, Mouton de Gruyter, 1993: 103-134.

* *Emanuele Banfi* (emanuele.banfi@unimib.it) è professore ordinario di Linguistica Generale presso l'Università di Milano-Bicocca. Si occupa prevalentemente di tipologia, linguistica storica (con particolare riferimento alle lingue classiche – greco e latino – e alle lingue dell'area balcanica), sociolinguistica, etnolinguistica. Negli ultimi anni ha dedicato particolare attenzione allo studio della lingua cinese.

Nomi e nominalizzazioni

[CORSO]

Antonietta Bisetto* e Livio Gaeta**

Il corso è articolato in due parti.

Parte A. Nomi d'azione a metà strada tra nomi e verbi (Livio Gaeta)

Per formare nomi d'azione, in italiano così come in molte altre lingue del mondo, sono presenti, con gradi diversi di produttività, vari procedimenti di natura morfologica (la suffissazione e la conversione) e di natura sintattica (la sostantivazione dell'infinito). Nel corso verranno discusse le proprietà dei diversi processi, che delincono un quadro variegato ma coerente di possibilità. In particolare si metterà in evidenza come, nel passaggio categoriale verbo > nome, proprietà della categoria di base come l'azionalità verbale vengano messe a frutto nella formazione dei derivati nominali.

Parte B. I nomi risultato (Antonietta Bisetto)

Le lezioni avranno per oggetto le nominalizzazioni. Sono comunemente definite nominalizzazioni i derivati (nominali) da verbi. Frequentemente una nominalizzazione 'eredita' dal verbo (parte del)la struttura di argomenti che quest'ultimo proietta. Un verbo d'azione come *tradurre*, ad esempio, dà origine ad un derivato nominale come *traduzione* che, quando accompagnato dagli argomenti del verbo (*Paolo ha tradotto il testo di Searle molti anni fa / La traduzione del testo di Searle da parte di Paolo è di molti anni fa*) viene definito nominalizzazione d'azione o evento.

L'uso della nominalizzazione priva di argomenti non si riferisce, di contro, all'evento espresso dal verbo ma a quello che si chiama un 'risultato': *la traduzione è stata pubblicata subito* indica l'oggetto 'traduzione' non il processo/evento del tradurre.

Nel corso delle lezioni si discuterà in particolare della nozione di nominalizzazione risultato in relazione ai vari tipi di verbi dai quali i nominali possono essere derivati.

Riferimenti bibliografici

Gaeta, Livio (2002), *Quando i verbi compaiono come nomi*, Milano, Franco Angeli.

Simone, Raffaele (2000), *Cycles lexicaux*, in "Studi italiani di linguistica teorica ed applicata" 29 (2), 259-287.

Altre letture verranno indicate prima del corso.

* *Antonietta Bisetto* (antonietta.bisetto@unibo.it) è professore associato di Linguistica Generale presso l'Università di Bologna. Si occupa prevalentemente di morfologia derivazionale, composizione e semantica lessicale.

** *Livio Gaeta* (lgaeta@unina.it) è professore associato di Linguistica tedesca presso l'Università di Napoli "Federico II". Si occupa prevalentemente di morfologia, fonologia, teoria del cambiamento linguistico, teoria della grammaticalizzazione e linguistica cognitiva.

Il corso è articolato in due parti.

Parte A. Approcci teorici alla struttura argomentale (Elisabetta Jezek)

Nella prima parte del corso si offrirà un'introduzione alla teoria della struttura argomentale e un'analisi dei parametri semantici e pragmatici che condizionano la codifica morfosintattica degli argomenti, con discussione di alcuni modelli di rappresentazione, costruzionisti (es. Pustejovsky 1995, 2000, Goldberg, 1995, 2009) e proiezionisti (es. Rappaport Hovav & Levin 1998, Levin & Rappaport Hovav 2005, van Valin 2005, 2009). Si daranno inoltre nozioni introduttive per quanto riguarda i tipi di argomenti (semantici, sintattici, espressi, non espressi ecc.) e la relazione tra struttura dell'evento e struttura argomentale.

Riferimenti bibliografici preliminari

- Goldberg Adele E. (2009), *Verbs, Constructions and Semantic Frames*, in M. Rappaport Hovav, E. Doron and I. Sichel (eds.). *Syntax, Lexical Semantics and Event Structure*. Oxford University Press.
- Goldberg, Adele E. (2005), *Argument realization: the role of constructions, lexical semantics and discourse factors*, in Jan-Ola Oostman and Mirjam Fried (eds.) *Construction Grammar(s): Cognitive and Cross-language dimensions*, Amsterdam: John Benjamins.
- Levin, B. & M. Rappaport Hovav (2005), *Argument Realization*, Cambridge: Cambridge University Press.
- Pustejovsky, J. (1995), *The Generative Lexicon*. Cambridge MA: MIT Press.
- Pustejovsky, J. (2000), *Lexical shadowing and argument closure*, in Y. Ravin & C. Leacock (eds), *Polysemy. Theoretical and computational approaches*. Oxford: Oxford University Press, 68-90.
- Rappaport Hovav, M. & B. Levin (1998), *Building verb meanings*, in M. Butt and W. Geuder (eds), *The Projection of Arguments: Lexical and Compositional Factors*, 97-134. Stanford (CA): CSLI Publications.
- Van Valin, R.J. (2005), *Exploration on the Syntax-Semantic Interface*. Cambridge: Cambridge University Press.
- Van Valin, R.J. (2009), *Lexical Representation, Co-Composition and Linking Syntax and Semantics*, in J. Pustejovsky et al. (eds) *New Developments in the Generative Lexicon*. Dordrecht: Kluwer.

Parte B. Variazione e uniformità nella realizzazione degli argomenti (Michela Cennamo)

Nella seconda parte del corso si approfondiranno alcuni aspetti della interazione tra l'aspetto inerente/idiosincratico (*root*) e strutturale (schemi eventivi) del significato dei verbi nel determinare la codifica morfosintattica degli argomenti, in relazione al dibattito contemporaneo sui principi generali che ne regolano l'integrazione (Levin & Rappaport Hovav 2005, Rappaport Hovav 2008, int.al.). Si presenteranno, in particolare, alcuni *case-studies* in prospettiva comparata: alternanza (anti)causativa (Schäfer 2008, Cennamo & Jezek 2009), intransitività scissa (Levin & Rappaport Hovav 1996, Sorace 2004, Aranovich 2007), codifica dell'oggetto (Levin 1999, Cennamo 2003, Malchukov 2008).

Riferimenti bibliografici preliminari

- Cennamo, M. (2003), *(In)transitivity and object marking: some current issues*, in G. Fiorentino (ed), *Romance Objects*, Berlin: Mouton de Gruyter, 49-104.
- Cennamo, M. & A. Sorace (2007), *Auxiliary selection and split intransitivity in Paduan*, in R. Aranovich (ed), *Split Auxiliary Systems. A Cross-linguistic Perspective*. Amsterdam: Benjamins, 65-99.
- Cennamo, M. & E. Jezek (2009), *The anticausative alternation in Italian: constraints and variation*, Workshop on Verb Typologies Revisited, Gent, 2-4 Febbraio 2009.
- Levin, B. (1999), *Objecthood: an event structure perspective*. CLS 35.1: 223-247.
- Levin, B. and M. Rappaport Hovav (1996), *Lexical Semantics and Syntactic Structure*, in S. Lappin (ed.), *The Handbook of Contemporary Semantic Theory*, Blackwell, Oxford, 487-507.
- Levin, B. & M. Rappaport Hovav (2005), *Argument Realization*, Cambridge: Cambridge University Press.
- Malchukov, A. (2008), *Animacy and asymmetries in differential case marking*, "Lingua" 118: 203-221.
- Rappaport Hovav (2008), *Lexicalized meaning and the internal temporal structure of events*, in S. Rothstein (ed), *Crosslinguistic and Theoretical Approaches to the Semantics of Aspect*. Amsterdam: Benjamins, 13-42.
- Schäfer, F. (2008), *The Syntax of (Anti-)causatives*. Amsterdam: Benjamins.
- Sorace, A. (2004), *Gradience at the lexicon-syntax interface: evidence from auxiliary selection*, in A.

Alexiadou, M. Everaert and E. Anagnostopoulou (eds.). *The Unaccusativity Puzzle*. Oxford: Oxford University Press, 243-268.

* *Michela Cennamo* (micennam@unina.it) è professore straordinario di Glottologia e Linguistica presso la Facoltà di Lettere dell'Università di Napoli Federico II. Si occupa prevalentemente di transitività e struttura argomentale, in prospettiva sincronico-diacronica e interlinguistica, di predicati complessi e di morfosintassi dei dialetti italiani.

** *Elisabetta Jezek* (jezek@unipv.it) insegna Sintassi e Semantica presso la Facoltà di Lettere dell'Università di Pavia. Si occupa prevalentemente di semantica lessicale, interfaccia lessico-sintassi, classificazione dei verbi, lessicografia computazionale.

Linguistica computazionale

[LABORATORIO]

Alessandro Lenci*

La disponibilità crescente di corpora di grandi dimensioni permette di avere a disposizione una quantità di dati prima inimmaginabile per l'indagine lessicale e la pratica lessicografica. La linguistica computazionale è il complemento essenziale a tale disponibilità di dati perché consente di estrarre informazione utile all'analisi linguistica e al tempo stesso permette di arricchire i testi con livelli di annotazione linguistica in grado di moltiplicarne le potenzialità di analisi. La piena consapevolezza dei metodi computazionali di rappresentazione ed estrazione dell'informazione lessicale è, dunque, una parte ineludibile del bagaglio di competenze del linguista e del lessicografo di nuova generazione. In questo corso, verranno presentati, in maniera chiara ed accessibile a tutti, i principali metodi computazionali per la ricerca semantico-lessicale e per la costruzione di risorse lessicografiche attraverso la ricerca e l'analisi di dati estratti da corpora testuali. Esistono oggi molti strumenti e ambienti per l'esplorazione dei corpora, di fronte ai quali però spesso il linguista si pone solo come mero utente passivo. È invece essenziale che il linguista comprenda a pieno anche il "dietro le quinte" di questi ambienti informatici. L'obiettivo del corso è infatti di formare un linguista come "fruitore attivo" delle nuove tecniche computazionali, in grado di acquisire una consapevolezza critica delle loro potenzialità.

Nel corso verranno affrontati i seguenti temi:

1. metodi per l'annotazione linguistica dei corpora (es. annotazione morfosintattica, lemmatizzazione, analisi sintattica);
3. metodi statistici per l'estrazione di informazione linguistica dai testi: es. collocazioni, espressioni polirematiche, strutture argomentali dei verbi, ecc.
3. induzione di classi paradigmatiche sulla base delle proprietà distribuzionali dei termini lessicali.

Il corso è integrato da un **laboratorio** nel quale gli studenti avranno la possibilità di applicare operativamente le nozioni di statistica lessicale all'analisi di dati estratti da corpora per la compilazione di entrate lessicali. Il laboratorio verrà realizzato in collaborazione con Francesca Masini e in coordinamento con il laboratorio di Lessicografia Computazionale da lei tenuto.

Riferimenti bibliografici

Lenci, A., Montemagni, S., Pirrelli, V. (2005), *Testo e computer. Elementi di linguistica computazionale*, Roma, Carocci.

* *Alessandro Lenci* (alessandro.lenci@ling.unipi.it) è ricercatore presso il Dipartimento di Linguistica "T. Bolelli" dell'Università di Pisa. Si occupa prevalentemente di linguistica computazionale e di semantica lessicale.

Lessicografia

[CORSO E LABORATORIO]

Vincenzo Lo Cascio*

Durante il corso sarà trattato il problema della redazione dei dizionari sia monolingui che bilingui o multilingui e verranno mostrati alcuni strumenti elettronici su cui essi si basano. Si presenteranno in particolare alcune teorie lessicologiche e lessicografiche che stanno alla base del trattamento del lessico, soprattutto in relazione alla costituzione di banche dati elettroniche e alla redazione dei dizionari del futuro. In particolare l'attenzione si poggierà sul lessico trattato non come un elenco statico di parole ma come una rete in cui a partire da una qualunque parola si costruisce il minisistema a cui essa appartiene.

Il lessico visto, dunque, non come elenco di forme isolate ma come formato da parole che si combinano in modo privilegiato con alcune della stessa lingua e non soltanto in base al rapporto categoriale e semantico ma soprattutto in base a scelte fatte dai parlanti nativi attraverso l'uso. I dizionari soprattutto quelli elettronici dovrebbero rispettare questa idea.

I partecipanti al corso oltre a studiare le diverse teorie lessicologiche e lessicografiche parteciperanno, come esercitazione, al lavoro di redazione di lemmi, secondo i criteri suddetti, per la costituzione di una possibile banca dati e la creazione di ipotetici dizionari (monolingue, bilingue o multilingue, delle collocazioni, per la produzione ecc.).

Riferimenti bibliografici

Lo Cascio, Vincenzo (a cura di) (2007), *Parole in Rete*, Novara, Utet Università.

Altre letture verranno indicate prima del corso.

* *Vincenzo Lo Cascio* (v.locascio@wanadoo.nl) è stato professore ordinario di Linguistica Italiana presso l'Università di Amsterdam fino al 2001. Si occupa prevalentemente di lessicologia, lessicografia, sintassi e semantica. Negli ultimi anni ha dedicato particolare attenzione alla realizzazione di dizionari bilingui elettronici e a stampa.

Lessicografia computazionale: introduzione allo SketchEngine

[LABORATORIO]

Francesca Masini*

Oggi giorno la lessicografia si serve di strumenti computazionali avanzati per la compilazione delle singole entrate dei dizionari e la raccolta dei dati in genere. Esistono infatti diversi *Dictionary Writing Systems* che permettono di ottimizzare la stesura dei dizionari e programmi di interrogazione e analisi di *corpora* che consentono di estrarre informazione in modo automatico e veloce.

In questo laboratorio i partecipanti impareranno ad usare uno dei principali programmi di interrogazione presenti attualmente sul mercato, lo SketchEngine (a cui si accede direttamente da internet:

<<http://www.sketchengine.co.uk/>>). Spiegheremo brevemente quali sono le principali funzionalità dello SketchEngine per poi passare alla parte pratica, in cui ogni partecipante dovrà formulare interrogazioni semplici e complesse sui *corpora* a disposizione e utilizzare i dati estratti per la compilazione di entrate.

Letture di riferimento

Atkins, Sue & Michael Rundell (2008), *The Oxford Guide to Practical Lexicography*, Oxford, Oxford University Press.

Kilgarriff, Adam, Pavel Rychly, Pavel Smrz & David Tugwell (2004), *The Sketch Engine, Proceedings of EURALEX*, Lorient, France, 105-116.

* *Francesca Masini* (fmagini@uniroma3.it) è dottore di ricerca in linguistica. Svolge attività di ricerca presso il TRIPLE (Università Roma Tre) e insegna Linguistica Generale presso l'Università di Bologna. Si occupa prevalentemente di lessico e semantica.

Basi teoriche di un dizionario analogico

[CONFERENZA]

Raffaele Simone*

Avendo progettato e diretto la realizzazione di un Dizionario Analogico (apparso per la UTET nel 2009) che si basa su una varietà di assunti teorici forti a proposito dei Formati Semantici (FS), illustrerà (a) il concetto generale di dizionario analogico, (b) la elaborazione teorica sottostante, (c) i problemi aperti nella definizione dei FS e delle relazioni tra FS diversi, (d) problemi non risolti nella rappresentazione dei FS.

* *Raffaele Simone* (simone@uniroma3.it) è professore ordinario di Linguistica Generale presso l'Università Roma Tre e direttore del laboratorio TRIPLE presso la stessa università. Si occupa prevalentemente di lessico, teoria della grammatica, teoria delle costruzioni e delle categorie, comparazione di lingue romanze, tipologia linguistica.