

La linguistica di *corpus* e sue applicazioni per la didattica della L2

Tommaso Raso (UFMG) e Cecilia Casini (USP)

La linguistica di *corpus* è una disciplina giovane, legata alle disponibilità tecnologiche che vengono costantemente sviluppate. Il Brasile, nonostante alcuni progetti pionieristici nel settore, è indietro rispetto al panorama internazionale più avanzato.

Saranno presentati due importanti progetti, uno già parzialmente concluso e uno in corso, che offrono alla ricerca e alla didattica linguistica brasiliane strumenti all'avanguardia nel settore.

Il primo progetto, che è già utilizzabile, è il *corpus* di parlato brasiliano C-ORAL-BRASIL (Raso e Mello, 2012), il quale, essendo comparabile con il progetto C-ORAL-ROM per portoghese europeo, italiano, spagnolo e francese (Cresti e Moneglia, 2005), permette un confronto immediato con altre quattro lingue.

Mostreremo alcuni studi che permettono di cogliere differenze importanti tra le varie lingue, tradizionalmente trascurate nell'insegnamento, anche per carenza di strumenti appropriati. In particolare, metteremo a confronto alcune delle principali misure del parlato in queste lingue, e soprattutto i dati relativi ai cosiddetti *discourse markers* nelle due varietà di portoghese, in italiano e in spagnolo, approfondendo soprattutto il confronto fra italiano e portoghese brasiliano, mettendo in rilievo le ricadute che questi dati possono avere nell'insegnamento dell'italiano a studenti brasiliani.

Il secondo è il progetto RIDIRE, finanziato dal MIUR, che prevede la realizzazione di una piattaforma online per la ricerca e l'utilizzazione didattica dell'italiano come L2. La piattaforma offre due importanti novità:

1. Lo scaricamento dalla rete di testi italiani per 2 miliardi di parole divise per ambiti semantici e funzionali. Lo scaricamento, contrariamente a quanto accade nei web corpora, prevede la pulizia di pubblicità, testi ripetuti, testi non adeguati agli obiettivi;
2. La realizzazione di un'interfaccia amichevole che venga realmente incontro alle esigenze della didattica (spesso senza strumenti adeguati per sfruttare le enormi potenzialità dei corpora).

Quest'ultimo obiettivo è affidato a due *user groups* esteri, di cui uno in Brasile (USP e UFMG). Nella conferenza si esporrà il lavoro fatto finora dal gruppo brasiliano.