

## Corpora CMC – Composizione e legenda

I corpora di nickname e testi sono stati raccolti per la seguente pubblicazione:

Matteo Casoni, *Italiano e dialetto al computer. Aspetti della comunicazione in blog e guestbook della Svizzera italiana*, Osservatorio linguistico della Svizzera italiana, Bellinzona, 2011.

www.ti.ch/olsi

### Corpus dei nickname

Il corpus copre un periodo di cinque anni, dal 2002 al 2007 ed è stato raccolto nel 2007 in *blog* e *guestbook* della Svizzera italiana. Accanto ai *nickname* sono indicate le seguenti caratteristiche.

*Nick*: la colonna comprende l'intera lista di *nickname*; l'acronimo [s.n.] indica i testi pubblicati senza *nickname*; nomi e cognomi propri sono stati abbreviati alle iniziali; versioni diverse del *nick* sono separate dal punto e virgola.

*Gen.*: genere del nome (m = maschile, f = femminile, x = genere non individuabile);

*Lin.*: lingua in cui è scritto il *nickname*; l'indicazione è data solo quando il *nick* è in una lingua diversa dall'italiano; *arab*, arabo; *dial*, dialetto della Svizzera italiana; *fr*, francese; *giap*, giapponese; *ing*, inglese; *it*, italiano; *lat*, latino; *mar*, marocchino; *nap*, napoletano; *por*, portoghese; *reg*, regionalismo, nome con apposizione dell'articolo (p. es. *il Gianni*); *rom*, romanesco; *sard*, sardo; *sic*, siciliano; *sla*, lingue slave; *spa*, spagnolo; *svted*, svizzero tedesco; *ted*, tedesco; *tosc*, toscano; *ven*, veneziano.

Accanto alla sigla della lingua è aggiunta a volte una lettera maiuscola che specifica:

*N*: *nick* contenenti il nome del gruppo di affiliazione dello scrivente quando il nome è in una lingua altra dall'italiano (per esempio *Bic Pincirö* è siglato *dialN* perché la parte dialettale del *nick* è data da *Pincirö*, nome del gruppo di carnevale di cui lo scrivente *Bic* è membro);

*P*: nomi di persona caratterizzati da una lingua diversa dall'italiano, per esempio per gli ipocoristici anglicizzati (*dany*), i nomi spagnoleggianti (*el pibe*) o dialettali (*Giuan* "Giovanni").

*&*: nomi contenenti il simbolo della *ampersand* o "e commerciale";

\*: l'asterisco indica una lingua maccheronica o una forma di parodia (p. es. *tetinum* è siglato *lat\**).

### Corpus dei testi con commutazione

Il corpus comprende i testi (*post*) che presentano fenomeni di commutazione in senso ampio; in grassetto sono evidenziate le parti di testo in lingue diverse dall'italiano. I materiali coprono un periodo di quattro, dal 2003 al 2007 e sono stati raccolti nel 2007 in *blog* e *guestbook* della Svizzera italiana. I testi sono stati resi anonimi al fine di garantire la protezione della sfera personale degli autori o di persone menzionate nei testi.

*Gb*: nome dei *Guestbook* in cui il *nick* è attestato.

*Zona*: zona della Svizzera italiana in cui ha sede l'associazione che pubblica il *Gb*.

*BL* = Blenio; *GR* = Grigioni; *LE* = Leventina; *LO* = Locarnese; *LU* = Luganese; *ME* = Mendrisiotto; *TI* = Ticino.

*L-LA*: lingue diverse dall'italiano attestate nel *post*:

*d*, dialetto della Svizzera italiana; *dber*, dialetto (o regionalismo) bergamasco; *dgen*, genovese; *dmil*, milanese; *dnap*, napoletano; *drom*, romanesco; *dromag*, romagnolo; *dsar*, sardo; *dsic*, siciliano; *dtos*, toscano; *dven*, veneto; *f*, Francese; *gia*, giapponese; *haw*, hawaiano; *ing*, inglese; *itstran*, italiano parlato da stranieri (in forme parodistiche); *lat*, latino; *nor*, norvegese; *por*, portoghese; *rum*, retoromancio; *rus*, russo; *s*, spagnolo; *st*, svizzero tedesco; *t*, tedesco. L'asterisco (per es. *lat\**) indica forme di pseudo-lingua.