Linguistica Computazionale – Lezione 4

Corpora linguistici & analisi automatiche

11 Marzo 2004
Cristiano Chesi, chesi@media.unisi.it

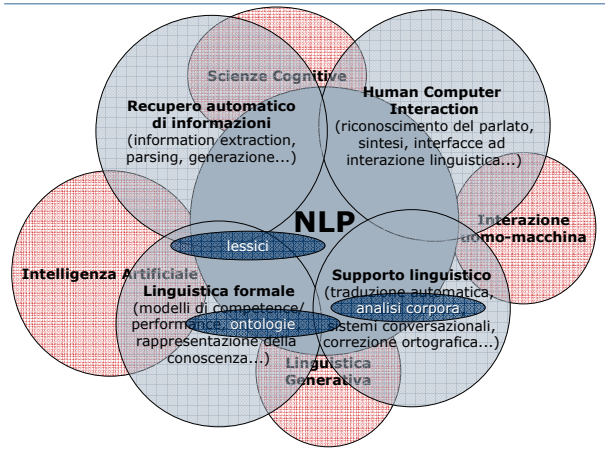
Corpora linguistici & analisi automatiche

- Indice
  - Corpora linguistici
    - motivazioni e struttura
    - database e corpora
  - Lessici computazionali
    - struttura generale
    - un esempio: Wordnet
  - Analisi automatiche
    - analisi morfologica
    - introduzione all'acquisizione del linguaggio

Letture, approfondimenti

- **Bibliografia essenziale**
  - Jurafsky & Martin (2000) *Speech & Language Processing*. Prentice Hall, NJ (Cap. 2)
  - Miller & al. (1993) *Introduction to WordNet: An On-line Lexical Database*. ms.
- **Approfondimenti**
  - Guasti (2002) *Language Acquisition*. MIT Press. (Cap.1-4)
  - MacWhinney & Snow (1985) *The child language exchange system*. Journal of Computational Lingustics, 12:271-296

Corpora, lessici e ontologie in rapporto al mondo del NLP



Cosa sono, che struttura hanno

- collezioni **finite** di informazioni, **omogenee** e **rappresentative** rispetto ad un dominio, raccolte in un modo **sistematico**, in **condizioni controllate** in modo da riflettere la **reale distribuzione** (quantitativa e qualitativa) dei fenomeni linguistici che si intendono studiare
  - **non strutturate** (unica informazione presente è l'informazione linguistica del testo)
    es. files di testo con formattazione non significativa (colonne, giustificazione...)
  - **strutturate** (convenzioni precise indicano la natura dei dati linguistici)
    es. database, testo taggato
  - **semistrutturate** (convenzioni implicite forniscono implicitamente informazioni (extra)linguistiche)
    es. pagine html, testi formattati (titoli, paragrafi, sottoparagrafi, grassetti, corsivi ecc.)

Esempio di corpus linguisticamente non strutturato

(inizio ...)

- **Brown corpus** (Francis and Kucera, 1964)
  - circa un milione di parole rappresentativo dell'inglese americano scritto (500 testi del 1961)
  - i testi sono raccolti in 15 categorie:
    - A. stampa: reportage (44 texts)
    - B. stampa: editoriali (27 texts)
    - C. stampa: periodici (17 texts)
    - D. religione (17 texts)
    - E. arti e mestieri (36 texts)
    - F. tradizioni popolari (48 texts)
    - ...
  - Esempio:
    A01 0010 The Fulton County Grand Jury said Friday an investigation
    A01 0020 of Atlanta's recent primary election produced "no evidence" that
    A01 0030 any irregularities took place. The jury further said in term-end
    A01 0040 presentments that the City Executive Committee, which had over-all
    A01 0050 charge of the election, "deserves the praise and thanks of the
    A01 0060 City of Atlanta" for the manner in which the election was conducted.

Esempio di corpus linguisticamente strutturato

(... continua ...)

- **Penn Treebank** (Marcus & al. , 1989-1992)
  - 1 milione di parole provenienti da articoli del 1989 del Wall Street Journal
  - Un campione di materiale proveniente da ATIS-3 (Automatic Terminal Information Service)
  - Etichettatura secondo lo "standard" Treebank II style
  - esempio:
    (S (PP (IN Of) (NP (NN course))) (, ,) (S (S (NP (DT some) (PP (IN of) (NP (PRP$ my) (NN color) (NNS values)))) (AUX (VBP do)) (NEG (RB not)) (VP (VB match) (NP (NP (DT the) (JJ old) (NN Master)) (POS 's)))) (CC and) (S (NP (DT the) (NN perspective)) (VP (VBZ is) (ADJP (JJ faulty)))) (CC but) (S (NP (PRP I)) (VP (VBP believe) (S (NP (PRP it)) (AUX (TO to)) (VP (VB be) (NP (DT a) (JJ fair) (NN copy))))))))

Esempio di corpus linguisticamente semi-strutturato

(... fine)

- **Childes** (MacWhinney & Snow, 1985)
  - (**Chi**ld **L**anguage **D**ata **E**xchange **S**ystem) è un archivio di trascrizioni spontanee di bambini (solitamente dai 14 mesi ai quattro anni di età) che interagiscono con adulti in varie situazioni. Generalmente ogni trascrizione si riferisce ad una conversazione di durata variabile dai 20 ai 60 minuti.
  - Le trascrizioni sono codificate secondo il formato standardizzato, detto **CHAT**

```
@UTF8
@Begin
@Participants:    CHI Cam Target_Child, DON Mother
@Age of CHI:      3;4.9
@Sex of CHI:      female
@Birth of CHI:    3-MAY-1988
@Date:            12-SEP-1991
*DON:             quale volevi ?
*CHI:             io volevo questo .
*DON:             si ma cosa, che canzoni ci sono, sopra .
*CHI:             non lo so .
*DON:             come non lo sai ?
                  [...]
@End
```

A cosa servono i corpora

- Informazioni quantitativamente significative per implementare **sistemi esperti** o per l'**estrazione di grammatiche**
  - registrazioni telefoniche (call center)
  - corpora taggati
  - ...
- **Analisi linguistiche** specifiche
  - acquisizione prima lingua
  - acquisizione seconda lingua
  - soggetti con disturbi linguistici (Specific Language Impairment, sordi, afasici ...)
  - ...

Database relazionali

- **tabelle**

studenti

ID	nome	cognome	mail	...
1	aldo	rossi	aldo@...	
2	giovanni	bianchi	gio@...	
3	giacomo	verdi	gia@...	
...				

- **relazioni (o link relazionali o join)**

studenti

ID	nome	cognome	stato
1	aldo	rossi	1
2	giovanni	bianchi	2

iscrizione

ID	descrizione
1	presente
2	assente

Database relazionali - normalizzazione

- **procedura di normalizzazione**

  far in modo che ogni informazione **occorra una sola volta** nella struttura dati.

  Si preferisce quindi far riferimento ad un'unica istanza di questa informazione (via **link relazionale**) piuttosto che copiare la stessa informazione in più posizioni (risolve il problema dell' "**update anomaly**": alcune istanze dell'informazione sono aggiornate, mentre altre sfuggono al controllo e diventano obsolete)

Interrogazioni su corpus testuali

- **Espressioni Regolari**
  - notazione algebrica per definire insiemi di stringhe di testo (linguaggi di **tipo 3**).
  - Il cuore dell'espressione regolare è il **pattern di identificazione** composto da caratteri alfanumerici (compresi segni di spaziatura e di interpunzione) e da segni speciali volti a stabilire le relazioni tra i caratteri del pattern.

Espressione Regolare	Corrispondenza	Es. pattern identificato
[Dd]uomo	Duomo oppure duomo	Il duomo è nella piazza
[^a-z]	tutto fuorché lettere minuscole	Il duomo è ...
sali?ta	salita oppure salta	Marco deve saltare
sal.ta	accetta ogni carattere tra le i e la t	Marco saluta
bu*	b seguito da un numero imprecisato (anche nullo) di u	buuuuuu! oppure b!
^L Vs. a$	^ = inizio stringa; $ = fine stringa	La casa
cas(a|e)	| è equivalente alla disgiunzione logica	Marco vive in un casale
*	il backslash è il simbolo di escape	A*

Note su Eliza

- **Espressioni Regolari e l'operazione di Sostituzione**
  - La **sostituzione** è un'operazione che permette di sostituire l'occorrenza di un'espressione regolare con un'altra espressione regolare e può essere definita come segue:
  - **s/espressione_regolare1/espressione_regolare2/**
  - **s/ www\.[a-z]*\.com / www\.pe{2}\.com/**
- **Registri**: se si usano più blocchi di operatori (ogni parentesi tonda delimita un blocco), si può riutilizzare l'espressione trovata da un determinato blocco nell'espressione da sostituire, facendo riferimento all'ordine dei blocchi nel pattern di ricerca:
  - **s/ la (casa|macchina) è stata comprata da (Maria|Gianni)/ \2 ha comprato la \1 /**
    permette di costruire la forma attiva (Gianni ha comprato la casa) della frase passiva (la casa è stata comprata da Gianni).
- operazioni di sostituzione in ELIZA:
  - s/ sono [.* |](depress[o|a]|triste)/sono spiacente di sapere che sei \1/
  - s/ sono tutt[i|e] (.*) /in che senso sono \1?/
  - s/ sempre / potresti far riferimento ad un esempio specifico?

Database relazionali - Interrogazione

- **Structured Query Language (SQL)**
  Linguaggio di interrogazione e manipolazione di database
  - **Data Manipulation Language (DML)**
    - **select**
      select * from studenti where id > 1
    - **insert**
      insert into studenti (nome, cognome) values ("mario", "rossi")
    - **delete**
      delete from studenti where id=10
    - **update**
      update studenti set nome="gianni" where id=11
  - **Data Definition Language (DDL)**
    - **create database**
    - **create table**
    - **drop database**

Strutture dati ed OGGETTI

- **Oggetto**
  entità costituita da **proprietà** (il **valore** assegnato a tale proprietà in un determinato istante di tempo determina lo **stato** dell'oggetto) e **comportamenti** (metodi o procedure di modificazione dei dati proprie dell'oggetto).
  Un oggetto complesso è un oggetto costituito da altri oggetti (funzione di **estensione**).
- **Identità**
  si dice **Object IDentifier** (**OID**) l'identificativo unico dell'oggetto, indipendente dai valori che tale oggetto assume (avere lo stesso identificativo è diverso dall'avere gli stessi valori!)

Strutture dati ed OGGETTI – EREDITARIETA'

- Gli oggetti possono essere definiti gerarchicamente: ogni oggetto gerarchicamente inferiore eredita proprietà e metodi dagli oggetti padri
- Nuovi metodi/proprietà possono essere inclusi nei figli, gli stessi metodi ereditati possono essere ridefiniti (**overriding**)
- si può parlare di ereditarietà semplice (classi>sottoclassi) oppure ereditarietà multipla (reticolo aciclico)

Lessico "generativo"

(inizio ...)

- perché ogni entrata lessicale non può essere registrata **singolarmente**?
  - sarebbe **inefficiente**:

	mangi -	**sogn -**	**corr -**	**puff -**
-are/ere	**mangi-**are	**sogn-**are	**corr-**ere	**puff**-are
-o	**mangi-**o	**sogn-**o	**corr-**o	**puff-**o
-ato	**mangi-**ato	**sogn-**ato	***corr-**ato **(**corso**)**	**puff-**ato

  - in Turco (lingua agglutinante) ci sarebbero circa 600x10^6 entrate lessicali da considerare. In Finlandese 10^7
  - sarebbe **non informativo**:
    - **nessuna relazione significativa** tra entrate lessicali (l'unica relazione possibile sarebbe l'ordine alfabetico, ma *casa* e *case* hanno intuitivamente una relazione più "intima" rispetto a quella tra *case* e *caso*)
    - non esisterebbe **nessun indizio per processare** in modo "particolare" ad esempio un verbo rispetto ad un nome

Lessico "generativo"

(... continua ...)

- classicamente un lessico computazionale era concepito in funzione delle applicazioni da cui doveva essere usato
  - **lessico mentale** - modelli psicologicamente plausibili di relazioni tra unità minime di significato
  - **modelli computazionali** - usati per creare i database lessicali efficienti.
- regole di efficienza computazionale:
  - la rappresentazione lessicale deve essere **esplicita** ed **indipendente** dalle applicazioni che la utilizzeranno
  - la **struttura globale** delle entrate lessicale è importante almeno quanto la **struttura interna** delle singole parole: la sua organicità e significatività serve a rappresentare una complessa base di conoscenza (**ontologia**)
  - il lessico deve essere in grado di coprire adeguatamente il suo **dominio** (approssimativamente 400.000 entrate lessicali di cui 5.000 entrate verbali, 30.000 nominali, 5.000 aggettivali, un migliaio di avverbiali, altrettanti termini logici, 2.000 composti e 300.000 nomi propri + vari termini dominio-specifici)

Lessico "generativo"

(... continua ...)

- regole di efficienza computazionale (...continua):
  - i lessici computazionali devono essere valutabili almeno su tre scale:
    - **copertura** (sia in estensione, che in profondità, a livello di ricchezza dell'informazione)
    - **estensibilità** (deve poter essere formalmente possibile arricchire il vocabolario con termini nuovi)
    - **utilità** (stavolta valutata a livello delle singole applicazioni/elaborazioni)
- da ricordare:
  - la **completezza** non assicura la **correttezza** (psicolinguistica e computazionale)
  - la **plausibilità psicolinguistica** non garantisce l' **efficienza computazionale** e viceversa

Lessico generativo - Struttura di una singola entrata lessicale

- informazioni **ortografiche/fonetiche** (devono insomma codificare l'input nel modo più adeguato possibile)
- **morfologiche** (tratti inerenti, quali plurale/singolare, massa/contabile, animato/inanimato...)
- **sintattiche** (categoria grammaticale ed eventualmente la sottocategoria)
- **semantiche** (sia a livello di selezione semantica, che di significato ai fini della traduzione ad esempio)

CASA:
<C,A,S,A>
{N, singolare, femminile ...}
{N comune ...}
[house]

Lessico "generativo"

(... fine)

- struttura **struttura globale del lessico**
  - correlare la sottocategorizzazione con la **classe semantica** (Levin 93 propone una vasta serie di **classi di alternanza** cercando di dimostrare che certi comportamenti sintattici verbali, quali l'assegnazione di ruoli tematici, sono prevedibili sulla base di certi tratti semantici minimalmente distintivi, quali la modificazione di stato, la causatività, la relazione tra gli elementi in azione ecc.)
  - trarre immediate **inferenze** in base all'organizzazione gerarchica degli items (**part_of**, **member_of**…)

Esempio di struttura globale di un lessico: le reti concettuali o semantiche

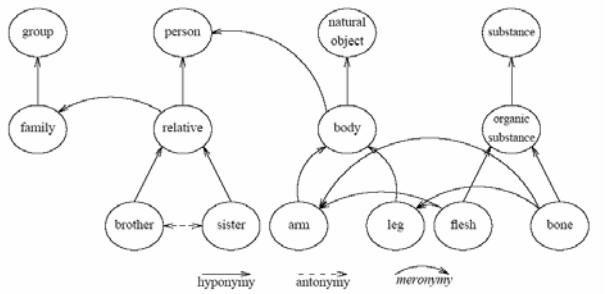
(inizio ...)

- **Wordnet** (Miller 90)
  - interessante esempio di rete semantica (scopo: organizzare il lessico sulla base del **significato** delle parole piuttosto che sulla base della loro **ortografia**) basata sui seguenti principi:
    - certe relazioni semantiche tra **nomi** (gerarchie ad eredità), **verbi** (implicazioni), **aggettivi** e **avverbi** (opposizioni) ma non tra **parole funzionali**, sono psicolinguisticamente significative
    - ogni concetto lessicale (**synset**) può essere **rappresentato dai suoi sinonimi** (altri synset)
    - es. di relazioni:
      - **iponimia** (relazione tra un concetto generale ed uno più specifico; ad esempio "pettirosso" è un iponimo di "uccello")
      - **iperonimia** (relazione inversa all'iponimia)
      - **meronimia** (parte_di)...
    - attraverso l'uso di synset distinti si affronta il problema della **polisemia** (*cane* = animale domestico e *cane* = parte metallica di una pistola saranno due nodi distinti di wordnet anche se si scrivono allo stesso modo)

Esempio di struttura globale di un lessico: le reti concettuali o semantiche

(... fine)

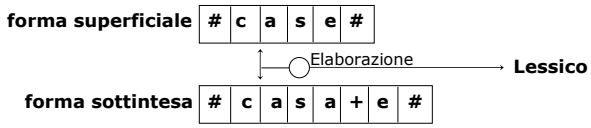
- **Esempio di relazioni semantiche** (Miller 1993)



Analisi morfologica – modello teorico

(inizio ...)

- **obiettivo**: riconoscere una stringa ben formata di caratteri e metterla in relazione con la struttura dei morfemi che la compongono; questo compito ci permette di introdurre tutti i problemi che si presenteranno nel parsing delle strutture frasali
- **modello teorico**:



Analisi morfologica – modelli computazionali

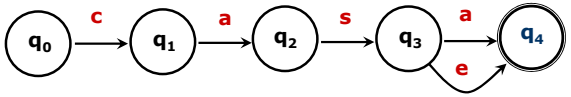
(... continua ...)

- **Finite-State Automata** (**FSA**)
  definiti come quintuple **<Q, Σ, q_0, F, δ>** dove:
  - **Q** = insieme finito e non nullo di stati
  - **Σ** = alfabeto finito e non nullo di caratteri accettabili in input
  - **q_0** = stato iniziale, con $q_0 \in Q$
  - **F** = insieme di stati finali, con $F \in Q$
  - **δ** = insieme delle regole di transizione definite in QxΣ su Q

Analisi morfologica – modelli computazionali

(... continua ...)

- un insieme di FSA non è solo un insieme di macchine che permettono di **riconoscere** o **rifiutare** un elemento lessicale, ma anche di **rappresentare** l'intero lessico.
- FSA che riconosce la parola *casa* ed il suo plurale:



Q = {q_0, q_1, q_2, q_3, q_4},
Σ = {c,a,s,e,#},
Q_0={q_0},
F ={q_4},
δ =

	q_0	q_1	q_2	q_3	q_4
c	q_1				
a		q_2		q_4	
s			q_3		
e				q_4	

Analisi morfologica – modelli computazionali

(... continua ...)

- **Limiti degli FSA**
  per **associare** una **descrizione strutturale** ad un elemento riconosciuto come appartenente al lessico, i semplici FSA non sono più sufficienti (non esiste una memoria esterna, se non la memoria implicita data dallo stato in cui si trova l'automa, in cui "conservare" il percorso e la struttura esaminata).
- Koskenniemi (83) propone un modello di morfologia a due livelli (**two-level morphology**): un **livello lessicale** ed uno **superficiale** che devono essere messi in una qualche relazione significativa dal punto di vista morfologico.
- Tale modello è implementabile utilizzando i **Finite-State Transducers** (**FST**, o **Trasduttori**)

Analisi morfologica – modelli computazionali

(... continua ...)

- **Finite-State Transducers** (**FST**, o **Trasduttori**)
  definiti come quintuple **<Q, Σ, q_0, F, δ>**, dove però sussistono alcune sostanziali differenze rispetto agli FSA:
  - **Σ** = alfabeto finito e non nullo di *caratteri complessi* accettabili in input della forma i:o dove i sono i simboli dell'alfabeto I di input e o simboli dell'alfabeto O di output. $\Sigma \in I \times O$. ε (l'elemento nullo) può essere incluso sia in I che in O
  - **δ** = è definita come (q , i : o) e rappresenta la matrice di transizione che mette in relazione uno stato q di partenza e uno stato q' di arrivo se la relazione i : o è definita. δ è quindi una relazione da Q x Σ su Q
- i trasduttori hanno funzioni più generali degli SFA: questi ultimi descrivono un linguaggio formale definendo un insieme di stringhe ben formate, gli FST definiscono invece **relazioni** tra insiemi diversi di stringhe.

Analisi morfologica – modelli computazionali

(... continua ...)

- In particolare gli FST possono essere usati come **riconoscitori**, **generatori**, **traduttori**, **correlatori tra insiemi**.
- alcune proprietà di cui gli FST godono sono:
  - l'**inversione**, definita come T^{-1}, scambia le etichette di input con quelle di output
  - la **composizione**, se T_1 mappa I_1 su O_1 e T_2 è un trasduttore da I_2 ad O_2, $T_1 \circ T_2$ mappa I_1 in O_2.

Analisi morfologica – esempi di FST

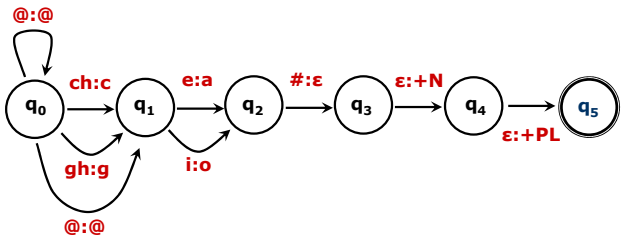
(... continua ...)

- problema di **morfologia flessiva**: definire un FST che descriva il fenomeno dei plurali in italiano.
  - **rappresentazione del problema**
    esempi: casa > case; donna > donne; gatto> gatti; ago > aghi; sacco > sacchi ...
  - **intuizioni e generalizzazioni**
    i nomi femminili prendono il plurale in *e*, i maschili in *i*. *c* e *g* diventano rispettivamente *ch* e *gh* al plurale.
  - **formalizzazione**
    caso regolare: nome maschile > @:@ c|g|@:ch|gh|@ o:i
    nome femminile > @:@ c|g|@:ch|gh|@ a:e
    caso irregolare: uomo > @:@ o:i #:n #:i
  - **implementazione**
    nome femminile > @:@ c|g|@:ch|gh|@ e:a #:ε #:+N #:+PL
    es. case > casa +N +PL (c:c a:a s:s e:a #:ε #:+N #:+PL)

Analisi morfologica – esempi di FST

(... continua ...)

- **FST** (approssimativo) per descrivere i plurali regolari in italiano:



Analisi morfologica – (in)adeguatezza di FSA e FST

(... continua ...)

- certe lingue mostrano fenomeni più problematici di quelli appena descritti. Tali fenomeni sono detti di **morfologia non-concatenativa**
- **Tagalog** (un dialetto parlato nelle Filippine), **infissi** nel mezzo della parola:
  um (marca l'agente dell'azione) + **hingi** (prestare) =
  h-um-ingi
- **Lingue semitiche**, morfologia a modelli (**templatic morphology**):
  radici verbali composte da consonanti (CCC) **lmd** (apprendere) + flessioni in schemi vocalici (CVCVC) =
  lamad (studiò)
  lumad (fu insegnato)

Analisi morfologica – (in)adeguatezza di FSA e FST

(... fine)

- **Problemi** incontrati:
  - **non-determinismo** (due o più percorsi possono essere innescati dallo stesso carattere allo stato q; transizioni ε)
  - **inadeguatezza** del modello per trattare fenomeni morfologici complessi
  - **ordine** di applicazione degli FSA (o delle regole a seconda dei punti di vista)

Analisi morfologica – Alcune applicazioni

(inizio ...)

- **Ricerca di informazioni**
  (web, archivio digitale strutturato e non)
  - **Keywords** combinate con operatori booleani
    (alberghi & Firenze)
  - **Stemming**
    si cerca di ricavare la radice (*stem*) delle parole da cercare in modo da effettuare ricerche più complete e tolleranti (es. da "alberghi & Firenze" si può generare una query (alberghi AND Firenze) OR (albergo AND Firenze)).

Analisi morfologica – Alcune applicazioni

(... fine)

- L'algoritmo di **Porter Stemming**
  (Porter dal nome del suo ideatore)

  semplice serie di FST a cascata per l'inglese del tipo:
  - ATIONAL -> ATE (es. relational -> relate)
  - ING -> ε (talking -> talk)

- pro e contro:
  - **ipergeneralizzazione** (Krovetz 93)
    es. organization > organ, generalization > generic,
  - **non cattura generalizzazioni** corrette:
    matrices > matrix o European > Europe.
  - vantaggio nell'uso dello stemming solo quando la ricerca è espansiva

Analisi morfologica – Plausibilità psicolinguistica

(inizio ...)

- **Come è strutturato il lessico mentale?**
  - **full listing hypothesis -** *correre*, *corre* e *ha corso*, sono entrate distinte nel lessico mentale (nessuna struttura morfologica interna)
  - **minimum redundancy** - solo i morfemi costituenti sono compresi nel lessico umano; quando si ha accesso ad una parola come corre in realtà si ha accesso a due morfemi (*corr-* radice ed *-e* terza persona sing. presente) che poi vengono combinati tra di loro

Analisi morfologica – Plausibilità psicolinguistica

(... fine)

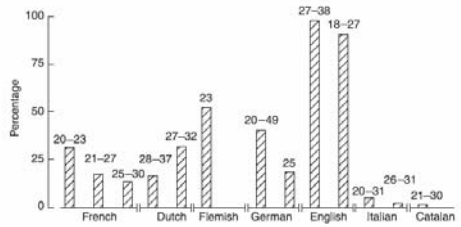
- **Evidenze sulla struttura del lessico mentale**
  - **Effetti di priming** (Stanners ad al. 79)
    flessioni irregolari: *happiness*, *happily* no priming con la radice *happy* Vs. flessioni regolari *pouring* > *pour*
  - **Affinità semantica** (Marslen-Wilson 94)
    government > *govern*
  - **Errori di pronuncia** (Fromkin e Ratner 98)
    easy enoughly*** invece di ˚*easi**ly** enough*
- Questo sembra suggerire che il lessico mentale debba contenere alcune informazioni sulla struttura morfologica delle parole rappresentate.

Morfologia e acquisizione del linguaggio

- **Cos'è l'acquisizione del linguaggio**
  - **dato:**
    - **S_0** (stato iniziale universale, biologicamente determinato)
    - **S_L** (stadio di sviluppo linguistico adulto, relativamente stabile)
    - **esempi empirici positivi** (Brown e Hanlon 1970, Pinker 1979:226)
  - **S_0 -> S_1 -> … -> S_L** (stadi di "sviluppo linguistico")
  - prime fasi dell'acquisizione del linguaggio:
    - solo nomi (Gentner 1982, Caselli & al. 1995)
    - tra i 10 ed i 12 mesi i bambini iniziano a combinare parole e significato
    - tra i 20 ed i 24 mesi si assiste ad un'esplosione linguistica (in termini di vocabolario, prime espressioni con più parole, distribuzione sistematica degli elementi all'interno della frase)

L'acquisizione del linguaggio – fenomeni interessanti

- **Iperregolarizzazione**
  cadde > cadò
- **Bare nouns**
  voglio la bambola > voglio bambola
- **Root infinitives** (Guasti 1993-94, 2002)



Concetti chiave della lezione di oggi

- Corpora linguistici
  - corpora strutturati/non strutturati e database
  - interrogazioni (espressioni regolari, SQL)
- Lessici computazionali
  - struttura delle entrate lessicali
  - struttura del lessico (relazioni tra entrate lessicali, es. Wordnet)
- Analisi morfologica
  - automi a stati finiti
  - finite-state transducers
  - applicazioni (es. stemming)
- Accenni all'acquisizione del linguaggio

Prossima lezione

(Giovedì 17 Marzo, ore 15-18, Aula Workshop 2, S. Francesco)

- Note sulla teoria formale dell'apprendibilità
  - formalizzazione del problema
  - l'apprendibilità nel limite
  - alcuni utili risultati
- L'apprendibilità delle lingue naturali
  - l'acquisizione della prima lingua
  - dagli algoritmi alle euristiche
  - modelli e problematiche