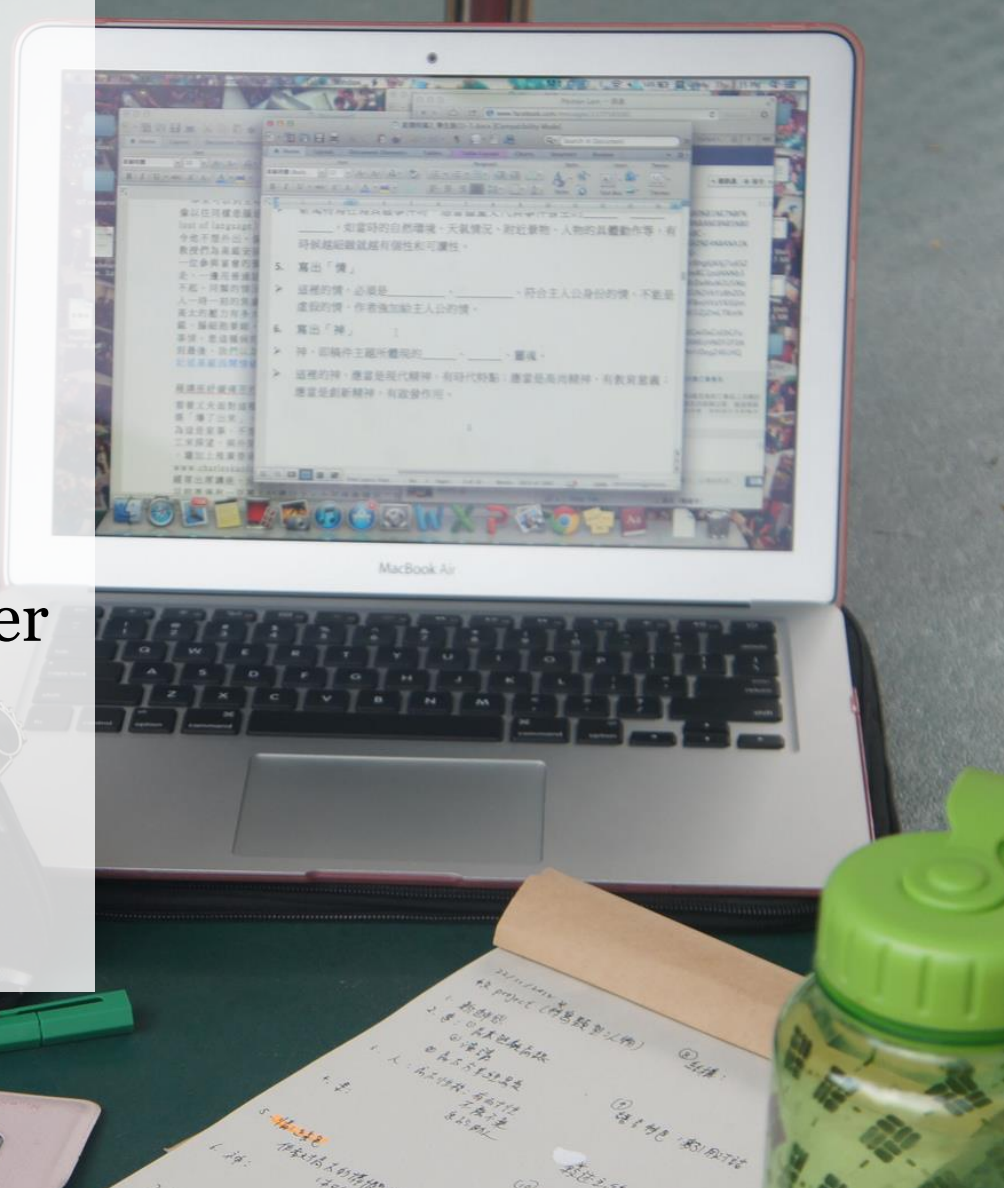


# Modulo A

## Valutare i sistemi automatici di interazione linguistica

17. Progettare un corpus per il futuro

1 dicembre 2016



Linguistica italiana II  
Mirko Tavosanis  
A. a. 2016-2017

# Calendario

- Fino a domani alle 13: scegliere tra Letto e Dialogo
- All'inizio della prossima settimana vi chiederò di scegliere anche i file (che rivedrò caso per caso)

# Progettare un corpus per le interazioni con assistenti digitali

- Come minimo, viene un dubbio: come si possono documentare interazioni che ancora non esistono?
- Una strada seguita è stata quella delle interazioni simulate
- Tuttavia, anche la simulazione deve essere ragionevole da due punti di vista:
  - Ciò che può fare il sistema
  - Ciò che è realistico dal punto di vista linguistico
- Vediamo altre metriche, poi torniamo sul corpus ADAM

# Metriche

- Innanzitutto: le prestazioni dipendono dal tipo di testo
- Alcune aree di attività sono collegate ad applicazioni dedicate con percentuali di successo molto alte; in particolare, sembra funzionino bene le trascrizioni di tipo:
  - Medico (referti)
  - Giudiziario (trascrizioni di testimonianze e simili)
- In questi settori il WER viene usato normalmente come indicatore di affidabilità
- In Italia adesso alcuni tribunali danno in affidamento le trascrizioni ad aziende che promettono WER superiore al 90%...  
Però su questo ho solo informazioni indirette

# Limiti del WER

- Xiaodong He, Li Deng, Alex Acero, *Why word error rate is not a good metric for speech recognizer training for the speech translation task?*, IEEE, 2011  
(<http://dx.doi.org/10.1109/ICASSP.2011.5947637>).
- In sostanza: non serve a molto cercare di abbassare la soglia d'errore affrontando parole che non dicono molto agli esseri umani
- Proposto: usare le metriche della traduzione
- Del resto, è ovvio che gli esseri umani sono capaci di ricavare moltissime cose da un testo anche molto danneggiato

# Ortografia devastata, testo leggibile

DXMXNXCX 2 mxrzx 2008 lx prxm Gxxrnxtx  
Nxzxxnxx dxll Fxrrxvxx Dxmntxcctx.  
L'nxzxxxtvx, prxmssx dxll Cxnfdxrxzxxnx pr  
lx Mxbxlxtx Dxlcx (CxMxDx), xntxndx prtxrx  
xll'xttnzxxnx dxll'xpxnxxnx pxbblxcx x dxglx  
xmmnxstrtxrx...

Per alcuni è più difficile, ma resta comunque  
leggibile, se cambiamo la forma della parola:  
DMNC 2 mrz 2008 l prm Grnt Nznl dll Frrv  
Dmntct. L'nztv, prmss dll Cnfdrzn pr l Mblt Dlc  
(CMD), ntnd prtr ll'ttnzn dll'pnn pbble dgl  
mmnstrtr..

# Per scopi pratici: interazione

- Per gli assistenti digitali, più che esaminare eventuali trascrizioni ha senso valutare le **percentuali di completamento del compito**
- Del resto, abbiamo visto che «sveglia otto» funziona con Siri tanto quanto «metti una sveglia per le otto»
- I dettagli linguistici possono essere considerati poco influenti, soprattutto in un contesto di interazione semplificata... come quelli che vedremo tra poco

# Il corpus ADAM

- «Il dominio applicativo e' quello turistico; inoltre i dialoghi raccolti sono due tipi: dialoghi che si realizzano tra un utente ed una macchina (denominati dialoghi persona-macchina) e dialoghi che hanno luogo tra due persone (denominati dialoghi persona-persona). Le conversazioni nei dialoghi persona-macchina vertono sulle richieste di orari di collegamento aerei o treni; le conversazioni nei dialoghi persona-persona riguardano l'organizzazione di un viaggio, comprensivo sia dei trasporti che di prenotazioni alberghiere e di informazioni turistiche.»
- In particolare, per l'interazione automatica è stato usato **FS-Inforna**, sistema telefonico sviluppato alla fine degli anni Novanta da Loquendo per le Ferrovie dello Stato all'interno del progetto europeo ARISE
- Il confronto tra i due tipi di dialogo è interessante, specie nella prospettiva della persona che chiede informazioni



Dialogue id: **dial\_001**

Dialogue Type: **hnm**

t\_001A: globetrotter viaggi buongiorno [*puff*]

---

t\_002B: [*breath*] buongiorno [...] senta io vorrei(i) [**vorrei**] prenotare(e) [**prenotare**] un(n) [**un**] viaggio in treno [...] [*breath*] da roma a verona [...]

---

t\_003A: un treno da roma a verona [*breath*] [...]

---

t\_004B: si` [...] per {*e*} [...] le(e) [**le**] otto [*puff*] cioe` partendo [...] [*cough*] verso le otto di lunedì` quindici settembre [...]

---

t\_005A: [*breath*] [...] allora lunedì` [...] quindici settembre [*puff*] [...] {*e*} otto di mattina [*puff*] [...]

---

t\_006B: si` [*puff*]

---

t\_007A: verso le otto di lunedì` quindici settembre [...] da roma [...] a verona [*breath*] beh avremmo un [...] treno diretto [*breath*] che parte all- [**alle**] otto un quarto [...] e arriva a verona alle dodici e cinquanta [*puff*] [*mouth*] [...] pensa che le vada bene

---

t\_008B: si` {*e*} mi puo` dire quanto costa(a) [**costa**] [*puff*]

---

t\_009A: [*breath*] in prima n [**in**] seconda classe [*puff*]

---

# Interazione con ARISE

t\_001S: il servizio e` automatico [...] e fornisce gli orari dei treni [...] che collegano le principali citta` italiane ed alcune citta` europee [...] risponda alle domande del sistema [...] dopo il segnale acustico [...] ora [...] mi dica da dove parte e dove vuole andare [...]

t\_002U: orte roma termini

t\_003S: vuole andare da orte a roma termini

t\_004U: si`

t\_005S: parte oggi

t\_006U: si`

t\_007S: attenda [...] a che ora desidera partire

t\_008U: sette

t\_009S: parte intorno alle sette

t\_010U: si`

t\_011S: attenda [...] attenda in linea

## Dial\_202

t\_002U: parto da torino e vorrei andare a bologna

t\_003S: vuole andare da torino a bologna

t\_004U: si'

t\_005S: parte oggi

t\_006U: no

t\_007S: puo' indicarmi **il giorno e l' orario** in cui vuole viaggiare

t\_008U: vorrei partire i l' otto ottobre verso sera

t\_009S: **in che giorno desidera partire**

t\_010U: l' otto ottobre

t\_011S: parte l' otto ottobre intorno alle otto di sera

t\_012U: si'

t\_013S: attenda [...] attenda in linea

# Interazioni persona-macchina

- Non è detto che debbano essere complesse
- Gli esseri umani sono abbastanza abituati a esprimersi in linguaggio semplificato e asimmetrico (padroneggiato pienamente dall'emittente ma non dal ricevente); due esempi sono:
  - Baby talk
  - Foreigner talk
- Questi tipi di lingua possono essere anche appresi spontaneamente
- È concepibile un “computer talk”? Ovviamente sì – ed esiste già; la domanda è quanto sia formalizzato da una parte, natural da usare dall'altra

# Baby talk

- La lingua semplificata con cui ci si rivolge ai bambini (v. Bernini, voce *Baby talk* nell'*Enciclopedia dell'italiano*)
- Non tutti la usano, e non è chiaro se sia più comprensibile ai bambini
- Comprende:
  - a. una componente affettiva (e protettiva) sul piano del rapporto sociale, contraddistinto dalle necessità di sostentamento del bambino – presente anche nel dialogo con piccoli animali, innamorati, ecc.
  - b. una componente linguistica derivata dall'asimmetria nella padronanza della lingua utilizzata, piena per l'adulto e solo agli esordi nel bambino.

- (a) *Livello prosodico*: volume relativamente alto, modulazione più alta della voce, articolazione lenta e accurata delle sillabe con enfaticizzazione della curva intonativa;
- (b) *Livello fonologico*: riduzione di nessi consonantici e vocalici, tendenza alla costituzione di sole sillabe CV e CVC, abbreviazioni (*baccio* «braccio», *quetto* «questo», *pangere* «piangere», *cappa* «scarpa», *elicotto* «elicottero»);
- (c) *Livello morfosintattico*:
- (i) diffusa presenza di suffissi diminutivi coi nomi (*acquina*, *cagnolino*, *lattuccio*, *orsettin*) e, in Toscana, desinenza -i per nomi masch. sing. (*chichi* «dolce», *cicci* «carne»);
- (ii) riduzione preferenziale della morfologia verbale alla terza pers. sing. e alla prima plur. del pres. e dell'imperf. indicativo:
- (1) Laura [la bambina destinatario del messaggio] schiaccia quello col ciufciuf «Laura! Schiaccia il tasto dove è raffigurato un treno!»
- (2) prendiamo quetto e facciamo una torre alta alta; voleva il dolcino il mio bambino?
- (iii) talvolta, omissione di elementi grammaticali (*bae fine / bae finito* «Vale ha finito»; *male manina?* «[ti sei fatto] male [alla] mano?»);
- (d) *Livello lessicale*: gli elementi lessicali sono riferiti agli ambiti rilevanti per il bambino (le parti e le funzioni del corpo, i rapporti di parentela, il cibo, i giochi, gli animali, certe qualità). Certi elementi lessicali sono costituiti dal raddoppiamento della stessa sillaba (*pappa*, *popò*, *pipì*) anche con valore iconico rispetto al referente, di cui riproducono proprietà sonore (*ciufciuf* «treno», *brumbrum* «automobile», *baubau* «cane»). Tra i verbi sono preminenti quelli più generici, come *fare* (*fare la nanna*, *fare la pipì*, *fare ciao ciao* «salutare»);
- (e) *Livello discorsivo*: frequenza di elementi deittici che collegano il discorso all'ambiente in cui esso è prodotto (avv. e pron. dimostr.); espressioni di rinforzo delle verbalizzazioni del bambino (*sì*, *bravo*) e di guida della sua attenzione (*guarda*, *sentì*); ridondanza (*dov'è la palla? ecco la palla, tirala qui*).

# Foreigner talk

- La lingua semplificata con cui ci si rivolge agli stranieri (v. Bernini, voce *Foreigner talk* nell'*Enciclopedia dell'italiano*)
- In italiano una volta aveva alcune caratteristiche tradizionali, come l'uso dei verbi all'infinito (eredità della lingua franca mediterranea)
- Non porta ovviamente agli stessi esiti del baby talk, ma è comunque abbastanza complesso

# Due “strategie di adattamento”

## 1. **Introdurre** nel messaggio alcune **peculiarità**:

- (a) allungamento delle frasi;
- (b) uso di parafrasi in luogo di parole ritenute difficili;
- (c) aggiunta di informazioni che permettano di contestualizzare meglio un elemento;
- (d) uso della deissi (per es., *quando tu + quando tu eri piccolo ++ giocavi a calcio cioè giocavi con una palla +++ football* «quando eravate ragazzi, giocavate a calcio?»).

## 2. **Ridurre la complessità** del messaggio a livello **lessicale** e **morfosintattico**. In tal modo si può arrivare a produrre perfino enunciati agrammaticali, contrassegno estremo del *foreigner talk* rispetto al modo di parlare che si usa coi nativi. Il seguente esempio, in cui un parlante nativo riformula il proprio enunciato per renderlo più comprensibile a due stranieri, illustra il particolare intreccio di strategie di elaborazione:

- (1) ma, se voi vi trovaste, o tu A. o tu M., nella stessa situazione che cosa fareste? mettiamo che voi arrivate in un campeggio, volete affittare una roulotte e non trovate più i documenti. Cosa fareste?

Qui si notano, per es.: contestualizzazione (*campeggio, roulotte* rispetto a *situazione*) e semplificazione (mancanza della subordinazione, indicativo al posto del congiuntivo in *arrivate*).



# Aspetti di semplificazione

- (a) sono cancellati il verbo *essere* e le preposizioni (*io Natale casa* «a Natale sono stato a casa»);
- (b) è usato il pronome tonico soggetto e i clitici sono sostituiti con pronomi tonici:  
(2) lui a me detto ieri «me ne ha parlato ieri»
- (c) l'articolo è omissivo:  
(3) aprire porta, fare luce «apri la porta e accendi la luce!»
- (d) l'infinito e il participio passato sono sovraestesi e la temporalità è espressa con avverbi:  
(4) cosa fare tu ora, adesso? «cosa stai facendo?»  
(5) piovuto tanto, vedi, acqua terra «deve aver piovuto molto, perché le strade sono ancora piene di pozzanghere»
- (e) *niente* e *no* sono generalizzati come particelle negative:  
(6) qui no fumare «in questo scompartimento non si può fumare»  
(7) niente visto uomo che tu dici «non ho visto l'uomo di cui stai parlando»
- (f) sono usati termini concreti invece che astratti e si ricorre a parole di lingue di maggiore diffusione, come francese e tedesco:  
(8) bisogna mettere il dito a gauche a sinistra links «bisogna premere il pulsante di sinistra e non quello di destra»
- (g) si usano onomatopee al posto di elementi lessicali:  
(9) drin drin drin ha visto te «mi ha detto al telefono che l'aveva vista»

# Computer talk?

- Idea frequente agli esordi dell'informatica (v. Licklider)
- In un certo senso, un linguaggio di programmazione è un computer talk, ma è molto diverso da una lingua naturale e non può essere imparato spontaneamente
- Questo vale anche per i linguaggi che sono stati sviluppati in modo da essere più intuitivi – dal Basic al Logo
- L'uso di parole della lingua inglese per dare istruzioni alla macchina era già considerato una forma di «parlare al computer»
- Per esempio: i nomi di istruzioni BASIC come «IF ... THEN ... ELSE»
- Oggi si può concepire qualcosa di più evoluto?

# Computer talk?

- Uno scenario: interrogazione di database... o di foglio Excel
- In generale, mi sembra che si sia creato un nuovo spazio per fare cose che con le interfacce grafiche attuali sono difficili ma che le capacità di elaborazione del linguaggio potrebbero rendere intuitive
- Le interfacce attuali sono per esempio difficili da usare quando richiedono di memorizzare comandi complessi ma usati solo occasionalmente
- Per esempio: dare un comando di formattazione nelle tabelle di Word, usare funzioni di calcolo in Excel

# Limiti delle interfacce

- Oggi c'è l'idea che un'interfaccia informatica debba essere intuitiva
- Va bene per i prodotti generici destinati a tutto il pubblico (3 miliardi di persone)... ma occorre che sia l'unica soluzione?
- Semplici istruzioni, facili da memorizzare, possono essere un primo esempio
- Alcune interfacce oggi consolidate sono in realtà molto difficili da usare
- Google Home ha, tra i difetti segnalati, l'assenza di «esplorabilità»

Aprire cartelle speciali di Windows	
Aprire la finestra Risorse del computer	"Apri Risorse del computer"
Aprire la cartella Documenti	"Apri Documenti"
Aprire la cartella Immagini	"Apri Immagini"
Aprire la cartella Musica	"Apri Musica"
Aprire la cartella Video	"Apri Video"
Avviare applicazioni e utilità	
Avviare un programma mediante nome	"Apri <nome programma>"

# Sviluppare un dialogo

- Vorrei avere la media dei voti di questo semestre
- *Per quali materie?*
- Per tutte le materie. Voglio calcolare il voto medio
- *Hai registrato 235 esami e il voto medio è stato 25,5*
- (Però) Mi serve anche la media ponderata... intendo ponderata per credito
- *Intendi dire che devo dividere il voto di ogni esame per i crediti dell'esame?*
- Sì
- *Hai registrato 235 esami e il voto medio ponderato è stato 24,9*