

Modulo A

Valutare i sistemi automatici di interazione linguistica

1. Storia delle interfacce vocali

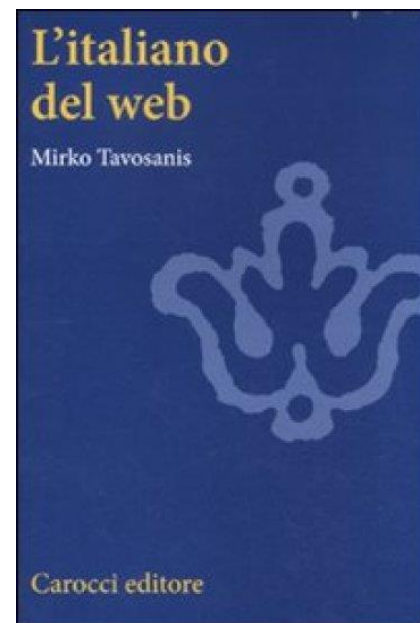
28 settembre 2016



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

Presentazione personale

- Laurea in Storia della lingua italiana (1991), dottorato di ricerca in Storia della lingua italiana (1997): *La prima stesura delle Prose della volgar lingua: fonti e correzioni*
- Dal 2014: professore associato di Linguistica italiana presso il Dipartimento di Filologia, letteratura e linguistica, Pisa
- Dal 2011 al 2016: direttore del Consorzio ICoN (www.italicon.it) per la diffusione della lingua e della cultura italiana via Internet
- Interessi di ricerca: Editoria e grammatica nel Cinquecento, Comunicazione professionale, Comunicazione mediata da computer, Lingua italiana all'estero...
- Uscito nel 2011: *L'italiano del web* (Roma, Carocci)
- Blog: linguaggiodelweb.blogspot.it...
- ... ma (aggiornato, spero, durante il corso): interfacevocali.blogspot.it



Presentazione del corso

- Una presentazione ancora un po' vaga
- Nel modulo A sicuramente **valuteremo** i «sistemi automatici di interazione linguistica»
- Nel modulo B probabilmente ci occuperemo della **generazione automatica di testi** (composizione automatica di notizie)
- Questa settimana e la prossima saranno introduttive
- Modalità di esame: stesura di relazioni scritte o preparazione di progetti informatici; discussione orale del programma
- Una relazione o un progetto per ogni modulo
- Per il modulo A, probabilmente, vi chiederò di sottoporre audio a un sistema come Google Translate e di analizzare i risultati

Intersezioni consigliate

- Linguistica computazionale (II, ma anche I)
- Elaborazione del linguaggio naturale (Magistrale in Informatica)
- In generale, tutto ciò che riguarda l'elaborazione automatica del linguaggio

Indicazioni pratiche

- Orario: 16-17:30
- Il mercoledì in aula F, il giovedì in aula O1
- Ricevimento: per ora il lunedì dalle 16 alle 18, ma probabilmente cambieremo
- Non ci sarà ricevimento lunedì 10 ottobre

- Iscrivetevi al corso (annuale) su Moodle:
<http://polo4.elearning.unipi.it/course/view.php?id=1539>
- Qual è il vostro corso di laurea?

Calendario

Lezioni:

- Settembre: 28 e 29
- Ottobre: 5, 6, 19, 20, 26, 27
- Novembre: 2, 3, 9, 10, 16, 17, 23, 24, 30
- Dicembre: 1

Attenzione! non ci sarà lezione **il 12 e il 13 ottobre**

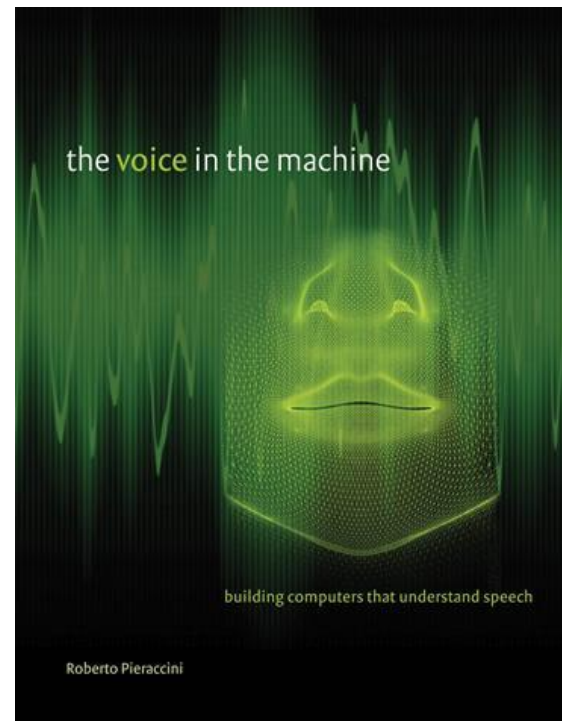
Probabilmente salteremo qualche lezione a novembre e recupereremo nei giorni disponibili a dicembre (7, 14, 15)

Oggi

- Partiamo dalla descrizione delle interfacce immaginate
- Alla nascita dei computer (anni Quaranta e Cinquanta) si pensava che lo strumento base di interazione sarebbe stato la voce
- In astratto era ragionevole pensarlo: la voce è uno strumento naturale di interazione, lo scritto serve a immagazzinare informazioni
- L'idea dell'interazione vocale era però irrealizzabile dal punto di vista tecnico
- Le possibilità si stanno concretizzando oggi, ma nel frattempo tutti si sono dimenticati di questa idea

Riferimento

- Roberto Pieraccini, *The Voice in the Machine: Building Computers That Understand Speech*, Cambridge e Londra, The MIT Press, 2012, pp. xxviii + 325, ISBN 978-0-262-01685-8.
- Disponibile (una copia) nella biblioteca LM1
- Usato come testo d'esame nel corso di Linguistica italiana II 2015-2016



Caratteristiche

- Si tratta di un testo divulgativo e di taglio storico
- Alcuni contenuti sono più tecnici e spesso richiederebbero approfondimenti; in particolare:
 - Alcuni concetti di fonetica («formanti»)
 - La descrizione generale del funzionamento del Dynamic Time Warp
 - Alcune note sui modelli di Markov e sui modelli di Markov nascosti
- Chi ha già studiato questi argomenti per Linguistica computazionale I o II o simili?

Argomenti

- La maggior parte del testo è dedicata ai sistemi di riconoscimento del parlato
- Alcune sezioni (tra cui il cap. 7, *An Interlude at the Other End of the Chain*) sono dedicate alla produzione artificiale di parlato
- Il lavoro riguarda soprattutto gli sviluppi degli ultimi sessant'anni, con alcune applicazioni pre-informatiche, dagli strumenti di von Kempelen al Voder
- La situazione è descritta da punti di vista che vanno dalle informazioni tecniche alla storia dello sviluppo e alla storia dello sviluppo delle aziende
- Un punto base: la comprensione del parlato è *difficile*, a tutti i livelli – e gli approcci che funzionano sono quelli basati sulla forza bruta, invece che sulla comprensione delle regole

Esempio: un dialogo telefonico (p. 220)

- *System*: Please say the name of your departure city.
- *User*: Boston.
- *System*: Okay, Austin. Please say the name of your arrival city.
- *User*: Er... no... it's not Austin, it's *Boston*.
- *System*: Okay. Going from Austin to Boston. At what time?
- *User*: No... I don't want to *go* to Boston, I am *leaving from* Boston!
- *System*: Okay. There is a flight leaving Austin at 7:15 p.m. Would that be okay?
- *User*: No, I am *not* leaving from Austin! I am going from *Boston...*
B-O-S-T-O-N... to San Francisco.
- *System*: Is the 7:15 p.m. flight okay? Please say yes or no.
- *User*: <hangs up>

Aspetti storici

- Prima di arrivare alla situazione attuale: è sempre stata sottovalutata la complessità del parlato...
- ... e non si è capito che gli esseri umani «comprendono» attivamente (il cervello sa più o meno che cosa sta per essere detto, ed è in grado di colmare le lacune dell'input)
- Finché non si è provato a farlo per davvero (inizi degli anni Sessanta), sembrava che capire il linguaggio umano fosse tutto sommato un'operazione meccanica e semplice
- Come nota anche Pieraccini, più difficile sembrava *produrre* il linguaggio: spesso i robot dei film agiscono, ma sono muti

Arcadia Darell esordì con voce ferma nel microfono del suo Transcriber:

“Sviluppi futuri del progetto Seldon, a cura di A. Darell.”

(...)

Strinse le labbra immaginandosi suo padre che alzava gli occhi dallo schermo di lettura...

(...)

Storse il collo e si guardò allo specchio con la coda dell'occhio, poi, parlando con voce profonda da donna matura, disse: - Veramente, papà, se credi che mi impressioni il giudizio di quegli stupidi giovani...

Poi si ricordò che aveva ancora il Transcriber aperto: - Per Giove! - esclamò e lo chiuse.

Sul foglio di carta leggermente violetto era scritto:

SVILUPPI FUTURI DEL PROGETTO SELDON

“Veramente, papà, se credi che mi impressioni il giudizio di quegli stupidi giovani...”

“Per Giove!”

Tolse il foglio della macchina e un altro scivolò al suo posto.

Guardò la macchina con soddisfazione. Era proprio elegante. Proprio l'ultimo grido in fatto di macchine scriventi (...). Il negoziante aveva aggiunto: - Non ne esiste uno migliore e così sensibile. Scrive la frase punteggiandola correttamente a seconda del senso. È di grande aiuto nello studio poiché induce chi lo usa a pronunciare in modo corretto le parole, dando la giusta tonalità alla frase, in modo che la trascrizione risulti perfetta.

Anche dopo le parole del negoziante, però, il padre aveva cercato di comperare un apparecchio manovrato a mano.

Per ottenerlo aveva dovuto versare qualche lacrimuccia anche se ciò non si addiceva proprio ad una donna di quattordici anni, ma ne era valsa la pena: la trascrizione era perfetta, femminile ed elegante, con le più aggraziate maiuscole mai viste (Isaac Asimov, *L'altra faccia della spirale*, pp. 72-74).

(oggi, l'inserimento della punteggiatura avviene solo attraverso comandi diretti in sistemi come Dragon o Siri)

J. C. R. Licklider, *Libraries of the Future* (1965)

- Commissionato dalla Fondazione Ford: uno dei rapporti tecnici che hanno condizionato più a fondo le politiche di finanziamento americane negli anni Sessanta.
- Licklider parte dalle idee di Vannevar Bush (“Memex”, 1945) ma propone per le biblioteche del futuro lo sviluppo di un “sistema procognitivo”: non solo archiviazione e recupero, ma anche rielaborazione di informazioni
- Alla base del progetto c’era la traduzione di articoli scientifici, ecc. in “unambiguous English” (p. 89):
 - From natural (technical) language to “unambiguous English” using machine-aided human translators
 - From “unambiguous English”, through “a purely machine translation”, into “the language(s) of the computer or of the data-base itself”
- Licklider stimava che per un articolo di 10 pagine la traduzione in “unambiguous English” potesse richiedere circa un’ora di lavoro... è uno dei punti più deboli di tutto il suo programma

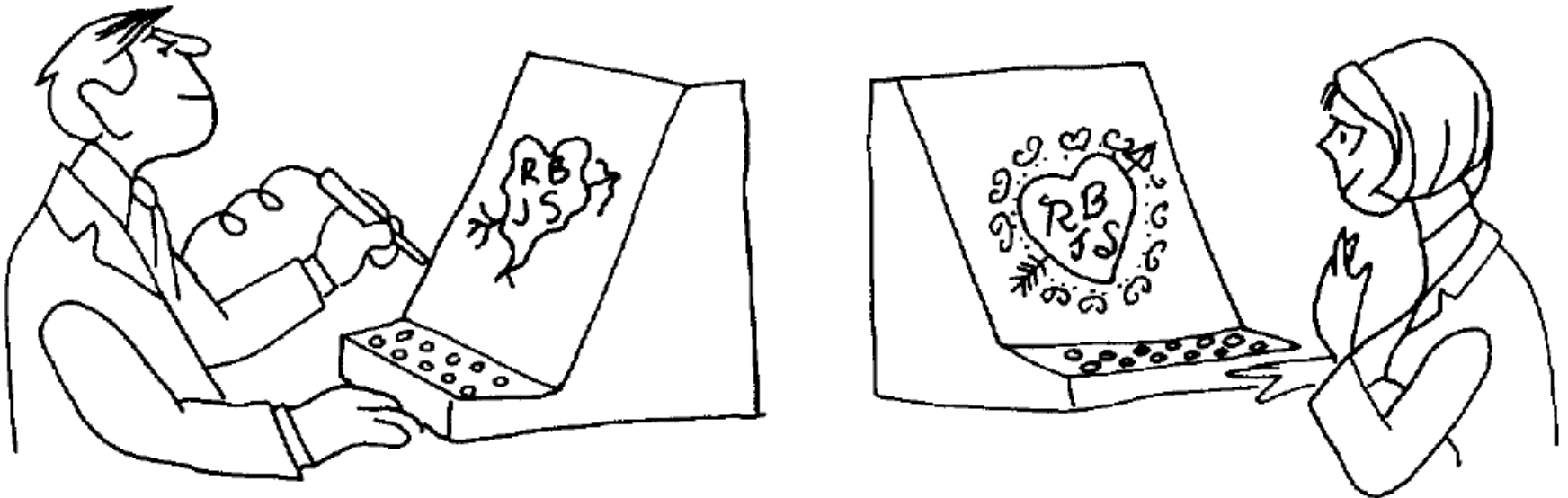
2001: odisea nello spazio (1968): il computer HAL 9000 comunica unicamente attraverso la voce - e sa leggere le labbra



The computer as a communication device

J. C. R. Licklider e Robert W. Taylor,
Science and Technology, aprile 1968: comprensione
dell'importanza e praticità dello schermo... anche se
basato più su penne che su tastiere

Vignette di Roland B. Wilson



La riunione del futuro



Douglas Engelbart

- Inizia a lavorare negli anni Cinquanta – anche in questo caso nella cornice concettuale di Bush
- Tuttavia Engelbart comprende presto che, nella pratica, l'unico modo per realizzare qualcosa di simile consiste nell'usare i computer... e la scrittura su schermo!
- Nel 1967 presenta la richiesta per il brevetto del mouse
- Nel 1968 realizza “la madre di tutte le demo”: mouse, posta elettronica, ipertesti...



monday afternoon

december 9

3:45 p.m. / arena

Chairman:

DR. D. C. ENGELBART

*Stanford Research Institute
Menlo Park, California*

a research center for augmenting human intellect

This session is entirely devoted to a presentation by Dr. Engelbart on a computer-based, interactive, multiconsole display system which is being developed at Stanford Research Institute under the sponsorship of ARPA, NASA and RADC. The system is being used as an experimental laboratory for investigating principles by which interactive computer aids can augment intellectual capability. The techniques which are being described will, themselves, be used to augment the presentation.

The session will use an on-line, closed circuit television hook-up to the SRI computing system in Menlo Park.

Following the presentation remote terminals to the system, in operation, may be viewed during the remainder of the conference in a special room set aside for that purpose.

Xerox Parc a Palo Alto

- Tra il 1970 e il 1973 un gruppo di lavoro coordinato da Taylor lavora su prototipi di computer con interfacce moderne - Wimp
- Nel 1973 vengono prodotti i primi esemplari di Xerox Alto, con commercializzazione limitata
- L'Alto presenta tutte le caratteristiche dei sistemi operativi moderni: videoscrittura, finestre multiple, mouse, metafora del desktop...
- Il suo successore, Xerox Star (1981), è la prima vera (e costosissima) commercializzazione del sistema
- Alcune scelte fatte all'epoca sono ancora con noi! (Metafora del desktop, cartelle...)



System Browser			
Collections-Sequence	Interval	accessing	collect:
Collections-Text	LinkedList	copying	do:
Collections-Array	MappedCollection	adding	doandBetweenDo:
Collections-Stream	OrderedCollection	removing	promoteFirstSuchT
Collections-Support	SortedCollection	shuffling	reverse
Graphics-Primitives		private	reverseDo:
Graphics-Display			select: Form Editor
Graphics-Media			
Graphics-Paths			
	instance	class	

collect: aBlock
"Evaluate aBlock with each of my elements as the argument. Collect the resulting values into a collection that is like me. Answer with the collection. Override superclass in order to use add:, not atPut:"

```

newCollection |
newCollection = self species new.
self do: [each | newCollection add: (aBlock value: each)].
+newCollection
  
```

User Interrupt

```

Paragraph>>characterBlockAtPoint:
Paragraph>>mouseSelectto:
CodeController(ParagraphEditor)>>processRedButton
CodeController(ParagraphEditor)>>processMouseButtons
CodeController(ParagraphEditor)>>controlActivity
CodeController(Controller)>>controlLoop
  
```

controlActivity

```

self scrollBarContainsCursor
ifTrue:
  [self scroll]
ifFalse:
  [self processKeyboard]
self processMouseButtons
  
```

```

[](<Robson>SF)*
[File](<Robson>SF)ScreenForm.st
[File](<Robson>SF)ScreenFormChanges.st
[File](<Robson>SF)WordGraphics.form
  
```

Rectangle fromUser origin
ScreenForm setFullPageWidth.

(Form readFrom: 'FiledSkate.form') edit

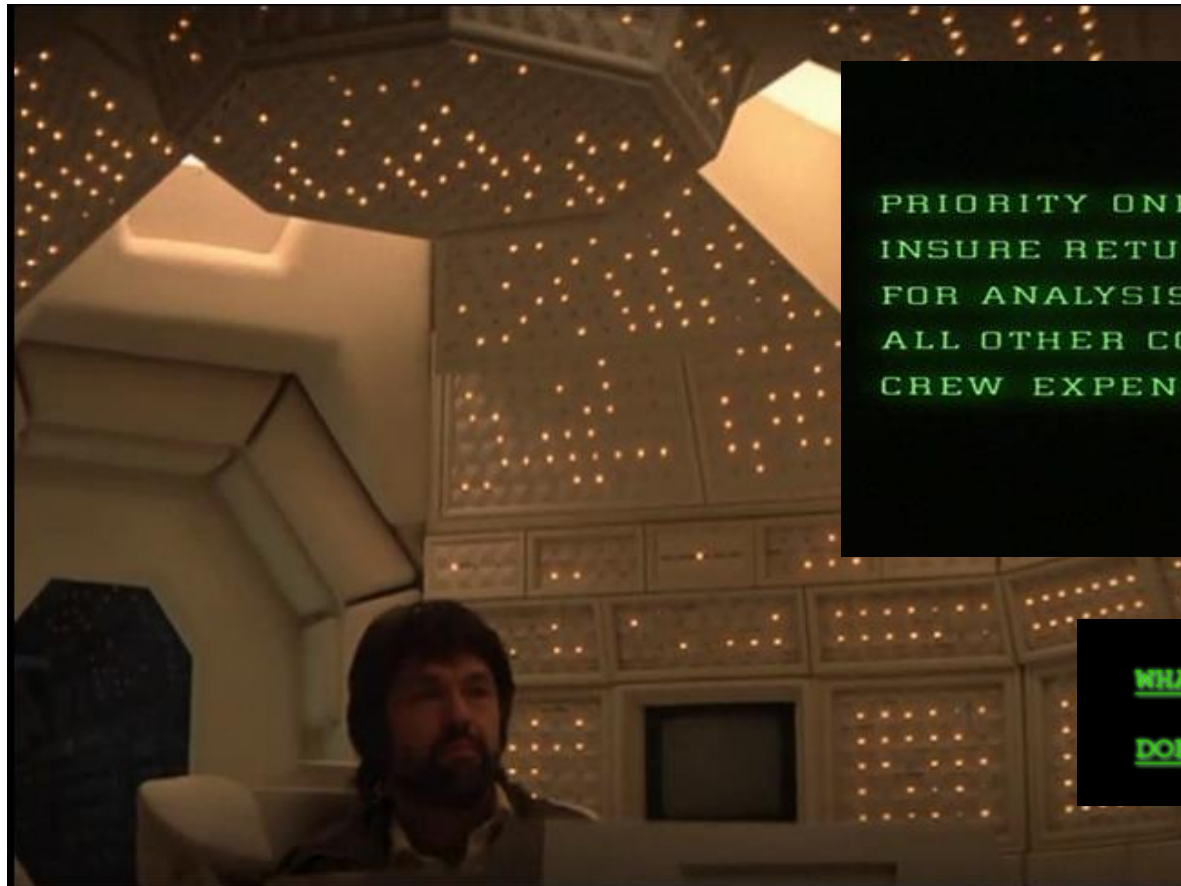
Fig. 1.

Diffusione

- Nel 1979 viene mostrata a Steve Jobs l'interfaccia Smalltalk (forse su un Alto)
- Dichiarazioni successive di Jobs: “I remember within ten minutes of seeing the graphical user interface stuff, just knowing that every computer would work this way some day. It was so obvious.”
- Altre storie sono più pittoresche...
- In seguito, sia Apple che Microsoft riprenderanno molti aspetti delle interfacce prodotte al Parc
- L'interazione vocale resta completamente a margine rispetto a questi sistemi – è molto più complessa da realizzare, ma a un certo momento semplicemente non se ne parla più (v. Pieraccini)

Alien, 1979

Schermo a fosfori verdi, tutto
maiuscole, frasi ellittiche...



PRIORITY ONE
INSURE RETURN OF ORGANISM
FOR ANALYSIS.
ALL OTHER CONSIDERATIONS SECONDARY.
CREW EXPENDABLE.

WHAT ARE MY CHANCES?

DOES NOT COMPUTE

Macintosh, 1984

