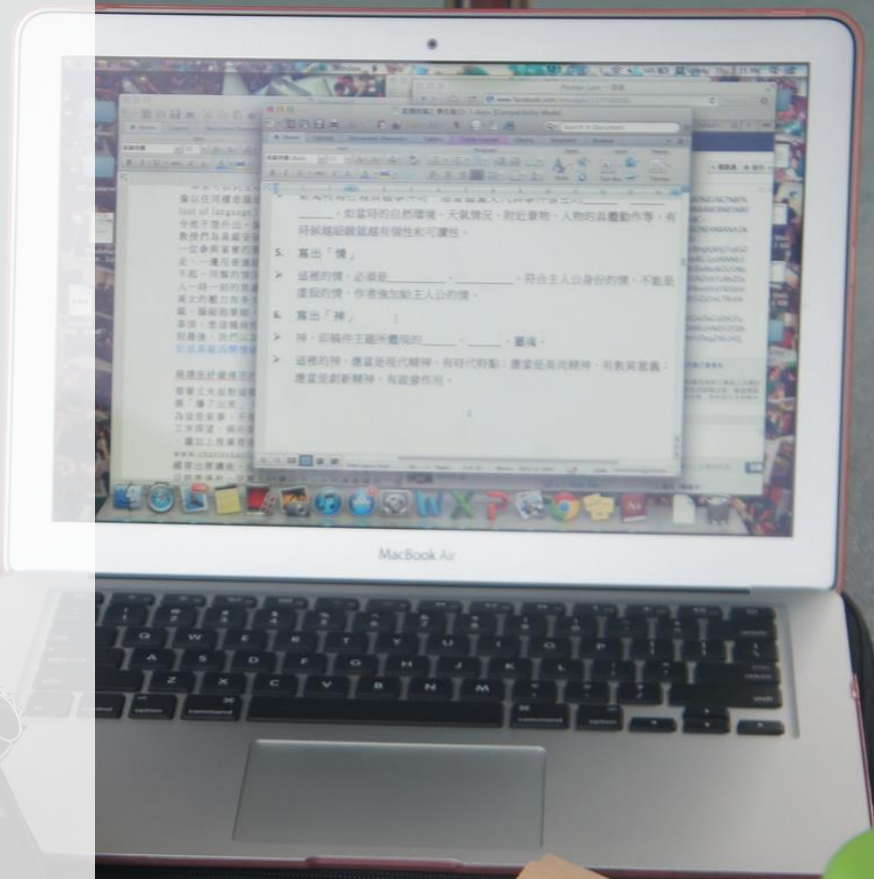


Modulo A

Valutare i sistemi automatici di interazione linguistica

5. Assistenti vocali

19 ottobre 2016



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

Oggi

Assistenti vocali

- Gli ideali
- La realtà:
 - Siri
 - Cortana
- Una valutazione delle prestazioni in ottica linguistica

Assistenti vocali

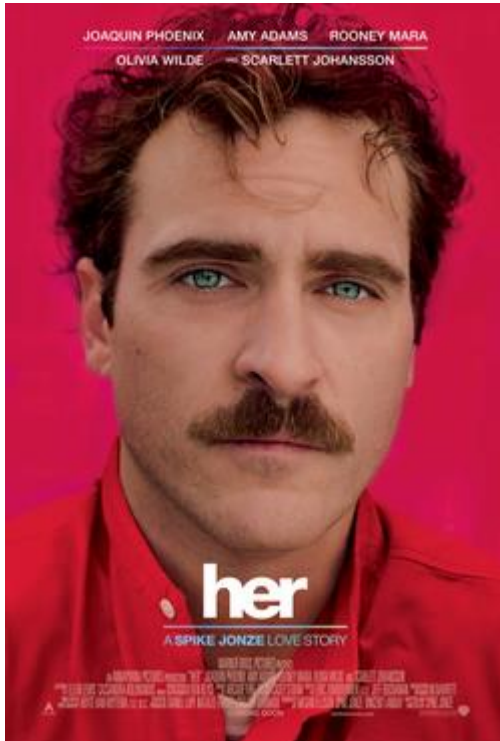
- In generale: strumenti capaci di interagire con l'utente attraverso la voce (a volte integrata con la scrittura)
- Gli esempi più importanti che simulano una personalità umana:
 - Siri (Apple)
 - Cortana (Microsoft)
 - Alexa (Amazon, ma solo per l'inglese)
- Non simula una personalità umana:
 - Google (Google), noto anche come «Google Now», forse in futuro come «Google Assistant» (per ora solo per l'inglese), ecc.
- Ambizioso ma ancora non disponibile (e solo per l'inglese?):
 - Viv, appena comprato da Samsung (<http://viv.ai/>)

Come dovrebbero funzionare?

- La conversazione è «l'interfaccia più naturale»
- Fin dagli inizi dell'informatica è stato un sogno: dimenticato dal pubblico con l'avvento delle interfacce grafiche, delle tastiere e del mouse, ma ogni tanto ricordato dai designer
- Il video sul «Knowledge navigator» per Apple mostra un'interfaccia vocale del 2011, immaginata nel 1987:
<https://youtu.be/hb4AzF6wEoc>



Her / Lei



- Un film di Spike Jonze (2013) interpretato da Joaquin Phoenix
- Racconta l'evoluzione di un'intelligenza artificiale – che interagisce solo via audio – nel prossimo futuro
- Le scene iniziali mostrano anche giornate di lavoro con interfacce vocali più realistiche:
 - Dettatura al computer
 - Comando vocale del computer
 - Ascolto e gestione di messaggi vocali in varie situazioni (per esempio, camminando)

Il mondo reale: funzionalità

- Per l'inglese, viene spesso dichiarato che la capacità di trascrizione del parlato è adeguata (lo verificheremo, ma sono abbastanza d'accordo anche per l'italiano): le lamentele riguardano gli aspetti di intelligenza artificiale, non quelli linguistici
- Un esempio di queste osservazioni: Walter Mossberg, *Why does Siri seem so dumb?*, "The Verge", 12 ottobre 2016
<http://www.theverge.com/2016/10/12/13251068/walt-mossberg-apple-siri-is-dumb>
- "In my recent experience, Siri has become quite good at transcribing what I'm asking, just not at answering it"
- Qui in parte si esce dall'ambito linguistico: molte delle critiche di Mossberg riguardano l'incapacità di Siri di usare informazioni correttamente registrate su iPhone
- Un altro esempio: Brian X. Chen, *Google Pixel Review: Assessing the New Smartphone*, "New York Times", 18 ottobre 2016

Esempi di interrogazione (17 ottobre)

“Chi è il Presidente del Consiglio?”

- Siri: presenta solo un rinvio alla voce di Wikipedia “Presidente del Consiglio dei ministri della Repubblica Italiana”
- Google: lo stesso.

“Quando c’è la Formula 1?”

- Siri: “Mi dispiace, Mirko, ma non riesco a trovare informazioni su questo campionato”, però mostra in accompagnamento link a siti web con il calendario corretto
- Google: rinvio al sito web con il calendario corretto

“Quando c’è la prossima partita del Pisa?”

- Siri: “La partita Torino-Pisa è il 29 novembre 2016 alle nove di sera” (risposta sbagliata)
- Google: “Pisa affronterà Latina sabato alle 15 e zero minuti” (risposta giusta)

“Chi è Walter Mossberg?”

- Siri: “Ecco cosa ho trovato cercando ‘Chi è Walter Moss Berg su Internet’” e ricerche web non pertinenti
- Google: profilo della persona corretta.

Follow-up (19 ottobre)

“Quando c’è la prossima partita del Pisa?”

- Siri: “La partita Torino-Pisa è il 29 novembre 2016 alle nove di sera” (risposta sbagliata)
- Google: “Pisa *giocherà contro* Latina sabato alle 15 e zero minuti” (risposta giusta)

E quella dopo?

- Siri: “Non so se ho capito bene ”
- Google: “Sono visualizzate immagini corrispondenti alla tua richiesta” (con ricerca su «dopo», «pillola del giorno dopo» e così via)

Possibile e impossibile

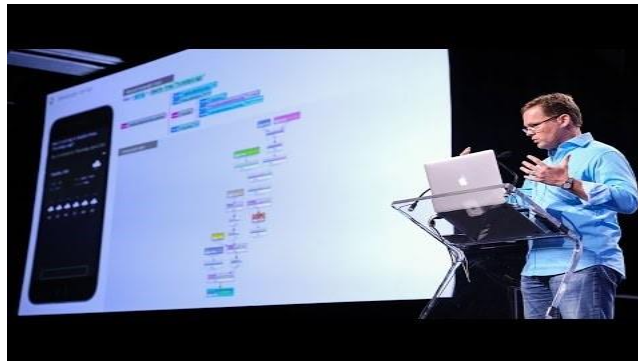
- Qualche attività è già possibile e sorprendentemente funzionale (*Dilbert*, 16 dicembre 2011)



- «Segna sul mio calendario tutte le partite del Pisa giocate a Pisa»: non sembra difficile, ma è ancora impossibile
- Vediamo il bicchiere mezzo pieno

Prototipi: Viv

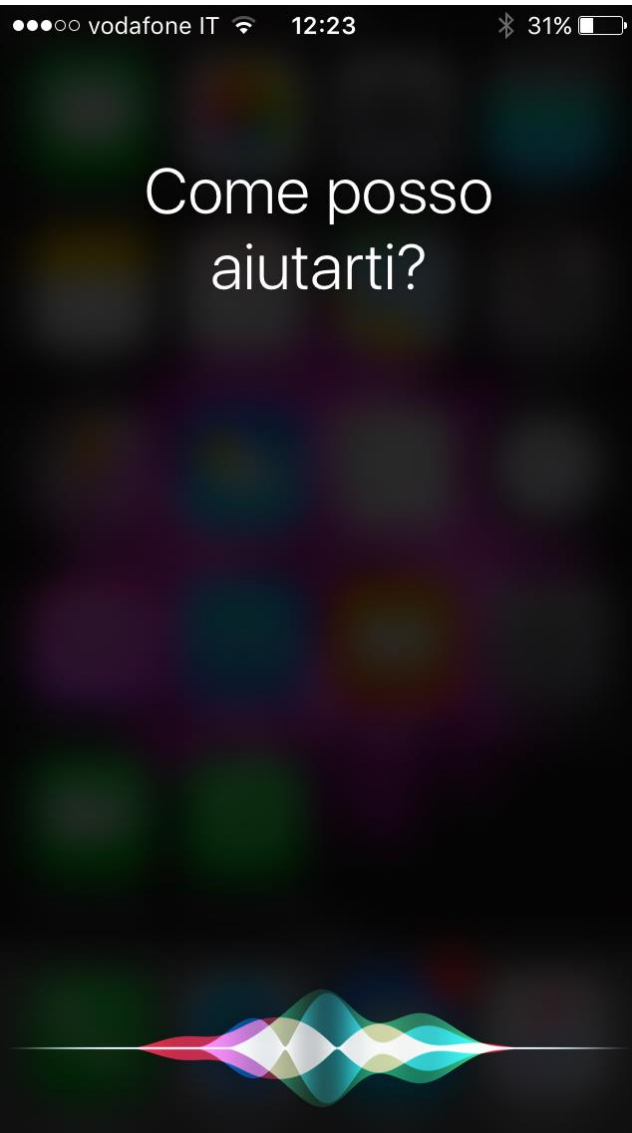
- Una dimostrazione di Viv fatta nel maggio 2016 (in particolare 4' 50" e 9' 10")
<https://youtu.be/Rblb3sptgpQ>
- Uno dei punti chiave: mantenere la memoria della conversazione per capire le «follow-up questions» e i riferimenti testuali
- Per le applicazioni pratiche, occorre ricordare che molte delle cose che facciamo sono fatte di parole
- L'obiettivo comunque è l'intelligenza artificiale, non una semplice interfaccia vocale



Personalità

- Gli assistenti digitali come Siri e Cortana sono stati dotati di una personalità
- Personalmente, credo che sia un errore (anche se è utile per creare titoli sui giornali, ecc.)
- Il discorso andrebbe approfondito, ma sembra evidente: gli esseri umani accettano i limiti delle interfacce ma sono infastiditi dalle cose che *cercano di sembrare umane* e non raggiungono capacità umane
- Vediamo le caratteristiche dei sistemi (integrati nei rispettivi sistemi operativi) e delle due personalità
- I sistemi sono complessi, e il riconoscimento del parlato ne rappresenta solo una parte
 - Siri: sistema di riconoscimento vocale sviluppato da Nuance
 - Cortana: elaborazione del parlato sviluppata da Tellme Networks

Siri (2011)



- Per un'informazione aggiornata consiglio le voci di Wikipedia in lingua italiana e in lingua inglese:
[https://it.wikipedia.org/wiki/Siri_\(software\)](https://it.wikipedia.org/wiki/Siri_(software))
<https://en.wikipedia.org/wiki/Siri>
- Le aree coperte vengono indicate (<http://www.apple.com/it/ios/siri/>) come:
 - Tenere i contatti (messaggi, mail, chiamate)
 - Organizzarsi (fissare promemoria, sveglie)
 - Sport (risultati sportivi)
 - Svago (musica, film)
 - In giro (indicazioni stradali, ricerca negozi)
 - HomeKit (domotica)
 - Trucchi e dritte (personalizzazione, scorciatoie)
 - Risposte su tutto (operazioni matematiche, conversazione)



Cortana (2014)

- Anche qui consiglio le voci di Wikipedia in lingua italiana e in lingua inglese:
[https://it.wikipedia.org/wiki/Cortana_\(software\)](https://it.wikipedia.org/wiki/Cortana_(software))
[https://en.wikipedia.org/wiki/Cortana_\(software\)](https://en.wikipedia.org/wiki/Cortana_(software))
- Le cose che Cortana può fare vengono indicate (<https://support.microsoft.com/it-it/help/17214/windows-10-what-is>) come:
 - Visualizzare promemoria in base a data/ora, luoghi o persone.
 - Monitorare pacchi, squadre, interessi e voli.
 - Inviare e-mail e SMS.
 - Gestire il tuo calendario e mantenerti aggiornato.
 - Creare e gestire elenchi.
 - Chiacchierare e giocare.
 - Trovare dati, file, luoghi e info.
 - Aprire qualsiasi app nel tuo sistema.

•

Confronto

- Francesca Bellante, Elisa Salatti ed Eleonora Zedda hanno confrontato le caratteristiche dei due assistenti come relazione di fine modulo per il corso di Linguistica italiana II del 2015-2016
- Titolo: *Siri vs. Cortana. Un'analisi pragmatica del linguaggio degli assistenti vocali*, in linea a:
<http://interfacevocali.blogspot.it/2016/10/le-personalita-di-siri-e-cortana.html>
- Come quadro teorico: Brown e Levinson, *Politeness: Some universals in language usage*, 1987
- Alla base, evitare strategie negative nell'interazione
- Nella verifica gli assistenti digitali sono stati trattati come persone, salutandoli, insultandoli, ringraziandoli... e valutando il modo in cui hanno risposto

Tipologie esaminate

- Saluti
- Cultura generale
- Insulti
- Riappacificazione e complimenti
- Aspetto intellettuale / professionale dell'assistente
- Sfera affettiva ed emotiva
- Domande sulla vita privata
- Opinioni personali
- Ringraziamenti

Risultati

- «Cortana è più aperta a sviluppare un rapporto cercando di inserire nel dialogo dei riferimenti comuni e cerca di creare questo common ground grazie all'uso di citazioni o imitazioni.»
- «Anche Siri cerca di creare un common ground, ma assume un comportamento più "freddo". In certi sensi, rispecchia di più il comportamento umano, dal momento che può provare imbarazzo e sentirsi a disagio.»
- «È interessante notare come i due software siano stati sviluppati con comportamenti "molto umani" ma molto differenti tra loro: Cortana dà più importanza alla propria faccia positiva, e si comporta da amica più che da assistente, mentre Siri sembra preferire la propria faccia negativa, e tende quindi a proteggere il proprio spazio personale e mantenere la relazione su toni più professionali.»
- Aggiungo: Siri varia molto le risposte

Funzionalità

- Se si guarda bene, le funzioni sono poche e rispondono a un numero piuttosto ridotto di comandi
- Forse anche perché gli sviluppatori non sono in grado di pensare a un numero sufficientemente ampio di combinazioni e non hanno (almeno inizialmente) corpora di riferimento
- Comunque, Siri è per esempio capace di rispondere a un gran numero di variazioni («Com'è il meteo?» «che tempo fa?») sulle aree molto ristrette coperte dalle sue competenze

Che ore sono?

- Che ore sono?
- Che ora è?

Siri e Cortana rispondono in modo corretto in entrambi i casi

- Quanto manca all'una?

Siri: Mirko, abbiamo un problema. Potresti provare di nuovo?

Cortana: rinvia a una ricerca sul web

- Metti una sveglia all'una

Siri e Cortana rispondono in modo corretto in entrambi i casi, ma Cortana mette la sveglia all'una di notte, Siri alle 13 (richiesta fatta alle 11:40)

- Metti una sveglia al tocco

Siri e Cortana chiedono chiarimenti; De Mauro considera «tocco» nel senso di «l'una del pomeriggio» come sostantivo «Comune»:

rintocco della campana | *il tocco*, l'una dopo mezzogiorno

Quanto ci vuole?

- Quanto costa un biglietto dell'autobus a Pisa?
- Quanto ci vuole per l'autobus a Pisa?

Ovviamente, i sistemi non riescono a fare nulla di più di una ricerca online

Una ricerca possibile: esaminare il modo in cui vengono gestite le più probabili variazioni linguistiche all'interno di uno o più sistemi

Gradazioni possibili: un campione

- non sono affatto a conoscenza di che cosa sia stato loro detto
- non sono affatto a conoscenza di che cosa abbiano loro detto
- non so affatto che cosa abbiano loro detto
- non so affatto che cosa abbiano loro detto
- non so affatto che cosa hanno loro detto
- non so mica che cosa gli hanno detto
- non so mica che cosa gli han detto
- non so mica cosa gli han detto
- so mica cosa gli han detto
- [eccetera]
- so mica cosa ci han detto...

(Gaetano Berruto, *Sociolinguistica dell'italiano contemporaneo*, 1987, p. 30)