

Modulo A

Valutare i sistemi automatici di interazione linguistica

2. Sintesi del parlato e Dragon

29 settembre 2016



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

Oggi

- Lo sviluppo dei sistemi commerciali per il trattamento dell'italiano
- Sintesi del parlato
- Comprensione del parlato e sistemi che elaborano in locale: Dragon

Strumenti

- Nel 2012, l'Istituto di Linguistica Computazionale di Pisa (ILC-CNR) ha attirato l'attenzione sulla necessità che l'italiano sia incluso tra le lingue per cui sono sviluppate le “tecnologie linguistiche” elettroniche
- L'italiano comunque è ben studiato, e vengono gestiti e aggiornati tutti gli strumenti necessari a descriverlo scientificamente (grammatiche, vocabolari, descrizioni fonologiche...)
- Di solito è la quarta o quinta lingua per lo sviluppo di interfacce, tecnologie linguistiche, eccetera; nel maggio del 2015 italiano e cinese mandarino sono stati terza e quarta lingua per Microsoft Skype Translator.
- Alcuni prodotti commerciali degli ultimi anni includono:
 - Siri in italiano (Apple)
 - Cortana in italiano (Microsoft)
 - Google Voice, Google Translate, etc. (Google)

Categorie

Alcune alternative

- Sintesi vocale («Text to speech», TTS) oppure comprensione del parlato («Speech to text», STT)
- Elaborazione in locale (possibile solo su computer) o elaborazione in remoto

I sistemi di comprensione del parlato hanno due categorie ulteriori:

- Con fase di addestramento oppure senza
- Con funzioni di dialogo oppure senza (includendo anche il riconoscimento di comandi)

Assistenti vocali

- Un settore importante per l'uso della voce è l'accessibilità
- Spesso gli assistenti vocali sono considerati strumenti «di accessibilità», non di produttività
- Windows 10 colloca i principali controlli per gli strumenti vocali nel menu «Accessibilità»

Programmi per sintesi vocale

- La lettura del testo viene fatta attraverso «voci», che possono essere condivise tra più programmi (ci sono standard come SAPI)
- Di solito le voci sono femminili; esistono anche voci maschili

Sintetizzatori

- Online: per esempio Ivona, con due voci in italiano, Carla e Giorgio, che possono essere scaricate
<https://www.ivona.com/>
- Locali: il sistema base di Windows 10 o quello di Dragon Dictate

Sintetizzatori: Microsoft Windows 10

- L'**Assistente vocale** (Microsoft Narrator) è uno strumento Microsoft che è in grado, tra l'altro, di **leggere a voce alta** elementi dell'interfaccia o testi articolati: «text to speech» o TTS
- L'Assistente vocale fa parte di Windows; è disponibile in inglese (Stati Uniti, Regno Unito e India), francese, italiano, tedesco, giapponese, coreano, mandarino (cinese semplificato e cinese tradizionale), cantonese (cinese tradizionale), spagnolo (Spagna e Messico), polacco, russo e portoghese (Brasile).
- Voci per l'italiano: Microsoft Elsa e (molto scarsa) Microsoft Cosimo
- Voci base per l'inglese: Microsoft Zira e Microsoft David
- Vediamo alcuni esempi di lettura ad alta voce di un testo in italiano standard: la pronuncia è di alta qualità, nei modi che trovate descritti nel libro di Pieraccini

SAPI

- Microsoft usa per le sue voci la Speech Application Programming Interface (SAPI, version 4 e 5) che consente di usare più voci
- Le voci SAPI 5 sono le più sofisticate, ma sono di solito distribuite a pagamento
- Alcuni esempi italiani sono disponibili sul sito di Loquendo, oggi assorbita da Nuance:
<http://italy.nuance.com/azienda/soluzione/soluzioni-assistenza-clienti/servizi-soluzioni/inbound/loquendo-small-business-bundle/tts-demo/italian/index.htm>
- Nuance Dragon usa come voce Silvia

Lettura ad alta voce

- Oggi è perfettamente possibile usare questi sistemi per una lettura professionale
- È anche possibile creare audiolibri o simili, senza troppi problemi, usando programmi come Balabolka, che si appoggiano al sistema e alle voci di Windows
- Esistono sicuramente molti utenti, ma l'impressione è che usino questi strumenti per interesse nei confronti delle tecnologie assistive (e che si tratti comunque di una nicchia)

Riconoscimento

- Speech to text: più complesso
- Il sistema di riconoscimento del parlato Microsoft **non** copre l'italiano
 - la funzionalità Riconoscimento vocale è disponibile solo per sei lingue: inglese, francese, spagnolo, tedesco, giapponese, cinese (con trascrizione in cinese semplificato o tradizionale)
- Si tratta di uno dei rari casi in cui l'italiano rimane fuori dal gruppo delle lingue più supportate
- Per le lingue supportate, il sistema Microsoft non ha fase di riconoscimento dedicata
- Comunque Windows ha una serie di funzioni simili, spesso poco note: per esempio il riconoscimento della scrittura a mano, in stampatello e in corsivo

Dragon NaturallySpeaking

- La storia aziendale è in parte raccontata in *The voice in the machine*: il programma è sviluppato e commercializzato da Nuance, nata dalla fusione Nuance e ScanSoft nel 2005
- ScanSoft aveva in precedenza acquisito il prodotto sviluppato da Dragon (fondata nel 1982): il primo sistema a dettatura continua, senza pause tra le parole e basato su catene markoviane nascoste, venne commercializzato da Dragon nel 1997
- Supporta l'italiano
- Elabora in locale, con una (breve) fase di riconoscimento e senza funzioni di dialogo, a parte il riconoscimento di comandi
- Al momento, è forse il prodotto più potente in assoluto per quanto riguarda il riconoscimento vocale in italiano – e anche in altre lingue; ma non sono sicuro che abbia vantaggi rispetto a Google
- Rigorosamente a pagamento, anche se a costi accessibili



Parla «come un libro stampato»?

Proviamo la trascrizione

- di un testo «scritto scritto»
- di un brano di conversazione

Le *lingue speciali*, invece, sono le lingue collegate ad alcune discipline e professioni. Per esempio, un chimico o un medico, per svolgere il proprio lavoro, devono conoscere molti concetti che non sono di dominio comune e che spesso vengono indicati con parole create appositamente. Termini come *granodiorite* o *trombosi* sono esempi tipici di queste categorie di parole (più che per differenze sintattiche o fonetiche, che pure sono presenti in misura ridotta, le lingue speciali si distinguono quindi dall'italiano standard soprattutto a livello lessicale).