

DOVE IL SÌ SUONA

GLI ITALIANI
E LA LORO LINGUA



www.doveilsisuona.it

Esposizione di sistemi sul

TAL

Trattamento Automatico della Lingua

a cura della  **Fondazione Ugo Bordononi**

Firenze, Sala delle Regie Poste

14 marzo / 30 settembre 2003



Le teste parlanti

Introduzione

La lingua, scritta o parlata, è la forma più naturale con cui l'uomo gestisce le informazioni che gli consentono di vivere nel mondo e di evolversi. Spesso ci piace dire che la lingua è un qualcosa di vivo e conseguentemente essa è destinata a morire quando non svolge più la sua funzione comunicativa. Con l'invenzione della stampa le lingue che non hanno beneficiato di questa nuova tecnologia sono retrocesse allo stato di dialetto. Solamente le lingue che hanno prodotto un numero significativo di testi scritti sono diventate lingue di comunicazione e di pensiero.

Lo sviluppo di adeguati metodi di TAL è la chiave per la promozione della lingua nella società dell'informazione: la lingua è uno strumento di comunicazione ma anche di progresso della comunità; attraverso la lingua stessa si formano le nuove generazioni e si diffonde la cultura.

Attualmente assistiamo ad una nuova rivoluzione dovuta all'impatto della tecnologia informatica. L'odierna trasformazione è anche più sorprendente di quella provocata dalla stampa poiché oggi il linguaggio non viene solamente replicato e conservato, ma viene trasformato, manipolato e adattato ad esigenze diverse mediante il trattamento automatico dei suoni e dei simboli.

Pensiamo alle tecniche informatizzate per il reperimento di documenti in base al loro contenuto e i centralini automatici in grado, 24 ore su 24, di fornire risposte utili a domande complesse formulate in linguaggio naturale. Inoltre grazie a questa tecnologia sono stati sviluppati nuovi strumenti di ausilio alle persone affette da alcune disabilità motorie e sensorie.

Una frontiera non troppo lontana e di grande interesse applicativo, soprattutto per i cittadini d'Europa, è la comunicazione tra interlocutori di lingue diverse grazie ad una macchina di traduzione simultanea.

Nella sala delle Poste Regie vi sono alcuni esempi di sistemi di Trattamento Automatico del Linguaggio (TAL): macchine parlanti che narrano la loro storia, macchine traduttrici, sistemi di riconoscimento del parlato e del parlante.

Il TAL dà vita a un'industria della lingua che valorizza i suoi prodotti: risorse linguistiche di base; sistemi di riconoscimento e sintesi, cioè strumenti di conversione da parlato a testo e viceversa; sistemi di analisi sintattica e semantica; sistemi di gestione del dialogo; sistemi di traduzione.

I beneficiari di questa industria potrebbero essere sia le amministrazioni pubbliche e private sia il grande pubblico: il trattamento automatico o comunque meccanizzato di documenti nelle amministrazioni, nell'industria e nel commercio è uno dei principali impieghi di manodopera in Europa.

Diverse azioni sono già state intraprese in Italia sul TAL: il progetto "TAL" - Infrastruttura nazionale per le risorse linguistiche nel settore del Trattamento Automatico della lingua naturale parlata e scritta (legge 46/82 art. 10) e il progetto LCRMM "Linguistica computazionale - ricerche monolingui e multilingui" (legge 488/90 cluster "Linguistica").

Recentemente è stato costituito, sotto gli auspici del Ministero delle Comunicazioni, un forum nazionale sul TAL, il cui obiettivo principale è quello di monitorare e stimolare ulteriormente le attività nel campo.

I sistemi esposti

LE MACCHINE PARLANTI

Fin dall'antichità l'uomo ha provato a costruire statue e macchine parlanti. I sacerdoti Caldei facevano parlare una testa di cera, il Teraphim, con un simpatico espediente: un uomo nascosto parlava in una trachea di gru collegata con la testa e la voce usciva magicamente dalla bocca dell'oracolo. Nel X secolo il Papa Silvestro II costruì una testa meravigliosa che diceva

sempre il vero, simile a quella conservata ancor oggi nella Cattedrale di Barcellona, che funziona grazie al vapore dell'acqua bollente.

Nel Settecento furono realizzati numerosi congegni in grado di pronunciare qualche suono o delle brevi frasi.

Nel Novecento, con l'invenzione del giradischi e del registratore magnetico, si sono resi disponibili meravigliosi sistemi in grado di "suonare". Questi dispositivi riproducono un suono precedentemente inciso e si discostano in questo dalle antiche teste parlanti.

Solo nel dopoguerra si sono finalmente realizzati, con l'uso dell'elettronica, apparati in grado di generare dei suoni a partire da un testo scritto.

Oggi, i risultati della ricerca di Loquendo dimostrano che è possibile sintetizzare una voce comprensibile, armoniosa e naturale, con timbro e intonazione molto simili a quelli di una voce umana.

Seguendo le dimostrazioni in questa sezione, potete apprezzare voci sintetiche sorprendentemente belle e naturali, esplorare le tappe storiche della sintesi vocale e riflettere sulle possibilità d'impiego di questa tecnologia.

IL RICONOSCIMENTO VOCALE

Le tecnologie vocali possono essere facilmente immaginate come la bocca e l'orecchio di un calcolatore elettronico, ovvero come componenti tecnologiche in grado di dotare un elaboratore di capacità di ascoltare e capire (tecnologia di riconoscimento automatico e di comprensione della voce) e di parlare (tecnologia di sintesi della voce). Il riconoscimento vocale in particolare consiste nell'identificare, all'interno di un segnale normale, costituito da voce e rumore, una o più parole appartenenti ad un vocabolario prefissato, in modo da comprendere correttamente un messaggio.

Le tecnologie vocali sono componenti alla base dei centralini automatici in quanto permettono di interpretare le richieste dell'utente abilitando la possibilità di fornire istruzioni, accedere alle informazioni ed ai servizi, in modo completamente automatico.

Conversare con una macchina sembra essere un obiettivo possibile. In questa sezione esploriamo la capacità di un calcolatore di capire un essere umano che parla.

Seguendo le dimostrazioni in questa sezione, potete apprezzare l'evoluzione del dialogo con le macchine e riflettere sulle possibilità d'impiego che il riconoscimento del parlato è in grado di offrire.

IL RICONOSCIMENTO DEL PARLANTE

Una postazione all'interno della mostra è relativa al tema del riconoscimento, o meglio, della caratterizzazione del parlante.

Tutti noi sappiamo che è piuttosto facile riconoscere la voce di un personaggio famoso, o di una persona a noi cara. In altre parole ognuno di noi è facilmente in grado di riconoscere voci "familiari". Sappiamo inoltre che dal solo ascolto di una voce possiamo "immaginare" delle caratteristiche della persona che ci sta parlando. Su alcune caratteristiche possiamo essere abbastanza sicuri (genere, età, zona di provenienza) su altre invece siamo generalmente meno sicuri (aspetto fisico, livello di cultura, ecc.).

Anche gli elaboratori sono in grado di operare queste analisi ma, come noi, anche le macchine hanno necessità di familiarizzare con una determinata voce per poterla riconoscere e questo richiede molto tempo. Non essendo perciò possibile addestrare il sistema con la vostra voce (è un processo lungo e delicato) abbiamo pensato di far conoscere al sistema le voci di una serie di attori/attrici famose del cinema italiano o di personaggi dei cartoni.

Per il divertimento di grandi e piccini, il sistema in dimostrazione confronterà la vostra voce con quella di una lista di personaggi che voi avrete prescelto mostrandovi quale attore, attrice o personaggio dei cartoni ha la voce più simile alla vostra; se tuttavia avrete registrato preventivamente la vostra voce, il sistema vi riconoscerà!

I CORPORA

Con il termine "corpora" si designano quelle collezioni di segnale vocale che hanno assunto sempre maggior importanza nello studio dell'italiano parlato e nella realizzazione dei prodotti basati sulla tecnologia vocale.

Man mano che la tecnologia ha offerto prodotti sempre più sofisticati quali i sistemi per la dettatura automatica o quelli per l'accesso vocale ad informazioni anche la raccolta di materiale vocale è diventata via via più raffinata orientandosi verso la raccolta di costrutti linguistici e vocali più complessi e delle varietà regionali.

Si è anche posto il problema di una migliore integrazione tra le comunità accademiche di orientamento linguistico e quelle di ispirazione ingegneristica che ha condotto a iniziative congiunte volte alla costruzione di risorse linguistiche riutilizzabili. Con questo termine si designano sia gli insiemi di dati linguistici sia gli strumenti software usati per costruire, ampliare, rendere operativi e valutare modelli, algoritmi, componenti dei sistemi per il Trattamento Automatico della Lingua.

Nella macchina in dimostrazione vengono descritti vari corpora tra i quali ricordiamo SIVA e CLIPS.

La base di dati italiani SIVA (Speaker Identification and Verification Archives) contiene i segnali di più di 2.000 parlanti, registrati sulla rete telefonica, che coprono le principali aree dialettali.

Il corpus CLIPS (Corpora Linguistici per l'Italiano Parlato e Scritto), rappresenta una delle maggiori iniziative per la costruzione di una base di dati delle varietà dell'italiano. Si tratta infatti di un corpus di circa 100 ore di parlato, articolato in parlato dialogico spontaneo, radiofonico e televisivo (notiziari, cultura e intrattenimento, pubblicità), telefonico e letto. Il materiale è stato raccolto in 15 località diverse, scelte secondo criteri linguistici, demografici e socioeconomici in modo da essere rappresentativo delle maggiori varietà dialettali.

INSEGNAMENTO A DISTANZA

L'apprendimento di una lingua mediante un sistema di e-learning viene favorito dalla possibilità di seguire percorsi didattici personalizzati definiti sulla base di test automatici di autovalutazione ed è supportato da una notevolissima quantità di materiale didattico multimediale disponibile in rete o su supporto ottico. Momenti di interazione tra studenti singoli e nell'ambito di gruppi di lavoro, consentiti dalle tecnologie di comunicazione sincrona, introducono la possibilità di praticare la conversazione e la comunicazione su temi di interesse reale e sotto il controllo di un tutor.

Il tutoraggio prevede momenti espliciti di valutazione ed eventualmente di certificazione. Il sistema in dimostrazione contiene un programma interattivo che fornisce alcuni esempi del percorso formativo e dei test automatici previsti dal corso.

LA TRADUZIONE AUTOMATICA

La traduzione automatica, cioè la possibilità di affidare ad un elaboratore la conversione da una lingua ad un'altra, riducendo al minimo la revisione da parte dell'utente, è stata per molti anni un'utopia e una sfida per le attività di ricerca nelle discipline coinvolte (informatica, linguistica computazionale, intelligenza artificiale, scienze cognitive).

Il Traduttore utilizza i risultati più avanzati di tali ricerche e dimostra che oggi la traduzione automatica è diventata una realtà, disponibile per chiunque ha necessità di tradurre comunicazioni e documenti tecnico/commerciali su un comune personal computer. La postazione, a cura della ditta Synthema, dedicata al tema della traduzione automatica contiene tutti i componenti linguistici necessari: una grammatica automatica e un analizzatore per l'inglese e l'italiano, un potente sistema di tra-

sformazione sia strutturale sia lessicale, una grammatica generativa per l'italiano e l'inglese che consente di "generare" le frasi tradotte, risolvendo tutti gli aspetti linguistici necessari (flessioni, concordanze, coniugazioni, pronomi clitici, frasi idiomatiche, ecc.).

Il sistema in dimostrazione presenta una breve storia della traduzione automatica e consente ai visitatori di tradurre testi scritti dall'italiano all'inglese (e viceversa). All'interno della mostra si possono apprezzare i risultati raggiunti da una tecnologia in costante progresso e divertirsi con traduzioni bizzarre.

CENTRALINO VOCALE

Le applicazioni vocali automatiche si rivolgono a un bacino d'utenza molto vasto, adottano tecnologie di alta qualità ed effettuano un'attività di riconoscimento indipendente dal parlatore. Non richiedono apparecchiature specializzate: il telefono diventa il terminale di accesso privilegiato ai servizi informatici. Le opzioni disponibili sono presentate per mezzo di un menù: basterà effettuare la scelta, pronunciando semplicemente il nome del servizio desiderato. Durante la lettura dell'elenco sarà sempre possibile interrompere il sistema con una nuova richiesta.

Il centralino automatico, che fa uso di un sistema di sintesi da testo e di un riconoscitore della voce è in grado di fornire informazioni sulla mostra, dimostrando contemporaneamente il livello tecnologico raggiunto dai sistemi di colloquio uomo-macchina.

Una postazione telefonica permette di interagire con un servizio automatico che dà informazioni sulla mostra. Il visitatore potrà così sperimentare direttamente come l'interazione vocale sia semplice ed intuitiva.

In particolare si hanno informazioni su:

- Tariffe e tipologia di biglietti
- Esistenza di servizi e modalità di fruizione
- Descrizioni sui temi della mostra
- Percorsi di avvicinamento alla mostra

Il servizio proposto si basa su tecnologie di "vocal browsing" e non prevede l'intervento di un operatore umano. In caso di necessità, è comunque possibile prevedere il trasferimento di chiamata verso un operatore di call center (se esistente).

Il numero verde da chiamare è: 800672233



La Fondazione Ugo Bordoni svolge attività di ricerca e studio in tutti gli attuali contesti della telecomunicazione e dell'informazione e coadiuva il Ministero delle Comunicazioni nella elaborazione strategica e nella soluzione delle problematiche tecnologiche, economiche, normative. Particolare attenzione è rivolta alle comunicazioni multimediali: codifica e valutazione dei segnali audio e video, studio delle interfacce vocali interattive, ovvero sintesi della voce e comprensione del parlato, traduzione voce-voce e riconoscimento e verifica del parlante. Nell'ambito di questa manifestazione, la Fondazione presentando un'esposizione incentrata sulle tematiche del Trattamento Automatico della Lingua Italiana, dà la possibilità di utilizzare dispositivi interattivi, per meglio comprendere i traguardi della tecnologia odierna. Le ricerche della Fondazione vengono spesso svolte in collaborazione con altre istituzioni italiane ed europee, sostenute dai principali operatori nel settore TLC e dal Ministero delle Comunicazioni. www.fub.it



L'Istituto Superiore delle Comunicazioni e delle Tecnologie dell'Informazione (ISCTI), organo tecnico-scientifico del Ministero delle Comunicazioni, ha ruolo di primaria importanza nel campo dello sviluppo delle tecnologie dell'informazione e delle telecomunicazioni. Nei suoi laboratori nascono soluzioni che contribuiscono allo sviluppo delle comunicazioni nel nostro Paese. In collaborazione con la Fondazione Ugo Bordoni, rivolge particolare attenzione alla tecnologia della comunicazione vocale. È stato il primo istituto, in Italia, ad aver raccolto una banca dati vocale pubblica (A.I.D.A). L'ISCTI organizza corsi di formazione in telecomunicazioni e cura formazione e aggiornamento, in campo informatico, per il personale del Ministero e di altre organizzazioni. Tra le altre attività da segnalare: gestione della banca dati per la numerazione nazionale, orologio nazionale di riferimento per la sincronizzazione delle reti di telecomunicazione, realizzazione del Centro per la valutazione della sicurezza informatica, convenzione con l'Istituto Italiano del Marchio di Qualità per la Certificazione dei Sistemi Qualità Aziendale. www.comunicazioni.it

Loquendo

Loquendo, leader nelle soluzioni vocali e forte delle competenze maturate in 30 anni di ricerca, propone una gamma completa di soluzioni, piattaforme e componenti, per la realizzazione di servizi automatici ad interazione vocale in applicazioni di Call Center, CRM, Intranet e Unified Messaging. Loquendo ha realizzato applicazioni a livello nazionale ed internazionale che gestiscono automaticamente più di un milione di telefonate al giorno grazie a tecnologie (sintesi, riconoscimento, speaker verification, speech suite), piattaforme e applicazioni disponibili nelle principali lingue. Loquendo - società di Telecom Italia - conta 180 dipendenti, ha sede a Torino, uffici e attività in Europa, Nord e Sud America e si avvale di importanti partner commerciali. www.loquendo.com



Synthema nasce nel 1994 dall'iniziativa di un gruppo di ricercatori del Centro di Ricerca IBM di Pisa. L'attitudine all'innovazione tecnologica, insieme con la consistente esperienza maturata in temi di ricerca applicata, hanno costituito un'importante base di partenza ed hanno permesso all'azienda di crescere e consolidarsi nel mercato dell'Information Technology come fornitori di prodotti e soluzioni software avanzate, mantenendo al tempo stesso una consistente attività di Ricerca Industriale e Sviluppo Tecnologico. www.synthema.it