



L'INDUSTRIA ITALIANA DELLA VOCE

ROBERTO BILLI,
Loquendo S.p.A.

Il settore delle tecnologie vocali è caratterizzato da una forte trasversalità e tocca uno degli aspetti più importanti e strategici per il successo di qualunque nuovo prodotto o servizio, ovvero l'interfaccia utente. Le tecnologie vocali giocano perciò un ruolo molto importante nella evoluzione di diversi settori, da quello dei terminali mobili alle telecomunicazioni, al settore automotive, alla casa intelligente, alle applicazioni desktop, a quelle per disabili, all'apprendimento delle lingue, al settore nel mondo Internet e dei contenuti digitali (fig. 1).

Il mercato delle tecnologie ed applicazioni vocali è strutturato in diversi segmenti in ciascuno dei quali operano attori più o meno specializzati (vedi fig. 2). Alla base della catena del valore sta l'attività di R&D che è fondamentale per assicurare un progresso su queste tecnologie che hanno ancora molto spazio di miglioramento. Su questo segmento gli attori principali sono le Università, i Centri di Ricerca nazionali, alcune società specializzate sulle tecnologie vocali.

Il primo segmento di mercato è quello delle tecnologie di base (speech recognition, text-to-speech, speaker verification, Spoken Dialogue) dominato da pochi fornitori specializzati (tra cui Loquendo) e da 2-3 grossi player nella Information Technology che possiedono proprie tecnologie vocali. In alcuni paesi sono presenti nel ruolo di fornitori anche cen-

tri di ricerca pubblici/privati che hanno però principalmente lo scopo di assicurare la disponibilità di tecnologie per la propria lingua.

Il segmento successivo è quello delle piattaforme hardware/software che supportano lo sviluppo delle applicazioni, in particolare quelle telefoniche che presentano requisiti molto stringenti di scalabilità ed affidabilità. Alle piattaforme basate su hw e sw proprietario si sono recentemente affiancate piattaforme che utilizzano hw commerciale (server, schede telefoniche) e che si basano su standard software come VoiceXML, SALT, SAPI.

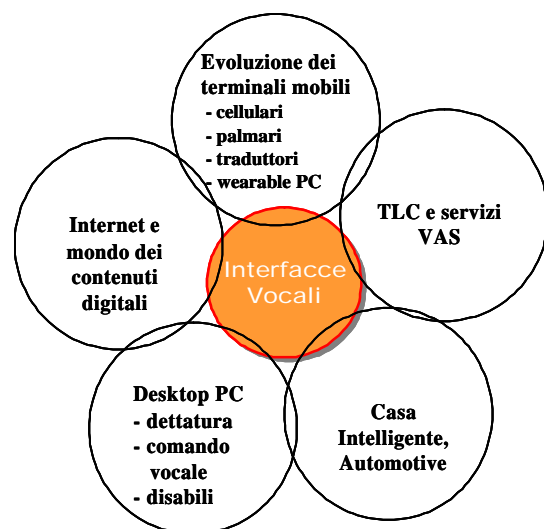
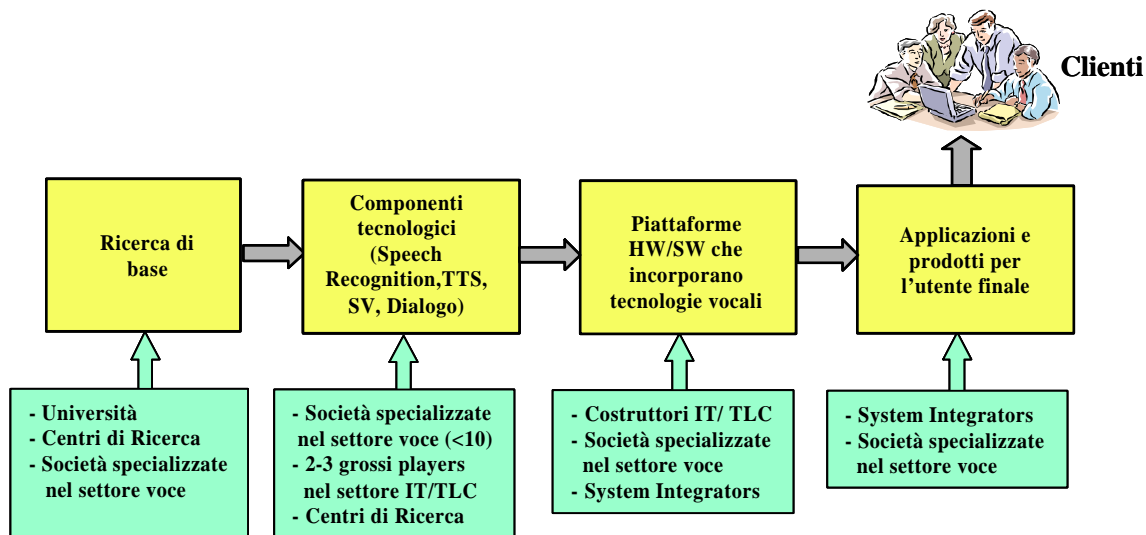


Fig. 1 - Aree di impiego delle tecnologie vocali

Su questo segmento operano costruttori di Information Technology/TLC, e società specializzate nel settore voce, oltre ad alcuni System Integrators che realizzano soluzioni utilizzando i componenti disponibili sul mercato.

resto del mondo . L'Italia rappresenta circa il 5,5% del mercato mondiale.

Si passerà ora ad analizzare la situazione italiana nei diversi segmenti, iniziando dal settore della ricerca.



Infine c'è il segmento delle applicazioni e prodotti per l'utente finale in cui operano prevalentemente System Integrators ed alcune società specializzate nelle applicazioni/soluzioni vocali.

Loquendo è una delle poche società al mondo che una propria offerta in tutti i segmenti ed è in grado quindi di soddisfare un ampio ventaglio di esigenze tecnologiche/applicative garantendo al tempo stesso un elevato livello di supporto su tutte le componenti.

La figura 3 riporta le previsioni di mercato nel settore delle tecnologie ed applicazioni vocali per ciascuno dei segmenti illustrati. Il mercato di gran lunga più significativo ed a più alto tasso di crescita è quello delle applicazioni telefoniche (automazione di Call Center, portali vocali, unified messaging, voice dialing, personal information management). Il mercato mondiale è suddiviso per il 46% al Nord America, per il 37% all'Europa e per il 17% al

Fortunatamente la ricerca italiana nel settore voce ha un'ottima tradizione. Nella sintesi vocale, ad esempio, le ricerche in Italia risalgono ai primi anni 70, ed il nostro paese è stato uno dei primi al mondo a realizzare un sistema di sintesi da testo (Musa 1979). Rispetto ad altri paesi il numero di ricercatori non è molto alto ma il livello scientifico è decisamente di primo piano come dimostrato dalla qualità delle pubblicazioni e dal prestigio di cui alcuni nostri ricercatori godono nella comunità internazionale.

L'attività di ricerca che sostiene lo sviluppo delle tecnologie di base è concentrata in pochi Centri di eccellenza nazionali tra cui vanno citati Loquendo, l'IRST, l'Istituto di Fonetica del CNR di Padova, il Politecnico di Torino, la Fondazione Bordoni, l'Università di Venezia, il CIRASS.

Molto alta percentualmente è stata la partecipazione a programmi di ricerca in ambito europeo (es. quinto programma quadro) e di

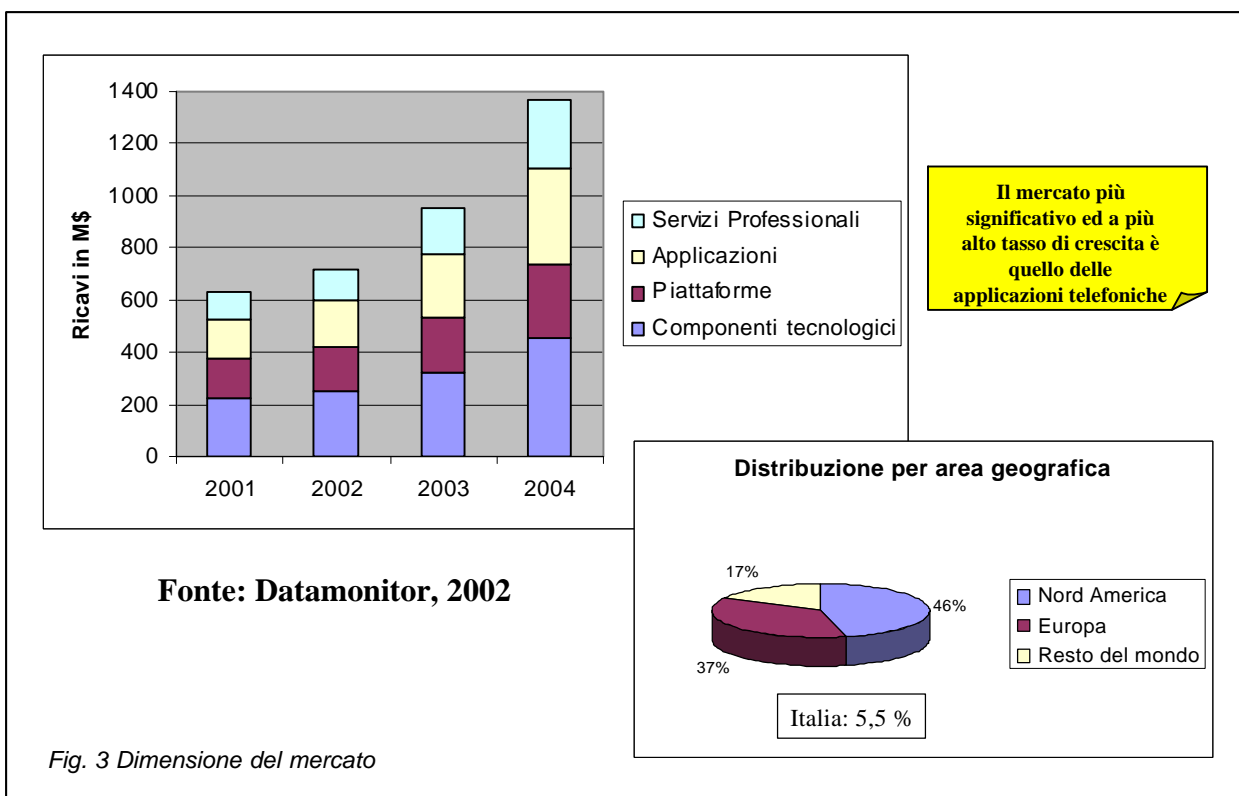
interesse nazionale (TAL) per la raccolta ed annotazione di risorse linguistiche, la sperimentazione di approcci innovativi, l'assessment di sistemi di sintesi e riconoscimento.

Tuttavia si rischia un progressivo impoverimento dei contenuti di ricerca innovativa a causa in parte alla esigenza di dimostrare ritorni a breve-brevissimo termine e conseguente focalizzazione su attività a basso rischio e basso livello di innovazione. Occorre equilibrare le iniziative industriali con iniziative volte a supportare/stimolare l'attività di ricerca innovativa per assicurare il mantenimento di un elevato livello di know-how dei nostri ricercatori ed il necessario progresso tecnologico senza il quale non si può guardare al futuro. La scarsità di finanziamenti alla ricerca, soprattutto nel settore pubblico, rischia di far assottigliare ulteriormente la popolazione di ricercatori e di far fuggire all'estero i migliori, problema purtroppo presente non solo in questo settore.

Alcune tematiche sfidanti nella ricerca della voce sono il miglioramento della capaci-

tà espressiva e simulazione di diversi stili comunicativi o emotivi nella sintesi da testo, le trasformazioni timbriche o prosodiche in grado di modificare in modo flessibile il segnale vocale preservandone la naturalezza, il riconoscimento in ambiente molto rumoroso e con voci sovrapposte, le strategie di dialogo ad iniziativa mista in domini applicativi complessi, l'adattamento automatico a livello acustico, lessicale e grammaticale, la traduzione automatica del parlato. Una grossa opportunità che richiede ancora sforzi di ricerca e' costituita infine dalla integrazione delle tecnologie del linguaggio scritto e del parlato per consentire la realizzazione di sistemi in grado di dialogare con l'utente in modo estremamente sofisticato e naturale. In questo campo il nostro paese puo' giocare un ruolo da protagonista dal momento che possiede tutte le tecnologie di base necessarie ed ha centri di eccellenza di livello internazionale.

Per quanto riguarda lo sviluppo di tecnologie vocali a livello industriale, l'Italia e' completamente autosufficiente in quanto possiede tecnologie vocali di ottimo livello con prestazioni ed efficienza comparabili, se non supe-



riori a quelle di fornitori stranieri. Sia Loquendo che IRST offrono su licenza tecnologia di riconoscimento vocale ingegnerizzata per le principali piattaforme elaborative.

Loquendo ha acquisito notevole esperienza su applicazioni di rete telefonica fissa e mobile e di conseguenza ha potuto ottimizzare in modo particolare il proprio riconoscitore su tali ambienti mentre IRST ha acquisito notevole esperienza su applicazioni di dettatura in parlato continuo ed ambiente desktop, ed ha quindi investito in modo significativo in tale direzione.

Loquendo offre anche una tecnologia di text/to/speech che si colloca al top di qualità per la lingua italiana ed è best in class a livello mondiale per diverse lingue.

Nel settore della comprensione e gestione del dialogo in linguaggio naturale abbiamo in Italia un background riconosciuto a livello internazionale e sono disponibili pacchetti software sviluppati dalle società ABLA, IRST e LOQUENDO. Loquendo ha infine intrapreso con determinazione lo sviluppo di sistemi multilingua offre TTS e speech recognition in 9 lingue, oltre ad un motore di comprensione e gestione del dialogo in linguaggio naturale.

Lo sviluppo di applicazioni vocali, in particolare nel settore dei servizi telefonici, è grandemente facilitato dall'utilizzo di piattaforme che offrono in forma integrata tutte le funzionalità che servono per sviluppare ad esempio un portale vocale o un risponditore automatico integrabile in un Call Center.

Grazie alle attività condotte nell'ambito del progetto TAL, è possibile utilizzare in modo indistinto tecnologie vocali provenienti da diversi fornitori ed interfacciarle con una varietà di piattaforme elaborative utilizzando interfacce comuni e protocolli standard.

Lo sviluppo di applicazioni diventa così accessibile a diversi soggetti e si creano le condizioni per uno sviluppo industriale su larga scala del settore dei servizi, condizione indispensabile per assicurare un ritorno sul fronte degli investimenti tecnologici e di R&D.

Una buona piattaforma per applicazioni TLC deve fornire un ambiente di sviluppo ser-

vizi estremamente ricco di funzionalità ed efficiente, fortemente orientato al riuso di moduli software, in grado di supportare standards come VoiceXML e SALT, scalabile da poche linee a migliaia di linee ed in grado di soddisfare requisiti di affidabilità estremamente stringenti, deve essere monitorabile e gestibile da remoto, deve poter generare con facilità dati di reportistica e tassazione, ecc.. La piattaforma VoxNauta di Loquendo offre tutto questo ed inoltre può vantare un ottimo livello di assistenza tecnica e manutenzione garantito direttamente dal fornitore.

La realizzazione di applicazioni vocali di successo, in grado di soddisfare gli obiettivi di business del Cliente e le aspettative di qualità degli utilizzatori finali è un fattore critico per lo sviluppo del mercato di questo settore. Mentre possiamo dire di essere allo stesso livello del Nord America come sviluppo delle tecnologie, la stessa cosa non può essere detta per quanto riguarda le applicazioni. Il nostro paese è indietro di almeno 3 anni nella consapevolezza delle opportunità di business offerte dall'impiego di sistemi conversazionali nella automazione di Call Center e nello sviluppo di servizi self-service in voce.

Loquendo sta investendo, attraverso lo sviluppo di servizi di grande diffusione, come i portali vocali (i-TIM Voice 4444, il servizio 412 di Telecom Italia) i servizi informativi (FS-Infoma: 892021) e dispositivi (IntesaTrade) a sviluppare questo settore ed abituare la grande massa degli utenti ad utilizzare correntemente servizi ad interazione vocale. C'è ancora molto lavoro da fare, in particolare in settori come la pubblica amministrazione ed i servizi al cittadino in cui l'utilizzo dell'accesso vocale può consentire di facilitare l'accessibilità delle informazioni attraverso un canale complementare al Web e di facile utilizzo come il telefono, estendere la disponibilità temporale dei servizi da operatore, ridurre i costi di gestione dei Call Center, abbattere la barriera di accesso alle nuove tecnologie come il Web e l'e-commerce, che possono trovare nel canale vocale una modalità di fruizione accessibile anche alle fasce non informatizzate.

Conclusioni

Il settore del trattamento del parlato è di importanza strategica sia per la promozione della lingua italiana che per lo sviluppo dell'industria e dell'occupazione in un settore tecnologicamente molto avanzato.

La situazione italiana in questo campo è nel complesso incoraggiante, soprattutto tenendo conto delle risorse limitate di cui si dispone rispetto ad altri paesi. Il settore della ricerca di base ha però ancora bisogno di sostegno finanziario, particolarmente per assicurare la promozione della lingua italiana e per stimolare una ricerca più coraggiosa ed avanzata che possa permetterci di mantenere la competitività nel futuro, affrontando le prossime sfide come la traduzione del parlato ed il dialogo in linguaggio naturale su ambiti linguistici e semantici ampi.

L'industria italiana ed il settore dei servizi ha bisogno di essere aiutata a crescere, il che può avvenire principalmente creando opportunità per progetti applicativi, ad esempio nella PA, stimolando lo sviluppo del mercato, sviluppando interventi per la formazione di personale skillato sulle problematiche vocali.

Profilo della società

Loquendo è stata fondata all'inizio del 2001 come spin-off del gruppo di ricerca di Telecom Italia Lab (ex Csel) sulla voce ed eredita quindi oltre 25 anni di background di R&D nel settore delle tecnologie ed applicazioni vocali interattive. La mission affidata alla nuova società è quella di valorizzare tali investimenti sfruttando industrialmente il proprio know-how e commercializzando i propri prodotti sui mercati internazionali. Loquendo ha oggi 174 dipendenti e si posiziona tra i player mondiali del settore, con una forte leadership di mercato in Italia, ed una penetrazione in Spagna e Sud America. Loquendo ha una offerta molto completa che va dalle tecnologie di base proprietarie (text-to-speech e riconoscimento in 9 lingue, verifica del parlatore), alle piattaforme vocali (VoxNauta), ai pacchetti applicativi, alla fornitura di servizi professionali per la realizzazione di soluzioni chiavi in mano offerte in modalità in-house o ASP. Tra le applicazioni basate sulle tecnologie Loquendo si possono citare i portali vocali I-TIM Voice 4444 e Telecom Italia 412-DimmiTutto, il servizio di informazioni su orari treni FS-Infoma, il servizio di Voice Trading per Intesa-Trade, i servizi 12 e 1412 automatici, l'automazione dei Call Center Targa, ecc..



Lo stand espositivo della Loquendo