

TECNOLOGIE INFORMATICHE DI SUPPORTO A DISABILI PER LA NAVIGAZIONE E L'APPRENDIMENTO

Minore di Dottorato di Ricerca in Ingegneria Informatica di
Dott. Ing. Thimoty Barbieri
Ciclo XV
Anno 2002

Politecnico di Milano

Relatore: Prof. L. Sbattella

-----REV 0.3 -----

1 Abstract

Utenti con disabilità fisiche, sensoriali e mentali possono trarre grande vantaggio dall'uso di tecnologie informatiche avanzate. La ricerca presenta alcuni studi svolti nell'ambito dell'impiego di tecniche sonore, vocali e musicali in contesti semi-immersivi, multimediali e distribuiti volti ad ampliare l'orizzonte funzionale o le opportunità di riabilitazione e apprendimento dei soggetti considerati.

In particolare verranno illustrati i risultati afferenti a tre diversi filoni di indagine:

- a) l'uso di strategie di ricerca e catalogazione gerarchizzata congiunta all'uso di tecnologie di sintesi e riconoscimento vocale che consentano un accesso più efficiente alle informazioni via Internet da parte di persone non udenti o con scarsa capacità motoria;
- b) l'uso del riconoscimento vocale e di tecnologie di rappresentazione semi-immersive a supporto della riabilitazione spazio-temporale in soggetti affetti da patologie congenite o acquisite (ritardo mentale o patologie psichiche e mentali);
- c) l'uso di tecniche multimediali, semi-immersive e cooperative per la creazione di storie musicali e musicate, a supporto di itinerari cognitivi rivolti a bambini affetti da gravi disturbi nella sfera della comunicazione verbale e della relazione.

2 Introduzione

L'uso delle tecnologie informatiche per aiutare le persone disabili è un campo in continua crescita, che consente di offrire sempre maggiori possibilità a terapeuti e pedagoghi, che si occupano della riabilitazione (fisica o cognitiva) di soggetti divenuti disabili in seguito ad una malattia o ad un incidente, oppure dell'assistenza a soggetti che soffrono di deficit cognitivi, sensoriali o motori congeniti o acquisiti.

Il progredire delle tecnologie informatiche apre inoltre opportunità alla società, che anche le persone disabili potrebbero cogliere: una per tutte la possibilità di accedere alle risorse informative della rete. Nella modalità standard di accesso, troppe barriere impediscono a soggetti non vedenti o con disfunzioni motorie, o con disordini di tipo cognitivo di sfruttare queste risorse.

Parte della ricerca in corso mira a fornire strumenti compensatori che consentano a queste persone di non rimanere escluse o isolate. Browser vocali, lettori con sintesi vocale, dispositivi braille, linee guida per una maggiore accessibilità dei siti, uso del suono e del rumore, periferiche tattili e con force feedback, sono solo alcune delle tecnologie che vengono in soccorso. Ultimamente la convergenza tra telefonia e reti di calcolatori fornisce nuove opportunità, quali aprire l'accesso alle informazioni della rete Internet anche attraverso un semplice apparecchio telefonico: una nuova opportunità da sfruttare per portare Internet vicino alle persone disabili.

Un altro interessante campo riguarda la possibilità di riprodurre in tempo reale immagini immersive o semi-immersive tridimensionali. La possibilità di rappresentare relazioni spaziali tra oggetti, e di collegarvi eventualmente suoni ed azioni, è di estremo interesse per la costruzione di esercizi per la riabilitazione di capacità cognitive collegate a disturbi dell'attenzione e dell'orientamento, o per il superamento di disagi anche gravi di tipo psicologico.. Combinare strumenti di rappresentazione tridimensionale con strumenti di controllo e feedback vocale potrebbe costituire un utile strumento in mano al terapeuta, che diventa in grado di realizzare - secondo alcuni criteri fondanti di tutti gli esercizi di tipo

riabilitativo - sistemi per misurare i progressi del proprio assistito e aiutarlo a realizzare dei progressi.

L'uso di suoni, controllo vocale, spazializzazione e scene tridimensionali sono veicoli che costituiscono un sicuro motivo d'interesse, stimolando memoria, motivazione, interesse, voglia di esplorare, e nel contempo offrendo un'alternativa sicura alla realtà, spesso rifuggita o vista con timore dal paziente. Per i soggetti più giovani, viene utilizzata la tradizionale tecnica della narrazione della fiaba per costruire un vero e proprio "mondo virtuale" in cui gli elementi della realtà, con le loro gioie, paure, difficoltà collegate, sono proposte in versione più controllabile e più sicura per il paziente. Le tecnologie multimediali ed in particolare la realtà virtuale semi-immersiva sono di ausilio per creare scene, suoni ed oggetti che siano difficilmente reperibili in realtà, o che esercitano un eccessivo impatto emotivo sul paziente.. Un sistema che consenta dunque di narrare storie nel senso più tradizionale del termine, ma offrendo anche possibilità di interazione con oggetti e suoni di natura diversa e configurabile da parte del terapeuta, e che offra un contesto percepito come sicuro in cui il paziente possa avventurarsi, può risultare estremamente prezioso.

Il presente lavoro si prefigge di approfondire questi tre temi:

- a) l'uso del controllo vocale e della sintesi vocale per consentire la navigazione delle informazioni presenti in rete, generalmente preclusa a persone non vedenti o con disabilità motorie, eventualmente anche attraverso l'utilizzo del telefono, senza costringere il soggetto a confrontarsi con l'elaboratore
- b) l'uso del controllo vocale e delle rappresentazioni tridimensionali per proporre un framework per la riabilitazione di disfunzioni cognitive legate a problemi di attenzione ed orientamento, che consenta al terapeuta di programmare esercizi di difficoltà crescente, misurabili e dalle caratteristiche programmate, e di realizzare situazioni virtuali protette che educino il paziente ad affrontare la corrispondente situazione nella realtà.
- c) L'uso della multimedialità, del suono, e dello spazio tridimensionale come supporto all'attività di narrazione, in cui il docente e l'allievo creano un contatto all'interno di

un mondo fantastico e virtuale, più ricco a livello espressivo e dialogico rispetto a quanto avviene in contesti tradizionali (che privilegiano il linguaggio verbale e la comunicazione visiva trascurando aspetti fondamentali il suono e l'animazione). .

I tre spunti, prendendo le mosse dalle proposte esistenti in letteratura, cercano di proporre delle direttive di marcia per nuove ricerche e nuovi lavori, cercando di esaminare le tecnologie disponibili e di proporre modalità per utilizzarli ed arrivare ad uno strumento prototipale concreto, che possa fungere da base su cui poggiare le idee per nuovi sviluppi.

Ben lungi dall'avere pretesa di completezza, tuttavia i tre punti trattati suggeriscono degli elementi interessanti, che sembrano avere buone prospettive future di sviluppo.

3 Natura delle problematiche dei soggetti portatori di handicap

3.1 *Disfunzioni sensoriali, motorie e cognitive*

Le menomazioni che possono affliggere una persona disabile sono di vario tipo, e possono essere congenite o acquisite a seguito di un incidente, di un decorso post-operatorio infausto, o di una malattia.

Una categoria di queste problematiche limita il soggetto nello svolgimento delle attività quotidiane dal punto di vista fisico: sensoriale (sordità, cecità), o motorio (paralisi parziale oppure totale, ...). Una seconda categoria riguarda invece i disturbi di natura mentale e psichica, derivanti da disfunzione o trauma cerebrale, oppure collegati a sindromi congenite o acquisite (ad esempio autismo, Alzheimer, , etc.

Dal punto di vista cognitivo, si possono individuare alcune categorie principali, più a scopo diagnostico e di studio, dato che la diversità dei processi mentali è tale, e dipendente dal soggetto e dalla natura degli stimoli a cui rispondono, da richiedere forzatamente una semplificazione della loro descrizione.

Le categorie riconosciute comprendono:

- La funzione esecutiva, che consente di risolvere problemi, pianificare, mettere in sequenza delle attività, decidere delle priorità, etc.
- L'utilizzo della memoria, nelle sue sotto-tipologie: a breve e a lungo termine, verbale, visivi, procedurale, dichiarativa, implicita, autonoetica.
- L'orientamento e la capacità di attenzione, che comprende la capacità di posizionarsi nel tempo e nello spazio, e di coordinare il proprio pensiero su determinati argomenti o seguendo determinate linee. In particolare si individuano diverse capacità di manifestare attenzione:
 - Attenzione focalizzata, che si riferisce alla capacità di rispondere coerentemente all'arrivo di uno stimolo esterno
 - Attenzione sostenuta (concentrazione), che consiste nella capacità di mantenere un comportamento coerente e consequenziale durante un'attività prolungata

- Attenzione selettiva, che riguarda la capacità di mantenere una successione di pensieri riguardo una singola attività in presenza di stimoli esterni di disturbo non pertinenti, che vengono scartati
 - Attenzione alternativa, che consiste nella capacità di passare da un'attività all'altra, con diversi presupposti, e poi tornare alla precedente con successo
 - Attenzione divisa, che riguarda la capacità di eseguire simultaneamente due attività che richiedono la risposta costante a stimoli esterni di natura diversa
 - Attenzione condivisa, che riguarda la capacità di svolgere azioni cooperando con altri
- La capacità di elaborare informazioni visive e spaziali, attraverso la percezione degli stimoli esterni e delle informazioni visive, derivandone elementi cognitivi
 - L'azione motoria e la risposta agli stimoli sensoriali
 - La capacità del linguaggio, e di espressione ai vari livelli di fluenza nella comunicazione.
 - La capacità di comprendere e controllare le proprie emozioni, la tolleranza al senso di frustrazione

Disturbi quali la dislessia o l'autismo comprendono disordini in una o più di queste aree, ma nessun soggetto, sebbene esista una classificazione in materia, è mai assolutamente comparabile con un altro, e per ciascuno è necessario un trattamento e un affiancamento personalizzato da parte del terapeuta.

Nel caso di disfunzioni fisiche o disturbi cognitivi in una delle aree sopra individuate, acquisiti a seguito di malattia od incidente, il paziente segue una cura di riabilitazione idonea, atta a valutare l'entità del danno e a cercare di recuperare parte delle funzionalità perdute (approccio ripristinativo), oppure ad insegnare al soggetto il modo migliore per superare le difficoltà date dalla disfunzione incorsa (approccio compensatorio). Nei soggetti affetti da malformazioni o sindromi congenite, è necessaria l'assistenza di terapeuti e pedagoghi specializzati, che insegnino fin dalla prima età al soggetto tecniche compensatorie alla propria disfunzione, o tentino di stimolare o di accrescere le capacità motili o cognitive del soggetto.

3.2 L'utilizzo di tecnologie informatiche

Le moderne tecnologie informatiche vengono in aiuto dei soggetti disabili da due punti di vista:

- a) offrono un sistema compensatorio alla disfunzione che affligge il soggetto, offrendo la possibilità di portare a termine un'attività che altrimenti per il soggetto sarebbe impossibile (lettura di una pagina di informazioni per un cieco, comunicazione telefonica per un sordo, utilizzo di sistemi di domotica per il controllo delle funzioni domestiche, etc.)
- b) offrono un supporto al terapeuta o al pedagogo durante il processo di riabilitazione o di educazione della persona disabile, fornendo strumenti che superano i limiti di quelli tradizionali e consentono dunque di esplorare e più efficaci, metodologie.

Considerando questi due punti di vista, le tecnologie informatiche accorrono in supporto in particolare riguardo a due punti nodali: consentire la fruizione delle informazioni (Basi di Dati, e, più in generale, Internet), anche a coloro che ne sono preclusi a causa di disfunzioni motorie o sensoriali, e aiutare nell'apprendimento o nella riacquisizione di nozioni perdute coloro che soffrono di disfunzioni cognitive.

3.2.1 Tecnologie a supporto della ricerca di informazioni

Per i non vedenti, sono oggi disponibili unità di visualizzazione braille, le quali sono in grado di trasformare testo ricevuto in una finestra di interazione con l'utente, in una serie di rilievi braille, ottenuti mediante il dislocamento successivo di aghi disposti in fila. Tuttavia, questo dispositivo è piuttosto costoso, e una valida alternativa è costituita dai sistemi di sintesi vocale (TTS, Text To Speech), qualora il soggetto non sia anche non udente.

Esistono in generale delle linee guida di Accessibilità dei Siti Web standardizzate dal W3C, nell'ambito del progetto WAI (Web Accessibility Initiative¹) le quali si prefiggono di fissare

¹ W3C WAI - <http://www.w3.org/WAI/>

dei parametri con i quali valutare la difficoltà che persone con vari livelli di disabilità possono incontrare nel consultare un particolare sito.

Tra le linee guida di accessibilità², ad esempio, risulta utile proporre siti unicamente contenenti testo (in questo modo più facilmente convertibili per il lettore braille o per il browser vocale), evitare video o immagini complesse non descritte, organizzare i form in modo che il tasto TAB segua un ordine sequenziale, mantenere un layout consistente, con colori ad alto contrasto (eventualmente personalizzabili), per facilitare persone affette da daltonia o con cecità parziale.

Alcune linee guida indirizzano anche su come ottenere una maggiore accessibilità per disabili con disfunzioni cognitive, quali dislessia, discalcolia, disturbi di attenzione o di memoria³. Tra gli accorgimenti principali, quello di fornire più modi alternativi per consumare la medesima informazione: ad esempio la possibilità di leggere lo scritto e di avere contemporaneamente la lettura in sintesi vocale, aiuta il dislessico nella comprensione del testo. Anche la presenza di immagini efficaci aiuta a comprendere quali siano i punti principali del testo. Tuttavia immagini animate causano problemi a persone con disturbi di attenzione, e risulta spesso utile evitare l'inserimento di questo tipo di immagini o fornire un meccanismo per fermarle. In generale, poi, un'organizzazione chiara e ben strutturata dei contenuti favorisce le persone con problemi di memoria o di apprendimento, evitando metafore particolarmente astruse (o fornendone accanto una spiegazione), o linguaggio inutilmente forbito o troppo complesso. In alcuni casi l'utilizzo di sistemi di schematizzazione dei contenuti (ad esempio Ygnius⁴) per la creazione di mappe mentali, riesce a produrre strutture di navigazione ben organizzate e coerenti, che possono evitare di indurre un senso di confusione in persone con disturbi alle funzioni di apprendimento.

² [Judy Brewer](http://www.w3.org/WAI/EO/Drafts/PWD-Use-Web/Overview.html) , How People with Disabilities Use the Web, <http://www.w3.org/WAI/EO/Drafts/PWD-Use-Web/Overview.html>

³ Lisa Seeman, Inclusion Of Cognitive Disabilities in the Web Accessibility Movement

⁴ Ygnius Mind Manager - http://www.ygnius.com/site/0_home/

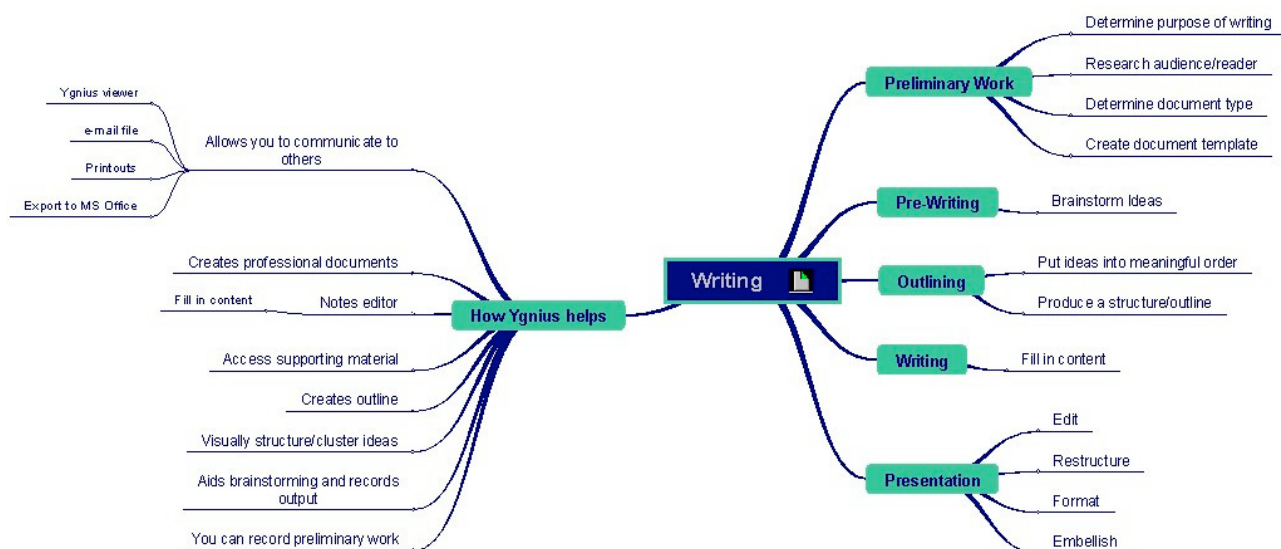


Fig. 3.1 : Un esempio di mappa mentale per aumentare l'accessibilità da parte di persone con disfunzioni cognitive, migliorando la struttura navigazionale di un sito web.

Esistono strumenti automatici, quali Bobby (<http://bobby.watchfire.com/bobby/html/en/index.jsp>), i quali consentono di raffrontare la struttura di un determinato sito con le linee guida W3C, e determinare automaticamente una valutazione di accessibilità.

Alcuni strumenti consentono di calare una interfaccia di controllo vocale sopra gli strumenti standard disponibili a tutti gli utenti (ad es. Window-Eyes), oppure offrono dei browser vocali (HomePage Reader, Sensus Internet Browser), i quali leggono mediante sintesi vocale il contenuto della pagina richiesta. In alcuni casi si ricorre a voci diverse nella sintesi per sottolineare caratteristiche diverse del testo, anche se risulta chiaro che questa tecnica deve rimanere limitata ad un ristretto numero di voci per non mettere in difficoltà l'ascoltatore.

Ulteriori aiuti provengono dal settore degli *haptics*, periferiche che consentono di trasmettere all'utilizzatore sensazioni tattili, quali mouse o joystick dotati di force-feedback (Immersion I-Feel...), oppure strumenti più raffinati quali dataGlove dotati di struttura haptics di force feedback. In particolare il mouse con force-feedback aggiunge alla metafora piana del "desktop" utilizzata nell'interazione con il sistema operativo, una dimensione tattile: scorrendo lungo i componenti delle finestre e delle icone, è possibile collegare una sensazione tattile da comunicare all'utente mediante la periferica haptics, consentendogli in

questo modo di percepire i contorni delle finestre, la posizione e la prossimità delle icone, e via dicendo, coadiuvando in questo modo le informazioni sonore.

Un'ulteriore tecnica per coadiuvare la navigazione ed aumentarne l'accessibilità è la *sonificazione*. Con essa si cerca di utilizzare suggerimenti sonori o musicali per coadiuvare la navigazione. Queste tecniche comprendono ad esempio l'associazione di suoni a link di diverso tipo (Beatnik, www.beatnik.com), oppure l'uso di piccoli brani musicali che alludano a un determinato argomento (ad es. se la pagina contiene un questionario, la sigla di un famoso quiz a premi televisivo⁵).

Un altro interessante plug-in (Speaks For Itself) consente di inserire script JavaScript per sonificare link o azioni web utilizzando un motore di sintesi TTS (Text To Speech). In questo modo è ad esempio possibile scrivere normali pagine web contenenti link “parlanti” o paragrafi “auto-leggenti” quando l'utente vi scorre sopra con il mouse (magari aiutato dai suggerimenti tattili di un mouse con force-feedback).

L'utilizzo di queste tecnologie, combinata con migliori sistemi di navigazione, basati sulla struttura del documento, e in generale una limitazione della profondità dei link, consentono di aprire l'accesso della grande quantità di informazioni a persone non vedenti o con limitate capacità sensoriali.

3.2.2 Tecnologie a supporto dell'apprendimento

Numerose tecnologie informatiche sono utilizzate in vario modo a scopo riabilitativo o di coadiuvazione all'apprendimento.

Nella riabilitazione da disturbi cognitivi di natura traumatica o acquisita, ad esempio, alcune “protesi cognitive”⁶, sistemi atti ad aiutare il paziente a superare il proprio disturbo cognitivo, per es. mediante sequenzializzazione delle attività più comuni, sono costituite da PDA contenenti programmi di assistenza al paziente⁷, oppure da programmi di uso comune espressamente sviluppati e semplificati per condurre l'utente a ritrovare la propria capacità di utilizzarli.

⁵ James F., Lessons from Developing Audio HTML Interfaces, in ASSETS 98 CA

⁶ Cole, E.; Dehdashti, P. – Computer-Based Cognitive Prosthetics: Assistive technology for the Treatment of Cognitive Disabilities in ASSETS 98

⁷ Gentry, T. – Cognitive Prosthetics: 21st century tools for the Rehabilitation of Thinking Skills – <http://www.cerebreon.com>



Fig. 3.2. Una “protesi cognitiva” sotto forma di PDA: ad intervalli regolari, il PDA, sempre indossato dal paziente, propone testo e piccole animazioni riguardanti la sequenza corretta con cui eseguire operazioni di uso comune: fare una telefonata, lavare il bucato, cucinare, etc.

Diversi casi di studio hanno dimostrato che pazienti con disfunzioni acquisite recuperano in molti casi gran parte delle proprie abilità, attraverso un miglioramento progressivo ottenuto mediante l’ausilio di questi sistemi.

Le caratteristiche salienti che un sistema informatico di supporto deve avere sono:

- possibilità di evidenziare chiaramente il passaggio da uno stato di funzionamento all’altro (ad es. da schermata a schermata, etc.)
- consentire all’utente di tornare in modo facile alla condizione corretta di funzionamento al verificarsi di un errore
- consentire all’utente di utilizzare tutte le funzioni principali senza sforzo
- richiedere breve tempo per l’addestramento all’utilizzo

In particolare il “linguaggio” visuale utilizzato dai programmi di supporto deve essere immediatamente evidente e chiaro, per consentire al paziente di riconoscere con sicurezza il messaggio che la “protesi cognitiva” deve comunicare al paziente.

Da questo punto di vista l'utilizzo della Realtà Virtuale e delle rappresentazioni tridimensionali immersive e semi-immersive possono risultare di deciso interesse nella realizzazione di strumenti che consentano la riabilitazione, ad esempio, di disfunzioni a livello di capacità di orientamento ed elaborazione spaziale, oltre che essere un veicolo che stimoli l'interesse del paziente per l'apprendimento di attività di vita comune.

L'utilizzo di rappresentazioni tridimensionali sintetiche aggiunge motivazione e stimola l'attenzione, consentendo a persone con disturbi nella capacità di concentrazione di rimanere più a lungo su un soggetto.⁸ Inoltre la possibilità di utilizzare forme e di esplorarle anche in prima persona consente un coinvolgimento diretto facilita i processi di memorizzazione. La memoria può essere inoltre esercitata allenando il paziente alla navigazione o alla ricerca di oggetti con caratteristiche ben definite.⁹

Spesso inoltre, è possibile ricostruire situazioni od oggetti che, se reali, creerebbero situazioni di difficoltà o disagio al paziente (si pensi a casi di agorafobia o acrofobia, che potrebbero essere curati passando prima per situazioni di realtà virtuale immersiva o semi-immersiva).

Anche simulazioni di vita quotidiana, come l'orientarsi in un supermercato e l'acquisto dei prodotti, possono essere riprodotti con sistemi tridimensionali semi-immersivi, dando al paziente una sensazione di maggiore padronanza di situazioni inizialmente percepite come difficili, e incoraggiandolo ad affrontarle anche nella vita reale¹⁰.

Di per contro, l'utilizzo di apparati poco confortevoli o inusuali (come head mounted display) può non essere accettato da tutti i pazienti, e le implicazioni tecnologiche della virtual reality richiedono l'assistenza di operatori tecnicamente preparati.

L'utilizzo della dimensione spaziale nella realtà virtuale immersiva e semi-immersiva consente di spazializzare anche gli effetti sonori. L'utilizzo della direzionalità del suono e della diversa intensità in relazione con la distanza possono costituire utili stimoli da collegare agli stimoli visivi: in questo modo persone con difficoltà di orientamento spaziale

⁸ Iso-Markku, P. & Seiler, O. 1998, *Virtual Reality in special Needs Education; Computers in Special Needs Education Seminar at the Department of Computer Science*, (1998, April 27 - last update), <http://www.cs.helsinki.fi/~isomarkk/vrml.html> [1999, Oct. 13].

⁹ Rizzo, A., & Buckwalter, J. 1995, *Center On Disabilities 1995 Virtual Reality Conference: Theoretical and Practical Issues for the use of Virtual Reality in the Cognitive Rehabilitation of Persons with acquired Brain Injuries - An Update*, [Online], (1998, Mar.10 - last update), Available: <http://www.csun.edu/cod/95virt/0004.html> [1999, Oct. 31].

¹⁰ Lannen T., Brown D. – Computer Interface Design to Virtual Environments For People With Learning and Physical Difficulties, in Proceedings ICCHP 2000

possono utilizzare la percezione sonora per imparare a collegare le due sensazioni e a comprendere l'informazione effettiva della posizione dell'oggetto, all'interno di esercizi di graduale difficoltà offerti in un ambiente virtuale.

4 Utilizzo di interfacce vocali per la navigazione

4.1 La convergenza tra telefonia e web

Sebbene siano ormai presenti numerosi software per la lettura sintetica delle informazioni, pochi di questi sfruttano le nuove possibilità offerte dalla convergenza tra la rete telefonica e quella informatica. Lo sfruttamento di questa opportunità consente di portare il browser vocale, con lettura vocale e a controllo vocale, direttamente dentro l'apparecchio telefonico, senza costringere l'utente a passare attraverso le fasi iniziali di attivazione del software di navigazione vocale. Del resto servizi automatici di telefonia di questo tipo sono di uso comune anche in applicazioni non espressamente dirette a persone disabili. Essi permettono agli utilizzatori di recuperare informazioni quali saldi di conti correnti, orari di aerei, treni e proiezioni cinematografiche attraverso un semplice telefono. La crescita esponenziale di Internet e delle tecnologie legate al Web ha fatto sì che i fornitori di servizi telefoni tradizionali si trovassero di fornire un nuovo insieme di clienti. Questi, invece di utilizzare le tecnologie tradizionali, preferiscono accedere alle medesime informazioni attraverso il web. Sebbene nella maggior parte dei casi i potenziali clienti accedano ai servizi automatizzati attraverso il telefono, gli sviluppatori e i fornitori di servizio trovano sempre più facile lo sviluppo di nuove applicazioni, sfruttando le potenzialità delle tecnologie Web. La Computer Telephony consiste nell'applicazione delle tecnologie IT per effettuare, ricevere e gestire le chiamate telefoniche o più in generale le comunicazioni. Il panorama della CT può essere segmentato in cinque categorie generali:

- Information Delivery
- Call Control
- Messaging
- Trasmissione
- Core Technology

Ogni categoria può essere ulteriormente segmentata per tipo di applicazioni e per utenti di soluzioni applicate ai problemi di ogni giorno. Per esempio, nel caso di Information Delivery, un utente può incontrare un servizio CT quando richiede informazioni a proposito del proprio conto corrente, utilizzando i comandi a tasto da un telefono. Allo scopo di riunire il mondo web con il mondo della telefonia, la CT si è sviluppata moltissimo negli ultimi anni. I primi tentativi si sono rivolti al trasferimento della voce sulla rete (Voice over IP). Lo sviluppo di questa tecnologia ha portato alla realizzazione di diversi protocolli, taluni molto complessi, tra i quali non si è ancora definito chiaramente uno standard. Tra questi si possono citare H.323 di Microsoft, e SIP (Session Initiation Protocol) delle IETF (Internet Engineering Task Force), con l'ultimo di questi il più probabile standard per il futuro. Un'altra importante capacità che viene resa disponibile dall'introduzione del VoIP è data dalla convergenza: esso infatti permette sia la comunicazione vocale tra due punti sulla rete, sia la comunicazione da PC a telefono, o addirittura tra due telefoni. Questo rende necessaria l'introduzione di un sistema del tutto generale per la comunicazione, dotato di portali che permettano la distribuzione del segnale vocale direttamente sulla rete.

Lo sviluppo delle tecnologie ha puntato quindi all'integrazione definita delle applicazioni Web e Telefoniche, con lo scopo di utilizzare lo stesso sistema per la gestione delle richieste sia provenienti dal "mondo IP", sia giunte dalla telefonia tradizionale. A questo scopo, si è sviluppato quello che si può considerare come lo standard de facto del settore della convergenza voce-IP, il linguaggio VoiceXML. VoiceXML poggia sul ben noto e diffuso linguaggio XML, definendo uno standard per applicazioni vocali. Pur non essendo particolarmente semplice, VoiceXML (anche detto VXML) risulta sufficientemente flessibile da poter gestire, in congiunzione con gli opportuni applicativi software, le stesse basi di dati usate dai tradizionali servizi basati sul Web. In questo modo è possibile realizzare sullo stesso sistema due versioni del servizio, accessibile in questo modo da entrambi i mondi.

Il poter accedere a servizi tradizionalmente identificati come Web direttamente dal proprio telefono ha come effetto collaterale la possibilità di fornire questi servizi a categorie che prima ne erano escluse. Tra queste categorie si trovano i disabili che, per qualche ragione non sono in grado di utilizzare le periferiche o le modalità tradizionali per l'accesso a Internet.

Lo scopo dei browser vocali è fornire un sistema telefonico di accesso alla rete, in modo da permettere l'accesso ai servizi Web a categorie disagiate. Tutto ciò deve poter avvenire con il solo utilizzo del telefono e nel modo più generale possibile, ossia senza la necessità di realizzare servizi ad-hoc. Il prototipo sviluppato a tal fine tenta di realizzare questo scopo.

4.2 Tecniche di navigazione vocale

Uno dei maggiori problemi nella realizzazione di un sistema di lettura vocale è rappresentato dalla politica di navigazione del testo, mediante controlli vocali, e dalle relative modalità di scansione del testo stesso.¹¹ Il testo visualizzato ha caratteristiche tali da poter essere consultato con un “accesso casuale”, dato dall’abitudine del lettore a saltare con l’occhio in porzioni diverse dal documento, aiutato in questo dal layout e dalla struttura del documento. Questo tipo di navigazione nel testo si mantiene anche quando il testo viene proposto dal convertitore braille, anche se l’analisi “a prima occhiata” della struttura del testo risulta più complessa per la mancanza di tutte le complesse convenzioni di layout disponibili per il normale testo scritto.

La lettura vocale, invece, è di natura strettamente sequenziale, e l’utente può solo muoversi avanti o indietro lungo lo scorrere della lettura. Questo aspetto rende la navigazione dell’informazione poco efficiente.

Uno degli approcci per risolvere questo problema è tentare di simulare la lettura “campionata” che l’occhio di un lettore esperto compie, cercando algoritmi in grado di estrarre le parole significative dal testo, e facendo leggere solo queste dal sintetizzatore vocale. Tuttavia, mentre questa tecnica appare efficace nella lettura visiva, nella lettura vocale il risultato è per lo più costituito da frasi sconnesse, scarsamente efficaci.

Un secondo approccio è quello di utilizzare una tecnica di lettura detta “gerarchica”. Questa tecnica cerca di comprendere la struttura del documento (titolo, paragrafo di primo livello, etc.), e consente di leggere il documento a uno dei livelli scelti, proponendo il testo completo di un dato livello gerarchico su richiesta del lettore.

Questo tipo di approccio si sposa bene con la struttura a tag annidati dell’HTML e l’uso in generale dei Cascading Style Sheets. Il prototipo vocale utilizza un analizzatore di HTML

¹¹ Choi, F. – Improving the Efficiency of Speech Interfaces for Text Navigation, in Proceedings of Conference on Educational Uses of Information and Communication Technologies, Pechino 2000

per tentare di determinare la struttura del documento e proporre una navigazione di base di tipo gerarchico.

Restano tuttavia aperti numerosi problemi, quali la navigazione vocale corretta delle rappresentazioni tabellari, o di layout complessi.

4.3 Architettura e funzionamento del browser vocale

Il browser vocale VioLin è un'applicazione destinata a permettere la navigazione su Internet a soggetti sprovvisti della vista, o comunque incapaci nell'utilizzo di un normale browser Internet come Internet Explorer o Netscape. Per rendere la navigazione più semplice possibile, si è pensato di utilizzare come interfaccia verso l'utente il mezzo di comunicazione più diffuso al mondo: il telefono. Grazie a VioLin è quindi possibile accedere alla rete componendo un numero ed eseguendo richieste vocali.

Il sistema deve essere idealmente composto da 3 parti:

1. un interprete vocale che gestisca la comunicazione con l'utente.
2. Un agente che si occupi di recuperare dalla rete i dati richiesti dall'utente.
3. Un agente che interpreti i dati letti e li renda di tipo vocale, in modo che l'utente possa accedervi

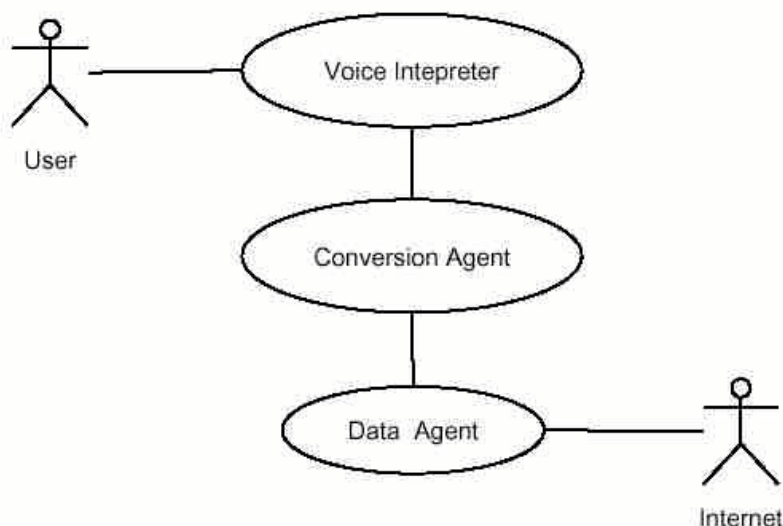


Fig. 4.1: lo schema concettuale di un browser vocale

Una analisi leggermente più dettagliata del problema porta a considerare una serie di possibili scenari di funzionamento, in modo da poter identificare con precisione le

funzioni che il sistema dovrà svolgere. Consideriamo un primo esempio: un non vedente desidera conoscere le previsioni del tempo per il giorno successivo.

1. Compone il numero associato a VioLin sul suo apparecchio telefonico
2. Una voce gli risponde chiedendo quale sito voglia visitare
3. L'utente detta l'URI del servizio meteo nazionale
4. Il sistema risponde chiedendo se la pagina deve essere letta per intero o se si vuole selezionare una sezione particolare
5. L'utente dice che gli interessa la sezione su Milano
6. Il sistema legge la sezione, chiedendo, dopo aver letto un certo numero di righe, se l'utente desidera proseguire
7. L'utente risponde di no e chiude la comunicazione.

Questo scenario evidenzia alcune delle funzioni più importanti che il sistema deve possedere: la capacità di interpretare gli URI specificati dall'utente e la lettura sequenziale o gerarchica delle pagine. Questi sono i requisiti fondamentali dell'applicazione, che deve permettere una navigazione, seppur limitata, delle pagine web in modo completamente vocale. Un altro scenario interessante potrebbe essere:

1. Comporre il numero associato a VioLin sul proprio apparecchio telefonico
2. Il sistema chiede se si voglia visitare un URI predefinito o inserirne uno nuovo
3. Si prosegue come nel caso precedente fino all'uscita
4. Il sistema chiede se si desidera visitare un altro URI
5. L'utente risponde di no e la comunicazione si chiude

Questo secondo scenario fornisce altre due funzioni molto importanti: la gestione dei segnalibri e la richiesta di una nuova operazione. Queste non sono funzioni indispensabili, ma possono risultare molto comode per l'utente rendendo in questo modo l'applicazione molto più usabile.

L'algoritmo generale del browser vocale può essere schematizzato come in Fig. x.x.

Questo semplice algoritmo descrive un primo sviluppo del browser vocale. Si potrebbero ovviamente aggiungere numerosi miglioramenti, come una funzione di back, un menu contestuale, ecc. Questi però non risultano indispensabili al funzionamento del sistema e possono anche essere di complessa implementazione, e pertanto non hanno trovato posto nel prototipo.

```
1: il sistema riceve una richiesta di connessione;
2: il sistema chiede se si vuole usare un segnalibro;
3: if viene richiesto un segnalibro then
4:   for all  $s \in$  segnalibri do
5:     Leggi  $t$ 
6:   end for
7:   segnalibro* := segnalibro richiesto;
8:   if  $\exists$  segnalibro* then
9:     ritorna all'inizio
10:  end if
11: else
12:    $u$  := URI richiesto all'utente
13:   if  $\exists u$  then
14:     ritorna all'inizio
15:   end if
16:   caricamento URI
17:   richiesta tipo di lettura
18:    $l$  := {sequenziale, gerarchica}
19:   if  $l$  = sequenziale then
20:     lettura sequenziale con conferma ogni  $n$  righe
21:   else if  $l$  = gerarchica then
22:     richiesta della sezione da leggere
23:     lettura sezione
24:     richiesta di una nuova sezione o uscita
25:   end if
26: end if
27: richiesta nuova operazione
```

Fig. 4.2: il funzionamento di un browser vocale

4.4 Tecnologie Utilizzate in VioLin

Il prototipo si basa sull'utilizzo di tool per la programmazione e la realizzazione di applicazioni basate su VoiceXML di W3C. Il motore vocale utilizzato è stato il WebSphere Voice Server di IBM, e come scripting engine per rendere dinamiche le applicazioni VoiceXML si è utilizzato Perl.

4.4.1 VoiceXML

VoiceXML riunisce i mondi della telefonia dello sviluppo Web, fornendo in pratica la possibilità di produrre pagine Web navigate attraverso la voce. I fornitori di servizi, utilizzando un linguaggio di markup come VoiceXML possono aprire i loro nuovi servizi Web utilizzando interfacce vocali, come ad esempio il telefono. Gli sviluppatori possono inoltre realizzare servizi vocali automatizzati usando le stesse tecnologie utilizzate per i servizi Web “visuali”, riducendo sensibilmente il costo di sviluppo e diffusione verso l'utente telefonico.

Le radici di VoiceXML affondano in un progetto di ricerca chiamato PhoneWeb di AT&T. Dopo la scissione di AT&T che ha dato vita a Lucent, entrambe le compagnie hanno cercato di sviluppare due linguaggi di markup telefonico indipendenti. Lucent ha proseguito nel progetto, ora noto come TelePortal, concentrandosi sulla creazione di servizi ed applicazioni in linguaggio naturale.

AT&T, invece, ha sviluppato un linguaggio di markup maturo, e una piattaforma utilizzata per lo sviluppo di diversi tipi di applicazioni, spaziando dai servizi di call center alla personalizzazione dell'offerta telefonica via Web. Motorola ha proseguito l'apporto di AT&T, allo scopo di fornire in tempo reale ai propri utenti mobili, consentendo un certo livello di interazione. In particolare, Motorola si è concentrata nel cosiddetto accesso a mani libere, che ha portato allo sviluppo del riconoscimento vocale più che all'utilizzo di sistemi touch-tone. Motorola ha basato il proprio linguaggio di markup su XML, e nel 1998 ha definito lo standard VoxML.

Nell'ottobre del 1998, grazie al crescente interesse rivolto all'idea di un linguaggio di markup per la definizione dell'accesso vocale ad applicazione Web, il W3C (world wide web Consortium) ha sponsorizzato un workshop sui Browser Vocali, che ha portato allo sviluppo di vari linguaggi, tra questi:

- SpeechHTML di Vocalis, che utilizza un sottoinsieme di HTML unito ad un insieme di politiche d'interazione
- SpeechML di IBM, volto a fornire un'interfaccia vocale ad applicazioni Web.

L'evoluzione di VoxML ha infine portato alla definizione dello standard W3C VoiceXML. VoiceXML offre un completo controllo della telefonia tradizionale (input, output, chiamata, segnalazione, ecc): riconosce comandi dati da tastiera, riconoscimento vocale, registrazione audio, la possibilità di riprodurre file audio, trasferimento di chiamata, sintesi vocale ed altre funzioni. Essendo basato su XML e simile ad HTML, Voice XML risulta anche facile da imparare per gli sviluppatori del settore. Per queste ragioni, VoiceXML è stato scelto come tecnologia da utilizzare nello sviluppo dell'applicazione VioLin.

4.4.2 IBM WebSphere Voice Server

IBM WebSphere Voice Server è un middleware che permette lo sviluppo rapido di applicazioni vocali. Questo sistema supporta standard quali Java e Voice-XML, e fornisce servizi di riconoscimento e sintesi vocale in diverse lingue. Il sistema è costituito da quattro componenti fondamentali:

- *Motore di riconoscimento vocale*: basato sulle librerie ViaVoice di IBM, ha lo scopo di interpretare e riconoscere i comandi vocali pronunciati dall'utente. Attualmente WebSphere riconosce diverse lingue, anche se il supporto completo è disponibile solo per la lingua inglese.
- *Motore di sintesi vocale*: serve per fornire informazioni all'utente attraverso la voce. Messaggi testuali vengono interpretati e inviati all'utente sotto forma di messaggi vocali sintetizzati in tempo reale. Anche qui esiste il supporto a diverse lingue.
- *Connessione alla Piattaforma Telefonica*: rappresenta l'interfaccia tra il server e la piattaforma utilizzata per la telefonia. Questa può essere di tipo VoIP, audio o in presa diretta o una connessione telefonica standard (tipo PBX).
- *Strumenti per lo Sviluppo*: è il livello più alto del sistema, composto da un insieme di API per lo sviluppo di applicazioni vocali tramite VoiceXML e Java

Il sistema di sviluppo (privo della parte di connessione alla piattaforma telefonica) è liberamente disponibile, e per questo è stato scelto per la realizzazione del prototipo. Per funzionare, WebSphere Voice Server necessita di un server Web, su cui effettuare le richieste di pagine VoiceXML. Il server interagisce con il sistema tramite i componenti di riconoscimento e di sintesi, permettendo l'accesso ai servizi offerti dalla rete attraverso un parser VoiceXML, che recupera le pagine da utilizzare da un server web accessibile. WebSphere è unicamente disponibile per le piattaforme Windows NT, 2000 e XP.

4.4.3 Scripting Serverside

Gli eventi ricevuti dal webserver e le API di sintesi e riconoscimento fornite dal kernel WebSphere sono utilizzati come ingressi della logica applicativa del browser vocale vero e proprio, realizzato utilizzando Perl. L'utilizzo di Perl ha consentito, mediante alcune librerie di utility, di scrivere rapidamente gli script di gestione degli input vocali e delle relative risposte di sintesi. Tra le librerie Perl utilizzate:

- Perl::CGI per la gestione dei parametri in arrivo via POST HTTP (pagine richieste, etc.)
- HTML::Parse per l'interpretazione degli URL HTML e la loro conversione in siti vocali via VoiceXML
- Perl::XML per la scrittura dinamica del codice XML-VoiceXML di risposta alle richieste di input.

4.5 Struttura del prototipo

Il prototipo risulta strutturato in 5 componenti principali:

- L'interprete vocale, costituito interamente dall'SDK WebSphere Voice
- L'agente di recupero delle pagine web, implementato mediante uno script Perl
- Il convertitore HTML-VXML, anch'esso realizzato mediante uno script Perl
- Il web server, che si occupa di gestire via HTTP le richieste e le risposte
- La logica di controllo dell'applicazione vocale (richieste, gestione risposte, etc.), realizzata in VoiceXML nativo.

L'interprete vocale consiste in un eseguibile da lanciare seguito dall'URI dell'applicazione vocale da eseguire (il browser VioLin). Il prototipo non realizza in realtà l'accesso alla piattaforma telefonica, tutto l'input e l'output audio avvengono in locale, attraverso l'utilizzo di microfono ed altoparlanti. Questo impedisce un effettivo test sulla rete, ma riduce notevolmente i problemi di riconoscimento (specialmente in caso di scarsa banda). La lingua utilizzata è l'inglese americano, in quanto risulta quella maggiormente supportata. L'agente di recupero è un semplice script che utilizza HTTP::Request per effettuare una richiesta di un file ad un qualsiasi sito web. Il file viene poi scaricato in una directory temporanea.

Il convertitore HTML-VXML utilizza le librerie HTML::Parse e HTML::TreeBuilder. Viene letto l'intero HTML e viene costruito un albero DOM (Document Object Model) corrispondente. L'albero viene poi esplorato alla ricerca di tag HTML supportati per il browsing vocale, o alla ricerca di testo libero. I tag non supportati vengono eliminati, mentre da quelli utilizzabili viene estratto il testo e vi viene associata la necessaria logica di controllo. Il prodotto della conversione è un file VoiceXML ben formato, che può essere passato all'interprete vocale per la navigazione.

La logica di controllo è realizzata in VoiceXML ed è costituita da segmenti di codice da inserire al posto dei tag conosciuti. Altre parti della logica di controllo sono il codice relativo alla lettura dei bookmark e degli URI.

Allo stato attuale, il prototipo è in grado di interpretare correttamente gli URI, di recuperare le pagine e leggerle all'utente in modo sequenziale o gerarchico. Sono previsti anche dei bookmark predefiniti da cui l'utente può scegliere.

4.6 Grammatiche di Riconoscimento

Per riconoscere l'input vocale da parte dell'utente sono utilizzate delle grammatiche che consentono di limitare il range delle possibili combinazioni di ingresso vocale e rendere pertanto il riconoscimento più rapido, affidabile e ripetibile. Sono state sviluppate diverse grammatiche: per la lettura degli URI, la gestione della navigazione, la selezione dei segnalibri, attraverso il linguaggio GSL (Nuance Grammar Specification Language).

Ad esempio, la seguente grammatica implementa la gestione della navigazione, collegando ad esso anche i toni telefonici, per un futuro utilizzo mediante telefono:

```

<grammar>
  <![CDATA[
    [
      [+ (dtmf-4) previous] {<option “previtem”>}
      [+ (dtmf-6) next] {<option “nextitem”>}
      [(dtmf-star dtmf-7) stop] {<option “stop”>}
      [(dtmf-start dtmf-5) repeat] {<option “repeat”>}
      [ (follow ?link) ] {<option “follow”>}
    ]
  ]]>
</grammar>

```

Un'altra utile grammatica consiste nell'aiutare la comprensione dello spelling delle parole mediante l'ausilio di parole. Questo si rivela particolarmente utile nella composizione degli URI, che spesso non sono riconducibili a parole note nel dizionario di lingua del kernel di riconoscimento. Un esempio di definizione di questa grammatica può essere il seguente:

```
#JSGF V1.0;
```

```
grammar com.ibm.url.en_US.url;
```

```

<a> = a [(like [in] | as [in] | for) (apple | abby | adam | alpha)];
<b> = b [(like [in] | as [in] | for) (boy | bob | baker | bravo)];
<c> = c [(like [in] | as [in] | for) (cat | cap | charlie)];
<d> = d [(like [in] | as [in] | for) (dog | david | delta)];
<e> = e [(like [in] | as [in] | for) (elephant | elevator | edward | echo)];
<f> = f [(like [in] | as [in] | for) (fred | forest | frank | foxtrot)];
<g> = g [(like [in] | as [in] | for) (gary | game | gold | george | golf | good)];
<h> = h [(like [in] | as [in] | for) (home | hello | henry | hotel | horse)];
<i> = i [(like [in] | as [in] | for) (insert | intruder | ida | india | island | indigo)];
<j> = j [(like [in] | as [in] | for) (job | janet | john | juliet)];
<k> = k [(like [in] | as [in] | for) (karen | kevin | king | kilo)];
<l> = l [(like [in] | as [in] | for) (long | lighter | larry | lima | lucy)];
<m> = m [(like [in] | as [in] | for) (mister | mother | mary | mike)];
<n> = n [(like [in] | as [in] | for) (norton | nora | nancy | november | nice)];
<o> = o [(like [in] | as [in] | for) (open | other | ocean | oscar)];
<p> = p [(like [in] | as [in] | for) (peter | paper | paul | papa)];
<q> = q [(like [in] | as [in] | for) (quarter | quit | queen | quebec)];
<r> = r [(like [in] | as [in] | for) (red | ralph | robert | romeo)];
<s> = s [(like [in] | as [in] | for) (summer | sandy | sam | son | sierra)];
<t> = t [(like [in] | as [in] | for) (top | time | tom | tango)];
<u> = u [(like [in] | as [in] | for) (under | unite | union | uniform)];
<v> = v [(like [in] | as [in] | for) (vibrate | vienna | villa | victor | victory)];
<w> = w [(like [in] | as [in] | for) (william | water | wilma | whiskey)];
<x> = x [(like [in] | as [in] | for) (xerox | xylophone | x ray)];
<y> = y [(like [in] | as [in] | for) (yellow | young | yankee)];
<z> = z [(like [in] | as [in] | for) (zipper | zest | zebra | zulu)];

```


<upper> = capital | uppercase | big;

<lower> = lowercase | little | small;

4.7 Utilizzo di VioLin

Il prototipo, alla partenza, pronuncia un messaggio di benvenuto, e chiede all'utente se desidera accedere al menu dei segnalibri. In caso di risposta negativa, si procede con la richiesta dell'URL che si intende visitare. In caso di risposta affermativa, viene richiesto all'utente di specificare una delle seguenti azioni:

- *Aggiungere un segnalibro*, specificando vocalmente l'indirizzo da registrare. Violin si occupa in questo caso di recuperare la pagina specificata, leggerne il titolo e registrare in modo persistente quest'ultimo associandolo all'indirizzo
- *Rimuovere un segnalibro*, pronunciando in linguaggio naturale il nome del segnalibro da eliminare.
- *Elencare i segnalibri*.
- *Visitare un indirizzo registrato tra i segnalibri*, pronunciando anche in questo caso il nome desiderato.

Merita particolare menzione il meccanismo di inserimento degli URL, sia per quel che riguarda i segnalibri, sia per la navigazione normale. Dato che gli URL sono formati da una grammatica molto complessa, con parole e numeri, non necessariamente appartenenti al linguaggio naturale esolitamente in diverse lingue, l'unico metodo possibile per l'inserimento prevede la pronuncia carattere per carattere degli indirizzi. Per minimizzare gli errori dovuti all'utilizzo del riconoscimento vocale, la grammatica degli URL prevede che i caratteri possano essere seguiti da una parola che li abbia come iniziale. Ad esempio, è possibile specificare la lettera "t" usando la locuzione "t like Tom". Ogni inserimento di un carattere è seguito da un tono di conferma.

Terminato l'inserimento non è raro che, nonostante le precauzioni, alcune lettere risultino sbagliate. Per questo il sistema chiede conferma dell'URL inserito e, in caso di risposta negativa, permette di riesaminare tutti i caratteri inseriti offrendo i seguenti comandi:

Add Before – aggiunge un carattere prima dell’ultimo carattere letto dal sistema, specificandolo nel modo anzidetto.

Add After – aggiunge un carattere dopo l’ultimo carattere letto

Change – Permette di cambiare l’ultimo carattere letto. Per facilitare la correzione viene proposto un carattere alternativo, dotato di fonema simili (e che quindi era facile scambiare) con una domanda, ad esempio “Did you mean L like Lucy?”

Erase – elimina il carattere letto

Finish – Conclude la modifica dell’URL

Inserito un URL o utilizzato un segnalibro, si procede alla lettura vera e propria delle pagine. Questa avviene per paragrafi: innanzitutto viene letto il titolo della pagina, seguito dal primo paragrafo.

4.8 Navigazione Gerarchica in VioLin

A questo punto si entra in un menu che offre la possibilità di passare al paragrafo successivo, di tornare indietro, di uscire o di elencare i collegamenti ipertestuali presenti nella pagina. Scegliendo l’opzione relativa ai collegamenti, è possibile scegliere uno di questi come prossima destinazione del browser con un menu del tutto simile a quello utilizzato per i segnalibri. Scelto un collegamento, le possibili azioni si ripetono da capo.

VioLin genera una versione di HTML semplificata mediante un’opportuna procedura interna, che gli consente di scartare tag HTML non significativi a determinare la struttura gerarchica di navigazione.

5 Navigazione vocale in ambienti tridimensionali: il modulo vocale di WebTalk

5.1 L'uso del comando vocale per l'orientamento nello spazio: opportunità riabilitative

La perdita delle capacità di localizzazione spaziale, o l'incapacità di combinare gli stimoli auditivi con quelli visivi per comprendere le caratteristiche geometriche-spaziali di una situazione richiedono un intenso allenamento per la riabilitazione o la acquisizione di queste capacità. Gli esercizi che vengono studiati a sostegno di queste problematiche, come altri esercizi volti al recupero o alla conquista di determinate capacità cognitive, devono essere strutturati con un'ottica di tipo "ripristinativo", e di conseguenza devono possedere le seguenti caratteristiche:

- devono proporre attività che si concentrino su una determinata abilità da recuperare, e che consentano un trattamento diretto del paziente
- devono essere tali da poterne livellare la difficoltà in proporzione alla caratteristiche del paziente, consentendo di cominciare da una versione di semplice esecuzione, e progredire verso un livello che viene considerato di capacità normale
- devono essere misurabili, in modo tale da poter valutare i progressi compiuti dal paziente tra più sedute
- devono poter essere ripetuti, identici, più e più volte, eventualmente con la possibilità di scomporlo in sottofasi più piccole
- devono poter dare al paziente un riscontro immediato dell'esito dell'attività, se andata a buon fine o meno

La possibilità di creare scene virtuali, in cui risulti possibile muoversi non mediante complicate interfacce (quali trackball, tastiera) di difficile uso anche per persone senza deficit cognitivi, o mediante interfacce semplificate che potrebbero essere rifiutate dal paziente, ma attraverso il controllo vocale, apre la possibilità di creare dei percorsi di training adeguati da sottoporre al paziente.

L'utilizzo di semplici comandi, che corrispondono alla volontà del paziente di spostarsi verso una determinata direzione, corrisponde ad una evoluzione della rappresentazione tridimensionale che può suggerire al paziente uno spostamento spaziale, che in situazioni reali potrebbe essere prematuro, difficile o indesiderato.¹²

5.2 Struttura di WebTalk

WebTalk è un progetto in fase di realizzazione presso il Politecnico di Milano, il cui obiettivo è di creare un sistema di rappresentazioni 3D semi-immersiva (senza l'ausilio di periferiche che consentano di conferire all'utente la sensazione di immersività nel modo, quali Head Mounted Displays o Guanti). Il sistema, basato su un server e su numerosi client, consente di proporre agli utenti una rappresentazione tridimensionale disegnata in VRML, e renderizzata mediante l'utilizzo delle API grafiche Java3D. Ciascun client, collegato alla medesima scena, può vedere visualizzata la posizione di altri utenti attualmente presenti nel mondo virtuale, creando una situazione di collaborazione. La finestra di lavoro dell'utente comprende un'ampia visualizzazione tridimensionale, una finestra di dialogo testuale con altri utenti, ed una lista degli altri utenti attualmente presenti nel mondo. Il sistema è studiato per valutare le possibilità date dalla rappresentazione tridimensionale per aumentare il grado di collaborazione possibile tra utenti remoti.

All'architettura corrente di questo progetto, è stato aggiunto un modulo di controllo vocale, che prelevando gli input vocali di un microfono in dotazione dell'utente, comanda opportuni spostamenti all'interno della scena tridimensionale. Questi spostamenti, come qualsiasi altro tipo di spostamento comandato al sistema con periferiche di input tradizionali (mouse, tastiera...), possono essere seguiti e notati in remoto da tutti gli altri utenti eventualmente collegati alla medesima scena.

L'utilizzo del kernel grafico di WebTalk e la possibilità di condividere azioni e movimenti possono farne un'utile cornice tecnologica per la realizzazione di esperimenti di riabilitazione o educativi, grazie alla possibilità del controllo vocale.

¹² Rizzo, A., & Buckwalter, J. 1995, *Center On Disabilities 1995 Virtual Reality Conference: Theoretical and Practical Issues for the use of Virtual Reality in the Cognitive Rehabilitation of*

5.3 Architettura e Funzionamento del Modulo di Controllo Vocale

Il modulo di controllo vocale è stato realizzato sotto forma di una finestra di controllo aggiuntiva, innestata nel client di navigazione tridimensionale. Per la navigazione, si è fornito un insieme molto limitato di comandi, in particolare corrispondenti alle principali sei possibili direzioni di navigazione nello spazio. Si sono poi aggiunti dei comandi di movimento (go, back, continuous, stop), che consentono rispettivamente di avanzare o retrocedere di un passo, della distanza desiderata (ed impostata mediante uno slider), e di navigare in modo continuo (e non un passo alla volta) nella direzione scelta, fino all'arrivo del comando di stop.

La modalità di navigazione continua, con impostazioni di velocità di navigazione contenute, ed un modo realizzato espressamente per una semplice navigazione (come ad esempio uno spazio aperto, circoscritto ai lati per impedire la fuoriuscita dell'utente dal campo di interesse dell'applicazione, e dotato di un oggetto evidente che possa costituire l'obiettivo del movimento), risulta molto semplice ed efficace. Il numero contenuto di comandi infatti, rende molto rapido e preciso il riconoscimento, anche dopo un training vocale breve del motore di riconoscimento.

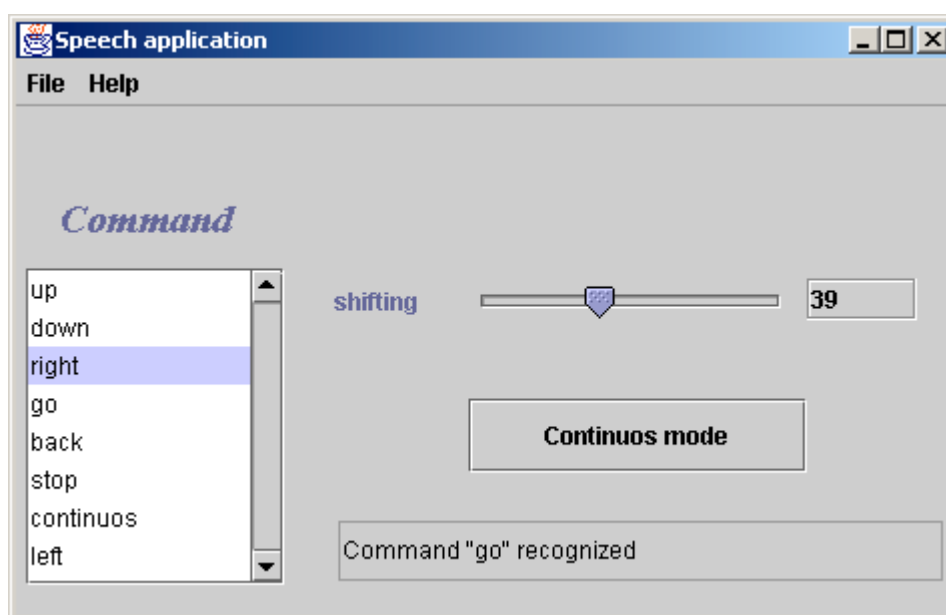


Fig. 5.1 : L'interfaccia di regolazione del modulo di comando vocale, che offre la possibilità di impostare la dimensione dei quanti di spostamento, entrare in modalità di spostamento continuo, e visualizza inoltre l'effettivo comando di navigazione riconosciuto (nell'immagine, il comando "go").

5.4 Tecnologie Utilizzate

Nello sviluppo del prototipo di controllo vocale si è cercato di utilizzare una tecnologia alternativa a quella utilizzata per VioLin, in modo da poter valutare anche le differenze riguardo agli strumenti messi a disposizione. In particolare si è scelta l'implementazione delle JSAPI realizzata da Cloud Garden (www.cloudgarden.com), che presenta il vantaggio di essere meno costosa in termini di licenza di utilizzo, e maggiormente adattabile rispetto all'intero kernel IBM di WebSphere, nella natura Client/Server dell'applicazione. Questa implementazione richiede l'uso delle SAPI di Windows come motore di riconoscimento/sintesi scaricabili liberamente dal sito www.microsoft.com.

In ambiente JSAPI esiste la possibilità di effettuare il riconoscimento vocale attraverso le due modalità canoniche: utilizzando parole all'interno dell'intero vocabolario, o attraverso una grammatica. Nel primo caso il riconoscimento risulta più lento, dato che la parola pronunciata deve essere ricercata all'interno di un set più ampio di vocaboli, nel secondo caso invece è possibile definire una grammatica ristretta (secondo le specifiche JSGF **Java Speech Grammar Format**), come fatto per il browser vocale. Essendo i comandi vocali di navigazione un insieme piuttosto ristretto, la grammatica risulta una brillante soluzione in termini di prestazioni e precisioni di riconoscimento, e contiene semplicemente i comandi di navigazioni nel seguente formato:

```
public <up> = Up {up};
```

dove <up> è il nome della regola, Up è la parola che deve essere pronunciata dall'utente per attivare questa regola, e {up} è il tag restituito dal motore quando viene riconosciuta la regola, da passare poi al motore grafico per la rappresentazione in rendering.

5.5 Realizzazione ed Utilizzo del Modulo di Controllo

Nella prova del modulo di controllo, si è utilizzato un ambiente pre-esistente che si adattasse ai criteri di semplicità navigazionale richiesti, rappresentante un cortile recintato da un portico e contenente un albero.

Il cortile, privo di ostacoli, consente di sperimentare in modo molto semplice la navigazione tridimensionale controllata vocalmente, sia in modo passo-passo che in modo continuo. La presenza del portico ai lati protegge l'utente da errori di navigazione, impedendogli di allontanarsi troppo dall'obiettivo della sua navigazione (l'albero posto in un angolo del cortile).

E' possibile selezionare diversi punti di vista: soggettivo, esterno rispetto alla propria posizione, o ripreso da un punto specifico dell'ambiente tridimensionale. Anche il punto di vista soggettivo risulta quello più indicato per questo tipo di esercizi, la possibilità di eseguire il movimento osservandolo da diversi punti di vista può costituire uno strumento aggiuntivo per comprendere la natura del movimento eseguito attraverso il comando vocale. In questo modo si è creato un semplice esercizio che consente di associare la volontà di movimento, espressa attraverso l'enunciazione di un verbo di comando appropriato, con una risposta visiva espressa in termini spaziali (l'effettivo avvicinamento verso l'albero in risposta al comando), il tutto all'interno di un ambiente protetto, ripetibile, e non rischioso, che consente di mettere in tutta tranquillità chi compie l'esercizio. Anche gli altri criteri risultano soddisfatti: l'esercizio è misurabile, in termini di tentativi compiuti o di grado di vicinanza raggiunto con la navigazione all'obiettivo, e dà a chi lo compie un immediato riscontro visivo dei propri risultati. Inoltre, l'esercizio può essere reso progressivamente più difficile, modificando opportunamente le caratteristiche della scena (più oggetti da raggiungere nella scena, più ostacoli, distanze maggiori, e via dicendo).

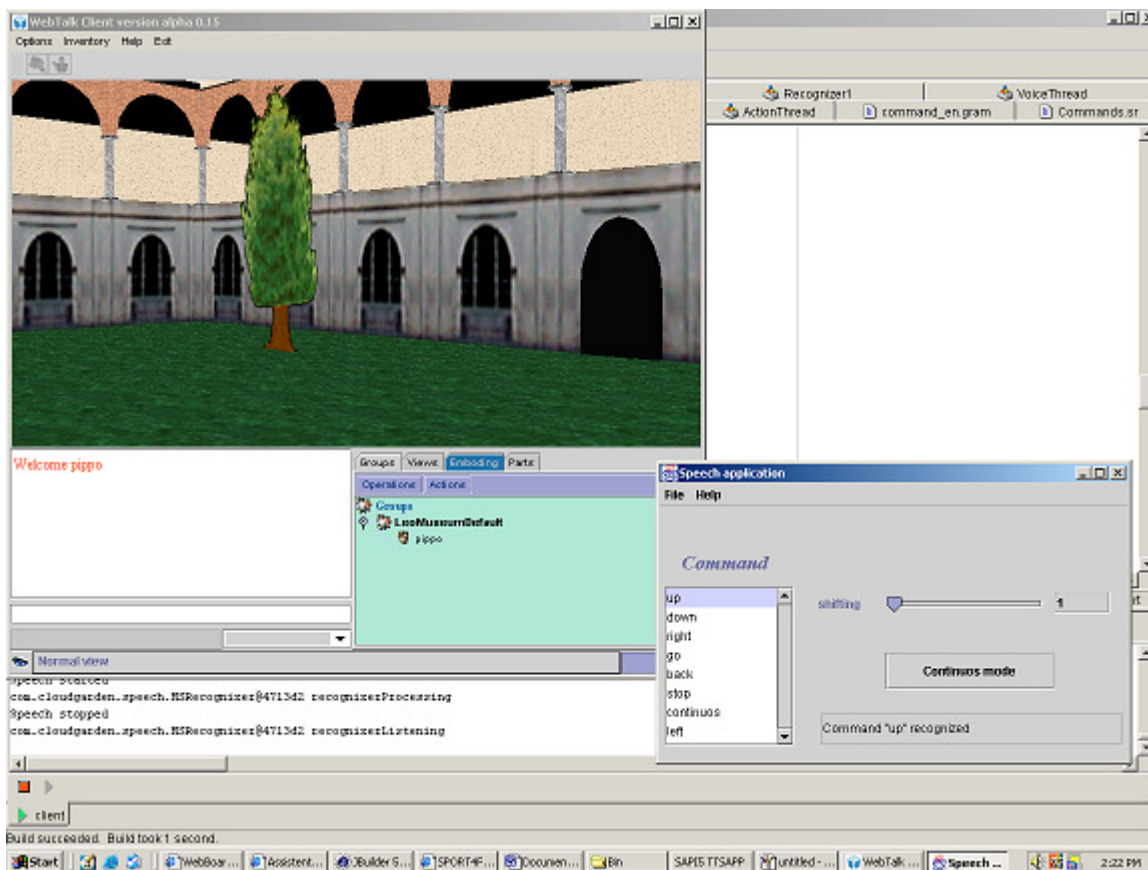


Fig. 5.2: l'esercizio di navigazione spaziale con controllo vocale inizia. Il modulo di controllo vocale è in basso a destra. In alto la scena dell'esercizio viene proposta all'utente. L'obiettivo è avvicinarsi all'albero mediante l'emissione del comando vocale 'go'.

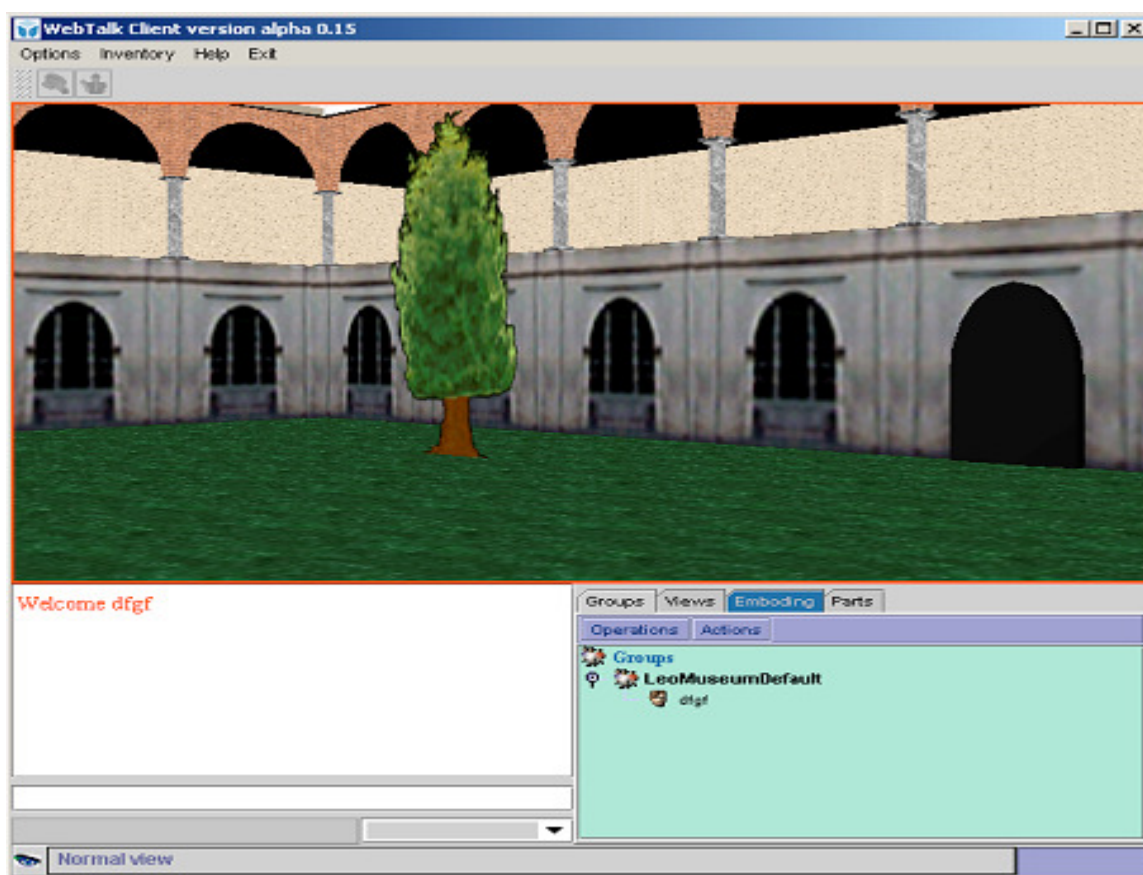


Fig.5.3: il comando vocale è stato emesso e la navigazione verso l'obiettivo è stata effettuata.

5.6 Sviluppi Futuri

La sperimentazione del prototipo di navigazione vocale ha suggerito alcuni sviluppi futuri che potrebbero rivestire un certo interesse. Ad esempio, nell'intento di facilitare ulteriormente la navigazione, potrebbe essere utile dare all'utente la possibilità di inserire dei bookmark per le posizioni salienti all'interno del mondo 3D, e di raggiungerli mediante un comando vocale definito dall'utente.

Un altro punto di interesse è quello di inserire degli stimoli vocali generati dal sistema attraverso una sintesi vocale, aumentando quindi le possibilità di interazione: ad esempio potrebbe essere ipotizzabile descrivere le caratteristiche del mondo attraverso un testo in sintesi vocale, o fornire dei suggerimenti navigazionali (troppo avanti, troppo in alto, scendi più in basso), per consentire all'utente di compiere l'esercizio anche senza l'aiuto del proprio insegnante.

6 Lo stimolo delle capacità cognitive mediante la narrazione interattiva

6.1 La narrazione interattiva

Quando il paziente affetto da un disturbo cognitivo linguistico o relazionale è un bambino, il racconto o la cos-costruzione di una favola può aiutare a creare occasioni che consentono di comprendere, misurarsi e vivere situazioni complesse o difficili sia dal punto di vista logico-deduttivo che dal punto di vista comunicativo o relazionale. I personaggi o le situazioni del racconto, consentono di affrontare e proporre in modo vivo, protetto e interessante le difficoltà e le soluzioni che si incontrano nella vita reale e che si desidera comunicare. I personaggi e le situazioni fantastiche vengono affrontate con maggiore interesse e disponibilità dal bambino che trova nelle sessioni di lavoro anche un buon supporto ai processi di memoria e di attenzione. Situazioni caratterizzate da gioia, tristezza, e ansietà, possono essere affrontati in modo attivo e cooperando con l'insegnante in contesti sicuri e sotto controllo.

Lo strumento del racconto viene anche utilizzato nei soggetti anziani affetti da sindrome di Alzheimer¹³, allo scopo di stimolare competenze cognitive e capacità di memoria.

La pedagogia e la riabilitazione tradizionale spesso utilizza a supporto della narrazione oggetti quali bambole, giocattoli e libri, o , incoraggia a disegnare usando carta e matite colorate. Anche la musica e i suoni possono contribuire ad attirare maggiormente l'attenzione e a rendere più vivida l'esperienza.

La possibilità di utilizzare lo strumento della favola non a senso unico, ma di consentire la creazione di variazioni, di diversi angoli di racconto, e di far interloquire l'ascoltatore nello svolgimento della storia e nella descrizione dei personaggi e delle loro azioni, è un ulteriore

¹³ **Patricia Heyn**, Functional Exercises for Alzheimer's Disease Based on Storytelling – Proceedings of 130th American Public Health Association Meeting, http://apha.confex.com/apha/130am/techprogram/paper_41790.htm

passo in avanti nello stimolare il bambino alla elaborazione autonoma delle informazioni e dei messaggi proposti dal terapeuta.

In questo contesto sembra ben adattarsi l'utilizzo dello strumento multimediale per creare situazioni narrative controllabili dal terapeuta e che consentano al paziente di sperimentare con suoni, forme, e colori, utilizzando strumenti nuovi e più coinvolgenti rispetto alla carta e alle matite colorate¹⁴.

La rappresentazione tridimensionale della scena, unita all'uso dei suoni e della musica, possono dare un'ulteriore occasione di stimolo. Nella scena 3D, il suono acquista una dimensione spaziale, e viene percepito come direttamente collegato a particolari oggetti. Navigando nello spazio, è possibile ascoltare come i suoni mutano in direzione ed intensità, aggiungendo alla possibilità di esplorare da ogni angolo gli oggetti proposti, anche quella di ascoltarli da ciascuna direzione. Queste possibilità possono risultare preziose in caso di bambini che preferiscono rimanere soli (e quindi hanno la possibilità di navigare o di "crearsi" la propria storia anche quando non sono direttamente affiancati), oppure di bambini i quali non gradiscono toccare fisicamente gli oggetti, e per i quali l'uso di giocattoli ed oggetti fisici durante la terapia o la narrazione è difficoltoso.

6.2 Struttura generale di uno strumento di narrazione interattiva

Prima della realizzazione tecnologica di un prototipo atto a dimostrare la fondatezza dell'uso di tecniche semi-immersive negli ambiti descritti, si è tentato di definire, in fase progettuale, le caratteristiche ideali cui dovrebbe corrispondere lo strumento da creare.

A tal fine si è cominciato con l'analizzare le esigenze principali dell'educatore, da una parte, e del bambino dall'altra.

Tra i principali requisiti individuati dall'educatore, i più importanti sono i seguenti:

- Avere la possibilità di visualizzare scene multiple, su cui è possibile una navigazione in avanti o in indietro. Questo consente all'educatore di scegliere gli scenari migliori in cui contestualizzare la propria narrazione.

¹⁴ Bellettini C., Bianchi A., Sbattella L. – Distributed and cooperative Story Telling for the treatment of cognitive and relational disorders – Proc. of 7th International Conference on Computers Helping People with Special Needs ([ICCHP 2000](#)), July 17-21, 2000, Karlsruhe, Germany

- Poter Inserire o Togliere oggetti, disponibili da una libreria di oggetti presenti. Attivarne o disattivarne le proprietà (renderli interattivi o meno), associarvi dei suoni o delle voci.
- Avere la possibilità di modificare i parametri legati agli oggetti, quali ad esempio il volume o la velocità di riproduzione del suono associato all'oggetto, per creare stimoli diversi da proporre associati alla narrazione.
- Avere la possibilità di registrare i commenti propri e del bambino durante la narrazione, in forma scritta, oppure in forma vocale.

Dal punto di vista del bambino, si è ritenuto che potesse essere importante fornire al bambino la possibilità di interagire con lo strumento nei seguenti modi¹⁵:

- Navigare nella scena, muovendosi nelle varie direzioni dello spazio, per mezzo di una tastiera semplificata o di un trackball
- Interagire con gli oggetti, toccandoli mediante uno switch button, oppure direttamente sullo schermo mediante un touch screen.
- Utilizzare alcuni oggetti appartenenti alla scena per "far sentire la propria voce", modularla in funzione dell'evento che si sta affrontando, renderla sempre più dialogica nei confronti delle altre "voci" protagoniste della storia arricchendo così sia l'articolazione della stessa che la qualità del proprio intervento e del proprio vissuto.

¹⁵ L. Sbattella, L'erbaMusicA

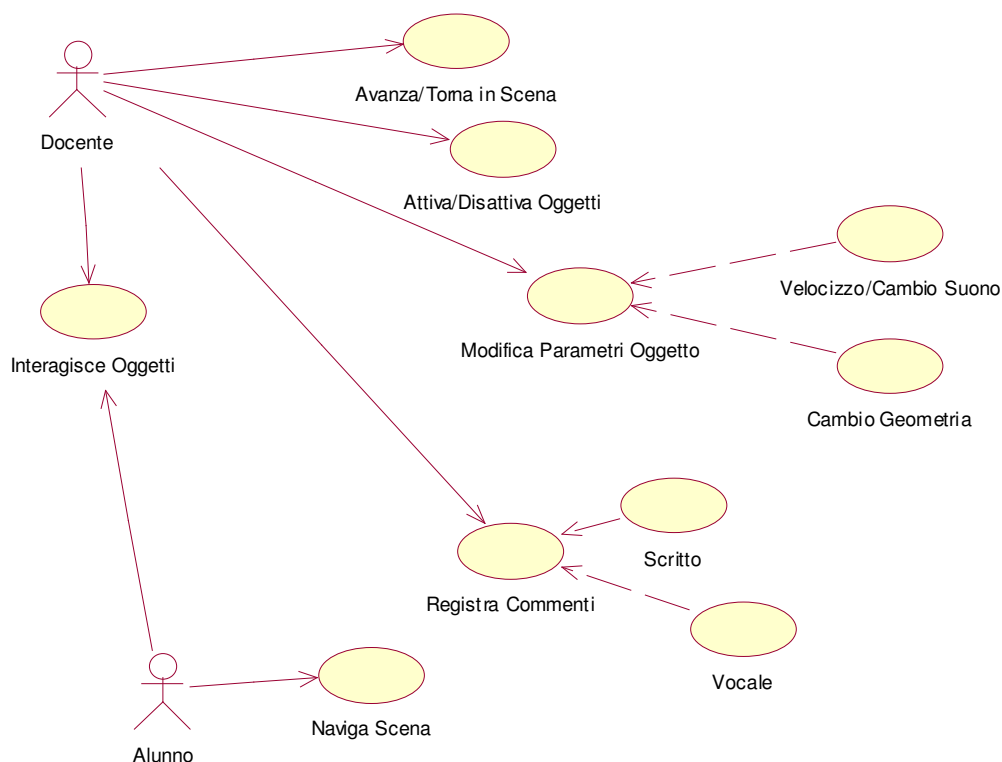


Fig. 6.1 La raccolta dei requisiti del Docente e dell'Allievo per uno strumento di narrazione interattiva.

Le necessità del Docente di poter variare la scena in funzione della narrazione, richiedono di poter offrire funzionalità avanzate di configurazione, che consentano idealmente di “costruire” per elementi le scene, assegnando ai componenti di ciascuna scena comportamenti e suoni. Oggetti, suoni, e comportamenti, potrebbero idealmente essere prelevati da una opportuna “libreria” messa a disposizione del Docente. Gli elementi della libreria dovrebbero tutti ricondursi ad un medesimo template (ad es. gli oggetti potrebbero essere tutti modellati in VRML, i suoni tutti .wav, ed i comportamenti modellati attraverso uno script JavaScript standard, che risponda esplicitamente ad una interfaccia di controllo di comportamento).

Pensando pertanto la configurabilità il prototipo attraverso il paradigma della “libreria”, è possibile ipotizzare la presenza di un terzo attore, denominato “Programmatore”, ma che potrebbe eventualmente coincidere con il Docente stesso, in qualità di configuratore della scena.

I requisiti del Programmatore della scena potrebbero essere i seguenti:

- Utilizzare una geometria come “fondale” di una scena. Collegare tra loro le scene in una sequenza navigazionale.
- Assegnare a ciascuna scena una voce narrante, un commento musicale
- Scegliere gli oggetti che devono comparire nella narrazione e disporli nelle scene.
- Assegnare a ciascun oggetto un comportamento predefinito (o programmarne di nuovi derivando dall’interfaccia di comportamento), e un suono predefinito.

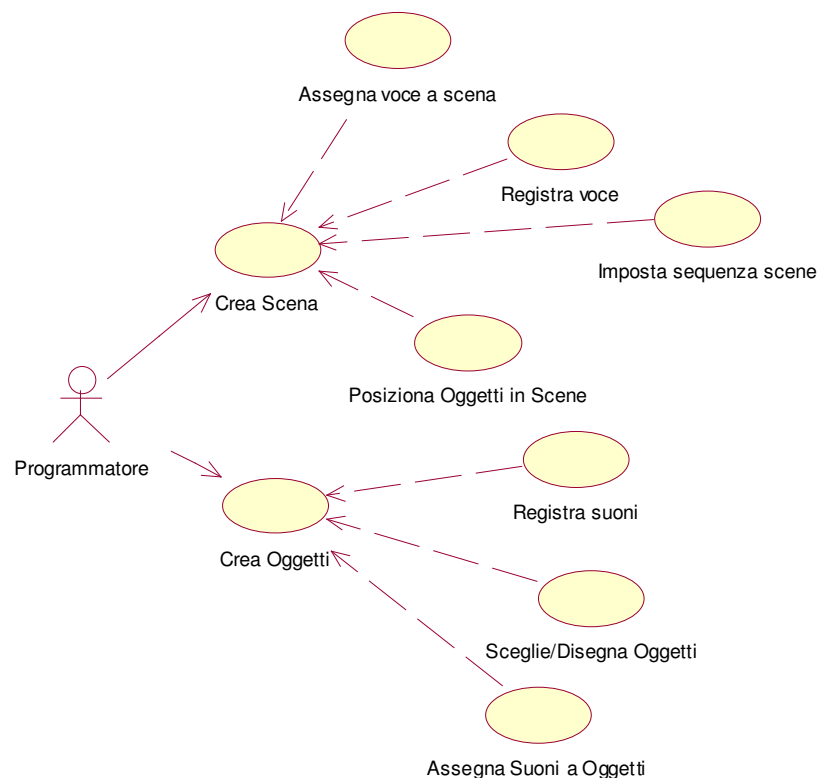


Fig. 6.2. I Requisiti di configurabilità delle scene di narrazione

Partendo da questi requisiti, lo strumento di narrazione dovrebbe presentare una struttura generale composta essenzialmente da tre elementi, raccolti in una struttura generale, che si potrebbe chiamare Canvas (o frame principale) applicativo:

- Logica di caricamento della configurazione della narrazione
- Presentazione delle opzioni run-time di variazione delle caratteristiche programmate di scene ed oggetti (una sorta di “pannello di controllo” a disposizione del Docente)

- Lo spazio di renderizzazione della scena 3D, che offre funzionalità di navigazione ed esplorazione

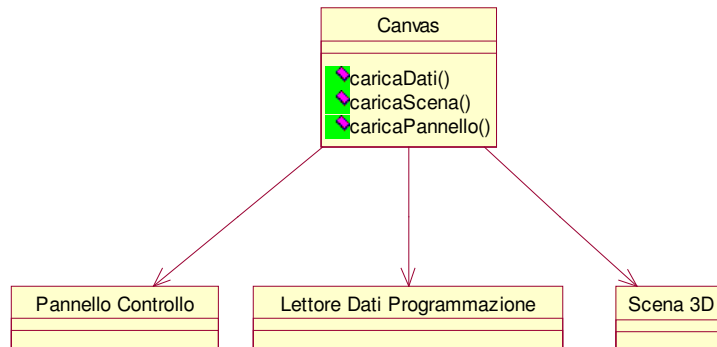


Fig 6.3. I tre elementi strutturali principali dello strumento di narrazione, raccolti nel Canvas

Per consentire la configurabilità della narrazione, la Scena 3D va pensata composta in Elementi. Gli Elementi sono entità geometriche che contengono una descrizione tridimensionale, ed una o più associazioni sonore e comportamentali. In questo modo, è possibile pensare sia ai fondali di scena che agli oggetti che li popolano negli stessi termini, mettendone a comune svariate proprietà di base, ad esempio la gestione degli elementi audio¹⁶. Gli elementi di tipo Scena hanno la possibilità di riferirsi uno all'altro nei ruoli di Scena Precedente e Scena Successiva, in modo da consentire la navigazione sequenziale necessaria alla narrazione.

Gli oggetti, intesi come elementi che è possibile disporre nella scena per completare la narrazione, possono derivare da una superclasse che descrive le capacità generali degli oggetti, successivamente espandibili per ereditarietà.

¹⁶ Gamma, E. et al. – Design Patterns, Prentice Hall, 1994

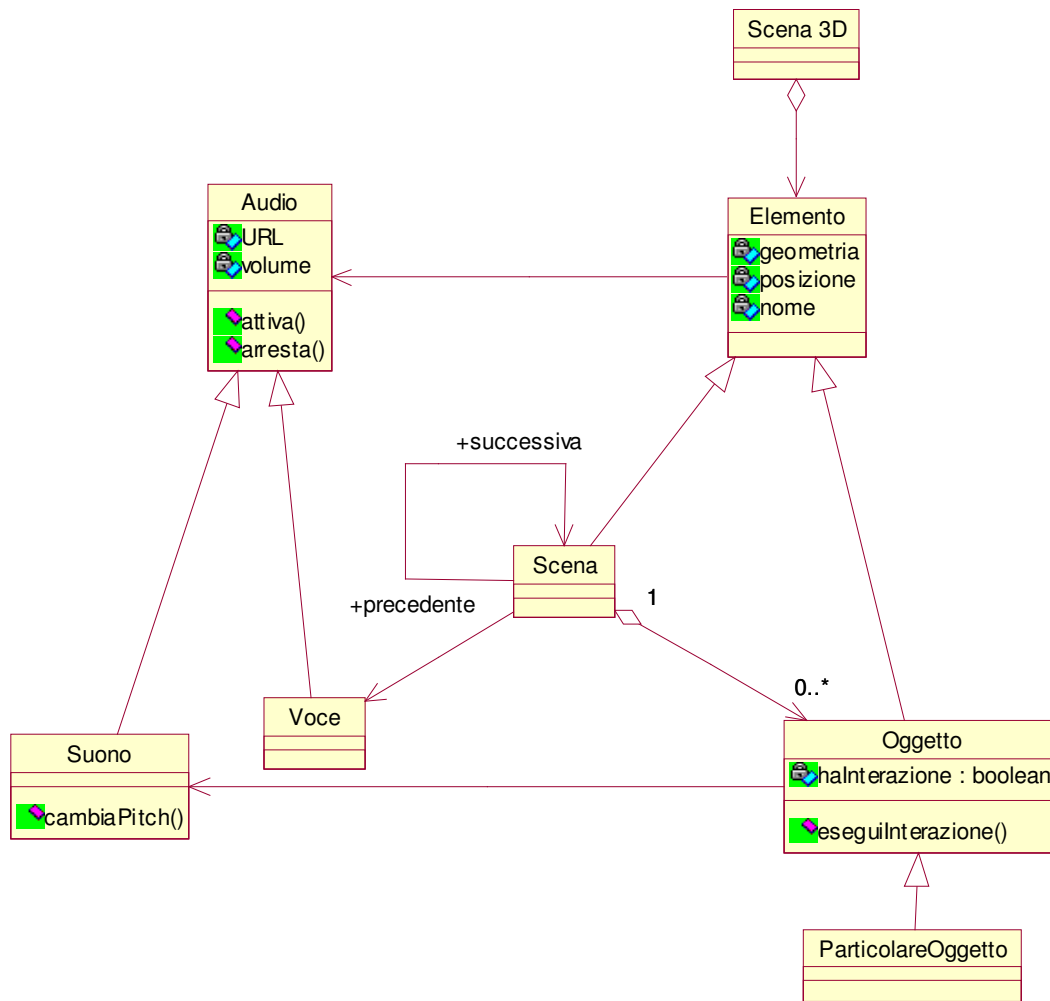


Fig. 6.4: La struttura concettuale degli elementi tridimensionali configurabili per la composizione della narrazione

In particolare si può pensare di utilizzare questo meccanismo di specializzazione per modellare una serie di comportamenti predefiniti che il Programmatore della narrazione può assegnare ai singoli oggetti. In particolare tutti i comportamenti dovrebbero supportare due metodi di tipo callback (ad invocazione, cioè, automatica): uno da invocare ad ogni porzione di tempo definibile, con un parametro che specifica la marca temporale di invocazione, ed uno da invocare ovunque non ci siano specificatamente altre azioni sull'oggetto (quando cioè l'oggetto si trova in stato di idle). Questi due metodi consentono di avere una struttura generale con cui specificare comportamenti temporali dell'oggetto (ad es. animazioni). Altri comportamenti possono derivare dal comportamento generale. Ad esempio, un comportamento che rende un oggetto sensibile al tocco può fornire due metodi

di interfaccia, uno all'inizio dell'evento di tocco, ed uno alla fine, entrambi che forniscano le esatte coordinate spaziali dell'evento scatenante.

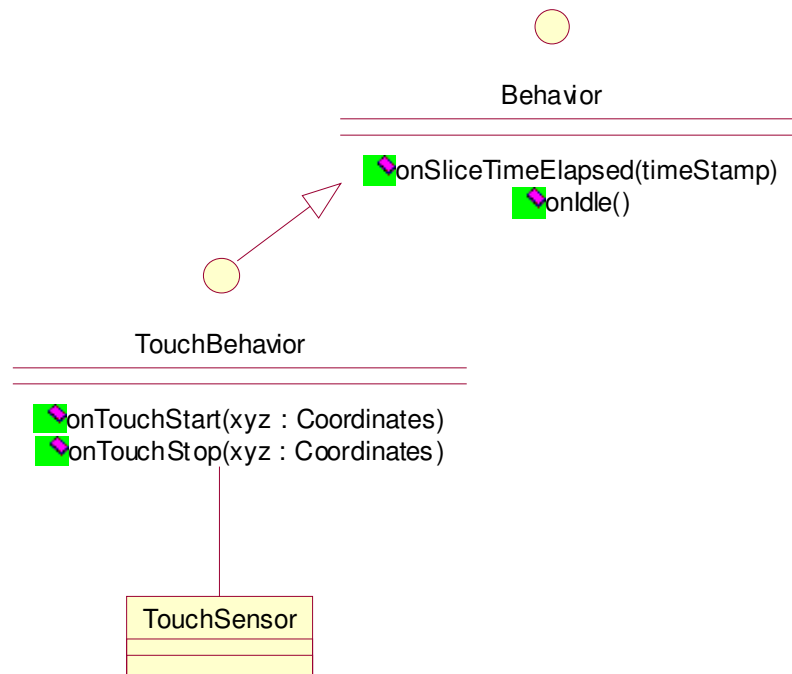


Fig 6.5. Una concettualizzazione del meccanismo ‘a comportamento’ necessario agli oggetti interattivi nello strumento di narrazione

6.3 MINUS-TWO¹⁷: un prototipo di StoryTeller per la narrazione interattiva di storie animate in VRML

La realizzazione del prototipo StoryTeller 3D è volta a sperimentare l'uso di oggetti tridimensionali e di suoni spazializzati per creare nuove occasioni di stimolo nell'ambito di una narrazione interattiva, in cui il terapeuta incoraggia il paziente ad intervenire e ad esplorare¹⁸.

¹⁷ ©ESAGRAMMA-ONLUS - Milano

¹⁸ Il disegno di MINUS-TWO - ambiente per la creazione di storie musicali interattive - è stato curato da Licia Sbattella e dagli specialisti del Centro ESAGRAMMA di Milano (Centro di terapia e formazione per il disagio psichico e mentale) e intende supportare le metodologie di Comunicazione Aumentativa Multimediale e di Musicoterapia Orchestrale ivi adottate.

Si desiderava creare uno strumento che fosse facilmente accessibile dal punto di vista tecnologico, non richiedesse particolari complessità di installazione, e che fosse al tempo stesso sufficientemente flessibile da consentire di creare e modificare gli oggetti tridimensionali presenti nella scen.

L'utilizzo del VRML e di un framework di pagine HTML consente di accedere ai contenuti della narrazione interattiva mediante un browser dotato di plug-in, ed un accesso alle risorse grafiche ed audio, fornite da una intranet interna o da un accesso Internet a larga banda (le risorse audio, infatti, per poter raggiungere un livello di qualità soddisfacente occupano uno spazio piuttosto voluminoso). Inoltre la struttura particolare del VRML ben si adatta ad implementare la struttura concettuale sopra descritta, mediante l'uso dei nodi Sensori, ad esempio collegato con il costrutto PROTO in VRML¹⁹.

La scena VRML si coordina con i suoni mediante l'uso di script JavaScript, collegati ai nodi VRML. Viene poi fornita una Applet di controllo, la quale, attraverso le API EAI di accesso al plug-in VRML, consente di creare un'interfaccia dotata di pulsanti, per consentire al docente di selezionare il commento musicale od il personaggio da far intervenire.

Oltre alla possibilità di muoversi all'interno della scena durante la narrazione, per esplorare meglio i paraggi, i personaggi, gli oggetti, ed ascoltare i suoni da prospettive diverse, è anche possibile manipolare gli oggetti spostandoli, ruotandoli o avvicinandosi, o selezionando direttamente l'oggetto dalla scena (richiedendo quindi l'uso del mouse e la capacità di selezionare l'oggetto di interesse all'interno dello spazio 3D), oppure scegliendo da un menu l'oggetto desiderato e premendo dei pulsanti di direzione, per consentire uno spostamento più semplice.

Altri oggetti possono essere attivati e produrre suoni o animazioni durante la narrazione, mediante l'uso del mouse, della tastiera, oppure di opportune periferiche (quali switchbutton, touchpad, etc.) atte a semplificare l'interazione con il mondo virtuale da parte dell'utente.

Le possibilità di movimento degli oggetti sono programmati mediante nodi VRML nella scena, mentre i pulsanti di controllo e la logica di movimento, animazione o suono che producono nella scena sono realizzati mediante codice Java, che interagisce sui nodi VRML attraverso lo strato EAI.

6.4 Struttura ed uso da parte del pedagogo degli strumenti StoryTeller

Come caso di studio per la narrazione tridimensionale interattiva, si è scelto di utilizzare “Pierino e il Lupo”, la fiaba musicale di Sergej Prokofiev.

La fiaba si adatta bene ad essere “implementata” tecnicamente, e grazie all’uso ben caratterizzato dei personaggi mediante i suoni, consente al terapeuta di creare spunti di esplorazione, di spiegazione e di narrazione interattivi con il bambino.

La narrazione è articolata in varie scene, una per ciascuna fase principale dello svolgimento della fiaba. E’ possibile passare da una scena all’altra seguendo i link HTML. Ciascuna scena è accompagnata dalla narrazione della storia, e dall’ingresso di personaggi diversi, che sono eventualmente manipolabili in modi diversi. A ciascuno dei personaggi è associato il noto motivo musicale assegnatogli da Prokofiev.

In ogni momento si può arrestare la narrazione audio automatica, e interagire con i suoni collegati con il singolo personaggio, oppure utilizzare gli strumenti messi a disposizione nella scena per produrre i propri suoni ed accompagnamento. Due strumenti messi a disposizione sono una coppia di tamburi, collegabili a degli switch button, ed una piccola tastiera di pianoforte. Agendo sugli switch button o sui tasti virtuali del pianoforte, è possibile suonare i due strumenti. Navigando in diverse posizioni, il suono sia della narrazione, che dei personaggi, che degli strumenti messi a disposizione dell’utente si spazializza.

¹⁹ Paolini P., Barbieri T., et al. – [Visiting a Museum Together: how to share a visit to a virtual world](#), in *Proceedings Museums & Web '99, New Orleans (USA)*, Marzo 99, pg. 27 e segg.

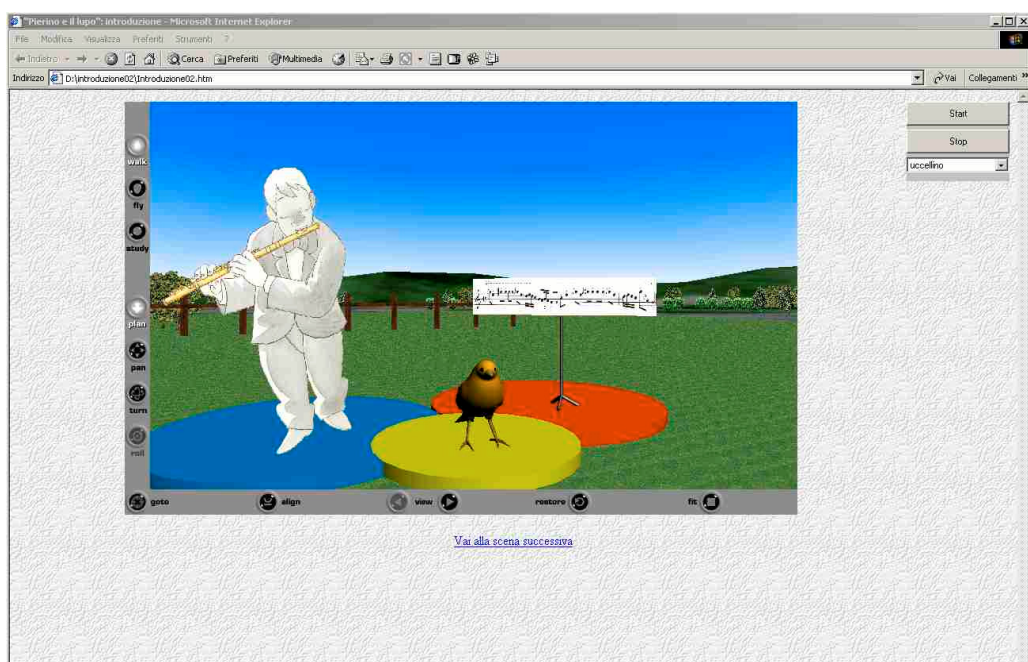


Fig. 6.6. Nell'introduzione, vengono presentati i personaggi della fiaba, associati al proprio motivo musicale.

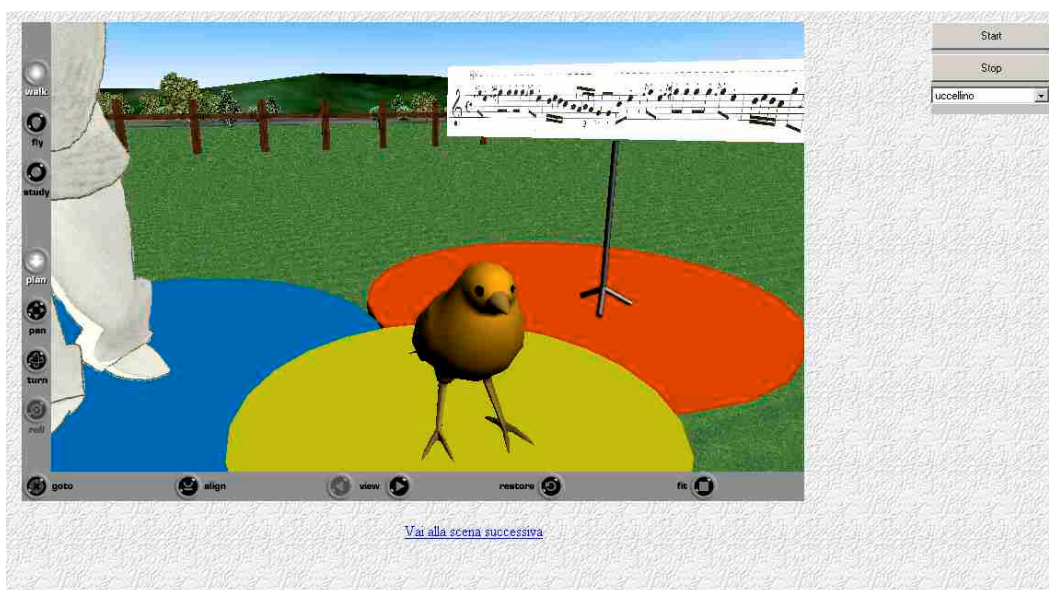


Fig. 6.7 Durante la narrazione, è possibile muoversi nella scena ed esplorare più da vicino, ascoltando le variazioni di suono durante la navigazione

Gli accompagnamenti messi a disposizione durante la presentazione delle scene sono di più tipi: uno, orchestrale, consente di ascoltare l'opera in modo integrale, e di muovere nel contempo i personaggi nella scena (o di osservarli muovere mentre seguono lo svilupparsi

della narrazione), l'altro propone solo un accompagnamento di pianoforte per consentire di cimentarsi nell'uso degli oggetti sonori interattivi (ritmo con i tamburi o con note dalla tastiera del pianoforte). E' anche possibile modificare la tonalità di riproduzione dei suoni legati agli oggetti della scena mediante un opportuno controllo. In questo modo si possono creare variazioni sonore (uso di pitch molto elevati o molto bassi, per ottenere effetti marcati) atte a proporre stimoli nuovi e ad allenare al riconoscimento delle variazioni sonore a ad offrire opportunità espressive varie ed articolate in dinamica ed agogica.



Fig. 6.8. La prima scena, con l'ingresso di Pierino, consente la navigazione attorno all'albero con l'uccellino e la casa del nonno. Due strumenti interattivi sono inoltre proposti in primo piano. Sulla destra una piccola console offre i controlli sonori, il comando di pitch per i suoni, e dei comandi di spostamento degli oggetti, alternativi alla scelta diretta dell'oggetto all'interno della scena 3D.



Fig. 6.9. I due strumenti interattivi, suonabili in qualsiasi momento, producono suoni spazializzati a comando di mouse o switch button. I tamburi (uno di tonica e l'altro di dominante) sono divisi in tre fasce di intensità (pp, mf, ff).



Fig. 6.10. I controlli a lato della scena consentono al docente di controllare le proprietà dei suoni e le posizioni degli oggetti principali

In generale l'uso di questo strumento prevede sempre che il docente affianchi il bambino (uno strumento viene suonato dal bambino, l'altro dall'insegnante), e controlli la narrazione, e decidendo di volta in volta se seguire la fiaba preregistrata proposta, o interromperla per assecondare gli interessi del bambino nel suonare, navigare, o riascoltare parti dei brani.

6.5 Sviluppi Futuri

Due futuri ampliamenti potrebbero essere previsti per un'applicazione di questo genere: il primo consiste nel fornire uno strumento visuale che consenta al docente di combinare in modo più semplice che non mediante la programmazione VRML le diverse scene. Sarebbe ad esempio sufficiente avere a disposizione una libreria (estendibile) di fondali, di personaggi, di oggetti e di suoni, e di consentire, attraverso un semplice editor, di inserire personaggi ed oggetti nei fondali, e di associarvi suoni e relative azioni scatenanti. Questo semplice editor potrebbe generare il codice VRML e JavaScript necessario per il caricamento nella struttura dello StoryTeller. In questo modo per il docente è possibile

rapidamente approntare nuove situazioni che possano aiutare a trasferire un determinato messaggio o sensazione al bambino.

Un altro aspetto potrebbe essere l'aggiunta di una dimensione collaborativa: l'uso di browser internet per la visualizzazione della scena rende questo passo relativamente semplice. Più StoryTeller potrebbero essere messi in comunicazione tra di loro. In questo modo si potrebbe – ad esempio – spingere soggetti che hanno difficoltà a socializzare tra loro ad interagire, interponendo uno “strato tecnologico” rivestito di forme animate e suoni. Le azioni eseguite all'interno di uno StoryTeller potrebbero trasformarsi in suoni e movimenti di oggetti riprodotti in altri StoryTeller. In questo modo potrebbe essere possibile per soggetti che lavorano in modo collaborativo scambiarsi reciprocamente stimoli di questo tipo sotto la direzione e la sorveglianza del docente.

La realizzazione tecnica di questo scenario prevede l'aggiunta di un piccolo strato di rete TCP/IP al modulo Java di controllo della scena, che si occupa di comunicare ad un server, che rimane in ascolto in rete, la lista degli eventi scatenati in una determinata istanza. Questo server si occupa poi di comandare agli altri client di replicare il medesimo evento scatenato in remoto.

7 Ringraziamenti

Questa ricerca, nell'indagare possibili direzioni da prendere per lo sviluppo di nuovi, più approfonditi progetti, è stata condotta con l'aiuto di numerose persone, che hanno contribuito con le proprie idee, con le proprie capacità ed il proprio lavoro. Fra questi:

- Browser Vocale
 - Giovanni Beltrame
- Modulo di Controllo Vocale 3D
 - Antonietta Lo Duca
 - Simone Basilico
- WebTalk
 - Roberto Ciatti
 - Matteo Foccoli
 - Davide Affaticati
 - Ferdinando Carella
 - Matteo Arru
- StoryTeller
 - Moreno Prencipe
 - Pinuccia Gelosa e Chiara Redaelli del Centro ESAGRAMMA

Il disegno di MINUS-TWO - ambiente per la creazione di storie musicali interattive - è stato curato da Licia Sbattella e dagli specialisti del Centro ESAGRAMMA di Milano (Centro di terapia e formazione per il disagio psichico e mentale) e intende supportare le metodologie di Comunicazione Aumentativa Multimediale e di Musicoterapia Orchestrale ivi adottate.

8 Bibliografia e Sitografia

Aggiornamento al Settembre 2002

8.1 Bibliografia

James F., Lessons from Developing Audio HTML Interfaces, in ASSETS 98 CA

Cole, E.; Dehdashti, P. – Computer-Based Cognitive Prosthetics: Assistive technology for the Treatment of Cognitive Disabilities in ASSETS 98

Lannen T., Brown D. – Computer Interface Design to Virtual Environments For People With Learning and Physical Difficulties, in Proceedings ICCHP 2000

Choi, F. – Improving the Efficiency of Speech Interfaces for Text Navigation, in Proceedings of Conference on Educational Uses of Information and Communication Technologies, Pechino 2000

Bellettini C., Bianchi A., Sbattella L. – Distributed and cooperative Story Telling for the treatment of cognitive and relational disorders – Proc. of 7th International Conference on Computers Helping People with Special Needs ([ICCHP 2000](#)), July 17-21, 2000, Karlsruhe, Germany
[L. Sbattella, L'erbaMusica](#)

Zajicek M, Powell C., Reeves C., - A Web Navigation Tool for the Blind – Proceedings of Assets 98

Gaura P. – Internal Document Structure of the Blind User Oriented Hypertext Audis - , in Proceedings ICCHP 2000

Miesenberger K. – Voice To Text Recognition Systems – Access to Spoken Language for People with Disabilities – Proceedings of Conference on Educational Uses of Information and Communication Technologies, Pechino 2000

Asakawa C., Takashi I, - User Interface of a Home Page Reader – Proceedings of Assets 98

Ressler S., Wang Q. – Making VRML Accessible for People with Disabilities – Proceedings of Assets 98

Song W., Ou S. – Using Virtual reality modeling to improve training techniques, Proceedings ICCHP 2000

Paolini P., Barbieri T., et al. – [Visiting a Museum Together: how to share a visit to a virtual world](#), in *Proceedings Museums & Web '99*, New Orleans (USA), Marzo 99, pg. 27 e segg.

Barbieri T., Loiudice P., et al. – [Il museo della Scienza Virtuale e WebTalk-I](#) in *atti conferenza IT-VRML '99* <http://www.lucia.it/vrml>, Giugno 99

Barbieri T., Paolini P. – [WebTalk: a 3D Collaborative Environment to Access the Web in Proceedings EUROGRAPHICS '99](#), Short Papers pg. 111-113, Settembre 99

Barbieri T., Paolini P., Alonzo F. – [Cooperative Visits to 3D Virtual Museums](#) , in *Proceedings International Conference for Cultural Heritage & MEDICI Day*, Milano, Sett. 99

Barbieri T., Paolini P. – [Cooperative Visits to WWW Museum Sites a Year Later: Evaluating the effect](#), in *Proceedings Museums & Web 2000*, Minneapolis (USA), April 2000

Barbieri T. – [Networked Virtual Environments for the Web: The WebTalk-I and WebTalk-II Architectures](#) , in *Proceedings IEEE for Computer Multimedia & Expo 2000 (ICME)* , New York, USA, July 2000

8.2 Sitografia

W3C WAI

<http://www.w3.org/WAI/>

Judy Brewer , How People with Disabilities Use the Web,
<http://www.w3.org/WAI/EO/Drafts/PWD-Use-Web/Overview.html>

Lisa Seeman, Inclusion Of Cognitive Disabilities in the Web Accessibility Movement
<http://www2002.org/CDROM/alternate/689/>

Ygnius Mind Manager - http://www.ygnius.com/site/0_home/

Beatnik – www.beatnik.com

ParallelGraphics Cortona – www.parallelgraphics.com

Java EAI – eureka.lucia.it/vrml/tutorial/eai/home.html

Java 3D – Sito Sun – java.sun.com/products/java3D

Speaks For Itself Plugin – www.speaksforitself.com

Immersion Ifeel – www.immersion.com

Gentry, T. – Cognitive Prosthetics: 21st century tools for the Rehabilitation of Thinking Skills –
<http://www.cerebreon.com>

Iso-Markku, P. & Seiler, O. 1998, *Virtual Reality in special Needs Education; Computers in Special Needs Education Seminar at the Department of Computer Science*, (1998, April 27 - last update),
<http://www.cs.helsinki.fi/~isomarkk/vrml.html> [1999, Oct. 13].

Rizzo, A., & Buckwalter, J. 1995, *Center On Disabilities 1995 Virtual Reality Conference: Theoretical and Practical Issues for the use of Virtual Reality in the Cognitive Rehabilitation of Persons with acquired Brain Injuries - An Update*, [Online], (1998, Mar.10 - last update),
Available: <http://www.csun.edu/cod/95virt/0004.html> [1999, Oct. 31].

VoiceXML – www.voicexml.org

IBM WebSphere – www.ibm.com/WebSphere

SIP Protocol – www.cs.columbia.edu/sip

Cloudgarden JSAPI – www.cloudgarden.com

Patricia Heyn, Functional Exercises for Alzheimer's Disease Based on Storytelling – Proceedings of 130th American Public Health Association Meeting,
http://apha.confex.com/apha/130am/techprogram/paper_41790.htm

9 Indice

<u>1</u>	<u>ABSTRACT</u>	2
<u>2</u>	<u>INTRODUZIONE</u>	3
<u>3</u>	<u>NATURA DELLE PROBLEMATICHE DEI SOGGETTI PORTATORI DI HANDICAP</u>	6
3.1	<u>DISFUNZIONI SENSORIALI, MOTORIE E COGNITIVE</u>	6
3.2	<u>L'UTILIZZO DI TECNOLOGIE INFORMATICHE</u>	8
3.2.1	<u>Tecnologie a supporto della ricerca di informazioni</u>	8
3.2.2	<u>Tecnologie a supporto dell'apprendimento</u>	11
<u>4</u>	<u>UTILIZZO DI INTERFACCE VOCALI PER LA NAVIGAZIONE</u>	14
4.1	<u>LA CONVERGENZA TRA TELEFONIA E WEB</u>	14
4.2	<u>TECNICHE DI NAVIGAZIONE VOCALE</u>	16
4.3	<u>ARCHITETTURA E FUNZIONAMENTO DEL BROWSER VOCALE</u>	17
4.4	<u>TECNOLOGIE UTILIZZATE IN VIO LIN</u>	20
4.4.1	<u>VoiceXML</u>	20
4.4.2	<u>IBM WebSphere Voice Server</u>	21
4.4.3	<u>Scripting Serverside</u>	22
4.5	<u>STRUTTURA DEL PROTOTIPO</u>	22
4.6	<u>GRAMMATICHE DI RICONOSCIMENTO</u>	23
4.7	<u>UTILIZZO DI VIO LIN</u>	25
4.8	<u>NAVIGAZIONE GERARCHICA IN VIO LIN</u>	26
<u>5</u>	<u>NAVIGAZIONE VOCALE IN AMBIENTI TRIDIMENSIONALI: IL MODULO VOCALE DI WEBTALK</u>	27
5.1	<u>L'USO DEL COMANDO VOCALE PER L'ORIENTAMENTO NELLO SPAZIO: OPPORTUNITÀ RIABILITATIVE</u>	27
5.2	<u>STRUTTURA DI WEBTALK</u>	28
5.3	<u>ARCHITETTURA E FUNZIONAMENTO DEL MODULO DI CONTROLLO VOCALE</u>	29
5.4	<u>TECNOLOGIE UTILIZZATE</u>	30
5.5	<u>REALIZZAZIONE ED UTILIZZO DEL MODULO DI CONTROLLO</u>	30
5.6	<u>SVILUPPI FUTURI</u>	33
<u>6</u>	<u>LO STIMOLO DELLE CAPACITÀ COGNITIVE MEDIANTE LA NARRAZIONE INTERATTIVA</u>	34
6.1	<u>LA NARRAZIONE INTERATTIVA</u>	34
6.2	<u>STRUTTURA GENERALE DI UNO STRUMENTO DI NARRAZIONE INTERATTIVA</u>	35
6.3	<u>MINUS-TWO©: UN PROTOTIPO DI STORYTELLER PER LA NARRAZIONE INTERATTIVA DI STORIE ANIMATE IN VRML</u>	41
6.4	<u>STRUTTURA ED USO DA PARTE DEL PEDAGOGO DEGLI STRUMENTI STORYTELLER</u>	43
6.5	<u>SVILUPPI FUTURI</u>	46
<u>7</u>	<u>RINGRAZIAMENTI</u>	48
<u>8</u>	<u>BIBLIOGRAFIA E SITOGRAFIA</u>	48
8.1	<u>BIBLIOGRAFIA</u>	48
8.2	<u>SITOGRAFIA</u>	50

2 **INDICE** **52**