



UNIVERSITÀ DEGLI STUDI DI PADOVA
DIPARTIMENTO DI INGEGNERIA DELL'INFORMAZIONE

TESI DI LAUREA MAGISTRALE IN
INGEGNERIA INFORMATICA

**FIABA MAGICA: SVILUPPO E SPERIMENTAZIONE DI
UN AMBIENTE INTERATTIVO MULTIMODALE PER
L'ATTUAZIONE DI PERCORSI DIDATTICI INCLUSIVI**

Relatore: prof. Sergio Canazza Targon

***Correlatori: prof. Antonio Rodà
prof.ssa Maja Roch
dott.ssa Serena Zanolla***

Laureando: Marco Calabrese

ANNO ACCADEMICO 2012 - 2013

Indice

Sommario.....	5
CAPITOLO 1 Stato dell'arte.....	7
1.1 Descrizione del problema.....	7
1.2 Stato dell'arte	8
1.2.1 Disabilità motorie: riabilitazione per i malati di Parkinson	9
1.2.2 Disabilità degli arti superiori e della mano: un'interfaccia per gestire l'interazione uomo-computer	13
1.2.3 Disabilità della vista: uno strumento per migliorare l'accesso alle informazioni web.....	21
1.2.4 Malati di Parkinson, quadriplegici, persone con ridotte capacità visive, persone senza braccia e anziani: un'interfaccia per facilitare l'accessibilità al World Wide Web	29
1.2.5 Stanza logo-motoria: apprendimento e comunicazione in un ambiente interattivo multimodale	40
CAPITOLO 2 Soluzione proposta	51
2.1 Soluzione.....	51
2.1.1 Fiaba Magica.....	52
2.1.2 Play&Learn	53
2.2 Motivazioni	54
2.2.1 Motivazioni connesse alla psicologia dello sviluppo del linguaggio.....	54
2.2.2 Motivazioni cliniche	62
2.2.3 Ambito normativo	64

CAPITOLO 3 Descrizione tecnica	67
3.1 Hardware.....	67
3.1.1 Requisiti hardware	67
3.2 Software.....	68
3.2.1 Requisiti software	68
3.2.2 Schema generale dell'interfaccia	68
3.2.3 Applicazione Kinect-based	69
3.2.4 Applicazione Flash	76
CAPITOLO 4 Sperimentazione	81
4.1 Sperimentazione	81
4.1.1 Protocollo di validazione dell'applicazione Play&Learn	81
4.2 Collaborazioni.....	83
4.3 Sviluppi futuri.....	84
Conclusioni.....	85
Bibliografia.....	87

Sommario

L'obiettivo che si pone il presente lavoro di tesi è sviluppare e sperimentare un ambiente interattivo multimodale in grado di supportare persone diversamente abili e normodotate nell'apprendimento.

Le motivazioni che giustificano l'impegno della ricerca in questo ambito sono molteplici. Dal punto di vista teorico, la creazione di un ambiente interattivo ha il suo fondamento nell'approccio *Learning by Doing*, nell'*Integrated Learning* e nella progettazione di ambienti sensibili. L'approccio multimodale è motivato dalla presenza in letteratura di molteplici interfacce, alcune delle quali verranno illustrate nel corso dell'elaborato. Il supporto a persone normodotate e diversamente abili è una conseguenza diretta dell'attuale contesto educativo europeo: l'aumento di studenti stranieri e diversamente abili, come definito dalle normative vigenti, rende necessaria la progettazione di adeguati strumenti di supporto. Infine l'apprendimento, di per sé, è una motivazione per questo lavoro di tesi; esso, avendo come prerequisito la comprensione dei testi, è sorretto da motivazioni specifiche che provengono dalla psicologia dello sviluppo del linguaggio.

Un ulteriore obiettivo che si prefigge il progetto è l'accessibilità dell'interfaccia: in questo contesto, l'hardware e il software necessario per il suo funzionamento devono essere obbligatoriamente *low-cost*.

Il lavoro è strutturato come segue: nel capitolo 1 verrà illustrato lo stato dell'arte relativo alle interfacce multimodali presenti in letteratura. Successivamente si descriverà la soluzione proposta focalizzando l'attenzione, nel capitolo 2, sulla dinamica di interazione e, nel capitolo 3, sui dettagli implementativi relativi ad hardware e software. Infine, dopo aver delineato nel capitolo 4 lo stato di avanzamento della sperimentazione, si esporranno le conclusioni.

CAPITOLO 1

Stato dell'arte

Nel presente capitolo verrà illustrato il campo di ricerca in cui è inserito questo lavoro di tesi, delineando il problema e presentando una raccolta bibliografica riguardo lo stato dell'arte.

1.1 Descrizione del problema

Il focus di questo lavoro di tesi è stata la seguente questione:

com'è possibile supportare persone normodotate o diversamente abili nell'attività dell'apprendimento?

Gli aspetti normativi (si rimanda alla sezione 2.2.3 per ulteriori dettagli) sono in grado di tradurre in modo più chiaro la medesima domanda: si discute della necessità di sviluppare uno strumento compensativo e/o dispensativo per supportare la didattica.

Riformulando le due considerazioni, il problema diventa quello di creare uno strumento in grado di supportare sia persone normodotate, sia diversamente abili nell'apprendimento. E in particolare esso dovrebbe:

- rivolgersi a utenti molto diversi tra loro con specifiche necessità: per esempio, tra i diversamente abili, studenti in situazione di multi-disabilità; tra i normodotati, invece, studenti stranieri o studenti che devono imparare la seconda lingua;
- essere flessibile e semplice da configurare (e personalizzare) per persone che non possiedono conoscenze informatiche specifiche.

Inoltre la problematica porta con sé due ulteriori questioni: da un lato, per poter aumentare il coinvolgimento, si richiede di ricorrere all'interattività; dall'altro le difficoltà specifiche delle persone costringono a utilizzare canali di input che si differenzino da quelli tradizionali (tastiera e mouse).

Nel cercare di definire dettagliatamente il problema, si è giunti a comprenderne, in parte, la chiave per poterlo risolvere: l'approccio multimodale¹, di cui avremo modo di parlare nello specifico nei prossimi paragrafi, sembra ben adattarsi al problema in questione.

1.2 Stato dell'arte

Prima di descrivere la soluzione proposta, si è ritenuto utile dare al lettore una panoramica generale delle interfacce che si prefiggono di rispondere alle medesime problematiche. È stata creata, pertanto, una raccolta bibliografica in cui si descrivono diverse interfacce multimodali; le prime sono rivolte a persone diversamente abili e, nello specifico, riguardano:

- disabilità motorie;
- disabilità degli arti superiori e della mano;
- disabilità della vista;
- malati di Parkinson, quadriplegici, persone con ridotte capacità visive e persone senza braccia.

Successivamente si è spostata l'attenzione sull'apprendimento, andando a descrivere un'interfaccia che rappresenta il "punto di partenza" di questo lavoro. La stanza logo-motoria ha finalità simili a quelle descritte precedentemente e si rivolge sia a bambini normodotati, sia a fanciulli con disturbi pervasivi dello sviluppo e disturbi specifici dell'apprendimento.

Naturalmente le interfacce che verranno illustrate non sono in grado di riassumere dettagliatamente l'intero panorama, ma permettono, in ogni caso, di comprendere le potenzialità dell'approccio multimodale.

Le interfacce sono state analizzate utilizzando uno schema comune di descrizione in grado di mettere in risalto i seguenti aspetti: chi sono i ricercatori coinvolti, qual è l'idea, il luogo e l'anno di realizzazione, gli obiettivi e gli aspetti tecnici-

¹ Il termine multimodalità, nel contesto dell'interazione uomo-computer (HCI), si riferisce alla creazione di interfacce che permettano l'interazione sfruttando più canali contemporaneamente. Nella maggior parte dei casi si utilizzano i sensi (tra cui la vista, l'udito e il tatto); tuttavia si potrebbe utilizzare qualsiasi altro canale (per esempio l'analisi delle attività cerebrali).

implementativi. In questo modo si è cercato di delineare le principali caratteristiche dei sistemi riproducendo lo schema logico, utilizzato soprattutto nel settore giornalistico, delle *5 W e 1 H* (chi, cosa, dove, quando, perché e come).

1.2.1 Disabilità motorie: riabilitazione per i malati di Parkinson

Il team

Un interessante progetto nell'ambito delle disabilità motorie è stato sviluppato da Camurri et al. (2003), ricercatori del laboratorio *InfoMus Lab* presso la facoltà di ingegneria dell'università di Genova.

I principali filoni di ricerca di questo laboratorio sono le interfacce multimodali uomo-computer e la comunicazione espressiva-emotiva non verbale. Ulteriori studi vengono inoltre compiuti sull'integrazione degli strumenti provenienti dalle scienze umanistiche nelle *performing arts* (arti sceniche o dello spettacolo) e sulla creazione di applicazioni nell'ambito *edutainment* (forma di intrattenimento che ha in sé un aspetto istruttivo) oltre a quello culturale, terapeutico e riabilitativo.

L'idea

L'interfaccia è basata sul concetto di risonanza estetica, vale a dire la possibilità di fornire ai pazienti un *feedback* multimediale, sulla base di un'analisi qualitativa del loro movimento. In particolare questa finalità viene raggiunta lasciando i pazienti liberi di muoversi nello spazio (senza alcun tipo di sensore sul corpo) ed eseguendo, in tempo reale, l'analisi del movimento e la conseguente produzione del *feedback*. La risposta multimediale (visiva e/o acustica) ha lo scopo di provocare reazioni ludiche nel paziente e, di conseguenza, stimolare fattori emotivi-motivazionali.

Il luogo e l'anno di realizzazione

L'interfaccia è stata prodotta presso i laboratori di *InfoMus* e gli esperimenti pilota sono stati eseguiti in collaborazione con il Centro di Bioingegneria presso l'ospedale La Colletta, ASL 3, Genova.

La prima parte della sperimentazione è terminata nel dicembre del 2003 con la pubblicazione dei relativi risultati scientifici.

Gli obiettivi

La ricerca possiede finalità specifiche sia nell'ambito tecnico, sia in quello clinico.

Nel primo settore si possono individuare due aspetti principali: da un lato la necessità di sviluppare un'architettura di calcolo aperta in grado di eseguire esercizi terapeutici che integrino l'analisi e riconoscimento gestuale con la generazione di

feedback multimediali. Dall'altro l'esigenza di sviluppare algoritmi precisi per l'analisi dei movimenti (pur non avendo a disposizione sensori sul corpo), che siano in grado di generare i *feedback* richiesti (secondo il principio della risonanza estetica) e di fornire ai terapisti gli strumenti necessari per valutare i progressi dei pazienti.

Invece per quanto concerne gli obiettivi clinici, il dispositivo, in grado di compensare la acinesia controllando le strutture ritmiche del movimento, permette ai pazienti di allenarsi a riconoscere e riprodurre tali strutture per aiutare a controllare obiettivi motori più complessi.

L'utilizzo di questi metodi produce un aumento di fluidità del movimento che può essere trasformato in stimoli visivi e/o acustici, che possono guidare la rappresentazione interna del controllo volontario del paziente.

L'implementazione

Le modalità

Il progetto si basa sull'utilizzo delle seguenti modalità:

- i movimenti del corpo (e la gestualità);
- l'udito.

L'hardware

La piattaforma comprende due tipologie di apparecchiature: quelle necessarie per acquisire i movimenti dell'utente e quelle che riproducono i *feedback* sonori.

Per quanto riguarda la prima parte di dispositivi, la piattaforma è decisamente versatile, perché il software *EyesWeb*, le cui caratteristiche verranno illustrate nel prossimo paragrafo, supporta moltissime tipologie di sensori e di videocamere. Anche per quanto riguarda il *feedback* sonoro non vengono richieste delle apparecchiature particolari: è sufficiente un semplice sistema di riproduzione.

Il software

I ricercatori hanno deciso di adottare il software *EyesWeb*², arricchendolo mediante dei moduli specifici che permettono l'analisi del movimento e la creazione di esercizi terapeutici. Tale programma gode di diverse caratteristiche interessanti: in primo luogo permette all'utente di interagire con modelli di calcolo di comunicazione espressiva non verbale e di mappare i gesti provenienti da diverse modalità in media sonori o video. Inoltre supporta l'integrazione di nuove tecniche di analisi e fornisce strumenti per creare, sviluppare e testare velocemente storie interattive ed esercizi

² Anch'esso sviluppato dal laboratorio InfoMus.

clinici. Infine è in grado di mostrare in tempo reale le misure fisiche associate ai movimenti e può essere programmato per eseguire analisi specifiche di singoli gesti.

Il processo

Il *feedback* multimediale, come descritto precedentemente, viene prodotto sulla base di determinati parametri motori che descrivono qualitativamente il movimento. In questa sezione si illustrerà il processo a livelli che permette di ottenere questi parametri a partire da alcune misure fisiche di basso livello.

La prima fase che viene eseguita è denominata *Motion Detection and Tracking* e consente (mediante l'utilizzo dell'algoritmo di *Lucas-Kanade*) di determinare la posizione e la traiettoria di alcune parti specifiche del corpo (come mani, piedi e testa) e di estrarre le immagini in movimento della *silhouette* (*SMI*).

Tali informazioni vengono poi rielaborate per determinare una collezione di caratteristiche motorie, chiamate *features*. Tra queste, quelle che rivestono maggiore importanza sono: l'indice di contrazione (*C.I.*) e la quantità di movimento (*QoM*).

Il *C.I.* equivale all'utilizzo dello spazio circostante da parte del paziente e permette, mediante campionatura dei dati, di valutare il tipo di movimento: contrazione o estensione.

La *QoM* corrisponde invece al rapporto tra l'area delle *SMI* e l'area della silhouette più recente e permette di suddividere il movimento in fasi di moto e fasi di quiete. Questo riconoscimento, conosciuto tecnicamente con il nome di *Motion Segmentation*, avviene analizzando i parametri caratteristici della curva di evoluzione nel tempo della *QoM*. Questa funzione può essere approssimata ad una sequenza di "campane" e alcuni dei parametri presi in considerazione sono il valore di picco e la durata. A partire dalle informazioni fin qui elaborate, mediante l'analisi dell'andamento temporale e della forma delle "campane", si estraggono due ulteriori parametri: la fluidità e l'impulsività del movimento.

Il processo termina con il trasferimento delle caratteristiche di alto livello, vale a dire la tipologia di movimento (contrazione/estensione) e la fluidità e impulsività, al modulo di produzione dei *feedback* che li trasformerà, sulla base del *mapping* desiderato, in suoni e/o immagini.

La sperimentazione clinica

Mediante la piattaforma *EyesWeb* sono stati sviluppati diversi esercizi per pazienti malati di Parkinson; una parte di questi sono stati testati presso l'ospedale La Colletta di Genova, in collaborazione con il Centro di Bioingegneria.

Un esempio interessante è stato realizzato utilizzando un grande schermo che viene dipinto (in tempo reale) in base ai movimenti eseguiti dal paziente.

Il *mapping* movimento-colorazione è basato su alcune delle caratteristiche descritte precedentemente. Per esempio il colore dipende dalla fluidità, l'intensità della traccia è correlato alla *QoM* e il *re-start* del processo (riassegnazione dei colori) viene eseguito in base alle pause nel movimento (si veda Figura 1).

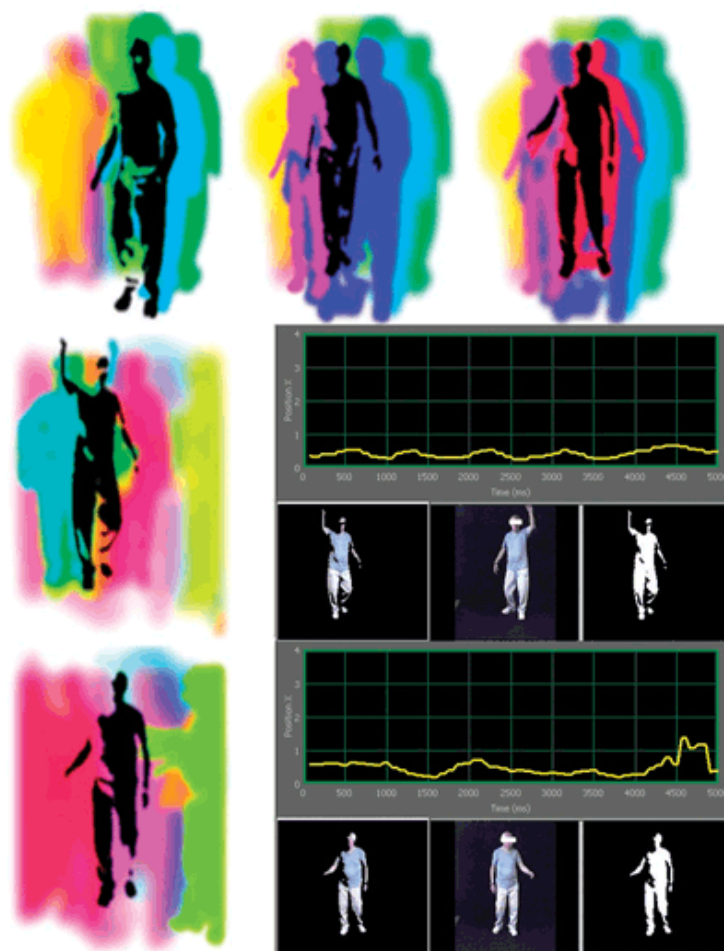


Figura 1. Mapping movimento-colorazione (Camurri et al., 2003, p. 274).

Gli obiettivi clinici sono stati raggiunti utilizzando un *mapping* di colori piacevoli associato a movimenti fluidi e senza esitazioni.

L'esperimento ha dimostrato che il *feedback* visivo incoraggia i pazienti a migliorare la qualità dei movimenti, fornendo pertanto risultati notevoli nell'ambito riabilitativo.

Inoltre durante l'esercizio i ricercatori hanno potuto sfruttare a pieno l'interfaccia monitorando i parametri motori ed eventualmente modificandoli per adattare lo strumento ai bisogni del paziente.

1.2.2 *Disabilità degli arti superiori e della mano: un'interfaccia per gestire l'interazione uomo-computer*

Il team

Nell'ambito della disabilità degli arti un progetto di particolare rilevanza è *ICANDO*, acronimo di *Intellectual Computer AssistaNt for Disabled Operators*. Questa interfaccia è stata realizzata da Alexey Karpov e Andrey Ronzhin (2007) dello *Speech and Multimodal Interfaces Laboratory*³ del *St. Petersburg Institute for Informatics and Automation of the Russian Academy of Sciences (SPIIRAS)*, in collaborazione con l'*University of West Bohemia (UWB)* e la *Tomsk State University of the Control Systems and Radio-electronics (TSUCSR)*.

Lo *Speech and Multimodal Interfaces Laboratory* è impegnato principalmente nella progettazione di sistemi di riconoscimento e comprensione automatica del parlato (per la lingua russa e non solo), nella creazione di interfacce in grado di instaurare conversazioni con l'uomo, nell'ambito della multimodalità, nei sistemi biometrici basati sul riconoscimento della voce e nella realizzazione di software per la traduzione automatica.

L'idea

ICANDO è una interfaccia grafica che gestisce l'interazione uomo-computer combinando il riconoscimento dei movimenti della testa, con la comprensione di comandi vocali in lingua inglese, francese e russa.

In particolare, l'utente interagisce con la piattaforma utilizzando contemporaneamente entrambi i canali di comunicazione: mediante lo spostamento della testa pilota il cursore del computer e tramite la voce impartisce le azioni da eseguire.

Luogo e anno di realizzazione

L'interfaccia è stata sviluppata presso i centri di ricerca di *SPIIRAS*, *UWB* e *TSUCSR*, che hanno sede, rispettivamente, a San Pietroburgo (Russia), Pilsen (Repubblica Ceca) e Tomsk (Russia).

La fase di test è stata eseguita dal team *SPIIRAS* in collaborazione con il *St. Petersburg Professional Rehabilitation Center (PRC)*, polo medico che si occupa principalmente della riabilitazione di persone con bisogni speciali.

³ Il gruppo di ricerca nacque nel 1984 con il nome *Speech Informatics Group*. Successivamente nel 2008 ha cambiato denominazione in *Speech and Multimodal Interfaces Laboratory of SPIIRAS*.

Per quanto concerne l'anno di realizzazione, la prima dimostrazione live dell'interfaccia è stata trasmessa nel 2005 dalla più importante emittente televisiva russa. Successivamente, nel 2006, il prototipo è stato insignito del premio *Best PC Multimodal User Interface Software* nel concorso *Low-cost Multimodal Interfaces Software Contest*, organizzato dalla rete di eccellenza *SIMILAR*. I primi risultati scientifici sono stati infine pubblicati nel 2007.

Gli obiettivi

Per poter descrivere le finalità del progetto in maniera più chiara, ho deciso di suddividerle in due parti: le motivazioni di contesto e quelle tecniche.

Motivazioni di contesto

In questa sezione vorrei mettere in risalto lo scopo del progetto e il ruolo che esso assume nel processo riabilitativo delle persone diversamente abili.

L'obiettivo dell'interfaccia è fornire alle persone senza mani (o con disabilità alle mani e/o braccia) uno strumento di interazione con il computer. Per compensare le abilità mancanti vengono utilizzate, come già detto in precedenza, due modalità: la voce e il movimento della testa. Questa soluzione permette di raggiungere dei risultati veramente notevoli facilitando, mediante l'ausilio della piattaforma web, anche la comunicazione tra più persone.

A riguardo invece del processo riabilitativo, la disabilità degli arti produce un'incapacità evidente, da parte dell'utente, di interagire con i dispositivi standard di *input* quali tastiera, mouse, touchpad, ecc. (si veda Figura 2).

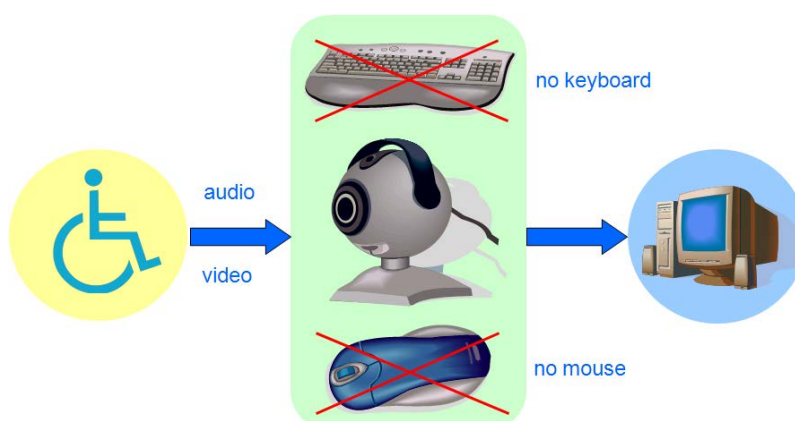


Figura 2. Lo scopo del progetto (Karpov & Ronzhin, 2007, p. 22).

Le conseguenze di questo fatto si ripercuotono nella diminuzione delle operazioni realizzabili con il computer e, in maniera più generale, sulle attività socio-economiche dell'individuo. La piattaforma in esame permette alle persone

diversamente abili di svolgere tutte le azioni principali con il computer, facilitando la loro uguale partecipazione nella società dell'informazione. Ma questo non è tutto: l'indipendenza raggiunta e il miglioramento dello stile di vita permette di ottenere, in molti casi, una carriera professionale di successo.

Motivazioni tecniche

Dal punto di vista tecnico si vuole creare un'interfaccia, sufficientemente accurata e robusta, in grado di funzionare senza utilizzare sensori di contatto. Questa scelta è strettamente connessa ai test eseguiti nelle prime fasi di progetto. Il *feedback* degli utenti nell'indossare un elmetto, dotato di diodi luminosi, si è rivelato essere estremamente negativo, soprattutto in relazione alla praticità e semplicità di utilizzo dello strumento. Si è preferito, pertanto, un approccio di tipo software, che facesse uso di una webcam non professionale in grado di catturare il segnale audio e video. La conseguente riduzione dei costi delle periferiche, ottenuta eliminando i sensori esterni, ha prodotto benefici rilevanti, facendo acquisire al progetto il titolo di interfaccia *low-cost*.

Dal punto di vista delle funzionalità, l'obiettivo principale è permettere, nel limite del possibile, che le persone diversamente abili possano svolgere tutte quelle azioni che un utente normodotato esegue quotidianamente con il computer.

L'implementazione

Le modalità

La piattaforma è stata progettata per combinare l'utilizzo di due modalità:

- il movimento della testa;
- la voce.

L'hardware

A garanzia del titolo di progetto *low-cost*, si è deciso di utilizzare un laptop "comune", dotato di processore Intel Core 2 Duo 1.66 Ghz (tramite il quale si è in grado di gestire l'elaborazione dei due flussi di informazione in parallelo), 1 GB di RAM, uno schermo da 17 pollici (con risoluzione 1280x1024 pixels) e una webcam non professionale con le caratteristiche tecniche riportate in Tabella 1.

Tabella 1. Caratteristiche tecniche webcam.

Nome prodotto	USB web-camera Logitech QuickCam for Notebooks Pro
Tipo webcam	Non professionale, <i>low-cost</i>
Segnali gestiti	Video e Audio (microfono integrato)
Risoluzione Video	640x480, 25fps
Campionamento Audio / SNR	16 KHz / Accettabile

Il software

Dal punto di vista applicativo, l'interfaccia è stata suddivisa in due moduli. Il riconoscimento vocale è stato affidato al software *SIRIUS (SPIIRAS Interface for Recognition and Integral Understanding of Speech)*, secondo il processo che verrà illustrato nei prossimi paragrafi. Per inciso, tale programma viene utilizzato in moltissime altre applicazioni, soprattutto in ambito neurochirurgico.

Il modulo di analisi dei movimenti della testa, invece, è stato implementato mediante la libreria *Intel Open-source Computer Vision Library OpenCV*, che permette di realizzare numerosi algoritmi per l'elaborazione di immagini e video.

Per quanto concerne le funzionalità, l'obiettivo illustrato in precedenza è stato raggiunto con successo. I comandi riconosciuti dall'interfaccia coprono tutte le azioni che si svolgono quotidianamente con un pc: l'interazione con il mouse (per es. il *click* con il tasto sinistro, con il destro, ecc.), la gestione della tastiera (per es. i tasti invio, esc, ecc.) e il controllo dell'interfaccia grafica (per es. i comandi apri, salva, copia, ecc.). In particolare è interessante sottolineare, ai fini della descrizione del processo, che solamente i comandi relativi alla gestione del mouse hanno una natura multimodale, vale a dire utilizzano entrambe le modalità. Invece sia l'interazione con la tastiera, sia quella con l'interfaccia utilizzano solo il riconoscimento vocale.

Il processo

Dal punto di vista procedurale, i due moduli di riconoscimento lavorano in parallelo, producendo dei risultati che vengono poi combinati mediante le fasi di sincronizzazione e fusione (si veda Figura 3).

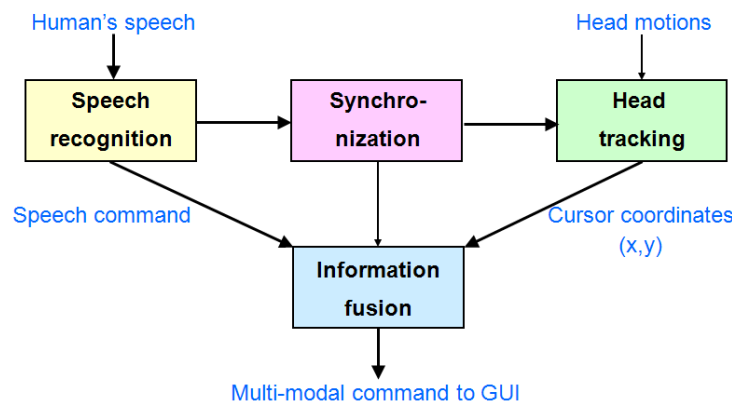


Figura 3. Il processo (Cadiou, Karpov, & Ronzhin, 2006, p. 3).

Riconoscimento automatico della voce

L'elaborazione eseguita dal modulo di riconoscimento della voce è composta da tre fasi: l'estrazione delle *features*, il riconoscimento dei fonemi e, infine, l'analisi morfemica.

La prima fase raggiunge il proprio obiettivo campionando il segnale audio registrato ad una frequenza di 16 KHz e combinando i campioni ottenuti in segmenti. In particolare, per eseguire la cosiddetta "parametrizzazione del parlato", si utilizzano i *Mel Frequency Cepstral Coefficients (MFCC)*, avvalendosi della derivata prima e seconda.

Il riconoscimento dei fonemi e delle parole del vocabolario avviene mediante delle *Hidden Markov Models* di trifoni⁴, miste a funzioni di densità di probabilità gaussiana.

L'analisi morfemica, infine, viene gestita con la medesima tecnica di riconoscimento dei fonemi e permette, riducendo la taglia del vocabolario, di individuare le parole in maniera più accurata e veloce. Quest'ultima fase è stata introdotta in un secondo momento, rispetto alle precedenti, per poter gestire in maniera più efficiente il riconoscimento della lingua russa, una tra le più flessive al mondo.

Riconoscimento dei movimenti della testa

Il modulo di analisi e riconoscimento dei movimenti della testa elabora le immagini registrate dalla webcam, ne determina le variazioni e, sulla base di queste, fornisce al computer le coordinate del cursore.

L'elaborazione prevede due fasi: la calibrazione e il riconoscimento vero e proprio (*tracking*). Nella prima fase si utilizza un rivelatore, chiamato *Haar Based Object*

⁴ Con il termine trifone si intende il fatto che ad ogni fonema vengono associati tre stati.

Detector, in grado di individuare la regione rettangolare che contiene il volto umano d'interesse. Successivamente la fase di *tracking* determina, all'interno del rettangolo, cinque riferimenti naturali: il centro del labbro superiore (1), la punta del naso (2), il punto di mezzo tra le due sopracciglia (3), l'occhio sinistro (4) e quello destro (5). Questi punti naturali, non a caso, formano due rette perpendicolari: rispettivamente una verticale (1-2-3) e una orizzontale (4-3-5) (si veda Figura 4).



Figura 4. I cinque riferimenti naturali (Karpov & Ronzhin, 2007, p. 24).

Per eseguire questo riconoscimento viene utilizzato l'algoritmo *Iterative Lucas-Kanade*, corredato da una funzione di ripristino automatico che, in caso di perdita, recupera i punti di riferimento⁵ utilizzando le due rette perpendicolari.

Una volta terminata l'identificazione dei punti, la coordinata del cursore sull'asse delle ascisse viene calcolata in questo modo: si valutano le differenze di posizione, tra l'ultimo frame e il precedente, di ogni punto appartenente alla retta orizzontale e si calcola la media dei risultati ottenuti. Per quanto riguarda l'asse delle ordinate si procede in maniera simmetrica, utilizzando, naturalmente, i punti della retta verticale.

La procedura appena descritta costituisce il nucleo centrale dell'elaborazione, tuttavia, durante lo sviluppo e il test dell'interfaccia, si è dovuto far fronte a due ulteriori difficoltà. La risoluzione ridotta della webcam, in relazione a quella dello schermo utilizzato, ha reso necessario l'introduzione di alcuni coefficienti moltiplicativi per poter "ri-mappare" le coordinate all'interno di uno spazio decisamente più grande (si veda Figura 5).

⁵ Mi riferisco ai cinque riferimenti naturali precedentemente elencati.

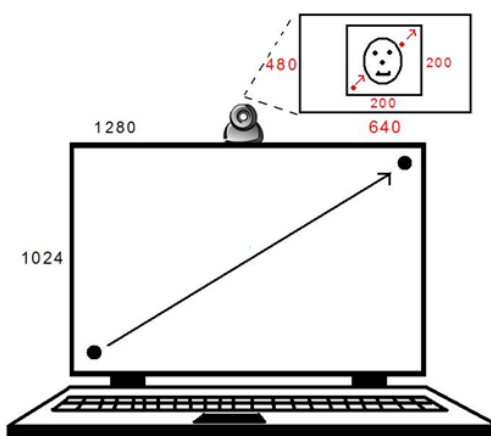


Figura 5. Risoluzione del monitor e della webcam (Karpov & Ronzhin, 2007, p. 25).

Inoltre l'instabilità del cursore, nel selezionare oggetti di piccola dimensione, ha comportato l'introduzione di altri fattori di correzione, in grado di diminuire la velocità del puntatore (e aumentare la precisione) nel caso in cui i movimenti della testa risultino più lenti.

Modalità di sincronizzazione e fusione delle informazioni

Come delineato precedentemente, i flussi d'informazione provenienti dai due moduli vengono combinati durante le fasi di sincronizzazione e di fusione.

Prima di descriverle in dettaglio, è necessaria una premessa. Le informazioni indispensabili per portare a compimento un determinato comando sono: la descrizione dell'azione che deve essere eseguita, le coordinate del cursore (che identificano dove operare) e la sua natura (unimodale o multimodale). Il modulo di riconoscimento vocale si occupa di determinare (e memorizzare) la descrizione dell'azione. La natura del comando è un'informazione conosciuta a priori dal programma: come descritto precedentemente i comandi associati al mouse sono multimodali. Le coordinate del cursore vengono determinate dal modulo di riconoscimento dei movimenti della testa, secondo il processo descritto precedentemente, tuttavia la loro memorizzazione avviene in momenti specifici, per evitare problemi legati a movimenti involontari dell'utente.

In questo contesto, la fase di sincronizzazione ha il compito, appunto, di salvare le coordinate un istante prima che l'utente pronunci un comando vocale. In questo modo le intenzioni effettive dell'utente vengono rispettate.

La fase di fusione, d'altra parte, riunisce le informazioni provenienti dai diversi moduli in un unico messaggio⁶ e le invia alla tastiera o al mouse, in base alla

⁶ Il termine viene utilizzato, non a caso, per sottolineare il fatto che si adoperano i messaggi standard di Windows.

pertinenza. La tecnica che viene utilizzata per combinare le informazioni è denominata metodo dei frame e prevede l'esecuzione di un comando solo quando tutte le informazioni⁷ sono disponibili.

Appena si verifica questa condizione, il messaggio viene inviato ad un destinatario differente a seconda della natura del comando. Se è di tipo multimodale viene inoltrato al mouse. Se invece è unimodale, le coordinate non vengono prese in considerazione e il messaggio viene spedito alla tastiera.

Un'ultima considerazione va riportata a riguardo della gestione dei messaggi (fase di *controlling*) e del *feedback* restituito all'utente. Per quanto riguarda il *controlling* si utilizza il meccanismo standard delle code di messaggi in Microsoft Windows. Invece, per quanto concerne il *feedback*, al completamento di ogni comando l'utente riceve una risposta multimediale sotto forma di musica, voce o video.

La sperimentazione clinica

Le persone coinvolte nei test sono state in totale sei: cinque normodotate (con poca esperienza nell'utilizzo del computer) e una con disabilità fisica conseguente all'assenza delle mani (proveniente dal *St. Petersburg Professional Rehabilitation Center*) (si veda Figura 6).



Figura 6. Test dell'interfaccia da parte di un disabile senza mani (Karpov & Ronzhin, 2007, p. 27).

Per valutare la bontà della piattaforma, gli utenti sono stati sottoposti ad una prova che prevedeva la ricerca di un sito internet, la selezione e la copia di una parte di testo in Microsoft Word e infine il salvataggio e la stampa del relativo file.

I parametri di riferimento sono stati: la misura del tempo di esecuzione e l'accuratezza del riconoscimento dei comandi vocali.

⁷ Si fa riferimento ai tre dati, sopra riportati, che sono necessari ad identificare un comando.

Per quanto riguarda il primo aspetto, le operazioni richieste sono state completate, dalle persone normodotate, impiegando mediamente il doppio del tempo rispetto ad un'interazione tradizionale (con mouse e tastiera). Dal momento che l'interfaccia è rivolta a persone con disabilità alle mani, che non possono interagire con il computer in nessun modo, il ritardo di esecuzione si può considerare più che accettabile.

Il secondo parametro, vale a dire il grado di precisione nel riconoscimento dei comandi vocali, si è rivelato essere estremamente elevato raggiungendo, mediante la modalità *speaker-dependent*⁸, delle percentuali di accuratezza oltre il 96%. Inoltre è stato possibile avanzare due ulteriori considerazioni a riguardo: in primo luogo si è osservato che la frequenza di utilizzo dei comandi non è regolare e rispecchia, quasi completamente, ciò che accade nell'interazione tradizionale (ad esempio il comando *left*, vale a dire l'equivalente del tasto sinistro del mouse, viene utilizzato più frequentemente degli altri).

Inoltre si è riscontrato l'utilizzo da parte degli utenti di alcuni comandi *out-of-vocabulary*: si tratta di una percentuale ridotta di comandi (8%) con semantica simile a quelli corretti (es. *back* al posto di *previous*). Per poter rendere l'interfaccia maggiormente *user-friendly* si è deciso di includere anche questi nel vocabolario dell'interfaccia.

I risultati ottenuti, in definitiva, hanno mostrato che *ICANDO* è un'ottima piattaforma utilizzabile, con successo, da persone con disabilità alle mani.

1.2.3 *Disabilità della vista: uno strumento per migliorare l'accesso alle informazioni web*

Il team

L'interfaccia nasce dalla collaborazione congiunta dei gruppi di ricerca *Sonic Arts Research Centre (SARC)* e *Virtual Engineering Centre (VEC)*, entrambi afferenti alla *Queen's University of Belfast (QUB)*. In particolare i ricercatori coinvolti sono stati R. Kuber, G. McAllister, E. Murphy, P. Strain e W. Yu (2005).

Per quanto concerne gli argomenti di ricerca, il *Sonic Arts Research Centre* è impegnato in moltissime aree, tra cui la creazione di strumenti acustici-virtuali, la gestione delle *performance*⁹, la progettazione di sistemi interattivi per la composizione musicale, la *Sonic Art* e il *rendering* spaziale del suono.

⁸ L'interfaccia prevede una fase preliminare di addestramento per ogni singolo utente, al termine della quale viene creato un profilo personalizzato.

⁹ Con il termine *gestione delle performance* si intende tutti gli aspetti connessi ad essa tra cui: l'utilizzo di nuove tecnologie, la creazione di nuovi sistemi, lo studio critico e l'analisi delle pratiche contemporanee, ecc..

Il *Virtual Engineering Centre*, invece, focalizza l'attenzione su tre macro-aree principali: la multimodalità, la realtà virtuale e il miglioramento di processi tecnici. In maniera più precisa la prima area riguarda l'accessibilità web mediante approcci di tipo multimodale¹⁰. Il secondo filone affronta l'ambito della realtà virtuale sia analizzando gli aspetti tecnici, sia progettando alcune implementazioni vere e proprie. Dal punto di vista tecnico si occupa, tra i tanti argomenti, della gestione delle comunicazioni nelle applicazioni distribuite per realtà virtuali. Invece per quanto riguarda le applicazioni propone delle soluzioni interessanti per la riproduzione dell'urbanistica delle città o per la simulazione delle operazioni chirurgiche. Infine, il terzo ambito cerca di rispondere, in maniera adeguata, alle difficoltà tecniche riscontrate in moltissimi settori: per esempio identificare (e gestire) i problemi dei motori durante la fase di combustione, evitare i danni relativi all'ambiente causati dai propulsori ad elica delle barche, scegliere i materiali corretti per costruire apparecchiature aerospaziali e moltissimi altri ancora.

L'idea

Il progetto prevede la realizzazione di un'interfaccia multimodale che sia in grado di rispondere, in modo soddisfacente, alle difficoltà riscontrate dai disabili della vista: in particolare, deve poter garantire la navigazione web e la lettura delle pagine visitate.

Luogo e anno di realizzazione

Il prototipo del sistema è stato sviluppato presso i laboratori della *QUB* in Irlanda del Nord nel 2005; nello stesso anno sono stati pubblicati i primi risultati scientifici.

Gli obiettivi

Le finalità del progetto possono essere suddivise in due sezioni: le motivazioni cliniche e quelle tecniche.

Motivazioni cliniche

La quantità di informazioni reperibili e i numerosi servizi offerti hanno reso il mondo web una piattaforma indispensabile nella società moderna. Tuttavia, una delle limitazioni principali di questo strumento rimane la sua accessibilità alle persone disabili (per esempio persone con ridotte capacità visive): il progetto, come già delineato in precedenza, cerca proprio di rispondere a questo tipo di problematica.

¹⁰ In questo ambito il laboratorio si occupa di numerosi progetti: uno dei più interessanti è sicuramente *ENABLED* (*Enhanced Network Accessibility for the Blind and Visually Impaired*) che possiede finalità simili al progetto illustrato in questo elaborato.

Le principali sfide da affrontare sono l'accessibilità delle pagine web e le limitazioni che riguardano gli attuali strumenti di assistenza.

Per quanto riguarda il primo ambito, esistono delle vere e proprie linee guida per i web designer, chiamate *W3C's Web Content Accessibility Guideline*. Tuttavia queste misure risultano essere insufficienti a garantire l'accessibilità vera e propria e, inoltre, sono pochi i siti internet che vengono costruiti secondo questi canoni. Gli studi Di Blas e di Brajnik del 2004 mettono in risalto, in maniera evidente, l'inadeguatezza delle *guidelines*. Invece i test eseguiti, nello stesso anno, dalla *Disability Rights Commission* confermano la seconda affermazione, affermando addirittura che quasi tutti i siti internet pubblicati non rispettano gli standard previsti. Ma questo non è tutto: pur analizzando i siti "meglio costruiti", si rilevano moltissime difficoltà di interazione sia per coloro soggetti a ridotte capacità visive, sia per le persone normodotate. Solo il 53% dei disabili della vista riesce a raggiungere gli obiettivi previsti nei test; la percentuale cresce al 76% per i normodotati, ma rimane comunque sorprendentemente bassa. I problemi riscontrati, a differenza di quanto ci aspetteremmo, sono dovuti solo nel 50% dei casi al mancato rispetto delle linee guida.

Per quanto concerne, invece, le tecnologie di assistenza attuali, lo strumento maggiormente utilizzato dai disabili della vista è sicuramente lo *screen reader*, un dispositivo in grado di descrivere i contenuti delle pagine web. Questo sistema, purtroppo, ha due grandi limitazioni: non è in grado di "illustrare" le immagini (se non quando esistono dei testi alternativi appropriati) e allunga i tempi di navigazione, perché legge in ordine sequenziale l'intero testo della pagina.

Nel corso del tempo molti ricercatori hanno cercato di trovare soluzioni a queste problematiche: per esempio, sono stati creati degli strumenti in grado di riassumere i contenuti dei siti internet (per es. *WebSpeak* e *BrookesTalk*) oppure dei *browser* audio che eseguono il *rendering 3D* del suono (le interfacce di *Goose* e *Donker*). Inoltre sono nati i cosiddetti *Sonified Hyperlinks* (il ricercatore Susini fu uno dei primi sostenitori), alcune interfacce basate sull'interazione sonora e tattile (il progetto *WebSound*) oppure solo su quella tattile (il lavoro del ricercatore Parente).

Il progetto in esame vuole fornire una risposta concreta ai problemi precedentemente trattati e costruire un'interfaccia che, facendo tesoro delle ottime idee avanzate in passato, sfrutti a pieno la multimodalità per rispondere, in modo soddisfacente, alle problematiche dei disabili della vista.

Motivazioni tecniche

In questo paragrafo si vuole fornire al lettore una breve panoramica delle motivazioni che hanno portato i ricercatori ad eseguire determinate scelte tecniche. La descrizione dell'implementazione vera e propria verrà illustrata nel prossimo paragrafo.

Il team di studiosi si è impegnato, in un primo momento, ad identificare le modalità che potessero sostituire e compensare la vista in maniera più proficua. Contemporaneamente, ha cercato di comprendere quali sono le problematiche attuali (tecniche e non) dei disabili e come possono essere risolte sfruttando le modalità alternative.

Il primo aspetto, vale a dire la scelta delle modalità, è stato affrontato analizzando alcuni studi di settore, in cui si mettevano in evidenza i benefici che si possono ottenere dall'utilizzo dei diversi sensi. In particolare queste ricerche raggiungevano una conclusione interessante: il tatto era la scelta più vantaggiosa, grazie alla sua capacità di estendere il panorama visivo rendendo gli oggetti più realistici (teoria di Brewster), di fornire una mappa mentale degli oggetti presenti in un ambiente (teoria di Lahav) e di migliorare l'orientamento dell'utente durante l'esplorazione dello spazio (teorie di Caffrey). Sulla base di quanto descritto, i ricercatori hanno focalizzato l'attenzione, pertanto, sull'utilizzo di questo senso. Inoltre, basandosi sui buoni risultati ottenuti dal progetto *WebSound*, già citato nel paragrafo precedente, hanno deciso di integrare l'utilizzo dell'udito, sfruttando i benefici ottenibili da un'interazione di tipo multimodale.

Per quanto concerne le problematiche, invece, si è scelto di utilizzare un approccio di tipo *user-centered*, vale a dire utilizzare dei test per individuare le principali difficoltà dei disabili della vista. Le problematiche riscontrate sono state diverse a seconda del livello di esperienza degli utenti. I “principianti” hanno mostrato una buona propensione a svolgere alcune attività specifiche, come per es. leggere la posta, gli orari dell'autobus, ecc.. Invece si sono rifiutati di eseguirne delle altre, tra cui gli acquisti online o il cosiddetto *surfing* nel web (a causa della difficoltà di interazione e della paura di sbagliare). Le difficoltà di navigazione osservate sono state di vario genere: l'utente si annoiava perché il tempo necessario per leggere una pagina era troppo elevato¹¹, perdeva l'orientamento durante la navigazione, non riusciva ad utilizzare più finestre contemporaneamente e aveva paura di eseguire download (paura dei virus, di alterare il sistema, ecc.).

¹¹ In molti casi veniva eseguita una lettura sequenziale di tutta la pagina.

Gli utenti “intermedi”, invece, si sono dimostrati molto abili nell'utilizzare alcuni servizi (tra cui gli acquisti online, i servizi bancari, la chat, ecc.), nel gestire lo *screen reader* e nell'accedere velocemente alle informazioni (utilizzando strumenti che riassumono il contenuto delle pagine). I problemi riscontrati sono stati di tipo “tecnico” e hanno riguardato l'utilizzo di flash player, di javascript (che causava il *refresh* della pagina e la conseguente perdita della posizione dello *screen reader*), la gestione delle immagini (testi alternativi mancanti o troppo lunghi), i *files embedded* (per es. i collegamenti a files PDF) e l'inserimento di dati all'interno di maschere. In generale tutti i soggetti con disabilità coinvolti nell'esperimento, a prescindere dalla loro esperienza di navigazione e dalle capacità visive più o meno ridotte, hanno identificato tre tipologie di “supporti” potenzialmente utili: ottenere maggiori informazioni (*feedback*) per comprendere il layout della pagina, essere aiutati da parte di persone vedenti e partecipare a dei training specifici.

A partire da questi risultati, i ricercatori hanno potuto identificare le principali esigenze dei disabili e gli aspetti tecnici da migliorare negli strumenti attuali. Le esigenze sono state riassunte in questo modo: gestire le problematiche degli utenti in maniera personalizzata¹² e garantire la compatibilità con gli *screen readers*. Per quanto riguarda gli aspetti da migliorare, invece, si è focalizzato l'attenzione sulla creazione di alcune funzionalità che permettessero di fornire un'*overview* generale della pagina (prima della lettura completa), di gestire una memoria a breve termine (per poter tornare indietro di alcune righe) e, infine, di velocizzare il raggiungimento delle sezioni di interesse.

L'implementazione

Le modalità

Le modalità utilizzate in questo progetto sono:

- l'udito;
- il tatto.

L'hardware

L'utente interagisce con l'interfaccia utilizzando il *Logitech Wingman Force Feedback mouse* (si veda Figura 7): si tratta di un dispositivo, compatibile con il browser Firefox, che ha la capacità, appunto, di produrre il cosiddetto *force feedback*. Questa particolare tipologia di risposta multimediale, utilizzata per esempio per

¹² Vale a dire tenendo in considerazione l'esperienza di navigazione e le capacità visive più o meno ridotte dell'utente.

simulare gli incidenti nei videogames, prevede l'esecuzione di vibrazioni sul dispositivo che vengono percepite dall'utente come degli "impedimenti" a svolgere un determinato movimento.



Figura 7. Logitech Wingman Force Feedback mouse (Kuber et al., 2005, p. 5).

Il software

La piattaforma è composta da tre componenti principali: il *content-aware web browser plug-in*, il modulo di gestione del *feedback* audio e quello relativo al *feedback* tattile (si veda Figura 8).

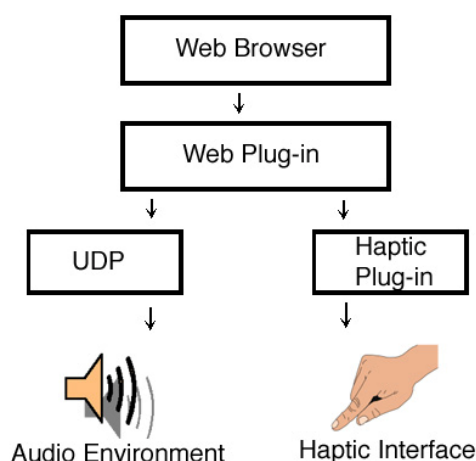


Figura 8. L'architettura di sistema (Kuber et al., 2005, p. 5).

Il *content-aware Web browser plug-in* è un modulo complementare a Mozilla Firefox: la scelta di questo *browser* non è casuale, bensì è legata alla natura *open-source* del prodotto e al fatto che rispetta gli standard W3C delle *Web Content Accessibility Guidelines 1.0*, di cui abbiamo già avuto modo di parlare in precedenza.

La riproduzione del *feedback* audio è stata gestita, invece, mediante l'ambiente (real-time) di programmazione audio *MAX/MSP*.

Infine la gestione dell'interazione tattile viene eseguita da un ulteriore plug-in, che ha la funzione di collegare il *content-aware Web browser plug-in* al *Logitech Wingman Force Feedback mouse*.

Il processo

Dal punto di vista operativo, l'interazione web viene gestita in maniera molto semplice: l'utente esplora le pagine web utilizzando il *Logitech Wingman Force Feedback mouse*. In modo assolutamente trasparente la posizione del cursore viene monitorata dal *plug-in*¹³, il quale si occupa di richiamare i moduli di produzione del *feedback* (audio e tattile) solo nel caso in cui venga rilevato un oggetto vicino alla posizione corrente.

La piattaforma, oltre alla modalità di utilizzo appena descritta, può trasformarsi, inoltre, in uno strumento d'apprendimento: l'utente viene guidato verso una certa destinazione e, durante il tragitto, deve compiere determinate azioni per poter proseguire.

Per quanto concerne la parte tecnica, si riportano di seguito i dettagli legati alla gestione delle due tipologie di *feedback*.

Il *feedback* tattile

Il mouse in dotazione è in grado di riprodurre, mediante il *force feedback*, tre primitive tattili: l'effetto *enclosure*¹⁴, *buzz*¹⁵ e *periodic*.

Il primo effetto viene utilizzato per aiutare l'utente a comprendere dove finisce il bordo delle immagini; questo *feedback* funziona molto bene se combinato con un ulteriore effetto che limita i movimenti del mouse quando ci si avvicina alla cornice di un'immagine.

L'effetto *buzz*, invece, realizzato in combinazione ad un *feedback* audio, fornisce all'utente una maggiore consapevolezza di trovarsi "sopra un'immagine".

Infine il *feedback periodic* viene utilizzato per identificare i collegamenti ipertestuali e aumenta il senso di orientamento degli utenti nel riconoscere la posizione relativa del cursore.

Il *feedback* audio

Dal punto di vista tecnico, l'ambiente di programmazione audio *MAX/MSP* si occupa di riprodurre il *feedback* audio sulla base delle coordinate del cursore¹⁶ ricevute, via *UDP*, dal *plug-in*.

¹³ Si intende il content-aware Web browser plug-in.

¹⁴ Il termine significa letteralmente recinzione.

¹⁵ Il significato letterale del termine è ronzio, brusio. Tuttavia il significato più appropriato in questo contesto è "effetto scalpore".

¹⁶ Le coordinate, relative all'asse x e y, vengono inviate sotto forma di oggetto esterno MSP.

Come per il *feedback* tattile, lo scopo delle risposte multimediali sonore è aumentare la consapevolezza dell'utente di essere in contatto con immagini o collegamenti ipertestuali.

Per fare questo *MAX/MSP* fornisce dei *feedback* diversi all'utente a seconda della posizione del cursore (si veda Figura 9). Quando l'utente passa al di sopra di un'immagine o di un collegamento ipertestuale vengono riprodotti degli *auditory icons*, che corrispondono rispettivamente al *click* di una fotocamera e ad un tintinnio metallico (quello che si percepisce quando un anello di una catena batte contro un vicino).

Invece, quando l'utente si trova al di fuori dalle aree di interesse viene riprodotto un sottofondo continuo (chiamato *earcon*), le cui caratteristiche audio variano a seconda della direzione del movimento. Se viene eseguito uno spostamento sull'asse delle ascisse ne segue una variazione del *panning*¹⁷, mentre sull'asse delle ordinate del *pitch*¹⁸.

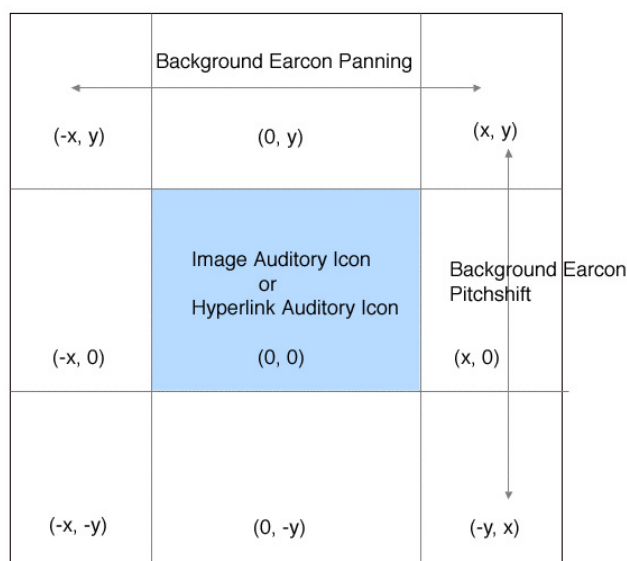


Figura 9. Feedback audio in relazione alla posizione del cursore (Kuber et al., 2005, p. 6).

La sperimentazione clinica

I risultati sperimentali hanno messo in evidenza due aspetti interessanti concernenti la riproduzione dei *feedback* audio e tattile.

L'utilizzo di queste due tipologie di risposte multimediali, in riferimento alle immagini e ai collegamenti ipertestuali, ha evidenziato due grandi benefici: l'utente è in grado di identificare correttamente la posizione delle regioni di interesse e, inoltre,

¹⁷ Con il termine *panning* si intende lo spostamento del segnale audio (nel tempo) su diversi canali di uscita.

¹⁸ Variare il *pitch* significa rendere il suono più o meno acuto.

riesce ad avere maggiore consapevolezza dello spazio esplorato (si parla anche di costruzione di una mappa spaziale mentale).

Come secondo aspetto, si è osservata una certa difficoltà (degli utenti) nel gestire più suoni contemporaneamente e nell'ascoltare riproduzioni audio di lunga durata. In termini tecnici si è arrivati alla conclusione che gli *auditory space* si sovraccaricano più velocemente dei *visual space*¹⁹ e, quindi, è consigliabile utilizzare pochi suoni di durata contenuta.

Il futuro

Nel futuro le principali aree di esplorazione riguarderanno la riduzione del carico cognitivo che viene sostenuto dall'utente nel creare la mappa mentale, la compatibilità con altri *browser*, la scelta di *feedback* più espressivi, la possibilità di riassumere il contenuto delle pagine e, infine, l'integrazione con lo *screen reader*, al fine di condurre automaticamente l'utente nelle sezioni che reputa di maggiore interesse.

1.2.4 Malati di Parkinson, quadriplegici, persone con ridotte capacità visive, persone senza braccia e anziani: un'interfaccia per facilitare l'accessibilità al World Wide Web

Il team

Nel settore della *web accessibility* è interessante soffermare l'attenzione su un progetto presentato dai ricercatori M.L. González González, A. Muñoz Muñoz e M.A. Valero Duboy (2008) del Dipartimento di Ingegneria e Architettura Telematica (DIATEL) dell'Università Politecnica di Madrid.

Il dipartimento in questione è impegnato in molteplici aree di ricerca tra cui la sicurezza delle reti, i sistemi di telecomunicazione per ambienti industriali, i servizi interattivi multimediali e le tecnologie *WEB/WAP*, la domotica e le piattaforme e-learning.

¹⁹ I due termini identificano rispettivamente gli spazi che si vengono a creare mediante l'utilizzo esclusivo di *feedback* audio e gli spazi tradizionali visivi (quelli che siamo abituati a vedere, per esempio, sullo schermo di un computer).

L'idea

La soluzione proposta è un'interfaccia multimodale *barrier-free*²⁰ a basso costo, in grado di semplificare la navigazione web ai malati di Parkinson, quadriplegici, persone con ridotte capacità visive, persone senza braccia e anziani.

La caratteristica peculiare della piattaforma è sicuramente l'adattabilità. Questa si manifesta, in primo luogo, nelle modalità di costruzione dell'interfaccia stessa, nata sulle basi di una piattaforma pre-esistente. D'altro canto, risulta essere il termine più appropriato per descrivere il processo che ha permesso ad un unico strumento di gestire le esigenze specifiche di molte disabilità.

Luogo e anno di realizzazione

Il progetto è stato sviluppato presso i laboratori del Dipartimento di Ingegneria e Architettura Telematica del Politecnico di Madrid nel 2008.

Dopo una prima fase sperimentale, svoltasi presso i laboratori, il prototipo è stato testato da un gruppo eterogeneo di persone con la partecipazione congiunta di alcuni importanti istituti: la *Parkinson Association in Madrid*, la *Federation of People with Intellectual Disability* e il *National Centre for Technical Aids and Personal Autonomy (CEAPAT)*.

Gli obiettivi

Le motivazioni che hanno spinto il gruppo di ricercatori ad intraprendere questo progetto sono di diversa natura. Come per i precedenti capitoli, si procede ad illustrarle suddividendole per categorie.

Motivazioni di contesto

Nel corso degli ultimi decenni si è assistito alla formazione della cosiddetta "società dell'informazione", che ha portato con sé lo sviluppo di sistemi e servizi "rivolti a tutti" e fruibili attraverso la piattaforma web. Oltre al "semplice" accesso e reperimento delle informazioni, si sono sviluppati il commercio elettronico, i servizi governativi e per la salute del cittadino, la teledidattica, l'intrattenimento e molti altri ancora.

Tali servizi, pur essendo estremamente eterogenei tra loro, possiedono una caratteristica comune: gestiscono l'interazione dell'utente solo mediante alcuni

²⁰ Il concetto di *barrier-free design* nasce nel campo dell'architettura e può essere considerato come il precursore dell'attuale paradigma di *universal design*. Il primo termine significa rimuovere le eventuali barriere architettoniche al fine di rendere un' ambiente fruibile da parte di persone diversamente abili. Il secondo invece propone di progettare e costruire gli ambienti in modo che possano adattarsi sia a persone disabili sia al resto della popolazione. Nell'ambito della computer science un'interfaccia *barrier-free* si pone l'obiettivo di rimuovere le eventuali barriere (o impedimenti) al fine di rendere gli strumenti attuali fruibili da parte di persone diversamente abili.

dispositivi di *input*. Nella maggior parte dei casi si parla solamente di tastiera e mouse.

Proprio quest'ultimo, negli ultimi anni, ha assunto un ruolo preponderante sugli altri, semplificando alcune operazioni di uso quotidiano come il movimento, la selezione e il *click*, ma, allo stesso tempo, producendo delle difficoltà enormi agli individui soggetti a problemi di vista, motori o cognitivi.

Le conseguenze di questo fatto si sono ripercosse pesantemente sulla vita sociale e lavorativa delle persone con disabilità, aumentando quello che viene chiamato il *digital gap*, vale a dire la difficoltà ad operare con strumenti digitali, misurata in termini di capacità e velocità di utilizzo rispetto a persone normodotate.

In questo contesto il progetto in esame, sfruttando l'interazione multimodale, cerca di oltrepassare i limiti dettati dall'utilizzo di un unico canale di comunicazione e di colmare il *digital gap*, rendendo l'interazione uomo-macchina un'esperienza più soddisfacente.

Un'ultima considerazione riguarda, infine, la volontà di creare dei servizi "rivolti a tutti": il motto *design for all*, oltre ad essere uno slogan promosso dalla *European Policy Centre*, diventa una finalità propria del progetto che si traduce nel soddisfare le richieste del mercato in termini di fruibilità, accessibilità ed economicità del prodotto.

Motivazioni teoriche

Oltre alle finalità precedentemente analizzate, è interessante comprendere in dettaglio le motivazioni teoriche che stanno alla base del concetto di accessibilità. Secondo la prima definizione ISO/TC 16027, con questo termine si intendeva l'utilizzabilità di un prodotto, servizio, ambiente o interfaccia da parte di persone normodotate²¹. In particolare si mettevano in risalto due aspetti importanti: per quanto riguarda l'azione in sé si rimandava al concetto di utilizzabilità mentre, in relazione al pubblico, si affermava chiaramente che l'accessibilità deve essere misurata nei confronti di persone normodotate, senza in alcun modo preoccuparsi di quelle con disabilità.

Procedendo quindi con l'analisi, cerchiamo di approfondire questi due punti chiave. Il termine utilizzabilità, nella *ISO 9241-11 Guidance on Usability* del 1998, è stato definito come "la misura entro cui un prodotto può essere utilizzato da utenti specifici, all'interno di un contesto particolare di utilizzo, per raggiungere obiettivi precisi in maniera efficace, efficiente e portando soddisfazione all'utente".

²¹ La definizione originale riporta "l'utilizzabilità di un prodotto, servizio, ambiente o interfaccia da parte di persone che possiedono un ampio numero di capacità", tuttavia credo che il termine utilizzato colga il concetto in maniera più concisa.

Per quanto riguarda invece il pubblico, la questione è resa più complessa dal fatto che la definizione di disabilità è cambiata notevolmente a partire dagli anni '80 in poi. In particolare alcune pubblicazioni importanti che hanno portato a maturare nuove consapevolezze a riguardo sono: *The International Classification of Deficiencies, Disabilities and Handicaps (ICDDH)* del 1980 e *The International Classification of Functioning, Disability and Health (ICF)* del 2001. Analizzare in dettaglio questi documenti sarebbe molto interessante, ma si rischierebbe di andare ben oltre lo scopo di questo elaborato. In generale l'obiettivo attuale è quello di gestire l'accessibilità a 360°, cercando di adattare sempre di più gli strumenti alle esigenze dei diversamente abili.

L'interfaccia in esame cerca di far proprie le idee precedentemente esposte fornendo, alle persone normodotate e a quelle diversamente abili, uno strumento potente in grado di rendere la navigazione web un'esperienza soddisfacente.

Motivazioni hardware e software

Lo scopo dei ricercatori è progettare un unico dispositivo hardware a basso costo, corredato da un software di interfacciamento con il computer che non richieda alte capacità di calcolo.

Ai fini di rendere il progetto effettivamente *low-cost*, è preferibile che sia l'hardware, sia il software siano già reperibili sul mercato e possano essere adattati "facilmente" a seconda delle esigenze.

Se possibile, è opportuno limitare i problemi spesso ricorrenti nelle applicazioni multimodali tra cui i costi relativi all'acquisto e al mantenimento di più dispositivi (e modalità), la gestione di diversi protocolli di rete, gli approcci elettronici specifici, l'eterogeneità degli utenti, l'elevato numero di dispositivi necessari e l'aumento di complessità del sistema.

Per quanto concerne le funzionalità, il dispositivo deve poter fornire procedure di puntamento facili (muovere il cursore e selezionare un oggetto con un *click*), un'interazione basata sull'accelerazione, deve facilitare l'inserimento di testo, utilizzare dei protocolli di comunicazione possibilmente wireless, fornire un *feedback* a vibrazione, avere delle batterie autonome e dei tasti configurabili.

Particolare attenzione va rivolta proprio alla configurazione del dispositivo che contribuisce, in buona parte, ad aumentare la flessibilità dell'intero sistema. Solo mediante questa caratteristica è possibile rispondere ad un ampio spettro di disabilità, fisiche e cognitive, e gestire le esigenze specifiche di ciascuna tipologia di utente.

Motivazioni cliniche

Come già delineato nei paragrafi precedenti, l'interfaccia vuole rispondere alle esigenze specifiche di quattro tipologie di disabilità. La difficoltà principale, però, consiste nell'adattare la piattaforma alle necessità di ognuna di esse. In questa sezione ci occupiamo di fornire una breve panoramica clinica, al fine di rendere più comprensibili le scelte d'implementazione dei ricercatori.

La malattia di Parkinson è un disordine degenerativo e progressivo del sistema nervoso centrale che produce difficoltà motorie e di comunicazione, lentezza (e a volte perdita) dei movimenti, rigidità e tremori a bassa frequenza. Questi ultimi si possono contenere, in parte, mediante una superficie di appoggio rigida e semiflessibile.

La quadriplegia è una forma di paralisi di tutti e quattro gli arti, dovuta, in linea generale, alla lesione di una delle vertebre della cervicale. Si manifesta in forma più o meno grave a seconda della vertebra compromessa e, nel caso peggiore, porta il paziente a poter muovere solo la testa.

Una persona con ridotte capacità visive (nei casi più gravi si parla di ipovisione) può riscontrare problemi di messa a fuoco di oggetti lontani o, in alcuni casi, di "annebbiamento" delle immagini stesse. In questo contesto è molto utile disporre di uno schermo che si possa regolare a seconda delle esigenze dell'utente.

Gli anziani, infine, pur non essendo affetti da disabilità motorie, soffrono spesso di disfunzioni di tipo cognitivo e necessitano, pertanto, di strumenti che li aiutino a svolgere con più semplicità le attività di ogni giorno.

Sulla base di quanto esposto precedentemente risulta chiara quindi la strategia che deve essere intrapresa. Per far fronte alle difficoltà dei malati di Parkinson è conveniente creare un dispositivo, costituito da una superficie solida e semiflessibile, che gestisca in modo semplice l'interazione dell'utente. Per coloro che non dispongono delle braccia lo strumento sopra descritto può essere pilotato mediante i piedi o, in alternativa, può essere rimpiazzato da un braccialetto, una collana o qualsiasi altro oggetto di uso comune (opportunamente riconosciuto dal computer).

Nel caso della quadriplegia l'interfaccia deve poter riconoscere i movimenti della testa e, sulla base di questi, pilotare il cursore del computer.

Per quanto riguarda i soggetti con ridotte capacità visive è consigliabile utilizzare un videoproiettore che possa ingrandire l'immagine a proprio piacere. La configurazione della grandezza dello schermo si può effettuare, per esempio, utilizzando uno schermo *touch screen*.

Infine per gli anziani (o le persone normodotate in generale) è utile costruire un dispositivo di puntamento semplice da utilizzare e di forma simile ad un telecomando della televisione.

L'implementazione

Le modalità

Le diverse modalità utilizzate in questo progetto sono riportate in Tabella 2.

Tabella 2. Modalità utilizzate.

<i>Scenario</i>	<i>Modalità</i>
Parkinson	Tatto
Quadriplegia	Movimenti della testa, voce
Ridotte capacità visive	Vista, Tatto
Senza braccia	Tatto (con i piedi)
Normodotati e anziani	Vista

La scelta della modalità di interazione viene eseguita dall'utente durante la fase iniziale di configurazione, ma può essere modificata in qualsiasi momento. Inoltre il sistema è in grado di riconoscere e attivare automaticamente la modalità in uso tramite delle interrogazioni cicliche.

L'hardware

Gli obiettivi presentati nella sezione precedente sono stati raggiunti, con successo, adattando uno dei componenti della *console Nintendo Wii*, chiamato *Wii Remote*. Le caratteristiche principali di questo dispositivo sono: una fotocamera ad infrarossi con risoluzione 1024x768 (che sfrutta la tecnologia *PixArt's Multi-Object Tracking* per monitorare quattro punti infrarossi 100 volte al secondo), accelerometri sui tre assi (che operano alla medesima frequenza), connessione bluetooth con consumo medio della batteria, *feedback* a vibrazione, 11 tasti configurabili (oltre al tasto di accensione/spengimento) e 2 batterie AA con autonomia di 15 ore. Un singolo dispositivo ha un prezzo commerciale che si aggira intorno ai \$39,99: si mantengono quindi le "promesse" di progetto *low-cost* di cui si è parlato precedentemente. Inoltre questa spesa si potrebbe ridurre ad un terzo costruendo una soluzione di tipo proprietaria.

Dal punto di vista hardware, oltre al controller *Wii Remote*, sono stati impiegati alcuni *led* a infrarossi (*LED IR*) collocati su supporti differenti a seconda del tipo di

disabilità. Questi *led* hanno il vantaggio di funzionare egregiamente anche in ambienti poco luminosi e di poter essere collocati potenzialmente su qualsiasi oggetto: braccialetti, collane, ecc. (si veda Figura 10).

Un'ultima considerazione riguarda, infine, coloro che hanno ridotte capacità visive. In questo contesto si è proposto l'utilizzo di un video-proiettore, il cui acquisto potrebbe essere considerato una spesa impegnativa. Tuttavia, visto il largo uso che si fa di questo strumento sia nelle scuole, sia nelle case per anziani, possiamo considerare questo problema facilmente risolvibile.



Figura 10. Hardware (González González et al., 2008, p. 72).

Il software

Naturalmente, una volta individuato l'hardware, è stato necessario creare uno strato *middleware* opportuno, in grado di interfacciare il dispositivo con un qualsiasi *browser web* installato su un computer.

La soluzione è stata quella di adattare il lavoro di Johnny Chung Lee dell'Università della Virginia, nel quale venivano mostrati una serie di modi originali per utilizzare la *Wii*; si è focalizzando però l'attenzione su un punto che non veniva trattato in questo lavoro, vale a dire renderla accessibile a persone con disabilità.

Dal punto di vista funzionale, per poter gestire le esigenze degli utenti si è scelto di creare dei profili completamente personalizzabili.

Il processo

In questa sezione ci occupiamo di descrivere come vengono impiegati gli strumenti hardware in ognuno dei diversi contesti di disabilità (si veda Figura 11).



Figura 11. Da sinistra a destra: a) persone con ridotte capacità visive; b) parkinson e persone senza braccia; c) quadriplegia; d) utenti normodotati e anziani (González González et al., 2008, p. 71).

Utenti normodotati e anziani

In questo scenario il dispositivo *Wii Remote* viene utilizzato in combinazione con due *led IR*, collocati agli angoli superiori dello schermo. Questi ultimi permettono alla fotocamera *IR* di riconoscere, in qualsiasi istante, dove l'utente vuole puntare l'apparecchio. I dati relativi alla posizione vengono trasferiti via Bluetooth al computer il quale, dopo averli elaborati, esegue il relativo movimento del cursore.

L'operazione di *click* può essere eseguita dall'utente utilizzando un tasto qualsiasi del dispositivo.

Quadriplegia

In questo contesto l'utente indossa un diadema su cui sono collocati due *led IR*, mentre il dispositivo viene fissato su un alloggiamento al di sopra dello schermo. Come nel caso precedente, *Wii Remote* è in grado di monitorare continuamente la posizione dei due *led* e riconoscere eventuali spostamenti dell'utente.

L'operazione di *click* può essere eseguita in due modi: muovendo la testa in avanti o, nel caso non fosse possibile, utilizzando un microfono appositamente posizionato sul diadema. Nel primo caso il dispositivo riconosce il movimento della testa in base all'aumento/diminuzione della lunghezza della linea immaginaria che connette i due *led IR*. Nel secondo, invece, le operazioni di *click* e di *drag&drop* (trascinamento della selezione) sono associate in modo diretto ad un soffio breve o lungo nel microfono. In particolare quest'ultima modalità è accompagnata dalla riproduzione di un *feedback* sonoro sotto forma di rumore bianco.

Parkinson e persone senza braccia

Per gestire le difficoltà dei malati di Parkinson e delle persone senza braccia, i ricercatori propongono l'utilizzo di un apposito apparecchio, costituito da due sezioni circolari separate da tre molle a compressione; la parte superiore di questo strumento è dotata di un alloggiamento, dove è possibile fissare il dispositivo *Wii Remote*.

L'utente interagisce con l'apparecchio in maniera simile a come si utilizzerebbe un "grande mouse" vale a dire appoggiandocisi sopra in modo da far flettere l'asse superiore. Per coloro che non dispongono delle braccia lo strumento può essere utilizzato in maniera equivalente con i piedi.

L'azione di *click* può essere eseguita secondo due modalità: i malati di Parkinson possono utilizzare un qualsiasi tasto del dispositivo, mentre chi è privo di braccia fa uso di un grande pulsante posizionato sulla superficie superiore.

Personne con ridotte capacità visive

L'ultima applicazione riguarda, invece, le persone con ridotte capacità visive. In questo caso il materiale necessario è un videoproiettore e una penna alla cui estremità è presente un *led IR*.

Il videoproiettore viene utilizzato alla maniera usuale, con l'unico accorgimento che la regolazione della grandezza dell'immagine può essere effettuata, a proprio piacere, mediante un apposito schermo *touch screen*. Per quanto riguarda invece il dispositivo *Wii Remote*, esso può essere posizionato ovunque, a patto che l'angolo visivo della sua fotocamera gli permetta di "vedere" tutto lo schermo.

Il funzionamento è molto semplice: l'utente utilizza la penna in dotazione sullo schermo proiettato. Il *led IR* e *Wii Remote* si occupano di gestire il movimento del cursore, in maniera del tutto simile a quanto accadeva nei casi precedenti.

Due ultime considerazioni vanno riportate, infine, a riguardo dell'attivazione della penna e della calibrazione dello schermo. La prima può avvenire in due modi: prendendo in mano lo strumento o premendo un apposito tasto. L'operazione di calibrazione, invece, che consiste nel toccare i quattro angoli dello schermo (mediante la penna *IR*), viene eseguita durante la prima fase di configurazione e risulta essere indispensabile dal momento che le dimensioni dello schermo non sono fisse.

La sperimentazione clinica

I test dell'interfaccia sono stati suddivisi in due fasi: la prima, eseguita in laboratorio, ha permesso di verificare i diversi scenari utilizzando come misura di riferimento la facilità di movimento del mouse.

La seconda, invece, ha cercato di cogliere le sensazioni che provavano gli utenti navigando su pagine web di diversa complessità.

Per rendere più significativo il campione statistico si è scelto di utilizzare due gruppi di cinque persone ciascuno: i primi con età compresa tra 20 e i 30 anni, i secondi invece oltre i 50. Come già detto in precedenza, inoltre, vi è stata la partecipazione congiunta di alcuni importanti istituti: la *Parkinson Association in*

Madrid, la *Federation of People with Intellectual Disability* e il *National Centre for Technical Aids and Personal Autonomy (CEAPAT)*.

La prima fase

Nel contesto degli anziani e delle persone normodotate, il dispositivo di puntamento è stato sperimentato a distanze variabili di 1, 2 e 3 metri e ha portato risultati molto soddisfacenti. Dal punto di vista tecnico, si è potuto osservare che l'accuratezza dello strumento aumenta sensibilmente al diminuire della distanza (si veda Figura 12a).

Nell'ambito delle persone con ridotte capacità visive si è arrivati alla medesima conclusione e, in particolare, si è osservato che l'azione di *click* risulta essere molto precisa fino a tre metri dallo schermo.

Il riconoscimento dei movimenti della testa per i quadriplegici si è rivelato essere estremamente accurato sia nell'identificare la direzione del movimento, sia nel permettere all'utente il movimento del cursore e i *click* (si veda Figura 12b).

Infine, in relazione a tutti gli scenari, la gestione delle batterie si è rivelata essere un punto fondamentale. Nell'ambito dei quadriplegici e di coloro soggetti a ridotte capacità visive, il dispositivo ha avuto un'autonomia delle batterie più che soddisfacente. Negli altri due casi invece, a causa del trasferimento dei dati via bluetooth, le batterie hanno avuto una durata complessiva decisamente più ridotta, compresa tra le 15 e le 20 ore. Per limitare questi problemi, in futuro, i ricercatori propongono due interessanti miglioramenti: una base, da posizionare sopra lo schermo, che permetta la ricarica del dispositivo e un collegamento USB per alimentare il controller.



Figura 12. Da sinistra a destra: a) utenti normodotati e anziani; b) quadriplegia (González González et al., 2008, p. 72).

La seconda fase

La seconda parte degli esperimenti ha cercato di valutare la facilità con il quale gli utenti interagivano con il web, utilizzando come base di riferimento tre pagine di diversa complessità: *CEAPAT*, il *Portal Mayores* e *Discapnet*.

Nel complesso l'interazione con il primo sito web è stata un'esperienza più che soddisfacente. La seconda pagina, pur risultando più complessa della prima, è risultata gradevole agli utenti, grazie alla sua semplicità di selezionare i dati. La terza, invece, ha causato molti problemi di navigazione, dovuti principalmente alle troppe voci presenti nei menù e al carattere eccessivamente ridotto dei testi.

Il *feedback* ricevuto dal gruppo più giovane è stato decisamente molto positivo, in termini di velocità e accuratezza dei movimenti. Gli utenti hanno imparato ad interagire con questa nuova modalità in maniera efficiente, senza errori e in pochissimo tempo grazie, in parte, alle esperienze pregresse con il mouse tradizionale.

Il gruppo di persone anziane invece, pur riscontrando qualche difficoltà iniziale, ha mostrato di saper interagire in maniera soddisfacente negli scenari più accessibili, raggiungendo livelli di velocità e accuratezza simili a quelli dei giovani.

Per quanto concerne le persone diversamente abili riporto qui di seguito i principali *feedback* ottenuti.

I quadriplegici hanno mostrato buone capacità di gestione della piattaforma, con dei tempi di interazione però molto lunghi.

Gli anziani e le persone soggette a ridotte capacità visive, invece, sono state in grado di navigare sul web con una velocità simile a quella dei dispositivi di puntamento tradizionali.

Nel futuro si prevede di eseguire ulteriori esperimenti in scenari reali (la casa digitale allestita all'*EUIT* delle Telecomunicazioni dell'Università Politecnica di Madrid) lasciando liberi gli utenti di interagire con qualsivoglia sito internet.

Il futuro

L'evoluzione di questa piattaforma potrebbe portare alla nascita di moltissime applicazioni tra cui, per esempio, i giochi di memoria per l'apprendimento e gli esercizi per stimolare la conoscenza.

Inoltre, allargando la prospettiva, si potrebbe pensare alla creazione di sistemi multimodali intelligenti (in grado di predire le intenzioni dell'utente), al supporto di altre disabilità e alla gestione di servizi telematici aggiuntivi, come per esempio l'e-mail.

1.2.5 Stanza logo-motoria: apprendimento e comunicazione in un ambiente interattivo multimodale

Il team

Il progetto, denominato stanza logo-motoria, nasce dalla collaborazione congiunta di sette ricercatori provenienti da diverse realtà interessanti. Antonio Camurri, Corrado Canepa e Gualtiero Volpe del laboratorio *InfoMus* dell'Università di Genova. Sergio Canazza e Antonio Rodà del gruppo di ricerca *Sound and Music Computing (SMC)* dell'Università di Padova e, infine, Serena Zanella dell'Università di Udine (Camurri et al., 2010).

I gruppi di ricerca sopra citati sono impegnati in molteplici aree di studio. Il laboratorio *InfoMus Lab* dell'Università di Genova focalizza l'attenzione su diversi filoni, che abbiamo già avuto occasione di descrivere nella sezione 1.2.1. Il gruppo di ricerca *Sound and Music Computing* dell'Università di Padova è impegnato principalmente nella sintesi e nel *rendering* sonoro, nel campo dell'espressività musicale, nel restauro audio, nelle tecniche di miglioramento dell'acustica e nel riconoscimento (e sintesi) vocale. Infine per quanto concerne i ricercatori dell'Università di Udine, Serena Zanolla è una dottoranda in Comunicazione Multimediale presso il Dipartimento di Scienze Umane e afferisce, inoltre, al Dipartimento di Matematica e Informatica. La complessità del progetto è ben rispecchiata dall'eterogeneità del gruppo di lavoro.

L'idea

L'interfaccia in questione è un sistema interattivo multimodale che ha lo scopo di migliorare l'apprendimento di alcune materie scolastiche e stimolare diversi aspetti comunicativi. La piattaforma risiede, in maniera permanente, in una scuola statale primaria a Gorizia e viene utilizzata sia da bambini normodotati, sia da fanciulli affetti da alcuni disturbi dello sviluppo e dell'apprendimento (per es. autismo e dislessia).

In particolare, il sistema, grazie ai sensori installati nell'ambiente, analizza i movimenti corporei e la gestualità dei bambini e vi associa la riproduzione di alcuni *feedback* sonori. Tra i gesti che vengono riconosciuti, quelli di tipo espressivo ricoprono un ruolo particolarmente rilevante, in relazione soprattutto al messaggio emotivo-affettivo che sono in grado di trasportare.

Luogo e anno di realizzazione

Il sistema è stato realizzato presso i laboratori delle Università di Genova, Padova e Udine; il nucleo centrale del software (*EyesWeb XMI*) proviene dal laboratorio *InfoMus* di Genova.

La fase di sperimentazione si è svolta presso la scuola statale primaria “E. Frinta” a Gorizia.

I primi risultati scientifici sono stati pubblicati nel 2010, mediante una licenza *Creative Commons*.

Gli obiettivi

Per migliorare la chiarezza espositiva, le finalità del progetto sono state raggruppate in quattro filoni principali: le motivazioni di contesto, quelle teoriche, quelle legate all'apprendimento e infine quelle cliniche.

Motivazioni di contesto

L'aumento degli studenti stranieri e di quelli disabili all'interno delle scuole primarie ha portato, nell'ultimo decennio, notevoli trasformazioni nel sistema educativo europeo.

Gli insegnanti, per far fronte a questi cambiamenti, hanno dovuto farsi “supportare” in diversi modi: per esempio hanno richiesto la collaborazione di maestri di sostegno, hanno iniziato a sfruttare materiali alternativi durante le lezioni e si sono avvicinati a nuovi strumenti ed equipaggiamenti tecnologici. Anche le modalità di insegnamento si sono trasformate notevolmente: dalla semplice lezione frontale si è passati a delle nuove metodologie che, sulla base della teoria delle Intelligenze Multiple di H. Gardner²², valorizzano nuove abilità sfruttando per esempio la musica, la cooperazione, l'arte, il gioco, la multimedialità. In questo contesto la stanza logo-motoria diventa uno strumento di supporto formidabile per l'insegnante, in grado di far sviluppare intelligenze “alla Gardner” che prima venivano trascurate.

Motivazioni teoriche

Dal punto di vista letterario il lavoro è inserito in una cornice molto variegata, che verrà descritta di seguito per sommi capi.

Un primo aspetto importante riguarda l'approccio *Learning by doing*, vale a dire imparare qualcosa di nuovo svolgendo una certa attività. Il concetto, pur essendo

²² La teoria delle Intelligenze Multiple di H. Gardner (1983) afferma l'esistenza di otto tipi diversi di intelligenze: linguistica, logica-matematica, spaziale, musicale, corporeo-cinestesica, interpersonale e intrapersonale. Secondo il ricercatore le scuole si focalizzano solo su due di queste: la linguistica e la logica-matematica.

intuitivamente semplice, ha il suo fondamento in una scoperta neurofisiologica ed è strettamente connesso alle teorie relative all'*Enactive knowledge* e all'*Embodiment*. La scoperta a cui facciamo riferimento è quella concernente i neuroni specchio. Questo tipo particolare di neuroni, scoperto in un primo momento nelle scimmie e identificato successivamente nell'uomo, permette il riconoscimento e la comprensione delle azioni eseguite da altri individui. Il fatto che sia proprio l'azione il motore della sua attivazione ci riconnette in maniera diretta al concetto iniziale. Le teorie relative alla conoscenza enattiva e all'*Embodiment* sono intrinsecamente connesse tra di loro e seguono il percorso logico di seguito riportato. È possibile identificare due tipologie diverse di conoscenza: quella scientifica e quella che deriva dall'esperienza. La prima ha come obiettivo raccogliere tutte le informazioni necessarie per rispondere ad una certa domanda. La seconda, invece, è quella che svolgiamo "inconsapevolmente" fin dalla nascita, vale a dire quella forma di conoscenza che ci permette di apprendere facendo qualcosa. Focalizziamo l'attenzione su quest'ultima tipologia. Secondo la teoria di Varela, Thompson e Rosch (1991)²³, l'esperienza è enattiva, vale a dire è un qualcosa (un'abilità) che ogni animale mette in atto (*enacts*) esplorando l'ambiente nel quale è situato (*embodied*). Il termine *Embodiment*, in questo frangente, mette in risalto due ulteriori concetti: in primo luogo, il fatto che la conoscenza cresce in maniera proporzionale all'esperienza (allenamento) che abbiamo "maturato" con le nostre capacità sensoriali-motorie. In seconda battuta, queste ultime risultano essere intrinsecamente radicate in un contesto di tipo biologico, psicologico e culturale: detto in termini più semplici "in quello che facciamo" (Avanzini, 2008). La connessione con l'idea espressa all'inizio del paragrafo dovrebbe risultare a questo punto evidente. La stanza logo-motoria cerca di far sue queste motivazioni e, in un certo senso, supera questa visione, costruendo un ambiente dove l'utente è "costretto" (indirettamente) a svolgere un movimento (*in motion*) per poter recepire nuovi contenuti (*learning*). Si parla quindi, in aggiunta a *Learning by doing*, di un secondo concetto importante: *Learning in motion*.

Procedendo nell'analisi, un'ulteriore motivazione teorica proviene dai lavori del ricercatore Krueger, considerato da molti il padre della "realtà virtuale". Egli focalizza l'attenzione su due aspetti interessanti: in primo luogo, mette in risalto la stretta connessione che intercorre tra ambiente e spazio fisico e, in virtù di questo, crea degli ambienti interattivi dove l'utente può interagire liberamente senza indossare alcun tipo di dispositivo esterno. È l'ambiente stesso che si occupa di recepire i movimenti dell'utente (mediante dei sensori a pavimento e delle

²³ Una trattazione più esaustiva dell'argomento è stata pubblicata nel 2001 dai ricercatori O'Regan and Noë, i quali si sono occupati di raggruppare, in un unico lavoro, le idee "simili" presenti in diverse opere.

telecamere) e di elaborare queste informazioni per rispondere in modo appropriato. Tale processo è basato fondamentalmente su due operazioni: la localizzazione della posizione della persona e il riconoscimento dei gesti eseguiti.

Il secondo concetto, evidenziato dal ricercatore, è quello di creare dei *responsive environments*, vale a dire degli spazi “intelligenti” in cui il computer è il mezzo di connessione tra due o più individui e fornisce un’interazione real-time non prevedibile dall’utente. Questi ambienti sono degli strumenti potentissimi attraverso cui l’uomo può riacquisire tutte quelle facoltà sensoriali che vengono trascurate a causa della supremazia dell’audio e del video nella nostra vita quotidiana. Il progetto in esame cattura entrambi gli aspetti eseguendo l’acquisizione video della posizione e dei gesti dell’utente (corredata da opportuni *feedback*) e rivelandosi un ottimo strumento per riacquisire le facoltà sensoriali trascurate. In aggiunta a tutto questo, la stanza logo-motoria riesce a catturare l’attenzione dei bambini e ad incrementare la loro soddisfazione personale.

Infine, ulteriori spunti vengono offerti dai progetti *SOUND=SPACE* e *CaDaReMi*. Il primo, nato nel 1984, proponeva una piattaforma multi-utente in cui gli ospiti erano in grado di produrre musica muovendosi semplicemente in una stanza. Il progetto ha avuto modo di evolversi notevolmente nel corso del tempo, tuttavia questo tipo di approccio è ancora oggi uno tra i più utilizzati, perché permette ai ragazzi con bisogni educativi speciali di ottenere notevoli benefici. Il progetto *CaDaReMi* (Gehlhaar, Girao, & Rodrigues, 2008), sviluppato più di vent’anni dopo, fornisce anch’esso una piattaforma multi-utente interattiva dove le persone sono libere di muoversi nello spazio. Tuttavia, a differenza del suo predecessore, lo scopo dell’utente è fare in modo che un *avatar*, proiettato su uno schermo, possa raggiungere determinati obiettivi. L’introduzione di un’informazione visiva (l’*avatar* sullo schermo) fornisce all’utente dei punti di riferimento nello spazio, risolvendo una delle più grandi difficoltà riscontrate dagli utenti di *SOUND=SPACE*. I “suggerimenti visivi”, inoltre, aiutano l’utente a comprendere il funzionamento della piattaforma, senza dover ricorrere ad alcun tipo di training. Il progetto riprende anche queste caratteristiche offrendo un ambiente espressivo, collaborativo, stimolante, intuitivo e accattivante (sia visivamente, sia socialmente). Come succedeva in *CaDaReMi*, le modalità di interazione con l’interfaccia vengono scoperte in maniera naturale dall’utente, senza partecipare ad alcun corso di formazione. Inoltre la piattaforma mette alla prova l’utente proponendogli sempre nuove sfide da affrontare; per aumentare il suo coinvolgimento, in particolare, si fa uso di un meccanismo “a livelli”, simile a quello che viene impiegato nei videogames.

Motivazioni legate all'apprendimento

Come affermato in precedenza, l'obiettivo principale dell'interfaccia è migliorare l'apprendimento di alcune materie scolastiche: nelle prime fasi si è focalizzata l'attenzione sull'inglese, l'italiano, la storia e la musica.

In particolare l'insegnamento della lingua inglese si è rivelato essere di notevole interesse rispetto alle altre materie, in virtù degli obiettivi potenzialmente raggiungibili mediante l'impiego della stanza logo-motoria. In primo luogo il riconoscimento di fonemi differenti dal proprio linguaggio poteva essere sfruttato come mezzo per comprendere gli strumenti non linguistici impiegati per comunicare (per es. il linguaggio del corpo). Inoltre la piattaforma poteva consentire ai bambini di parlare l'inglese in maniera spontanea, divertente e naturale, avvalendosi della metodologia *Integrated Learning*, vale a dire imparare "inconsapevolmente" frasi e parole di un linguaggio perché queste sono il mezzo attraverso cui raggiungere un obiettivo nel gioco.

In aggiunta a queste motivazioni specifiche, il sistema si prefigge un obiettivo più generale, vale a dire abbracciare trasversalmente più linguaggi: quello musicale (per es. tramite il riconoscimento dei suoni), quello teatrale (con il movimento, la mimica, ecc.), quello legato alla tecnologia (tramite l'utilizzo di nuove tecnologie, appunto) e quello dei *community languages* (per es. i *Content and Language Integrated Learning (CLIL)*²⁴).

Motivazioni cliniche

Nelle motivazioni di contesto sono state descritte le notevoli trasformazioni del modello educativo recente: l'aumento degli studenti disabili nelle scuole primarie è stato sicuramente uno dei motori principali di questo fenomeno. Proprio per questo motivo, la piattaforma si pone l'obiettivo di supportare bambini con bisogni educativi speciali, focalizzando l'attenzione, in prima battuta, su soggetti autistici, dislessici, ipovedenti e con disturbi comportamentali.

Lo scopo dell'interfaccia, a prescindere dal tipo di disabilità, è quello di facilitare l'apprendimento delle discipline scolastiche, compensando le abilità mancanti oppure semplificando i metodi attraverso cui raggiungere gli obiettivi prestabiliti.

In particolare, per i bambini autistici si rivolge l'attenzione alle abilità comunicative della persona, invece per i soggetti con dislessia si cerca di facilitare l'apprendimento costruendo strumenti di valutazione specifici. Nel corso

²⁴ Si tratta di una metodologia di insegnamento che punta alla costruzione di competenze linguistiche e abilità comunicative in lingua straniera insieme allo sviluppo e all'acquisizione di conoscenze disciplinari.

dell'elaborato verranno illustrate le principali tecniche (ed esercizi) utilizzate per raggiungere queste finalità.

Per quanto riguarda i bambini ipovedenti e coloro soggetti a disturbi comportamentali, sfruttando la semplicità di utilizzo della piattaforma (che richiede solo l'abilità di ascoltare e di muoversi), si vuole raggiungere il duplice obiettivo di sostenerli nell'apprendimento e di facilitare la loro integrazione nel gruppo.

L'implementazione

Le modalità

Il progetto prevede l'utilizzo delle seguenti modalità:

- i movimenti del corpo e la gestualità;
- l'udito.

L'hardware

L'architettura del sistema è suddivisa in due parti: le apparecchiature necessarie per registrare i movimenti dell'utente e i sistemi con i quali riprodurre i *feedback* sonori.

Per quanto riguarda la prima parte, la stanza è stata "sensorizzata" utilizzando delle videocamere e dei microfoni: sviluppi futuri prevedono l'impiego di sensori kinect e, per applicazioni specifiche, di pressione a pavimento. Le prime eseguono l'acquisizione delle immagini, mentre i microfoni e i sensori forniscono ulteriori informazioni relative alle parole pronunciate dall'utente e alle zone occupate.

Il *feedback* sonoro viene naturalmente riprodotto mediante delle casse audio.

Il software

L'interfaccia è stata implementata utilizzando il software *EyesWeb XMI*, variante del sistema *EyesWeb* "tradizionale" già descritto nella sezione 1.2.1.

In particolare l'applicazione è stata classificata nella categoria *Interactive Multimedia Systems and Augmented Action and Perception* del *SMC Roadmap*.

Il processo

Il flusso di elaborazione delle informazioni è composto da tre fasi: la ricezione degli *input*, l'estrazione delle *features* e la riproduzione real-time dei contenuti audio (si veda Figura 13).

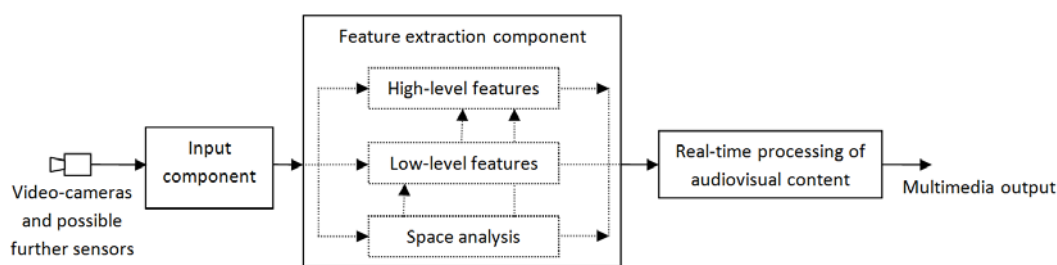


Figura 13. Il flusso di elaborazione dei dati (Camurri et al., 2010, p. 356).

La prima fase ha il compito di immagazzinare i dati provenienti dalle videocamere e dai sensori e di eseguire un *pre-processing* iniziale per eliminare gli eventuali rumori ed estrarre la silhouette dell'utente dallo sfondo. In particolare, quest'ultima operazione viene eseguita mediante il *brightness/chromaticity distortion method*.

Il secondo stadio, vale a dire l'estrazione delle *features*, permette di valutare come l'utente occupa lo spazio e le caratteristiche di espressività dei suoi gesti. L'approccio utilizzato è basato su tre livelli di profondità: l'analisi dello spazio, l'estrazione delle *features* di basso livello e quelle di alto livello. L'analisi dello spazio viene eseguita monitorando la traiettoria dell'utente e determinando le regioni maggiormente occupate (*occupation rate*). Sulla base di queste informazioni si costruisce un profilo personalizzato che descrive il comportamento (le abitudini) dell'utente. Le *features* di basso livello provengono direttamente dai sensori e, in particolare, comprendono le *features* cinematiche (es. posizione, velocità, ecc.) e quelle relative alla silhouette (es. *Motion Index*, *Contraction Index*²⁵, orientamento del corpo, ecc.). Le caratteristiche di alto livello, invece, vengono calcolate, a partire dalle precedenti, utilizzando diverse teorie, modelli ed esperimenti tra cui la *Rudolf Laban's Theory of Effort* e gli esperimenti di Wallbott, De Maijer, Boone e Cunningham. Alcuni esempi di *features* di alto livello sono il *Directness Index*, l'*Impulsivity Index* e la *Fluidity*.

L'ultima fase, infine, si occupa di controllare e riprodurre il materiale sonoro, sulla base delle *features* precedentemente descritte.

Un'applicazione concreta: Resonant Memory Application

Nel paragrafo precedente è stato descritto il funzionamento tecnico della stanza logo-motoria. Ora, per comprendere meglio il processo, si analizzerà un esempio pratico di applicazione: la *Resonant Memory Application*. In particolare, in questa sezione verrà spiegata la dinamica di interazione, mentre nella successiva focalizzeremo l'attenzione sugli aspetti relativi alla sperimentazione clinica.

²⁵ Di cui abbiamo già avuto modo di parlare nella sezione 1.2.1.

Nella *Resonant Memory Application* lo spazio viene suddiviso²⁶ in nove regioni: una centrale (a cui viene associata la narrazione di una storia) e otto periferiche (dove vengono riprodotti rumori, suoni e musiche che sono collegati in qualche modo al racconto).

Il bambino, inizialmente, può esplorare l'ambiente in maniera libera (*exploration mode*) muovendosi in qualsiasi regione dello spazio: tutte le zone gli restituiscono il relativo *feedback*, ad eccezione di quella centrale. Quest'ultima diverrà operativa, vale a dire la narrazione della storia avrà inizio (*story mode*), solo nel momento in cui il bambino avrà terminato di visitare tutte le regioni esterne (si veda Figura 14).

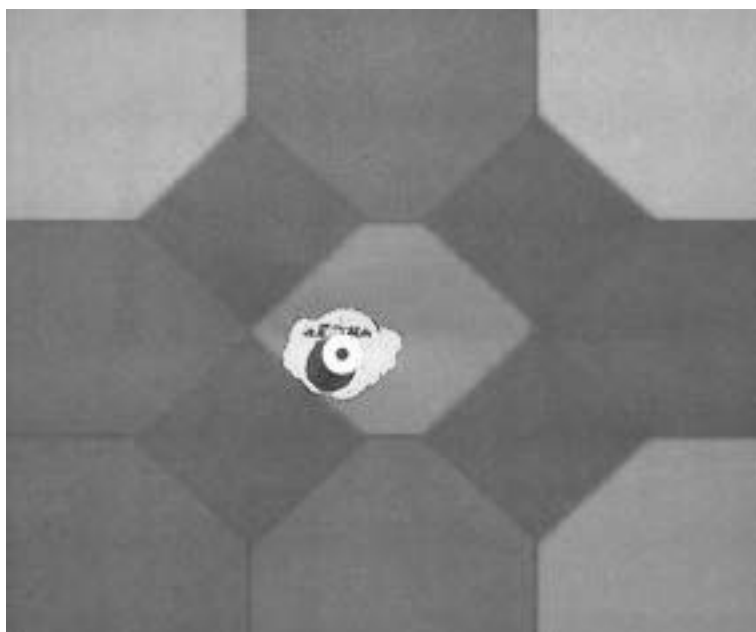


Figura 14. Story mode (Camurri et al., 2010, p. 356).

Le dinamiche più interessanti, dal punto di vista dell'apprendimento, avvengono proprio durante la narrazione: infatti, in questa fase, il bambino deve rintracciare i suoni precedentemente ascoltati spostandosi nelle zone corrette (stimolando di fatto l'utilizzo della memoria).

In qualsiasi istante l'utente può interrompere il gioco aprendo semplicemente le braccia (*pause mode*) e riprenderlo chiudendole. Se la piattaforma rimane inattiva per un lungo periodo di tempo l'applicazione si riavvia (*reset mode*) e la cronologia degli eventi viene cancellata.

²⁶ In modo virtuale, non fisicamente.

La sperimentazione clinica

Le applicazioni

Il collaudo della piattaforma è stato effettuato coinvolgendo dieci classi della scuola statale primaria “E. Frinta” a Gorizia (per un totale di 170 bambini). I principali ambiti di applicazione sono stati l’inglese, la storia, la musica, il progetto Infolibro e il supporto a bambini diversamente abili. A titolo esemplificativo, analizziamo le applicazioni riguardanti l’insegnamento della lingua inglese, la soluzione proposta per i bambini dislessici e le metodologie di lavoro adottate per gli ipovedenti e coloro soggetti a disturbi comportamentali.

Per quanto concerne l’insegnamento dell’inglese si è utilizzata una variante della *Resonant Memory Application*, dove la storia riprodotta era naturalmente in lingua inglese. In particolare, rispetto al modello tradizionale, si è introdotta un’ulteriore difficoltà: alcune delle aree periferiche riproducevano delle parti di racconto dove alcuni nomi, aggettivi o verbi risultavano mancanti. Il bambino doveva individuarli correttamente pronunciandoli ad alta voce (si veda Figura 15).



Figura 15. L'insegnamento della lingua inglese (Camurri et al., 2010, p. 358).

Per aiutare i bambini dislessici, la *Resonant Memory Application* è stata modificata in questo modo: il testo da studiare è stato suddiviso in parti più piccole²⁷, ognuna delle quali è stata descritta mediante una sequenza di suoni e, alternativamente, di immagini. In una prima sessione di interazione lo studente aveva il compito di ascoltare la sequenza di suoni e posizionare le relative immagini su una *billboard*²⁸ secondo l’ordine suggerito. Nella seconda sessione, invece, doveva ascoltare nuovamente la sequenza e ripetere i contenuti, servendosi delle immagini precedentemente posizionate.

²⁷ Tante quante le aree della stanza logo-motoria.

²⁸ In ogni area era stata preventivamente piazzata una *billboard*.

Infine, per i bambini ipovedenti e con disturbi comportamentali non si è utilizzata alcuna applicazione specifica: questi hanno partecipato alle normali attività insieme ai compagni.

Principali risultati ottenuti

Per poter valutare le performance della piattaforma si sono utilizzate tre metodologie differenti: l'osservazione diretta degli esercizi, l'analisi dei video e l'utilizzo di test di valutazione "tradizionali"²⁹.

In tutte le aree di sperimentazione si è potuto notare un notevole aumento del coinvolgimento, della motivazione, dell'entusiasmo e del periodo di attenzione. I bambini hanno raggiunto tutti gli obiettivi proposti a partire da quelli più semplici, fino ad arrivare a quelli più complessi.

Focalizzando l'attenzione sull'insegnamento della lingua inglese, a differenza dei metodi tradizionali si è potuto osservare un notevole aumento della motivazione nell'ascoltare i dialoghi (e di conseguenza nell'imparare parole nuove) e un apprendimento "più duraturo" nel tempo. Inoltre l'utilizzo di un ambiente in cui la percezione del suono era assoluta (e l'utente era completamente immerso nel contesto) ha facilitato la comprensione della lingua e il miglioramento della pronuncia e della produzione orale.

Anche i risultati ottenuti nel campo delle disabilità sono stati notevoli: vi è stato un aumento delle abilità comunicative, dell'espressività dei gesti, dell'interazione e dell'autonomia personale. Nell'ambito specifico della dislessia, inoltre, si è riusciti a superare il grande ostacolo della scrittura, con il conseguente aumento della motivazione ad imparare nuovi contenuti. Gli studenti sono stati in grado di padroneggiare gli argomenti a tal punto da saperli riassumere, con sicurezza, all'insegnante e ottenere i voti più alti nei test. Infine i bambini ipovedenti e con disturbi comportamentali, hanno raggiunto con successo gli obiettivi clinici precedentemente illustrati: hanno portato a termine tutti gli esercizi proposti, integrandosi in modo costruttivo nel gruppo.

Concludendo i risultati ottenuti si possono considerare molto positivi.

Il futuro

Nel futuro si prevede di creare un gruppo di controllo per valutare in modo ancora più approfondito i benefici della piattaforma. Dal punto di vista delle modalità, si vogliono iniziare dei nuovi esperimenti per includere il *feedback* video e il *rendering* spaziale del suono; inoltre, per risolvere le difficoltà dei bambini autistici a

²⁹ Si trattava di compiti scritti con i quali si valutavano gli argomenti appresi mediante la stanza logomotoria.

relazionarsi, si sta studiando un nuovo approccio in cui il *feedback* ricevuto da un utente sia influenzato dalle azioni degli altri.

CAPITOLO 2

Soluzione proposta

Nel presente capitolo si descriverà la soluzione ideata e le motivazioni che sorreggono il progetto.

2.1 Soluzione

La soluzione proposta consiste nella creazione un ambiente interattivo multimodale in cui l'utente, senza essere dotato di alcun tipo di sensore sul corpo, è libero di muoversi nello spazio. Il sistema si occupa di analizzare e riconoscere in tempo reale i movimenti e i gesti dell'utente, associando la riproduzione di opportuni feedback sonori e visivi.

Le interfacce illustrate nella sezione 1.2 motivano la scelta di questa tipologia di soluzione.

La definizione che è stata data precedentemente è decisamente generica: è necessario precisare, infatti, che la tipologia di gesti e i relativi feedback variano in base agli obiettivi che ci si prefigge e agli utilizzatori dell'interfaccia.

Per questo motivo sono state sviluppate, nel dettaglio, due interfacce che hanno finalità e "pubblico" differente: Fiaba Magica è costruita per studenti in situazione di multi-disabilità, mentre Play&Learn è pensata per studenti stranieri oppure per l'insegnamento della seconda lingua a bambini italiani.

Riprendendo la struttura già utilizzata precedentemente, nel presente capitolo si descriveranno le due interfacce focalizzando l'attenzione solo sulla dinamica di interazione. Nel capitolo 3, invece, si analizzeranno i dettagli implementativi relativi ad hardware e software.

Entrambe le interfacce condividono l'utilizzo delle medesime modalità³⁰:

- i movimenti del corpo: in particolare lo spostamento avanti / indietro e l'apertura delle braccia;
- l'udito;
- la vista.

2.1.1 *Fiaba Magica*

Descrizione sintetica

Fiaba Magica è un'interfaccia interattiva multimodale nella quale l'utente ha la possibilità di essere il "regista" della storia di Cappuccetto Rosso, vale a dire decidere come e quando far continuare la storia.

Come avviene l'interazione

Come rappresentato in Figura 16, lo spazio di gioco è stato suddiviso virtualmente in tre aree.

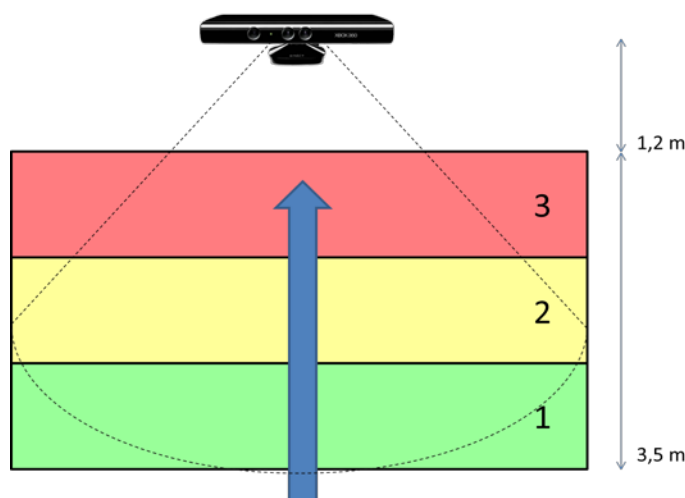


Figura 16. Suddivisione dello spazio di gioco.

Il bambino inizia il gioco entrando nella prima zona (quella più lontana dal sensore): questa azione attiva la riproduzione audio della prima porzione di storia e la visualizzazione a schermo dei due personaggi che ne sono protagonisti (come raffigurato in Figura 17).

³⁰ In questo contesto il termine modalità fa riferimento alle diverse "modalità di interazione" che caratterizzano la cornice multimodale.



Figura 17. I protagonisti della prima porzione di storia.

Una volta ascoltata l'intera porzione di storia, il bambino ha la possibilità di "animare" (vale a dire far parlare e muovere) il personaggio a sinistra o a destra della scena: per farlo è sufficiente aprire il braccio corrispondente.

Animando almeno una volta entrambi i personaggi, il bambino può accedere alla zona successiva dove la narrazione continua con il pezzo successivo di storia e le rispettive animazioni dei personaggi.

La dinamica si ripete in modo assolutamente analogo anche nell'ultima zona, dove la storia termina con il consueto finale (Zanolla et al., 2011).

2.1.2 Play&Learn

Descrizione sintetica

Play&Learn è un'interfaccia interattiva multimodale nella quale i bambini devono rispondere ad alcune domande per poter procedere di livello. L'utilizzo di un feedback audio e video permette all'utente di memorizzare nuove parole in modo quasi inconsapevole.

Come avviene l'interazione

Lo spazio è suddiviso virtualmente in tre aree: la zona più lontana dal sensore, la zona intermedia e la zona più vicina; ad ognuna di queste sono associate tre domande.

Il bambino entra nella zona più lontana, dove gli vengono proposte le prime domande: se risponde correttamente può proseguire al livello successivo (vale a dire muoversi in avanti accedendo alla zona successiva). Altrimenti gli viene chiesto di riprovare.

La particolarità più interessante riguarda la modalità attraverso cui il bambino risponde alle domande. Il bambino ha a disposizione, per ogni domanda, due possibili risposte che vengono mostrate nella parte sinistra e nella parte destra dello schermo. La scelta della risposta avviene semplicemente aprendo il relativo braccio.

2.2 Motivazioni

In questa sezione si descriverà la complessa cornice che caratterizza le motivazioni del progetto. Per poter avere una panoramica complessiva sono stati affrontati diversi ambiti di studio e diverse discipline: la struttura che segue, quindi, ricalca le aree analizzate, fornendo al lettore delle immagini parziali che acquisteranno un significato complessivo solo al termine della lettura.

Si può comprendere molto riguardo la natura e le motivazioni di un'interfaccia, analizzando l'idea che ne ha permesso il concepimento: il progetto stanza logomotoria, illustrato nell'exkursus relativo alle interfacce multimodali (sezione 1.2.5), rappresenta il punto di partenza di questo lavoro. La lunga trattazione riguardo le motivazioni di contesto, teoriche, di apprendimento e cliniche, che è stata illustrata durante la descrizione di questo progetto, forniscono una cornice altrettanto valida per il presente lavoro di tesi. Pertanto evitiamo di riproporle facendo riferimento alla relativa sezione.

In aggiunta a queste, tuttavia, esistono delle ulteriori motivazioni specifiche che andremo ad illustrare qui di seguito: alcune sono connesse alla psicologia dello sviluppo del linguaggio, altre riguardano l'aspetto clinico e infine le ultime fanno riferimento alle normative vigenti.

2.2.1 *Motivazioni connesse alla psicologia dello sviluppo del linguaggio*

Il processo di sviluppo del linguaggio è un argomento di studio decisamente vasto. In questo breve paragrafo si è cercato, pertanto, di focalizzare l'attenzione solo su uno dei molteplici aspetti, vale a dire la comprensione dei testi. Al termine, inoltre, si è inserito un piccolo approfondimento relativo all'insegnamento della lingua straniera che discende, per diversi motivi, dall'argomento principale.

La trattazione ha come obiettivo principale quello di costruire un percorso logico che copra le tematiche principali, evidenziando, durante la strada, gli elementi di contatto con le interfacce Fiaba Magica e Play&Learn.

La comprensione dei testi

La comprensione dei testi è un'abilità che assume notevole importanza per diversi motivi: in primo luogo perché la comunicazione umana avviene, appunto, attraverso discorsi e testi, vale a dire sequenze di frasi caratterizzate da un'organizzazione³¹ (Levorato & Nesi, 2001). Inoltre perché essa è presente ovunque e riveste un ruolo fondamentale nella vita di ogni persona (adulti e tanto più bambini): si pensi, per esempio, a quelle azioni quotidiane come la lettura di un giornale, la spiegazione di un evento ad un amico o il racconto di un film (Levorato & Roch, 2007). Come si potrebbe farne a meno?

Detto questo è però necessario sottolineare che si tratta di una competenza complessa alla quale concorrono diverse abilità di natura sia linguistica, sia cognitiva (Levorato & Roch, 2007). Inoltre, per comprendere a fondo l'argomento, lo si deve analizzare sia dal punto di vista emotivo, sia da quello cognitivo. Le emozioni dei personaggi e l'empatia (nei testi narrativi) sono solo alcuni esempi della competenza emotiva, mentre il processo di elaborazione dell'informazione e della conoscenza è il concetto principe del funzionamento cognitivo. La suddivisione dei paragrafi che seguirà ricalca, appunto, quanto anticipato.

Gli aspetti affettivo-emotivi e quelli cognitivi, inoltre, sono basilari per lo sviluppo dell'alfabetizzazione (*literacy*), intesa non solo come le attività necessarie per imparare a leggere e scrivere, bensì in senso più ampio come la conoscenza dei testi della propria cultura: comprendere il modo in cui sono scritti, i contesti, le relazioni che intercorrono tra autore e lettore e molto altro ancora. Tale processo inizia già a due anni grazie alle narrazioni (di cui avremo modo di parlare nel prossimo paragrafo), viene consolidato durante l'età scolare e continua per tutto il resto della vita. L'obiettivo "a lungo raggio" è quello di comprendere che ogni testo viene prodotto da un autore con specifiche intenzioni comunicative: ciò non significa che esse vengano comprese univocamente, anzi, nella maggior parte dei casi, un testo si può prestare a diverse interpretazioni (Levorato & Nesi, 2001).

Aspetti affettivo-emotivi

Al contrario di quanto potremmo aspettarci, il concetto base da cui si parte per affrontare questo argomento è decisamente "pratico" e "naturale": si tratta della narrazione.

³¹ Con il termine organizzazione ci si riferisce ai principi di coerenza e di organizzazione tematica con cui devono essere costruite le sequenze di frasi. Ciò presuppone che al loro interno sia riconoscibile un tema e venga sviluppato un argomento.

La nostra vita, anche se spesso non ce ne rendiamo conto, è caratterizzata dal susseguirsi di narrazioni. Da bambini i cosiddetti “librini” ci fanno conoscere fiabe, favole e storie più o meno legate alla nostra cultura di appartenenza. Successivamente i racconti di esperienze quotidiane finiscono per diventare una delle principali forme di comunicazione utilizzate, in modo spontaneo, dalle persone. L’importanza delle narrazioni è messa in luce in modo ancora più evidente da Bruner, il quale afferma l’esistenza di una forma di pensiero chiamata, appunto, pensiero narrativo (Levorato & Nesi, 2001). Secondo lo psicologo, si tratterebbe di una predisposizione delle mente ad utilizzare la narrazione come mezzo per comunicare ciò che cogliamo con l’esperienza. Quest’ultima frase è decisamente fondamentale per capire la vera natura del pensiero narrativo: si tratta di uno strumento per comunicare e per stimolare l’interazione con gli altri esseri umani. Non a caso si parla anche di comunicazione narrativa. Al posto di elencare semplicemente le caratteristiche chiave di questo tipo di comunicazione, vorrei tentare di farlo in modo diverso utilizzando il gergo culinario. La comunicazione narrativa può essere paragonata, a mio avviso, al sale da cucina: è di uso comune, grazie alla sua semplicità di utilizzo, e dà sapore alle cose, grazie all’interesse e alla carica emotiva che la contraddistingue. Focalizzando l’attenzione sulle caratteristiche appena delineate e considerando il fatto che la comunicazione narrativa è il mezzo attraverso cui gli adulti entrano spontaneamente in comunicazione con i bambini (fiabe, favole, racconti, come si diceva all’inizio), riusciamo a collegarci in modo diretto con un’altra affermazione di Bruner. Lo psicologo afferma che la comunicazione narrativa svolge un ruolo fondamentale nell’acquisizione del linguaggio. Fondamentalmente esso viene acquisito dai bambini per poter raccontare le proprie esperienze personali, la propria soggettività (Levorato & Nesi, 2001).

Le osservazioni riportate fin qui forniscono delle fondamenta solide per l’applicazione Fiaba Magica, giustificando la scelta di utilizzare la narrazione come mezzo di comunicazione principale.

Un attento lettore potrebbe chiedersi, a questo punto, dove siano finiti gli aspetti affettivo-emotivi di cui si era promesso di parlare all’inizio. Quanto appena detto fornisce le basi per poterne discutere. Il testo narrativo, nei primi anni dell’età evolutiva, viene interpretato dai bambini come un semplice racconto in cui si susseguono delle azioni. Con il trascorrere degli anni, però, essi sono in grado di rileggere gli eventi alla luce degli stati interni (emozioni, affetti, ecc.) dei personaggi, sviluppando la cosiddetta comprensione della soggettività (Levorato & Nesi, 2001).

Sicuramente non si tratta di una capacità semplice da sviluppare. Per fortuna ci vengono in aiuto alcune risposte partecipatorie che, determinando un coinvolgimento

emotivo del lettore, ci aiutano in questa complessa attività. Tra queste, la principale è sicuramente l'empatia: uno stato emotivo che ci permette di provare le emozioni che vengono osservate in qualcun altro. Si viene a creare un vero e proprio circolo vizioso: l'emozione provata dal personaggio della narrazione viene trasferita al lettore per mezzo dell'empatia. Essa stimola, a sua volta, la ricerca di informazioni per comprendere come sia nata l'emozione nel personaggio. Quest'ultima affermazione si inserisce in continuità con gli esperimenti di Phillips (Levorato & Nesi, 2001), che dimostrano la connessione tra l'empatia e il miglioramento della comprensione del testo. Inoltre con quelli di Feshbach e Feshbach (Levorato & Nesi, 2001), che hanno osservato un collegamento con lo sviluppo delle abilità di lettura. Infine l'empatia contribuisce a dare motivazione alla lettura, è fonte di piacere e, oltre ad aiutare a comprendere la soggettività altrui, permette di conoscere meglio anche la propria. Il lettore, individuando il collegamento tra stati interni ed eventi nella narrazione, comprende indirettamente come funzionano le proprie emozioni (Levorato & Nesi, 2001).

In Fiaba Magica questa risposta partecipatoria si manifesta in modo decisamente evidente: il bambino, essendo completamente immerso nella storia e "pilotando" i personaggi, dovrebbe mostrare degli atteggiamenti empatici nei loro confronti.

Processi cognitivi

Se consideriamo gli aspetti fin ora delineati la faccia superiore di una medaglia, voltandola ce ne appare un'altra, altrettanto complessa, costituita dai processi cognitivi: vale a dire gli eventi che si susseguono all'interno della nostra memoria durante (e dopo) la lettura dei testi.

Può risultare in parte sorprendente, ma la memoria umana e quella di un computer condividono una struttura decisamente simile. Per l'uomo viene definita memoria di lavoro, la parte responsabile di riconoscere ed elaborare le informazioni in entrata (le singole frasi di un testo); per i computer la medesima funzione viene svolta (con le debite semplificazioni) dalla memoria RAM. D'altro canto la memoria a lungo termine, che ci permette di "ricordare" i fatti, non si differenzia quasi per nulla dall'hard disk che utilizziamo quotidianamente sul nostro computer. Questa premessa introduce, in modo velato, il vero motivo per cui si attivano i processi cognitivi: le informazioni che leggiamo sono troppo complesse per poter essere memorizzate e gestite dalla memoria così come sono (Levorato & Roch, 2007). Pertanto il cervello avvia un processo di elaborazione delle informazioni che si può descrivere analizzando le interazioni tra le due memorie. Le idee che verranno illustrate di seguito sono il frutto di una rielaborazione delle ipotesi di alcuni importanti psicologi

tra cui Kintsch, Van Dijk (Levorato & Nesi, 2001) e Graessner (Levorato & Roch, 2007).

Le interazioni tra la memoria di lavoro e quella a lungo termine danno vita ad un processo molto articolato che è suddiviso in più fasi:

- 1) la memoria di lavoro analizza le singole parole e proposizioni, unisce alcune sequenze di frasi tra loro e procede ad estrarne un significato (“un riassunto”);
- 2) in seguito viene interpellata la memoria a lungo termine, che fornisce tutte le conoscenze “già presenti” necessarie per la comprensione del testo, vale a dire l’attribuzione di un nuovo significato. Questa operazione viene eseguita facilmente perché le conoscenze sono organizzate sotto forma di una rete di nodi: durante la lettura, a seconda del contesto, si attivano solo le connessioni “di pertinenza”;
- 3) una volta ottenute le informazioni, la memoria di lavoro procede alla costruzione della cosiddetta rappresentazione semantica: un insieme di significati (vale a dire riassunti/informazioni importanti) organizzati in modo gerarchico e coerente in base alla tematica e collegati tra loro mediante relazioni di tipo temporale e causale. Questo tipo di elaborazione non può essere eseguita sull’intero testo in un unico momento perché, come si diceva precedentemente, la memoria non è sufficientemente ampia per farlo. Pertanto la memoria di lavoro si occupa di mantenere attive le informazioni provenienti dalla memoria a lungo termine solo per il tempo necessario ad integrarle con quelle in ingresso e a ricavare il significato di una singola porzione del testo;
- 4) non appena si termina la rappresentazione semantica di una parte di testo, essa viene trasferita immediatamente nella memoria a lungo termine per “fare posto” alle informazioni successive.

A questo punto è interessante sottolineare che il frutto della comprensione del testo sono proprio le rappresentazioni semantiche che vengono immagazzinate nella memoria a lungo termine. Dal momento che le rievocazioni (ciò che noi siamo in grado di ricordare) derivano direttamente da queste, è quindi impossibile ricordare qualcosa che non sia stato compreso. Ma c’è di più: la qualità del processo di comprensione (vale a dire l’attribuzione di un significato più o meno “azzeccato”) influenza la memorizzazione rendendo il ricordo più o meno persistente e fedele al testo. Infine un’altra caratteristica interessante riguarda il fatto che il processo si basa su un continuo aggiornamento, ampliamento e riorganizzazione delle nostre conoscenze (Levorato & Roch, 2007).

La descrizione fin qui riportata ha delineato quasi completamente il funzionamento dei processi cognitivi; rimane da sottolineare un ultimo aspetto: quello dei processi inferenziali. Essi sono fondamentali nella comprensione del testo perché ci permettono di integrare quelle informazioni, concetti, collegamenti che non sono espliciti nello scritto e tramite cui è possibile garantire la coerenza (del testo e delle nostre conoscenze). In particolare i processi inferenziali si suddividono in due gruppi: quelli che forniscono la coerenza locale, vale ad dire quelle semplici inferenze che permettono di collegare tra loro brevi sequenze di frasi. E quelli che garantiscono la coerenza globale: si tratta di inferenze complesse che permettono di identificare i nessi causali, vale a dire connettere tra loro gli eventi della storia. È interessante osservare che il testo narrativo, di cui abbiamo avuto modo di parlare precedentemente, è costituito da un rete di eventi, azioni e stati interni che sono connessi tra loro mediante, proprio, una relazione causale (per semplicità si parla anche di catena causale della storia). Gli eventi che appartengono ad una catena causale aiutano a mantenere la coerenza della storia, vengono considerati più importanti e ricordati di più (Levorato & Nesi, 2001). Questo aspetto ci permette di comprendere, in modo ancora più evidente, perché le narrazioni abbiano un ruolo così fondamentale nella nostra vita.

Durante la descrizione dei processi cognitivi, si è già messo in evidenza il ruolo fondamentale delle conoscenze nella comprensione del testo: in particolare più conoscenze possediamo, migliore è la nostra comprensione (Levorato & Roch, 2007). Consci di questo fatto, diversi ricercatori hanno cercato di approfondire l'argomento cercando di classificare le conoscenze. Un compito di questo genere è decisamente molto ampio da analizzare. Tuttavia, soffermandosi nuovamente sui testi narrativi, si è scoperto che le conoscenze che guidano la comprensione di questi testi sono organizzate in uno schema mentale denominato script (copione, sceneggiatura). Questo schema contiene tutte le azioni che "ci aspettiamo" vengano compiute³² per raggiungere un determinato scopo; un esempio tra i tanti potrebbero essere le azioni che caratterizzano l'evento "andare al ristorante": ordinare al cameriere, mangiare, ecc.. Ma come si formano gli script nella nostra mente? Fin da bambini, in realtà, impariamo inconsapevolmente centinaia di script: per esempio mediante l'ascolto di fiabe e favole, tramite i racconti della vita quotidiana oppure semplicemente partecipando ad azioni di routine (Levorato & Roch, 2007).

L'esistenza degli script porta con sé l'idea che nella nostra mente siano presenti anche delle aspettative sulla forma che una storia dovrebbe assumere: in effetti è

³² Le azioni che costituiscono uno script dipendono, solitamente, dalla cultura di appartenenza.

proprio così, la nostra mente possiede un vero e proprio schema delle storie. Ogni narrazione è sempre caratterizzata da un'ambientazione e da un episodio. Quest'ultimo prevede il susseguirsi di fatti ed avvenimenti che nella nostra mente sono raggruppati in categorie chiamate, appunto, categorie della storia: l'evento iniziale che causa una reazione del protagonista e lo motiva a raggiungere un determinato obiettivo, i tentativi per raggiungerlo, le conseguenze dei tentativi e il finale della storia. Al contrario di quanto potremmo pensare, durante l'età evolutiva lo schema delle storie evolve e diventa più flessibile, più completo e più adatto a compiere inferenze; si attua, inoltre, un processo di differenziazione da altri tipi di testi (descrittivi, scientifici, argomentativi, ecc.) a cui vengono associati naturalmente schemi differenti (Levorato & Nesi, 2001).

I processi cognitivi nel loro complesso delineano il contesto generale in cui opera l'interfaccia Fiaba Magica contribuendo, in parte, ad evidenziare le sue peculiarità. Inoltre la dinamica strutturata di Fiaba Magica che prevede, come illustrato nella sezione 2.1.1, la suddivisione della storia in tre parti, ognuna delle quali è caratterizzata dal dialogo di due personaggi, dovrebbe facilitare il lettore a richiamare "nella propria mente" le diverse categorie della storia.

Aspetti sia affettivo-emotivi sia cognitivi

Infine ci sono due ultimi argomenti che si trovano a cavallo tra gli aspetti affettivo-emotivi e quelli cognitivi e, forse anche per questo, influenzano in modo notevole la comprensione del testo: l'interesse e il piacere.

L'interesse del lettore determina, dal punto di vista cognitivo, uno stimolo ad elaborare le informazioni disponibili e a ricercarne di nuove. Inoltre stimola il lettore emotivamente perché lo porta a riconoscere le informazioni che legge come rilevanti (Levorato & Nesi, 2001).

Anche il piacere ha un ruolo duplice: la lettura determina di per sé un piacere della mente, perché fornisce una collezione di emozioni distribuite nel tempo. Inoltre il nostro cervello trova piacere e attivazione emotiva nel ricercare continuamente il consolidamento e l'ampliamento delle rappresentazioni semantiche (definite nel paragrafo precedente) (Levorato & Nesi, 2001).

Uno degli obiettivi principali di Fiaba Magica è quello di coinvolgere l'utente durante l'interazione in modo tale da stimolarne l'apprendimento: è chiaro, quindi, che in questo contesto l'interesse e il piacere sono due aspetti che vengono ricercati attivamente.

Difficoltà di comprensione

La comprensione dei testi, come già detto in apertura, è una competenza complessa alla quale concorrono diverse abilità di natura sia linguistica, sia cognitiva. Di conseguenza anche le difficoltà di comprensione saranno dovute a più fattori. Sono stati svolti diversi studi per cercare di individuare quali possano essere; i risultati hanno messo in luce la compartecipazione di diverse abilità (Levorato & Roch, 2007) tra cui:

- l'intelligenza generale: in particolare le abilità di tipo verbale;
- la conoscenza lessicale: possedere un vocabolario di parole più o meno ricco;
- la competenza sintattica: comprendere le singole frasi;
- le abilità inferenziali: connettere due parti di testo che sono separate da alcune frasi; possedere e saper metter in relazione la propria conoscenza con le informazioni presenti nel testo;
- la capacità di inibire informazioni irrilevanti, evitando un sovraccarico della memoria di lavoro;
- le abilità metacognitive: saper monitorare il proprio processo di comprensione e mettere in atto delle strategie per poterlo migliorare;
- la memoria: la capacità di ricordare e contemporaneamente elaborare le informazioni in ingresso;
- il riconoscimento della struttura del testo: il cosiddetto schema delle storie, di cui abbiamo già avuto modo di parlare precedentemente;
- le difficoltà di comprensione del testo orale che hanno una ripercussione diretta anche sulla comprensione di quello scritto.

In questo contesto Fiaba Magica, con la suddivisione della narrazione in tre parti e l'ulteriore separazione dei dialoghi dei personaggi, può, in parte, facilitare il bambino nella comprensione della struttura del testo, lo può stimolare a creare inferenze, ad acquisire nuove conoscenze oppure ancorarle a quelle esistenti.

L'insegnamento della lingua straniera

In continuità con le ipotesi di Bruner³³, la psicologa Taeschner (1993) ha proposto un modello psicolinguistico per l'insegnamento della lingua straniera nella scuola dell'infanzia e primaria, denominato Format.

L'idea da cui si parte è la seguente: i bambini imparano una lingua in modo più efficace se il contesto in cui sono inseriti è naturale e spontaneo. La psicologa ha ritrovato questa "naturalità" nei vissuti a carattere ripetitivo che madre e figlio condividono durante l'età evolutiva (es. bagnetto, cambio pannolini, ecc.) e ha cercato di riproporre la medesima struttura (chiamata appunto format) a scuola. Il format, pertanto, è un'avventura in lingua straniera condivisa da insegnante e bambini. Essa assume una connotazione potremmo dire ripetitiva perché gli alunni ripercorrono la stessa storia in modalità diverse: in primo luogo sono i protagonisti di una mini rappresentazione teatrale in cui tutti i bambini, contemporaneamente, impersonano i vari personaggi (facendo uso abbondante di espressioni mimiche-gestuali). Successivamente la possono riascoltare sotto forma di *mini-musicals*, nei quali i cantanti sono di lingua madre: in questo modo imparano la pronuncia corretta delle parole e delle frasi. Infine le storie vengono fissate visivamente mediante dei fumetti e dei lucidi.

Fiaba Magica riprende, in parte, le teorie sopra esposte, cercando di far rivivere al bambino un'avventura. Una possibile evoluzione dell'interfaccia potrebbe permettere, inoltre, un'interazione multi-utente che andrebbe nella medesima direzione ipotizzata dalla Taeschner.

L'insegnamento della lingua straniera, d'altra parte, è uno degli obiettivi che si propone di raggiungere l'interfaccia Play&Learn. Come descritto precedentemente, essa costituisce la naturale evoluzione di Fiaba Magica, proponendo di costruire un percorso formativo in cui il bambino impara nuove parole rispondendo "con il corpo" a delle domande.

2.2.2 Motivazioni cliniche

All'interno del panorama complessivo, altre motivazioni interessanti riguardano l'ambito clinico. In particolare, da questo punto di vista, soffermeremo l'attenzione solo sull'interfaccia Fiaba Magica, perché, come già detto in precedenza, essa riguarda il supporto a studenti in situazione di multi-disabilità. Delineeremo il profilo clinico di questi utenti evidenziando, durante la trattazione, i principali punti di

³³ Le ipotesi di Bruner sono state illustrate in dettaglio nel paragrafo relativo agli aspetti affettivo-emotivi che intervengono nella comprensione del testo.

contatto con l'interfaccia; ciò dovrebbe aiutare il lettore a comprendere perché sia sensato che l'interazione avvenga secondo le modalità descritte nella sezione 2.1.1.

Gli studenti in situazione di multi-disabilità, dal punto di vista clinico, sono soggetti principalmente a sindrome da disequilibrio, anartria e ritardo mentale (Zanolla, 2008). Queste tipologie di disturbi causano difficoltà su più fronti: quello relativo alla comunicazione verbale e quello delle capacità motorie (Zanolla et al., 2011).

La comunicazione verbale risulta essere particolarmente limitata, in quanto può realizzarsi solo mediante modulazioni vocaliche e gesti di indicazione (Zanolla et al., 2011). Ciò provoca un grosso impedimento: i bambini non riescono a soddisfare i propri bisogni primari mediante i naturali gesti espressivi o lallazioni e, di conseguenza, non sviluppano la cosiddetta intenzionalità comunicativa, vale a dire la comprensione che i propri messaggi/gesti producono degli effetti (Searle, 1983).

L'unico aspetto semi-positivo riguarda l'ambito della comprensione: pur non avendo buone capacità a comprendere frasi e discorsi, l'utilizzo dei gesti di indicazione permette la comprensione del materiale iconico (Zanolla, 2008).

Fiaba Magica si inserisce in questo contesto come uno strumento in grado di fornire una modalità di espressione alternativa (corporea, gestuale) che permetta, alle persone in situazione di multi-disabilità, di comunicare ed esprimere le proprie intenzioni. Inoltre l'utilizzo abbondante di materiale iconico (i personaggi che si animano nelle scene della storia) risulta essere decisamente efficace in base alle motivazioni sopra esposte.

La seconda area problematica riguarda le difficoltà motorie. In questo contesto si riscontra l'incapacità di portare a termine compiti motori in modo efficace, alcuni disturbi dell'equilibrio dinamico (con relative alterazioni dei movimenti) e diverse difficoltà nella deambulazione (Zanolla, 2008).

I gesti motori, come si può facilmente comprendere, sono fondamentali nella vita di ogni persona e si basano, in generale, sulla fusione di qualità fisiche (resistenza, forza, velocità) e capacità coordinative. Per le persone in situazione di multi-disabilità, pertanto, è necessario creare degli esercizi di riabilitazione specifici che, ripetuti più volte, possano permettere di lavorare su entrambi gli aspetti.

In particolare, uno dei metodi più naturali per poter sviluppare la coordinazione è quello di utilizzare il gioco. Questo svolge un ruolo fondamentale nella crescita dei bambini: infatti ogni gioco permette di vivere delle esperienze attraverso cui si sviluppano moltissime abilità tra cui quelle motorie, cognitive e affettive (Romano, 2010).

Fiaba Magica, mediante il movimento avanti/indietro e l'apertura delle braccia, offre un incentivo allo spostamento autonomo, permette l'acquisizione dei concetti spazio-temporali e aiuta a prendere consapevolezza del gesto che si sta eseguendo. Inoltre, dal momento che è stato costruito sotto forma di gioco, può essere di aiuto anche nello sviluppo della coordinazione.

2.2.3 *Ambito normativo*

Infine l'aspetto normativo completa il quadro delle motivazioni, illustrando gli obblighi che gli istituti scolastici devono garantire in materia di apprendimento.

Vista l'ampiezza dell'argomento, si è deciso di focalizzare l'attenzione su due aspetti: in primo luogo sull'orientamento dell' "Agenzia Nazionale per lo sviluppo dell'autonomia scolastica"³⁴ (ANSAS) riguardo le tecnologie digitali. Successivamente si è spostata l'attenzione sui diversamente abili, analizzando il contenuto della Legge n. 170 del 08/10/2010³⁵ e i consigli dell' "Agenzia Europea per lo sviluppo dell'Educazione Speciale".

L' "Agenzia Nazionale per lo sviluppo dell'autonomia scolastica" è un organo atto a promuovere la ricerca educativa e l'innovazione didattica nella scuola. Tra i molteplici obiettivi, ha lo scopo di: "promuovere l'innovazione nel campo delle tecnologie didattiche, con progetti di formazione, di sostegno alla diffusione delle tecnologie digitali nelle scuole, di ricerca in ambito nazionale e internazionale sull'evoluzione degli ambienti di apprendimento"³⁶. In questo contesto, l'introduzione per esempio della Lavagna Interattiva Multimediale è solo uno dei molteplici progetti che sono stati attuati.

La Legge n. 170 del 08/10/2010 definisce, in modo chiaro, i diritti degli studenti con diagnosi di DSA, vale a dire con dislessia, disgrafia, disortografia e discalculia. La normativa naturalmente è molto ampia, tuttavia è possibile individuare al suo interno due tematiche principali: in primo luogo si afferma la necessità di mettere in atto ogni possibile azione utile a garantire il diritto all'istruzione. Tra le tante ricordiamo la necessità di fornire una formazione adeguata agli studenti, favorire il

³⁴ Per dovere di precisione, dal 1° settembre 2012 l' "Istituto Nazionale di Documentazione, Innovazione e Ricerca Educativa" (INDIRE) è subentrato all' "Agenzia Nazionale per lo sviluppo dell'autonomia scolastica" (ANSAS). Essendo un cambiamento recente, si è deciso di mantenere in questo lavoro di tesi il vecchio riferimento, consapevoli del fatto che gli obiettivi perseguiti dall'ANSAS verranno mantenuti e ampliati anche dal nuovo Istituto.

³⁵ Legge 8 ottobre 2010, n. 170, Nuove norme in materia di disturbi specifici di apprendimento in ambito scolastico.

³⁶ Atto di indirizzo 3 ottobre 2011, n. 89 emanato dal Ministero dell'Istruzione, dell'Università e della Ricerca. Disponibile online sul sito internet: http://hubmiur.pubblica.istruzione.it/alfresco/d/d/workspace/SpacesStore/04910a45-30c7-4d88-b58f-d85a3236f23f/atto_di_indirizzo89_11.pdf.

successo scolastico, incrementare la comunicazione (tra scuola, famiglie, servizi) e tanto altro ancora. D'altro canto, per raggiungere l'obiettivo di cui sopra, si afferma la necessità di utilizzare ogni tipo di strumento dispensativo e compensativo (vale a dire mezzi di apprendimento alternativi e tecnologie informatiche) in grado di rendere la didattica il più possibile personalizzata.

L' "Agenzia Europea per lo sviluppo dell'Educazione Speciale", in continuità con la Legge n. 170, amplia per così dire il "raggio di azione", affermando la necessità di creare politiche educative di integrazione per gli studenti con bisogni educativi speciali. Come già delineato nelle motivazioni di contesto (sezione 1.2.5), ciò può avvenire solo se gli insegnanti hanno a disposizione diverse tipologie di supporti: personale scolastico ausiliario (insegnanti di sostegno), nuovi materiali di insegnamento, corsi di formazione adeguati e nuovi strumenti tecnologici (Camurri et al., 2011).

Fiaba Magica e Play&Learn costituiscono "attività di ricerca in ambito nazionale riguardo l'evoluzione di ambienti di apprendimento". Pertanto, si inseriscono in perfetta continuità con le disposizioni dell' "Agenzia Nazionale per lo sviluppo dell'autonomia scolastica".

Inoltre Fiaba Magica, facendo riferimento alla Legge n. 170 e ai suggerimenti dell' "Agenzia Europea per lo sviluppo dell'Educazione Speciale", è definibile come uno strumento compensativo e dispensativo in grado di migliorare l'apprendimento a studenti in situazione di multi-disabilità.

CAPITOLO 3

Descrizione tecnica

Nel presente capitolo verrà illustrato il funzionamento tecnico delle interfacce Fiaba Magica e Play&Learn, delineando gli aspetti hardware e software.

3.1 Hardware

3.1.1 Requisiti hardware

Le due applicazioni richiedono l'utilizzo, da parte dell'utente finale, di un Personal Computer di fascia medio-bassa che soddisfi i requisiti minimi (Microsoft Corporation, 2012c) riportati in Tabella 3.

Tabella 3. Requisiti hardware minimi.

Processore	<ul style="list-style-type: none">• Dual-core 2.66-GHz o superiore• 32 bit (x86) o 64 bit (x64)
Memoria RAM	2 GB
Scheda video	Supporto DirectX 9.0c
Audio	Scheda audio con relative casse acustiche

Il computer, inoltre, deve essere equipaggiato con un sensore a scelta tra i seguenti:

- Microsoft Kinect for Windows Sensor;
- Microsoft XBOX 360 Kinect Sensor.

Infine, per aumentare il coinvolgimento dell'utente, si consiglia di proiettare l'immagine a parete utilizzando un videoproiettore.

3.2 Software

3.2.1 Requisiti software

I principali requisiti software (Microsoft Corporation, 2012c) sono riportati in Tabella 4.

Tabella 4. Requisiti software.

Sistemi operativi supportati	<ul style="list-style-type: none"> • Windows 7 • Windows 7N, 7 KN (è richiesto il Media Feature Pack) • Windows 8 • Windows Embedded Standard 7 • Windows Embedded POSReady 7
Software necessario	<ul style="list-style-type: none"> • .NET Framework 4 o 4.5 • Kinect Runtime (per il sensore Microsoft Kinect for Windows); Kinect For Windows SDK (per il sensore Microsoft XBOX360 Kinect)

3.2.2 Schema generale dell'interfaccia

Il funzionamento delle applicazioni Fiaba Magica e Play&Learn può essere sintetizzato mediante lo schema generale riportato in Figura 18. In particolare esse sono costituite da due moduli: il primo, sviluppato in linguaggio C++, permette l'acquisizione e l'elaborazione dei dati provenienti dal sensore Kinect. Il secondo,

scritto in linguaggio ActionScript 3 (piattaforma Flash), gestisce i contenuti audio e video. Il software FLOSC (Google Code, 2011) rende possibile la comunicazione tra i due moduli.

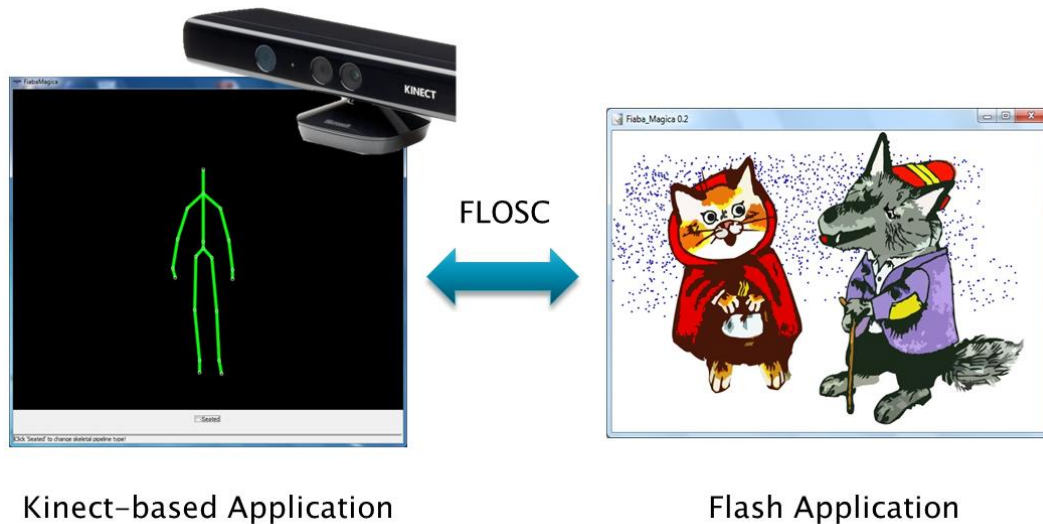


Figura 18. Schema generale delle interfacce.

Riprendendo quanto descritto nella sezione 2.1, è interessante sottolineare che le due applicazioni richiedono all'utente di eseguire la medesima tipologia di movimenti (spostamento in avanti/indietro e apertura delle braccia); tuttavia hanno dinamiche di interazione e contenuti molto differenti.

Ciò implica, dal punto di vista tecnico, che l'implementazione del modulo Kinect-based sia identica per entrambe, mentre lo sviluppo dell'applicazione Flash sia differente. Nella trattazione che segue verranno illustrati i dettagli relativi ad entrambi i moduli e nel secondo si differenzierà la descrizione in base all'interfaccia.

3.2.3 Applicazione Kinect-based

Il primo modulo è stato sviluppato utilizzando il paradigma di programmazione *event-based*. Gli eventi principali che vengono gestiti sono:

- l'evento che segnala l'arrivo di nuovi dati provenienti dal sensore³⁷;

³⁷ Il sensore Kinect è in grado di fornire nuovi dati con una frequenza di aggiornamento pari a 30 frame per second (fps). Un pacchetto di dati contiene informazioni relative a sei persone: le due persone che si trovano in primo piano vengono riconosciute completamente (vale a dire si riesce ad identificare le coordinate x, y, z di venti articolazioni che formano il loro scheletro). Eventuali altre persone (per un totale di quattro) vengono identificate solo parzialmente, vale a dire se ne determina solo il baricentro.

- gli eventi generati dall'interazione dell'utente con la finestra principale dell'applicazione (es. pressione dei tasti, modifica dei parametri, chiusura della finestra, ecc.).

Ai fini di comprendere il flusso logico del programma è interessante soffermare l'attenzione sulla prima tipologia di eventi. Nel prossimo paragrafo, approfondiremo in dettaglio questo argomento.

Schema dell'applicazione

L'arrivo di nuovi dati dal sensore (che genera il relativo evento) avvia un processo di elaborazione che si articola in quattro fasi: la ricezione dei dati, l'estrazione delle features, la creazione di pacchetti in formato OSC e l'invio di questi ultimi via UDP (si veda Figura 19).

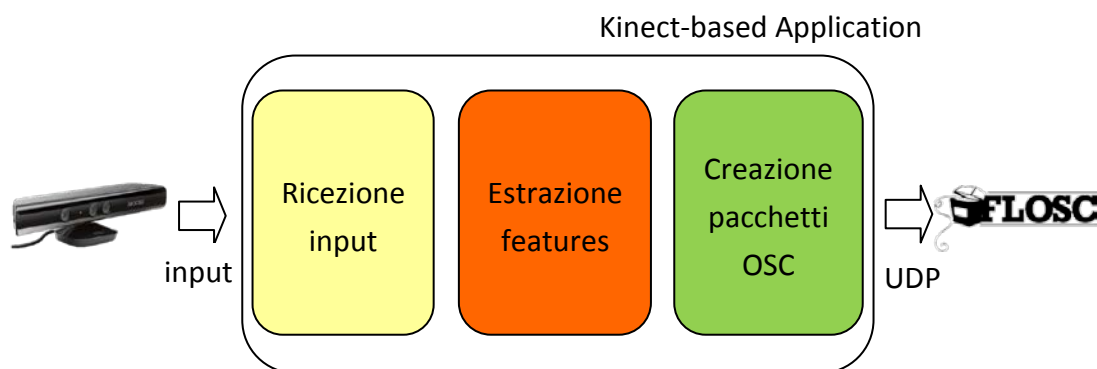


Figura 19. Schema applicazione Kinect-based.

Ricezione degli input

La ricezione dei dati si articola in tre ulteriori fasi. In primo luogo si ottengono dal sensore i nuovi dati (mediante il metodo `NuiSkeletonGetNextFrame`).

```
NUI_SKELETON_FRAME skeletonFrame = {0};

// è arrivato il nuovo frame (contenente i dati degli skeleton)
// lo prendo in carico
HRESULT hr = m_pNuiSensor->NuiSkeletonGetNextFrame(0, &skeletonFrame);
```

Successivamente si procede con il loro filtraggio (mediante il metodo `NuiTransformSmooth`³⁸).

```
// filtro i dati che ho ricevuto
m_pNuiSensor->NuiTransformSmooth(&skeletonFrame, NULL);
```

³⁸ Il metodo `NuiTransformSmooth` si basa sull'*Holt Double Exponential Smoothing method*: un metodo utilizzato per lo più per analisi statistiche di dati finanziari che permette di ottenere i dati filtrati con una latenza limitata (Microsoft Corporation, 2012a).

Infine si analizzano i dati ricevuti e si verificano quali scheletri sono stati identificati dal sensore (NUI_SKELETON_TRACKED)³⁹. Per tutti gli scheletri riconosciuti si eseguono le operazioni che verranno illustrate nel prossimo paragrafo.

```
// individuo gli scheletri riconosciuti dalla kinect
for (inti = 0 ; i< NUI_SKELETON_COUNT; ++i)
{
    NUI_SKELETON_TRACKING_STATE trackingState =
        skeletonFrame.SkeletonData[i].eTrackingState;

    if (NUI_SKELETON_TRACKED == trackingState)
    {
        // estrazione delle features (paragrafo successivo)
    }
}
```

Estrazione delle features

I dati precedentemente ottenuti vengono rielaborati al fine di ottenere le seguenti features: la posizione dell'utente e il riconoscimento dell'apertura del braccio destro e sinistro.

La posizione dell'utente

Premessa

Il sistema di riferimento utilizzato dal sensore Kinect è centrato sulla posizione del dispositivo stesso (si veda Figura 20).

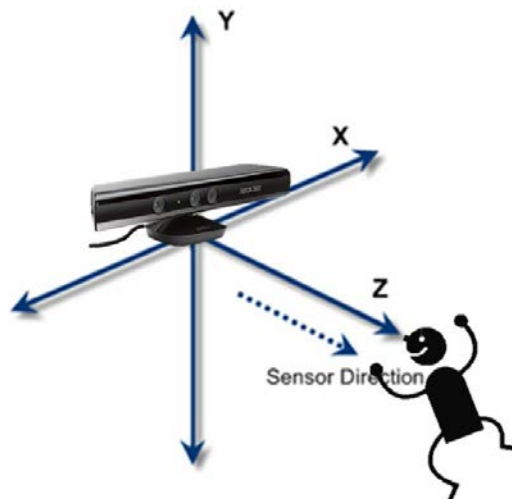


Figura 20. Sistema di riferimento del sensore Microsoft Kinect.

³⁹ Il pacchetto di dati che restituisce il sensore ha sempre una struttura fissa: vale a dire prevede lo spazio per tutti i dati relativi ai possibili sei individui (due scheletri e quattro baricentri, come spiegato precedentemente). Nel caso in cui il sensore non rilevi alcuni individui, i relativi scheletri vengono segnalati come non riconosciuti (NOT TRACKED) e le coordinate delle articolazioni rimangono naturalmente nulle.

Pertanto la coordinata z di un qualsiasi punto dello scheletro corrisponde alla distanza dell'utente dal sensore. Si è scelto di utilizzare la testa (NUI_SKELETON_POSITION_HEAD) come punto di riferimento.

```
head[2]=(int)(skeletonFrame.SkeletonData[i].  
              SkeletonPositions[NUI_SKELETON_POSITION_HEAD].z*1000);
```

In questo modo l'implementazione è svincolata dall'utilizzo della parte bassa dello scheletro (dalle gambe, per esempio): questo aspetto è decisamente importante perché permette l'utilizzo dell'applicazione anche da parte di persone diversamente abili che si trovano su sedia a rotelle. Questa modalità di utilizzo alternativa è stata nominata modalità seduto e può essere attivata mediante un apposito tasto (si veda Figura 21).

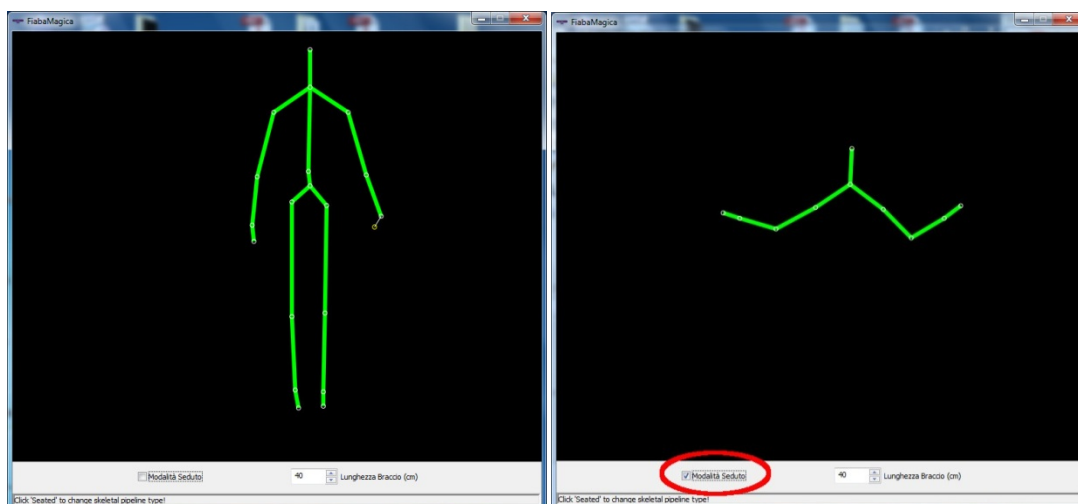


Figura 21. Da sinistra a destra: a) modalità normale; b) modalità seduto.

Determinazione della posizione

Lo spazio di gioco è stato suddiviso in tre zone, utilizzando come limiti superiori e inferiori il campo visivo del sensore (Microsoft Corporation, 2012b) (si veda Figura 22).

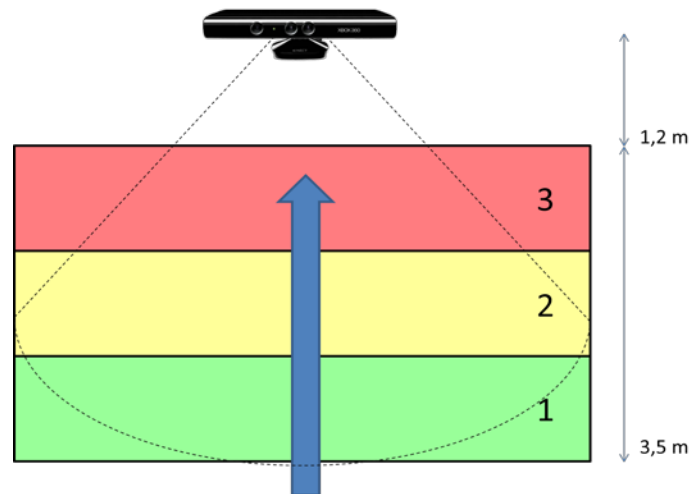


Figura 22. Campo visivo del sensore Microsoft Kinect.

Sulla base di questa mappa, il programma verifica se l'utente si trova in una delle tre zone e, in caso affermativo, procede con la creazione di un messaggio opportuno (si veda il paragrafo "Creazione dei pacchetti OSC").

```

distanceMin = 1200;
distanceMax = 3500;
zones = 3;
zoneWidth = (int)(distanceMax - distanceMin) / zones;

// zona più vicina
if ( (distanceMin < head[2]) && (head[2] < distanceMin + zoneWidth) )
{
    // creazione del messaggio in formato OSC
}

// zona di mezzo
if ( (distanceMin + zoneWidth < head[2]) &&
      (head[2] < distanceMin + 2*zoneWidth) )
{
    // creazione del messaggio in formato OSC
}

// zona più lontana
if ( (distanceMax - zoneWidth < head[2]) && (head[2] < distanceMax) )
{
    // creazione del messaggio in formato OSC
}

```

L'apertura del braccio destro (o sinistro)

L'analisi dell'apertura del braccio destro si basa sull'utilizzo dei due punti evidenziati in Figura 23 (ciò vale simmetricamente anche per il braccio sinistro). In particolare si tratta della:

- mano destra (NUI_SKELETON_POSITION_HAND_RIGHT);
- spalla destra (NUI_SKELETON_POSITION_SHOULDER_RIGHT).

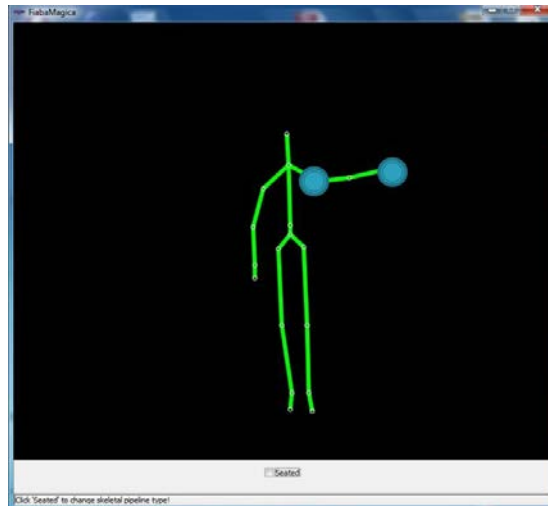


Figura 23. Riconoscimento dell'apertura del braccio.

In primo luogo si estrae la coordinata x di entrambi i punti (si faccia riferimento a Figura 20).

```
handLeft[0]=(int)(skeletonFrame.SkeletonData[i].
                  SkeletonPositions[NUI_SKELETON_POSITION_HAND_RIGHT].x*1000);

shoulderLeft[0]=(int)(skeletonFrame.SkeletonData[i].
                      SkeletonPositions[NUI_SKELETON_POSITION_SHOULDER_RIGHT].x*1000);
```

Successivamente si calcola la differenza delle due coordinate, determinando in questo modo la lunghezza del braccio. Per comprendere se il braccio risulta aperto o chiuso tale lunghezza viene confrontata con un parametro di riferimento (armLength).

```
armLength = 400;

// se la lunghezza del braccio > armLength -> braccio aperto
if ( (handRight[0] - shoulderRight[0]) > armLength )
{
    // creazione del messaggio in formato OSC (paragrafo successivo)
}
else // altrimenti il braccio è chiuso
{
    // creazione del messaggio in formato OSC (paragrafo successivo)
}
```

Il parametro è stato configurato di default a 40 cm, tuttavia è possibile modificarlo in qualsiasi istante operando sulla finestra principale dell'applicazione (si veda Figura 24).

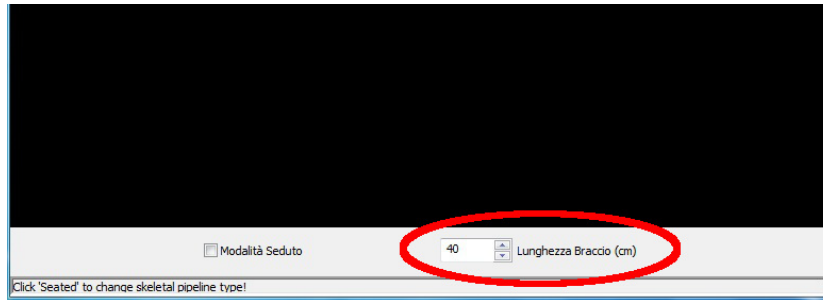


Figura 24. Configurazione del parametro lunghezza braccio.

Creazione dei pacchetti OSC

Le features estratte precedentemente determinano la creazione di altrettanti messaggi che sono costruiti secondo lo standard OSC (The Center For New Music and Audio Technology (CNMAT), 2002).

Riportiamo, qui di seguito, un esempio di creazione di pacchetto in cui si comunica che l'utente è entrato in zona 2 (analoghe considerazioni valgono naturalmente per i pacchetti relativi all'apertura delle braccia).

```
zone = gcnew OSCPacket<float>( "/zoneNumber", (float)2 );
```

La classe OSCPacket implementa gli standard OSC e, in particolare, permette di gestire messaggi dotati di un solo argomento.

```
template <typename T>
ref class OSCPacket
{
public:
    // metodi

private:
    String ^address;
    String ^typeTag;
    T argument;
}
```

Il tipo di argomento (int, float o string⁴⁰) viene definito dall'utente al momento della creazione dell'istanza della classe, sfruttando i template.

```
zone = gcnw OSCPacket<float>( "/zoneNumber", (float)2 );
```

Invio dei pacchetti

L'ultima fase esegue l'invio del pacchetto, precedentemente creato, via UDP.

```
udp->send( zone->toBytes() );
```

Il metodo send è stato implementato mediante le librerie presenti nel pacchetto Microsoft .NET Framework 4.0 (Microsoft Corporation, 2012d). Qui di seguito si riporta un estratto del codice.

```
#using <System.dll>
using namespace System;
using namespace System::Net;
using namespace System::Net::Sockets;
using namespace System::Text;
using namespace System::Runtime::InteropServices;

void UDPClient::send(array<Byte>^ msgBytes)
{
    (...)

    udpclient->Send(msgBytes, msgBytes->Length,
                   (gcnw String(hostname))->ToString(), port);

    (...)
}
```

3.2.4 Applicazione Flash

Fiaba Magica e Play&Learn sono caratterizzate, come già detto in precedenza, da due implementazioni differenti del modulo Flash. Nella trattazione che segue verranno illustrate, pertanto, separatamente.

Fiaba Magica

L'applicazione Fiaba Magica è costituita da:

- un'interfaccia grafica tramite cui l'utente è in grado di scegliere quale storia selezionare⁴¹;

⁴⁰ La classe, per come è stata implementata, può accettare qualsiasi tipo di argomento. Tuttavia i tipi testati sono stati int, float e string.

⁴¹ Nell'implementazione attuale è stata costruita solamente una storia (Cappuccetto Rosso); tuttavia, mediante l'inserimento di nuove animazioni (audio e video), è possibile creare nuove storie senza modificare in alcun modo il programma.

- una componente che gestisce l'interazione dell'utente durante il divenire della storia: in termini tecnici l'arrivo dei pacchetti (contenenti le informazioni relative alla zona occupata e all'apertura delle braccia) e la riproduzione dei relativi feedback.

La componente che riveste maggiore importanza è sicuramente la seconda, di cui andremo a delineare i dettagli implementativi.

Schema dell'applicazione

All'arrivo dei dati da FLOSC, vengono eseguite tre operazioni: la ricezione degli input, la decodifica dei pacchetti ricevuti e la riproduzione dei feedback audio/video (opportuni) (si veda Figura 25).

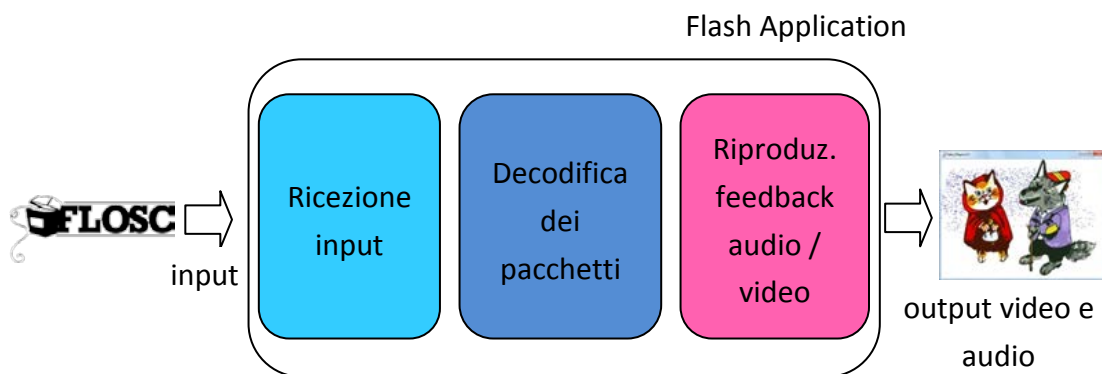


Figura 25. Schema generale dell'applicazione Flash.

Ricezione input

I messaggi OSC provenienti da FLOSC vengono ricevuti mediante l'apertura di un socket di ricezione su una porta prestabilita (9998). All'arrivo di un messaggio viene richiamata la funzione `dataReceived` che si occupa di decodificarlo in maniera opportuna (si veda paragrafo successivo).

```
oscConnection.mSocket.addEventListener(DataEvent.DATA, dataReceived);

function dataReceived(e:DataEvent):void
{
    // decodifica dei pacchetti (paragrafo successivo)
}
```

Decodifica dei pacchetti

In base al tipo di messaggio ricevuto, si identifica se l'utente è avanzato di zona oppure se ha aperto il braccio sinistro (o destro). Se una delle precedenti condizioni viene soddisfatta, si procede con la riproduzione dei relativi feedback audio/video (paragrafo successivo).

```

// l'utente è avanzato nella zona successiva
// e ha già animato entrambi i personaggi
if ( (oscConnection.zonaNumero == zone + 1) && (leftDone) && (rightDone) )
{
    // riproduzione feedback audio/video (paragrafo successivo)
}

// l'utente ha aperto il braccio sinistro
if ( (oscConnection.braccioSinistro == 1) && (...) )
{
    // riproduzione feedback audio/video (paragrafo successivo)
}

```

Riproduzione feedback audio/video

Prima di procedere con la descrizione di questa parte è necessario fare una premessa. I percorsi dei file audio (o video), utilizzati dall'applicazione, sono stati salvati in opportuni array il cui indice corrisponde alla zona a cui sono associati.

Per esempio, l'array `audioIntroPaths` contiene i percorsi dei file audio (voce del narratore) che vengono lanciati nel momento in cui l'utente entra in una specifica zona: `AudioIntroPath[i]` conterrà il percorso del file relativo alla zona `i`.

In questo modo il programma è in grado di conoscere, in ogni momento, quali file audio o video deve richiamare quando l'utente si trova in una determinata zona.

Questa introduzione ci permette di capire meglio il funzionamento del programma: la riproduzione di un file audio prevede, quindi, l'individuazione del percorso corretto in base alla zona (che avviene secondo la dinamica precedentemente illustrata) e la riproduzione vera e propria con il comando `play`.

```

// imposto il percorso del file audio corretto
// l'indice corrisponde alla zona in cui mi trovo
reqNextIntro = new URLRequest(audioIntroPaths[zone]);
nextIntro = new Sound(reqNextIntro);

// inizio la riproduzione del file audio
channel = nextIntro.play();

```

La riproduzione dei file video funziona in modo decisamente simile: si esegue il caricamento del video associato alla zona corrente e si procede alla visualizzazione a schermo.

```

// imposto le azioni da intraprendere
// quando verrà terminato il caricamento video
loaderLeft.contentLoaderInfo.addEventListener
(Event.COMPLETE, videoLeftLoaded);

```

```
// inizio il caricamento video
loaderLeft.load(new URLRequest(videoLeftPaths[zone]));

// il video è stato caricato, lo visualizzo sullo schermo
function videoLeftLoaded(event:Event):void
{
    loaderLeft.contentLoaderInfo.removeEventListener(Event.COMPLETE,
                                                    videoLeftLoaded);
    addChild( loaderLeft );
}
```

Play&Learn

L'applicazione Play&Learn, in modo analogo a Fiaba Magica, è costituita anch'essa da un'interfaccia grafica e da una componente che gestisce l'interazione.

Tale componente si può definire utilizzando lo stesso schema riportato in Figura 25, precisando però che esistono alcune differenze relative alla decodifica dei pacchetti e alla riproduzione dei feedback audio e video. Le due fasi in questione sono state riportate qui di seguito.

Decodifica dei pacchetti

Le differenze riguardanti la decodifica dei pacchetti rispecchiano la dinamica di interazione di Play&Learn (sezione 2.1.2): si richiede che l'utente risponda correttamente a tre domande prima di poter accedere alla zona successiva.

Questa funzionalità è stata implementata utilizzando il codice sotto riportato, il quale si occupa di caricare la domanda successiva solo se:

- la posizione dell'utente non è cambiata (vale a dire si riceve un pacchetto la cui posizione è la stessa dell'ultima ricevuta);
- l'utente non ha ancora risposto a tutte e tre le domande.

```
// carico la prossima domanda
if ( (oscConnection.zonaNumero == zone) && (question < 3) && (...) )
{
    // riproduzione relativo feedback audio/video
}
```

L'avanzamento alla zona successiva è consentito solo quando l'utente ha risposto correttamente a tutte e tre le domande.

```
// vado avanti alla scena successiva
if ( (oscConnection.zonaNumero == zone + 1) && (correctAnswers == 3) )
{
    // riproduzione relativo feedback audio/video
}
```

```
// hai sbagliato alcune risposte
if ( (oscConnection.zonaNumero == zone) && (question == 3)
      && (correctAnswers < 3) )
{
    // riproduzione relativo feedback audio/video
}
```

L'apertura del braccio sinistro (o di quello destro) permette di rispondere alla domanda corrente.

```
// apertura braccio sinistro
if ( (oscConnection.braccioSinistro == 1) && (question < 3) && (...) )
{
    // riproduzione relativo feedback audio/video
}
```

Riproduzione feedback audio/video

I percorsi dei file audio o video utilizzati dall'applicazione Play&Learn sono stati memorizzati in opportuni array bidimensionali in cui:

- il primo indice corrisponde alla zona;
- il secondo alla domanda.

Accedendo, per esempio, a `AudioIntroPath[i][j]` otterremo pertanto il percorso del file audio che è associato alla domanda `j` che viene chiesta all'utente quando si trova in zona `i`.

In questo modo il programma è in grado di conoscere, in ogni momento, quali file audio o video deve richiamare in base alla zona in cui si trova l'utente ed in base alla domanda a cui deve ancora rispondere.

Il resto dell'implementazione, se non per via di alcune piccole modifiche, è analoga a quella precedente illustrata.

CAPITOLO 4

Sperimentazione

Nel presente capitolo si illustrerà lo stato di avanzamento della sperimentazione, le collaborazioni che sono attualmente in atto e i possibili sviluppi futuri del progetto.

4.1 Sperimentazione

Le interfacce Fiaba Magica e Play&Learn sono attualmente in fase di sperimentazione presso la Scuola Primaria “E. Frinta” dell’Istituto Comprensivo “GORIZIA 1” a Gorizia, sotto la supervisione della dott.ssa Serena Zanolla dell’Università di Udine.

In particolare, la validazione dell’applicazione Fiaba Magica inizierà non appena sarà individuato uno studente in situazione di multi-disabilità che possa usufruire dell’interfaccia.

Per quanto concerne Play&Learn, invece, il protocollo di validazione stilato da Serena Zanolla, in collaborazione con il Centro di Sonologia Computazionale, è già in fase di attuazione. I primi risultati scientifici saranno disponibili nell’arco di alcuni mesi e si baseranno sul confronto tra un gruppo sperimentale e un gruppo di controllo.

Qui di seguito viene riportato il protocollo di validazione opportunamente dettagliato.

4.1.1 Protocollo di validazione dell’applicazione Play&Learn

1. Tipo di sperimentazione: disegno sperimentale a gruppi (*Experimental Design between Subjects*).
2. Soggetti: 23 bambini di classe prima tra i quali, per selezione casuale, sarà formato il gruppo sperimentale e il gruppo di controllo.

3. Ipotesi sperimentale: si intende verificare se l'apprendimento di nuove parole e significati avviene in modo più efficace attraverso la stanza logo-motoria in modalità PlayAndLearn rispetto l'utilizzo di una modalità tradizionale come la presentazione powerpoint.
4. Fattore sperimentale: stanza logo-motoria in modalità PlayAndLearn.
5. Variabile indipendente: utilizzo della stanza logo-motoria in modalità PlayAndLearn.
6. Variabile dipendente: l'apprendimento di nuove parole avviene in modo più efficace attraverso l'utilizzo del fattore sperimentale; la variabile dipendente è la misura della qualità dell'apprendimento dei soggetti in seguito alla somministrazione del fattore sperimentale.
7. Modalità:
 - A) Metodo sperimentale: il gruppo sperimentale utilizzerà la stanza logo-motoria in modalità PlayAndLearn; lo spazio interattivo viene suddiviso in tre zone; la presenza del bambino in ognuna delle zone attiva la riproduzione audio di tre parole (e delle relative definizioni) e la riproduzione video di due animazioni; una delle due animazioni corrisponde correttamente alla definizione mentre l'altra no. Il bambino ascolta le definizioni e sceglie le corrispondenti animazioni alzando un braccio piuttosto che l'altro a seconda della posizione (a destra o a sinistra) dell'immagine sullo schermo.
 - B) Metodo tradizionale: il gruppo di controllo utilizzerà una presentazione powerpoint (ogni slide presenterà due immagini e il file audio della parola/definizione); la scelta dell'immagine corrispondente alla definizione sarà eseguita dal bambino attraverso il gesto dell'indicazione.

Entrambe le modalità di apprendimento prevedono una fase di addestramento e una fase di test. Nella prima fase si somministreranno 9 parole/definizioni semplici che il bambino deve dimostrare di conoscere selezionando l'immagine corretta; la fase di test invece prevede la somministrazione di 9 parole sconosciute e i relativi significati che i bambini devono imparare.

La sperimentazione può prevedere più fasi di test (da definire).

8. Materiale:

I contenuti della sperimentazione (le parole nuove da imparare) sono ricavati dal PPVT test (<http://www.libriscientifici.com/dettview.php?id=11296>); le prime 9 parole del PPVT test sono state individuate dalle insegnanti di classe tra quelle semplici da utilizzare per la fase di addestramento; le 9 parole da utilizzare per la fase di test sono state individuate dalle insegnanti di classe tra quelle proposte dal PPVT test secondo questi criteri:

- parole utili al percorso di apprendimento dei bambini;
- parole che probabilmente i bambini non conoscono.

Si verificherà poi se effettivamente tutti i bambini non conoscono le parole individuate dalle insegnanti (circa 30); a tal fine, utilizzando solo queste parole, ai bambini sarà somministrato il PPVT test. Con la somministrazione del PPVT test si ricaverà la lista delle parole che tutti i bambini non conoscono.

Le definizioni delle parole-test saranno elaborate ad hoc dalle insegnanti di classe perché non esiste un vocabolario da cui si possano trarre delle definizioni sufficientemente semplici (i.e. che non contengano ulteriori parole sconosciute).

9. Valutazione degli apprendimenti: la verifica dell'effettivo apprendimento da parte dei bambini sarà effettuata dopo circa una settimana dalla somministrazione del metodo sperimentale e del metodo tradizionale. Lo strumento di valutazione sarà una scheda cartacea contenente le coppie di immagini (non le stesse, ma simili, utilizzate per il test) delle parole apprese nella fase-test. L'insegnante in classe leggerà il nome e la definizione delle parole e i bambini indicheranno con la matita l'immagine che corrisponde alla definizione. Ogni fase di test sarà seguita da una valutazione degli apprendimenti.

4.2 Collaborazioni

Il presente lavoro di tesi è stato sviluppato presso il Centro di Sonologia Computazionale del Dipartimento di Ingegneria dell'Informazione dell'Università di Padova.

Il progetto ha una connotazione multidisciplinare e ha coinvolto, in qualità di correlatrici, la prof.ssa Maja Roch del Dipartimento di Psicologia dello Sviluppo e della Socializzazione (DPSS) dell'Università di Padova e la dott.ssa Serena Zanolla, dottoranda in Comunicazione Multimediale presso il Dipartimento di Scienze Umane

dell'Università di Udine e afferente, inoltre, al Dipartimento di Matematica e Informatica (DIMI) della medesima università.

4.3 Sviluppi futuri

I progetti Fiaba Magica e Play&Learn fanno parte delle molteplici iniziative promosse del Centro di Sonologia Computazionale dell'Università di Padova: come tale sono soggetti a continui aggiornamenti e migliorie.

Alcuni sviluppi futuri potrebbero riguardare la realizzazione di un'interfaccia multimodale multi-utente, costruita sulle basi di Fiaba Magica, in grado di sfruttare l'approccio collaborativo. Dal punto di vista tecnico, a seconda del numero di utenti, questo tipo di interfaccia potrebbe essere realizzata mediante l'utilizzo di uno o più sensori Kinect contemporaneamente.

Conclusioni

Le interfacce Fiaba Magica e Play&Learn, illustrate nel presente lavoro di tesi, forniscono una soluzione *low-cost* al raggiungimento dell'obiettivo iniziale, vale a dire lo sviluppo di un ambiente interattivo multimodale in grado di supportare persone diversamente abili e normodotate.

Le due applicazioni condividono l'approccio multimodale, argomento introdotto facendo riferimento a diversi casi in letteratura, il quale si è rivelato uno strumento eccellente per la risoluzione del problema, mettendo in luce i molteplici vantaggi che si ottengono combinando le informazioni provenienti da diversi canali sensoriali.

La creazione di due applicazioni è stata dettata dall'esigenza di progettare percorsi didattici inclusivi, rivolti sia a persone diversamente abili, sia normodotate, in grado di sviluppare specifiche abilità. Tra i diversamente abili, si è scelto di fornire un supporto a studenti in situazione di multi-disabilità; tuttavia, modificando le dinamiche di interazione, si potrebbe fornire sostegno a qualsiasi altra tipologia di disabilità.

In questo contesto di flessibilità, le applicazioni sono state costruite seguendo due principi fondamentali: la personalizzazione dei contenuti e la modularità. Il primo permette all'utilizzatore di creare nuove storie (che condividono la medesima dinamica di interazione) aggiungendo semplicemente dei contenuti audio e video differenti. La modularità, invece, facilita gli eventuali sviluppi futuri del software.

L'obiettivo *low-cost*, come affermato in apertura, è stato ampiamente raggiunto: le interfacce richiedono l'utilizzo di un personal computer di fascia medio-bassa e l'impiego del sensore Microsoft Kinect, reperibile a costi molto bassi. Il sensore in questione, pur essendo molto economico, possiede al suo interno una tecnologia abbastanza avanzata: un sensore di profondità, una fotocamera RGB e un array di microfoni. Il sensore Microsoft Kinect non è sicuramente paragonabile, per via di precisione, ai sistemi professionali di *Motion Capture*: tuttavia le informazioni che è in grado di fornire sono stati sufficienti a permettere l'interattività delle applicazioni.

Le interfacce sviluppate in questo lavoro di tesi forniscono degli spunti interessanti per poter proseguire sul medesimo percorso. In particolare il sostegno a persone diversamente abili richiama la necessità di adattare i software in base agli obiettivi specifici che gli insegnanti (oppure logopedisti o persone che si occupano di riabilitazione) vogliono perseguire. Ciò si traduce nello sviluppo di applicazioni che permettano di essere sempre di più personalizzabili. In questa direzione, un possibile ampliamento potrebbe riguardare la creazione di sistemi in grado di riconoscere una serie “limitata” di movimenti (per esempio il movimento delle braccia, delle gambe, della testa, ecc.) e di riprodurre *feedback* audio e video che possano essere configurati dall’utente. L’insegnante (o persona che si occupa di riabilitazione) avrebbe la possibilità di creare esercizi specifici decidendo sia la tipologia di gesti da richiedere all’utente, sia i *feedback* audio e video associati. La sfida diventerebbe, in questo modo, la progettazione di esercizi appropriati al raggiungimento di determinati obiettivi (diversi di caso in caso); tuttavia, dal punto di vista tecnico, un’unica interfaccia sarebbe in grado di fornire tutte le funzionalità richieste.

Ulteriori sviluppi futuri potrebbero riguardare la realizzazione di un’interfaccia multimodale multi-utente in grado di sfruttare l’approccio collaborativo. Dal punto di vista tecnico, a seconda del numero di utenti, questo tipo di interfaccia potrebbe essere realizzata mediante l’utilizzo di uno o più sensori Kinect contemporaneamente.

La trattazione si conclude con queste interessanti riflessioni sul futuro che permettono di inserire il lavoro all’interno di un percorso di ricerca in continua evoluzione, con la consapevolezza che alcuni spunti si potranno rivelare utili per la costruzione delle interfacce che verranno.

Bibliografia

Cadiou, A., Karpov, A., & Ronzhin, A. (2006). *A MULTI-MODAL SYSTEM ICANDO: Intellectual Computer AssistaNt for Disabled Operators* [Powerpoint slides]. Retrieved from: <http://www.spiiras.nw.ru/speech/demo/Interspeech2006.ppt>.

Camurri, A., Canazza, S., Canepa, C., Rodà, A., Volpe, G., Zanolla, S., & Foresti, G. L. (2010). The stanza logo-motoria: An Interactive Environment for Learning and Communication. In E. Gómez, P. Herrera, & R. Ramírez (Eds.), *Proceedings of 7th Sound and Music Computing Conference* (pp. 353-360). Barcellona: Universitat Pompeu Fabra.

Camurri, A., Mazzarino, B., Volpe, G., Morasso, P., Priano, F. & Re, C. (2003). Application of Multimedia Techniques in the Physical Rehabilitation of Parkinson's Patients. *Journal of Visualization and Computer Animation*, 14(5), 269-278. doi: 10.1002/vis.324.

Gehlhaar, R., Girao, L. M., & Rodrigues, P. M. (2008). CaDaReMi - An Educational Interactive Music Game. In P. Sharkey, P. Lopes dos Santos, & P. L. T. Weiss (Eds.), *Proc. 7th ICDVRAT with ArtAbilitation*, (pp. 355-360). Maia: Reading University Press.

González González, M. L., Muñoz Muñoz, A., & Valero Duboy, M. A. (2008, June). A Low-Cost Multimodal Device for Web Accessibility. Paper presented at the 6th COLLECTer Iberoamérica, Madrid. Retrieved from: <http://www.imaginar.org/ecollecter/fullpapers/p66-ALowCostMultimodalDevice.pdf>.

Google Code. (2011). Flosc. Retrieved from <http://code.google.com/p/flosc/>

Karpov, A., & Ronzhin, A. (2007). ICANDO: Low-cost multimodal interface for hand disable people. *Journal on Multimodal User Interfaces*, 1(2), 21-29. doi: 10.1007/BF02910056.

Kuber, R., McAllister, G., Murphy, E., Strain, P., & Yu, W. (2005, July). Developing Multimodal Interfaces for Visually Impaired People to Access the Internet. Paper presented at the 11th International Conference on Human-Computer Interaction, Las Vegas. Retrieved from: <http://citeseerx.ist.psu.edu/viewdoc/download?doi=10.1.1.100.2131&rep=rep1&type=pdf>.

Levorato, M. C., & Nesi, B. (2001). Imparare a comprendere e produrre testi. In L. Camaioni (Ed.), *Psicologia dello sviluppo del linguaggio* (pp. 179-211). Bologna: il Mulino.

Levorato, M. C., & Roch, M. (2007). TOR: Test di Comprensione del Testo Orale. Firenze: Giunti O.S..

Microsoft Corporation. (2012a). NuiTransformSmooth Method. Retrieved from <http://msdn.microsoft.com/en-us/library/jj663874.aspx>

Microsoft Corporation. (2012b). Skeletal Tracking. Retrieved from <http://msdn.microsoft.com/en-us/library/hh973074>

Microsoft Corporation. (2012c). System Requirements. Retrieved from <http://msdn.microsoft.com/en-us/library/hh855359.aspx>

Microsoft Corporation. (2012d). UdpClient.Send Method. Retrieved from <http://msdn.microsoft.com/en-us/library/acf44a1a.aspx>

Romano, F. (2010). Progettazione di un ambiente interattivo multimodale da utilizzare nella prassi didattica come strumento compensativo e/o come misura dispensativa (Unpublished master's thesis). Università degli Studi di Udine, Italy.

Searle, J. (1983). *Intentionality: An Essay in the Philosophy of Mind*. Cambridge, England: Cambridge University Press.

Taeschner, T. (1993). *Insegnare la lingua straniera con il Format: Un modello psicolinguistico per la scuola materna ed elementare*. Roma: Anicia.

The Center For New Music and Audio Technology (CNMAT), UC Berkeley. (2002). *The Open Sound Control 1.0 Specification*. Retrieved from http://opensoundcontrol.org/spec-1_0

Zanolla, S. (2008). *La stanza logo-motoria: un ambiente interattivo per la comunicazione e l'apprendimento* (Unpublished master's thesis). Università degli Studi di Udine, Italy.

Zanolla, S., Rodà, A., Romano, F., Scattolin, F., Canazza, S., & Foresti, G.L. (2011). *When Sound Teaches*. In S. Zanolla, F. Avanzini, S. Canazza, & A. De Götzen (Eds.), *Proceedings of 8th Sound and Music Computing Conference* (pp. 64-69). Padova (Italy): Padova University Press.