



MAIS MultiLezi: un ambiente adattabile, multicanale e multimodale per l'e-learning

**MAIS MultiLezi: an adaptable
multichannel and multimodal
environment for e-learning**

Dipartimento di Elettronica e Informazione

Gruppo di ricerca
Research group

Licia Sbattella
Luca Mainetti
Thimoty Barbieri
Antonio Bianchi
Stefano Bruna
Barbara Pernici

Anno di avvio attività
Beginning of activity

2002

Cooperazioni nazionali e internazionali
National and international collaboration

Università di Lecce

Sito web
Website

<http://www.mais-project.it>

L'obiettivo del progetto MAIS (Multichannel Adaptive Information Systems) è stato sviluppare modelli, metodi e strumenti che consentano la realizzazione di sistemi informativi adattativi multicanale, in grado di fornire servizi anche in condizioni di forte variabilità dell'accesso alla rete e dei dispositivi di accesso.

MAIS, come progetto di ricerca di base, ha coinvolto molte Università italiane, fra cui il Politecnico di Milano, sui temi della comunicazione e dell'accesso all'informazione distribuiti, con particolare

The goal of the MAIS (Multichannel Adaptive Information Systems) project has been the development of models, methods and tools that allow the implementation of multichannel adaptive information systems able to provide services with respect to different types of networks and of access devices.

MAIS, as base research project, has involved many Italian Universities, among which Politecnico di Milano, about the themes of distributed communication and computing, taking in particular

attenzione all'interazione di persone in situazione di disabilità con questi sistemi. Le applicazioni di e-learning sono particolarmente adatte a dimostrare l'utilità e il potenziale dei servizi context-aware multicanali e multimodali in un quadro come quello definito per l'architettura di MAIS. La preparazione di materiali didattici e un adeguato ambiente di e-learning, che possano essere usati attraverso i diversi canali e contesti e le modalità differenti di interazione (permettere ad esempio l'accesso anche a utenti con disabilità specifiche, di carattere motorio e sensoriale), richiede il disegno e lo sviluppo di interfacce multimodali, di strutture adattive di navigazione e contenuti sviluppati e progettati in modo tale da poter essere facilmente ristrutturato per i diversi stili di accesso e di interazione.

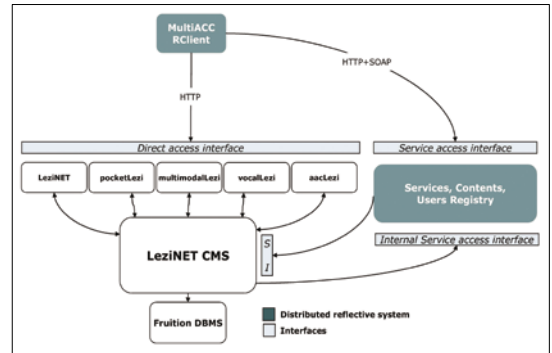
Qui brevemente presentiamo MultiLezi (Figura 1), un prototipo di un ambiente di apprendimento multicanale e multimodale, che utilizza il modello utente di MAIS per profilare i bisogni degli utenti e l'architettura riflessiva di MAIS per fornire all'utente la selezione del canale e della modalità più efficiente per garantire usabilità e accessibilità ottimale del contenuto di apprendimento.

Attualmente i canali supportati dal prototipo sono il Web, il telefono e dispositivi palmari, con la possibilità di utilizzare contenuti testuali e audio-video. È possibile leggere il contenuto testuale su un PDA, ascoltare il prodotto di un VoiceXML attraverso sintesi e riconoscimento vocali che strutturano un sistema di interazione accessibile dal telefono, utilizzare un canale multimodale che si avvale dell'uso del mouse, ma che utilizza anche controllo e sintesi vocale in modalità complementare (multimodale). Qui metteremo a fuoco in particolare l'interazione multimodale desktop e attraverso il telefono.

- MultimodalLezi: un'interfaccia multimodale per l'accesso al sistema di e-learning, che supporta un paradigma standard point-and-click di interazione dell'utente, e la possibilità di pronunciare specifici comandi e richiedere (attraverso TTS) una lettura sintetica della struttura di navigazione dell'ambiente e del contenuto stesso.

- VocalLezi: un'interfaccia di navigazione vocale per l'accesso al contenuto che permette di spostarsi all'interno della struttura ad albero secondo cui

1.
L'architettura generale di
MultiLezi
General architecture of
MultiLezi

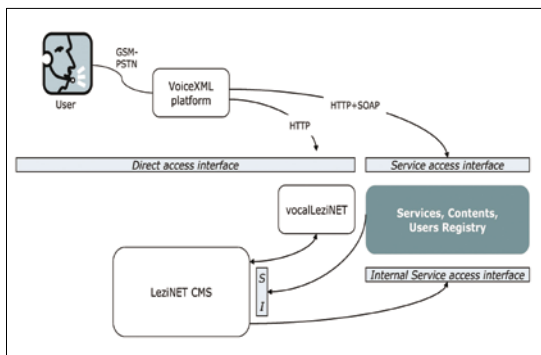


account the interaction with such systems by people in situation of disability.

E-learning applications are particularly well suited to demonstrating the usefulness and potential of multichannel and multimodal context-aware services within a framework such as that defined for the MAIS architecture. Providing teaching materials and a suitable learning environment, which can be used across various channels and contexts, and with different interaction modalities (e.g., to allow access also to users with specific disabilities, such as blindness or motor difficulties), requires the design and development of multimodal interfaces, adaptable navigation structures, and content developed and designed in such a way that it can be easily repurposed for different styles of access and interaction.

Here we briefly present MultiLezi (Figure 1), a prototype of a multichannel, multimodal learning environment, which leverages the MAIS user model to profile the needs of users in a learning environment, and the MAIS reflective architecture to provide the user with the most efficient selection of

2.
 Il servizio di VocalLezi permette all'utente di interagire con MultiLezi attraverso il telefono
The VocalLezi service allows user interaction with MultiLezi over the phone



questo è organizzato, attraverso una normale conversazione telefonica.

MultimodalLezi

Questo canale offre la possibilità di una vera interazione multimodale, unendo la modalità vocale e la normale modalità visiva di point-and-click (attraverso mouse e tastiera). Gli utenti target di questo canale sono utenti con limitata acuità visiva. L'uso di un'interfaccia utente appropriata e accessibile è un requisito fondamentale e deve soddisfare i seguenti criteri.

- Input e output: deve essere possibile interagire con l'interfaccia sia attraverso la voce sia attraverso le periferiche tradizionali come mouse e tastiera.
- Elementi di interfaccia autoesplicativi: le caratteristiche principali dell'interfaccia possono essere lette portando il mouse sopra di loro, secondo una modalità esplorativa. La sintesi vocale legge le etichette dei tasti o dei campi delle form per supportare gli utenti con ipovisione o con dislessia.
- Regolazione del contrasto e della dimensione del font: un elevato contrasto cromatico permette una

channels and access modalities, to guarantee optimal usability and accessibility of the learning content. Currently the channels supported by the prototype are the Web, the telephone, and hand-held devices, with the possibility to use text and audio-video content. It is possible to read text content on a PDA, to listen to VoiceXML with a text to speech/ automatic speech recognition navigation system accessible over the phone, and to use a multimodal channel which allows use of the mouse, but also voice control and speech synthesis in a combined (multimodal) mode. Here we will focus on desktop multimodal and phone interaction:

- MultimodalLezi: a multimodal interface for accessing learning content, which allows a standard point-and-click user interaction paradigm, which can be mixed with the possibility to speak specific commands, and to request a text-to-speech (TTS) synthetic reading of the navigational structure of the environment and of the content itself.
- VocalLezi: a voice-based navigation interface for the content, which provides interaction through a tree-based structure controlled by means of a vocal conversation across a regular telephone.

MultimodalLezi

This channel offers the possibility of a true multimodal interaction, combining the vocal modality and the regular visual point-and-click modality (mouse and keyboard). The target users of this channel are users with vision impairments (in particular, low vision). The use of an appropriate, accessible user interface is a fundamental requirement and must satisfy the following criteria.

- Input and output: it must be possible to interact with the interface both by speech and traditional peripherals as the mouse and the keyboard.
- Self-explaining interface elements: the main features of the interface can be read by rolling the mouse over them. A speech synthesizer reads the labels on buttons or fields to help users with low vision or dyslexia.
- Setting of visual contrast and font size: a high color contrast allows better visualization. Through the use of cascading style sheets, the user can consistently change the color appearance of the

migliore visualizzazione. Attraverso l'uso di fogli di stile l'utente può modificare consistentemente la combinazione di colore dell'interfaccia e del formato del testo.

- TTS: il contenuto può essere letto da un sintetizzatore integrato (nessuna esigenza di un costoso screen reader), per supportare utenti con ipovisione e con dislessia (il testo che viene letto viene inoltre evidenziato per permettere all'utente una migliore attenzione).

VocalLezi

L'uso del normale telefono come mezzo di accesso al sistema estende il numero di utenti che possono accedere al contenuto, sia in termini di contesto (da posizioni in cui un PC non è disponibile) sia in termini di possibilità dell'utente (l'utente potrebbe non sapere utilizzare un PC o potrebbe non volerlo usare). VocalLezi (Figura 2) è in grado di generare dinamicamente una resa alternativa (un processo denominato transcodifica) da pagine Web standard che siano aderenti alle linee guida di riferimento principali di accessibilità disposta in W3C WCAG 1.0. Viene definita un'interfaccia di navigazione del contenuto e della struttura usando il linguaggio VoiceXML, che permette la generazione di menu e la lettura del contenuto attraverso la sola interazione vocale. Nello sforzo di determinare la migliore strategia di accesso al contenuto da offrire all'utente, un metodo semplice e intuitivo di navigazione vocale che possa sostituire la complessa navigazione visuale dei canali sopra descritti, sono state valutate varie strategie di interazione vocale.

Per fornire una resa vocale del contenuto, in particolare per utenti con disabilità cognitive o visive, è stato necessario effettuare diversi esperimenti per verificare il ruolo della semplificazione del testo e della prosodia. Anche se gli attuali sintetizzatori vocali hanno fatto dei significativi progressi sia nella qualità della voce sia nella precisione della resa prosodica, c'è ancora spazio per il miglioramento. Rispetto a questa materia, durante le nostre prime valutazioni si è evidenziato che gli utenti con disabilità visive considerano meno importante la qualità della sintesi vocale, abituati da tempo al livello mediocre di sintesi degli screen reader degli scorsi anni e ai

interface and the text size.

- TTS: content can be read by an integrated synthesizer (no need for an expensive screen reader), to help users with visual impairments or with dyslexia (the content which is being read is also simultaneously highlighted to allow the user to read along).

VocalLezi

The use of a regular telephone as mean of access as a channel broadens the number of users who can access the content, both in terms of context (from places in which a PC is not available) and in terms of the user's abilities (the user may not use a PC or might not know how to use it).

VocalLezi (Figure 2) can dynamically process an alternative rendition (a process called transcoding) from standard Web pages which are compliant with the main accessibility guidelines set forth in W3C WCAG 1.0, creating a navigational structure and content rendition using the VoiceXML language, which allows one to create menus and read content using solely voice-based interaction. In an effort to determine the best strategy to offer the user, which would consist of a simple, intuitive method of vocal navigation that could represent the complex, visual-based navigation of the channels described above, various vocal-interaction strategies have been evaluated.

To provide vocal rendition of content, in particular for users with cognitive or visual disabilities, it was necessary to carry out various experiments to test the role of prosody and text simplification. While current speech synthesizers have made significant advancements both in the quality of the voice and in the precision of the rendition of human prosody in reading text, there is still room for improvement. In relation to this matter, during our early tests, we found that users with visual disabilities consider the quality of the speech to be less important (they get used to the synthesizer of the screen reader and adapt to its flaws), whereas the importance of the quality of speech increases when speech is used as a prosthesis for young readers with dyslexia. A bad synthesis can compromise irreparably the usefulness of the platform.

Text simplification showed a twofold potential:

loro problemi), mentre l'importanza della qualità della sintesi aumenta quando il discorso è usato come supporto per giovani lettori con dislessia. Una sintesi difettosa in questo caso può compromettere irrimediabilmente l'utilità della piattaforma di apprendimento.

La semplificazione del testo ha mostrato un duplice potenziale: (a) il sintetizzatore potrebbe leggere il contenuto con meno errori, sia nella resa della prosodia sia nella pronuncia delle parole complesse o insolite e (b) aumenta la leggibilità e l'accessibilità del contenuto, sia per i lettori con dislessia, sia per quelli con disabilità visiva, rendendo il testo più facile da leggere o navigato da un punto di vista morfologico, fonetico, sintattico e strutturale. I lettori giovani con dislessia e i loro insegnanti hanno inoltre apprezzato il ruolo pedagogico che il sistema ha assunto in letture assistite in attività di apprendimento: l'ambiente tecnologico si è dimostrato meno frustrante rispetto agli ambienti di apprendimento tradizionali e ha motivato l'utente nell'affrontare nuove esercitazioni. La possibilità di fornire contenuto ricco e accessibile in formati multimediali è sembrato essere un ulteriore importante vantaggio. La versione corrente del prototipo richiede una semplificazione strutturale manuale del contenuto testuale originale del corso e usa un indice di complessità basato su GULPEase. Ciò aiuta il redattore del contenuto, cioè la persona incaricata di intervenire sul testo, nel trovare la lunghezza ottimale delle frasi per la semplificazione della lettura. Il redattore inoltre controlla il lessico usato, confrontandolo col vocabolario di base italiano, una lista delle parole più frequentemente usate nella lingua parlata, per ottenere un più alto livello di comprensione.

A questo fine abbiamo aggiunto il vocabolario italiano di base come dizionario per il programma di word processing usato dai redattori dei contenuti. In questo modo si possono ottenere risposte immediate e feedback visivo quando viene utilizzata una parola che non appartiene al dizionario di base. L'indice di leggibilità utilizzato è basato su GULPEase, che prende in considerazione le lunghezze di parola e di frase e a questo si è aggiunto il risultato di una analisi "shallow" dei documenti, una tecnica di analisi che produce rappresentazioni sintattiche piatte (chunks),

(a) the synthesizer could read the content with fewer mistakes, both in the rendition of prosody and in the pronunciation of complex or unusual words and (b) it enhanced the readability and accessibility of the content, both for readers with dyslexia and those with visual disability, making text easier to read or navigated from a morphological, phonetic, syntactical and structural point of view. Young readers with dyslexia and their teachers also appreciated the pedagogical role that the system assumed in tutored reading and learning activities: the technological environment was less frustrating than traditional learning environments and could motivate the user to attempt new exercises. The opportunity to provide accessible, rich content in multimedia formats also appeared to be a major advantage.

The current version of the prototype requires a manual structural simplification of the original textual content of the course, and uses a complexity index based on GULPEase. This helps the content editor, i.e., the person in charge of rewriting the content, to find the optimal length of sentences for simplifying the text. The content editor also checks the lexicon used, matching it to the Italian Basic Vocabulary, a list of the most frequently used words in the spoken language, in order to obtain a higher understandability.

For this purpose, we have added the Italian Basic Vocabulary as a dictionary for the word processor used by the content editors. In this way they can obtain immediate, visual feedback when a word not belonging to the basic dictionary is being used. The readability index that we use is based on GULPEase, which takes phrase and word lengths into account, and uses in addition the results of a shallow parsing of the documents, an analysis technique that produces flat syntactical representations (chunks), with low computational requirements and reliable results. This kind of syntactical analysis produces ten types of chunks: adjectival, predicative, adverbial, subordinating, nominal, prepositional, finite verbal, gerundival, infinitival, and participial. To determine what weights to assign to the presence of various type of chunk, an analysis of a reference corpus of simplified articles was conducted, and coefficients were assigned to each kind of chunk. The indexes

con modesti requisiti di calcolo e buona robustezza dei risultati. Questa analisi sintattica produce dieci tipi di chunk: aggettivo, predicativo, avverbiale, subordinativo, nominale, preposizionale, verbale, gerundio, infinito, participio. Per determinare quali pesi attribuire alla presenza dei diversi tipi di chunk, è stata condotta un'analisi di un corpus di riferimento di articoli semplificati, ricavando coefficienti per ciascuno dei chunk. Gli indici sono stati scalati in modo da poter essere armonizzati con il sistema generale di indicizzazione di GULPEase. Durante le nostre prove di produzione, i redattori potevano riscrivere manualmente il contenuto per realizzare un alto indice di accessibilità, a un tasso di circa 160 parole all'ora, a partire da un documento esistente. Il requisito di leggibilità per il documento dopo la semplificazione manuale operata da un redattore umano è stato fissato in un indice minimo di 65 (più alto l'indice, più leggibile il contenuto). Nei nostri esperimenti iniziali, su un totale di 30.000 parole, i redattori hanno ottenuto facilmente un indice medio di leggibilità di GULPEase di 70. Ciò è stato realizzato lasciando nel testo soltanto il 16% di parole che non appartengono al vocabolario di base italiano. Se confrontiamo la resa della sintesi vocale della versione originale, non semplificata, del documento con quella semplificata, risulta chiaramente evidente che un utente che ascolta la lettura sintetizzata di una pagina per la prima volta ha comprensione del 100% della versione semplificata e un tasso molto più basso di comprensione dello stesso contenuto per la versione originale.

were devised in a way so that they could be harmonized with the general GULPEase indexing system. During our production tests, the editors were able to manually rewrite content to achieve a high accessibility index, at a rate of approximately 160 words per hour, starting from an existing document. After manual simplification by a human content editor, we required the document to have a minimum complexity index of 65 (the higher the index, the more readable the content). In our early experiments, on a total of 30,000 words, content editors easily obtained an average GULPEase readability index of 70. This was achieved by leaving in the text only 16% of words not belonging to the Italian Basic Vocabulary. If we compare the synthesized-speech rendition of a nonsimplified document with that of a simplified one, it is clearly evident that a user who listens to the synthesized reading of a page for the first time has 100% comprehension of the simplified version, and a much lower comprehension rate of the same content in the nonsimplified version.