



FBL S.r.l.
Società di informatica
Mortara PV

VoiceTranscribe 9.5

**Sistema di trascrizione
in diretta e in differita
Versione 9.5**



1 - Un po' di storia

Qualche anno fa ci siamo posti l'obiettivo di realizzare un sistema, destinato alla scuola, per la sottotitolazione di una lezione per alunni sordi. Abbiamo trovato nella Commissione Europea lo stesso interesse ed abbiamo collaborato con il Centro Comune di Ricerca di Ispra al progetto Voice.



Se la scuola ed i problemi degli alunni non udenti restano al primo posto nei nostri obiettivi VoiceMeeting ha dimostrato una versatilità applicativa non ancora del tutto scoperta: è stato utilizzato nella realizzazione di uno sportello informativo che copre le esigenze di disabili, in particolare sordi, ed anziani; è stato proposto come sistema per l'insegnamento delle lingue straniere e come compendio, in aula, per corsi di formazione tecnica tenuti da docenti stranieri.

Nell'esperienza scolastica l'insegnante parla rivolto alla classe utilizzando un microfono. Una telecamera riprende l'insegnante, la lavagna, l'ambiente scolastico. Microfono e telecamera sono collegati ad un personal computer che, utilizzando **VoiceMeeting** crea sul monitor una immagine (quella della ripresa televisiva) sottotitolata dalle parole pronunciate dall'insegnante. L'immagine può essere proiettata su grande schermo e messa a disposizione dell'intera classe.

L'alunno sordo ed i suoi compagni di classe concentrano la loro attenzione sull'immagine creata dal personal computer.

Si raggiunge così l'obiettivo di trasmettere **a tutti gli alunni della classe, le stesse informazioni, con le stesse parole, nello stesso istante.**

Con la versione 9.5 abbiamo introdotto a fianco di **VoiceMeeting** due altri pacchetti applicativi:

VoiceTranscribe

Un applicativo che consente di ottenere un nuovo punto di vista: dotare l'alunno di personal computer portatile ed affidare all'insegnante un microfono.

La lezione sarà trascritta ad uso esclusivo dell'alunno. Il testo generato verrà salvato sul personal computer insieme alla voce dell'insegnante permettendo una revisione dei sottotitoli. L'insegnante avrà sempre la possibilità di chiudere il microfono quando ritiene lo necessario.

Contemporaneamente alla trascrizione in diretta è possibile ottenere una trascrizione in differita: la voce, precedentemente registrata, viene elaborata dal sistema di riconoscimento vocale e trascritta frase per frase, cioè da pausa a pausa. La trascrizione potrà, in seguito, essere controllata, eventualmente corretta, editata. La gestione della frase facilita la rilettura del testo e la sua eventuale correzione.

VoiceReader

Un applicativo che permette di controllare ed eventualmente correggere il testo generato da VoiceMeeting e da VoiceTranscribe ascoltando, in sincronismo, l'audio originale.

VoiceMeeting e VoiceTranscribe utilizzano il sistema di riconoscimento vocale Dragon NaturallySpeaking, VoiceReader non ha prerequisiti.

2 - Requisiti hardware e software di VoiceTranscribe

Prerequisito all'utilizzo di VoiceTranscribe è la presenza sul personal computer (installazione, attivazione e addestramento di almeno un utente) del motore di riconoscimento vocale Nuance™ Dragon NaturallySpeaking 9.5 nelle versioni Preferred o Professional.

La configurazione hardware consigliata è la seguente:

- Processori Intel Pentium IV, Centrino o equivalenti
- Memoria RAM 1Gb o superiore.
- Schede audio di qualità.
- Sistema operativo MS Windows XP Pro, Vista.

Il personal computer può essere integrato con i seguenti strumenti:

- Microfono di qualità: particolarmente consigliati i radiomicrofoni.
- Registratore digitale che supporti i formati wave e wma.

3 – Il prodotto

Il prodotto VoiceTranscribe è inviato all'Utente su CD accompagnato da una unità USB (Rockey2) che ne permette il corretto funzionamento.

4 - Assistenza

L'Utente potrà chiedere informazioni sul prodotto mettendosi in contatto direttamente con

FBL S.r.l. Contrada della Torre 16 - 27036 Mortara PV

inviando le richieste all'indirizzo di posta elettronica fbl@lomellina.it

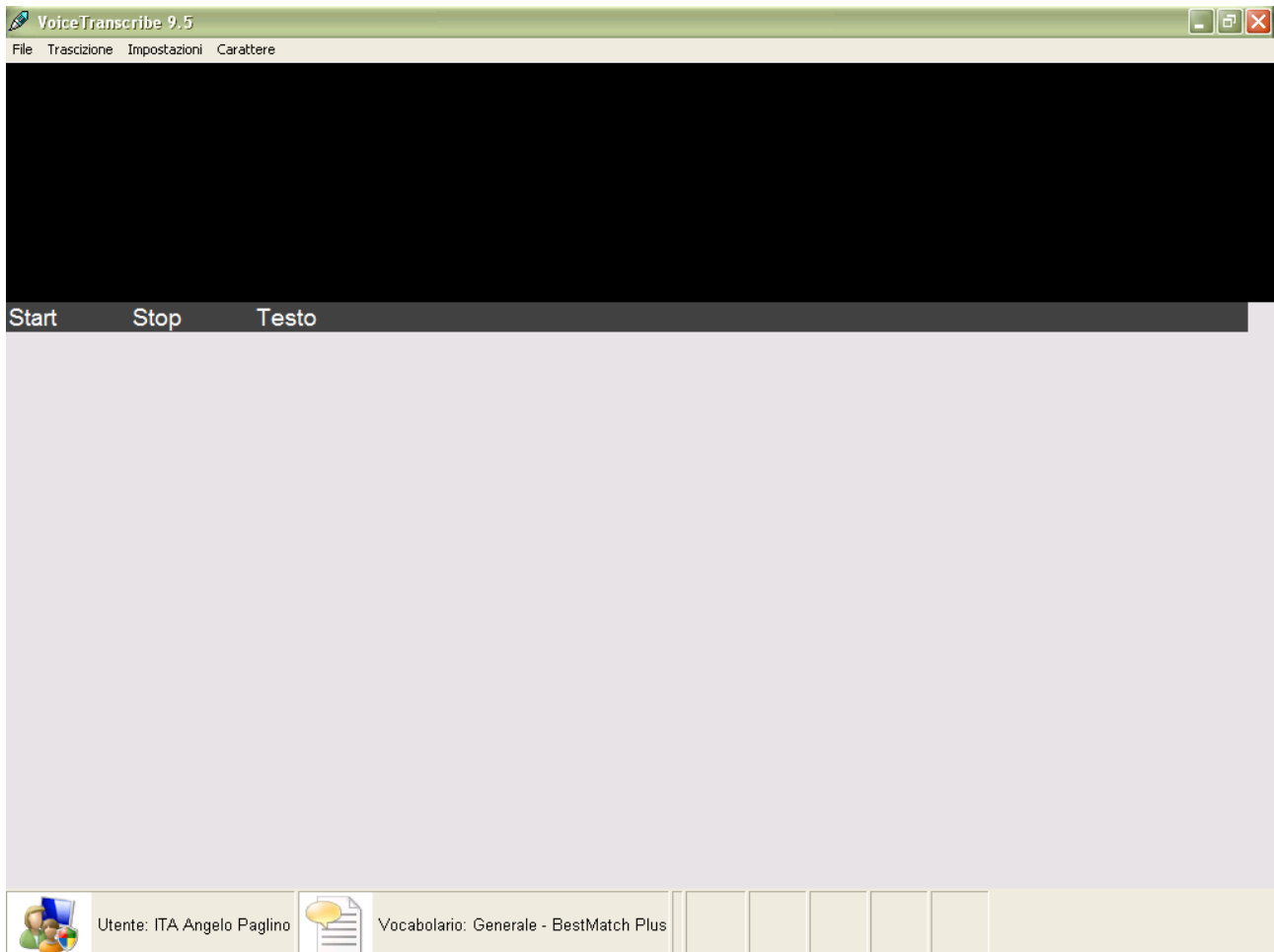
oppure telefonando al numero +39 0384 91399 nell'orario d'ufficio
o inviando fax al numero +39 0384 91362.

Indice degli argomenti

1. La barra dei comandi	pag. 2
1.1 Barra dei comandi: file	pag. 3
2. Barra dei comandi: trascrizione	pag. 4
3. Barra dei comandi: impostazioni / utente	pag. 11
4. Barra dei comandi: carattere	pag. 12
5. Allegato A) generazione di profilo vocale	pag. 13

VoiceTranscribe: La barra dei comandi

All'apertura del programma si presenta la seguente videata:



La barra dei comandi è caratterizzata dai bottoni:

File:

- Salva
- Esci

Trascrizione:

- Diretta
- Differita

Impostazioni:

- Utente

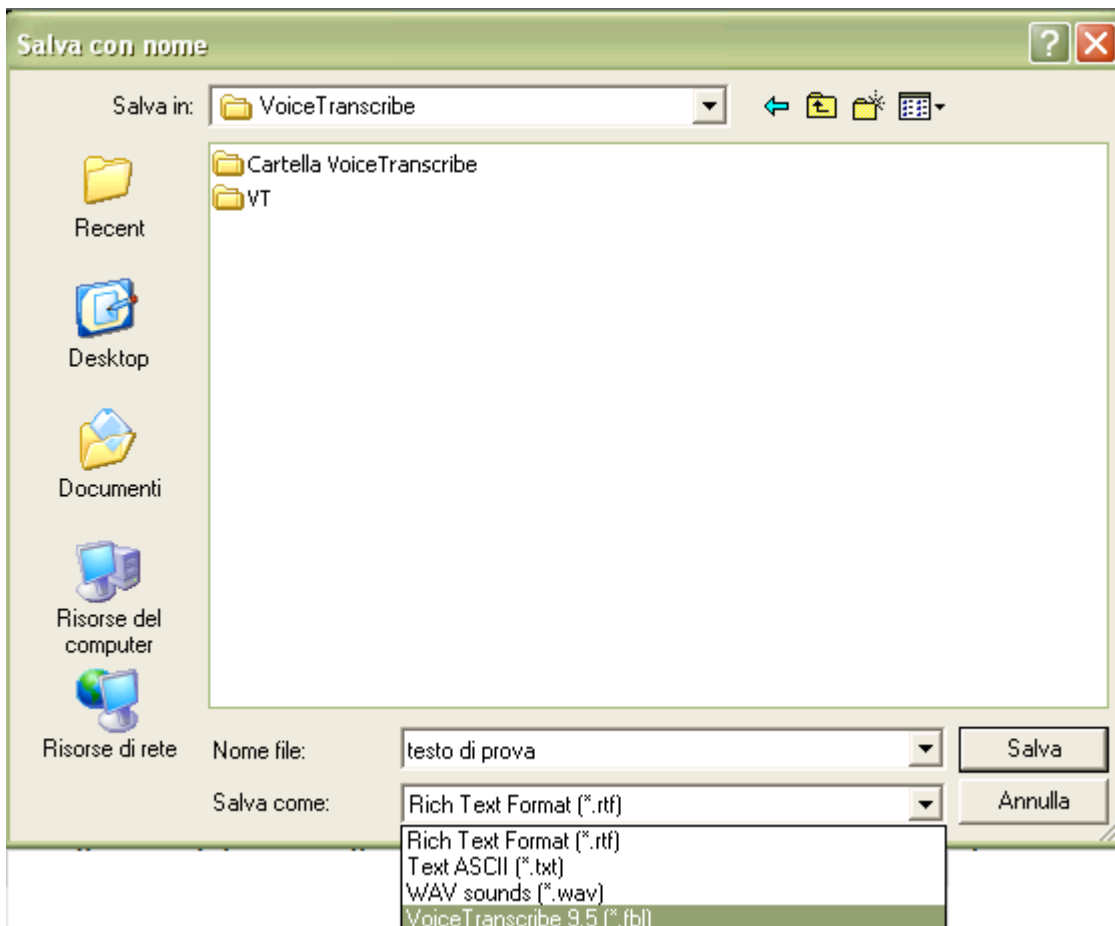
Carattere:

- Modifica

La barra di fondo riporta l'indicazione dell'utente attivo, quello il cui profilo vocale è disponibile per il riconoscimento, e del vocabolario utilizzato. Per poter utilizzare

vocabolari diversi da quello standard (Generale – BestMatch Plus) occorre aver installato la versione Professional di Dragon NaturallySpeaking.

1. Barra dei comandi: file



Il comando **SALVA** permette di “salvare con nome” il lavoro in tre formati:

- **Formato Testo** nelle modalità file.rtf e file.txt per la gestione del testo prodotto con programmi di word processing.
- **Formato Voce** (WAV sounds) nella modalità file.wav per un successivo riascolto della voce tramite il registratore di suoni.
- **Formato Proprietario** (VoiceTranscribe 9.5) nella modalità file.fbl che permette un successivo utilizzo, tramite **VoiceReader**, di testo e voce in sincronismo.

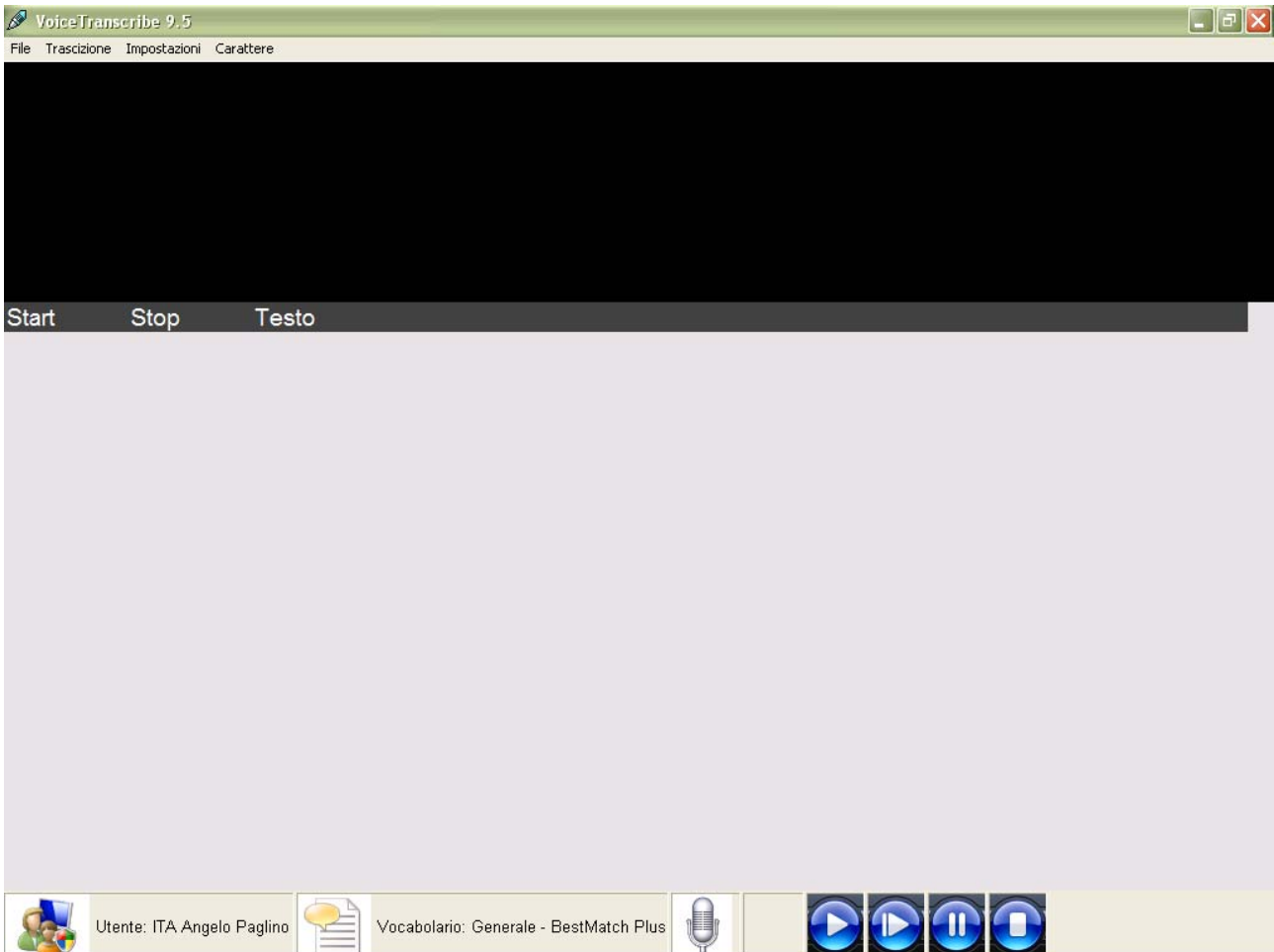
Il comando **ESCI** chiude il programma.

La chiusura del programma senza salvataggio comporta la perdita delle informazioni.

2. Barra dei comandi: Trascrizione

Trascrizione Diretta

Selezionando trascrizione diretta compare la videata:



È importante controllare il nome dell'utente selezionato ed il vocabolario attivo (questa seconda funzione è disponibile solo per gli utenti Dragon NaturallySpeaking Professional). Se l'utente non fosse quello desiderato occorre selezionarlo tramite **impostazioni / utente** (vedi capitolo 3).

Per iniziare la trascrizione è sufficiente cliccare sul microfono ...



microfono spento



microfono acceso

ed iniziare a parlare utilizzando il microfono collegato al personal computer.

Il sistema inizia a trascrivere frase per frase (una frase è rappresentata da una spezzona audio compreso fra due pause). Attenzione alla prima frase: se il personal computer non fosse sufficientemente potente potrebbe apparire in ritardo e non completa!

Nella parte alta dello schermo compaiono delle informazioni grafiche che danno una informazione visiva della prosodia.

- **In alto a sinistra: VU Meter** indica il volume del microfono.
- **In alto a destra: Forma Audio** mostra l'audio sotto forma di una serie di creste e avvallamenti di ampiezza. Evidenzia le pause del parlatore.
- **In basso a sinistra: Oscilloscopio** è uno strumento che visualizza su uno schermo l'andamento di una tensione in funzione del tempo. Con questo strumento si possono eseguire misure di tempo e di ampiezza della forma d'onda presentata all'ingresso.
- **In basso a destra: Spettroscopio** elaborazione numerica dei segnali, altresì nota come Digital Signal Processing (DSP).

Nella parte centrale, nel campo TESTO, la trascrizione della voce, frase per frase (da pausa a pausa).

Start	Stop	Testo
		.\punto
		VoiceTranscribe
		sistema la trascrizione di lezioni o conferenze
		a partire da una registrazione audio digitale
		la trascrizione integrale di una lezione o di una conferenza
		può essere eseguita utilizzando un personal computer ed un software standard di
		riconoscimento vocale
		VoiceTranscribe
		agevola le operazioni di creazione
		del profilo vocale
		di trascrizione e di correzione assistita e semplificata del testo
		di inserimento della punteggiatura
		e di formattazione del testo finale
		le operazioni di correzione
		sono facilitate dalla trascrizione con il controllo dalla pausa
		frase per frase
		da pausa a pausa

Il sistema ascolta la voce del parlatore e quando riconosce una pausa (assenza di segnale per un tempo predeterminato – min. 0,1 secondi, max 1,0 secondi -) elabora il segmento audio e lo trascrive.

La precisione della trascrizione dipende dalla qualità dell'audio e dall'addestramento del profilo vocale dell'utente. I lemmi, per essere riconosciuti, devono essere presenti nel vocabolario attivo. In assenza del lemma il sistema propone il termine che ritiene più assonante.

Al termine della dettatura o ad una pausa decisa dal parlatore, spegnendo con un click il microfono, il sistema elabora il file vocale ed inserisce, per ogni frase, il tempo di inizio (start) e di fine (stop).

Prima di chiudere il microfono attendere che il sistema abbia terminato la trascrizione.





Nella parte alta scompaiono gli indicatori, sostituiti dalla forma d'onda di tutta la registrazione.

Nell'ultima riga compaiono il tempo totale di registrazione ed i pulsanti di riascolto.

Start	Stop	Testo
0.000	2.953	VoiceTranscibe\voitranscraib
2.953	4.918	VoiceTranscibe
4.918	9.317	sistema per la trascrizione di lezioni o conferenze
9.317	13.028	a partire da una registrazione audio digitale
13.028	16.988	la trascrizione integrale di una lezione
16.988	28.291	o di una conferenza può essere eseguita utilizzando un personal computer ed un software standard di riconoscimento vocale
28.291	30.924	VoiceTranscibe
30.924	35.483	agevola le operazioni di creazione del profilo vocale
35.483	40.780	di trascrizione di correzione assistita e semplificata del testo
40.780	47.684	di inserimento della punteggiatura e di formattazione del testo finale
47.684	50.297	le operazioni di correzione
50.297	55.724	sono facilitate dalla trascrizione con il controllo dalla pausa
55.724	57.898	frase per frase
57.898	60.122	da pausa a pausa
60.122	64.172	e dal riascolto dallo spezzone audio
64.172	68.781	il risultato sarà una riduzione dei tempi di lavoro
68.781	72.731	ed un'aumentata produttività dell'operatore
72.731	74.476	grazie

Utente: ITA Angelo Paglino Vocabolario: Generale - BestMatch Plus Durata: 00:01:14.476

Opzioni di riascolto:

- Doppio click su una **frase**: si ottiene il riascolto dello spezzone audio che l'ha generata. Lo spezzone audio viene evidenziato.
- Click su **play**  riascolto della registrazione dall'inizio.
- Click su **pausa**  per interrompere temporaneamente il riascolto.
- Click su **stop**  per chiudere il riascolto.
- Click su **play selezione**  dopo aver selezionato, con il mouse o un doppio click su una frase, una parte della registrazione (forma audio) si riascolta l'audio selezionato.

La chiusura del programma senza salvataggio comporta la perdita delle informazioni.

Trascrizione differita

Prima di selezionare trascrizione differita è opportuno controllare che il profilo vocale attivo sia quello del parlatore.

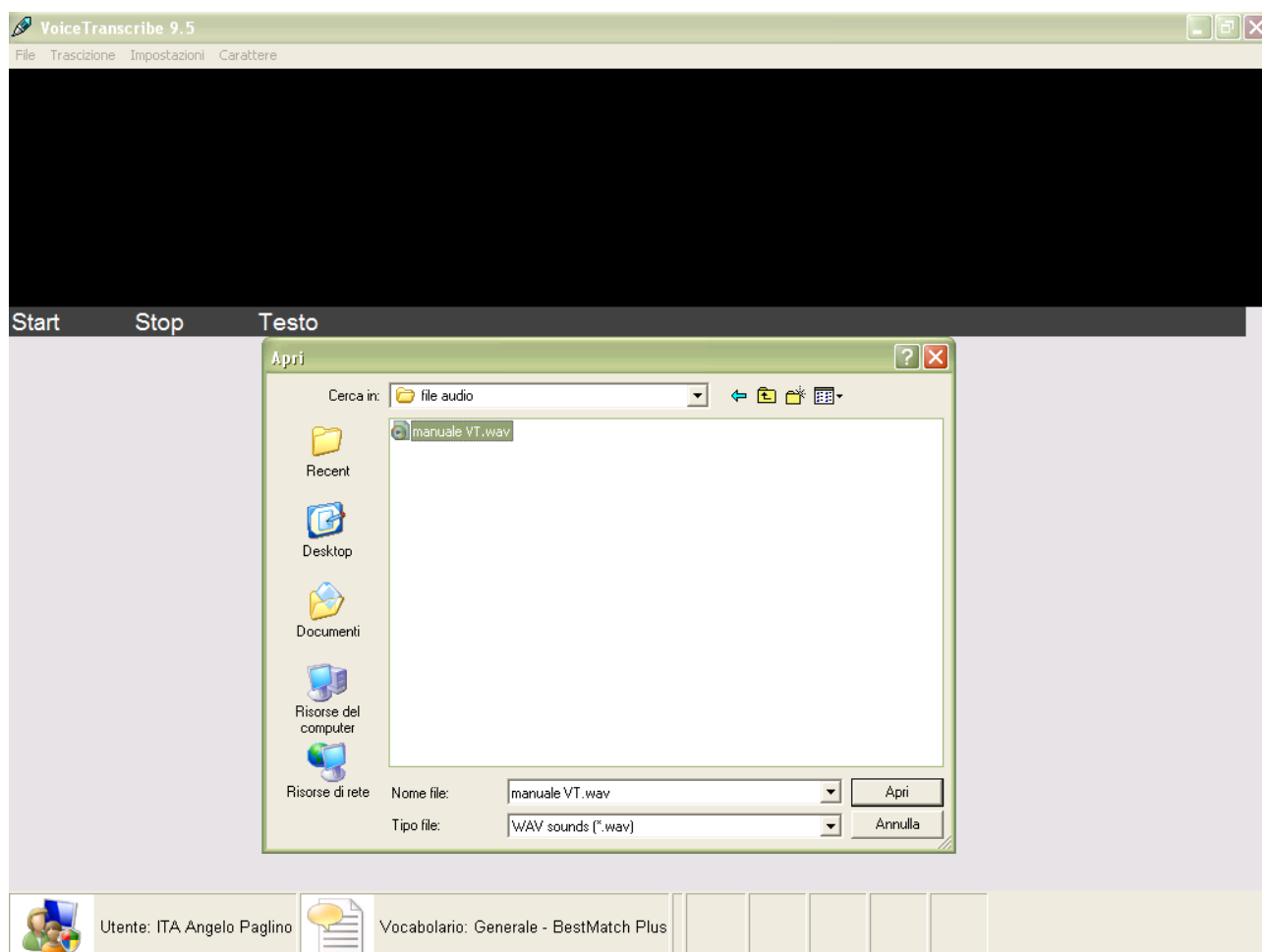
Il profilo del parlatore può essere conosciuto (è stato addestrato) o sconosciuto (non è stato addestrato e non esiste la possibilità di addestrarlo). In questo secondo caso sarà possibile, seguendo le istruzioni **dell'allegato A**, costruire un profilo attendibile.

Solo da una buona registrazione si potrà ottenere una buona trascrizione.

Selezionando trascrizione differita l'applicativo richiede l'apertura di un file audio.

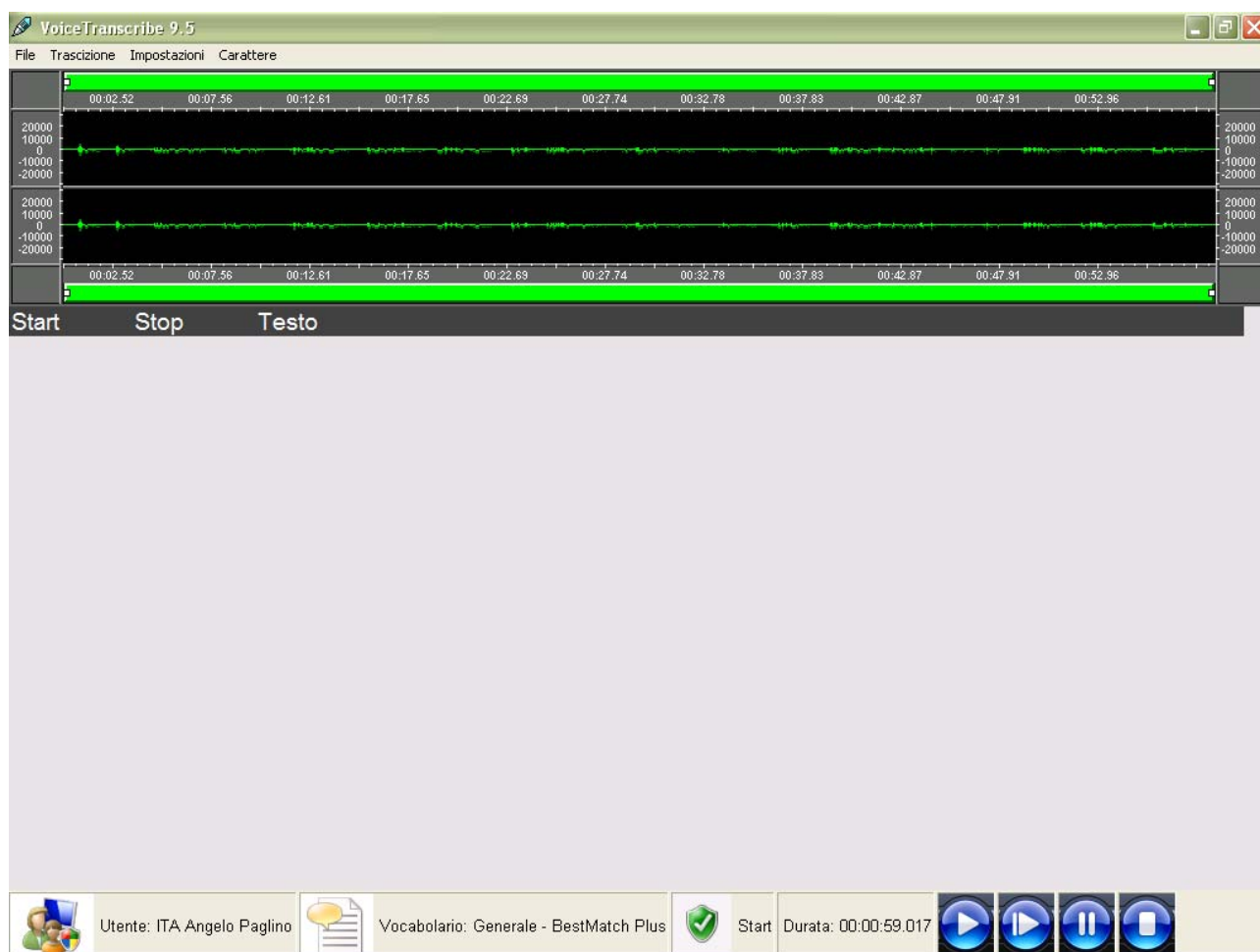
I formati supportati sono Wave (11KHz, 16bit, mono) e WMA.

Si presenta la seguente videata:



Occorre evidenziare con un click del mouse il file audio da trascrivere e selezionare **APRI**.

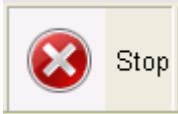
Il programma caricherà il file audio selezionato. Durante il caricamento una barra rossa indicherà l'avanzamento. Al termine, nelle parte alta dello schermo, comparirà la forma d'onda del file selezionato.



Nella parte bassa compariranno le informazioni:

Utente selezionato, Vocabolario attivo, pulsante di Start, Durata del file audio, Tasti di riascolto.

Cliccando sul bottone  inizia la trascrizione.

In qualsiasi momento si potrà arrestare l'elaborazione cliccando sul bottone .


Al termine della trascrizione il risultato sarà evidenziato nella consueta forma.

Start	Stop	Testo
0.049	2.363	VoiceTranscribe\voistranscraib
2.474	4.060	VoiceTranscribe
4.150	7.871	sistema per la trascrizione di lezioni ho conferenze
7.971	11.293	a partire da una registrazione audio digitale
11.404	15.494	la trascrizione integrale di una lezione di una conferenza
15.494	22.607	può essere eseguita utilizzando un personal computer ed un software standard di riconoscimento vocale
22.708	24.404	VoiceTranscribe
24.504	28.275	agevola le operazioni di creazione del profilo vocale
28.375	33.253	di trascrizione di correzione assistita e semplificata del testo
33.334	38.911	di inserimento della punteggiatura e di formattazione del testo finale
39.011	45.136	le operazioni di correzione sono facilitate dalla trascrizione con il controllo dalla pausa
45.237	46.803	frase per frase
46.903	49.008	da pausa a pausa
49.008	51.841	e dal riascolto dallo spezzone audio
51.872	55.413	il risultato sarà una riduzione dei tempi di lavoro
55.453	58.924	ed un'aumentata produttività dell'operatore

Utente: ITA Angelo Paglino | Vocabolario: Generale - BestMatch Plus | Start | Durata: 00:00:59.017


Opzioni di riascolto:

- Doppio click su una **frase**: si ottiene il riascolto dello spezzone audio che l'ha generata. Lo spezzone audio viene evidenziato.

- Click su **play**  riascolto della registrazione dall'inizio.

- Click su **pausa**  per interrompere temporaneamente il riascolto

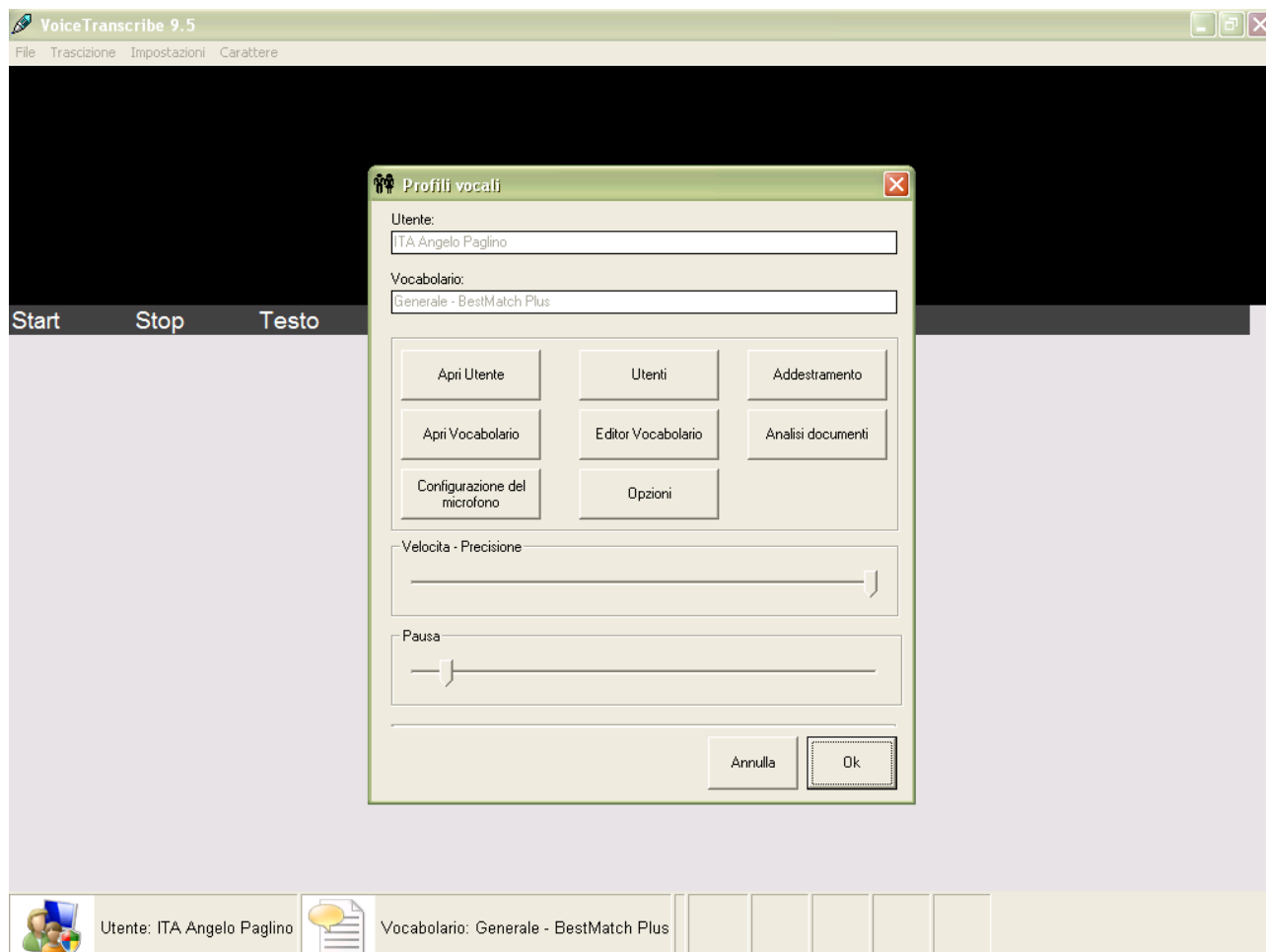
- Click su **stop**  per chiudere il riascolto

- Click su **play selezione**  dopo aver selezionato, con il mouse o un doppio click su una frase, una parte della registrazione (forma audio) si riascolta l'audio selezionato.

La chiusura del programma senza salvataggio comporta la perdita delle informazioni.

3. Barra dei comandi: impostazioni / utente

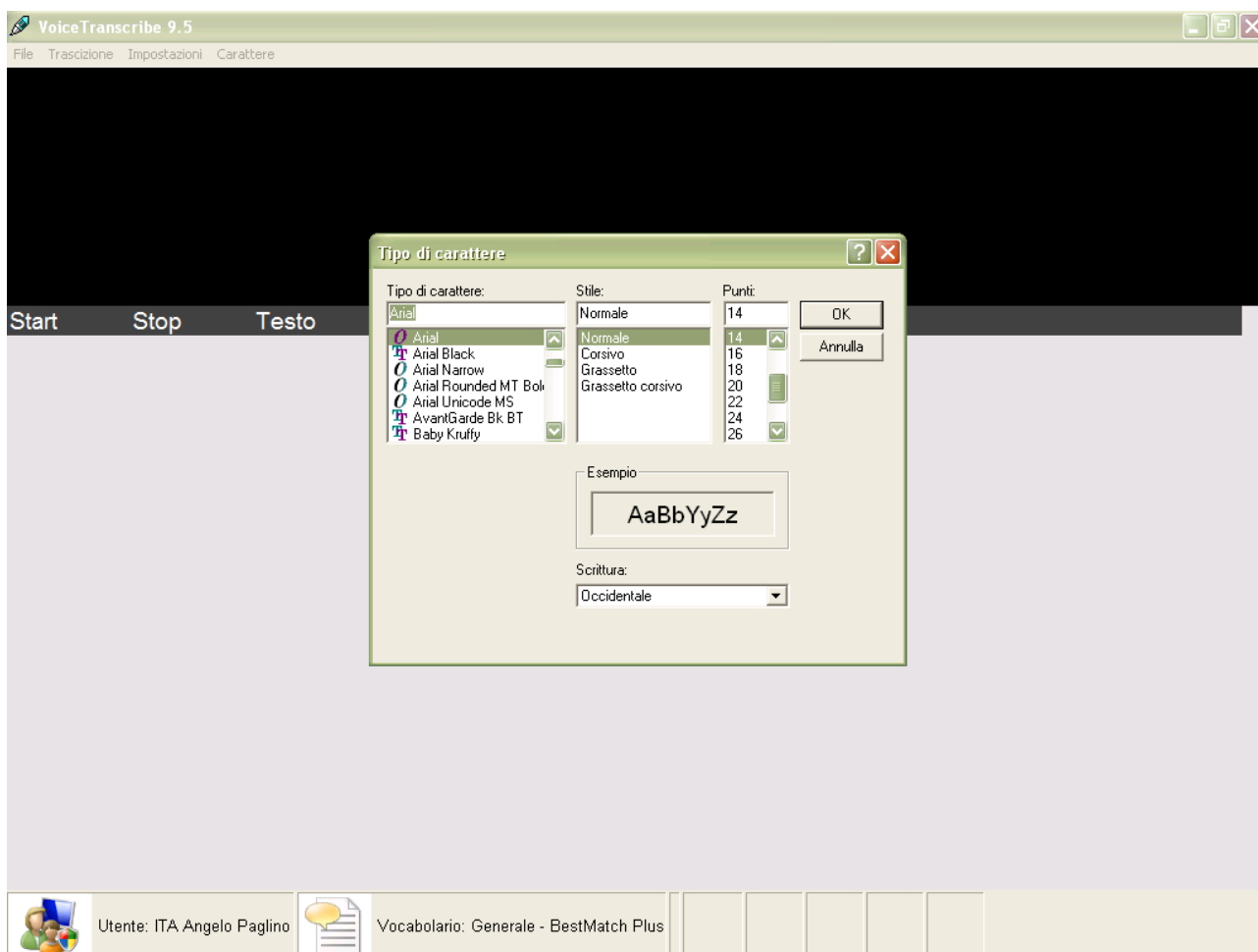
Questa funzione permette la scelta dell'utente attivo, del vocabolario da utilizzare (solo per gli utenti che utilizzano Nuance Dragon NaturallySpeaking Professional).



VoiceTranscribe aggancia automaticamente le funzioni di Dragon NaturallySpeaking 9.5: per ogni ulteriore informazione fare riferimento al manuale di Dragon.

4. Barra dei comandi: carattere

Questa funzione permette la selezione del carattere da utilizzare nella presentazione del testo.

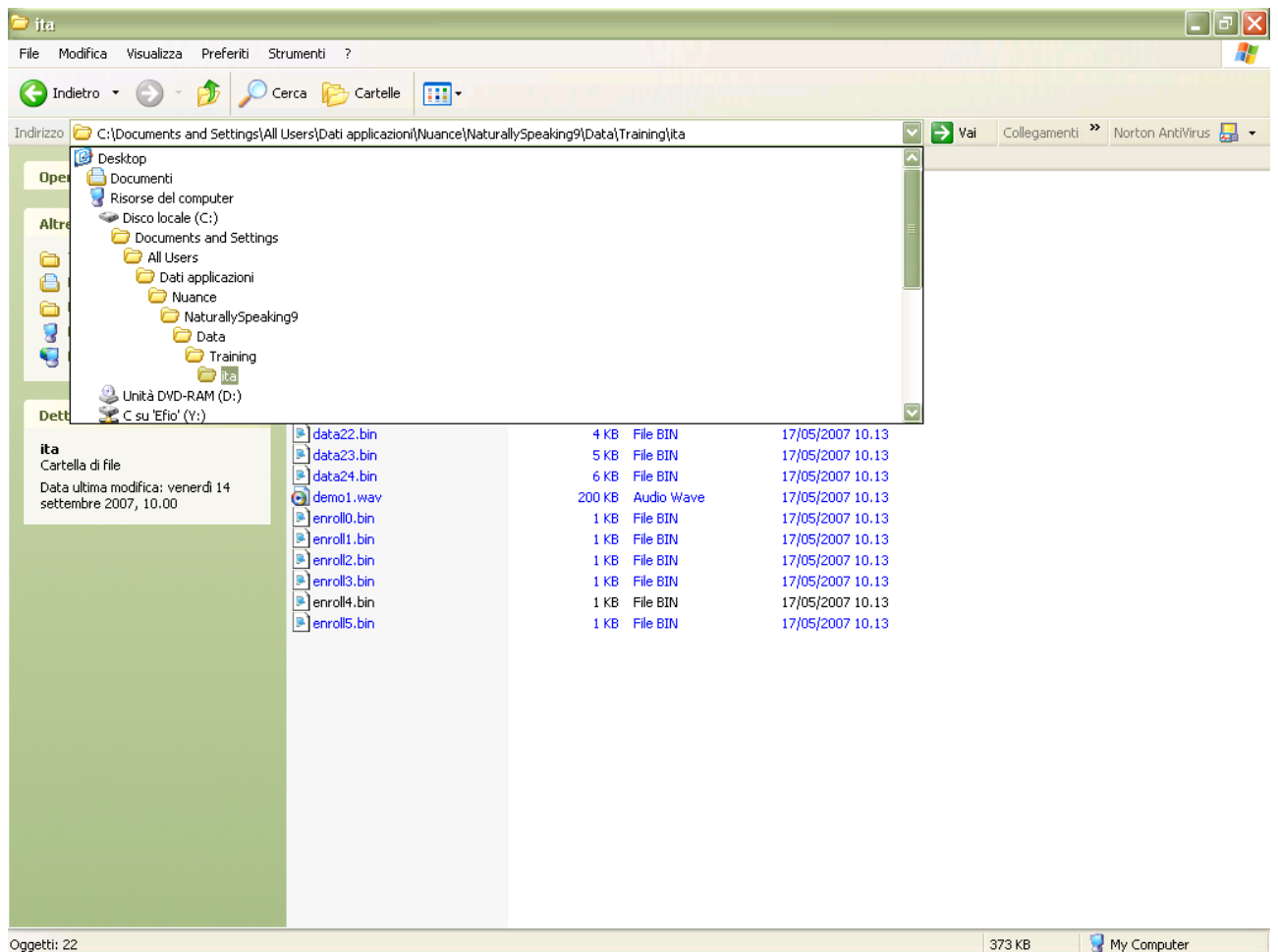


Selezionato il “tipo di carattere”, lo “stile” ed i “punti” occorre confermare la scelta cliccando sul bottone **OK**

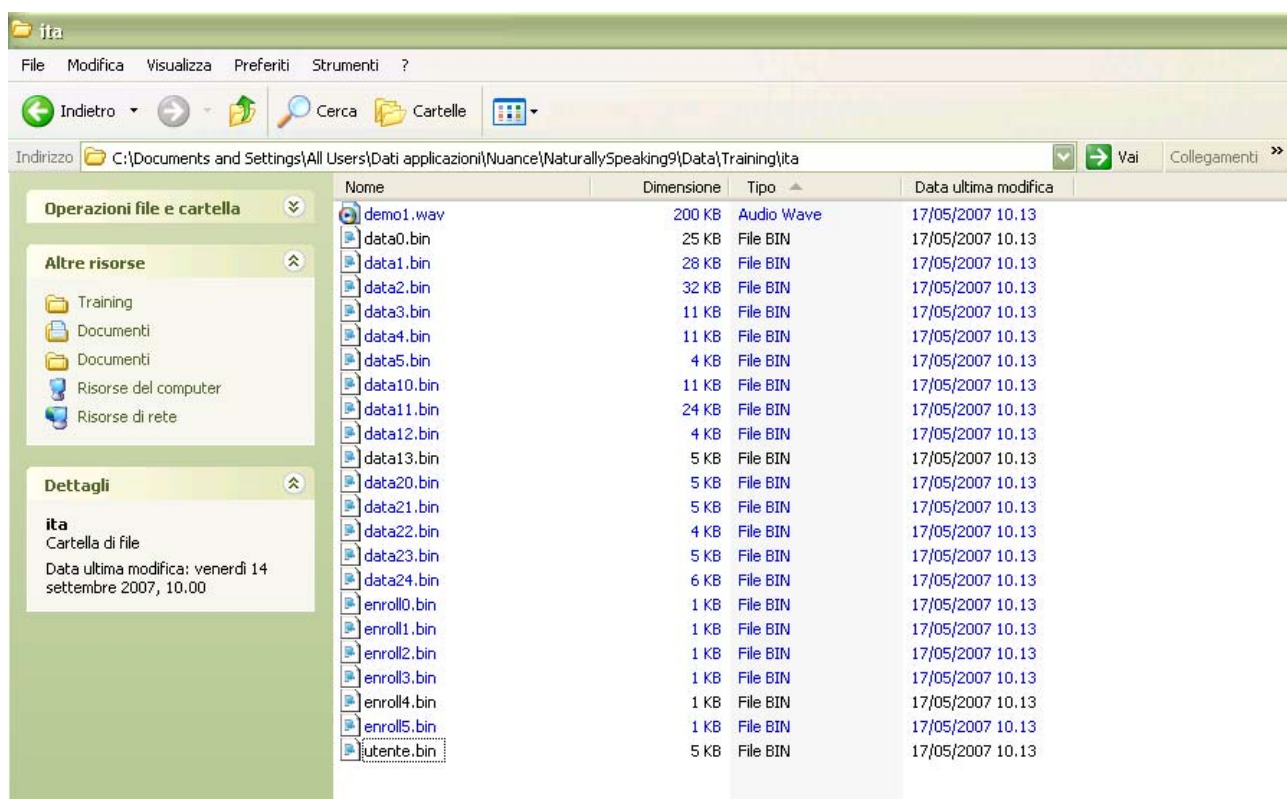
Appendice A): Creazione di un profilo vocale in assenza del parlatore

Questa funzione viene utilizzata normalmente per due differenti motivi: la non disponibilità dello speaker al momento della generazione dei sottotitoli e una migliore risposta del motore di trascrizione in presenza di un utente il cui addestramento è stato fatto direttamente da un file audio, con caratteristiche simili a quelle del file da trascrivere.

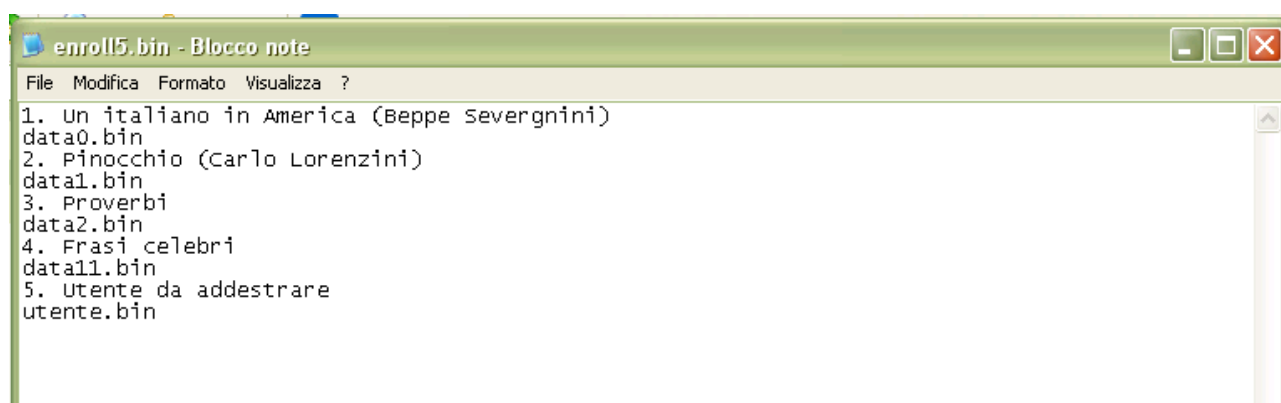
- A) estrarre dal filmato il file audio e convertirlo al formato PCM 11KHz 16bit Mono.
- B) **creare un nuovo file audio con i primi 4 minuti** della registrazione.
- C) duplicare quattro volte il file audio (durata totale > 15 minuti)
- D) **trascrivere i primi 4 minuti**. La trascrizione può essere effettuata utilizzando un utente con caratteristiche vocali affini e poi correggendo il testo oppure ascoltando e ridettando in Dragon.
- E) duplicare quattro volte il testo trascritto che deve essere allineato con il file audio.
- F) Posizionarsi sulla cartella:



- F) inserire il testo, dopo averlo salvato nel formato .bin, nella cartella (nel nostro esempio il file sarà denominato “utente.bin”):



- G) aggiornare il file “enroll5.bin” che si trova nella stessa cartella inserendo il richiamo per il nuovo testo (nell’esempio creare il punto 5. ed inserire nella riga successiva il richiamo al file “utente.bin”).



- H aprire Dragon NaturallySpeaking e iniziare la procedura per la creazione di un nuovo utente: **APRI UTENTE / NUOVO**.

Creazione guidata utente

Creazione utente

Nome:

Lingua:

Sorgente:

Vocabolario:

Salta addestramento iniziale di questo utente

Fare clic su 'Avanti' per continuare.

< Indietro **Avanti >** Annulla Guida

Dopo aver introdotto il **nome** dell'utente ed aver selezionato la **lingua**, selezionare nel campo **sorgente** “**registratore digitale utilizzando file audio**”.

Click su **Avanti**



Proseguire la creazione guidata scegliendo, nel nostro esempio, il punto 5. **“utente da addestrare”**

Click su **Avanti>**



Alla richiesta del file audio, utilizzando **Sfogli...**

Selezionare il file da utilizzare e cliccare su **Apri**



Dragon NaturallySpeaking inizierà ad elaborare le informazioni e predisporrà un profilo vocale che sarà utilizzabile solo da registrazione.

Per maggiori informazioni fare riferimento al manuale Nuance Dragon NaturallySpeaking.