

VoiceManager

Guida all'uso



Sommario

Sommario	3
Voice Manager	4
1 Introduzione	4
2 Uso del sistema di riconoscimento vocale	5
3 Funzionamento del riconoscitore vocale	12
4 Personalizzazione	18

Voice Manager

1 Introduzione

La comunicazione verbale è una delle abilità umane più utilizzate nella vita di tutti i giorni. Forse è per questo che l'uomo ha sempre sognato di poter interagire con la tecnologia parlando normalmente, allo stesso modo in cui si comunica tra persone. Questo desiderio è stato il traino delle ricerche sul riconoscimento vocale che hanno portato aziende come Google e Apple, solo per citarne alcune, a raggiungere dei risultati applicabili ai dispositivi attuali. L'utilizzo dei comandi vocali si è rivelato molto pratico per i dispositivi mobile e in futuro lo sarà anche per quelli indossabili. Non solo la voce permette di tenere le mani libere ma abbassa anche il numero di operazioni che separano l'utente dalle informazioni, rendendo le applicazioni più semplici e naturali da usare.

Dalla versione 13.0 di Instant Developer, nelle applicazioni web è integrato un sistema per il riconoscimento e l'interpretazione di comandi vocali, controllabile con la libreria Voice Manager descritta in questo capitolo. La libreria utilizza le Speech API di Google per trasformare il parlato in un testo che viene poi analizzato dall'interprete dei comandi dell'applicazione, con il risultato di eseguire il comando vocale. Attraverso i comandi vocali si possono aprire e chiudere videate, navigare i pannelli, inserire, modificare e cancellare i dati, fare delle ricerche ed eseguire procedure.

Dalla versione 13.5 il riconoscimento vocale è disponibile anche per le applicazioni mobile offline.

Le applicazioni per iOS usano le API di Nuance per la parte di riconoscimento vocale e le API di Apple per il sintetizzatore vocale. Le applicazioni per Android e Windows Phone 8.1 utilizzano rispettivamente le API di Google e di Microsoft sia per il riconoscimento vocale che per il sintetizzatore vocale. Per Windows il riconoscimento vocale è disponibile solo in Windows Phone 8.1.

Questa funzionalità è ancora in fase sperimentale. Per questo motivo consigliamo di non attivarla in progetti già esistenti ma di creare dei progetti ad hoc che permettano di esplorarne le possibilità. Per dimostrare le potenzialità di Voice Manager all'indirizzo <https://www.progamma.com/VoiceCrm/> è disponibile un'applicazione di prova. Il sorgente del progetto è scaricabile all'indirizzo <http://blog.progamma.com/wp-content/uploads/2014/04/VoiceCrm.zip>.

2 Uso del sistema di riconoscimento vocale

Il sistema di riconoscimento vocale non è attivo per impostazione predefinita ma deve essere esplicitamente abilitato impostando il parametro *Abilita comandi vocali* nella sezione *Generale* dei parametri di compilazione. Inoltre, solo per le applicazioni sviluppate per iOS, occorre:

- 1) Registrarsi sul sito di Nuance all'indirizzo <http://nuancemobiledeveloper.com>
- 2) Ottenere le credenziali sandbox: Host, AppId e AppKey.
- 3) Personalizzare la Shell, scrivendo nel file *Service.m* le proprie credenziali.
- 4) Quando l'applicazione è pronta, registrarla sul sito di Nuance in modo da ottenere le credenziali di produzione.
- 5) Sostituire le credenziali sandbox con quelle di produzione.
- 6) Sottomettere l'applicazione a Nuance.

Una volta che il riconoscimento vocale è abilitato, occorre utilizzare il parametro *Listening* per attivare il microfono e iniziare ad ascoltare quello che dice l'utente. L'utilizzo del parametro *Listening* è spiegato nel paragrafo 3.2. In alternativa, nelle applicazioni web, può essere usata la combinazione di tasti *ctrl + spazio* -. L'utilizzo del microfono non è libero ma occorre che l'utente lo autorizzi. Nelle applicazioni web, l'autorizzazione si concede cliccando sul bottone *Consenti* che comparirà in una barra nella parte superiore della pagina. Se l'applicazione è installata su un server che ammette il protocollo https, sarà necessario autorizzare l'uso del microfono soltanto la prima volta. In caso contrario, l'autorizzazione verrà richiesta tutte le volte che viene attivato il riconoscimento vocale.

Nella parte bassa dello schermo compare un'icona rossa che rappresenta un microfono e una nuvoletta con il testo *Cosa posso fare per te?* A questo punto si può iniziare a parlare. Nelle applicazioni web, le parole pronunciate vengono trascritte dal sistema in tempo reale all'interno della nuvoletta. Mentre l'utente sta parlando il testo è di colore nero, per indicare che il comando non è ancora concluso e per il momento non sarà inviato al riconoscitore. Quando l'utente smette di parlare, il testo diventa di colore blu e viene inviato per l'interpretazione. Nelle applicazioni mobile, invece, il testo verrà mostrato soltanto quando il comando è concluso. Se il riconoscitore riesce a interpretare il comando vocale, esegue l'azione corrispondente, scrive la risposta all'interno della nuvoletta con un testo di colore verde e la legge. In caso contrario non viene eseguita nessuna azione e il riconoscitore comunicherà che non ha capito il comando.

Nell'immagine che segue si può vedere come cambia il testo nelle varie fasi del riconoscimento in un'applicazione web. La prima nuvoletta rappresenta il momento in cui il sistema è in attesa di ricevere il comando. Quando l'utente inizia a parlare il testo viene trascritto in nero e quando smette di parlare il testo viene inviato al riconoscitore e diventa blu, come mostrato nella seconda e nella terza nuvoletta. Infine, il comando

riconosciuto viene eseguito e la risposta restituita nella nuvoletta è di colore verde, come nell'ultimo caso.



Interfaccia del riconoscitore vocale

I comandi riconosciuti dal sistema sono formati da due parti: un verbo, che indica l'azione da eseguire, e zero o più predicati, che descrivono meglio l'azione o indicano quali oggetti sono coinvolti. In base al comando che si vuole fare eseguire, una delle due parti può essere omessa. In questo caso è il sistema che, in base al contesto, capisce quale azione eseguire o individua l'oggetto sul quale eseguirla. Il verbo, se presente, deve essere la prima parola del comando.

Tutte le azioni eseguite in seguito a un comando vocale corrispondono esattamente alla stessa azione eseguita con il mouse. L'utilizzo dei comandi vocali non permette di bypassare i comportamenti e le proprietà definiti per gli oggetti.

2.1 Comandi di navigazione

I comandi di navigazione permettono di spostarsi tra i record di un pannello o tra le pagine di un book e agiscono soltanto sui pannelli e sui book visibili. Per questo tipo di comandi non è obbligatoria la presenza del verbo e del pannello o del book su cui deve essere eseguita l'azione, ma basta specificare dove ci si vuole posizionare o in che stato si vuole portare il pannello. Se il nome del pannello viene omesso il sistema esegue l'azione sul pannello con il fuoco. Se viene omesso il nome del book il sistema esegue l'azione sul book a video, se è l'unico.

I comandi disponibili per i pannelli sono i seguenti:

- 1) Portarsi nella prima o nell'ultima pagina
Es. *“Vai all'inizio del pannello Aziende”*; *“Ultima riga”*.
- 2) Andare avanti o indietro tra le pagine
Es. *“Mostra la pagina successiva”*; *“Indietro”*.
- 3) Portare un pannello in lista o in dettaglio
“Mostra il dettaglio del pannello Persone”; *“Torna in lista”*.
- 4) Spostarsi tra le pagine di un pannello paginato
Es. *“Vai alla pagina Persone del pannello Aziende”*; *“Pagina Persone”*.

- 5) Portarsi su una specifica riga del pannello, indicando il contenuto di un campo o il numero della riga.
Es. *“Vai all’azienda Pro Gamma”*; *“Riga 36”*.

I comandi disponibili per i book sono i seguenti:

- 1) Portarsi nella prima o nell’ultima pagina
Es. *“Vai alla prima pagina del book Aziende”*; *“Ultima pagina”*.
- 2) Andare avanti o indietro tra le pagine
Es. *“Mostra la pagina successiva”*; *“Indietro”*.
- 3) Portarsi su una specifica pagina del book
Es. *“Vai a pagina 3”*.

2.2 Comandi di pannello

In questo gruppo rientrano tutti i comandi presenti normalmente nella toolbar di un pannello e che non sono già compresi tra i comandi di navigazione. Anche questi comandi agiscono soltanto sui pannelli visibili e, come nel caso precedente, se il nome del pannello viene omesso il sistema prende in considerazione il pannello con il fuoco. A differenza dei comandi del gruppo precedente, il verbo deve essere pronunciato perché questi comandi siano riconosciuti. I comandi disponibili sono i seguenti:

- 1) Bloccare e sbloccare un pannello
Es. *“Sblocca”*; *“Blocca il pannello Persone”*.
- 2) Duplicare una riga
Es. *“Copia la riga”*; *“Duplica”*
- 3) Cancellare una riga
Es. *“Elimina la riga”*; *“Cancella”*.
- 4) Salvare le modifiche
Es. *“Salva i dati in Aziende”*; *“Salva”*.
- 5) Annullare le modifiche
Es. *“Annulla”*
- 6) Ricaricare i dati
Es. *“Aggiorna i dati”*; *“Ricarica”*.
- 7) Esportare i dati in Excel
Es. *“Esporta il pannello Persone”*.
- 8) Selezionare una riga
Es. *“Seleziona Colorado nel pannello Aziende”*; *“Seleziona Pro Gamma”*.
- 9) Selezionare tutte le righe
Es. *“Seleziona tutto”*.

- 10) Annullare la selezione
Es. “*Annulla selezione*”.
- 11) Invertire la selezione
Es. “*Inverti selezione*”.

2.3 Apertura e chiusura di una videata

Per permettere l’apertura di una videata attraverso l’utilizzo di un comando vocale, occorre che la videata da aprire sia collegata a una voce di menu con lo stesso nome. L’apertura della videata avviene pronunciandone il nome. Il verbo Apri in questo caso non è obbligatorio.

Per chiudere una videata aperta è disponibile il comando Chiudi. Se non viene indicato il nome della videata da chiudere, verrà chiusa quella attiva. Se si vogliono chiudere tutte le videate aperte si può utilizzare il comando “*Chiudi tutto*”.

2.4 Ricerca di dati

I comandi per la ricerca dei dati possono essere usati sia nei pannelli sia nei book. Per quanto riguarda i pannelli, quando viene pronunciato un comando per la ricerca dei dati, il testo da cercare può essere preceduto o seguito dal nome del campo in cui si vuole cercare. Se il nome del campo viene omissso, il sistema cerca prima nelle liste valori del pannello e poi, se non trova risultati, nei campi descrittivi. Si controlla inizialmente se il campo contiene l’intera parola e, se non vengono trovati risultati, si continua la ricerca togliendo ogni volta la lettera finale fino a quando rimangono almeno tre lettere. A questo punto la ricerca termina, anche se non sono stati trovati risultati. Questo tipo di ricerca non avviene per i campi numerici e per i campi di lookup, per i quali viene cercato soltanto il numero o la parola per intero.

Se, ad esempio, si ha una videata *Aziende* con un campo descrittivo *azienda* e una lista valori *stato*, i comandi vocali per effettuare una ricerca potrebbero essere i seguenti:

- 1) “*Cerca azienda Pro Gamma*”. In questo caso il sistema riconosce *azienda* come nome del campo e cerca i record che contengono *Pro Gamma* in quel campo.
- 2) “*Cerca azienda Pro Gamma spa*”. Anche in questo caso il sistema riconosce *azienda* come nome del campo e cerca i record che contengono *Pro Gamma spa* in quel campo. Dal momento che l’azienda è memorizzata come *Pro Gamma*, la ricerca dell’intera stringa non restituisce alcun risultato. Il sistema allora fa delle ulteriori ricerche togliendo ogni volta l’ultima lettera del testo da cercare fino a quando

non arriva ad ottenere *Pro Gamma*, che è la stringa che restituisce dei risultati e fa terminare la ricerca.

- 3) “*Cerca le aziende della California*”. In questo caso non è stato specificato il nome di nessun campo. Il sistema inizia a cercare nella lista valori *stato*, trovando tutte le aziende californiane.
- 4) “*Cerca Pro Gamma*”. In questo caso non è stato specificato il nome di nessun campo. Il sistema inizia a cercare nella lista valori *stato* e, non trovando risultati, continua la ricerca nel campo *azienda* che è un campo descrittivo.

La ricerca può essere effettuata anche in videate non aperte che siano collegate ad una voce di menu. Occorre però che voce di menu, videata e pannello abbiano lo stesso nome ed è inoltre indispensabile indicare nel comando vocale su quale videata si vuole agire.

Per quanto riguarda i book, invece, la ricerca viene effettuata sulla base del contenuto testuale delle box. Se una box contiene le parole dette dall’utente allora il book si sposterà sulla pagina contenente quella box.

2.5 Inserimento e modifica di dati

I comandi vocali possono essere usati per l’inserimento o la modifica dei dati di un pannello.

Per inserire una nuova riga occorre utilizzare il comando *inserisci*. Se non si specifica niente il pannello si porta semplicemente su una nuova riga, se invece pronuncia del testo preceduto o seguito dal nome di un campo, il pannello inserisce una nuova riga riempiendo quel campo con il testo detto nel comando. Il nome del campo può essere anche omesso e in questo caso il testo viene scritto nel primo campo descrittivo.

Usando il comando *modifica* è possibile modificare la riga corrente del pannello. Nel comando occorre specificare il nuovo testo per il campo e, facoltativamente, il nome del campo. Se quest’ultimo non viene pronunciato, il sistema modifica il primo campo descrittivo del pannello. Se il pannello è in stato modificato, per cambiare il valore della riga corrente, è sufficiente pronunciare il nome del campo e il testo da scrivere omettendo il verbo.

Se, ad esempio, si ha una videata *Aziende* con un campo descrittivo *azienda* e una lista valori *stato* i comandi vocali per inserire i dati potrebbero essere i seguenti:

- 1) “*Inserisci azienda Pro Gamma*”. Il sistema inserisce una nuova riga e scrive “*Pro Gamma*” nel campo *azienda*.
- 2) “*Inserisci Pro Gamma*”. Con questo comando si ottiene lo stesso risultato del comando precedente, perché anche se il nome del campo è stato omesso, *azienda* è un campo descrittivo ed è quindi quello scelto in automatico dal sistema.

Se invece si vogliono modificare i dati, alcuni esempi di comandi potrebbero essere i seguenti:

- 1) “*Modifica stato Florida*”. Il valore del campo regione della riga corrente viene modificato in *Florida*.
- 2) “*Stato Arizona*”. Se il pannello è in stato modificato, questo comando permette di assegnare al campo stato della riga corrente il valore *Arizona*.

2.6 Esecuzione di una procedura

Abbiamo visto che tramite i comandi vocali è possibile comandare le funzioni dell'interfaccia utente, che sono le stesse funzioni chiamabili usando la tastiera e il mouse. I comandi vocali possono però essere utilizzati anche per chiamare delle procedure che non sono collegate a un'interfaccia utente, in modo da poter attivare funzionalità dell'applicazione anche senza avere un'interfaccia dedicata. Per poter chiamare e interagire con una procedura attraverso un comando vocale, occorre che il parametro *Voice Cmd* sia attivato nella videata relativa ai parametri della procedura. Se nell'applicazione non c'è nessuna videata aperta, possono essere chiamate le procedure globali e quelle collegate a voci del menu principale che abbiano lo stesso nome. Se invece nell'applicazione sono aperte delle videate, è possibile chiamare anche le procedure contenute in menu di videata e di pannello.

È possibile definire delle procedure che accettano dei parametri. Occorre però che i parametri siano tutti opzionali. Quando viene pronunciato un comando per l'esecuzione di una procedura, le parole che fanno parte di un parametro possono essere precedute o seguite dal nome del parametro stesso. Se questo viene omissso, il riconoscitore prova a dividere il testo in tutti i parametri, a partire dal primo, basandosi sul tipo e sulla lunghezza se il parametro è di tipo *character*.

Facciamo un esempio utilizzando la procedura *Ordine* che prende come parametri un intero che rappresenta la quantità da ordinare, una stringa che rappresenta il nome dell'articolo e la cui lunghezza massima è 6 caratteri e una seconda stringa che rappresenta la descrizione dell'articolo e la cui lunghezza massima è di 255 caratteri.

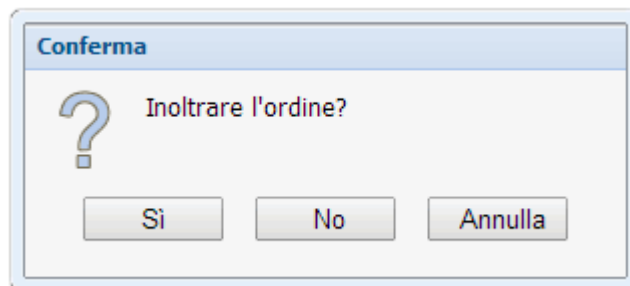
```
public voicecmd void VoiceCrm.Ordina(  
    optional int Quantità = 0 //  
    optional string Articolo = "0" //  
    optional string Descrizione = "0" //  
)
```

Esempio di procedura attivabile con un comando vocale

Per chiamare questa procedura, un possibile comando vocale è il seguente: “*Ordina quantità 3 articolo mouse descrizione rossi*”. In questo caso sono stati specificati i nomi dei parametri e il sistema assegna facilmente i valori esatti a ognuno di essi. Un comando possibile però potrebbe essere anche il seguente: “*Ordina 3 mouse rossi*”. Dal momento che i nomi dei parametri sono stati omissi, è il riconoscitore che si occupa di assegnare il valore corretto ad ogni parametro. I parametri vengono considerati nell’ordine: viene quindi analizzato prima *Quantità*, poi *Articolo* e infine *Descrizione*. Anche le parole del comando vengono prese nell’ordine in cui sono state pronunciate. La prima parola ad essere analizzata è quindi *3*. Il sistema capisce che *3* è un numero e, analizzando il parametro *Quantità*, si accorge che anch’esso è numerico e quindi decide di assegnare questo valore al parametro. Si passa quindi al parametro successivo, *Articolo*, e alla prossima parola, *mouse*. Dal momento che il parametro è di tipo *string* e la parola *mouse* è più corta di 6 caratteri, che è la lunghezza massima del parametro, il sistema assegna *mouse* al parametro *Articolo* e passa alla parola successiva, *rossi*, per verificare se anch’essa può essere assegnata al parametro. Assegnando anche *rossi* al parametro *Articolo* però si supererebbe la sua lunghezza massima, il sistema quindi passa al parametro successivo, *Descrizione*, per verificare se *rossi* può essere assegnata ad esso. Visto che *Descrizione* è di tipo *string* e *rossi* non supera la sua lunghezza massima, il valore può essere assegnato al parametro.

2.7 Gestione della conferma utente tramite comando vocale

A volte è necessario fare in modo che le procedure interagiscano con l’utente. A questo scopo si possono usare le message box, le message confirm e le input box che però, nel caso dei comandi vocali, hanno un comportamento diverso da quello standard. Nel caso della message box, non compare la videata modale standard ma il testo impostato per la message box appare nella nuvoletta del microfono del riconoscitore vocale. Nel caso della message confirm e della input box, invece, si vede la videata modale standard e, quando l’utente parla, il testo compare in questa videata invece che nella nuvoletta del riconoscitore vocale. Per attivare i bottoni di queste videate basta pronunciarne il nome. Per esempio, per usare in una procedura attivata con la voce una message confirm come quella in figura, si può pronunciare “*si*” per confermare, “*no*” e “*annulla*” per annullare l’azione.



2.8 Lettura di dati

L'introduzione del sintetizzatore vocale permette la lettura dei dati dei pannelli e dei book. Per fare in modo che l'applicazione legga i dati occorre usare il comando *leggi*.

Per quanto riguarda i pannelli, la lettura dei dati può essere effettuata anche in videate non aperte che siano collegate ad una voce di menu. In questo caso il sistema aprirà la videata, cercherà i dati da leggere e richiuderà la videata. Se il nome del pannello non viene specificato, il comando verrà applicato al pannello con il fuoco. Allo stesso modo se non viene specificato il nome del campo da leggere, il sistema leggerà il campo con il fuoco. Si può selezionare la riga da leggere specificando il valore del campo descrittivo. Se, ad esempio, si ha una videata *Aziende* con un campo descrittivo *ragione sociale* e un campo *indirizzo*, i comandi vocali potrebbero essere i seguenti:

- 1) "*Leggi*". Verrà letto il contenuto del campo con il fuoco.
- 2) "*Leggi indirizzo*". Verrà letto il contenuto del campo *indirizzo* della riga corrente.
- 3) "*Dimmi l'indirizzo dell'azienda Pro Gamma*"; "*Qual è l'indirizzo dell'azienda Pro Gamma*". Il sistema cercherà nel pannello un record la cui ragione sociale contiene "*Pro Gamma*" e ne leggerà il campo *indirizzo*.

Per quanto riguarda i book, occorre specificare la box da leggere indicandone il tooltip. Il sistema cercherà nel book la box il cui tooltip corrisponde alle parole del comando vocale e leggerà il testo contenuto in essa.

3 Funzionamento del riconoscitore vocale

Per fare in modo che nell'applicazione si possano usare i comandi vocali, occorre attivare il parametro di compilazione *Abilita comandi vocali* che si trova nella sezione *Generale* della videata relativa ai parametri di compilazione. In questo modo, al momento della compilazione viene creato il file *prjVoice.xml*. Questo file rappresenta il contesto in cui i comandi vocali devono essere interpretati e contiene gli oggetti dell'applicazione utili a tale scopo (videate, pannelli, menu, voci di menu, procedure, campi statici). Se si sta compilando in Java, il file viene creato nella cartella di compilazione dell'applicazione dentro *Servlet/web*. Se invece si compila l'applicazione in C#, il file si trova sempre nella cartella di compilazione ma nel percorso *CSharp/DB*.

Nella stessa cartella sono presenti anche i file *voice_ita.xml* e *voice_eng.xml* che descrivono i comandi accettati dal riconoscitore quando usato rispettivamente in italiano e in inglese.

Quando l'utente pronuncia il comando, per prima cosa scatta l'evento di *OnCommand* con parametro *Command* uguale a *VOICE*. Se il riconoscimento vocale non viene interrotto, la frase viene trasformata in testo e presa in carico dal riconoscitore che prova a interpretarla confrontandola con ogni comando presente nella grammatica. L'interpretazione avviene in due fasi: nella prima fase le parole del testo vengono confrontate con le parole del comando della grammatica, nella seconda fase viene preso in considerazione anche il contesto per individuare quali oggetti dell'applicazione, se ce ne sono, sono coinvolti nel comando. Al termine delle due fasi viene assegnato un punteggio e si passa al comando della grammatica successivo. Alla fine di tutti i confronti viene eseguito il comando che ha il punteggio maggiore, se questo punteggio supera la soglia minima accettata, e viene mostrata a video una risposta testuale. Altrimenti il riconoscitore comunicherà che il comando non è stato capito.

Per vedere il debug relativo alle operazioni compiute dal riconoscitore vocale occorre attivare i messaggi verbosi al momento della compilazione all'interno della videata delle opzioni di compilazione.

3.1 Grammatica

I file *voice_ita.xml* e *voice_eng.xml* definiscono la grammatica dei comandi accettati dal riconoscitore vocale quando l'applicazione è compilata rispettivamente in italiano e in inglese. Questi file contengono innanzitutto la lista delle parole del linguaggio, cioè parole che non devono essere prese in considerazione nell'interpretazione del comando vocale. Fanno parte di questa categoria, per esempio, gli articoli e le preposizioni.

In seguito vengono descritti tutti i comandi riconosciuti. Per esempio, il comando che permette di andare alla prima riga di un pannello è il seguente:

```
<COMMAND tag="gotop" text="Vado alla prima riga nel pannello |1">
```

```

<SYN>
{VAI, MOSTRA, TORNA, * }; {CIMA, INIZIO}; {LISTA, ELENCO, PANNELLO, TABELLA, GRIGLIA, *}
</SYN>
<SYN>
{VAI, MOSTRA, TORNA, * }; {PRIMA}; {RIGA, RECORD}; {LISTA, ELENCO, PANNELLO, TABELLA, GRIGLIA, *}
</SYN>
</COMMAND>

```

I comandi sono descritti nell'elemento `<COMMAND>` che ha due attributi tag e text. Il primo è il nome del comando, mentre il secondo è il testo che viene mostrato a video come risposta dopo l'esecuzione del comando stesso.

Ogni comando è costituito da uno o più sinonimi, descritti all'interno dell'elemento `<SYN>`, che rappresentano le frasi che si possono pronunciare per fare in modo che quel particolare comando vocale sia riconosciuto. L'elemento `<SYN>` contiene dei gruppi di parole racchiuse tra parentesi graffe. Ogni gruppo rappresenta una parola del sinonimo e le parole contenute al suo interno possono essere usate in modo interscambiabile per impartire quel comando. Affinché il comando venga riconosciuto, il riconoscitore deve essere in grado di trovare una parola per ogni gruppo di un sinonimo del comando stesso. Fanno eccezione i gruppi al cui interno è presente un asterisco. Questo indica che quel gruppo è opzionale e, anche se non viene trovata nessuna parola che ne fa parte, il comando viene preso in considerazione lo stesso.

Facciamo un esempio sulla base del comando mostrato in precedenza. Per fare in modo che il cursore si porti sulla prima riga del pannello, l'utente può pronunciare diverse frasi, tra cui:

- 1) *“Vai all’inizio del pannello Aziende”*. In questo modo il comando viene riconosciuto grazie al primo sinonimo: le parole *all* e *del* non verranno prese in considerazione perché fanno parte delle parole del linguaggio; la parola *vai* viene trovata nel primo gruppo, la parola *inizio* viene trovata nel secondo gruppo e la parola *pannello* viene trovata nel terzo gruppo. È stata trovata una parola per tutti i gruppi quindi il comando viene riconosciuto. La parola *Aziende* è, per il momento, memorizzata come parola sconosciuta e viene presa in considerazione nella seconda fase del riconoscimento per la valutazione del comando nel contesto.
- 2) *“Prima riga”*. In questo modo il comando viene riconosciuto grazie al secondo sinonimo: non viene trovata nessuna parola nel primo gruppo, ma questo è opzionale quindi va bene lo stesso. Viene trovata la parola *prima* nel secondo gruppo, viene trovata la parola *riga* nel terzo gruppo, non viene trovata nessuna parola nel quarto gruppo, ma questo era opzionale quindi va bene lo stesso. È stata trovata una parola per tutti i gruppi obbligatori, quindi il comando viene riconosciuto.

I file `voice_lang.xml` possono essere modificati per personalizzare il comportamento del riconoscitore vocale. Inoltre si possono aggiungere dei nuovi file che descrivano la

grammatica dei comandi in lingue diverse dall'italiano o dall'inglese e che permetteranno l'utilizzo dei comandi vocali anche in applicazioni compilate in altre lingue.

3.2 Contesto

Il contesto dell'applicazione è costituito dalle sue videate, pannelli, campi di pannelli, menu, voci di menu e procedure. È rappresentato nel file *prjVoice.xml* che viene generato al momento della compilazione se è attivo il parametro di compilazione *Abilita comandi vocali*. Se l'applicazione viene compilata in C#, il file si trova nel percorso *CSharp/DB*, mentre se l'applicazione è compilata in Java il file si trova in *Servlet/web*. Per ogni oggetto i dati importanti per definire il contesto sono i seguenti:

- 1) *Guid*: identificativo dell'oggetto.
- 2) *Name*: corrisponde al nome dell'oggetto specificato nell'IDE.
- 3) *Description*: corrisponde alla descrizione dell'oggetto specificata nell'IDE.
- 4) *Title*: corrisponde al titolo specificato nell'IDE.
- 5) *ParentGuid*: identificativo dell'oggetto padre.
- 6) *CustomData*: dati specifici per ogni tipo di oggetto.

Sulla base di queste informazioni, nella seconda fase del riconoscimento del comando vocale, si individuano gli oggetti sui quali eseguire i comandi. È importante impostare con attenzione i campi *titolo*, *nome* e *descrizione* perché le parole usate all'interno di essi vengono utilizzate per l'individuazione degli oggetti. In particolare, viene attribuita maggiore importanza alle parole contenute nel titolo, perché sono quelle che vengono visualizzate dall'utente. Seguono le parole contenute nel nome e le parole contenute nella descrizione. Le altre informazioni vengono impostate automaticamente da InDe.

Nella valutazione del contesto si tiene conto anche dello stato in cui si trova l'applicazione al momento dell'esecuzione del comando. Vengono analizzate le proprietà di ogni pannello, in che stato si trova e se ha il fuoco. Queste informazioni sono utili per decidere se il pannello può essere oggetto dell'azione del comando e, nel caso in cui il pannello non è stato indicato nel comando, servono anche per decidere su quale pannello agire.

3.2 La libreria Voice Manager

Le proprietà e le funzioni contenute nella libreria *VoiceManager* sono le seguenti:

<i>Listening</i>	Se impostata a <i>true</i> attiva il riconoscimento vocale, altrimenti lo disattiva.
------------------	--

Skip	Se impostato a <i>true</i> il riconoscimento vocale standard non avrà luogo.
Matching	Interpreta ed esegue un testo, trattandolo come un comando vocale
SetResponse	Permette di impostare la risposta del riconoscitore vocale.
CanSpeak	Se impostata a <i>true</i> attiva il sintetizzatore vocale, altrimenti lo disattiva.
Say	Permette di utilizzare il sintetizzatore vocale per far dire qualcosa all'applicazione.

Nelle applicazioni web, per fare in modo che l'applicazione inizi ad ascoltare quello che viene detto dall'utente, occorre premere la sequenza di tasti *ctrl + spazio*. In questo modo, il parametro *Listening* viene impostato automaticamente a *true*, compare l'icona del microfono con la nuvoletta del testo e l'utente può iniziare a parlare. Per fare in modo che l'applicazione smetta di ascoltare occorre premere *esc* oppure pronunciare il comando *Basta*. In questo modo il parametro *Listening* viene impostato a *false*, il microfono e la nuvoletta del testo scompaiono e l'audio non viene più acquisito.

Nelle applicazioni mobile, invece, il parametro *Listening* deve essere valorizzato dallo sviluppatore, per esempio in una procedura attaccata a un bottone. Si può fare lo stesso anche nelle applicazioni web per permettere all'utente di iniziare il riconoscimento vocale anche attraverso modalità diverse rispetto alla pressione dei tasti predefiniti.

Se si vuole far partire il riconoscimento vocale alla pressione di un bottone, si può scrivere il seguente codice:

```

public void Welcome.StartButton()
{
    if (!(VoiceManager.listening))
        VoiceManager.listening = true
}

```

Come abbiamo visto all'inizio del paragrafo 3 *Funzionamento del riconoscimento vocale*, la prima cosa che succede quando il sistema rileva un comando vocale è l'evento *OnCommand*. Quest'evento si può usare per cambiare il comportamento standard del riconoscitore, intercettando per esempio il comando pronunciato dall'utente e decidendo in base a questo se continuare con il riconoscimento o no. Per interrompere il riconoscimento di un comando vocale e impedire al sistema di interpretarlo, occorre impostare la proprietà *Skip* a *true*.

Per ottenere il testo del comando pronunciato dall'utente, occorre leggere il parametro *text* attraverso la funzione *GetUrlParameter*. Si può anche modificare il testo restituito come risposta dal riconoscitore con *SetResponse*.

La funzione *Matching* può essere usata per fare in modo che il riconoscitore vocale interpreti ed esegua un testo, trattandolo come un comando vocale pronunciato dall'utente. A differenza delle proprietà viste prima, non deve per forza essere chiamata nell'evento *OnCommand*.

Se, ad esempio, si vogliono memorizzare i comandi impartiti e la relativa risposta, si può scrivere il seguente codice:

```
event VoiceCrm.OnCommand(  
  string Command //  
)  
{  
  if (Command == "VOICE")  
  {  
    // Stop the standard recognizer  
    VoiceManager.skip = true  
    //  
    // Get the voice command transcription  
    string text = VoiceCrm.getURLParam("text")  
    //  
    // Get the response of the recognizer  
    string res = VoiceManager.matching(text)  
    //  
    // Save in the database the voice command and the response  
    insert values into Results (last value variable)  
    f(set VoiceCommand = text  
    f(set ExecutedCommand = res  
    //  
    // Show the response  
    VoiceManager.setResponse(res)  
  }  
}
```

Dopo l'esecuzione di un comando vocale l'applicazione leggerà la risposta restituita. Il sintetizzatore vocale è attivo di default quando si utilizzano i comandi vocali. Per disattivarlo occorre impostare a false la proprietà *CanSpeak*.

È possibile far leggere all'applicazione un testo qualsiasi utilizzando la procedura *Say* che prende come parametro il testo che il sintetizzatore dovrà riprodurre. Questa procedura può essere utilizzata anche se i comandi vocali non sono abilitati.

L'utilizzo delle proprietà e dei metodi presenti nella libreria non è indispensabile per il funzionamento del riconoscitore vocale ma serve soltanto se si desidera modificare il suo comportamento standard.

4 Personalizzazione

L'interfaccia del riconoscitore vocale delle applicazioni web può essere personalizzata modificando le classi css *voice-microext* e *voice-microint* per quanto riguarda l'icona del microfono e *voice-bubble* e la *voice-bubblewhisker* per la nuvoletta.

È inoltre possibile modificare il testo iniziale mostrato dalla nuvoletta personalizzando nel file *custom3.js* la proprietà `IDV_WELCOME_MSG` personalizzando la funzione `RD3_CustomInit` come nell'esempio seguente:

```
function RD3_CustomInit()
{
  ClientMessagesSet['ITA'].IDV_WELCOME_MSG = 'nuovo_testo';
  ClientMessagesSet['ENG'].IDV_WELCOME_MSG = 'nuovo_testo';
}
```

Il riconoscitore aspetta un massimo di 2 secondi dall'ultima parola rilevata prima di inviare il comando al riconoscitore. Il tempo massimo di attesa può essere modificato personalizzando la proprietà *VoiceAutoCommitDelay*. A questo scopo occorre personalizzare nel file *custom3.js* la funzione `RD3_CustomInit` come nell'esempio seguente:

```
function RD3_CustomInit()
{
  RD3_ClientParams.VoiceAutoCommitDelay = numero_di_millisecondi;
}
```

Per quanto riguarda il funzionamento, oltre alla personalizzazione del comportamento standard, che si può ottenere con i metodi e le procedure della libreria, un'ulteriore personalizzazione può essere compiuta sulla grammatica. Aggiungendo o togliendo delle parole nei sinonimi di ogni comando, si può permettere il riconoscimento di ognuno di essi anche con parole non previste dalla versione standard della grammatica.

È possibile inoltre aggiungere dei nuovi file *voice_lang.xml*, dove *lang* deve essere sostituito con il codice della lingua, per permettere l'utilizzo del riconoscitore vocale anche con lingue diverse dall'italiano o dall'inglese. Il sistema seleziona la grammatica corretta sulla base della lingua in cui è stata compilata l'applicazione.