

Sistema vocal WireLess TelNet voiXtreme

Descripción

El sistema vocal WireLess TelNet de SofToGo permite a la PDA leer textos en voz alta y realizar capturas ASR utilizando una aplicación telnet estándar (VT / 5250).

Sistema TTS

El texto a leer a través de TTS puede provenir de:

- comandos de impresión TnVT (**ESC 5i** / **ESC 4i**)
- texto de la pantalla, determinado por etiquetas de inicio y fin
- texto constante y variable perteneciente a los "Voice Triggers"

Sistema ASR

Ingreso vocal del usuario. El sistema ASR reconoce la voz del usuario y la convierte en texto utilizando "gramáticas". El ingreso es procesado por la aplicación local y luego es enviado al host de acuerdo con la configuración de programación.

Las gramáticas son algoritmos de reconocimiento diseñados para producir textos. Las gramáticas disponibles son:

- "Digits": acepta números decimales del "0" al "9", y las órdenes "Anular" y "Repetir". Las palabras clave *Anular* y *Repetir* son procesadas por la biblioteca de la voz y realizan transmisiones al host siguiendo el algoritmo correspondiente al ingreso.
- "Commands": acepta las órdenes "Aceptar", "Anular" y "Repetir". Las palabras clave *Aceptar*, *Anular* y *Repetir* son procesadas por la biblioteca de la voz y realizan transmisiones al host siguiendo el algoritmo correspondiente al ingreso.

Las palabras clave difieren según el idioma elegido para el ASR.

Voice Triggers

Los Voice triggers están compuestos por tres secciones:

- *Condiciones*: contiene condiciones que deben cumplirse para realizar un evento
- *TTS*: permite dar una instrucción vocal al usuario
- *ASR*: permite obtener información del usuario

Los "Voice triggers" arrancan el proceso de reconocimiento vocal. Su ejecución es determinada por la presencia de condiciones específicas en la pantalla del terminal, tales como:

- texto hallado en una parte específica de la pantalla
- posición del cursor

Se pueden definir uno o más triggers para realizar eventos vocales. Dichos eventos pueden ser exclusivamente TTS, exclusivamente ASR o TTS + ASR.

Dentro del Voice Trigger, puede definirse un anuncio vocal para decir antes de que comience el reconocimiento ASR. Este anuncio puede ser una mezcla de texto constante (definido en el archivo de configuración) y texto variable (leído de la pantalla del terminal). El texto variable debe ser definido por posición de inicio y longitud: tres valores numéricos (fila y columna de la posición de inicio del texto en la pantalla y su longitud) separados por comas y encerrados entre llaves: {f, c, l}. Por ejemplo, {10, 7, 14} define un texto ubicado en la fila 10, columna 7, de una longitud de 14 caracteres.

Para el reconocimiento ASR es necesario definir la gramática y el algoritmo ASR a utilizar.

Las gramáticas vocales estándar, "Digits" y "Commands", están incluidas en las bibliotecas voiXtreme. Más adelante se desarrollarán gramáticas personalizadas.

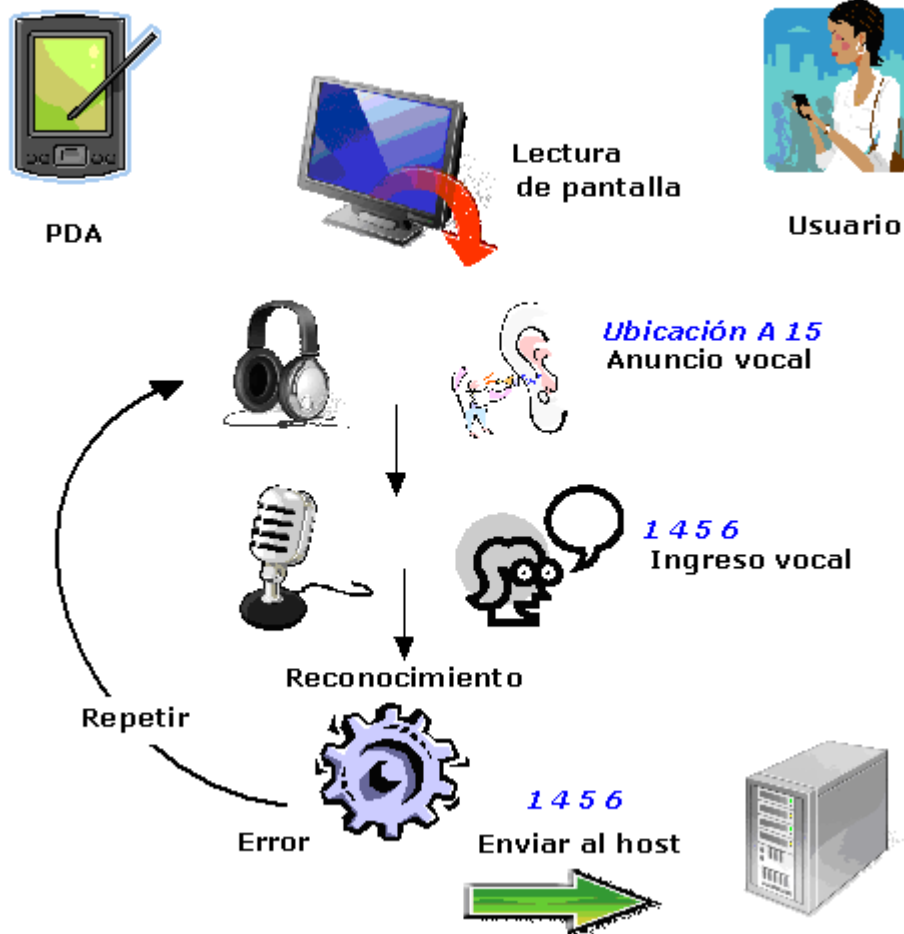
El algoritmo ASR puede ser:

- *Simple* (Confirm=No): el motor reconoce el ingreso del usuario, luego lo envía al host.
- *Simple con repetición* (Repeat=Yes): el motor reconoce el ingreso del usuario, lo repite a través de TTS, luego lo envía al host.
- *Validar* (Confirm=Yes): el motor reconoce el ingreso del usuario, lo repite y espera un comando del usuario (“*aceptar*” / “*anular*”):
 - Si el usuario dice "anular", la aplicación reinicia el trigger desde el comienzo (anuncio vocal).
 - Si el usuario dice "aceptar", la aplicación envía el ingreso del usuario al host.
- *Validar con repetición* (Repeat=Yes): el procedimiento es el mismo que en *Validar*, salvo que el comando de validación es repetido al usuario.

Modos de ingreso

Las siguientes imágenes describen los diferentes modos de ingreso de datos.

Ingreso Vocal Simple



Algoritmo de Ingreso Vocal Simple:

1/ Hay una coincidencia entre los datos del trigger y los datos en pantalla y la posición del cursor.

2/ Se procesa la sección TTS del trigger para hacer el anuncio vocal al usuario.

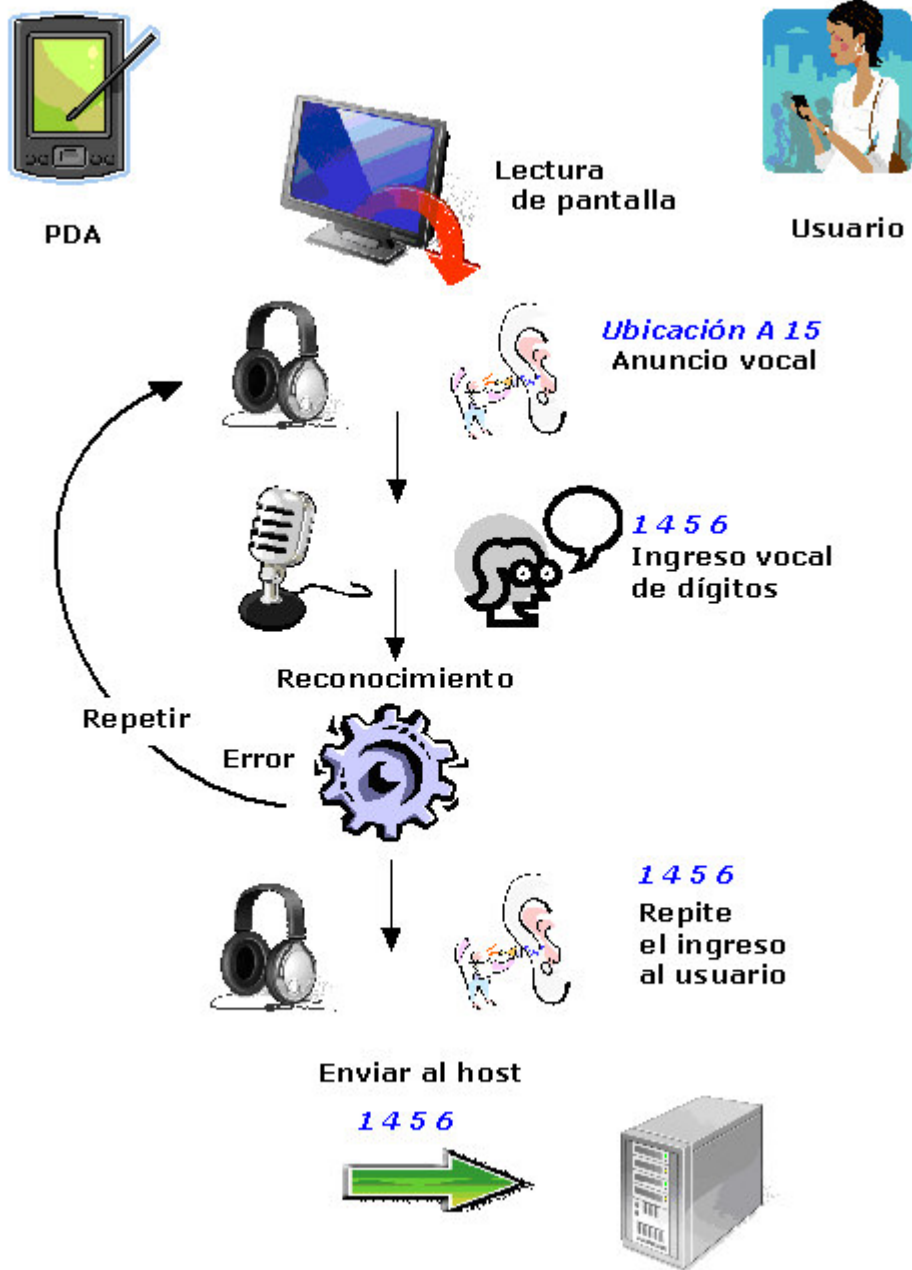
3/ Espera que termine el anuncio TTS.

4/ Se procesa la sección ASR del trigger para realizar una captura ASR, utilizando la gramática indicada en esta sección (en este caso, "Digits").

5/ Fin del reconocimiento:

- Repetir, vuelve a (2).
- Error (sin reconocimiento, bajo nivel de confianza, min / max) vuelve a (2).
- Anular, envía la cadena [\[VOICETRIGGER ASR XX\]](#)/Cancel al Host, luego sale.
- Dígitos, envía el ingreso, seguido de la cadena [\[VOICETRIGGER ASR XX\]](#)/Terminator, al Host, luego sale.

Simple + Repetición del Ingreso Vocal



Algoritmo del Ingreso Vocal Simple + Repetición del Ingreso:

1/ Hay una coincidencia entre los datos del trigger y los datos en pantalla y la posición del cursor.

2/ Se procesa la sección TTS del trigger para hacer el anuncio vocal al usuario.

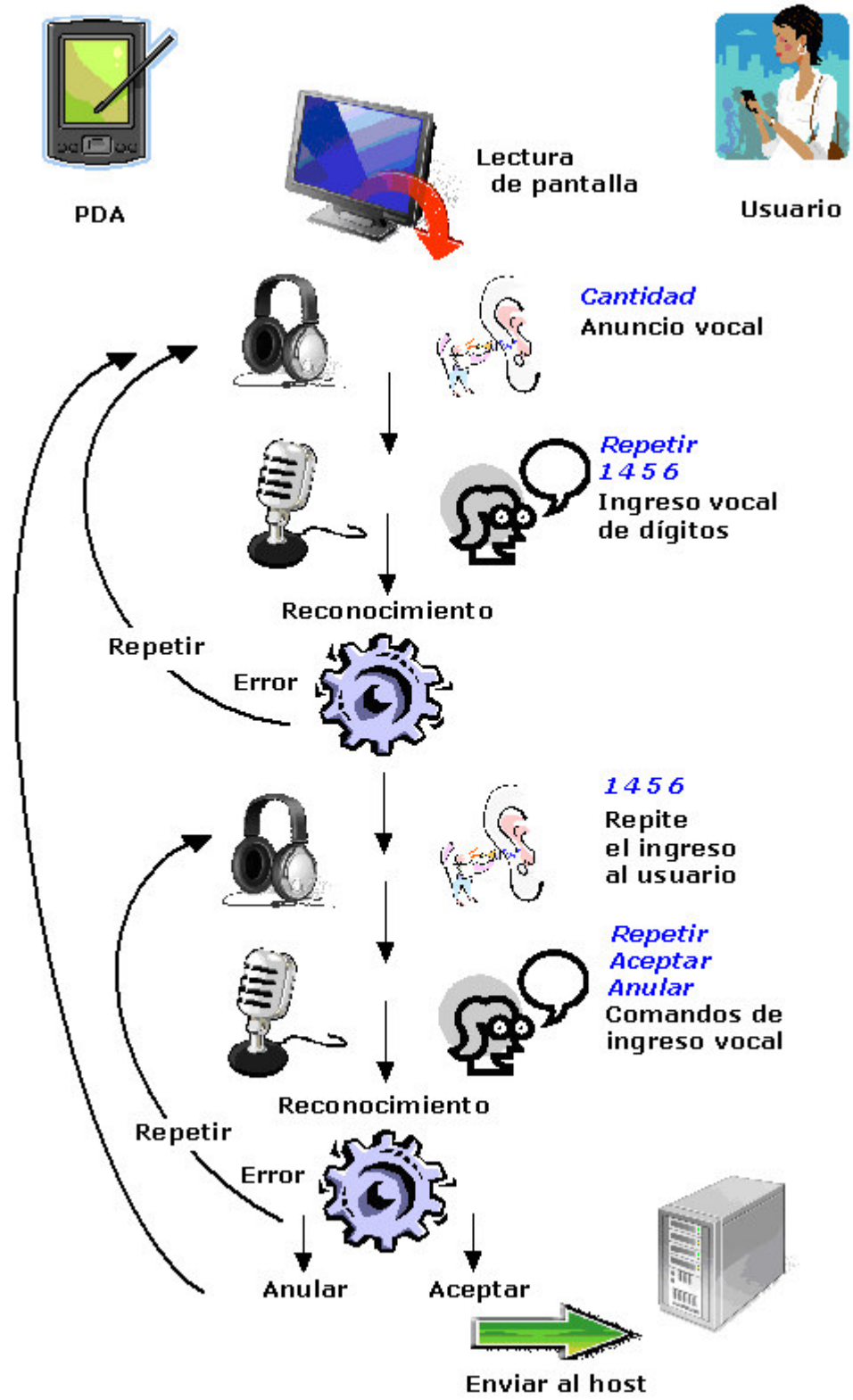
3/ Espera que termine el anuncio TTS.

4/ Se procesa la sección ASR del trigger para realizar una captura ASR, utilizando la gramática indicada en esta sección (en este caso, "Digits").

5/ Fin del reconocimiento:

- Repetir, vuelve a (2).
- Error (sin reconocimiento, bajo nivel de confianza, min / max) vuelve a (2).
- Anular, envía la cadena [\[VOICETRIGGER ASR XX\]](#)/Cancel al Host, luego sale.
- Dígitos, envía el ingreso, seguido de la cadena [\[VOICETRIGGER ASR XX\]](#)/Terminator, al Host, y le repite al usuario el valor del ingreso, luego sale.

Validar Ingreso Vocal (ingreso de dígitos)



Algoritmo de Validación del Ingreso Vocal (dígitos):

1/ Hay una coincidencia entre los datos del trigger y los datos en pantalla y la posición del cursor.

2/ Se procesa la sección TTS del trigger para hacer el anuncio vocal al usuario.

3/ Espera que termine el anuncio TTS.

4/ Se procesa la sección ASR del trigger para realizar una captura ASR, utilizando la gramática indicada en esta sección (en este caso, "Digits").

5/ Fin del reconocimiento:

- Repetir, vuelve a (2).
- Error (sin reconocimiento, bajo nivel de confianza, min / max) vuelve a (2).
- Anular, continúa en (9).
- Dígitos, continúa en (6).

6/ Se repite el ingreso al usuario a través de TTS.

7/ Se realiza el reconocimiento ASR, utilizando la gramática "Commands".

8/ Fin del reconocimiento:

- Repetir, vuelve a (6).
- Error (sin reconocimiento, bajo nivel de confianza, min / max) vuelve a (6).
- Anular, continua (6).
- Aceptar, dice "Aceptar" al usuario, envía el ingreso, seguido de la cadena [\[VOICETRIGGER_ASR_XX\]](#)/Terminator, al Host, luego sale.

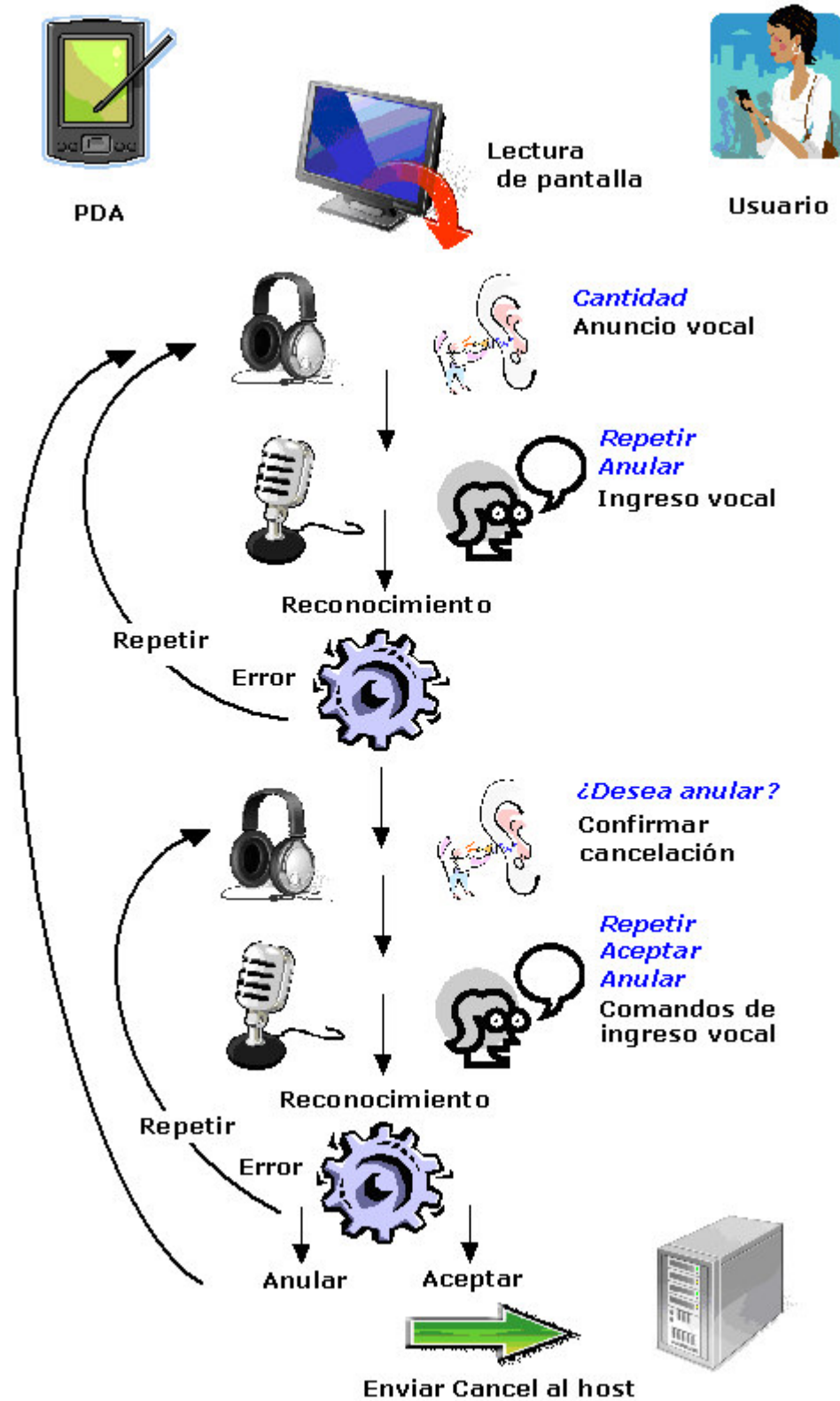
9/ Dice la frase *CancelConfirmation* al usuario a través de TTS.

10/ Realiza el reconocimiento ASR utilizando la gramática "Commands".

11/ Fin del reconocimiento:

- Repetir, vuelve a (9).
- Error (sin reconocimiento bajo nivel de confianza, min / max) vuelve a (9).
- Anular, dice "Anular" al usuario, vuelve a (2).
- Aceptar, dice "Aceptar" al usuario, envía la cadena [\[VOICETRIGGER_ASR_XX\]](#)/Cancel al Host, luego sale.

Validar Ingreso Vocal (anular ingreso)



Algoritmo de Validación del Ingreso Vocal (anular):

1/ Hay una coincidencia entre los datos del trigger y los datos en pantalla y la posición del cursor.

2/ Se procesa la sección TTS del trigger para hacer el anuncio vocal al usuario.

3/ Espera que termine el anuncio TTS.

4/ Se procesa la sección ASR del trigger para realizar una captura ASR, utilizando la gramática indicada en esta sección (en este caso, "Digits").

5/ Fin del reconocimiento:

- Repetir, vuelve a (2).
- Error (sin reconocimiento, bajo nivel de confianza, min / max) vuelve a (2).
- Anular, continúa en (9).
- Dígitos, continua en (6).

6/ Repite el ingreso al usuario a través de TTS

7/ Se realiza el reconocimiento ASR utilizando la gramática "Commands".

8/ Fin del reconocimiento:

- Repetir, vuelve a (6).
- Error (sin reconocimiento, bajo nivel de confianza, min / max) vuelve a (6).
- Anular, vuelve a (6).
- Aceptar, dice "Aceptar" al usuario, envía el ingreso, seguido de la cadena [\[VOICETRIGGER ASR XX\]](#)/Terminator, al Host, luego sale.

9/ Dice la frase *CancelConfirmation* al usuario a través de TTS.

10/ Se realiza el reconocimiento ASR utilizando la gramática "Commands".

11/ Fin del reconocimiento:

- Repetir, vuelve a (9).
- Error (sin reconocimiento, bajo nivel de confianza, min / max) vuelve a (9).
- Anular, dice "Anular" al usuario, vuelve a (2).
- Aceptar, dice "Aceptar" al usuario, envía la cadena [\[VOICETRIGGER ASR XX\]](#)/Cancel al Host, luego sale.

Lectura de pantalla

Ejemplo de lectura de pantalla

En este ejemplo, una parte de la pantalla (delimitada por SAY: / :SAY) es leída al usuario a través de TTS.

```
** Order Picking OPL-LOC  
  
Order Number   : 1259-AA0  
Type: Mandatory  
Bin Location   : A-L-008  
Bin-Loc ID:      
SAY: Go to ALPHA LIMA 8 :SAY
```

El texto delimitado por las etiquetas “SAY: - :SAY” es leído en voz alta al usuario; se oirá “*go to alpha lima eight*”.

Comandos de impresión Escape

Este ejemplo muestra un texto enviado al terminal dentro de un comando de impresión DEC-VT.

El caracter de escape (0xc1B) es representado como <ESC>.

```
<ESC>[5 IR A EQUIS ELE 5 <ESC>[4i
```

El usuario oirá “*ir a equis ele cinco*”.

Archivo de configuración

A continuación se realiza una descripción del archivo de configuración Voice.

El archivo de configuración es un archivo "punto ini" conformado por secciones (delimitadas por corchetes), claves (palabras clave dentro de la sección, ubicadas junto al signo "=") y valores (texto ubicado luego del signo "="). Las secciones y las claves están en una fuente especial: **SECCIÓN**. Los valores se encuentran en verde.

Definiciones globales del motor vocal

Sección TTS (Text-to-speech)

[TTS_DEF]

Esta sección contiene las opciones de configuración del Text-To-Speech (TTS) y permite activar el procesamiento de la pantalla o la impresora en acciones vocales.

VoiceSpeed=70

Velocidad relativa del habla (0 - 100). El usuario puede modificar este valor.

TtsPrinterOn=Yes

Permite que los comandos de impresión de TnVT (Esc 5 I) sean capturados por el motor TTS en lugar de ser enviados a la impresora, de manera que todo texto enviado a la impresora será leído en voz alta por TTS.

TtsScreenOn=Yes

Permite a la PDA decir a través de TTS algunos de los textos mostrados en la pantalla. La aplicación busca en el display ciertos patrones de inicio y fin (definidos en el archivo de configuración) para que el motor TTS lea en voz alta el texto delimitado.

Delimitadores de texto en pantalla

TtsScrSayTagBeg=:SAY:

Patrón de inicio del texto a buscar en el display del terminal para leerlo por TTS.

TtsScrSayTagEnd=:SAY

Patrón de fin del texto a buscar en el display del terminal para leerlo por TTS.

TtsScrSplTagBeg=:SPL:

Patrón de inicio del texto a buscar en el display del terminal para **deletrearlo** por TTS.

TtsScrSplTagEnd=:SPL

Patrón de fin del texto a buscar en el display del terminal para **deletrearlo** por TTS.

Mensajes al usuario

Mensajes que serán dichos al usuario bajo condiciones específicas.

MsgWelcome=Bienvenido

Mensaje a decir al arrancar.

MsgConnecting=Conectando...

Mensaje a decir al conectarse.

MsgConnected=Conectado

Mensaje a decir al haberse conectado.

MsgDisconnect=Fin de la conexión

Mensaje a decir al desconectarse.

Teclas para el control del motor vocal

KeyRepeat=2070

Tecla de repetición del último mensaje TTS.

KeyVolUp=2071

KeyVolDn=2072

Teclas para cambiar el volumen del TTS.

KeySpeedUp=2073

KeySpeedDn=2074

Teclas para cambiar la velocidad del TTS.

Opciones del traductor de caracteres especiales

TranslatorUse=No

TranslatorName=application\WTnVt\Fnt_PC850.cfg

Traductor de caracteres utilizado para codificar caracteres especiales. Permite realizar traducciones carácter por carácter.

Sección ASR (Automatic Speech Recognition)

[ASR_DEF]

Esta sección contiene las opciones de configuración del motor ASR.

ReliabilityLvl=4000

Nivel de confianza mínimo para aceptar el reconocimiento de una palabra. Al reconocer una palabra, el motor ASR devuelve un "nivel de confianza", valor que representa la confianza que asigna el motor al reconocimiento. Un nivel de confianza bajo implica un reconocimiento confuso que será rechazado.

CustomGrammar=<custom>

Nombre de la gramática referida por el parámetro de modo de entrada de datos en [VOICETRIGGER_ASR](#) / Grammar.

ResCancel=Anular

Esta es la palabra clave perteneciente a la gramática "commands" que el usuario debe decir para anular una acción.

ResAccept=Aceptar

Esta es la palabra clave perteneciente a la gramática "commands" que el usuario debe decir para aceptar una acción.

ResRepeat=Repetir

Esta es la palabra clave perteneciente a la gramática "commands" que el usuario debe decir para pedir una repetición de un mensaje (anuncio vocal antes del ASR).

SyncTts=Yes

Este parámetro hace que el ASR espere hasta que el TTS (el discurso de la PDA) finalice antes de comenzar el reconocimiento de la voz del usuario.

Voice Triggers

Los Voice triggers son secciones que permiten configurar la realización de acciones vocales según ciertas condiciones a hallar en la pantalla del terminal. A su vez, un trigger está compuesto por secciones:

- la sección de condiciones (condiciones que deben ser a satisfechas para que se ejecuten las acciones del trigger): [VOICETRIGGER_XX].
- la sección TTS (acciones TTS a realizar): [VOICETRIGGER_TTS_XX]
- la sección ASR (acciones ASR a realizar): [VOICETRIGGER_ASR_XX]

Es posible definir uno a más triggers.

Trigger 1 ...

[VOICETRIGGER_01]

[VOICETRIGGER_TTS_01]

[VOICETRIGGER_ASR_01]

Trigger 2 ...

[VOICETRIGGER_02]

[VOICETRIGGER_TTS_02]
[VOICETRIGGER_ASR_02]

Trigger 3 ...

[VOICETRIGGER_03]
[VOICETRIGGER_TTS_03]
[VOICETRIGGER_ASR_03]

...

Sección de condiciones del Voice trigger

En esta sección se definen las condiciones que se buscarán en la pantalla del terminal para ejecutar el trigger.

Hay dos tipos de condiciones: posición del cursor y condiciones propias del texto. Estas últimas son condiciones excluyentes ("Y" lógico), es decir, deben cumplirse todas para que el trigger sea ejecutado.

[VOICETRIGGER_XX]

On=Yes

Activar / desactivar este trigger.

Condiciones del primer texto a encontrar en la pantalla del terminal

Row_1=4

Fila en la que se buscará el texto. 0 (cero) significa todas las filas.

Col_1=4

Columna en la que se buscará el texto. 0 (cero) significa todas las columnas.

Fila=0 y Columna=0 significa en cualquier parte de la pantalla.

Match_1=Show System Summary

Texto a buscar en la posición Row_1 / Col_1.

Para que los espacios finales sean tomados en consideración, el texto debe ser delimitado por comillas "".

Si no se ingresa ningún texto, no se evaluará la condición (VERDADERO).

Condiciones del segundo texto a encontrar en la pantalla del terminal

Row_2=5

Col_2=4

Match_2=Main Menu

Condiciones del tercer texto a encontrar en la pantalla del terminal

Row_3=0

Col_2=0

Match_2=

Condiciones de la posición del cursor

CursorAtRow=8

Fila en la que está el cursor a evaluar. 0 (cero) significa en cualquier fila.

CursorAtCol=20

Columna en la que está el cursor a evaluar. 0 (cero) significa en cualquier columna.
Fila=0 y Col=0 significa en cualquier parte de la pantalla.

Sección de las acciones TTS del Voice trigger

[VOICETRIGGER_TTS_XX]

Éstas son las acciones TTS que se realizarán al ejecutar el trigger.

La aplicación leerá al usuario texto constante (definido en el trigger) y texto variable (extraído de la pantalla de la terminal).

DoTts=Yes

Habilita /deshabilita las acciones TTS en este trigger.

TtsSay01=any constant text{r,c,l} another constant text {r,c,l} final text.

Combinación de texto constante (texto normal) y texto variable (encerrado entre llaves {}). El texto variable es indicado por tres valores numéricos que señalan su ubicación por fila y columna, y su longitud.

Todo el texto es leído en voz alta, salvo las partes delimitadas por etiquetas especiales. Las etiquetas especiales <SPL></SPL> delimitan texto a deletrear, por ejemplo, el texto “*emplazamiento* <SPL>XL5</SPL>” será leído como “*emplazamiento equis ele cinco*”.

TtsSay02= any constant text{r,c,l} another constant text {r,c,l} final text.

Es posible configurar varias acciones TTS.

Sección de las acciones ASR del Voice trigger

[VOICETRIGGER_ASR_XX]

Éstas son las acciones ASR que se realizarán en el trigger:

DoAsr=Yes

Habilita /deshabilita el ASR en este trigger.

Grammar=1

Gramática a utilizar en este trigger. Una gramática es un conjunto de reglas y palabras utilizado por el motor ASR para el reconocimiento vocal. El motor ASR estándar incluye dos gramáticas: *commands* (Anular / Aceptar) y *digits* (0, 1, 2, 3, 4, 5, 6, 7, 8, 9). Se desarrollarán gramáticas personalizadas más adelante.

Valores: 0=commands, 1=digits, 2=custom.

LenMin=0

Mínima longitud de ingreso vocal esperada. Los ingresos inferiores a este valor serán rechazados y el sistema solicitará un nuevo ingreso.

LenMax=32

Máxima longitud de ingreso vocal esperada. Los ingresos superiores a este valor serán rechazados y el sistema solicitará un nuevo ingreso.

Confirm=Yes

Activa el algoritmo de ingreso de datos *Validar* o *Simple* para el ASR.

Para una explicación más detallada de los algoritmos de ingreso de datos, vea las descripciones de los algoritmos y las ilustraciones del apartado [Modos de ingreso](#).

Repeat=Yes

Activa el algoritmo Repetir para el ASR.

Para una explicación más detallada de los algoritmos de ingreso de datos, vea las descripciones de los algoritmos y las ilustraciones del apartado [Modos de ingreso](#).

CancelDo=Yes

Permite al usuario cancelar el ingreso de datos y envía la secuencia de cancelación al host.

Cancel=\1B\41

Secuencia enviada al host cuando el usuario cancela el ingreso.

CancelConfirmation=Desea anular?

Pregunta que la aplicación le hará al usuario para confirmar el comando de cancelación.

Terminator=\0D

Cadena de término que será enviada al host luego del ingreso ASR (añadida como sufijo) para validar el ingreso (en este caso, Enter).

Ejemplo de un archivo de configuración

```
//-----  
// Tts definitions  
[TTS_DEF]  
VoiceSpeed=70  
TtsPrinterOn=Yes  
TtsScreenOn=No  
TtsScrSayTagBeg=SAY:  
TtsScrSayTagEnd=:SAY  
TtsScrSplTagBeg=SPL:  
TtsScrSplTagEnd=:SPL  
MsgWelcome=Bienvenido  
MsgDisconnect=Fin de la conexión  
MsgConnected=Connectado  
MsgConnecting=Conectando...  
KeyRepeat=2070  
KeyVolUp=2071  
KeyVolDn=2072  
KeySpeedUp=2073  
KeySpeedDn=2074  
TranslatorUse=No  
TranslatorName=application\WTnVt\Fnt_PC850.cfg  
  
//-----  
// Asr definitions  
[ASR_DEF]  
ReliabilityLvl=4000  
ResCancel=Anular  
ResAccept=Aceptar  
ResRepeat=Repetir  
SyncTts=Yes  
  
//-----  
// Triggers  
//-----  
  
//-----  
// Voice trigger conditions  
[VOICETRIGGER_01]  
On=Yes  
Row_1=4  
Col_1=4  
Match_1=Show System Summary  
Row_2=0  
Col_2=0  
Match_2=  
CursorAtRow=0  
CursorAtCol=0  
  
// Voice trigger TTS to do  
[VOICETRIGGER_TTS_00]  
DoTts=Yes  
TtsSay01=Goto
```

```

TtsSay02=Location {1,4,7}
TtsSay03=Please

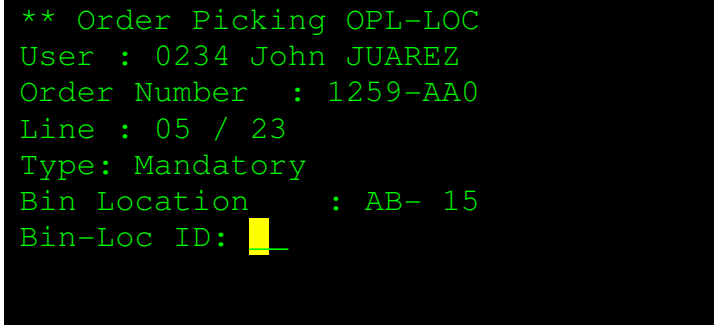
// Voice trigger ASR todo
[VOICETRIGGER_ASR_00]
DoAsr=Yes
Grammar=1
LenMin=4
LenMax=4
Confirm=Yes
Repeat=Yes
CancelDo=Yes
CancelConfirmation=Desea anular?
Cancel=\1B\41
Terminator=\0D

```

Ejemplos de pantallas y triggers

Este ejemplo explica, por una parte, la configuración del trigger necesaria para realizar el ingreso vocal para una operación de picking Bin-Location (emplazamiento de preparación de pedidos), por otra, describe la configuración del trigger para realizar un ingreso de cantidades en dos pantallas.

Pantalla de Bin location (emplazamiento)



```

** Order Picking OPL-LOC
User : 0234 John JUAREZ
Order Number : 1259-AA0
Line : 05 / 23
Type: Mandatory
Bin Location : AB- 15
Bin-Loc ID:  _

```

En este caso, se utiliza un algoritmo simple como ID de ubicación de dos dígitos conocida por el sistema. Luego, el ingreso es enviado sin confirmación y el host envía una nueva pantalla (si el ingreso es el esperado) o envía la misma pantalla (si no es el esperado).

```

PDA  "go to location ey be fifteen please"
User  "five seven"
PDA  "fifty seven"
PDA to host  "57<ENTER>"

```

```

[VOICETRIGGER_01]
On=Yes
Row_1=1
Col_1=1

```

```

Match_1=** Order Picking OPL-LOC
Row_2=0
Col_2=0
Match_2=
CursorAtRow=7
CursorAtCol=0

[VOICETRIGGER_TTS_01]
DoTts=Yes
TtsSay01=Go to
TtsSay02=Location <SPL>{6,19,2}</SPL>{6,22,3}
TtsSay03=Please

[VOICETRIGGER_ASR_01]
DoAsr=Yes
Grammar=digits
LenMin=2
LenMax=2
Confirm=No
Repeat=Yes
CancelDo=Yes
CancelConfirmation=Do you wish to cancel?
Cancel=\1B
Terminator=\0D

```

Pantalla de cantidad en la ubicación

```

** Order Picking OPL-QTY
User : 0234 John JUAREZ
Order Number : 1259-AA0
Line : 05 / 23
Type: Mandatory
Bin Location : AB- 15
Product : 1238767653490
Quantity: █_____

```

En este caso, se utiliza un algoritmo de validación para confirmar la cantidad ingresada. Sólo se dicen los últimos tres dígitos del producto.

```

PDA "product four hundred ninety, give quantity"
User "one four four"
PDA "one hundred and forty for"
User "accept"
PDA to host "144<ENTER>"

```

```

[VOICETRIGGER_02]
On=Yes
Row_1=1
Col_1=1
Match_1=** Order Picking OPL-QTY
Row_2=0

```

Col_2=0
Match_2=
CursorAtRow=8
CursorAtCol=0

[VOICETRIGGER_TTS_02]
DoTts=Yes
TtsSay01=Product {7,23,3}
TtsSay02=Give quantity

[VOICETRIGGER_ASR_02]
DoAsr=Yes
Grammar=digits
LenMin=1
LenMax=6
Confirm=Yes
Repeat=No
CancelDo=Yes
CancelConfirmation=Do you wish to cancel?
Cancel=\1B
Terminator=\0D