**UNIVERSIDAD MAYOR DE SAN SIMÓN**

**FACULTAD CIENCIAS Y TECNOLOGÍA**

**JUEGO DE PELEAS UTILIZANDO SPHINX4**

**(RECONOCIMIENTO   
ACÚSTICO FONÉTICO)**

Integrantes: Cari Aguilar Luis Alberto

Gumucio Suarez Fabiola

Miranda Rodríguez Ronald Rodrigo

Veliz Morales José

Materia: Reconocimiento de voz

Fecha: 4 de junio del 2019

Cochabamba – Bolivia

Contenido

[1 Reconocimiento de voz aplicada a un videojuego 2](#_Toc10532128)

[1.1 Introducción 2](#_Toc10532129)

[1.2 Reconocimiento acústico fonético 3](#_Toc10532130)

[1.3 Descripción breve del juego 3](#_Toc10532131)

[1.4 Diccionario 4](#_Toc10532132)

[1.5 Gramática 4](#_Toc10532133)

[1.6 Código del algoritmo 5](#_Toc10532134)

[1.7 Código principal 5](#_Toc10532135)

[1.8 Conclusiones 5](#_Toc10532136)

# 1 Reconocimiento de voz aplicada a un videojuego

## 1.1 Introducción

En la actualidad, existe una auténtica revolución conceptual en la informática. Podemos ver como algunos sistemas, Wii de Nintendo, han revolucionado el mundo de las consolas y videojuegos simplemente modificando el mando para que reconociese movimientos. Un reciente proyecto de Microsoft, proyecto Natal, nos permite a través de un sistema de webcams, interactuar directamente con la consola tan solo moviéndonos delante de ella. Cada vez más ordenadores y teléfonos móviles se dotan de pantallas táctiles que permiten directamente tocar lo que queremos.

En este contexto no podíamos dejar de lado los sistemas de reconocimiento de habla. Vemos cada día sistemas de atención telefónica con los que interactuamos a través de órdenes vocales. Nuevos sistemas operativos como Windows 7 o Mac OS X incluyen opciones de reconocimiento por voz. No podemos dejar de lado a Google, con su importante papel marcando tendencias tecnológicas. Actualmente ha mostrado una fuerte inclinación a este tipo de sistemas. Podemos ver en los últimos modelos de móviles del gigante como integran búsquedas en internet empleando tan solo la voz o búsquedas en su sistema Google Maps. En resumen, el reconocimiento de voz facilita en muchos aspectos la vida cotidiana de las personas.

## 1.2 Reconocimiento acústico fonético

El enfoque acústico fonético está basado en la teoría que postula que existe un número finito y diferenciado de unidades fonéticas en el lenguaje hablado, y que pueden ser caracterizadas por un conjunto de propiedades que se manifiestan en la señal de voz a lo largo del tiempo.

Engloba todos aquellos procesos destinados a realizar una decodificación de palabras a partir de las características diferenciadoras que la voz presenta y de un conjunto de reglas, dispuestas en forma de sistema experto, que existen en el habla.

Los fonemas se estructuran en palabras, y éstas en frases, que son las que modelan las ideas. Aunque las propiedades acústicas de las unidades fonéticas son altamente variables debido tanto a una variedad infinita de hablantes como a la vecindad de las unidades fonéticas (conocidas como las co-articulaciones de los sonidos), se asume que las reglas que gobiernan su variabilidad son sencillas y pueden ser aprendidas de manera rápida para poder implementarse en situaciones prácticas.

## 1.3 Descripción breve del juego

Super Smash Bros. utiliza un sistema de batalla distinto al de los tradicionales juegos de lucha. 2 o más jugadores (hasta 4) eligen personaje, tras lo cual pelean en diferentes escenarios mientras tratan de sacar a sus oponentes del escenario. En vez de usar barras de energía como la mayoría de los juegos de lucha, este juego presenta medidores de porcentaje. Estos medidores comienzan en 0%, y se incrementan cuando los personajes reciben daño, llegando a un límite de 999%. A medida que el porcentaje de un personaje incrementa, éste puede ser arrojado más lejos con mayor facilidad. Cuando un personaje es arrojado fuera de los límites de la pantalla pierde una vida o un punto dependiendo del modo de juego en que se esté participando. Los jugadores pueden usar ataques u otras técnicas para volver al escenario (generalmente la combinación de botones ↑+B), o como algunos saltos de ciertos personajes, que tienen un mayor alcance o mayor repetición, y que pueden hacer que la recuperación sea más fácil. Además, algunos personajes son más pesados que otros, haciendo más difícil para un adversario conseguir sacarlo del escenario, pero haciendo más difícil la recuperación. Los personajes pueden pelear usando una gran variedad de ataques resultantes de combinar los botones A y B del control con la dirección que se da a la palanca de control.

El juego contiene 9 escenarios distintos. Los escenarios son arenas de tres dimensiones, basadas en niveles de las distintas sagas de Nintendo y son interactivos con el jugador. Aunque los escenarios están dibujados en tres dimensiones, los personajes sólo se pueden mover en un plano bidimensional. No todos los escenarios están disponibles desde el inicio; un escenario está oculto, y puede ser desbloqueado si se siguen unos requerimientos. ​ Los escenarios van desde plataformas flotantes hasta áreas de tierra firme. Cada escenario posee un límite, el cual si es traspasado hará que el personaje pierda una vida o punto. Además, diversos objetos aparecen de forma aleatoria y de distintas formas y consecuencias al usarlo, como bates de béisbol, caparazones, martillos, y su uso comprende desde recuperar energía hasta aportársela al contrario.

## 1.4 Diccionario

Para llenar parcialmente nuestro diccionario se registraron combinaciones de las palabras que reconoce nuestro proyecto, para el proyecto el diccionario es pequeño lo que reduce la probabilidad de que el reconocedor de voz interprete erróneamente las palabras. Las palabras que utilizamos son las siguientes:

* abajo a b a j o
* combinar k o m b i n a r
* combinó k o m b i n o
* derechito d e r e ch i t o
* derecho d e r e ch o
* derechos d e r e ch o s
* golpe g o l p e
* golpea g o l p e a
* golpeada g o l p e a d a
* golpeado g o l p e a d o
* izquierda i z k i e r d a
* izquierdo i z k i e r d o
* jugado j u g a d o
* jugador j u g a d o r
* jugadora j u g a d o r a
* jugadoras j u g a d o r a s
* jugadores j u g a d o r e s
* pata p a t a
* patada p a t a d a
* patagonia p a t a g o n i a
* salto s a l t o
* saltó s a l t o
* salud s a l u d
* saluda s a l u d a

## 1.5 Gramática

La gramática nos permite acotar las palabras que pueda reconocer la herramienta sphinx, lo que aumenta en gran medida el porcentaje de precisión y rapidez en el reconocimiento de voz. La gramática que utiliza el reconocedor de voz, y que definimos consta de algunas palabras y son las siguientes:

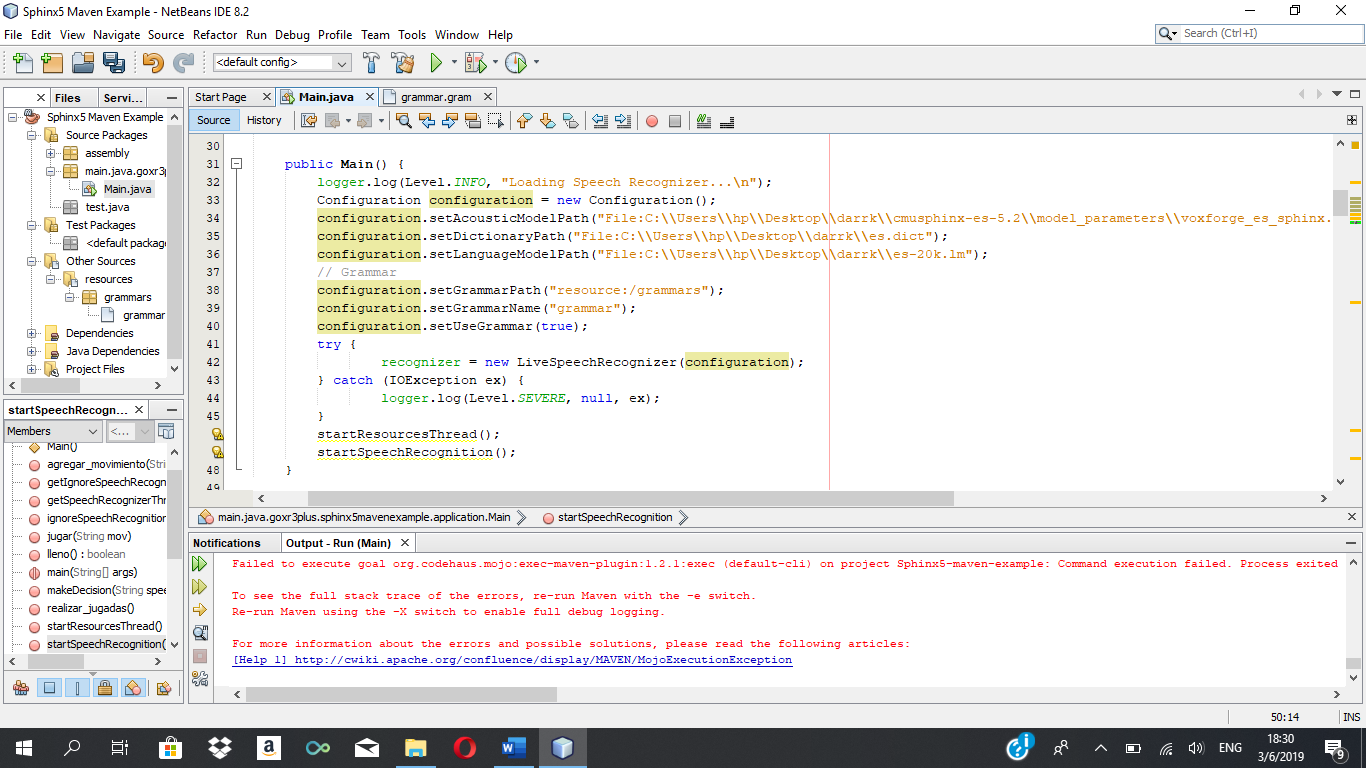
* Abajo
* Combinar
* Derecha
* Golpe
* Izquierda
* Jugador
* Patada
* Salto

## 1.6 Código del algoritmo

En el caso del reconocimiento acústico fonético, el algoritmo importa los diciones, modelos de lenguaje y

## 1.7 Código principal

En la clase principal el framework sphinx 4 se configuran los 3 archivos necesarios, que son modelo de lenguaje, modelo acústico y diccionario. Estos 3 prodrian incluirse en la librería o importarlos directamente como se ve muestra en la imagen.



## 1.8 Conclusiones

El reconocimiento de voz en cualquier tipo de proyecto puede parecer un tema complicado pero con las herramientas indicadas todo el proceso que se requiere puede ser mas sencillo de lo esperado, la herramienta empleada y evaluada es sphinx 4 que se adecua al reconocimiento acustico fonético, además de ser simple para usar y cuenta con diccionarios, modelos de lenguaje y sus librerias proporcionadas pueden ser muy útiles ya que estan disponibles en varios idiomas y estan al alcance, sin embargo no se recomienda hacer el uso sin entrenamiento, por los diferentes motivos como vocabularios extensos, variabilidad de los hablantes. Por todo esto y mas se pudo verificar que el reconocimiento sin entrenamiento o restricción con gramática y/o diccionarios se pueden obtener una gran cantidad de errores en el reconocimiento de las palabras.