

## INTERAÇÃO POR VOZ NO MUNDO DE JOGO

João Pedro Marques Pereira Ferreira

*Orientador*

Alberto Manuel Brandão Simões

Trabalho de Projeto apresentado(a) ao  
Instituto Politécnico do Cávado e do Ave  
para obtenção do Grau de Mestre em Mestrado Engenharia em  
Desenvolvimento de Jogos Digitais:

*Este trabalho não inclui as críticas e sugestões feitas pelo júri.*

Abril, 2020



## INTERAÇÃO POR VOZ NO MUNDO DE JOGO

João Pedro Marques Pereira Ferreira

*Orientador*

Alberto Manuel Brandão Simões

Trabalho de Projeto apresentado(a) ao  
Instituto Politécnico do Cávado e do Ave  
para obtenção do Grau de Mestre em Mestrado Engenharia em  
Desenvolvimento de Jogos Digitais:

*Este trabalho não inclui as críticas e sugestões feitas pelo júri.*

Abril, 2020



## RESUMO

Os sistemas informáticos cada vez estão, mais autónomos mais inteligentes, com uma capacidade de processamento muito maior e mais disponível, criando uma ligação com o humano cada vez maior. Surge a oportunidade de repensar novas formas de interação aplicadas no mundo de jogo.

Já todos nos cruzámos com sistemas inteligentes capazes de nos guiar, aconselhar e apontar-nos para a direção mais adequada, usando apenas a nossa voz como interface. Um gps, um motor de pesquisa, um lembrete no calendário. Um sistema inclusive que integra todas as soluções anteriormente descritas. Exemplo desses sistemas seriam a Siri (Apple), a Alexa (Amazon), Google Assistente (Google), entre outros.

A complementar estes sistemas, começam a aparecer em plataformas sociais, serviços que com uma base de dados nos conseguem esclarecer perguntas frequentes sem que seja necessária a intervenção de um ser humano no processo, os chamados ChatBots.

Numa análise mais virada ao mundo dos videojogos. As interfaces que o jogador usa no mundo de jogo, desempenham um grande papel na forma como o jogador experiencia o jogo. Se pensarmos nos vários tipos de controlos existentes: Comando, teclado/rato, guitar hero, Wii remote, etc a dedução torna-se óbvia. As interfaces têm um papel crucial naquilo que o jogo transmite ao jogador.

Após ter feito estas três análises, é fácil que perceber que ao adicionar um assistente pessoal/chatbot como uma interface de jogo, expande as possibilidades do mesmo bem como altera a forma como o jogador interage com o mundo de jogo. Quais as implicações da sua inclusão? É um tema ainda por explorar, muito devido ao facto de não haver actualmente a combinação dos vários sistemas.

Por fim a pergunta a que me proponho responder com a elaboração desta tese é: O que será necessário para incluir um assistente pessoal num mundo de jogo?

**Palavras-chave:** Assistente pessoal, ChatBot, Interface, Experiência de jogo.

# ÍNDICE

Este índice é automático. Para o atualizar selecione-o e pressione a tecla F9. De resto, o índice é automaticamente atualizado quando necessário (por exemplo, para imprimir). Não se esqueça de remover este parágrafo.

APOIOS .....	ERROR! BOOKMARK NOT DEFINED.
RESUMO .....	IV
ABSTRACT .....	ERROR! BOOKMARK NOT DEFINED.
AGRADECIMENTOS .....	ERROR! BOOKMARK NOT DEFINED.
DEDICATÓRIA.....	ERROR! BOOKMARK NOT DEFINED.
LISTA DE ABREVIATURAS E/OU SIGLAS .....	ERROR! BOOKMARK NOT DEFINED.
ÍNDICE .....	V
ÍDICE DE FIGURAS.....	VII
ÍNDICE DE TABELAS.....	IX
ÍNDICE DE EQUAÇÕES.....	IX
INTRODUÇÃO .....	1
1. UTILIZAÇÃO DO MODELO .....	3
1.1 UTILIZAÇÃO EFICIENTE DO WORD .....	3
1.1.1 MARCAS DE FORMATAÇÃO .....	ERROR! BOOKMARK NOT DEFINED.
1.1.2 ESTILOS .....	ERROR! BOOKMARK NOT DEFINED.
1.1.3 UTILIZAÇÃO DE AUTOMATISMOS.....	ERROR! BOOKMARK NOT DEFINED.
1.1.4 VERIFICAÇÃO ORTOGRÁFICA .....	ERROR! BOOKMARK NOT DEFINED.
1.2 ESTRUTURA DO MODELO .....	4
1.3 CABEÇALHOS E RODAPÉS.....	4
1.4 FIGURAS, TABELAS E EQUAÇÕES.....	5
1.4.1 FIGURAS .....	ERROR! BOOKMARK NOT DEFINED.
1.4.2 TABELAS.....	ERROR! BOOKMARK NOT DEFINED.
1.4.3 EQUAÇÕES.....	ERROR! BOOKMARK NOT DEFINED.
1.5 NOTAS DE RODAPÉ .....	ERROR! BOOKMARK NOT DEFINED.
1.6 OUTRAS REFERÊNCIAS CRUZADAS.....	ERROR! BOOKMARK NOT DEFINED.
1.7 REFERÊNCIAS BIBLIOGRÁFICAS .....	ERROR! BOOKMARK NOT DEFINED.
2. EXEMPLO DE SEGMENTAÇÃO .....	7
2.1 EXEMPLO DE SEGMENTAÇÃO DE NÍVEL 2 .....	7
2.1.1 EXEMPLO DE SEGMENTAÇÃO DE NÍVEL 3.....	ERROR! BOOKMARK NOT DEFINED.
2.1.1.1 Exemplo de Segmentação de Nível 4 .....	Error! Bookmark not defined.
2.1.1.1.1 Exemplo de Segmentação de Nível 5.....	Error! Bookmark not defined.
2.1.1.1.2 Outra segmentação de Nível 5.....	Error! Bookmark not defined.
2.2 OUTRAS SEGMENTAÇÃO DE NÍVEL 2.....	8
3. MAIS UM PONTO SECÇÃO EXEMPLO.....	11
CONCLUSÕES .....	ERROR! BOOKMARK NOT DEFINED.
REFERÊNCIAS BIBLIOGRÁFICAS .....	13
ANEXOS .....	ERROR! BOOKMARK NOT DEFINED.



## ÍDICE DE FIGURAS

Tal como o índice geral, este índice também é automático. Se utilizar esta secção, não se esqueça de remover este parágrafo. Se não a utilizar, selecione desde o início do título desta secção até ao início do título da secção seguinte e remova tudo (tecla *Del*).

*Figura 1:* Mostrar ou Ocultar pequenas marcas de formatação.**Error! Bookmark not defined.**

*Figura 2:* Remoção da numeração dos Títulos ..... **Error! Bookmark not defined.**

*Figura 3:* Como alterar o rodapé do trabalho ..... **Error! Bookmark not defined.**

*Figura 4:* Inserção de uma Figura..... **Error! Bookmark not defined.**

*Figura 5:* Referenciar uma figura ..... **Error! Bookmark not defined.**

*Figura 6:* Como Inserir uma tabela ..... **Error! Bookmark not defined.**

*Figura 7:* Os mesmos dois passos usados nas figuras ..... **Error! Bookmark not defined.**

*Figura 8:* O grupo de comandos "Citações e Bibliografia" ..... **Error! Bookmark not defined.**

*Figura 9:* Criação ou edição de Fonte Bibliográfica..... **Error! Bookmark not defined.**

*Figura 10:* Como inserir uma quebra de secção ímpar ..... **Error! Bookmark not defined.**





## ÍNDICE DE TABELAS

Tal como o índice geral, este índice também é automático. Se utilizar esta secção, não se esqueça de remover este parágrafo. Se não a utilizar, selecione desde o início do título desta secção até ao início do título da secção seguinte e remova tudo (tecla *Del*).

*Tabela 1:* Alguns Número Aleatórios ..... **Error! Bookmark not defined.**

## ÍNDICE DE EQUAÇÕES

Tal como o índice geral, este índice também é automático. Se utilizar esta secção, não se esqueça de remover este parágrafo. Se não a utilizar, selecione desde o início do título desta secção até ao início do título da secção seguinte e remova tudo (tecla *Del*).

*Equação 1* ..... **Error! Bookmark not defined.**

*Equação 2* ..... **Error! Bookmark not defined.**

## **INTRODUÇÃO**

Neste capítulo é definido e delimitado o espectro deste projeto. As necessidades, os problemas existentes ao implementar esta solução e qual diagrama técnico necessário desenvolvimento do projeto.



## **1. UTILIZAÇÃO DO MODELO**

Neste capítulo é definido e delimitado o espectro deste projeto. As necessidades, os problemas existentes ao implementar esta solução e qual diagrama técnico necessário desenvolvimento do projeto.

### **1.1. ESPECTRO DO PROJECTO**

O objectivo deste projecto é implementar uma solução que permita a um jogador, num determinado contexto, interagir com o mundo de jogo usando apenas a voz. Entenda-se que a ideia não é a de mudar as actuais interfaces já existentes (comandos, sensores, outros) mas antes complementá-las.

O idioma escolhido é o Inglês. Para além de haver mais recursos disponíveis, novamente sendo o fator tempo uma limitação, numa fase posterior será mais fácil difundir o projeto.

Será criado um pequeno mundo de jogo, um personagem controlado com o tradicional comando, e um NPC com o qual podemos interagir. É nessa interação que incide a investigação e desenvolvimento deste projeto.

Os objectivos são:

- Conseguir efectuar uma transação de um item e de informação, usando apenas a voz. (Exemplo: Comprar um item de jogo, pedir informação sobre qual a próxima missão)
- Dar inteligência ao NPC de forma a que o mesmo consiga, caso necessário, fazer perguntas despiste, de forma a melhor contextualizar a resposta.

- Criar uma base de dados, de forma a que a informação fornecida e adquirida seja dinâmica.
- Conseguir perceber o tom com que o jogador está a comunicar com o NPC e reagir de forma adequada.

Fora do espectro do projecto estão:

- Aspeto gráfico
- Performance
- Interfaces gráficas
- Sonoridade

## **1.2. NECESSIDADES**

De forma a conseguir produzir uma demo que contemple aquilo que é proposto no tempo disponibilizado para o efeito, é necessário usar várias ferramentas que de forma singular resolvem um problema específico.

É necessário portanto arranjar soluções para:

- Traduzir voz para texto em tempo real, para tratar a informação pedida pelo jogador.
- Traduzir texto para voz.
- Interpretar e contextualizar o que foi dito, para perceber qual a necessidade do jogador e encaminhá-lo numa conversa de forma a despistar questões pertinentes ramificar soluções para o mesmo problema.
- Consulta e criação de base de dados, de forma a arquivar dados que nos permitam formular uma resposta dinâmica e adequada a vários contextos.
- Conduzir o utilizador numa conversa de perguntas-respostas, de forma a melhor formular uma resposta adequada e contextualizada naquilo que foi o pedido o jogador.

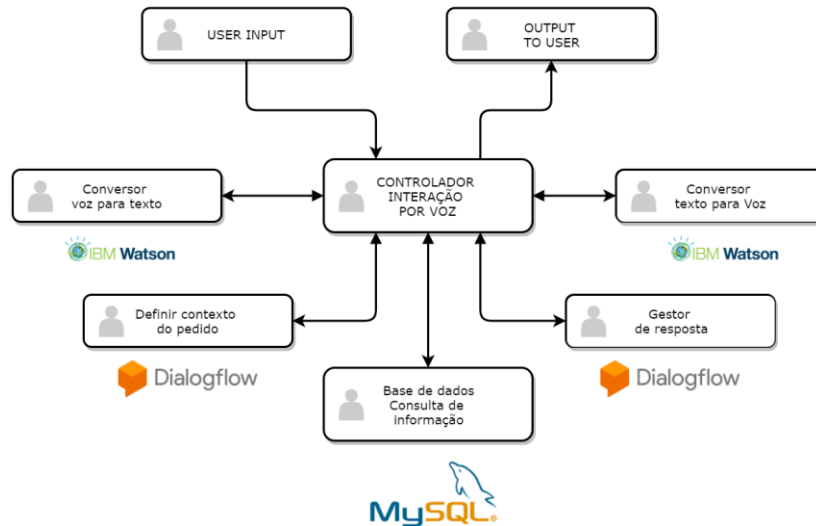
## **1.3. PROBLEMAS DE IMPLEMENTAÇÃO**

Estão detetados à partida, alguns problemas que vão aparecer na implementação da solução proposta fora de ambiente controlado.

Dado que estamos a tratar voz humana, tornam-se óbvios os problemas que podem e vão eventualmente aparecer na implementação. Idiomas, sotaques, ruído externo, regionalismos são alguns dos problemas que vêm ao de cima quando usamos sistemas de compreensão de linguagem natural.

#### 1.4. DIAGRAMA TÉCNICO

A figura seguinte apresenta a arquitetura da solução a ser implementada.



**Diagrama técnico de desenvolvimento**

Como se pode observar, a solução exige uma polivalência de tecnologias o que faz com que tanto a complexidade como as possibilidades aumentem tanto quanto a especificidade e complexidade do contexto em que queremos usar a solução.





## **2. ESTADO DE ARTE**

De forma a melhor estruturar este tópico, é agora apresentado o estado de desenvolvimento de cada um dos sistemas, as várias alternativas que podem ser usadas no projeto, quais as suas utilizações mais recorrentes e os custos associados a cada um deles.

### **2.1. SISTEMAS DE RECONHECIMENTO E PROCESSAMENTO DE VOZ PARA TEXTO E VICE-VERSA**

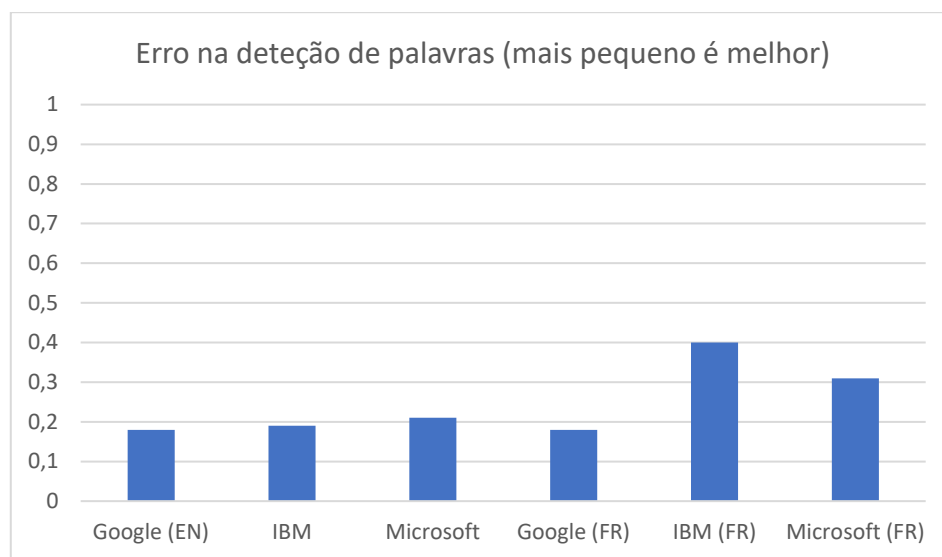
Estes serviços têm como função, traduzir a voz humana para texto e vice-versa. Diferem maioritariamente no número de idiomas que suportam e na possibilidade, ou não, de converterem voz para texto em tempo real.

IBM Watson

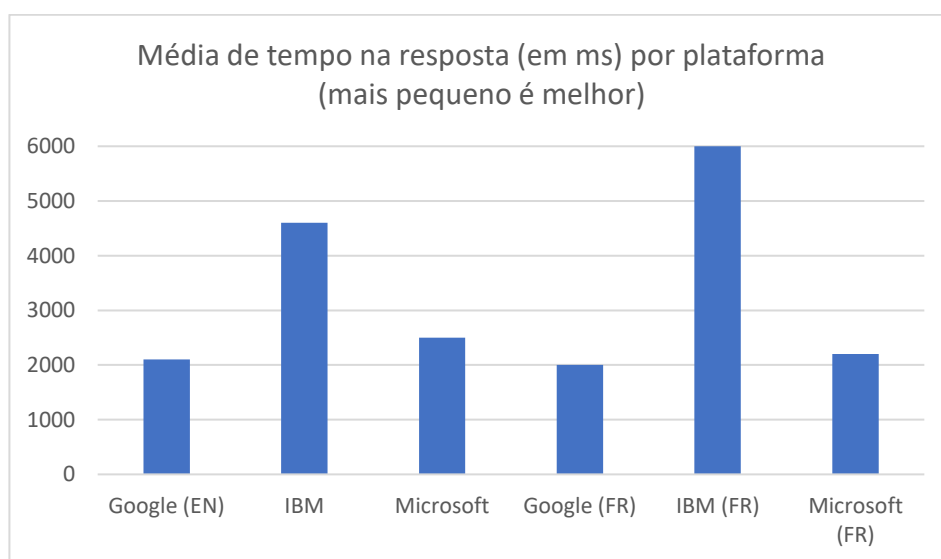
Microsoft Cognitive services

Google Platform

Comparativos:



Erros detectados em cada plataforma (maior barra = pior)



Velocidade na resposta de pedidos à API (maior = melhor)

Fonte: <https://recast.ai/blog/wp-content/uploads/2017/08/wer.png>

## 2.2. CHATBOTS

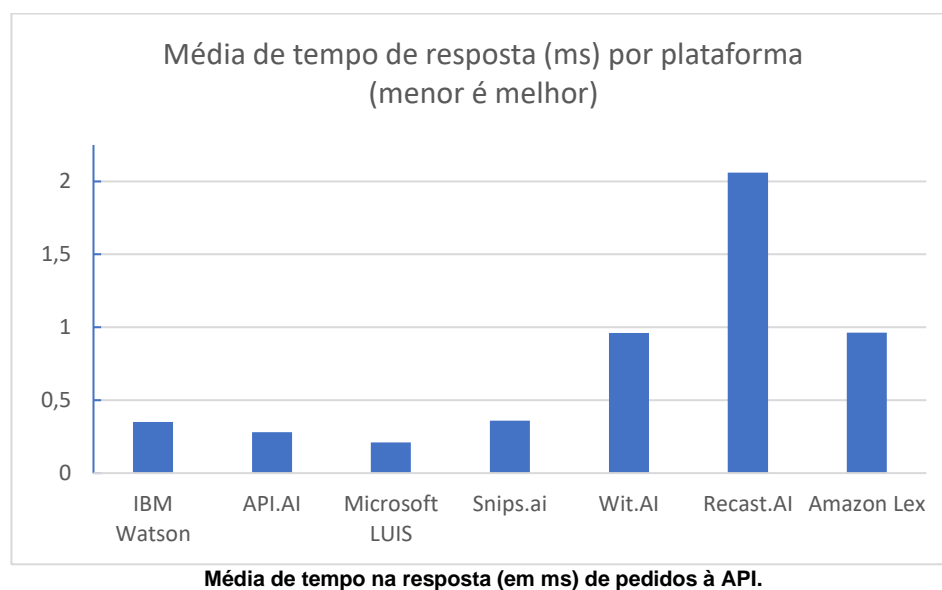
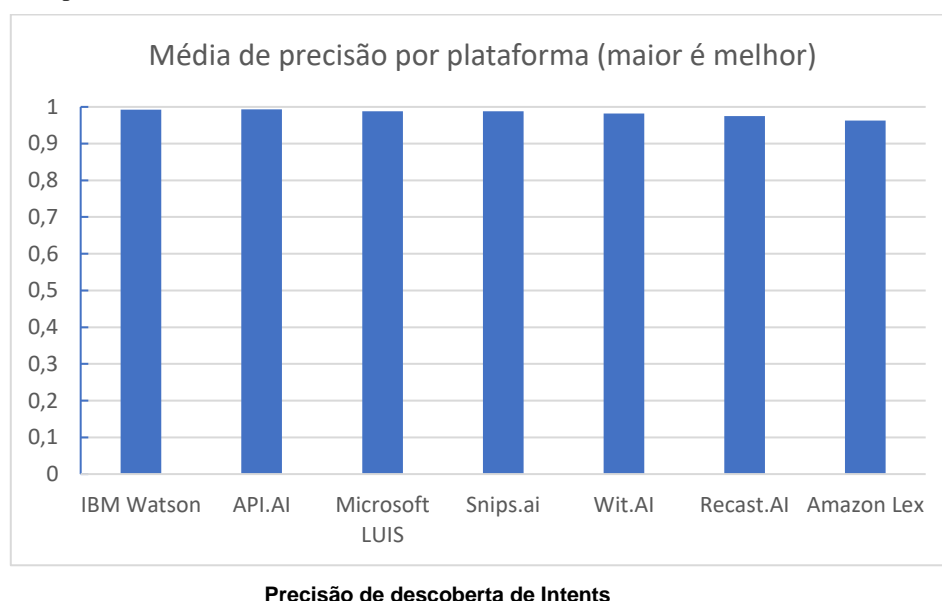
Sistemas que pretendem perceber contexto de uma conversa e encaminhar a mesma da forma mais adequada. Analisando uma frase, o objectivo destes sistemas é perceber qual a intenção (Intent) do utilizador, e formular uma resposta que possa dar seguimento a um diálogo que eventualmente irá resultar no esclarecimento de uma informação.

A maior diferença entre estes sistemas está, na deteção dos Intents, nos falsos positivos (quando o sistema assume de forma errada, que determinado contexto está a ser chamado) e na velocidade de resposta.

Sistemas mais populares:

- DialogFlow (2010, inicialmente chamado de api.ai, adquirido pela Google)
- Microsoft LUIS (2015)
- IBM Watson Conversation (2016)

Comparativos:



Fonte: <https://www.slideshare.net/KonstantinSavenkov/nlu-intent-detection-benchmark-by-intento-august-2017>

### **3. FERRAMENTAS**

O ambiente de desenvolvimento escolhido é o Unity3D com uso de linguagem de programação C#.

Das ferramentas mencionadas anteriormente, há algumas que já disponibilizam um plugin para Unity. Apesar de haver várias propostas de plugins para determinados problemas, as escolhas tomadas foram baseadas nos estudos acima mencionados e em: Tempo de implementação, facilidade de implementação e extensibilidade.

Concluiu-se portanto que as ferramentas que melhor servem o nosso propósito são:

Tradução voz para texto, e texto para voz: IBM Watson.

Chatbot: DialogFlow



## REFERÊNCIAS BIBLIOGRÁFICAS

American Psychological Association. (2010). *Publication manual of the American Psychological Association* (6th ed.). Washington: American Psychological Association.

Dorigo, M., & Blum, C. (2005). Ant colony optimization theory: A survey. *Theoretical Computer Science*, 243-278.

Epps, R. (2006). *Corporate Governance and Earnings Management: The Effects of Board Composition Size, Structure and Board Policies on Earnings Management*. Obtido em 20 de dezembro de 2006, de <http://ssrn.com/paper=929312>

Guimarães, R. C., & Cabral, J. A. (2011). *Estatística*. Lisboa: Verlag-Dashofer.

Makridakis, S., Wheelwright, S. C., & Hyndman, R. J. (1998). *Forecasting: Methods and Applications*. New York: John Wiley & Sons, Inc.

Iizuka, Tsujino (2012). Speech recognition technology and application for improving terminal functionality and service usability.

Santosh, K.Gaikwad & Bharti, W.Gawali & Yannawar, Pravin. (2010). A Review on Speech Recognition Technique. *International Journal of Computer Applications*. 10 10.5120/1462-1976.

Machin, Sam (2017). Bots and AI: Voice vs. Text for Your AI Bot Interface

Mannes, John (2017). TalkIQ raises \$14 million Series A to give enterprises AI insights into voice communication. *Techcrunch*

Fonte: <https://techcrunch.com/2017/09/19/talkiq-raises-14-million-series-a-to-give-enterprises-ai-insights-into-voice-communication/>

Tuttle, Tim (2015). The Future Of Voice-Activated AI Sounds Awesome

Fonte: <https://techcrunch.com/2015/03/06/the-future-of-voice-activated-ai-sounds-awesome/>

Embick, David (1998). Voice Systems and the Syntax/Morphology Interface

## ANEXO A