

# **IA, AI & Voice:**

## **How can we make voice interactions easier?**

a talk by

**Lade Tawak**

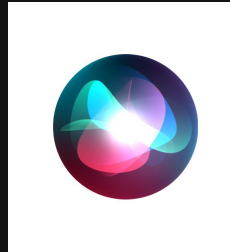
## Introducing the Vodafone-Idea Phone Line with the Google Assistant

Dial 000 800 9191 000 toll-free  
and connect with the Google Assistant  
to ask questions and get information



No data or an internet connection needed

Google for  
**INDIA**



# The cognitive considerations for voice interfaces

Miller's law = the span of immediate memory is limited to around 7 pieces of information

Avoid lengthy responses, chunk responses, confirm attention and understanding

**CONSIDERATION 1**

**Low digital literacy +  
no/low visual context  
= fear and anxiety**

## No direct manipulation



**I press something and  
something happens**



**I'm talking to someone?**

## What does this do?

*“The main bottleneck of using a voice assistant system is that the user cannot know all the commands.”* – **Park, Park, Song (2020)**

## Conversational cues

You need oil, salt...  
(still listing)

...

What do I need to  
make a marinade?

Okay  
(affirming)

...

Low digital literacy

# How do we come in?

Set context, explain to the user what the agent can do





Low digital literacy

# How do we come in?

Proper data structures  
which allow AI properly  
parse user requests to  
map to attributes

Low digital literacy

# How do we come in?

Create Dialog Flows and maps that include, clarification paths, error messages, and error recovery

How can I help you? You can say **report a stolen card, card activity, insurance**, or I need **help with something else**

That's **card activity**, is that correct?

What card activity can I help with? Say: **balance, recent transactions, dispute a charge** or I need **help with something else**.

Card activity

Yes

...

CONSIDERATION 2

# Communication styles and language

# Identifying user intent

Book me an **uber** to **the airport**

I'd like a **pizza**

Order a **medium triple decker bbq chicken pizza**

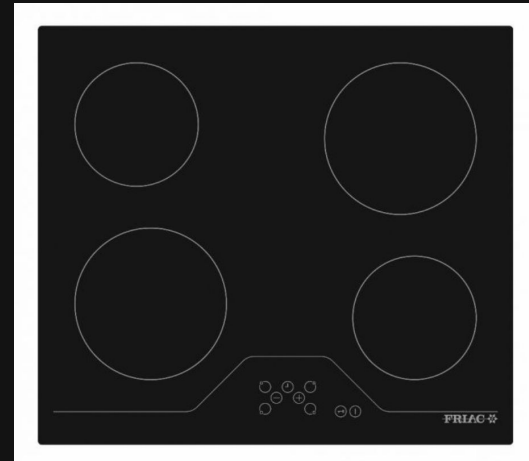
I need to go pick up **Tunta** and then **go to church**

## **Context and content**

- **Language is more than grammar and sentence structure**
- **Slangs, social norms, or cultural information add meaning**

**Cultural cognition**

**When I say stove, what do you picture?**



## Cultural cognition

When I say **cooker**, what do you picture?



- **pressure cooker?**
- **rice cooker?**
- **slow cooker?**
- **nothing?**



## Low context, high context

# Can the agent match my interaction style

Two chatbots to teach about consent in the context of sexual activities. One LC, one HC. (Fromet, 2021)

	HC Culture Participants	LC Culture Participants
Low context interaction	Neutral	Slight preference
High context interaction	Strong preference	Neutral

## Accent recognition

I didn't get that

Try again

Call Onyinye

Call Onyinye

...

## Accent recognition

What would you like to learn

How to make stock

I'm sorry I can only teach how to make a marinade or curry or stock

I want to make stock

Okay, we can resume whenever you're ready

*“Having to adapt our way of speaking to interact with speech recognition technologies is a familiar experience for people whose first language is not English or who do not have conventionally American-sounding names. I have even stopped using Siri because of it” – Claudia Lopez Lloreda*

## **Multilingual behaviour**

# **Can the agent understand me if I speak multiple languages at the same time**

My neighbours where fighting yesterday and my boyfriend went to intervene and asked them to stop.

During that. He fall and broke their TV. He was badly injured and I took him to hospital. Today the couples came and asked me ukuthi ubuya nin esbhedlela azobhadala iTV yabo.

Communication styles

# How do we come in?

Ensure data sets  
represent a wide range  
of accents, dialects, and  
cultural references.

Communication styles

# How do we come in?

Ensure taxonomies, metadata, and data labelling are free from biases, but also consider cultural contexts

Communication styles

# How do we come in?

Customise agent  
personality based on  
location and other signals

My Google assistant voice changed from the Nigerian voice and I'm so upset.

19:48 · 26/05/2020 From Earth

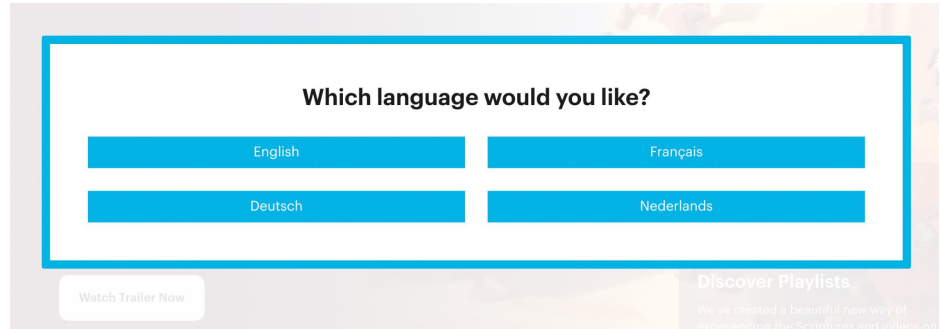




Communication styles

# How do we come in?

Allow personalisation and customization



Communication styles

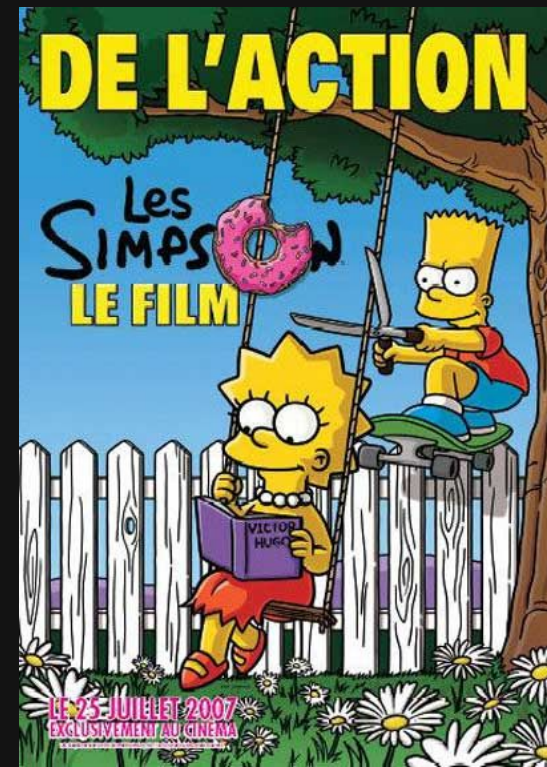
# How do we come in?

In creating DialogFlows, include multiple sample utterances that cover diverse cultural approaches



Communication styles

**This is not  
new**



Communication styles

**This is not  
new**

Different food menu  
items based on location

CONSIDERATION 3

# Environmental context

## Unavailable in my country

You need oil, soy sauce,  
garlic, pepper, and  
lemon juice

Should I place an  
order for you?

What do I need to  
make a marinade?

I don't have  
lemon juice

....

Communication styles

# How do we come in?

Create flexible interaction structures such as branching choices and decision points that can respond based on predetermined attributes

Communication styles

# How do we come in?

Customize what agent can do and not do based on what is available to the user in their context (e.g. based on location)



Communication styles

# How do we come in?

Map user journeys and  
understand the  
surrounding context  
outside of your product

**Wrapping up**

HELP THE AGENT

**...understand user (intent)**

**...retrieve relevant information**

**...generate the right output**

BY...

**...customizing agent based on sociocultural context**

**...using inclusive taxonomies and data labelling**

**...designing multiple branching choices in the dialogflow**

TO...

**make voice  
interactions more  
inclusive, usable,  
and accessible**

THANK YOU

**Lade Tawak**

