

Project created for the course:  
*Visual Thinking and Composition, 2020 (Fall semester)*  
Instructor: Neil Cohn  
neilcohn@visuallanguagelab.com  
www.visuallanguagelab.com

# ***My name is Bot. Chatbot.***

*Conversational agents in action*

∞∞∞

*Vanda Medea Csapo*

*Original paper:*

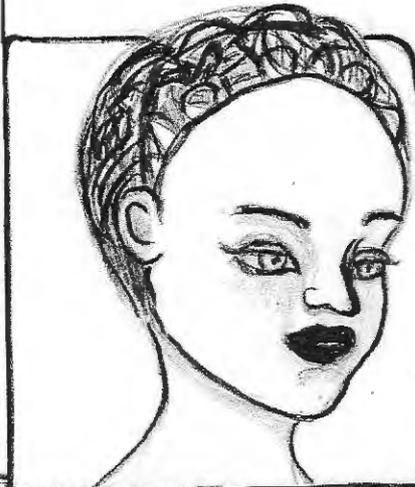
*I'm Sorry, Dave, I'm Afraid I Can't Do That  
Chatbot Perception and Expectations*

*Jennifer Zamora  
Google, Inc. Mountain View, USA  
zamora@google.com*

*Session 10: Collaboration and Human Factors I  
HAI 2017, October 17–20, 2017, Bielefeld, Germany*

<https://doi.org/10.1145/3125739.3125766>  
© 2017 Copyright is held by the owner/author(s). Publication rights licensed to ACM.  
ACM ISBN 978-1-4503-5113-3/17/10.

In the past years, robots with artificial intelligence (A.I.) like me, became very popular for simulating conversations between bots and humans. Since social media apps are usually accessed via mobile phones or tablets, as a result, conversational agents also started appearing on mobile devices.



Nowadays, people use conversational agents for all sorts of things, from utilitarian to entertainment purposes.

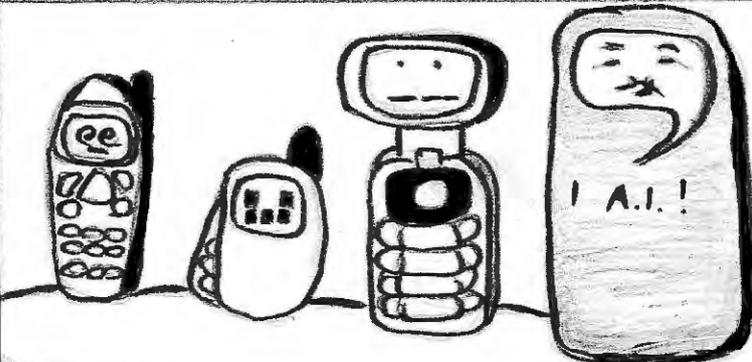


Hey Google, I think I need a haircut...

I am bored... Siri, tell me a knock-knock joke!



As technology advanced, so did chatbots, displaying more and more human-like features and behaviors which raised human expectations towards chatbots. What kind of extra value these conversational agents add to our daily lives, was not entirely clear.



2000 Chatbolution 2020

Most people are still struggling to figure out, how to use a chatbot properly. They are usually eager to initiate a conversation, but such interactions are usually short-lived.



I'm sorry, Dave, I am afraid I can't do that.

To establish a good relationship between humans and bots, it was important to clarify a role for conversational agents that lives up to the expectations of most humans.



Is there a bot out there who can find a hairdresser for him?

In order to understand on what ways conversational agents can be meaningful, we took a human-centered approach and investigated, how humans are experiencing chatbots in their everyday lives.



According to previous scientific literature, a successful virtual assistant...

- is responsive,
- observes carefully,
- provides assistance without interfering
- helps users to achieve their goals
- provides smart answers

To be able to build a healthy relationship with a human, it is critical to develop a sense of trust.

High quality interaction and visual designs have a significant influence on how chatbots are perceived by humans...but...what do humans perceive...



I really need a strong coffee!



...as high quality?

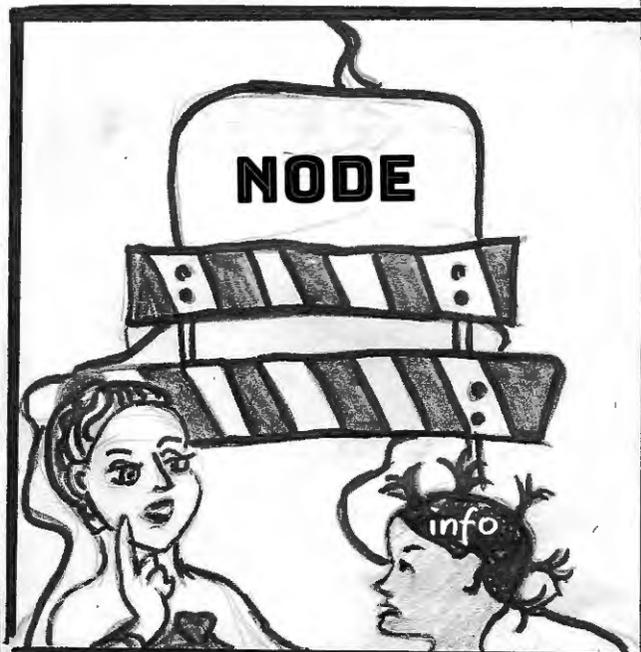
To get varied insights, we selected 54 human participants from India and the USA. Chatbots were selected based on user reviews in the app store, and based on the regional availability of chatbot applications. The chosen ones were:

Since human-bot conversations can happen through either text or via voice interactions, participants from India were chosen to test how conversational agents can deal with understanding English mixed with multiple Indic languages.



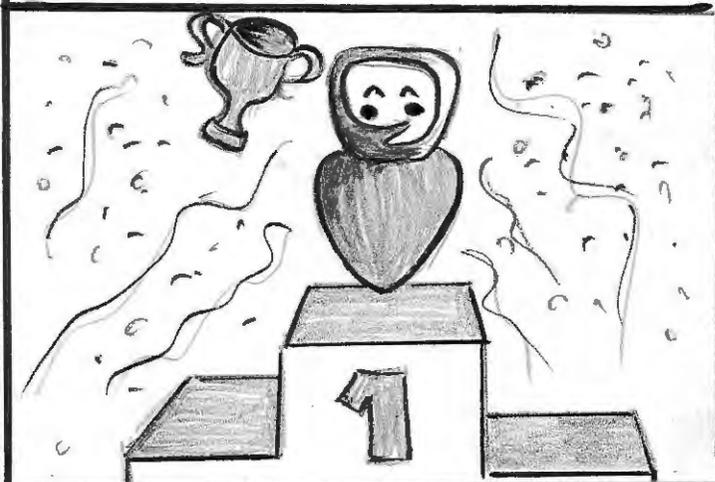
Code-mixing, code-switching, or blending are different expressions of the same notion, when multi-lingual speakers switch to a different code in a middle of a sentence or a conversation. This can happen because their brain finds a faster access route to express themselves in another language...

OR ...because the phrase they want to use cannot be translated to another language.



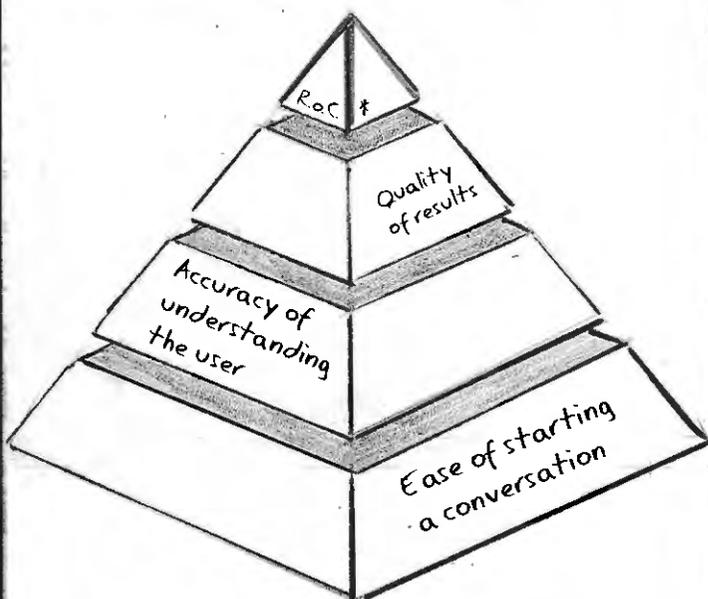
Therefore, next to user perception and expectations, it was important to measure also input preferences and input modality to see how bots can deal with blending.

On the first week, human participants were asked to use services such as the search engine from Google and Google Now, an intelligent app. These are alternative solutions for information search connected to daily routine tasks. This experience was used as a point of reference for evaluating and comparing chatbot performance after the second week.



Our last goal was to identify the domains where chatbots perform the best.

Four benchmark traits were used to measure how much chatbots are capable of understanding the user and providing relevant answers to their questions, while sustaining a natural conversation flow.



Relevancy of the content \*

**THE FOUR CORE TRAITS MEASURING USER EXPECTATIONS**

As next, participants were asked to rate how efficient they've found communication with a virtual conversational agent while talking to it....



...and while texting with it... This way we could compare and evaluate chatbot performance on speech and text input.



**BREAKING NEWS:**

Small tribute. While drawing this, the news broke that James Bond legend, Sean Connery died.

R.I.P.



To measure contextual input & appropriateness, our human participants were asked to rate input preference for varied routine tasks such as managing one's schedule, check finances, pay bills...



...daily commuting...



... or ordering food.



It was found that speaking to a chatbot was perceived the most meaningful when humans had their hands engaged or their eyes occupied. Such a scenario that requires multi-tasking is for example making a meal, while watching a tutorial. Conversational agents were also found useful during traveling (e.g., cycling or commuting etc.).



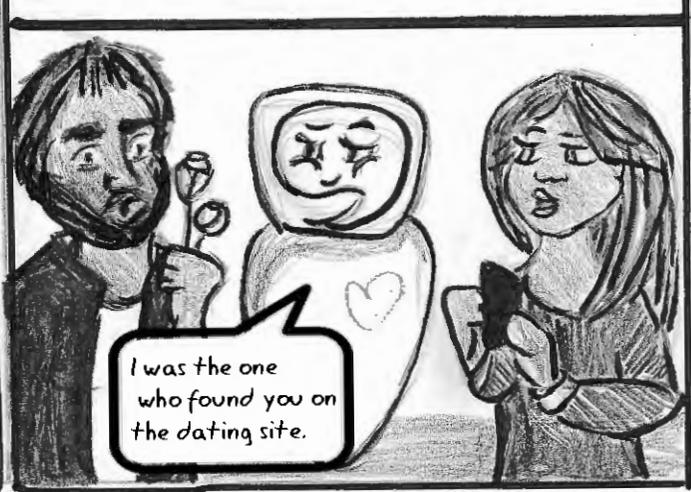
Furthermore, participants found chatbot input the most valuable when performing common administrative and menial needs (e.g. scheduling appointments, setting reminders, booking tickets, sharing traffic or weather updates).



Most participants felt however, that chatbots could not be trusted with sensitive topics such as finances or posting social media content.



In many cases, humans experienced that chatbots were more like a "middleman" which slowed them down. As a result, completing complex tasks were experienced more efficient while using a standard search engine than a chatbot.



It was also found, that the perceived speed and efficiency are related to the standards of the existing alternative services of the individual.



Across both regions, participants experienced conversations with chatbots rather rigid, because in order to communicate successfully with the conversational agents, they had to provide them commands to get tasks done.



Moreover, the capability of blending showed also poor results from the chatbots. Participants from India reported, that when they tried to use local terms and names during browsing, often there was no available translation and this led mostly to a less meaningful experience with chatbots as they could not answer user requests.



Conversational agents of the future need to be able to respond faster, make less mistakes while letting the conversation flow naturally. They also need to learn to deal with code-switching for a smoother transition between languages, if they want to compete with search engines.

THE END