When Voice is the Interface:  
Conversation Design Insights Based on a Use Case

Giuseppe Aceto, Federica Tazzi, Elisabetta Schiatti, Olivia Kubanski, Vittoria Chiappa, Chiara Martino, Eleonora Bruttini  
Assist Digital, Italy  
{giuseppe.aceto;federica.tazzi;elisabetta.schiatti;olivia.kubanski;vittoria.chiappa;chiara.martino;eleonora.bruttini}@assistdigital.com

Abstract. The following paper is placed at the conclusion of a project to improve the customer experience by Eni Gas and Luce (EGL). The activities described below lasted 12 months and led to the development of Lucilla, an Italian conversational IVR based on Natural Language Understanding that has both informative and dispositive skills. The activities saw the joint activity of Assist Digital, Google, the EGL team and other consultants for the technical implementation. In order to design a usable, effective, pleasant and inclusive voicebot, a user-centered approach was applied. Thus, a preliminary user research investigated EGL customer base, their familiarity with the energy topics and their attitudes towards virtual assistants. The collected insights led to the creation of Lucilla’s personality. In the second stage of the project, newly designed conversational flows were tested through the Wizard of Oz methodology, together with some flows that had been already implemented. The aim was investigating the usability of the flows and users’ perception of Lucilla’s voice and personality. The collected insights led to the definition of best practices for voicebot conversation design, such as the need to reciprocate kindness and avoid ambiguous words and technicalities, help users build a correct mental model of the system and suggest proactive prompts only if users have already accomplished their primary task.

Keywords: Conversational Agent, Voicebot, User Test, Wizard of Oz.

1 Conversational Approach

1.1 The Context of the Project and the Design Process

The project discussed in this work aimed at developing “Lucilla”, a Conversational IVR, i.e. a Virtual Assistant (VA) that people can talk to via telephone. Lucilla is based on Speech-to-Text (STT) and Natural Language Understanding (NLU) technologies, thus uses AI to interpret what users say using natural language. The first of Lucilla’s goals is to identify users in order to determine whether they are customers. Secondly, it tries to understand why they are calling Customer Support. Finally, it guides users towards the automated service or the person that can solve
their problems or answer their questions. We used a user-centred approach throughout the different stages of Lucilla’s design process. The first stage was a user research aimed at mapping the context and understanding customers’ needs and pain in order to define specific design guidelines. The second stage aimed at designing, prototyping and testing specific use cases, in order to assess their usability and effectiveness before implementing them.

1.2 The First Stage: Research and Conceptualization

To set a context for the Lucilla project, we identified two prevalent attitudes towards virtual assistants among the diverse customer base of EGL (7M customers with different backgrounds, use of language and digital skills). The “basic” attitude was the predominant one and consists of people of an advanced age and low culture who find virtual assistants unsatisfying because of the lack of “human touch” and the difficulty to reach their goals. On the contrary, the “evolved” attitude consists of informed, technological people who have easy interactions with VA and align their expectations to their previous experiences with smart speakers or personal assistants.

Our research thus highlighted the key areas to focus on in order to design a virtual agent with effective and pleasant conversational skills. First, the use cases selected should be relevant to users’ needs. Second, the VA’s personality and word choices should be designed to represent the brand’s tone of voice and values and to resonate with the customer base. Third, conversation design must provide for both a “happy path”, which enables users to reach their goal effortlessly, and “fallbacks”, which help users go back on track when miscommunication happens. Fourth, all potential users expect empathy and a mix of human components during the conversation.

1.3 The VA Personality

As a result of the research phase, we defined the VA personality and some conversational pillars. The research phase consisted in a series of workshops with all the main EGL stakeholders, qualitative interviews with customers and final approval tests.

On the basis of workshops results, we were able to define the name, the gender and the voice type for our VA: all the results converged on a specific name, a feminine presence, a warm and fluid voice, very distant from an impersonal, robotic voice.

To realize these characteristics, the right balance between speed and intonation of the voice was identified. Specifically, the tests have shown that a speed of 1.15 and Pitch 0.14 of the female voice offered by Google Text-to-Speech is the right balance between comprehensiveness and expressiveness.

---

1 On the web and with the bots the same real life's rules of behavior and communication are valid, as demonstrated in “The Man who Lied to His Laptop” by Clifford Nass (Current 2012), professor at Stanford. This book is about years of experiments on subjects that deal with technology and explains how people transfer the behavioral rules from the analog to the digital communication, through the use of interfaces.
The name chosen was “Lucilla” (as “luce” in Italian means “light”) a name that tells the story of the brand, putting in evidence the defined personality for this virtual assistant.

Finally, on the basis of the analysis results of the qualitative interviews taken, was consequently defined that the Eni Gas e Luce VA has a friendly but also authoritative personality; it communicates in a simple and direct way; it chooses the second person - “you” - and calls the customer by the name, to reinforce the relationship with the customer without affecting the professional VA image.

1.4 The Second Stage: Prototyping and User Test

During the second stage of the design process, we designed and prototyped some Conversational Flows related to specific use cases. Then we conducted a user test session to test two artifacts: some flows which had already been implemented and our new prototyped flows. In particular:

• welcoming, user identification and call steering were performed with the live system
• vertical conversational flows were performed using the Wizard of Oz methodology. Such flows allowed users to change their phone number on Eni database (voice-only flow), delete their phone number (voice-only flow), and get bills or other documents by communicating their email (multichannel flow: voice and Whatsapp).

The purpose of the test was twofold. On one hand, we aimed at investigating whether the flows which had already been implemented were usable and how users reacted to the VA's voice and personality. On the other hand, our goal was to investigate whether the newly designed flows, prototyped with a tool that allowed us to replicate the same voice, pitch and speed of the live VA, corresponded to users’ needs and were thus eligible for a subsequent development.

General Objectives of the User Test. We recruited 16 users, in order to validate the new prototyped flows, collect participants’ actual utterances of intents, and investigate users’ impressions about Lucilla’s voice and perceived personality. Additionally, we aimed at comparing the voice-only and multichannel experiences, investigating the effectiveness and desirability of an additional proactive request in two different scenarios (one with a functional aim, the other with a commercial aim), and researching uncovered needs, expectations and desires of the target users.

Conduction of the Test Session.

A. Initial interview: during the test, an expert moderator briefly interviews the participant, gathering information on the objectives and the context of use of the users interested in the energy and consumption theme.

2 In the field of human–computer interaction, the Wizard of Oz is a testing method in which users interact with a prototype that is operated by a researcher hidden behind the glass of the test room. Such research methodology aims at testing a service or a product before its implementation in order to understand users’ reactions.
B. Scenarios and task execution (Wizard of Oz): the moderator asks the participant to identify themselves in some realistic scenarios and perform some tasks using the prototyped IVR. The impressions of the participants are observed throughout the test. Before each task, the participants share their real previous experience with the proposed task, their strategy to solve it and their expectations towards the service. During the tasks, we investigate the perceived success and if the task was completed. At the end of each task, the researcher collects comments on the ease of use and judgments regarding some characteristics of the scripts (i.e. speed, clarity...).

C. Concluding interview: at the end of the test, the moderator asks the participant for an overall opinion on the user experience, together with any suggestions for future developments and improvements.

1.5 Collected Insights

The Wizard of Oz test enabled us to collect insights about the perceived affordance, usefulness and pleasantness of Lucilla’s conversational flows. Moreover, we gathered the main drivers that led users behavior when interacting with the VA for customer services purposes (see Table 3). We used such insights both as a resource to redesign the conversational flows, thus improving their quality, and to consolidate a set of conversational design best practices (see chapter 3).

Table 3. Details about main drivers and related descriptions.

<table>
<thead>
<tr>
<th>Topic</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>Speed in resolution</td>
<td>Participants recognize the value of speed in Lucilla. Compared to a traditional IVR and the long waits to be able to speak to an operator in some cases, Lucilla, mechanical, artificial and intelligent, speeds up.</td>
</tr>
<tr>
<td>Friendliness and Kindness</td>
<td>The way in which Lucilla addresses the user is extremely pleasant. Users appreciate its ToV (Tone of Voice) and word choices.</td>
</tr>
<tr>
<td>Effectiveness / Human Handover</td>
<td>Users appreciate when Lucilla reformulates questions in cases of misunderstanding. This fallback is available for a finite number of times, then a smooth transition to human agents completes the experience to help users solve their problem even in the most complex scenarios.</td>
</tr>
<tr>
<td>Conversational paradigm learning curve</td>
<td>Being able to speak freely with the VA left users surprised at first, as they were unsure how to formulate their requests. After a few interactions, however, they learned how to have a conversation with the system.</td>
</tr>
</tbody>
</table>
2 Evidence and Key Takeaways

Friendliness is Expected. Virtual assistants give the illusion of agency, so users project a personality onto them and expect the virtual assistant to reciprocate their politeness. For example, during the Wizard of Oz test users noticed if Lucilla did not reply to greetings, pointing out that it was unkind. Conversation design should thus encompass personality design in order to carefully choose a voice, set of reactions and vocabulary that show empathy and mimic human manners, personify the values of the company and fit the peculiarities of the target users.

Clarity, no ambiguity. During a conversation with an automatic system, people cannot ask for detailed explanations, as it happens when they speak to a human. Therefore, a VA must be precise in describing the operations it carries out. During the test, users expected Lucilla not to be ambiguous when a dispositive operation was carried out. As the script used the vague word “replace”, that means “delete”, “add” or “change” in Italian, users were left unsure about the action they performed and its outcome, thus being unsatisfied.

Multichannel with reserve. Users want speed of resolution, so they prefer to stay on one channel, since every change can slow down their journey. They can agree to change if the reason is clear, if there’s no way to solve it with the VA and if the second channel is familiar. When Lucilla asked users to change channel to collect an email address, most of them were surprised and couldn’t guess why (NLU performance). Overall, the speed and effectiveness of the response must be guaranteed. But users are frustrated when there are more than 2 channels, even when the third one is the hand-over to a human.

There is a learning curve. Users find it difficult to understand how to start a conversation with virtual assistants. For example, we noticed that, when asked to perform complex tasks, users tended to formulate a very long and detailed initial request. Even those users who were familiar with virtual assistants wondered whether to express their request with a structured sentence or speak by keywords in order to be understood. However, users tended to learn how to interact with Lucilla quickly.

Security and privacy, not always relevant. When it isn’t possible to recognize users from their phone number without asking for other personal information, Lucilla asks for either their name or client code. We tested the two requests and the results show that users don’t have any preferences between them because they were both fast. However, all users agree that in case of a dispositive operation, they would like to have a higher level of security and action confirmation.

The role of proactivity. Users do not appreciate proactive questions when unrelated to the task which they are trying to accomplish. They rather finish their task and then listen to the proposal. Moreover, when the proactive question is related to a commercial proposal, users are very annoyed because it breaks the flow, and they must spend energy to skip it without errors (fear to buy or subscribe by mistake).

Hand-over to human. When designing a Virtual Assistant, it is important to consider its behavior in case of errors: when Lucilla does not understand or is unable to respond, a human handover is possible. During the test we observed how people reacted to this transition, and we found that they prefer it when it happens on the same channel of the request. For example, if the conversation is on WhatsApp, the human support should be there in order not to lose the conversation already had with the VA and not to waste time.
3 Design the Conversation

After the research and the test phase, some design principles can be defined for a better vocal experience with VA in the customer care field.

Syntax and mental model. Choosing a simple and explicit syntax allows you to communicate more clearly and immediately, as discussed in the linguistic-centric prompt design study by Meck and Precht, who also state that “Prompts should be written results- and information-oriented and leave out unnecessary elements.” However, we found that you have to explain to users why you ask for certain things, so that users can build a correct mental model. Thus, you must usually be concise, but there are points of the flows where you must spend words to communicate more clearly.

Terminology. Avoiding the use of technicalities helps to make communication more accessible for even less experienced users. When it is necessary to resort to sector terminology, it is suggested to contextualize it.

Follow the learning curve. Use example utterances and give perimeter at the beginning of the flow and for complex tasks.

Only task related proactivity. In order to be useful, proactive messages should be pertinent to users’ context of use and to their needs, thus anticipating them. Put them at the end of the flow, after users have accomplished their task. Avoid commercial proactivity, both at the start and at the end of the flow. If you must include it, design a message with a very clear and safe skip option.

Hand-over and multichannel. When VA understands and executes tasks, users are very happy with speed and effectiveness. Always explain to users why they have to wait for a human operator or to change channel, since they are not aware of NLU limits or company process (i.e. I’m not good at understanding email by voice, so would you like to continue with me on WhatsApp?).

4 Conclusions

When designing a virtual assistant, it is paramount to gather as much information as possible about the target users. Conversation designers should use these insights both to give the VA a personality that resonates with users’ world of reference and to design conversational flows that enable users to reach their goal as effortlessly as possible. To do so, the designer should infuse the VA’s script with empathy, thus including small talk, greetings and exclamations. Such scripts should also use familiar words and short, simple sentences in order to be easily accessible and minimize cognitive load. Additionally, prompts should set users’ expectations about the VA’s skills and help them understand how to interact with it. However, they should not exceed in proactivity: they should suggest only those next steps that are relevant to the task users are trying to achieve.
References