Enabling Naturalness and Humanness in Mobile Voice Assistants

Sanjay Ghosh, Jatin Pherwani

Enabling Naturalness and Humanness in Mobile Voice Assistants

Sanjay Ghosh and Jatin Pherwani

Samsung R&D Institute, Bangalore, India
(sanjay.ghosh , j.pherwani)@samsung.com

Abstract. Voice Assistant applications are preferred to be designed with some form of personalization to enhance user experience. However, the question remains that what is the kind of Voice Assistant that users would prefer. As part of this research we tried formalizing the notion of naturalness and humanlike in the context of voice assistants by defining its optimal personality, creating guidelines for natural dialogues and expressions. We decomposed the design problem into four aspects, communication style, personality, speech or dialogue, and appearance along with non-verbal gestures. To investigate on each of these four aspects, we performed four different user elicitation techniques.

Keywords: Voice interface, natural language, voice interactions

1   Introduction

Our research goal was not to introduce yet another new candidate into the world of Siri, Cortana, Google Now and various other Voice Agents. Instead our attempt was to design a very natural and humanlike Voice Assistant from scratch. With modulations in language styles, speech parameters, and visuals it is possible to impart quasi-human characteristics in new versions of voice assistant. Earlier research reported in this area includes explorations of the personality for a robotic assistant for television [1], robot for giving advice [2], including emotions in a conversational agent [3]. Beyond all the existing work in this area, our contributions lies in incorporating the user's perspective as an important element in formulating a personality for the voice avatar, creating dialogues and expressions for the voice agents. We decomposed the design problem into four, (i) Communication Style, (ii) Personality, (iii) Speech/Dialogue, (iv) Appearance and Non-verbal gestures. Contextual interviews, survey questionnaires and participatory design were conducted to get an insight into the problem from user's view point. We studied human personal assistants to understand their communication style and behavior with their boss. Then we used a personality elicitation survey to formulate the attributes of personality desired by the users. This followed by the study of linguistic, speech and characteristics of user created dialogue library for a voice assistant across different scenarios. Users' direct perception regarding these were taken into account through a co-creation activity.
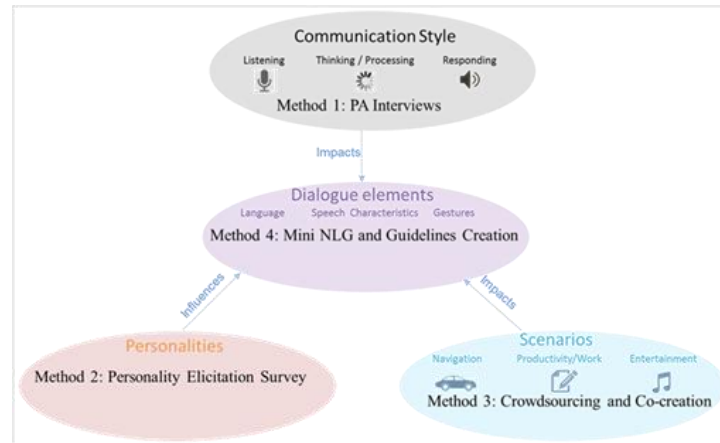
Fig. 1. Illustration of the scope of research

2   Defining the Communication Style

In the first phase, our interview sessions with seven professional assistants were focused mainly to identify various personality traits of the assistants, understand how they handle various situations and observe explicit cues that are exhibited during conversation. We formulated the social behavior of the assistants through our observations on emotional, behavioral and functional aspects.
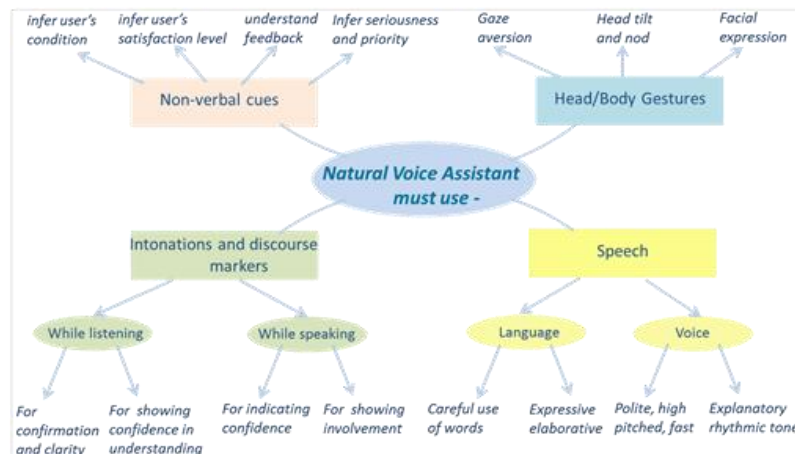

Fig. 2. Illustration of outcome from Assistant Interviews

We also observed the typical verbal expressions such as, intonations and exclamations and non-verbal expressions such as, gaze aversion, facial expression, head tilt and nod, on the assistants. In terms of the speech constructs, assistants used intonations and discourse markers to provide feedback on its level of understanding while listen-

ing to others but these are not to be repeated very frequently. While speaking, assistant provide feedback on its confidence level of the response through use of intonations and discourse markers.

3 Decoding the Personality

Our next objective was to formulate an appropriate and user desirable personality for the voice assistant. The definition of voice assistant personality in our exploration consisted of 9 attributes considering various aspects of personality traits. We derived these from various prior work [4], [5] as well as included few additional attributes relevant to the case of Voice Assistant like, Spontaneous, Inventiveness and Similarity based on our insights from the earlier personal assistant interviews. Next we performed the Personality Elicitation session through a brief questionnaire, a face to face survey discussion session with 30 participants to understand the importance and desirability of each of those attributes for a voice assistant. Our intent was to quantify, how much of each of these traits was desirable to a user. The evolved personality based on the 9 attributes is illustrated in fig.3. Interesting facts emerged out of the personality analysis; such as, users preferred their Voice Assistant to *Control* them to some extent, but they would prefer not be fully *Dependent* on it. Therefore, perhaps the users wish to see it as a Companion than as an Assistant.
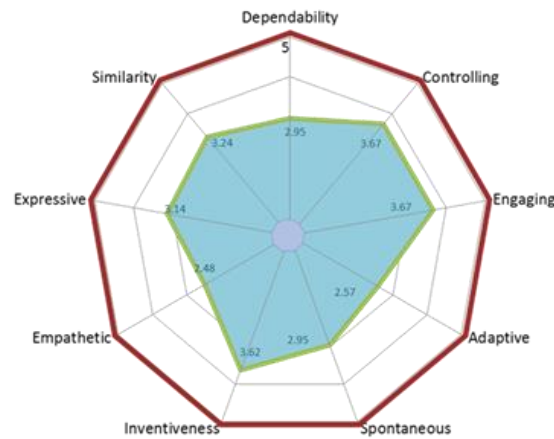

Fig. 3. Evolved Personality from elicitation activity

4 Developing Speech and Dialogue

In the next phase we conducted crowdsourcing and co-creation sessions with 32 participants which included language enthusiasts and avid readers of literature in order to create natural dialogues for the voice assistant in various situations. The intent of this experiment was to extract the underlying common patterns in responses from the users. The participants were given a set of user queries and corresponding responses from an existing voice assistant application covering various usage scenarios.

They were asked to frame the response of the voice assistant as natural and in the way they would prefer to hear. Participants could specify speech variables like pitch, speech rate and stress along with their language responses.
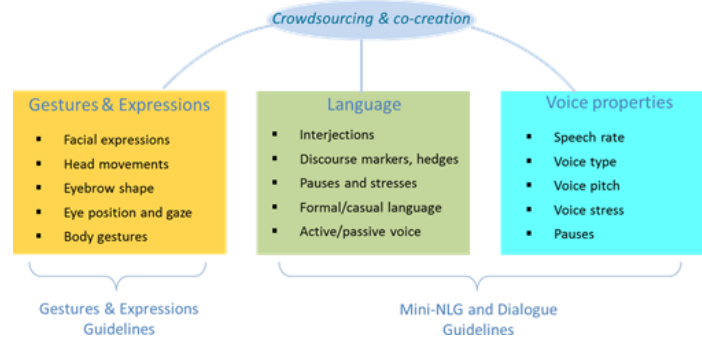

Fig. 4. Goals parameters for the co-creation activity

5 Creating Appearance and Gestures

As the next step, few of these raw dialogues from the participants in this earlier session were given to 6 animation designers with the intent of generating the visuals of the gestural aspects. This helped us formulate various gestural cues like smiles, frowns, eyebrow shapes, head nods, head positions, body postures, etc. [5].

6 Conclusions

The design guidelines, dialogues and personality definitions evolved from our research is being used as a framework for creating natural voice assistant avatar. Going forward, we plan to extend this research in exploring the notion of Voice as a Companion, beyond being an Assistant.

References

1. Meerbeek, B., Hoonhout, J., Bingley, P., Terken, S.: Investigating the relationship between the personality of a robotic TV assistant and the level of user control. In Robot and Human Interactive Communication, pp. 404-410, 2006.
2. Torrey, C., Fussell, S., Kiesler, S.: How a robot should give advice. In: 8th ACM/IEEE Human-Robot Interaction (HRI), pp. 275-282, (2013).
3. Ball, G., and Breese, J.: Emotion and personality in a conversational agent. Embodied conversational agents pp.189-219 (2000).
4. Costa, P., Robert, M.: Revised NEO Personality Inventory and NEO Five-Factor Inventory: Professional Manual. Psychological Assessment Resources, (1992).
5. Breese, J., Ball,G.: Modeling emotional state and personality for conversational agents. Rapport technique MSR-TR-98-41, Microsoft research (1998).