



Challenging the Voice Hype

Intro

Hype?

*Human
Interaction*



Use it

*Maintain
it*

*Understand
it*

Feel it

Trust it

*Where
next?*

*Final
Thoughts*

Nick Marfleet

Cofounder & CTO -
Lumin





Video Calling



Home Sensor Monitoring



Zigbee Connectivity



Health & Wellness Monitoring



Entertainment & Games



Medication Reminders & Delivery



Wearables



Doctors Appointments




Voice Control

powered by Sirv



Emergency Alerts



"The most profound technologies are the ones that disappear. They weave themselves into the fabric of everyday life until they are indistinguishable from it"

Marc Weiser

Challenging the voice hype



Is it hype?

No.

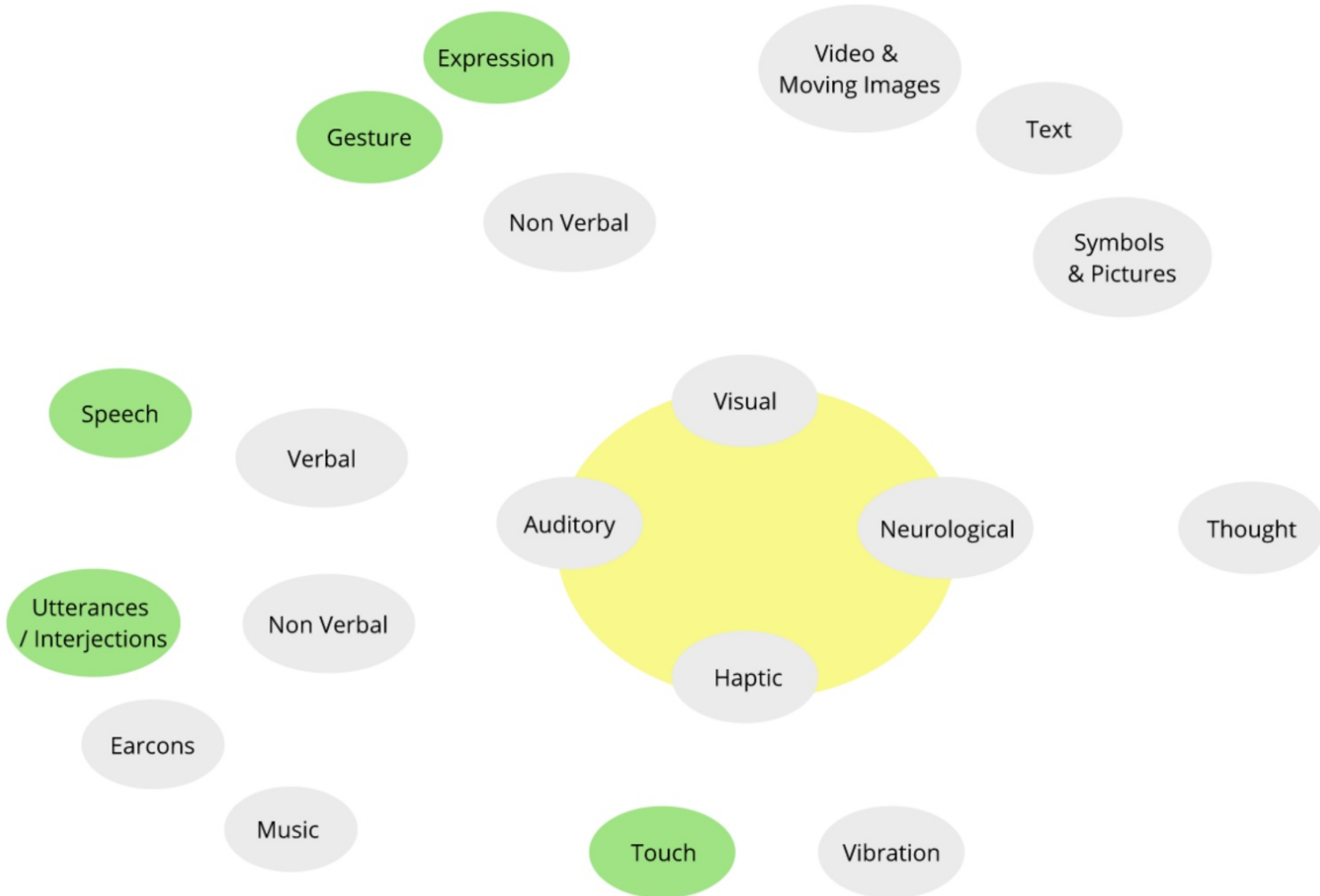


If designed and executed properly, as part of a complete, secure ecosystem.

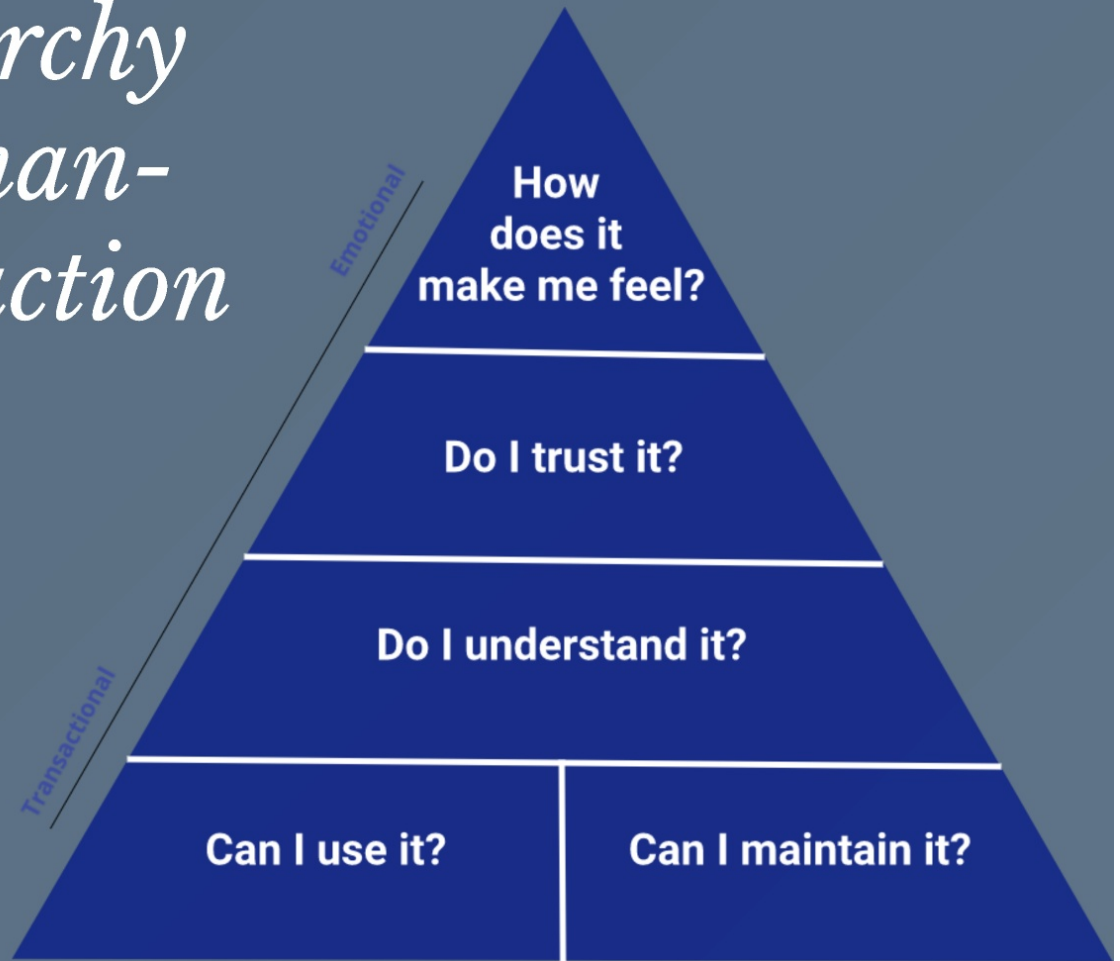
It's all about Interaction

*Modes of
Interaction*





Marfleet's hierarchy of needs in human-computer interaction





“Adults as they get older experience a wide range of age related impairments including loss of vision, hearing, memory and mobility, the combined effects of which contribute to loss of confidence and difficulties in orientation and absorption of information.”

“voice input is the most effective and preferable input modality for older adults.”

"The accuracy of ASR decreased not only for people with AD, but also for those with age-related cognitive decline."

Weiner et al

Results show that the Automatic Speech Recognition (ASR) Word Error Rates (WERs) for elderly voices are significantly higher than those of adult voices. The word error rate increases gradually as the age of the elderly speakers increase

Longitudinal study of ASR performance on ageing voices
Vipperla, Renals, Frankel, 2008

"Pauses, hesitations, and other disfluency features increase with cognitive decline."

Effects of Age-Related Cognitive Decline on Elderly User
Interactions with Voice-Based Dialogue Systems.
Kobayashi et al, 2019

“memory retention at the interface works differently for older and younger adult users.

While younger adults are able to accommodate differences of length of output message , older adults are not”

Zajicek & Morrissey

"The ephemeral nature of output in auditory interfaces places **significant cognitive demands on the user.**"

"Multimodal interfaces that combine speech with other modalities have the potential to mitigate some of these problems. *Even a small screen, when combined effectively with a speech interface, can **significantly reduce the cognitive demands on the user**"

Voice User Interface Design

Jennifer Balogh; James P. Giangola; Michael H. Cohen.
Addison-Wesley Professional, 2004

Can I Maintain it?



#fail

WHO SUPPORTS OLDER ADULTS WITH ICT USAGE?

“The results show, that it is mostly their family members who help older adults solve technology-related problems. Moreover, the usage of ICT influences the daily structure and variety of everyday life.”

#fail

#fail

Do I understand it?



How do I interact with this thing?


What expectations should I have about it?

Do I know it's limitations?


Enhancing the explanatory power of usability heuristics - Jacob Nielsen, 1994



Visibility of System Status




Match system to real world



User Control and Freedom




Consistency and Standards




Recognition rather than Recall




Error Prevention




Flexibility and Efficiency of Use



Aesthetic & Minimal Design



Help Users with Errors



Help and Documentation

Image created by @the-ux-evangelist — Icons: Freepik from www.flaticon.com

An Aging-focused Heuristic Evaluation of Home Automation Controls

Lopez, Jeremy; Textor, Claire; Hicks, W. Braxton; Pryor, Makenzie; McLaughlin, Anne Collins; et al.

Proceedings of the Human Factors and Ergonomics Society Annual Meeting, Vol 63 (1): 5 – Nov 1, 2019



Accessibility



Reliance on
timing



Trust in
the system



Privacy /
disclosure



Accessibility

Can I Use it?



Reliance on
timing

Can I Use it.



Trust in
the system

Can I Trust it?



Privacy /
disclosure

Can I Trust it?



Reliance on
timing

Google assistant allows approximately 7 seconds to deliver a clear and concise command.

yet ...

"Older adults may experience increased difficulty with forced deadlines since word retrieval degrades with age"

Burke and Shafto, 2004



Discovery

"Language is a form of **Joint Action**.

A joint action is one that is carried out by an ensemble of people acting in coordination with each other"

Using Language,
Herbert H Clarke.



“The critical property of these systems is that, since different modes offer different affordances for expressing content, through multimodality the user (or system) is empowered to use the mode or combination of modes best suited to the specific information to be conveyed or task to be completed.”

Extensible Multimodal annotation for
Intelligent Interactive Systems,
Michael Johnson

Do I trust it?



Last Week...

Devices unintentionally activate between 1.5 and 19 times
per day
(24 hour period)

Mon(iot)r Research Group @ Northeastern University, 2020

“The Echo Dot did record and stream private conversations to AVS without utilizing the wakeword.”

Alexa, are you listening to me? An analysis of Alexa voice service network traffic
Ford, Marcia; Palmer, William
Personal and Ubiquitous Computing, Volume 23 (1) – Jun 28, 2018

“Due to the long-held “third-party doctrine,” consumers relinquish their right to privacy when they engage a third party such as Amazon and Alexa”

**Alexa is
laughing
at users
& creeping
them out**

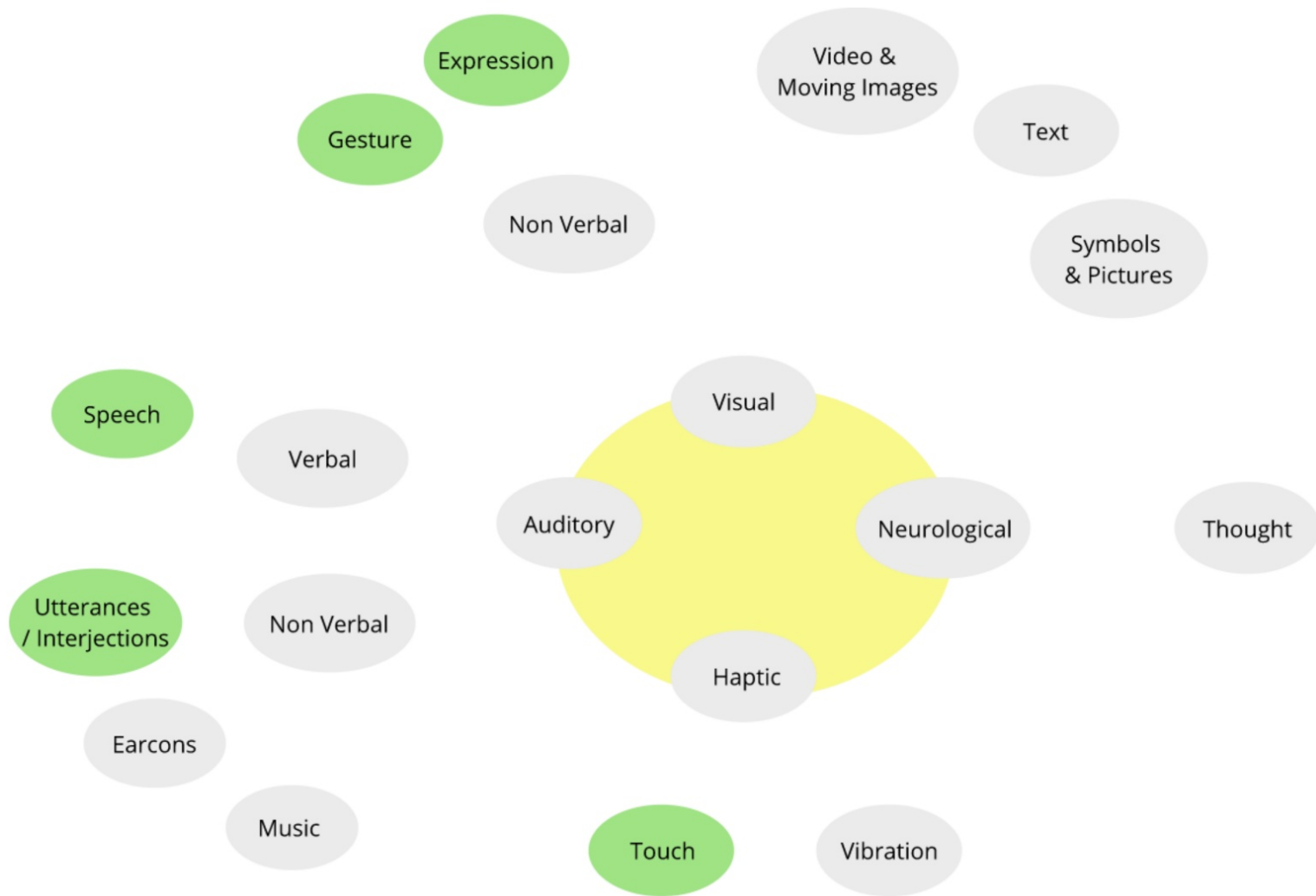


**Alexa is
laughing
at users
& creeping
them out**



How does it make me feel?





Parasocial Relationship Theory

a kind of psychological relationship experienced by members of an audience in their mediated encounters with certain performers in the mass media

Viewers or listeners come to consider media personalities as friends

Interpersonal Attraction

“An individual’s tendency or pre-disposition to evaluate another person or the symbol of the person in a positive (or negative) way”

Berscheid & Hatfield 1978

“Like a face-to-face relationship, PSR can be enhanced by interpersonal attraction”

Rubin & McHugh 87, Lee & Kwon 2013, Rubin & Step 2000



Task Attraction



Social Attraction



Physical Attraction

Oh, and security

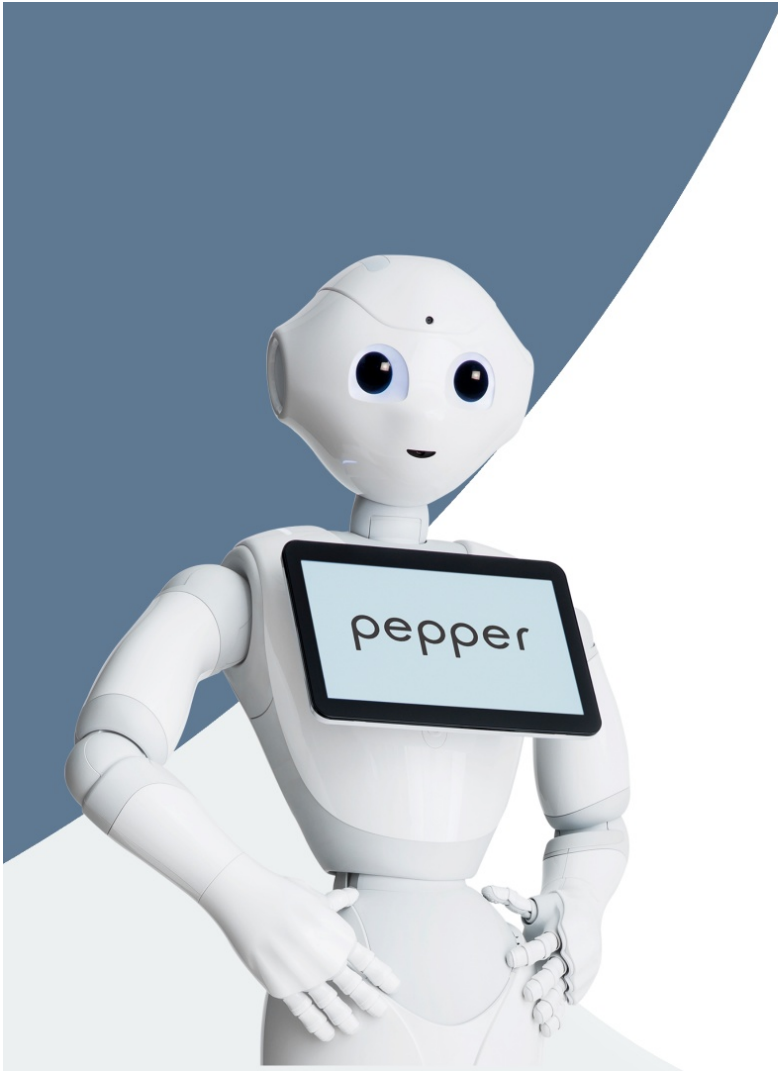
"Extending PSR, this study finds that Interpersonal Attraction AND security/privacy risk are important factors affecting the adoption of IPAs"

Understanding adoption of intelligent personal assistants – A parasocial relationship perspective
- Sangyeal Han, Heetae Yang



A Multimodal Future

Robots

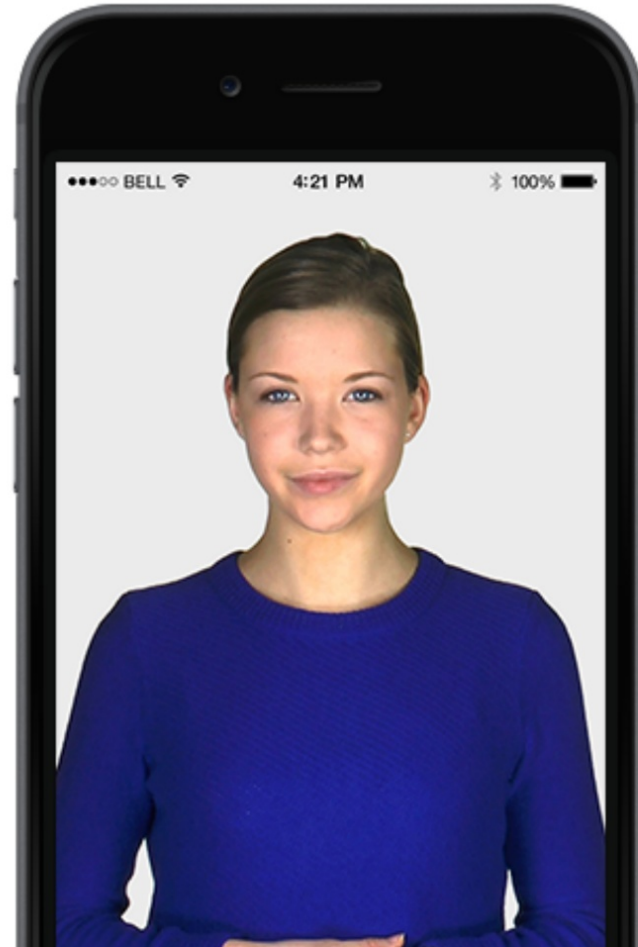


The problem with robots.

Fixed format of hardware.

Limited potential to provide anthropomorphic variations
(gender, ethnic variations etc)

Avatars



True Image Interactive

Final Thoughts

Don't just think that a consumer grade product will solve all your problems.

Do understand how you will design, deploy, manage and maintain a system - with careful consideration for the demographic.

Look to use multimodal solutions

Thank you.

Nick Marfleet
nick@mylumin.org



Challenging the Voice Hype

Intro

Hype?

*Human
Interaction*



Use it

*Maintain
it*

*Understand
it*

Feel it

Trust it

*Where
next?*

*Final
Thoughts*