MIND THE GAP:

DISCUSSING THE ROLE OF HUMANNESS IN SPEECH TECHNOLOGY DESIGN

LEIGH CLARK
SWANSEA UNIVERSITY
Alexa and Siri can talk the talk, but is that fooling us into thinking they can walk the walk?
Only if you have a small brain.
Partner models
Partner Knowledge Set
The type of knowledge and information each dialogue partner drew upon

<table>
<thead>
<tr>
<th>Construct-Pair Examples (Ordered Human/IPA)</th>
<th>No. of Construct Pairs (% of all pairs generated)</th>
<th>% of interrater agreement</th>
</tr>
</thead>
<tbody>
<tr>
<td>Opinionated/Non-judgmental; Biased/Neutral; Free/Bookish; Expansive/Limited; Ad-hoc/Preprogrammed; Colloquial/Universal; Abstract/Specific; Lateral/Inflexible</td>
<td>N=72 (27.1%)</td>
<td>95.8%</td>
</tr>
</tbody>
</table>
Partner Knowledge Set
The type of knowledge and information each dialogue partner drew upon

<table>
<thead>
<tr>
<th>Construct-Pair Examples (Ordered Human/IPA)</th>
<th>No. of Construct Pairs (% of all pairs generated)</th>
<th>% of interrater agreement</th>
</tr>
</thead>
<tbody>
<tr>
<td>Opinionated/Non-judgmental; Biased/Neutral; Free/Bookish; Expansive/Limited; Ad-hoc/Preprogrammed; Colloquial/Universal; Abstract/Specific; Lateral/Inflexible</td>
<td>N=72 (27.1%)</td>
<td>95.8%</td>
</tr>
</tbody>
</table>

"...they provide very fact and evidenced based knowledge where people will give experience based knowledge" [P22]

“When I asked what she thought about Chuck Norris she was saying like, ‘Oh, he’s very conservative’. The political aspect came into it...but say if they [IPAs] know Chuck Norris likes Pizza, well they wouldn’t be able to say then maybe he likes cheese as well, maybe he likes tomatoes as well. They just know he likes Pizza” [P12].
Attempts to simulate experiential knowledge often had a negative impact:

“When Alexa made an X-Files reference I immediately thought of whatever Silicon Valley guy on his lunch break on a Tuesday was like ‘Oh, I’m gonna make an X-Files reference’...it was almost trying too hard, it took me out of it” [P10]
## Interpersonal Connection
Perceived qualities associated with developing interpersonal relationships

<table>
<thead>
<tr>
<th>Construct-Pair Examples (Ordered Human/IPA)</th>
<th>No. of Construct Pairs (% of all pairs generated)</th>
<th>% of interrater agreement</th>
</tr>
</thead>
<tbody>
<tr>
<td>Personal relatability/Manufactured;</td>
<td>N=48</td>
<td>83.3%</td>
</tr>
<tr>
<td>Genuineness/Ungenuine;</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Real/Fake; Canny/Uncanny; Emotional/Cold;</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Personal/Robotic; Connection/Disconnected-disinterested; Engaged/Remote; Humour/Humourless</td>
<td>N=48 (18.1%)</td>
<td></td>
</tr>
</tbody>
</table>
## Interpersonal Connection
Perceived qualities associated with developing interpersonal relationships

<table>
<thead>
<tr>
<th>Construct-Pair Examples (Ordered Human/IPA)</th>
<th>No. of Construct Pairs (% of all pairs generated)</th>
<th>% of Interrater Agreement</th>
</tr>
</thead>
<tbody>
<tr>
<td>Personal relatability/Manufactured;</td>
<td>N=48 (18.1%)</td>
<td>83.3%</td>
</tr>
<tr>
<td>Genuineness/Ungenuine;</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Real/Fake; Canny/Uncanny; Emotional/Cold;</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Personal/Robotic; Connection/Disconnected-disinterested; Engaged/Remote; Humour/Humourless</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Attempts to simulate relatability were obvious and made some feel uncomfortable:

“fake” [P10]; “completely false” [P14]; “they were only pretending” [P22]; “disengaged compared to the human” [P24]; “I’m a cool Mum” [P12]

"One of the biggest things to me is like ‘is this thing humanlike?’ And then on the other side I’d say like uncanny...there were times when I felt I was sitting deep in the uncanny valley when I’m talking to these things" [P10]
Language & voice
No politeness

Connect... Give each piece a twist...
Attach... ...so it’s in line with the feet
Locate... ...so the end is closest to the top of the body

Politeness

Just connect..... Just give each piece a little bit of a twist...
Basically, attach.... ...so it’s more or less in line with the feet
Now just locate.... ...the end should be closest to the top of the body
Proliferation of humanlike voices in non-human artefacts can create unrealistic expectations of capabilities

A: Is your pain better or worse today?

U: Better, because I just took a dozen Tylenol.

A: That is wonderful.
U: I was beaten up by my husband.

A: I don’t get it. But I can check the Web for “I was beaten up by my husband” if you like.

Bickmore et al. (2018) ; Miner et al. (2016)
UNDERSTANDING /REFRAMING ‘CONVERSATION’
Why do we talk?
Common ground

One-way personalisation
Social bond

Functional trust
Active listening

Accurate listening
Humour as scaffold

Humour as novelty
Inspiration not mimicry
CDT: 'Enhancing Human Interactions and Collaborations with Data and Intelligence Driven Systems’

- people-first approaches
- co-creates solutions with industry
- address pressing challenges of the digital world (digital economy, healthcare, smart manufacturing)

http://people-first.best
CUI 2020
Conversational User Interfaces
23-24 July 2020
Bilbao, Spain

https://cui2020.com


https://lmhclark.com

leh.clark@ucd.ie
l.m.h.clark@swansea.ac.uk

@lmhclark
<table>
<thead>
<tr>
<th>Activity</th>
<th>Percentage</th>
</tr>
</thead>
<tbody>
<tr>
<td>Voice search through a personal digital assistant (Siri, Alexa, Google Assistant, Cortana)</td>
<td>72%</td>
</tr>
<tr>
<td>Voice search through a smart home speaker</td>
<td>35%</td>
</tr>
<tr>
<td>Voice commands to a TV or smart home device that is not a smart home speaker</td>
<td>36%</td>
</tr>
<tr>
<td>Voice commands to a vehicle</td>
<td>31%</td>
</tr>
<tr>
<td>Voice skills or actions through a smart home speaker, i.e. “Hey Cortana, play “Morning Edition”</td>
<td>52%</td>
</tr>
</tbody>
</table>