Empathy consideration in the design of natural language interfaces for future vehicles

Ben Anyasodo, Gary Burnett
Human Factors Research Group,
Faculty of Engineering,
University of Nottingham,
Nottingham, NG7 2RD, UK

ABSTRACT

While the future of transportation paints a picture of seamless understanding of the passenger goals by the vehicle, it also exposes a gap in understanding what the human-machine engagement must become for a more natural in-car experience. Through natural language interfaces, a human-like interaction is possible. This raises the question *is there a way that machines can be more "empathic"?* If so, would this make for a more natural human-machine interaction? And how can we design usable natural language interfaces (specifically speech systems) to achieve this? Especially because although humans are emotional beings, machines are not. This paper explores the concept of empathy for speech systems by investigating the human-human empathy model and proposing design considerations. To achieve this, we interviewed professional persons who have to show empathy as part of their work. Seven themes were generated from the responses that form a usable framework for a human-machine empathy which could be applied to natural language speech system design.

Keywords

Empathy, Driving, Natural language, Speech system, HCI limitations, dialogue management, future vehicles

INTRODUCTION

Natural language input is widely applied in vehicle speech systems more and more. However, system inaccuracies such as task complexity, naturalness of speech synthesis, system responsiveness/feedback, context awareness (given differing user/driving conditions), dialogue management etc); have continued to generate human-computer interaction (HCI) limitations (Weng et al, 2016). Such HCI limitations impact the user experience or any effect of human-like interaction on the user, regardless of the intelligence shown by natural language speech systems e.g., giving correct answers, predicting input, machine learning (Jenness et al, 2016). In an in-vehicle context this gets compounded further as most interaction with these speech systems happen as a secondary task, while driving is the primary task. Therefore, any secondary task must not increase cognitive load as that compromises driving safety. However, it has been shown that these HCI limitations increase user's cognitive load (Becker et al, 2006). According to Shneidermann (2000), the interaction becomes less empathic as cognitive demand increases on the user. In future vehicles, although users might be involved in something else other than driving as a primary task, reducing safety concerns, enhancing trust and improving user experience remain core priorities (Rödel et al, 2014).

This study is focused on understanding elements of human-human interaction that may inform a natural and empathic human-machine interaction in natural language speech interfaces/systems, as a way of improving the overall user experience and system goals. The central question considered here is: "*how*

can an in-vehicle Human-machine Interface (HMi) show empathy?" In other words, "how can the user interaction with the HMi be adapted in a human-like manner based on interpreting user emotion?".

The Empathy Model

Empathy is defined as "reactions of one individual to the observed experiences of another (Davis, 1983)." In a human-human interaction, it is that thing which makes it possible to vicariously experience the other person's feeling as if it were one's own, and as such respond or react accordingly. With this in mind, the question above can be rephrased as *"how can a machine communicate with feelings that mirror some of the feelings of the human interacting with it in a natural manner?"*

Empathy – as a word comes from Greek *em – into* and *pathos – passion, feeling.* It was originally a German word 'Einfühlung' (meaning "feeling into") which was used in the philosophy of art in the 18th century to describe a woman's projection of her entire personality on an object (Kouprie & Visser, 2009). Two commonly described components of empathy are: Affective empathy and Cognitive empathy. Affective empathy is also known as Emotional empathy and describes an immediate automatic response to another person's emotional state (Davis, 1983). It is '*the capacity to share another's emotional state, and to experience the feelings of the observed person*' (Cohen & Strayer, 1996). Cognitive empathy describes the role playing/perspective aspect of empathy, in which the empathiser takes the role of the other person by imagining it from their own point of view (Mead, 1934). The perspective-taking (or role taking) element is key in this case since that is how the shared experience is created. Cognitive empathy is about understanding the other person's perspective and is more than an emotional contagion (Stotland & Dunn, 1963), since it also depends on higher cognitive functions and requiring cognitive flexibility (Decety & Jackson, 2004).

In practice, affective and cognitive elements of empathy are interrelated and occur together. This is captured in Cohen & Strayer's definition of empathy as "*the ability to understand and share in another's emotional state and context'* (Cohen & Strayer, 1996). In that definition, the "understand" element is cognitive, and "share" is affective. This shared emotional experience concept forms the basis of the Perception-action model (PAM) of empathy, which describes the resulting process as on in which '*the representation of the subject's emotional states is automatically activated when the subject pays attention to the emotional state of the object'* (Preston, 2007: 428-447). In the PAM model, empathy has a linear relationship with past experience, similarity, and familiarity (Barnet & McCoy, 1989).

From the foregoing, empathy employs a vicarious effect since it implies one person "entering" or "experiencing" another person's feeling – and this feeling might be past, present, or future. (Low, 2012). Experiencing this feeling therefore involves being sensitive to the other person's psychological state, and in some cases even anticipating what next state they may get into. Empathy between two persons in an interaction implies a give and take relationship, thus defining it as being both a process and an outcome (Davis, 1983). As a process, it is the actual "perspective/role taking" that happens to the "empathiser" as they make sense of the other person's feelings. As an outcome, it is the result of the process, and is either an affective outcome or a cognitive outcome.

Scott et al (2008) view empathy as either Parallel or Reactive. In parallel empathy, the empathiser exhibits exactly the same emotion as the target. The empathiser must be able to identify this state. Therefore, this empathy is self-oriented. In reactive empathy however, the empathiser shows an emotion different to that of the target and can enhance or change the target's affective state. Scott et al consider reactive empathy as a "higher level of empathic behaviour" since it focuses on the target.

Empathic communication was also explored in the Barrett-Lennard model (1981) and Suchman et al model (1997). They saw empathic communication as a matter of opportunity and response – i.e., the empathee initiates the cycle, creating an "empathic opportunity". This empathic opportunity is then recognised by the empathiser, who then completes the cycle by providing an "empathic response" (Lapointe, 2014). Lapointe demonstrated this in a clinical study of empathic communication between nurses and patients. Ioannidou & Konstantikaki (2008) described the following characteristics as 'basic conditions' relevant to a fruitful empathic communication: Emotional understanding (i.e., health care professionals understand the problem through the patient's point of view), Respect, Authenticity (i.e., honesty, real expression of views without hypocrisy), Warmth and unconditional positive recognition, Self-exposure and Resolution.

To investigate aspects of the application of the empathy process in human-human interaction that might be applicable to an empathy framework for human machine interaction, an interview study was conducted which aimed to investigate how empathy is understood, expressed and interpreted by people who work in social professions that require empathy. This user group is of significance because the nature of such profession as therapy, health care, teaching etc**.** puts empathy right at the core of the work (Hojat, 2007). As such people in these professions master how to express empathy at will and this is achieved through numerous training programs and practice (Altmann et al, 2015). For example, a driving instructor must demonstrate understanding of their student's stress in order to be able to calm them down and take them through a new learning experience. Therefore, their day-to-day way of working can be said to be a good human-human interaction model involving empathy. Also, through their training, these professionals who have to consciously empathise as part of their work, are hoped to be able to articulate how they demonstrate empathy in their experience. Hypothetically, a preliminary understanding of empathy in human-human interaction could serve as a basis for a framework encouraging empathy in vehicle-based human-computer dialogues. This means understanding what informs a successful empathic outcome in the professions of interest, and what that implies for the empathy cycle. An outcome of the literature review in this section is a framework that suggests the following important empathiser conditions for empathy: adapting behaviour, awareness (how conscious the empathiser is of their own empathy actions and emotional state, at every point during the interaction) and appropriate syntax. In addition to these, this study also explores the impact of training and empathiser personality in the empathy process.

METHOD

Participants

The study was carried out with 15 participants (7 male, 8 female; age range 37 to 63) – ten driving instructors and five telephone councillors. The sample size was adopted in line with recommendation by Mason (2010) and Bertaux (1981, cited by Mason, 2010). Driving Instructors were of interest, particularly because their role (teaching, assuring, protecting, and encouraging) involves empathy in the confinement of the car context, where they can often get involved with small talk and chats that may be useful in creating their empathic interaction, but without much form of direct eye contact, given that they sit side-by-side – this is a relevant consideration in the operation of a natural language interface. Unlike driving instructors who occupy same physical space with their clients, telephone councillors achieve empathic outcomes remotely i.e., non-face-to-face. Again, this is of interest, as it mimics the context of a potential natural language speech system and the human-machine interaction that could arise. All participants were experienced in their profession (all >10 years' experience).

Design Method

This study utilised the interviewing method, and a thematic analysis of the data was undertaken. The study took place as a series of interviews of 30 to 50 minutes each.

The study is more focused on understanding the process by which empathy is achieved and maintained in the chosen case study areas, rather than just the outcome of the interactions, hence the qualitative approach was considered most appropriate for a study of this nature (Creswell, 1998). Thematic analysis is a common methodology used in Human Factors, and in this case, has been chosen since the study seeks to understand patterns and themes in human-human empathic interactions.

A pilot study was carried out with three participants who were not "professional empathisers". Through the pilot, the researcher built more confidence in the methodology and flow of the interview questions. Also, the phrasing of the questions was improved based on the pilot interviews. The questions used in the final study were a 3rd iteration from the original questions.

Data collection was undertaken via audio recording of the interviews (and subsequent transcription of the recording) as well as note taking during the sessions.

Procedure

In a semi-structured format, participants were asked the following questions in turn:

1. In the context of your job, what do you think it means to show empathy?
2. What is important to you when you show empathy?
3. What 3 things do you notice first with a new client?
4. What 3 aspects of your training help you create an empathic interaction with your client?
5. What 3 things do you become aware of in yourself during your interaction?
6. What 3 things do you change in your behaviour during your interaction?
7. What 5 words/phrases do you use to achieve a positive outcome? Are there any words you avoid?
8. What would you hear from your client that lets you know you are in rapport?
9. How do you know the interaction is empathic?
10. How much of your personality do you bring into your work? How much do you modify?

The questions were designed to open up relevant conversations while focussing the participant on the exact situations where they could recall their experience and so provide a more informative response. The use of open questions ensured that best possible data was received from the respondents. Follow-up questions were also asked where a deeper-dive necessary.

RESULTS AND ANALYSIS

Notes and transcription from audio recording were analysed to find common themes relating to how the professionals show, assess and maintain empathy, as well as any associated behaviour of their clients in the specific context. Having familiarised with and coded the data, themes were created to describe groups of similar observations from the participants. The themes were reviewed and refined into a total of seven named themes. These themes are grouped into two main categories – mental model themes (defined by patterns in mental models of the empathy transaction) and vocal attributes themes (defined by communicating empathy verbally). Most of the themes apply to assessing and maintaining empathy

in the interaction. In this report, "empathiser" is used interchangeably for the professional and "target" is used interchangeably with client.

Mental Model Themes (MMT)

Four Themes emerged that describe patterns in mental models of the empathy transaction:

Theme one (MMT1)– the empathiser has a pre-conceived internal state, which is then modified constantly during the interaction to achieve the outcome set out initially.

- There is a "preparation to empathise", as well the awareness of the need to modify language, tone, posture, personality as required during the interaction
- This deliberateness, which is formulated through an understanding of previous interactions
- The empathiser maintains a positive motivation to show empathy
- the professional decides on choice of words to use – learnt from training and experience.

Theme two (MMT2) – empathy is about target's expectations

- It is also expected that the target's negative emotions, though shared, are converted to positive ones – by the empathiser.
- rapport is broken when the empathiser provides a feedback different to what is expected
  - *"they usually expect you to always remain calm, calculated and positive regardless of how negative or angry they feel. It's almost like I'm expected to always make them feel better"*
  - *"it doesn't matter how the client feels, they expect that I will make them feel more positive"*
  - *"even when they make a mistake, they expect that I will understand and still make them feel good"*
- Through experience and training, the professional develops a breath of understanding to the possible emotional/empathy expectations of the target
  - *"I can always tell from that glint in their eyes or their body language"*
  - *"the sort of phrases they use helps me assess their expectation. For example, if someone constantly says, "I don't know, I guess it is…." Straightaway I know they want me to assure them a bit more"*

Theme three (MMT3)– acknowledging the target's emotional state creates empathy

- *"they tend to relax more when I voice their feelings. For example, acknowledging that I know they are nervous"*

Theme four (MMT4) – an empathic interaction is really more of an outcome-based interaction rather than full 2-way transaction and the shared experience/emotions serves as a means to achieving the overall outcome.

- E.g., the outcome in a driving lesson context is learning & safety – i.e., the instructor creates and maintains an empathic interaction so that the student learns to drive safely. In a telephone counselling context, the outcome is that the target stops to panic and feels better regardless of their situation.
- *"even when they are saying something annoying, I have to still make them feel comfortable otherwise I won't be able to control the situation or stop them doing something stupid"*

Vocal Attribute Themes (VAT)

Three themes emerged to describe on communicating empathy vocally

Theme one (VAT1)– The target's tone of voice is a key input for the professionals, and serves as a feedback factor to help assess the empathy balance

- The professional assesses empathy by listening with intent
    - *"I use their tone of voice to gauge how we are doing"*
    - *"I match my tone of voice to theirs. Sometimes I have to deliberately make my voice sound brighter or darker"*
    - *"I know they feel understood as soon as they get comfortable enough to engage more and speak more"*
    - "*I always listen for a shift in tone, example an enthusiastic 'yes', 'ok' or just a generally less agitated tone"*

Theme two (VAT2) – small talk helps assess and maintain empathy

- Professionals use small talk to gain more insight into the target's linguistic style with an intention to mirror it
    - *"while making small talk, I get to hear the phrases the student uses naturally, and I can use them back as well"*
    - "*with small talk I can use open questions to make the student talk more and I observe how they speak – whether they speak faster or slower than me. This way I know how best to change my own tone to sound like them. For example, if they sound a bit squeaky, I make my voice more high pitched"*

Theme three (VAT3) – empathy is achieved when there is an overall positive feeling. This is described in relation to language and tonal expression

- *"I keep my language unambiguous, and in context"*
- *"they engage better when I maintain a friendly tone and assertiveness"*
- *"that feeling of empathy happens when I use words that are both confirming and affirming, so when I say things like "that's right" or "let me check that I understand you""*
- "*the client feels better when I communicate in a way that matches their level of urgency but without showing panic"*
- *"I won't say I'm normally a friendly person as such, but the client is not interested in that. They expect me to sound nice and friendly to them, full stop. When I don't, you see them begin to shut down or withdraw from any active engagement"*

DISCUSSION

From the themes, it is clear that the professionals demonstrate reasonable emotional understanding, by understanding their client's emotional state, from the client's point of view. They then take a reactive approach to shift this emotional state to a more positive outcome (Scott et al, 2008). The professionals apply cognitive flexibility and look out for the empathic opportunity from the client, which becomes the input for the feedback loop. This loop is then completed through an empathic response, thus further demonstrating the Lapointe model, and in a non-clinical context. With both driving instructors and telephone counsellors, the basic conditions as posited by Ioannidou & Konstantikaki (2008) are also observed. Driving instructors used 2 modalities (vision i.e., client's body language and sound i.e., words and tone) as input for the interaction, while telephone counsellors used only speech/sound. This

difference in input modality had no effect in the professional's empathy response, which in both cases was delivered as speech. Both contexts rely more on client's speech attributes than physiological attributes.

The human-human framework in discussion translates to a machine framework that therefore requires the system to:

1. demonstrate a motivation to show empathy - perhaps through the conversation opening (MMT1 & VAT1). For instance, ensuring the dialogue is initiated in a manner that communicates an openness to listen (where the machine initiates the dialogue). This may be achieved through calm tone and appropriate language. E.g., “I am open to your requests and can assist you in any way possible” (or a similar phrase) may communicate this open internal state better than “hello, how may I help”. The key here being that “deliberateness” that is expressed.
2. be able to understand the user's emotional state, and adapt to it through user speech modalities and conversation predicates. This means being able to detect elements of empathy in user speech input and natural language understanding and sharing in it - by mapping the input to a suite of natural language calibrations (in the Automatic speech recognition engine perhaps), while generating a response (MMT3 & VAT2). This adaptability is demonstrated as a flexibility in subtly modifying language/speech attributes throughout the interaction, based on user emotions. This suggests real-time switching within a suite of calibrations that are developed through understanding basic human emotions. This suite thus becomes a core of the inherent intelligence that the system uses (e.g., through a listening mode) to identify an empathic opportunity from the user, and complete the loop by returning an empathic response.
3. demonstrate that understanding by returning an empathic response, using appropriate unambiguous, positive, affirmative language, that also matches (or at least represents) user level of urgency (VAT3).

HCI limitations have been identified as a main factor that negatively impacts user experience with speech interfaces (Reisinger et al, 2005). These "deficiencies in the completeness with which the interface models human conversation", lead to undesirable user experience and causing frustration. A direct consequence of this is a breakdown in the dialogue - i.e., a situation where user input is misinterpreted or even unrecognised. From MMT2, it is important that the empathiser recognises how the interaction can break down with incorrect response. This implies that detecting potential dialogue breakdown indicators (such as repetition of command, certain phraseology including expletives, sudden variance in tone) is important in a natural language system. Beyond detecting a potential breakdown, the system should be able to repair the dialogue and then apply the event as an input to increase machine intelligence - thereby proactively improving subsequent interactions. In addition to safety, one of the key considerations in future vehicles is trust (Antrobus et al, 2018), which has also been shown to be an outcome of empathy (Ickes et al, 1990). As such, an application of the themes in designing a natural language speech interface could enhance trust and user experience, as well as safety. MMT4 also suggests that this outcome is non-negotiable, and therefore an empathic approach is a key way for the system to utilise the rest of the interaction as important inputs that serve as a means to achieving the end goal.

CONCLUSION AND FUTURE WORK

This research considered the possibility of a machine showing empathy, as a way to improve user experience and system design of the vehicle natural language interface. This was achieved through investigating how empathy is understood, expressed, and interpreted by people who work in social

professions that require empathy, to identify key aspects of human-human interaction that are core to an empathic interaction. The findings suggest that empathy considerations could design out some of these issues around response generation in a natural language interface and enhance natural language understanding. Doing so improves user experience and safety in an in-vehicle context.

This paper concludes that the human-human empathy framework presents useful design implications for a natural language interface. While it is believed that the themes generated are applicable to natural language interfaces as a whole, this study has been carried out with an in-vehicle context in mind. Therefore, further study should test the framework in wider contexts outside of in-vehicle, as well as how much empathy the machine can show in specific in-vehicle use cases.

REFERENCES

1. Altmann, Tobias, Victoria Schönefeld, and Marcus Roth. "Evaluation of an empathy training program to prevent emotional maladjustment symptoms in social professions." *Psychology* 6.14 (2015): 1893.
2. Antrobus V., Burnett G., Large D. 2018. 'Trust me – I'm AutoCAB': Using natural language interfaces to improve the trust and acceptance of level 4/5 autonomous vehicles. Proceedings of the 6th Humanist Conference, The Hague, Netherlands, 13-14 June 2018
3. Barnett, M. A., & McCoy, S. J. (1989). The relation of distressful childhood experiences and empathy in college undergraduates. The Journal of Genetic Psychology; Child Behavior, Animal Behavior, and Comparative Psychology, 150(4), 417–426.
4. Baron-Cohen, S., Wheelwright, S., Hill, J., Raste, Y. & Plumb, I. (2001). The 'Reading the mind in the eyes' test revised version: a study with normal adults, and adults with Asperger syndrome or high-functioning autism. Journal of Child Psychology & Psychiatry & Allied Disciplines 42, 241–251.
5. Cohen, Douglas, and Janet Strayer. "Empathy in conduct-disordered and comparison youth." Developmental Psychology 32 (1996): 988–98.
6. Creswell, John W., and Cheryl N. Poth. Qualitative inquiry and research design: Choosing among five approaches. Sage publications, 1998.
7. Davis, M. H. (1980). A multidimensional approach to individual differences in empathy. JSAS Catalog of Selected Documents in Psychology, 10, 85.
8. Davis, M. H. (1983). Measuring individual differences in empathy: Evidence for a multidimensional approach. Journal of Personality and Social Psychology, 44, 113– 126.
9. Decety, Jean, and Philip L. Jackson. "The functional architecture of human empathy." *Behavioral and cognitive neuroscience reviews* 3.2 (2004): 71-100.
10. F. Weng, P. Angkititrakul, E. E. Shriberg, L. Heck, S. Peters and J. H. L. Hansen, "Conversational In-Vehicle Dialog Systems: The past, present, and future," in IEEE Signal Processing Magazine, vol. 33, no. 6, pp. 49-60, Nov. 2016, doi: 10.1109/MSP.2016.2599201.
11. Ickes, W., Stinson, L., Bissonnette, V., and Garcia, S. 1990, Naturalistic social cognition: empathic accuracy in mixed-sex dyads. *Journal of Personality and Social Psychology*, 59, 730-742.
12. Ioannidou, F., and V. Konstantikaki. "Empathy and emotional intelligence: What is it really about?" *International Journal of Caring Sciences* 1.3 (2008): 118.

13. Jenness, James W., et al. In-Vehicle Voice Control Interface Performance Evaluation. No. DOT HS 812 314. 2016.
14. Kouprie, Merlijn, and Froukje Sleeswijk Visser. "A framework for empathy in design: stepping into and out of the user's life." *Journal of Engineering Design* 20.5 (2009): 437-448.
15. Lapointe, Stephanie. A corpus study of the verbal communication of empathy/sympathy by anglophone nurses in Quebec. Diss. Université du Québec à Chicoutimi, 2014.
16. Low, K. C. "Being Empathetic, the Way of Confucius." (2012).
17. Mason, Mark. "Sample size and saturation in PhD studies using qualitative interviews." *Forum qualitative Sozialforschung/Forum: qualitative social research*. Vol. 11. No. 3. 2010.
18. Preston, Stephanie D. "A perception-action model for empathy." *Empathy in mental illness* (2007): 428-447.
19. Reisinger, J., Viswanathan, M., and Comerford, L. 2005. “Conversational Assistant for In-car Task Management,” in Proceedings of the 2005 AAAI Spring Symposium on Persistent Assistants, Stanford CA, USA
20. Rödel, C., Stadler, S., Meschtscherjakov, A., and Tscheligi, M.. Toward Autonomous Cars: The Effect of Autonomy Levels on Acceptance and User Experience. In Automotive UI ‘14: Proceedings of the 6th International Conference on Automotive User Interfaces and Interactive Vehicular Applications, Association for Computing Machinery, New York, 2014, pp. 1–8.
21. Shneiderman, Ben. "The limits of speech recognition." *Communications of the ACM* 43.9 (2000): 63-65.
22. Stotland E, Dunn RE. Empathy, self-esteem, and birth order. J Abnorm Soc Psychol 1963; 66: 532–40
23. Suchman, A. L., Markakis, K., Beckman, H. B., Frankel, R. (1997). A model of empathic communication in the medical interview. JAMA: The Journal of the American Medical Association, 277(8), 678-682. doi:10.1001/jama.l997.03540320082047