🏛 **UCL**

How people living with MND disease use personalized automated speech recognition technology to support conversation

Richard Cave, SLT, MRes RCSLT

Supervisors: Dr Steven Bloch, Professor Catherine Holloway

Richard.Cave17@ucl.ac.uk

@SpeedofTrust27

Language & Cognition, Division of Psychology & Lang Sciences, UCL

Disclosure

- Richard Cave is a Speech and Language Therapist (SLT) and PhD candidate at UCL.

- Richard provides specialist SLT consultancy to Google's Project Relate, Project Euphonia, Look To Speak projects amongst others, facilitating co-design with people living with ALS, head and neck cancer, laryngectomy, Cerebral Palsy, Stroke and many other conditions.

- Richard works with the MND Association - a non-profit in the UK for people living with ALS

- Clinical Ambassador to the Mouth Cancer Foundation – a non-profit in the UK

- National adviser for voice banking to the Royal College of Speech and Language Therapists.

- Richard by the contributes to the International Alliance of ALS Associations – a non-profit and was awarded the Allied Health Professional recognition award for 2022.

Presentation

Overview

1. Background
2. PhD aims
3. Longitudinal study
4. Results
5. Reflections and Recommendations

Amyotrophic Lateral Sclerosis (ALS)

- Motor Neurone Disease (MND), Lou Gehrig disease

- Affects the nerves in the brain and spinal cord

- Loss of function of limbs, weakness of muscles of the trunk and neck

- 15% plwMND may have symptoms similar to frontotemporal dementia (Bäumer et al., 2014)

- Average life expectancy of two to three years (Oliver et al., 2017)

Speech and Identity

- 80%+ plwMND eventually become unable to communicate using natural speech (Beukelman et al., 2011),

- Speech is a powerful medium of identity (Bucholtz & Hall, 2005)

- Mood, humour, geographical, social and educational background, health status, gender - as well as the content of the message (Nathanson, 2016)

- Sense of self, allowing listener to derive multiple levels of meaning. Identity constructed & mediated through social interaction (Wertsch, 1991)

# Augmentative and alternative communication (AAC)

- Up to 90% of people living with ALS may rely on AAC (Ball et al., 2004)

- AAC is slow compared to conversation and people may use other modalities to speed things up (Smith & Murray, 2016)

- Use of speech often preferred even if hard to understand (Smith, 2018)

- AAC design may not reflect how people want to communicate (Clarke et al., 2020)

Voice Banking

- A process for creating a 'personalised synthetic voice' (PSV), a synthetic approximation of a person's natural voice (Costello, 2016)

- Main reason for voice banking to preserve a sense of identity (Cave & Bloch, 2021b)

- PSV will not change the ease or speed of communication once they use AAC (Cave & Bloch, 2021b), and cannot replace the emotion and context reflected in natural speech (Pullin & Hennig, 2015)

[Mike] Yes, everything is changing. But I will always be me.

UCL
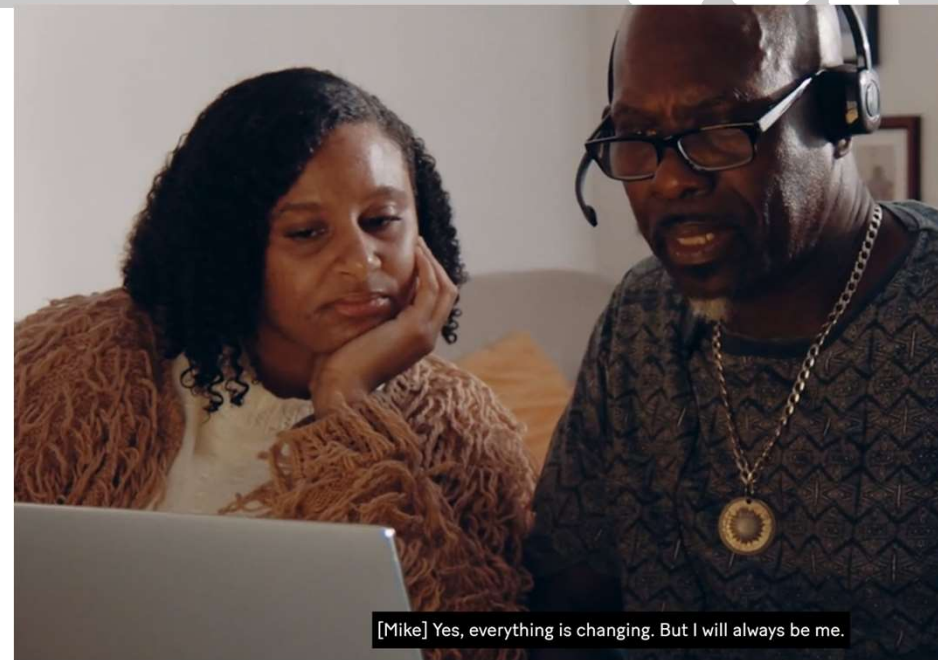
- Free Android app that aims to help people with non-standard speech communicate more easily with others and interact with the Google Assistant (Cattiau, 2021).

- Goal to improve personalized ASR accuracy for non-standard speech (Cattiau, 2019).

- May support the goal of AAC to preserve the speed and naturalness of spoken communication as far as possible (Hawley et al., 2013).

- May support 'social closeness' the preference to use own speech to communicate rather than through AAC (Murphy, 2004).


Helping people be better understood
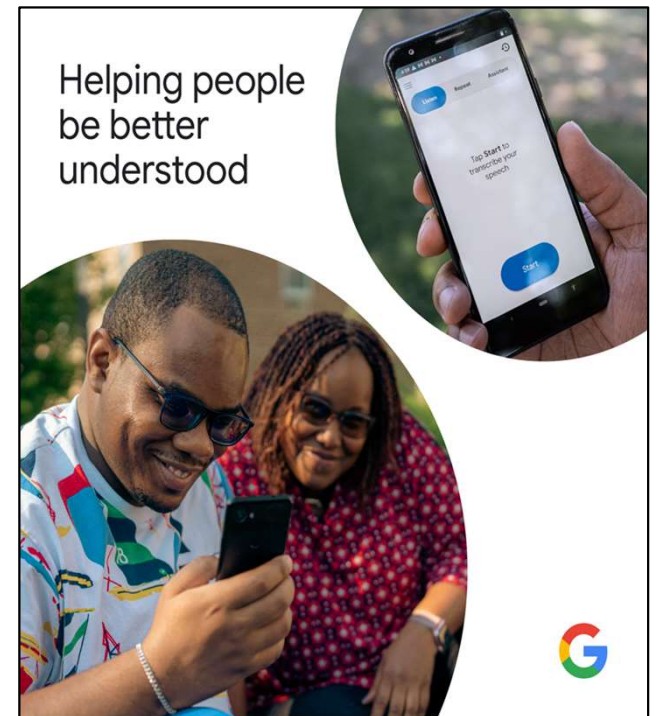
Google Project Relate - RECORD

- Record 500 phrases pre-set

- And/or your own "custom cards"
  (personal words – people, places and things, phrases).

- Skip any phrases, short, long, simpler, high frequency

- Record 6000+ if you wish

- Within a few days of recording 500, a personalized ASR
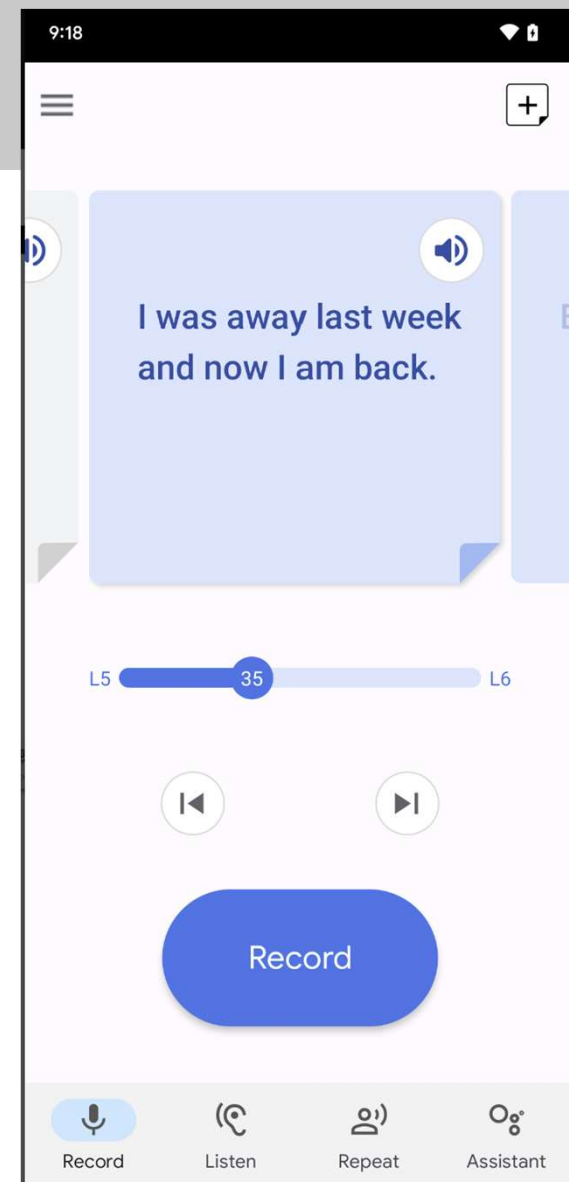
Google Project Relate - LISTEN

- Attempts to transcribes speech to text in real time

- Let people read what you want to tell them

- Enables copy-paste and share text into other apps

- Manual edit too.
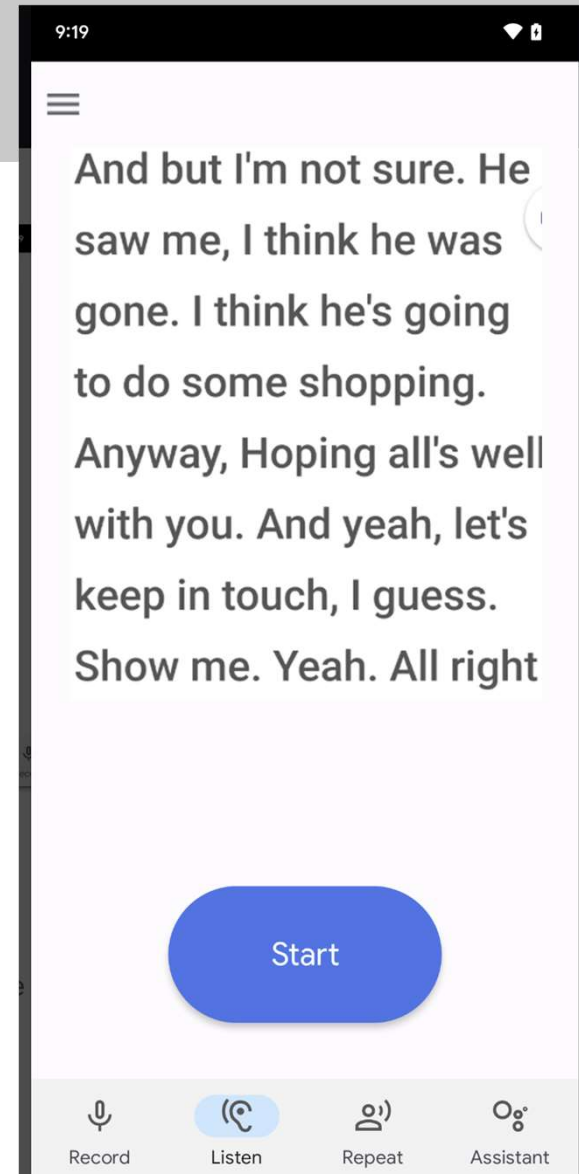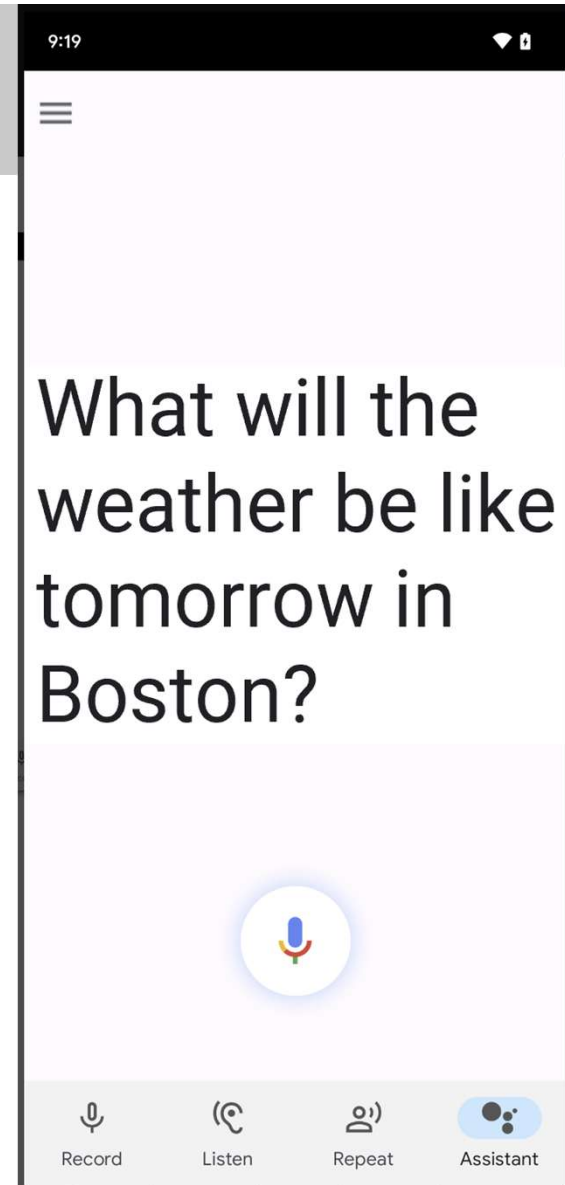
Google Project Relate - REPEAT

- Restate what is said using a clear, synthesized voice.

- Used in face-to-face conversation or

- speak a command to a digital assistant for example 'Turn bedroom lights on'
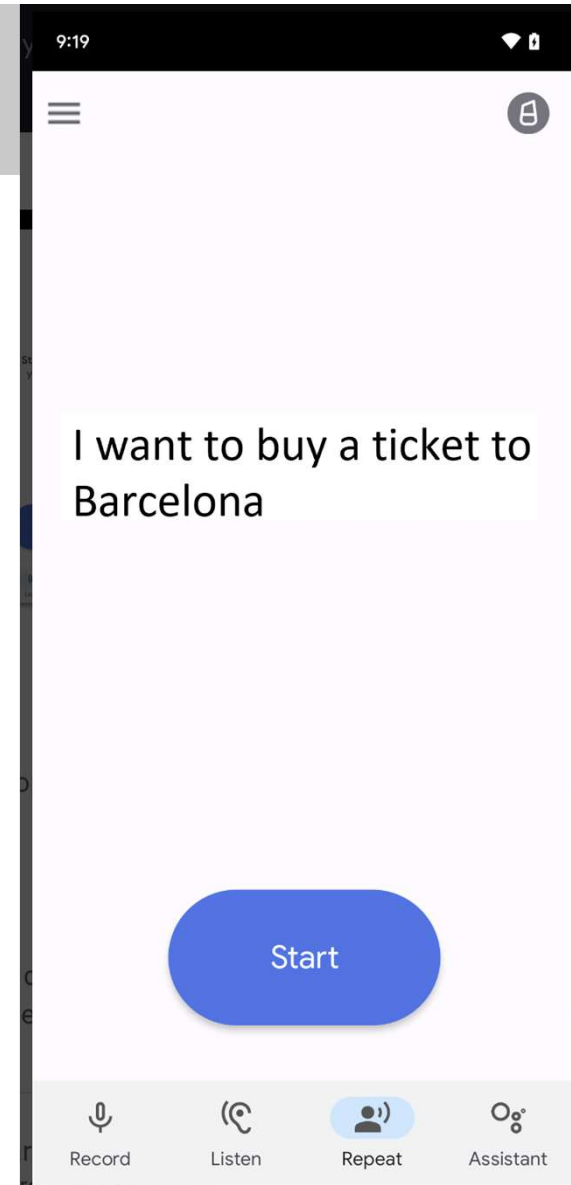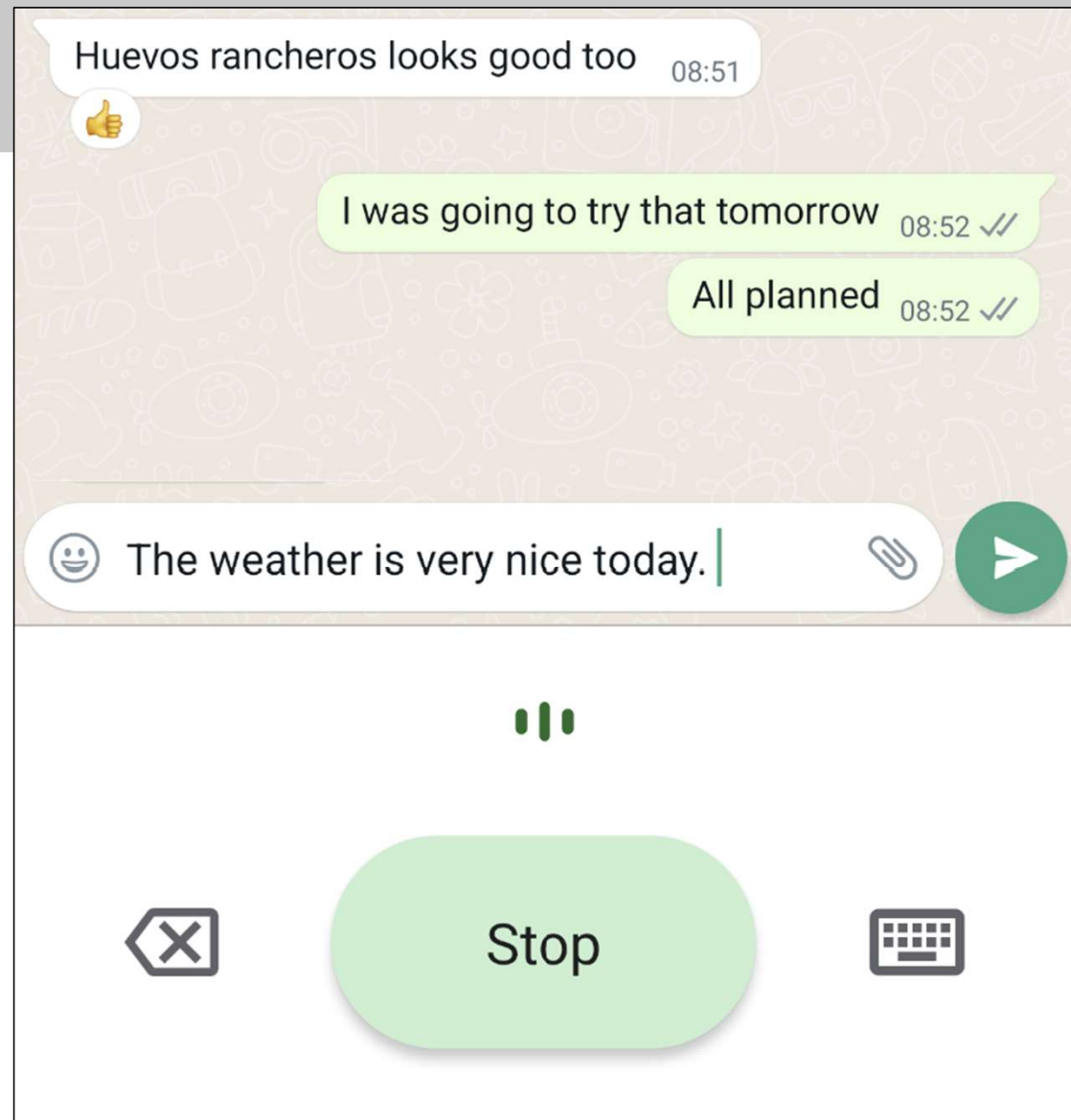
Google Project Relate - ASSISTANT

- Speak directly to Google Assistant from within the Relate app

- Potentially thousands of uses (Laricchia, 2020)

Google Project Relate - KEYBOARD
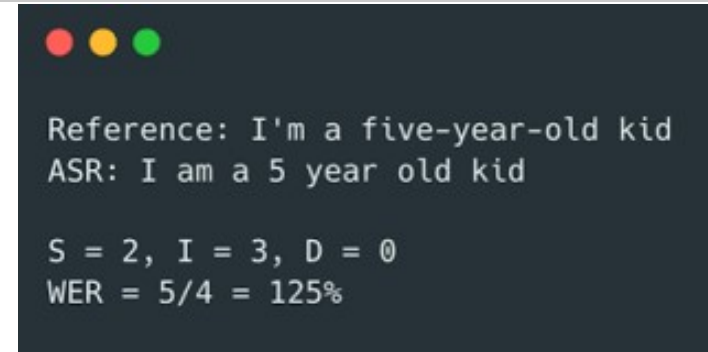
- Use voice wherever there is an app that uses a keyboard

- How plwMND and significant communication partners use Project Relate to support everyday spoken conversation.

- How a progressive decline in speech intelligibility may change the practical use of Project Relate to support spoken conversation
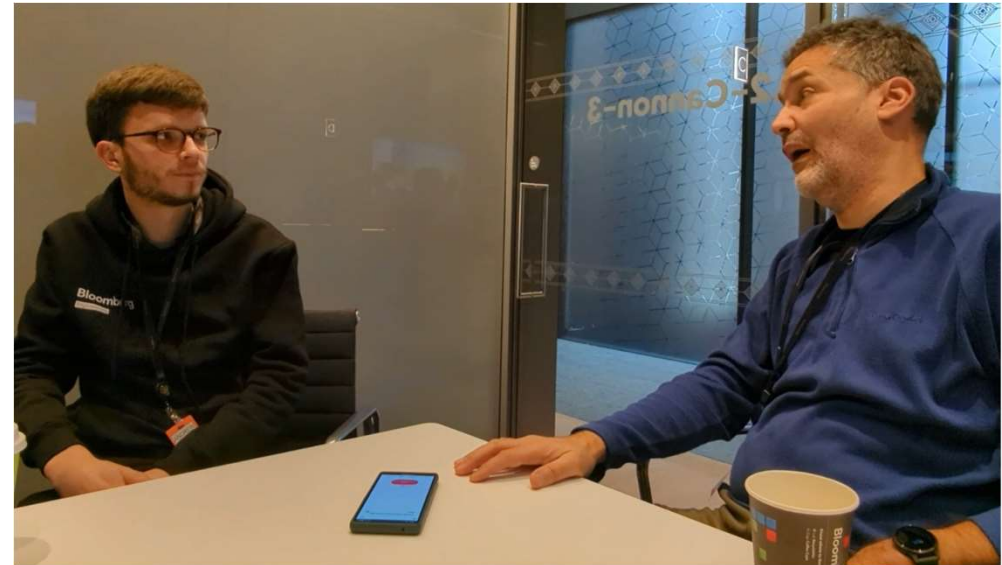
Assessment of ASR accuracy
- Word Error Rate (WER)
- Meaning Preservation

```
Reference: I'm a five-year-old kid
ASR: I am a 5 year old kid

S = 2, I = 3, D = 0
WER = 5/4 = 125%
```

Conversation analysis
- How naturally occurring interaction is structured and organised between speakers
- Based on recordings and detailed transcriptions of spontaneous naturally occurring interaction

Thematic Analysis
- Analysis of semi-structured interviews for some level of patterned response or meaning within the data set (Braun & Clarke, 2006)

Data collection

- Video recordings of three plwMND and conversation partners at home/work

- 8-10 conversations per plwMND (monthly) video-recorded and transcribed

- 26 recordings collected over approx. 12 months

- Frenchay Dysarthria Assessment – 2nd edition (FDA-2) (Enderby & Palmer, 2008)

Results – context, modality, repair

- Specific people, situations, contexts

- Speech was always preferred

- Relate was used in combination with other AAC

- Hierarchy of repair strategies that may reflect predictability of interaction.

Results – accuracy and functionality

- WER was higher than the Relate forecast, probably because conversational not phrase

- Meaning loss of phrases as measured by two SLPs was significantly greater when custom cards not utilised

- However, the functional impact of ASR accuracy is complex

Results – Custom Cards

- plwMND felt Relate was more accurate in everyday conversation when custom cards recorded

- Proper nouns may need extended strategies for repair than with other classes of word (Bloch & Saldert, 2020).

- Analysing 50 proper nouns before and after recording custom cards. 70% meaning loss before, 24% after.
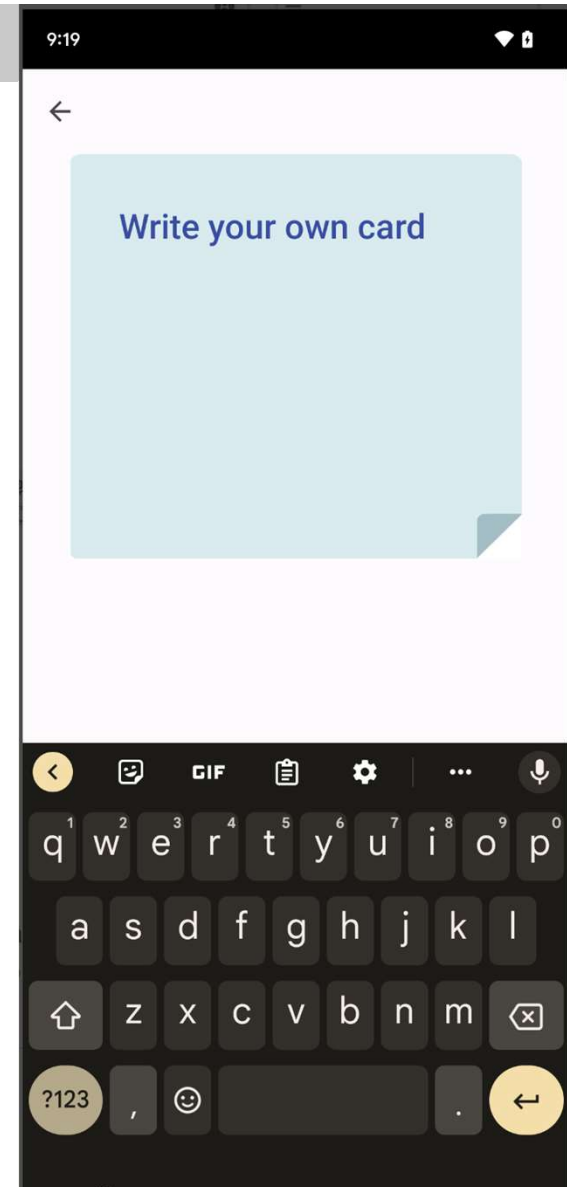
Results – use in interaction

plwMND and communication partners used Relate in different ways.

- One communication partner rarely looked at Relate, 7% on average, another 40%

- One plwMND almost never (1%), another increasingly from 4% to 30% in later sessions as speech declined.

Results - reflection

- All plwMND felt Relate helped in specific circumstances.

- Some communication partners felt they understood plwMND better than Relate.

- Personalisation of ASR appeared critical to perception of usefulness. Custom cards seem a key contributor.

Results - reflection

- Relate ASR supported multi-modality communication, speed of conversation

- Considerations of less control / accuracy vs other AAC

- preferential methods of repair of a misunderstanding

- Accurate proper noun captioning supported contextual understanding
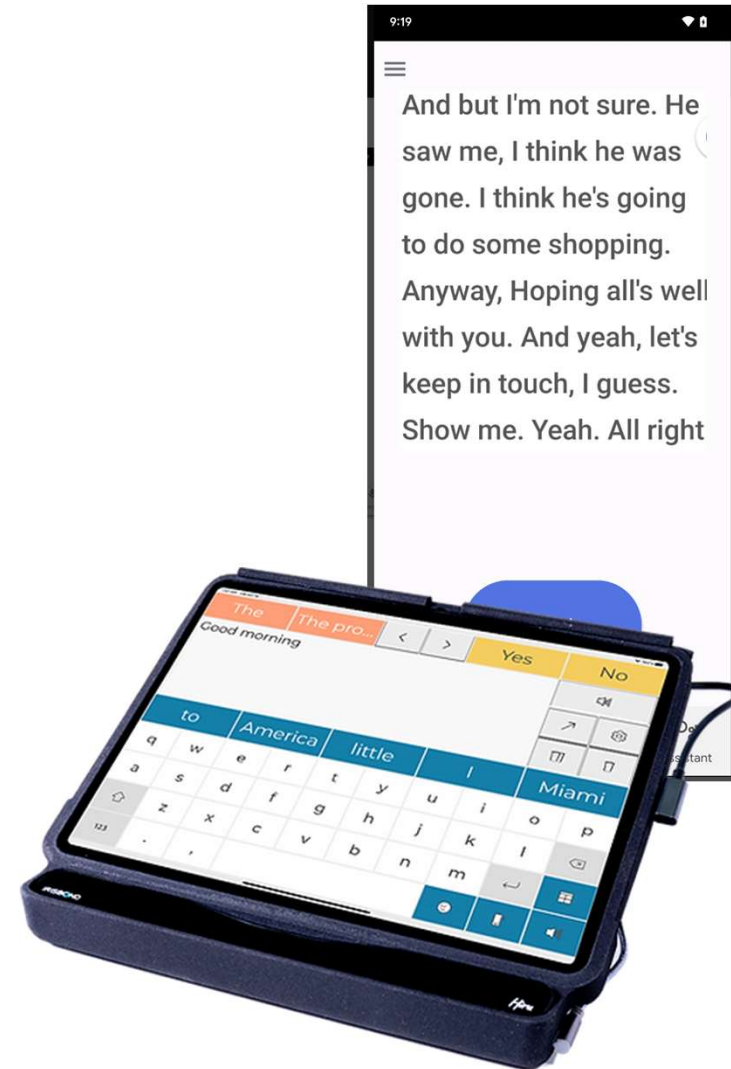
Responsibility of misunderstanding

- Two plwMND said they blame themselves for ASR incorrect captioning

- Attributing fault to speech change and lack of confidence with technology.

- Sense of negative psychological well-being, similar to other research findings

- Communication partners often acted to 'save face' of plwMND

Learnings, next steps and Recommendations

- Expectations Setting. Accuracy, Time to record, AAC and attitudes to AAC.

- Custom Cards.

- Guidelines for communication partners as well as plwMND.

- Recognition that ASR can be used with for different use cases.

- As with all AAC – it can be used anytime with any other AAC and in any context.

- Conversation partner training.

Lyn says hello!

🏛 **UCL**
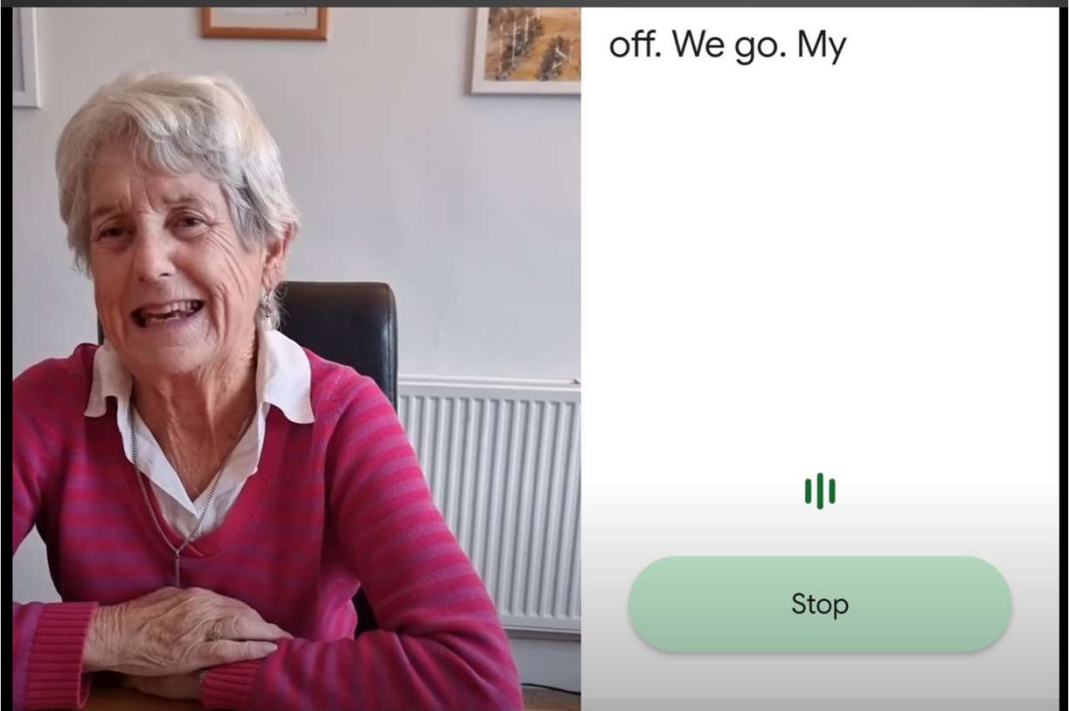
How people living with ALS disease use personalized automated speech recognition technology to support conversation

Richard Cave
Richard.Cave.17@ucl.ac.uk
@SpeedOfTrust27

Language & Cognition, Division of Psychology & Lang Sciences, UCL