



The use of speech recognition technology by people living with amyotrophic lateral sclerosis: a scoping review

Richard Cave & Steven Bloch

To cite this article: Richard Cave & Steven Bloch (2021): The use of speech recognition technology by people living with amyotrophic lateral sclerosis: a scoping review, Disability and Rehabilitation: Assistive Technology, DOI: [10.1080/17483107.2021.1974961](https://doi.org/10.1080/17483107.2021.1974961)

To link to this article: <https://doi.org/10.1080/17483107.2021.1974961>



© 2021 The Author(s). Published by Informa UK Limited, trading as Taylor & Francis Group.



Published online: 11 Sep 2021.



Submit your article to this journal [↗](#)



Article views: 867



View related articles [↗](#)



View Crossmark data [↗](#)



Citing articles: 1 View citing articles [↗](#)

The use of speech recognition technology by people living with amyotrophic lateral sclerosis: a scoping review

Richard Cave  and Steven Bloch 

Language and Cognition, University College London, London, UK

ABSTRACT

Background: More than 80% of people living with Amyotrophic Lateral Sclerosis (plwALS) develop difficulties with their speech, affecting communication, self-identity and quality of life. Automatic speech recognition technology (ASR) is becoming a common way to interact with a broad range of devices, to find information and control the environment.

Background: ASR can be problematic for people with acquired neurogenic motor speech difficulties (dysarthria). Given that the field is rapidly developing, a scoping review is warranted.

Aims: This study undertakes a scoping review on the use of ASR technology by plwALS and identifies research gaps in the existing literature.

Materials and Methods: Electronic databases and relevant grey literature were searched from 1990 to 2020. Eleven research papers and articles were identified that included participants living with ALS using ASR technology. Relevant data were extracted from the included sources, and a narrative summary of the findings presented.

Materials and Methods: Outcomes and Results: Eleven publications used recordings of plwALS to assess word recognition rate (WRR) word error rate (WER) or phoneme error rate (PER) and appropriacy of responses by ASR devices. All were found to be linked to severity of dysarthria and the ASR technology used. One article examined how speech modification may improve ASR accuracy. The final article completed thematic analysis of Amazon.com reviews for the Amazon Echo and plwALS were reported to use ASR devices to control the environment and summon assistance.

Conclusions: There are gaps in the evidence base: understanding expectations of plwALS and how they use ASR technology; how WER/PER/WRR relates to usability; how ASR use changes as ALS progresses.

ARTICLE HISTORY

Received 15 March 2021
Revised 25 August 2021
Accepted 27 August 2021

KEYWORDS

ALS; motor neurone disease; dysarthria; automatic speech recognition; human-computer interaction; environmental control

► IMPLICATIONS FOR REHABILITATION

- Devices that people can interact with using speech are becoming ubiquitous. As movement and mobility are likely to be affected by ALS and progress over time, speech interaction could be very helpful for accessing information and environmental control.
- However, many people living with ALS (plwALS) also have impaired speech (dysarthria) and experience trouble using voice interaction technology because it may not understand them.
- Although advances in automated speech recognition (ASR) technology promise better understanding of dysarthric speech, future research needs to investigate how plwALS use ASR, how accurate it needs to be to be functionally useful, and how useful it may be over time as the disease progresses.

Background

Amyotrophic lateral sclerosis (ALS), also known as motor neurone disease (MND), is a progressive, ultimately fatal disease causing increasing muscular weakness resulting in loss of function of limbs, weakness of muscles of the trunk and neck [1]. Between 80–95% of people living with ALS (plwALS) experience progressive dysarthria and increasing difficulty communicating their daily needs using natural speech [2], and in time, most will be unable to speak at all [3]. Up to 90% of plwALS have been identified as eventually relying on augmentative and alternative communication (AAC) to support daily communication [4]. As speech for

communication becomes increasing difficult, plwALS can adapt the way they speak to compensate, for example, for reduced breath control or articulatory reserve [5].

For many plwALS, there is nothing that can replace the ease or speed of natural speech, citing the extended time it takes to spell a message with AAC [6] and AAC users often use vocalisations and gesture to try to speed up communication [7]. Speech is a powerful medium of identity [8] and communicates mood, humour, geographical, social and educational background, health status, gender as well as the content of the message [9]. It contributes to a sense of self, while allowing the listener to derive multiple levels of meaning [10].

CONTACT Richard Cave  richard.cave.17@ucl.ac.uk

The data that support the findings of this study are available on request from the corresponding author.

This article has been corrected with minor changes. These changes do not impact the academic content of the article.

© 2021 The Author(s). Published by Informa UK Limited, trading as Taylor & Francis Group.

This is an Open Access article distributed under the terms of the Creative Commons Attribution-NonCommercial-NoDerivatives License (<http://creativecommons.org/licenses/by-nc-nd/4.0/>), which permits non-commercial re-use, distribution, and reproduction in any medium, provided the original work is properly cited, and is not altered, transformed, or built upon in any way.

Automatic speech recognition (ASR) technology is the process by which a machine, such as computer or smartphone, receives a speech signal and converts it into a series of words in text form [11]. ASR technology has many uses: it can convert spoken words into text to dictate emails, find information online, from restaurant bookings to directions, weather and news, make appointments and send messages, and provide real-time spoken translations, among many other uses. It can be also used for environmental control, such as changing the channel on a television or control the lighting and heating [12].

ASR technology has rapidly developed over the past three decades. Launched in 1990, Dragon Dictate is regarded as the first consumer speech recognition product [13]. In 1993, “Speakable items”, was introduced as a built-in speech recognition and voice enabled control tool for Apple computers [14]. Microsoft introduced ASR in the Windows operating system in 2007 [15], and In 2011 Apple’s Siri ‘digital personal assistant’ was integrated into the iPhone smart phone [16]. In addition to being able to recognise speech, Siri was claimed to understand the meaning of what was said and take an appropriate action [17]. Amazon launched Alexa in 2014, claimed to be similar in functionality to Siri, as did Microsoft with the Cortana in the same year [18]. “Google Assistant” was launched in 2016 [19]. As at 2019, Amazon’s Alexa digital assistant provided more than one hundred thousand different activities that can be controlled by ASR technology [12].

The STARDUST (Speech Recognition for People with Severe Dysarthria) project developed a speech recognition model for people with severe dysarthria to understand a small vocabulary of ten to twelve individual words that were selected for easier access to environmental control devices [20]. None of the participants were reported to have ALS. Speaker dependent recognition models were trained with the speaker’s own output and each target word could be represented by any phonetic realisation as long as sufficient phonemic contrast was present to distinguish it from any other target word. A computerised training package was also used to build consistency of utterance of each target word production utilising visual and auditory feedback. In this way, STARDUST did not require intelligibility of speech for machine recognition and was limited in the number of words it could recognise. The results showed that participants experienced a significant decrease in the time taken to complete control tasks, such as switching on the TV, and the results of the training package increased computer recognition rates by an average of 5%, though some participants observed no gain from training.

Hosom *et al.* [21] devised Supplemented Speech Recognition (SSR), combining ASR optimised for dysarthric speech, alphabet supplementation, and word prediction. The first letter of the target word was typed and then the whole word spoken. Coupling ASR with first-letter identification decreased the dictionary set from which an ASR system had to match the acoustic signal. The set was further reduced by using word prediction algorithms. SSR put the most likely word into the line of text and provided five further probable word choices in word prediction boxes. The seven participants in the study had mixed or spastic dysarthria with variation in intelligibility ratings from mild to severe. None was reported to have ALS. The results were measured in the number of keystrokes saved compared the total number of keystrokes to type each word fully. Keystroke savings using SSR typical averaged 68%, though the researcher gave assistance to accurately select from the touch-screen computer.

The VIVOCA project [22] combined training of an ASR for a small vocabulary of up to 47 individual words with predictive

message building based on the spoken input. The results were passed to a speech synthesizer, producing intelligible spoken output. To increase the accuracy and consistency of the ASR model, participants were given training to motivate speech consistency and modification where helpful – for example encouragement to replicate their “best attempt so far”. The participants in the study had moderate-to-severe dysarthria and low conversational intelligibility. No participant was reported to have ALS. The study achieved a word recognition accuracy in excess of 85% for the words in the study, and subsequently, the “Vocatempo” app was developed to support accessibility and usability of the VIVOCA concept [23,24].

Voiceitt has a goal to integrate their ASR software for people with “non-standard speech” with smart devices, enabling better access voice recognition tools including Amazon Echo or Google Assistant, and are beta-testing their release [25,26]. Voiceitt are a partner in the Nuvoic project, a European Union funded initiative to develop ASR software to support people with mild dysarthria and ageing voices and better access connected home devices [27]. Google’s Euphonia research project is also focussing on ways to improve ASR accuracy in recognising dysarthric speech [28]. Both projects may present an opportunity for people with dysarthric speech to be better understood by others, supporting the goal of AAC to retain, as far as possible, the speed and, ideally, the naturalness of spoken communication [22].

The potential benefits of ASR technology for plwALS may be significant. Hands-free interaction has been cited as the most popular reason for using ASR devices [29] and for plwALS, this is an important consideration as around two-thirds of plwALS experience limb weakness as an initial and progressive symptom [30]. ASR may help plwALS maintain some control of their home environment as more devices incorporate ASR technology, for example, using voice commands to switch lights on or off, or adjust the thermostat [31], or controlling their wheelchair [32].

For speakers without impaired speech, Google Assistant is estimated to make around five word recognition errors per hundred, equating to a Word Error Rate (WER) of 4.9% [33]. However, ASR devices have been of limited benefit to those with severe speech impairments [34], they are trained from typically non-disordered speech [35]. ASR systems are estimated to have as much as 80-90% WER for people with severe dysarthria, resulting in little or no functional usability [33].

The challenge of ASR better understanding dysarthric speech is a difficult problem to solve, as dysarthric speech has highly variable articulation, tempo, rhythm, and volume [36]. While some levels of ASR accuracy for dysarthric speakers have been reported, they involved speakers who had mild dysarthria or tasks that included very limited vocabulary sets [34]. The speech of plwALS may be particularly difficult for ASR technology to interpret due to progressive dysarthria and the changing strategies plwALS may use to be understood [37]. Fatigue, frequently experienced by plwALS also increases speech variability [38].

Enabling people living with dysarthria to better access voice technologies is timely because voice interaction with devices is becoming ubiquitous. By 2024, it is forecast people will be able to interact with over 8.4 billion devices using their voice – larger than the world’s population, and double the number of devices forecast for 2020 [39]. Although smartphones are the largest group of devices that people will interact with using voice, cars and household devices such as televisions are forecast to have the highest rate of growth [39]. As voice interaction with devices becomes more normalised, individuals with impaired speech

increasingly risk being excluded from these current, commercially available technologies - even though they could benefit extensively [34]. At a societal level, inequality in access to technology risks increasing the “digital divide”, worsening existing divisions and increased social exclusion [40].

Objectives

The field of ASR is rapidly developing in terms of technology and applications and this scoping review will provide a state of the art overview to highlight significant gaps in the evidence [41] and point the way to future research in the field [42].

The objectives of this scoping review are to understand:

- What is known, from the existing literature, about the use of ASR technology by plwALS
- What gaps there are in the existing ASR-ALS evidence base.

Materials and methods

The PRISMA (Preferred Reporting Items for Systematic Reviews and Meta-Analysis) guidelines were utilised for this study [43].

Literature eligibility criteria

All relevant literature regardless of study design was identified, as well as reviews, editorials, and relevant grey literature between 1990 - when the first consumer ASR product, Dragon Dictate was launched [13] and 2020.

Evidence from diverse sources can complement the findings from experimental studies and address questions beyond intervention effectiveness in emerging fields such as ASR [44].

Given the very limited peer-reviewed research in this field, this open approach is essential to capturing and cataloguing the full range of evidence of how plwALS utilize ASR technology. Relevant data were extracted from the included sources, and a narrative summary of the findings created.

For purpose of this investigation, the term ASR is used to refer to voice recognition software that uses natural language input from a speaker and carries out predetermined actions.

Information sources

The following relevant electronic databases were searched in separate stages using identified keywords and indexed terms: PsycINFO; PubMed; Medline; CINAHL; Scopus; LLBA; Cochrane Library; Embase; Web of Science Core Collection; ACM digital Library.

To identify further studies, reference lists were scanned, and the grey literature was also searched. The review considered any existing primary research studies of quantitative, qualitative or mixed-methods design, as well as reviews, editorials, and relevant grey literature in English.

The following sources of grey literature were search: Global Health, HMIC – Health Management Information Centre, OpenDOAR, OpenGrey, PsycEXTRA, Social Care Online, OpenAIRE, Semantic Scholar, BASE: Bielefeld Academic Search Engine, Google Scholar.

Additionally, broader internet searches were undertaken using Google Chrome.

The searches were undertaken in October and November 2020 and the most recent search was completed on 15th November 2020.

Search strategy

The following search terms were used, representing the most commonly used brands of voice assistants using ASR on home-based devices [45], frequently used names and abbreviations for automatic speech recognition technology, known research projects developing ASR technology for dysarthric speech specifically including plwALS, and frequently used names and abbreviations for ALS.

Alexa, ALS, Amazon Transcribe, Amazon AWS, Amyotrophic lateral sclerosis, ASR, Automatic Speech Recognition, Bixby, BlackBerry Assistant, Cortana, Dragon Assistant, Dragon Naturally Speaking, Euphonia, Google Assistant, Intelligent personal assistant, Lou Gehrig’s disease, Microsoft Azure, Microsoft Speech Studio, MND, Motor neuron disease, Motor neurone disease, Siri, Speech Recognition, Voiceitt, virtual assistant, voice-activated environmental control, voice assistants, voice recognition.

The search string was constructed in the following way:

("amyotrophic lateral sclerosis" OR " ALS" OR "motor neuron? disease" OR " MND" OR "Lou Gehrig?") **AND** ("ASR" OR "automat* speech recognition" OR "speech recognition" OR "intelligent personal assistant" OR "voice assistant*" OR "voice recognition" OR "Alexa" OR "Amazon AWS" OR "Amazon Transcribe" OR "Google assistant" OR "Microsoft Speech Studio" OR "Microsoft Azure" OR "Siri" OR "Cortana" OR "Bixby" OR "BlackBerry assistant" OR "virtual assistant*" OR "Dragon assistant" OR "Dragon naturally*" OR "Euphonia" OR "Voiceitt" OR "voice activated environmental control")

The final search results were exported into Mendeley and duplicates were removed.

Selection of sources of evidence

Literature screening was performed by the first author, a UK-registered Speech and Language Therapist (SLT) using the description of the ASR and how it related to plwALS, to understand if the study was eligible according to the criteria of this scoping review. If the SLT could not decide the eligibility, the second author was consulted.

Summarizing the results

The following data extraction proforma has been adapted from the template provided in the Joanna Briggs Institute methodology manual [46] to capture relevant information from grey literature as well as peer-reviewed articles. This proforma was trialled on the eligible studies from the search, and the categories were slightly revised. The charting was completed by the first author and reviewed by the second.

Results

From a total of 177 records identified from electronic databases and grey literature searches, 11 articles met the eligibility criteria. Figure 1 shows the search and selection process.

Table 1 contains the summary data extraction proforma for the 11 identified studies.

Tables 2–4 provide summaries of the data extracted from the 11 studies, comments on their outcomes, and suggestions for future studies.

Hird and Hennessey [47] examined the relative benefit modifying a person’s own speech to improve speech recognition software performance. Fifteen adults with dysarthria associated with a variety of aetiological conditions participated, including two

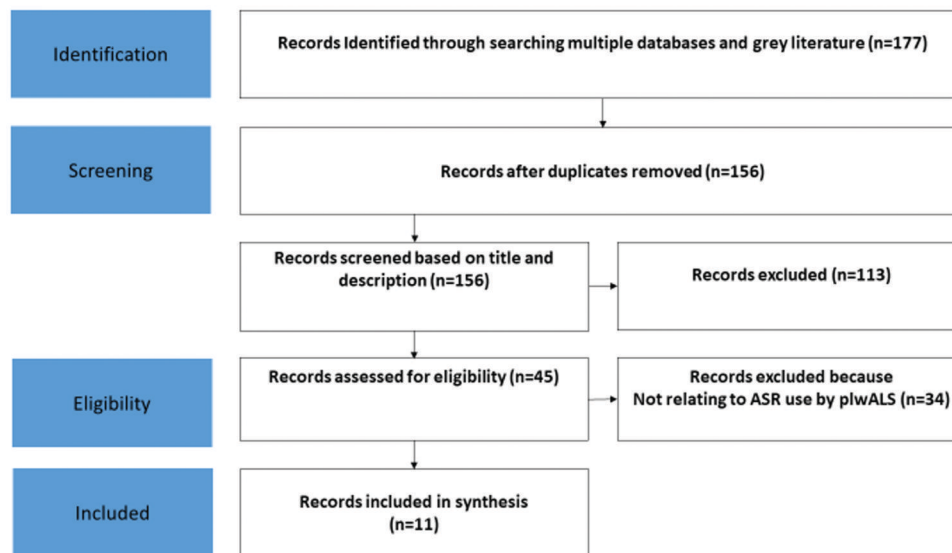


Figure 1. PRISMA flow diagram of record selection.

Table 1. Data extraction proforma.

Publication Details
Citation, country
Study type: primary research, review article, other
Study Objectives
Participant details: ALS type, age/sex, number
ASR Technology used
Study method / design
Details/Results extracted from study
Participant communication needs – dysarthria, AAC
Participant physical and/or cognitive impairment
Outcomes for Communication support and Environmental control
Summary: advantages, disadvantages, and future studies

plwALS. The plwALS were described as having moderate mixed dysarthria and low intelligibility following assessment using the Frenchay Dysarthria Assessment [56]. Type of ALS was not specified. For the group as a whole, overall mean rate of dictation increased from 4.3 words per minute (wpm) to 6.3 wpm as a result of therapy aimed at better regulation of flow of air and volume loss during speech, practice at increasing loudness of speech to mitigate growing environmental noise. The authors concluded that the clinical significance of the treatment effect may be doubtful given the difference in mean dictation rates was small – suggesting modest practical benefit.

Rudzicz [48] identified significant improvements to ASR accuracy for dysarthric speech when theoretical and empirical knowledge of the vocal tract for atypical speech were included in ASR model training, concluding that this approach should benefit assistive software for speakers with dysarthria. The recordings used were mainly from the MOCHA database of recordings of people without dysarthria and the authors own collection of dysarthric speech (it is unclear if any of these recordings were of plwALS). Recordings from the TORGO database were also used. The author highlighted that the assessed recordings were brief and reflected a snapshot in time, and some fundamental aspects of dysarthria such as dysfluency and reduced pitch control could not be reflected in the methods described in the paper.

Rudzicz [49] later described a portable communication device that could potentially increase the intelligibility of speech with dysarthria by removing repeated sounds, inserting deleted sounds, devoicing unvoiced phonemes, and adjusting the tempo

and frequency of speech. Recordings from the TORGO database were used. In the paper, ASR word recognition rate increased from 72.7% to 87.9% relative to the original speech, and the recognition rate of human listeners increased from 21.6% to 41.2%.

Hahm et al. [50] used 316 recordings of short phrases from five plwALS, three females and two males. All had been diagnosed a year or less, and their average speech intelligibility was 95%. There was no indication of what intelligibility assessment was used. Each plwALS produced up to four repetitions of 20 sentences used in daily conversation (for example, “How are you?”) or related to their medical condition (“This is an emergency”). The best reported ASR phoneme error rate (PER) of 31% was a result of combining articulatory and acoustic data with hidden Markov model (HMM) and deep neural network (DNN) technology. The authors acknowledged that the data in the study contained a small number of unique phrases from a small number of plwALS, and further research was needed with a larger dataset.

Kim et al. [51] collected recordings from nine plwALS. Four of the plwALS made recordings on two or more occasions on average six months apart. Speech data were also collected from seven people with normal speech and no history of speech or language impairment. The speech intelligibility of each plwALS was measured by a speech pathologist for each session and varied between normal and severely unintelligible. For every recording session, each plwALS produced up to four repetitions of twenty unique sentences typical of daily conversation (for example “How are you doing?”). Significant variation in speech intelligibility and ASR performance was measured across sessions within individual speakers, likely because of disease progression. For example, one participant declined from a 99% intelligibility rate to 0% and the phoneme error rate (PER) of the ASR increased from 14% to 75%. Factors were detailed that negatively affected ASR accuracy as ALS progressed, including increased dysarthria, long pauses between words, and reduced rate of speech. The authors predicted that as ALS progresses the capability of plwALS to effectively use ASR would reduce. The best ASR performance observed with a PERs of 30.6% for plwALS and the authors argue their technical approach presents a possibility in effectively modelling dysarthric speaker-independent speech. The authors hypothesized that the ASR performance may improve with more recordings from plwALS displaying increased diversity of dysarthria severity.

Table 2. Study types, participants' characteristics, technology, and study method/design.

Publication Details and Characteristics			
Citation	Study Type	Objectives	Participant details (e.g. type of ALS age/sex and number), Technology used, Method
Hird and Hennessey (2007) [47] Australia	Primary Research, peer reviewed	Therapy aimed at modifying a person's speech output to improve speech recognition software performance.	15 adults with dysarthria. Two had ALS. Type of ALS was not specified. One male age 46 and one female age. Power Secretary ASR technology was used
Caves et al. (2007) [48] USA	Conference Proceeding	To improve ASR accuracy for plwALS with mild-severe dysarthria.	30 min of speech therapy treatment followed by 30 min of dictation. 15 sessions overall. 50 plwALS (21 male, 29 Female). No ALS type or age of participant reported. An adapted military training simulation ASR model was used. Participants read digits 0 to 9 and one hundred "zip codes".
Ballati et al. (2018) [49] Italy	Conference Proceeding	Investigated the extent that Siri, Cortana and Google Assistant can comprehend words and sentences spoken by plwALS.	Eight participants (four female, four male), aged 67-83, ALS type not stated. Native Italian speakers. Google Assistant, Cortana, Siri were used. 34 Italian sentences were recorded from 8 people and audio files from each participant were presented to each voice assistant.
Ballati et al. (2018) [50] Italy	Conference Proceeding	Investigate the extent that Siri, Google Assistant, and Amazon Alexa can comprehend words and sentences spoken by people with dysarthric voices and the accuracy and consistency of the answers generated.	Recordings from eight participants from the TORGO database. One participant had been diagnosed with ALS (remainder diagnosed with Cerebral Palsy). Siri, Google Assistant, Amazon Alexa were used. 5 sentences were created from the TORGO dataset pronounced by 7 speakers. Performance of each system was measured by WER and appropriacy of answer.
De Russis and Corno (2019) [23] Italy	Primary Research, peer reviewed	<ol style="list-style-type: none"> (1) Are ASR platforms suitable for recognising dysarthric speech? (2) What is the attained recognition rate? (3) What kinds of transcription errors are more frequent, in case of imperfect/partial recognition? (4) Can transcription alternatives (as provided by ASR platforms) improve the overall recognition result? 	Recordings from eight participants (five males, three females) in English from the TORGO database. One participant had been diagnosed with ALS (remainder diagnosed with Cerebral Palsy). IBM Watson Google Cloud Speech, Microsoft Azure Bing Speech were used. 38 sentences recorded by all females, 13 sentences by all males. Performance measured by WER of each system, and the WER of the transcription alternative.
Pradhan et al (2020) [51] USA	Conference Proceeding	<ol style="list-style-type: none"> (1) To examine the accessibility of off-the-shelf Intelligent personal assistants (IPAs) (e.g. Amazon Echo) (2) To understand how users with disabilities are making use of these devices. 	(1) Analysis of 346 Amazon Echo reviews on Amazon.com that include users with a range of disabilities, including ALS (the number, type, and gender unspecified). Amazon Alexa was used. (2) Semi structured interviews of 16 people with visual impairments (none had ALS).
Shor et al. (2020) [52] USA	Conference Proceeding	(1) To improve ASR accuracy for plwALS with mild-severe dysarthria, and for people with heavy accent.	17 plwALS. No details provided on type of ALS, age or sex. Personalised ASR models were used.
Hahm et al. (2015) [53] USA	Conference Proceeding	(1) speaker independent recognition of dysarthric speech due to ALS using articulatory data.	The participants were given sentences to read from The Cornell Movie-Dialogs Corpus, phrases used by text-to-speech voice actors, and a modified selection of sentences from the Boston Children's Hospital. 5 plwALS. 3 females, 2 males. Average age 59. Recently diagnosed (within a year of the research) No details provided on type of ALS. Gaussian mixture model (GMM) - hidden Markov model (HMM), deep neural network (DNN) - HMM combined with articulatory data were used. plwALS produced up to four repetitions of twenty sentences used in daily conversation (for example, 'How are you?') or related to their medical condition ('This is an emergency').

(continued)

Table 2. Continued.

Publication Details and Characteristics			
Citation	Study Type	Objectives	Participant details (e.g. type of ALS age/sex and number), Technology used, Method
Kim et al. (2018) [54] USA	Conference Proceeding	(1) speaker independent recognition of dysarthric speech due to ALS using spectral and temporal speech data.	9 plwALS. 6 females, 3 males. Average age 62.8. Recently diagnosed (within a year of the research). No details provided on type of ALS. Gaussian mixture model (GMM) - hidden Markov model (HMM), deep neural network (DNN) - HMM combined with articulatory data were used. plwALS produced up to four repetitions of twenty sentences used in daily conversation (for example, 'How are you?') or related to their medical condition ('I need to make an appointment'). 4 plwALS recorded with an average gap of 6 months between recordings.
Rudzicz F. (2011) [25] USA	Conference Proceeding	(1) Improve ASR for dysarthric speech by including articulatory information.	Recordings from MOCHA database of people without dysarthria, the authors own collection of dysarthric speech (not specified if any were of plwALS) and recordings from the TORGO database (One participant had been diagnosed with ALS and remaining seven diagnosed with Cerebral Palsy). Hidden Markov model (HMM), deep neural network (DNN), Support Vector Machines (SVMs), Dynamic Bayes Networks (DBNs) combined with articulatory models were used. Recordings from two speakers with dysarthria (male and female) and two speakers without (male and female). It is unclear if any were plwALS.
Rudzicz F. (2012) [55] UK	Primary Research, peer reviewed	(1) This paper presents a system that transforms the speech signals of speakers with physical speech disabilities into a more intelligible form that can be more easily understood by listeners.	Recordings from the TORGO database (One participant had been diagnosed with ALS and remaining seven diagnosed with Cerebral Palsy). human listeners attempt to identify words in sentence-level utterances under a number of acoustic scenarios. Sentences are either uttered by a speaker with dysarthria, modified from their original source acoustics, or produced by a text-to-speech synthesiser.

Caves et al. [52] adapted ASR technology designed for military training simulation to create four ASR models trained on recordings from fifty plwALS (21 male, 29 Female). The participants were described as having spastic, flaccid, or mixed dysarthria. There was no indication of severity, current communication needs or AAC use. Each participant recorded the digits 0 to 9 and read one hundred zip codes (postal codes). Based on an 11-item digital recognition task, word accuracy was reported as 86% for individuals with mild dysarthria, 70–75% for moderate dysarthria and 51% for severe dysarthria. The word recognition rate for mild and moderate dysarthria models were claimed to be in line with the initial performance of off-the-shelf ASR systems for dysarthria.

Ballati et al. [57] investigated the extent that Siri, Cortana and Google Assistant accurately interpreted 34 sentences in Italian spoken by eight plwALS and the appropriacy and consistency of the generated answers. Three of participants were categorized as having flaccid dysarthria, two spastic dysarthria, two unilateral UMN dysarthria and one mixed dysarthria. All participants were described as having moderate dysarthria, six scored level 3 – detectable speech disturbance on ALS Functional Rating Scale (FRS-r) and two scored level 2 (intelligible with repeating). All were native Italian speakers.

Each participant recorded 34 phrases, selected from a collection of frequently asked questions for Amazon Echo and Google Home, modified to include all the phonemes of the Italian language (for example “When will the next Juventus match be?”). The recordings were then played to Siri, Cortana and Google Assistant.

Overall average word error rates (WER) for Google Assistant, Cortana, and Siri were 24.88%, 39.39% and 70.89% respectively. Correct responses were 63.96%, 39.7%, and 15.81%. Consistency in answers were 54.02%, 24.07%, and 60.47%.

The authors reported that WER is highly dependent on the level of dysarthria of the plwALS. In addition, the type of ASR device used also affected WER. In this study Google Assistant had the highest number of properly transcribed questions and provided a higher number and proportion of correct answers. Siri provided the highest consistency of correct answers across plwALS but for significantly fewer correct answers compared to Google Assistant

They concluded that Google Assistant was the only ASR device usable in functional tasks by people with dysarthria. However, they reported that usability is closely related to each individual's level of dysarthria and level of fatigue. ASR question comprehension was found not to be related to the different types of

Table 3. Participants' communication needs and studies' outcomes.

Details/Results extracted from study		
Citation	Communication needs: dysarthria, AAC Physical or cognitive impairments.	Outcomes for communication support / environmental control
Hird and Hennessey (2007) [47] Australia	ALS participants had mod mixed UMN and LMN dysarthria, and low or intelligibility. AAC needs not detailed. Any physical or cognitive impairments were not specified.	The overall mean rate of dictation increased from 4.3 wpm to 6.3 wpm.
Caves et al. (2007) [48] USA	The participants were reported to have including spastic, flaccid, or mixed dysarthria. There was no grouping of severity, indication of current communication needs, or AAC use. Any physical or cognitive impairments were not specified.	Based on an eleven-item digital recognition task, word accuracy was 86% for individuals with mild dysarthria, 70-75% for moderate dysarthria and 51% for severe dysarthria.
Ballati et al. (2018) [49] Italy	Six participants had detectable speech disturbance, and two were intelligible with repeating. Three participants had flaccid dysarthria, two spastic dysarthria, two unilateral UMN dysarthria and one mixed dysarthria. Any physical or cognitive impairments were not specified.	Overall average word error rates (WER) for Google Assistant, Cortana, and Siri were 24.88%, 39.39% and 70.89% respectively. WER dependent on the level of dysarthria and ASR type
Ballati et al. (2018) [50] Italy	No details provided about the single plwALS and unable to identify which of the seven participants the plwALS is. AAC needs not detailed Any physical or cognitive impairments were not specified.	Overall average (all participants) WER for Siri 69.41%, Google Assistant 15.38%, Amazon Alexa could not be measured. Overall average appropriate answers from Siri 58.82%, Google Assistant 64.7%, Amazon Alexa 41.17%.
De Russis and Corno (2019) [23] Italy	Six participants had detectable speech disturbance, and two were intelligible with repeating. Any physical or cognitive impairments were not specified.	Google Cloud Speech average WER 59.81%, Microsoft Azure Bing Speech 62.94% and IBM Watson 67.35%. Using the most accurate transcription alternatives improves WER though they were manually selected based on knowledge of the TORGO transcriptions.
Pradhan et al (2020) [51] USA	Communication / AAC needs were not specified Any physical or cognitive impairments were not specified.	The study did not detail which participants had ALS. Users with impaired speech were generally positive about their experience with ASR technology. One plwALS able to control lights / thermostat.
Shor et al. (2020) [52] USA	Three participants had flaccid dysarthria, two spastic dysarthria, two unilateral UMN dysarthria and one mixed dysarthria. Any physical or cognitive impairments were not specified.	The overall WER for ALS speakers, on a test set of message bank phrases (not disclosed), was reduced from 33% to 11% for mild dysarthria and from 60% to 20% for moderate-severe dysarthria.
Hahm et al. (2015) [53] USA	Severity of dysarthria was mild with average speech intelligibility of 95%. No indication of assessment used or who conducted the assessment. Any physical or cognitive impairments were not specified.	The best reported phoneme error rate (PER) 31%, reduced from 46% baseline.
Kim et al. (2018) [54] USA	Severity of dysarthria was mild to profound with average speech intelligibility ranging from 100% to 0%. Assessed by a speech pathologist. Any physical or cognitive impairments were not specified.	Significant variation in speech intelligibility and ASR performance was measured across sessions likely due to ALS progression. Factors that may have negatively affected ASR accuracy included increased dysarthria, long pauses between words, and reduced rate of speech.
Rudzicz F. (2011) [25] USA	Level of dysarthria was not described Any physical or cognitive impairments were not specified.	Significant improvements to ASR accuracy for dysarthric speech when theoretical and empirical knowledge of the vocal tract for atypical speech were included in ASR model training.
Rudzicz F. (2012) [55] UK	Level of dysarthria was not described Any physical or cognitive impairments were not specified.	Human listener recognition rates increased from 21.6% to 41.2%) and ASR word recognition rate increased from 73% to 88%) relative to the original speech. Potential gains to EC derived from increased functional use of ASR device.

dysarthria. The authors observed that when there was an interval of silence due to the effort of speaking, Google Assistant frequently stopped "listening", missing or misinterpreting the question.

The authors acknowledged some factors that may have affected the study outcomes: the phrases were pre-recorded, did not include the call phrases, for example "Hey Google", and the recordings were replayed from a speaker to the ASR devices.

In a related study, **Ballati et al.** [53] used recordings from seven participants from the University of Toronto's database of acoustic and articulatory speech from speakers with dysarthria

(TORGO). The recordings were used to examine how Siri, Google Assistant, and Amazon Alexa comprehends words and sentences spoken by people with dysarthric voices and the appropriacy of the generated answers.

One participant had been diagnosed with ALS and the remainder diagnosed with cerebral palsy. No details were provided about the type of ALS or type and severity of dysarthria.

Five sentences were created from the TORGO dataset pronounced by seven speakers. The test phrases were limited by available recordings from TORGO that could be adapted to be possible requests to ASR devices, for example "Please open the

Table 4. Studies' related advantages and disadvantages and suggestions for future studies.

Summary			
Citation	Advantages	Disadvantages	Future Studies
Hird and Hennessey (2007) [47] Australia	A marginal increase in WPM was observed.	The authors concluded that the clinical significance of the treatment effect was doubtful.	Further research using a between groups design and control group with no concurrent treatment may show material gains in WPM from clinical intervention.
Caves et al. (2007) [48] USA	word recognition rate for mild/mod dysarthria was claimed to be in line with off-the-shelf ASR systems for dysarthria.	A small phraseset was tested, potentially of limited functional use.	Further research to assess the ASR model with a larger dysarthric speech dataset.
Ballati et al. (2018) [49] Italy	Google Assistant was the only ASR argued to be usable in functional tasks by people with dysarthria.	ASR useability related to dysarthria severity and fatigue. Few subjects involved in the experiment. Sentences played from recordings - not reflecting variation in speech.	Further research to evaluate a larger number of phrases. Compare the same speech samples spoken by people without dysarthria as a control. Research accessibility of ASR devices.
Ballati et al. (2018) [50] Italy	Siri, Google Assistant and Alexa have comparable performance.	Limited recordings assessed. The recordings were adapted from phrases not used with ASR.	Further research to evaluate more sentences and participants. Assess usefulness or ASR for different severities of dysarthria, environmental control and accessibility of ASR devices.
De Russis and Corno (2019) [23] Italy	Google Cloud Speech had the lowest average WER of 59.81%.	Even the lowest WER assessed was reported as significantly worse than the average WER of speakers without dysarthria.	Future work to identification of an algorithm able to improve the selection of accurate transcription alternatives, without reliance on a priori knowledge of the original text sentence.
Pradhan et al (2020) [51] USA	Potential for increased autonomy and reduced burden on caregivers.	Very limited data for plwALS use of IPAs in the study.	Further research to explore the use of IPAs for speech therapy, memory aids, and improved design of feature rich voice-based interfaces and apps.
Shor et al. (2020) [52] USA	Significantly lower WER is argued to enable plwALS to more successfully interact with smart home devices	WER remains significantly higher than for speakers without dysarthria.	The study identified common speech patterns and suggested research for early ALS detection techniques. Also research for building ASR models with limited data of dysarthric speech.
Hahm et al. (2015) [53] USA	Significantly lower PER was achieved by articulatory data to the ASR technology.	The data set used in the experiment contained only a small number of unique phrases collected from a small number of ALS patients.	Further studies with a larger vocabulary from more ALS patients are necessary to explore the limits of the proposed approach.
Kim et al. (2018) [54] USA	Best ASR performance observed a PERs of 30.6% for plwALS. Authors argue their approach may effectively modelling dysarthric speaker-independent speech.	The data set used in the experiment contained only a small number of unique phrases collected from a small number of ALS patients.	Further studies with a larger vocabulary from more ALS patients are necessary to explore the limits of the proposed approach.
Rudzicz F. (2011) [25] USA	The authors reported that improvements in ASR accuracy are possible through discriminative methods such as latent-dynamic conditional random fields.	The author highlighted the recordings were brief and a snapshot in time, and some fundamental aspects of dysarthria could not be reflected in the methods described in the paper.	Further studies with a larger vocabulary from more ALS patients are necessary to explore the limits of the proposed approach.
Rudzicz F. (2012) [55] UK	Potentially significant improvement to ASR functional usefulness for people with dysarthria.	Not tested on a broad range of people living with dysarthria.	Further studies with a larger vocabulary from more ALS patients.

window quickly". The participants did not all use the same phrases or words due to the limitations of the TORGO recordings. The recordings made by each participant were not detailed.

The average WER for all participants was for Siri 69.41% and Google Assistant 15.38%. The WER for Amazon Alexa could not be measured as Alexa does not show a transcription. The overall average for appropriateness of answers was for Siri 58.82%, Google Assistant 64.7%, and Amazon Alexa 41.17%.

The authors concluded the three ASR devices assistants have comparable performance with a recognition percentage of around 50–60%. They recommended future research to evaluate a larger number of sentences pronounced by people with dysarthria and a detailed profile of type and severity of dysarthria for each participant.

De Russis and Corno [33] examined the WER of IBM Watson Speech-to-Text, Google Cloud Speech and Microsoft Azure Bing Speech ASR platforms for recognising dysarthric speech. They assessed the most frequent transcription errors and if the transcription alternatives provided by each ASR platforms could be used to improve the overall recognition result. Recordings from the TORGO database were used from eight participants. One participant had been diagnosed with ALS and the remainder were diagnosed with cerebral palsy. No details were provided about which participants had ALS and which had cerebral palsy. Thirty-eight specific sentences were used and had been recorded by all five males and 13 sentences recorded by all three females. The speech intelligibility of the eight speakers with dysarthria ranged from "no abnormalities" (for two females and one male) to

“severely distorted” (the remaining speakers, four males and one female).

The average WER for all participants for Google Cloud Speech was 59.81%, Microsoft Azure Bing Speech 62.94% and IBM Watson 67.35%. The most frequent transcription errors were “one or many” word errors with undetermined impact on semantic meaning. Using the most accurate transcription alternatives significantly improves WER (Google 5%, IBM 9%), although the alternatives were selected by the authors based on prior knowledge of the TORGO transcriptions.

The semantic impact of the ‘single or multiple’ word errors was reported as having little impact on understanding of the sentence, though the authors acknowledged the TORGO database phrases lacked ecological validity.

Pradhan et al. [54] examined how users with disabilities use the Amazon Echo (a device that uses ASR). The authors completed thematic analysis of 346 Amazon Echo reviews on Amazon.com that include reference to users with a range of disabilities. Although the study did not detail which participants had ALS there were two quoted reviews in the paper that related to plwALS. They also conducted semi structured interviews of 16 people with visual impairments (none had ALS). The summary results indicated that users with impaired speech were generally positive about their experience with the Amazon Echo. Where the study did attribute a quotation from plwALS, one highlighted the ability of the device to control smart home appliances such as lights or thermostats in order to reduce caregiver burden. Another indicated Alexa was used to alert others that help was needed by switching the lights off and on in another room.

The authors acknowledged that as two-thirds of the reviews were written by third-parties (significant others, carers for example) they may not be as accurate as first-person reviews in reflecting the experience of users with disabilities. The authors also hypothesized a potential sampling bias: users with more severe speech or physical impairments may not have thought to try the device and thus to write a review.

Shor et al. [55] completed research on how to improve ASR accuracy for plwALS with mild-severe dysarthria, and for people with heavy accent. Seventeen plwALS completed a combined total of 22 hours of audio recordings. The participants were given sentences to read from The Cornell Movie-Dialogs Corpus, phrases used by text-to-speech voice actors, and a modified selection of sentences suggested for message banking from the Boston Children’s Hospital. The plwALS had FRS-r speech scores between 1–3, categorized in the paper as having mild–severe dysarthria. No details were provided on an individual basis.

Using a test set of recorded phrases suggested for message banking (the list was not disclosed), overall WER was reduced from 33% to 11% for mild dysarthria (FRS-r score 3) and from 60% to 20% for moderate-severe dysarthria (FRS-r score 1–2). The study also identified patterns in phoneme deletion, insertion and substitution in the speech recorded by the group of plwALS and suggest further research might lead to early ALS-detection techniques.

Discussion

The objectives of this scoping review were to establish what is known, from the existing literature, about the use of ASR technology by plwALS, and identify gaps the existing ASR-ALS evidence base

What do plwALS use ASR technology for?

As only four studies identified which participants had ALS, and which outcomes referred to them [50, 52–54], it is difficult to ascertain which of the outcomes related to plwALS in the majority of the studies.

Pradhan et al. [54] investigated how useful the Alexa Echo may be to people with disability by assessing the contents of Amazon reviews but it is difficult to draw any conclusions for plwALS as the authors did not indicate which of the participants had ALS beyond citing two quotes. They added that many of the reviews were written by a third-party or product adverts which may impact on the validity of the review. The Amazon Echo is one of many devices that use ASR and this study design excluded the opinions of users of other devices, and of those that choose not to use ASR. Only those that had provided written feedback to Amazon about the Amazon Echo were in this study, which is likely to be a sub-section of the people who are using or have tried to use an ASR, given the challenges of plwALS progressive movement decline.

Ballati et al. [57] presented participant plwALS with a list of frequently used commands for Google and Alexa and asked if they would be useful in everyday life. The plwALS indicated that they generally would. However, this is not the same as understanding the ASR commands that plwALS actually use and how they use this technology. As an example, some plwALS check how much Alexa recognises their speech over time as a way of monitoring ALS progression [58]. **Kim et al.** [51] recorded samples of everyday conversation, phrases that may not bear a relation to how plwALS actually use ASR devices.

Two of the studies utilized recordings from the TORGO database in their research [33, 53]. However, as TORGO has no recordings of people interacting with ASR devices, samples were edited to approximately resemble ASR commands. These commands are potentially going to be different from those that plwALS actually use. **Caves et al.** [52] used recordings of zip codes and number repetition, and it is difficult to interpret potential functional use of ASR from this dataset.

Without a clear understanding of what plwALS/ALS use ASR technology for, it is difficult to measure how successful ASR could be at meeting their needs. There remains a significant gap in understanding how effective ASR technology is for plwALS in a functional setting.

What communication, physical and cognitive needs do participants have?

Only four studies described the type of dysarthria, and only four detailed the assessments used [25, 36, 50, 55]. Two described the participant’s level of communication effectiveness as measured by intelligibility [53,54], and none detailed any physical or cognitive impairment. Without an understanding of the communication, physical and cognitive needs of the participant plwALS, it is difficult to interpret research outcomes.

A recent study showed that ASR technology is used more frequently on smartphones and mostly for very simple tasks [59]. There has been no identified research on what platforms plwALS use ASR technology on, an important consideration for physical, visual and auditory accessibility.

How does ASR use and usefulness change as ALS progresses?

Only one study assessed how ASR recognition error rates may change as ALS progresses (for some but not all participants) [51].

However, no assessment was made of the change in usefulness or perceived benefit to plwALS as ASR error rates increased. Eight of the 11 studies used recordings of speech to assess ASR recognition. Although using recordings provides a more consistent basis for comparison across ASR devices, this approach does not reflect how the speech of plwALS may change or deteriorate as the condition progresses, even during the course of a day, due to fatigue for example [60]. Physical and psychological fatigue can affect the voice, mind, and body and are known to cause degradation in ASR performance [11]. Research is needed to understand the changes in use and benefit of ASR for plwALS as their condition progresses.

Is WER a useful measure for functional use of ASR for plwALS?

Five of the 11 studies used WER (word error rate) as a measure of success of ASR, two studies used PER (phoneme error rate) [51] and two WRR (word recognition rate), but no study examined how these measures relate to the usability of ASR devices for plwALS.

There has been no research to assess what level of WER, PER or WRR is acceptable to plwALS using ASR. No study assessed if a change in these measures was significant enough for plwALS to use ASR devices any more or less. ASR devices may still be useful to people with disability even with a high WER where the alternative method to complete a task is perceived to be slower [61]. Ballati points out that even with a relatively high WER, ASRs may leverage the context or some specific recognized keywords to provide a suitable response [53]. This supports the case for assessing the usefulness of ASRs by measuring the appropriacy of response to the commands plwALS would use. If the ASR responses are appropriate, it may be that the level of WER is less significant.

Research is needed on which device may provide the highest rate of appropriate responses for the commands used by plwALS. This is because of the reported variability in the appropriacy of responses between ASR technology providers, even for speakers with unimpaired speech. Google and Alexa scored significantly higher than Siri and Cortana in a study assessing the correctness of answers and how natural the responses feel to users [62].

No identified ASR technology has a zero WER even for speakers with unimpaired speech, for example the WER for English speakers with unimpaired speech in an optimal environment using Google Assistant was estimated at 4.9% [63]. However, environmental noise, unfamiliar words (to Google Assistant), and accent are all likely to affect WER negatively [64]. In addition, WER can vary significantly for an individual and between individuals due to the type and severity of dysarthria, levels of fatigue, environmental noise as well as a speaker's accent [35].

A speech recognition rate between 90–95% is considered satisfactory for people with non-disordered [65,66] and 80% speech recognition rate has been argued to be satisfactory for dysarthric speech [67], but more research is needed to understand if this level applies plwALS in daily life, and Ballati argues that the usability of an ASR device is strictly related to how an individual's level of dysarthria and the effect on WER and the appropriateness of a response [57]. It is likely that additional factors should be considered to measure ASR usability for plwALS, beyond WER/PER/WRR, as individuals may balance the increased speed and ease of use ASR against the frustration arising from lower accuracy when deciding whether to use ASR device [31].

Could changing how plwALS speak support functional use of ASR technology?

Hird and Hennessey [47] identified small gains in dictation speed using voice recognition technology as a result of therapy aimed at better regulation of flow of air and volume loss during speech, and practice at increasing loudness of speech to mitigate growing environmental noise. However, it is difficult to assess how this overall result applied to the plwALS in the study as individual results were not published. The paper did not explore the participant's own views on the relative functional benefits of the increase in speed, or assess how speech recognition may change as the disease progresses. Some researchers argue that there are functional benefits by training a speaker to alter speech production to improve speech recognition of dysarthric speech [22, 68]. No identified research was found to measure the usefulness of this approach using more recent and popular ASR devices.

This prompts the wider question of how far should a person be obliged to change their own behaviour to mitigate poor design that lacks inclusiveness [69]. More research is indicated to understand how acceptable it is to plwALS to need to adjust their own speech to bring about improved ASR performance.

Could ASR technology support diagnosis and ongoing management of ALS?

Research into automatic detection of ALS from speech samples and the feasibility of detection even in pre-symptomatic stages is ongoing [70–72]. Connaghan et al. [73] used a smartphone app to remotely identify and track speech decline in ALS. Shor et al. [35] identified patterns in the speech of plwALS relating to phoneme deletion, insertion and substitution, and suggest further research might lead to early ALS-detection techniques using ASR technology. Stegmann et al. [74] reported detection of early speech changes and track speech progression in ALS via automated algorithmic assessment of speech collected digitally. This application of ASR is also reflected in research with patients with Parkinson's disease, where on-the-shelf ASR devices are used to measure changes in speech intelligibility to estimate neurological state [75].

Conclusions

The objectives of this scoping review are to understand what is known, from the existing literature, about the use of ASR technology by people with ALS, and what the gaps are in the existing ASR-ALS evidence base.

Eleven articles were identified. Eight used recordings of plwALS and people with other aetiologies to assess WER/PER/WRR and appropriacy of responses by different ASR devices. These measures were found to be linked to severity of dysarthria and the ASR technology used. One article examined how strategies to modifying speech output may improve speech recognition accuracy and concluded the clinical gain was small. The final article completed thematic analysis on Amazon.com for how users with disabilities use the Amazon Echo and two plwALS were reported to use ASR devices to control the environment and to summon carer assistance.

Significant gaps in the research evidence base were identified. There is a paucity of research to understand how effective ASR technology is for plwALS in a functional setting, and what measures are most appropriate to measure benefits. Most studies used WER, PER or WRR as a measure of success of ASR but no study examined how these measures relate to the usability of ASR

devices for plwALS. Similar gaps in the evidence base were also identified when reviewing interventions using high technology communication aids generally [76].

No studies investigated the commands plwALS use with an ASR device. This is important because ASR technology utilises the context of key words to assist with accurate interpretation of commands, therefore WER/PER and the rate of appropriate responses may differ based on the key words in a command.

All but one of the identified studies focussed on a single snapshot in time, and research is needed to understand how the functional use and usefulness of ASRs changes as ALS progresses and the severity of dysarthria changes.

No study examined the views and expectations of plwALS themselves around ASR technology, how they use ASR devices in daily life, how effectively their needs are met, and when they choose not to use it. It is important to understand the expectation of plwALS from ASR technology. Further work is needed exploring how effective the technology is in aiding communication in a functional setting.

Much of the identified research did not detail which participants had ALS or another condition, and it was often not possible to conclude which research outcomes applied to plwALS. No paper provided a detailed profile of the individual participant's type of ALS, rate of progression, dysarthria type and severity.

More research is needed to understand the accessibility and usability of ASR devices for plwALS. No paper discussed the participants level of physical or cognitive needs.

One study examined the potential benefits of plwALS adjusting their own speech delivery to improve speech recognition. Research is needed to understand if this approach can be functionally beneficial, or if the effort of being forced to change speech delivery outweighs any gain.

Most studies used recordings of speech to assess ASR performance, not reflecting how natural voice may vary over the course of a day and over the course of the disease progression.

Implications for research

To better understand how plwALS use ASR technology the following research areas are proposed:

- To understand how effective ASR technology is for plwALS over time, in a functional setting.
- How the use and usefulness of ASR technology may change as ALS progresses.
- The attitudes and expectations of plwALS towards ASR technology, and when where and how they would use it in daily life or would choose not to.
- The interactions plwALS have with ASR devices, and the commands they use. How ASR devices respond to variations in rate of speech, the number and length of pauses, speech volume.
- The accessibility and usability of the different ASR devices for plwALS.
- How WER/PER/WRR relates to the usability of ASR devices for plwALS.
- The potential for ASR technology to support early diagnosis and ongoing management of ALS.

Disclosure statement

Richard Cave is funded to provide advice as a speech and language therapist to the Google Euphonia project.

Steven Bloch has no potential conflict of interest.

ORCID

Richard Cave  <http://orcid.org/0000-0002-6410-8200>

Steven Bloch  <http://orcid.org/0000-0002-5355-8134>

References

- [1] Oliver D, Radunovic A, Allen A, et al. The development of the UK national institute of health and care excellence evidence-based clinical guidelines on motor neurone disease. *Amyotroph Lateral Scler Frontotemporal Degener.* 2017; 18(5-6):313–323.
- [2] Beukelman D, Fager S, Nordness A. Communication support for people with ALS. *Neurol res int.* 2011;2011:714693.
- [3] Beukelman DR, Garrett KL, Yorkston KM. *Augmentative communication strategies for adults with acute or chronic medical conditions.* Baltimore; London: Paul H. Brookes Pub. Co; 2007.
- [4] Ball LJ, Beukelman DR, Pattee GL. Communication effectiveness of individuals with amyotrophic lateral sclerosis. *J Commun Disord.* 2004;37(3):197–215.
- [5] Bloch S, Beeke S. Co-constructed talk in the conversations of people with dysarthria and aphasia. *Clin Linguist Phon.* 2008;22(12):974–990.
- [6] Murphy J. I prefer contact this close": perceptions of AAC by people with motor neurone disease and their communication partners. *AAC Augment Altern Commun.* 2004;20(4): 259–271.
- [7] Higginbotham J, Fulcher K, Seale J. *The Silent Partner? Language, Interaction and Aided Communication,* J&R Press [Internet]. 2016. [cited 2020 Nov 1]. Available from: <https://www.jr-press.co.uk/the-silent-partner.html>.
- [8] Bucholtz M, Hall K. Identity and interaction: a sociocultural linguistic approach. *Discourse stud.* 2005;7:585–614. Available from: <https://www.jstor.org/stable/24048525sociocu>.
- [9] Nathanson E. Native voice, self-concept and the moral case for personalized voice technology technology. *Disabil Rehabil.* 2016;8288:73–81.
- [10] Wertsch JV, Wertsch JV. *Voices of the mind: a sociocultural approach to mediated action.* London: Harvester Wheatsheaf; 1991.
- [11] Young V, Mihailidis A. Difficulties in automatic speech recognition of dysarthric speakers and implications for speech-based applications used by the elderly: a literature review. *Assist Technol.* 2010;22(2):99–112.
- [12] Statista Research. • Amazon Alexa: skill growth 2016-2019 | Statista [Internet]. 2020. [cited 2020 Jul 13]. Available from: <https://www.statista.com/statistics/912856/amazon-alexa-skills-growth/>.
- [13] Pinola M. Speech Recognition Through the Decades: How We Ended Up With Siri | PCWorld [Internet]. 2011. [cited 2020 Jul 12]. Available from: https://www.pcworld.com/article/243060/speech_recognition_through_the_decades_how_we_ended_up_with_siri.html?page=2.
- [14] Neuburg M. Bossing Your Mac with PlainTalk - TidBITS [Internet]. 2000. [cited 2020 Jul 12]. Available from: <https://tidbits.com/2000/08/28/bossing-your-mac-with-plaintalk/>.
- [15] Shinder D. Speech recognition in Windows Vista - TechRepublic [Internet]. 2007. [cited 2020 Jul 12]. Available from: <https://www.techrepublic.com/article/speech-recognition-in-windows-vista/>.
- [16] Golson J. Siri Voice Recognition Arrives On the iPhone 4S - MacRumors [Internet]. 2011. [cited 2020 Jul 12]. Available

- from: <https://www.macrumors.com/2011/10/04/siri-voice-recognition-arrives-on-the-iphone-4s/>.
- [17] Daw D. What Makes Siri Special? | PCWorld [Internet]. 2011. [cited 2020 Jul 12]. Available from: https://www.pcworld.com/article/242479/what_makes_siri_special_.html.
- [18] Warren T. The story of Cortana, Microsoft's Siri killer | The Verge [Internet]. 2014. [cited 2020 Jul 12]. Available from: <https://www.theverge.com/2014/4/2/5570866/cortana-windows-phone-8-1-digital-assistant>.
- [19] Wallace L. The future is AI, and Google just showed Apple how it's done [Internet]. 2016. [cited 2020 Jul 12]. Available from: <https://www.cultofmac.com/447898/google-home-google-assistant-siri-ai/>.
- [20] Parker M, Cunningham S, Enderby P, et al. Automatic speech recognition and training for severely dysarthric users of assistive technology: the STARDUST project. *Clin Linguist Phon*. 2006;20(2-3):149–156.
- [21] Hosom J-P, Jakobs T, Baker A, et al. Automatic Speech Recognition for Assistive Writing in Speech Supplemented Word Prediction [Internet]. 2010. [cited 2020 Apr 26]. Available from: <https://aac.unl.edu/reference/SITmanual.pdf>.
- [22] Hawley MS, Cunningham SP, Green PD, et al. A voice-input voice-output communication aid for people with severe speech impairment. *IEEE Trans Neural Syst Rehabil Eng*. 2013;21(1):23–31.
- [23] Bright R, Hawley M, Fryer K, et al. From VIVOCA to VocaTempo: Development and evaluation of a voice-input voice-output communication aid app. *Technol Disabil* [Internet]. 2019. 31:S86–S87. Available from: <https://ovidsp.ovid.com/ovidweb.cgi?T=JS&PAGE=reference&D=emexb&NEWS=N&AN=630726085>.
- [24] Therapy-Box. VocaTempo | Therapy Box [Internet]. 2020. [cited 2020 Apr 26]. Available from: <https://therapy-box.co.uk/vocatempo>.
- [25] Firth A. Practical web inclusion and accessibility: a comprehensive guide to access needs. 1st ed. 2019.
- [26] Voiceitt. Voiceitt - Home [Internet]. 2020. [cited 2020 Nov 8]. Available from: <https://www.voiceitt.com/>.
- [27] Voiceitt. Nuvoic Project [Internet]. 2020. [cited 2020 Nov 13]. Available from: <https://www.voiceitt.com/nuvoic-project.html>.
- [28] Cattiau J. How AI can improve products for people with impaired speech [Internet]. 2019. [cited 2020 Jul 6]. Available from: <https://www.blog.google/outreach-initiatives/accessibility/impaired-speech-recognition/>.
- [29] Pew Research. Voice assistants used by 46% of Americans, mostly on smartphones | Pew Research Center [Internet]. 2017. [cited 2020 Jan 20]. Available from: <https://www.pewresearch.org/fact-tank/2017/12/12/nearly-half-of-americans-use-digital-voice-assistants-mostly-on-their-smartphones/>.
- [30] Bäumer D, Talbot K, Turner MR. Advances in motor neurone disease. *J R Soc Med*. 2014;107(1):14–21.
- [31] Hawley MS, Enderby P, Green P, et al. A speech-controlled environmental control system for people with severe dysarthria. *Med Eng Phys*. 2007;29(5):586–593.
- [32] Simpson R, LoPresti E, Cooper R. How many people would benefit from a smart wheelchair? *J Rehabil Res Dev*. 2008; 45(1):53–72.
- [33] De Russis L, Corno F. On the impact of dysarthric speech on contemporary ASR cloud platforms. *J Reliable Intell Environ*. 2019;5(3):163–172.
- [34] Koch Fager S, Fried-Oken M, Jakobs T, et al. New and emerging access technologies for adults with complex communication needs and severe motor impairments: State of the science. *Augment Altern Commun*. 2019;35(1): 13–25.
- [35] Shor J, Emanuel D, Lang O, et al. Personalizing ASR for dysarthric and accented speech with limited data. *Proceedings of the Annual Conference of the International Speech Communication Association, INTERSPEECH*. 2019. p. 784–788.
- [36] Moore M, Venkateswara H, Panchanathan S. Whistle-blowing ASRs: evaluating the need for more inclusive automatic speech recognition systems. *Proceedings of the Annual Conference of the International Speech Communication Association, INTERSPEECH*. 2018. p. 466–470. DOI:10.21437/Interspeech.2018-2391
- [37] Caballero-Morales SO, Trujillo-Romero F. Evolutionary approach for integration of multiple pronunciation patterns for enhancement of dysarthric speech recognition. *Expert Syst Appl* [Internet]. 2014;41(3):841–852. DOI:10.1016/j.eswa.2013.08.014
- [38] Sanders E, Ruitter M, Beijer L, et al. Automatic recognition of dutch dysarthric speech a pilot study. *7th International Conference on Spoken Language Processing, ICSLP*. 2002. p. 661–664.
- [39] Moar J, Escherich M. Voice Assistant Devices in Use to Overtake World Population by 2024 [Internet]. 2020. [cited 2020 Jul 12]. Available from: <https://www.juniperresearch.com/press/press-releases/number-of-voice-assistant-devices-in-use?ch=voiceassistants>.
- [40] United Nations. Leaving no one behind: the imperative of inclusive development Report on the World Social Situation 2016. 2016.
- [41] Arksey H, O'Malley L. Scoping studies: towards a methodological framework. *Int J Soc Res Methodol Theory Pract*. 2005;8(1):19–32.
- [42] Armstrong R, Hall BJ, Doyle J, et al. Cochrane Update. 'Scoping the scope' of a cochrane review". *J Public Health (Oxf)*. 2011;33(1):147–150..
- [43] Tricco AC, Lillie E, Zarin W, et al. PRISMA extension for scoping reviews (PRISMA-ScR): checklist and explanation. *Ann Intern Med*. 2018;169(7):467–473.
- [44] Levac D, Colquhoun H, O'Brien KK. Scoping studies: Advancing the methodology. *Implement Sci*. 2010;5:1–9.
- [45] Computerlove.com. Home voice assistant survey data 2019. How are people using their voice assistants? [Internet]. 2019 [cited 2020 Jul 26]. Available from: <https://www.code-computerlove.com/blog/voice-assistant-survey-2019>.
- [46] Peters MD, Godfrey CM, McInerney P, et al. The Joanna Briggs Institute Reviewers' Manual 2015: Methodology for JBI scoping reviews. *Joanne Briggs Inst* [Internet]. 2015;1–24. Available from: https://joannabriggs.org/assets/docs/sumari/ReviewersManual_Mixed-Methods-Review-Methods-2014-ch1.pdf.
- [47] Hird K, Hennessey N. Facilitating use of speech recognition software for people with disabilities: a comparison of three treatments. *Clin Linguist Phonetics*. 2007;21(3):211–226. [cited 2020 Aug 23] Available from: <https://ovidsp.ovid.com/ovidweb.cgi?T=JS&PAGE=reference&D=emed10&NEWS=N&AN=46957688>.
- [48] Rudzicz F. Articulatory knowledge in the recognition of dysarthric speech. *IEEE Trans Audio Speech Lang Process*. 2011;19(4):947–960.

- [49] Rudzicz F. Adjusting dysarthric speech signals to be more intelligible. *Comput Speech Lang* [Internet]. 2013;27(6): 1163–1177. DOI:10.1016/j.csl.2012.11.001
- [50] Hahm S, Heitzman D, Wang J. Recognizing dysarthric speech due to amyotrophic lateral sclerosis with across-speaker articulatory normalization. 2015. p. 47–54.
- [51] Kim M, Cao B, An K, et al. Dysarthric speech recognition using convolutional LSTM neural network. *Proceedings of the Annual Conference of the International Speech Communication Association, INTERSPEECH*. 2018. p. 2948–2952. DOI:10.21437/interspeech.2018-2250
- [52] Caves K, Cope B, Boemler S. Development of an automatic recognizer for dysarthric speech. *Proceedings of the RESNA Annual Conference*. 2007. Available from: <https://www.resna.org/sites/default/files/legacy/conference/proceedings/2007/Research/CAC/Caves.html>.
- [53] Ballati F, Corno F, Russis LD. Hey siri, do you understand me?": virtual assistants and dysarthria. *Intell Environ* [Internet]. 2018;0:557–566. Available from: <https://www.researchgate.net/publication/325466714>.
- [54] Pradhan A, Mehta K, Findlater L. "Accessibility came by accident": Use of voice-controlled intelligent personal assistants by people with disabilities. *Conf Hum Factors Comput Syst - Proc* [Internet]. Association for Computing Machinery; 2018. [cited 2020 Aug 14].
- [55] Shor J, Emanuel D. Project Euphonia's Personalized Speech Recognition for Non-Standard Speech. 2019.
- [56] Enderby P. Frenchay dysarthria assessment. *Int J Lang Commun Disord*. 1980;15(3):165–173.
- [57] Ballati F, Corno F, De Russis L. Assessing virtual assistant capabilities with italian dysarthric speech. *ASSETS 2018 - Proc 20th Int ACM SIGACCESS Conf Comput Access* [Internet]. 2018 [cited 2020 Aug 14];93–101.
- [58] MND Association Forum. Search Result - Motor Neurone Disease Association Forum [Internet]. 2020. [cited 2020 Oct 18]. Available from: <https://forum.mndassociation.org/search?q=Alexa&searchJSON=%7B%22keywords%22%3A%22Alexa%22%7D>.
- [59] Gillet F. Getting Consumers Beyond Simple Tasks On Smart Speakers Is Challenging [Internet]. *Forrester.com*. 2020. [cited 2020 Sep 13]. Available from: <https://go.forrester.com/blogs/getting-consumers-beyond-simple-tasks-on-smart-speakers-is-challenging/>.
- [60] Raheja D, Stephens HE, Lehman E, et al. Patient-reported problematic symptoms in an ALS treatment trial. *Amyotroph Lateral Scler Frontotemporal Degener*. 2016; 17(3-4):198–205.
- [61] Hawley MS. Speech recognition as an input to electronic assistive technology. *Br J Occup Ther*. 2002;65(1):15–20.
- [62] Berdasco L, Diaz , et al. User experience comparison of intelligent personal assistants: Alexa, google assistant, siri and cortana. *Proceedings*. 2019;31:51.
- [63] Protalinski E. Google's speech recognition technology now has a 4.9% word error rate [Internet]. *VentureBeat*. 2017. [cited 2020 Oct 18]. Available from: <https://venturebeat.com/2017/05/17/googles-speech-recognition-technology-now-has-a-4-9-word-error-rate/>.
- [64] Fuller D. Google's Word Recognition Error Rate Is Down To 4.9% [Internet]. *Android Headl*. 2017. [cited 2020 Oct 18]. Available from: <https://www.androidheadlines.com/2017/05/googles-word-recognition-error-rate-is-down-to-4-9.html>.
- [65] Microsoft. Evaluate accuracy for Custom Speech - Speech service - Azure Cognitive Services | Microsoft Docs [Internet]. *Azur. Prod. Doc*. 2019. [cited 2020 Oct 23]. Available from: <https://docs.microsoft.com/en-us/azure/cognitive-services/speech-service/how-to-custom-speech-evaluate-data>.
- [66] Rosen K, Yampolsky S. Automatic speech recognition and a review of its functioning with dysarthric speech. *AAC Augment Altern Commun*. 2000;16(1):48–60.
- [67] Ferrier LJ, Shane HC, Ballard HF, et al. Dysarthric speakers intelligibility and speech characteristics in relation to computer speech recognition. *Augment Altern Commun*. 1995; 11(3):165–175. [Internet]. [cited 2020 Oct 23] Available from: DOI:10.1080/07434619512331277289
- [68] Kotler AL, Thomas-Stonell N. Effects of speech training on the accuracy of speech recognition for an individual with a speech impairment. *AAC Augment Altern Commun*. 1997; 13(2):71–80.
- [69] Norman DA. *The design of everyday things* (Rev. and expanded ed.). Boulder: Basic Books; 2013.
- [70] Vieira H, Costa N, Sousa T, et al. Voice-based classification of amyotrophic lateral sclerosis: where are we and where are we going? A systematic review. *Neurodegener Dis*. 2019;19(5-6):163–170.
- [71] Wang J, Kothalkar PV, Cao B, et al. Towards automatic detection of amyotrophic lateral sclerosis from speech acoustic and articulatory samples. *Proceedings of the Annual Conference of the International Speech Communication Association, INTERSPEECH*. 2016. p. 1195–1199. DOI:10.21437/Interspeech.2016-1542
- [72] An KH, Kim M, Teplansky K, et al. Automatic early detection of amyotrophic lateral sclerosis from intelligible speech using convolutional neural networks. *Proceedings of the Annual Conference of the International Speech Communication Association, INTERSPEECH*. 2018. p. 1913–1917. DOI:10.21437/Interspeech.2018-2496
- [73] Connaghan KP, Green JR, Paganoni S, et al. Use of beibe smartphone app to identify and track speech decline in amyotrophic lateral sclerosis. *Interspeech*. 2019. p. 4504–4508. DOI:10.21437/Interspeech.2019-3126
- [74] Stegmann GM, Hahn S, Liss J, et al. Early detection and tracking of bulbar changes in ALS via frequent and remote speech analysis. *npj Digit Med*. 2020;3. DOI:10.1038/s41746-020-00364-6
- [75] Orozco-Arroyave JR, et al. Towards an automatic monitoring of the neurological state of Parkinson's patients from speech. *ICASSP*. 2016. p. 6490–6494. DOI:10.1109/ICASSP.2016.7472927
- [76] Baxter S, Enderby P, Evans P, et al. Interventions using High-Technology communication devices: a state of the art review. *Folia Phoniatr Logop*. 2012;64(3):137–144. [Internet]. Available from: <https://search.proquest.com/docview/1268739363?accountid=14511>.