

Evaluating *Camera Mouse* as a computer access system for augmentative and alternative communication in cerebral palsy: a case study

Lauren E. MacLellan, Cara E. Stepp, Susan K. Fager, Michelle Mentis, Alyssa R. Boucher, Defne Abur & Gabriel J. Cler

To cite this article: Lauren E. MacLellan, Cara E. Stepp, Susan K. Fager, Michelle Mentis, Alyssa R. Boucher, Defne Abur & Gabriel J. Cler (12 Sep 2023): Evaluating *Camera Mouse* as a computer access system for augmentative and alternative communication in cerebral palsy: a case study, *Assistive Technology*, DOI: [10.1080/10400435.2023.2242893](https://doi.org/10.1080/10400435.2023.2242893)

To link to this article: <https://doi.org/10.1080/10400435.2023.2242893>



Published online: 12 Sep 2023.



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



Article views: 16



View related articles [↗](#)



View Crossmark data [↗](#)



Evaluating *Camera Mouse* as a computer access system for augmentative and alternative communication in cerebral palsy: a case study

Lauren E. MacLellan, MS, CCC-SLP^a, Cara E. Stepp, PhD ^{a,b,c,d}, Susan K. Fager, PhD, CCC-SLP ^e, Michelle Mentis, PhD, CCC-SLP^a, Alyssa R. Boucher, PhD, CCC-SLP^a, Defne Abur, PhD ^a, and Gabriel J. Cler, PhD ^{a,b,f}

^aDepartment of Speech, Language, and Hearing Sciences, Boston University, Boston, Massachusetts, USA; ^bGraduate Program for Neuroscience, Boston University, Boston, Massachusetts, USA; ^cDepartment of Biomedical Engineering, Boston University, Boston, Massachusetts, USA; ^dDepartment of Otolaryngology – Head and Neck Surgery, Boston University School of Medicine, Boston, Massachusetts, USA; ^eInstitute for Rehabilitation Science and Engineering, Madonna Rehabilitation Hospitals, Lincoln, Nebraska, USA; ^fDepartment of Speech & Hearing Sciences, University of Washington, Seattle, Washington, USA

ABSTRACT

Camera Mouse is a freely available software program that visually tracks the movement of facial features to allow individuals with motor impairments to control a computer mouse. The goal of this case study was to provide an evaluation of *Camera Mouse* as a computer access method as part of a multiple modality communication system for an individual with cerebral palsy. The participant was asked to reproduce sentences and respond to ethical dilemmas for language sampling. Tasks were completed using natural speech and an AAC solution consisting of *Camera Mouse* paired with an orthographic selection interface and speech synthesis. The participant completed a questionnaire for satisfaction with the introduced assistive technology. *Camera Mouse* resulted in higher intelligibility than natural speech, while natural speech had a higher rate. She used more complex language with her natural speech. The participant rated *Camera Mouse* as at least 3/5 on all measures, including 5/5 on weight and safety. The results of this case study suggest *Camera Mouse* is a promising computer access system for communication supported by the participant's satisfaction rating, expressive language, and synthesized speech production capabilities.

ARTICLE HISTORY

Accepted 26 July 2023

KEYWORDS

Augmentative and alternative communication (AAC); cerebral palsy; camera; cursor control; speech; language

Introduction

Many people with motor impairments use adaptive and assistive supports to access augmentative and alternative communication (AAC) technologies (Fager, 2018). Access methods leverage voluntary movements to control a device and produce communication, and can include switches, head pointers, and eye-tracking devices. These access methods allow individuals to communicate their thoughts and needs through spelling or expression-building programs, use the internet, and more actively participate in recreational activities (Beukelman & Light, 2020; Buxton et al., 1986).

One common access option is head-tracking. Head-tracking systems use video or infrared cameras to track selected body features or reflective dots placed on the forehead, glasses, or head; movements of those body parts are translated into cursor control for computer access (Fager et al., 2012). These systems work well with small targets and require a short learning and training time to utilize (Bates & Istance, 2003). Commercial head-trackers typically utilize external hardware (producing infrared lights), usually require the user to wear a reflective dot, and cost \$500–\$1000 (e.g., SmartNav 4:AT by NaturalPoint; HeadMouse® Nano by Origin Instruments; TrackerPro by AbleNet). More recently, software that works on mobile devices has been made available for head tracking (e.g., Open Sesame by Sesame Enable; Ramirez, 2018). One

available software package that runs on consumer personal computers and tablets is called *Camera Mouse*.

Camera Mouse was developed to serve as a non-intrusive and inexpensive access method (Betke et al., 2002). *Camera Mouse* is free software that allows the user to control the mouse pointer on their computer via a webcam that tracks a predetermined body feature (available at <http://www.camera-mouse.org/>). *Camera Mouse* has been tested by the developers in a series of studies with participants with and without motor impairment (Betke et al., 2002; Gips et al., 2000). These studies have involved individuals using the system to select circular targets and to repeatedly type out prescribed messages (e.g., “Boston College”). Of the 15 participants with motor impairments tested in these studies, ten were able to use the system reliably, three could use it unreliably, and two could not use the system due to poor muscle control (Betke et al., 2002; Gips et al., 2000).

Objectives

Many individuals use multiple modalities, devices, techniques, and strategies to meet their communication needs throughout the day. *Camera Mouse* may be a promising addition to an AAC system for some individuals due to the fact that it can be used with a standard computer or tablet and is free.

The participant in this case study had severe dysarthria and used a range of modalities to support her daily communication, including natural speech for most face-to-face interactions, mobile devices for texting and some e-mail, and an eye-gaze speech-generating device (SGD) for written communication (i.e., e-mails and academic reports). This study aimed to assess whether *Camera Mouse* could effectively meet complex communication demands at the sentence and discourse level as a tool to augment face-to-face interactions that she typically managed with her natural speech.

Methods

Participant

The participant was a motivated female college student with cerebral palsy (CP). The participant was recruited for the study from a specialized rehabilitation hospital. She experienced gross and fine motor movement deficits secondary to CP. She used multiple modalities to communicate, including natural speech, texting and emailing via smartphone, and a SGD. Prior to participation, written consent was obtained from the participant, as approved by the Boston University Institutional Review Board.

Speech characteristics

She had a history of severe mixed spastic-dyskinetic dysarthria secondary to cerebral palsy. Her speech was characterized by impaired articulation, abnormal breath support, and abnormal prosody (pitch, loudness, rate) and resonance. As a result, her intelligibility was very low to unfamiliar listeners and moderate to even very familiar listeners. Even so, she chose to communicate almost entirely via speech in face-to-face interactions, even with unfamiliar listeners in contexts where AAC would be beneficial (e.g., job interviews).

Prior and current AAC use

Our participant had a long history of AAC use. She used Dynavox (3100) and Dynvox Vmax with Eye Gaze accessory and Tobii-Dynavox (Tobii I12) devices with touchscreen access, eventually progressing to eye-gaze access. She did not use low tech communication options. She was a direct selector with a stylus (hand access) and with the tip of her nose and eventually utilized eye-gaze for writing support, as the touchscreen approaches were very fatiguing for large amounts of text. She rejected other head-tracking systems as she did not want to wear a reflective dot on her head. All devices and access methods used were chosen specifically by her and represented her strong personal preferences; she was actively engaged in decision-making regarding what technology she used to support all forms of communication.

At the time of the study, the participant used natural speech for nearly all of her in-person communication, texting and emailing via smartphone accessed with the tip of her nose, and a TobiiDynavox I-12 eye tracking SGD with Communicator software, QWERTY keyboard interface with word prediction, and dwell time of 1 sec. She used the SGD primarily to support extended written communication and academic needs (e.g., communication for medical

appointments and health-related needs, writing reports, papers, giving pre-programmed speeches, and e-mail communication with teachers and peers). She accessed her device primarily in her room where it was optimally positioned to support her written communication needs. She indicated that the size of the device often obstructed her vision for safe ambulation when mounted to her power wheelchair and that it was difficult to operate in different lighting conditions.

Camera Mouse to text to speech

In this study, the participant completed tasks using her natural speech and the *Camera Mouse* access system, running on an Acer Aspire laptop with an integrated webcam. The laptop was equipped with other free software: the Click-N-Type Keyboard (Version 3.03, Lake Software) and Natural Reader Text-to-Speech (AT&T Co. 2016 NaturalSoft Limited). Click-N-Type allowed the participant to type out orthographic messages using movements from her head via *Camera Mouse*, and Natural Reader Text-to-Speech was used to generate speech output from those messages. *Camera Mouse* could be used with any virtual keyboard and text-to-speech system, but these were chosen as they were also free and compatible with the Windows laptop used. The system will henceforth be referred to as *Camera Mouse* when referring only to the access method and *Camera Mouse+* when paired with virtual keyboard and text-to-speech.

Camera Mouse made selections via dwell time, generating a mouse click if the user directed the mouse pointer over a specified location for more than 0.5 seconds. The participant produced her responses on a blank page of Natural Reader Text-to-Speech and selected “play” to initiate speech synthesis. The *Camera Mouse+* setup is shown in Figure 1.

Data collection

All acoustic recordings were made in a quiet room using a portable digital audio recorder (Zoom Corporation Handy Recorder model H4n) and a headset microphone and/or a video camera (Sony). The participant first completed two tasks with her natural speech and then completed those tasks

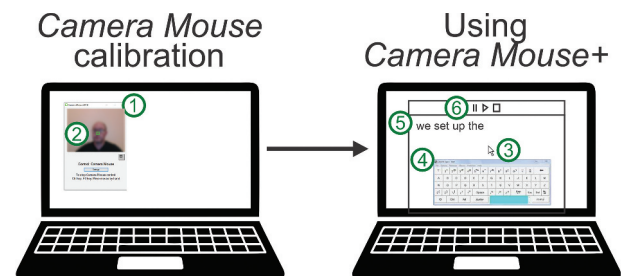


Figure 1. Visualization of computer setup. Left panel shows (1) the calibration window for *Camera Mouse*, in which the participant was centered in the window and the experimenter clicked her nose (marked with 2 here) so that the software would know what feature to track. The right panel shows the full *Camera Mouse+* package, with (3) the cursor now being controlled by the user's head, without any additional windows on the screen. The Click-N-Type software is shown in (4), with the natural Reader text-to-speech window in (5), where the participant would enter text. The natural Reader text-to-speech controls are shown in (6).

Table 1. Ethical dilemma prompts.

Modality	Dilemma
Camera Mouse	The cashier at the grocery store forgot to charge you for some fruit you had in your cart. You realize this as you are leaving the store. It was an honest mistake. When you look back at the register you see that there is already a long line. What would you do, and why?
Natural Speech	You are working at a bank. One day, one of the bank tellers who has become one of your best friends tells you that her daughter is extremely ill and needs to undergo a \$10,000 operation to survive. She has no insurance and no money left because of the medications and doctors visits she has been paying for. A few weeks later you ask your friend how her daughter is doing. She confides in you that she took \$10,000 from a dormant account that hasn't been touched in a few years. Her daughter was able to get the surgery and is now healthy. Your friend assures you that she has already begun paying the money back and will continue to do so until it is all returned. What would you do, and why?

using *Camera Mouse+* as an AAC system. She used *Camera Mouse+* for less than five minutes before data collection began.

Two tasks were chosen to evaluate her ability to use *Camera Mouse+* to support complex communication needs: a sentence replication task and an ethical dilemma task. As previous studies using *Camera Mouse* had focused on clicking dots or very simple text generation, these tasks were designed to probe more simple (sentence replication) and complex (ethical dilemma responses) language usage.

Sentence replication task

Speech samples were elicited via the Sentence Intelligibility Test (SIT; Yorkston et al., 1996). SIT software was used to generate 43 random sentence stimuli ranging from 5 to 11 words in length. The first author read the sentence to the participant, and said, “Go,” after which the participant replicated it using the given communication method. All productions were recorded and timed.¹

Self-generated language samples

The participant responded to hypothetical ethical dilemma scenarios designed to elicit speech that mimicked the participant's language in daily life (see Table 1). The participant was asked to explain what she would do in a given situation and provide reasoning to support her decision. The first author transcribed and timed the participant's responses using acoustic and video recordings.

User satisfaction survey

After using *Camera Mouse+*, the participant completed the Quebec User Evaluation of Satisfaction with Assistive Technology (QUEST 2.0) about her experience. The QUEST 2.0 is a 12-item outcome measure that examines an individual's satisfaction with features of the device and other related services (Demers, Monette, et al., 2002; Demers, Weiss-lambrou, et al., 2002). The first author read each of the 12 items to the participant and presented her with a rating scale ranging from 1 (not satisfied at all) to 5 (very satisfied). The participant responded verbally.

Data analysis

Data were analyzed in terms of intelligibility, rate, expressive language, and user satisfaction. Intelligibility was assessed via transcription by unfamiliar listeners. Rate was measured as words per minute. Expressive language was measured with

SALT software, and user satisfaction was derived from the QUEST 2.0 survey.

Intelligibility

Unfamiliar listeners were recruited to provide orthographic transcriptions as a measure of intelligibility. Four individuals who reported no history of speech, language, or hearing disorders (two men, two women) participated. Transcribers were 20–22 year old ($M = 21.5$ years, $SD = 1.0$) native speakers of North American English, and they demonstrated normal hearing via pure tone hearing screening (25 dB HL at 250, 500, 1000, 2000, and 4000 Hz). Transcribers provided written consent as approved by the Boston University Institutional Review Board.

Transcribers heard recorded SIT sentences and provided orthographic transcriptions. Each SIT sentence (natural speech and synthesized speech) was peak normalized such that the loudness would remain approximately consistent for listeners. Recordings were presented in a pseudorandom order via a custom program written in MATLAB (“MATLAB,” 2013). Transcribers wore Sennheiser HD 280 pro over-the-ear headphones with the output adjusted to their most comfortable loudness level. They were instructed to listen to the sentence and type what they heard into a text box on the screen. They were only given one opportunity to listen to each sentence and were instructed to make their best guess even if they were unsure of the intended message. This method was chosen to prevent context bias and familiarization if they had been allowed repeated listenings.

Intelligibility was calculated as the total words matching between the listener transcription and the target sentences divided by the total number of words (Garcia & Dagenais, 1998; Hustad et al., 2012). All minor misspellings, homonyms, and contractions (e.g., “it's” for “it is”) were counted as correct (Hustad et al., 2012). The intelligibility score for each modality was averaged over all listeners and sentences. Inter-transcriber reliability was calculated with ICC(2,k) in R (R Core team, 2015) with the package *irr*, version 0.84.1. For *Camera Mouse+* sentences, inter-transcriber reliability was .85 (95% confidence interval = 0.65–0.95). For natural speech sentences, inter-transcriber reliability was .81 (95% confidence interval = 0.63–0.91).

Rate

Rate was calculated from the SIT sentences as words per minute. Based on the video of the session, the first author calculated the amount of time it took to produce the sentence

¹The participant was instructed to produce half of sentences as fast as possible, and the remaining sentences as accurately as possible. However, initial analysis showed no differences between the conditions. Thus data from both conditions were ultimately pooled to calculate accuracy and rate values.

from when she said “Go” after presenting the stimulus until the end of the produced sentence. The end was defined as when the participant stopped speaking or when the text-to-speech finished.

Expressive language

The ethical dilemma language samples were analyzed using the Systematic Analysis of Language Transcripts (SALT) software (Miller & Iglesias, 2012). The samples were transcribed by the first author and segmented into C-units, which consist of a main clause and all embedded subordinate clauses (Miller et al., 2011). The transcripts were then coded using standard SALT transcription conventions. All transcripts were reviewed by a second investigator. Interscorer agreement for both C-unit segmentation and clausal coding exceeded .95. A third investigator reviewed the discrepancies and resolved the disagreements. Raters were allowed to listen to each sample as many times as they wished. Although the number of repetitions that were needed to transcribe the text were not tallied, it took many repetitions due to the low level of intelligibility.

The Standard Measures Report was generated using SALT software to yield measures of transcript length (total number of words), intelligibility, syntax/morphology (MLU in words and morphemes, Subordination Index), semantics (number total words (NTW), number different words (NDW), Type Token Ratio (TTR)), verbal facility (words per minute, pauses within and between utterances, maze words as a percentage of total words, and abandoned utterances), and errors.

Results

The participant completed all tasks using both her natural speech and *Camera Mouse+*. Objective and subjective measures of performance are presented in terms of intelligibility, rate, language, and the participant’s satisfaction ratings.

Sentence intelligibility and rate

The participant produced a total of 34 sentences (12 *Camera Mouse+*, 22 natural speech). Intelligibility improved considerably with the introduction of assistive technology (Figure 2a). The participant’s average intelligibility was 26.6% using natural speech (SD = 19.9%, range = 1.8–87.5%). Of the 22 sentences and 4 transcribers, only one sentence was completely correctly understood by any transcriber. That sentence, the most intelligible sentence in this modality at 87.5%, was “we hope to meet them again.” The next most intelligible sentence was “if you read the fine print you’ll find that most brands must be defrosted first,” at 48.3%, which was transcribed as “If you read the fine print you will find,” “if you read the fine print you will find most be fine will fit,” “if you read the fine print you will find more information,” and “if you dream you will find more happiness.” Her least intelligible sentence was “merely defining the risks is not enough for a jockey to keep his job,” which was transcribed as “we cannot hear it go,” “In order,” “i don’t know,” and “for the street edge.”

Intelligibility improved to 95.8% (SD = 6.6%, range = 80–100%) with *Camera Mouse+*. All sentences with *Camera Mouse+* had at least one transcriber with 100% intelligibility

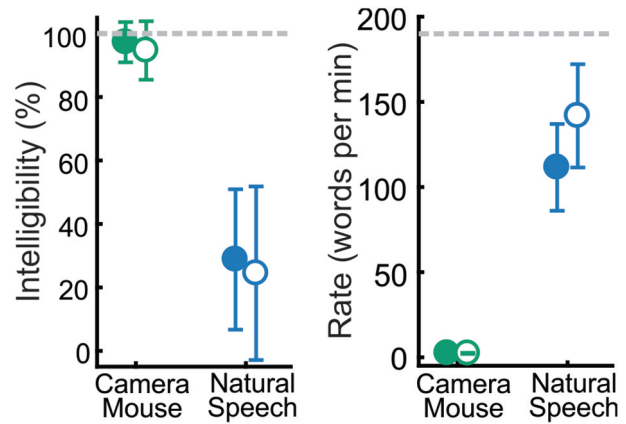


Figure 2. (a) Average intelligibility (%) of *Camera Mouse+* and natural speech, as transcribed by 4 recruited listeners in a sentence intelligibility testing task. Horizontal dotted line at 100% is typical speaker intelligibility. (b) average rate, as measured in words per minute. Horizontal dotted line at 190 is typical speaker intelligible words per minute. All plots: error bars represent standard deviation.

except the sentence, “The first step is to realize that one is proud,” which was transcribed as “The first step is to realize I is proud (three transcribers)” and “The first step is to realize I was proud.” The six sentences that were transcribed correctly by all raters included, “Their output was mostly in the second half,” and “We do not regard him as a financial wizard.”

Natural speech resulted in a higher speech rate: 126.7 WPM (SD = 32.2), compared to 2.4 WPM (SD = 0.7) with *Camera Mouse+* (Figure 2b).

Language sampling analysis

Results of the SALT analyses of the ethical dilemma language samples showed meaningful differences in transcript length, syntax and morphology, verbal facility, and errors (see Table 2). The participant produced a longer transcript utilizing natural speech, producing a total of 67 words versus 39 words using *Camera Mouse+*. Verbal facility, or speech rate, was 55.1 WPM with natural speech and 2.1 WPM with *Camera Mouse+*. The mean length of utterance (MLU) in words and morphemes for natural speech was higher at 12.0 words and 12.8 morphemes, versus 6.5 words and 6.7 morphemes with *Camera Mouse+*. The discourse sample resulted in subordination index scores of 1.25 (natural speech) and 0.8 (*Camera Mouse+*). Finally, natural speech resulted errors in 60% of utterances, including one omitted function word (“by”) and one omitted bound morpheme signaling past tense. *Camera Mouse+* resulted in 66.7% of utterances with errors, including four omitted words (“and,” “I,” and the auxiliary verb “would”

Table 2. Ethical dilemma standard measures report.

	Natural Speech	Camera Mouse
SI-Composite	1.25	0.8
Rate (WPM)	55.1	2.1
Total # Words	60	39
Total # C-Units	5	6
MLU Morphemes	12.8	6.7
MLU Words	12.0	6.5
# Of Omissions	2	4
% Utterances with Errors	60.0	66.7

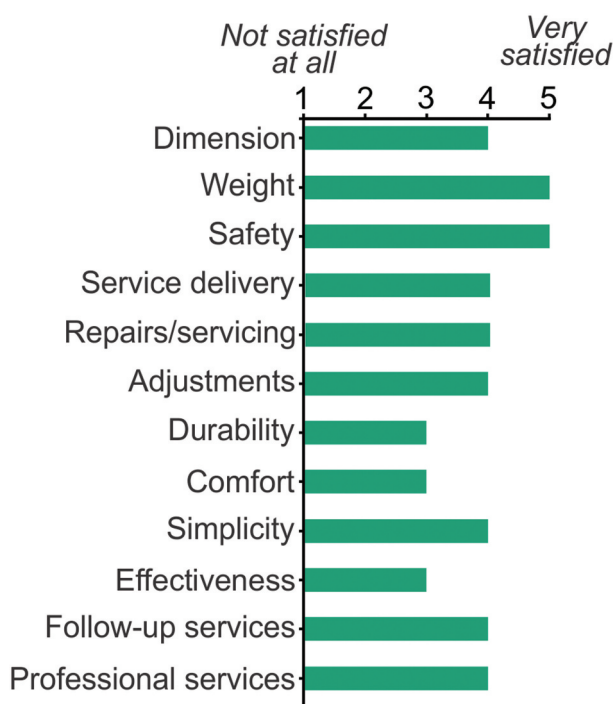


Figure 3. Results from QUEST 2.0 survey about the participant's experience with Camera Mouse+. Participant rated each item on scale of 1–5, ranging from not satisfied at all to very satisfied.

twice). The participant made one word-level error where she used an incorrect auxiliary verb (“don’t” for “wouldn’t”) and one utterance-level error (a sentence fragment).

User satisfaction

Figure 3 shows item scores for the QUEST 2.0 related to the Camera Mouse+ system. Average scores for Camera Mouse+ were 3.92 (SD = 0.67). These values suggest that, after her fairly brief use of the system, the participant is “quite satisfied” with Camera Mouse+. The participant that the satisfaction factors of adjustments, ease of use, and comfort were the most important.

Discussion

Many who experience dysarthria and use AAC continue to use their natural speech in some contexts, and they find themselves relying on a wide range of technologies, for various practical and personal preference reasons, to support all communication needs. The aim of the present study was to evaluate the impact of a camera-based access strategy on communication ability (intelligibility, rate, and measures of expressive language) and satisfaction in an individual with CP and compare its impact with that of her natural speech. The participant demonstrated the ability to use Camera Mouse+ after a very short familiarization in order to produce sentences and short sequences of discourse to transmit information to listeners.

Intelligibility and rate

The participant in this study had a history of severe dysarthria secondary to CP. As with many adults with CP, her speech was

characterized by impaired articulation, abnormal breath support, and abnormal prosody and resonance (Cockerill et al., 2014; Schölderle et al., 2016). As a result, her intelligibility was severely impaired, which was captured in the accuracy results from the SIT task: using natural speech, unfamiliar listeners could only identify 23.7% of the message. However, this participant still includes her natural speech as part of a multimodal communication system. Despite reported high rates of unintelligibility, many individuals with CP use natural speech as a primary mode of communication, particularly with familiar listeners (Chung et al., 2012; Cockerill et al., 2014). Unintelligible natural speech may negatively impact an individual in their daily life in terms of peer attitudes, employment, and participation (Beck et al., 2000; Bryen et al., 2007; Shikako-Thomas et al., 2008). In the current study, the participant's average intelligibility improved substantially using the assistive technology, from 23.7% to 96.7%.

AAC devices and technologies are inherently slow compared to natural speech interactions (Higginbotham, 2009). In this study, the participant utilized dwell selection of .5 seconds to select each letter. This automatically imposes an additional time constraint. Common methods used to enhance rate and efficiency of communication (e.g., word prediction, frequently use pre-stored messages) were not used in this study, but would be vital as part of a multimodal AAC strategy.

Individuals who rely on assistive technologies to support communication are often faced with this intelligibility/rate trade-off, which motivates their use of multiple modalities to support communication. For some individuals in some contexts, the intelligibility of the message trumps the speed of communication (e.g., communication of complex medical needs, communication with unfamiliar listeners in community and vocational settings). In other settings, with listeners who are highly familiar or with messages that are context-dependent, a high intelligibility may not be required.

Expressive language

The overall results of the evaluation of the participant's discourse language samples indicate that expressive language was more complex for natural speech in terms of transcript and C-Unit length and clausal density. The participant exhibited expressive language with a variety of grammatical structures using Camera Mouse+. However, expressive language may also be impacted by AAC use, particularly in terms of utterance and transcript length and complexity. AAC users may produce utterances that are shorter and simpler than what would be expected given age and developmental level and may produce more grammatical errors (Binger & Light, 2008; Yorkston et al., 1990). This was observed in our study: the participant produced fewer words per transcript and utterance and maintained a lower MLU and clausal density using assistive technology compared to her natural speech. Given the time required to transmit a message via AAC, using short utterances and omitting nonessential information (e.g., bound morphemes and function words) is likely an effective strategy to reduce the time needed to create a message (Smith et al., 1989).

The subordination index analysis showed that the participant achieved subordination composite scores of 1.25 using

her natural speech and 0.8 using *Camera Mouse+*. Subordination index composite scores less than 1.0 suggest very few or no complex sentences, suggesting that the participant was generating simplified language when using the assistive technology.

User satisfaction and comparison to other technology

Although her exposure to the *Camera Mouse* system was brief, the participant rated it as satisfactory in all dimensions. She particularly noted that she was very satisfied (5/5) with *Camera Mouse*'s weight and safety, and quite satisfied (4/5) with *Camera Mouse*'s dimensions, simplicity, adjustments, and several metrics of servicing (service delivery, repairs/servicing, follow-up services, professional services). It has been well documented that factors like these are important to people who use high-tech AAC systems (Cooper et al., 2009; Dattilo et al., 2008; Hodge, 2007).

Through her many years of AAC use, the participant has used many different devices and is likely comparing *Camera Mouse* to all of these options and to her natural speech. She has presumably noted those components of the *Camera Mouse* system that its developers also emphasize: it uses a webcam in a typical laptop, rather than as part of a custom AAC device. It does not require the external hardware often needed for other head- and eye-trackers, nor a reflective dot placed directly on the skin or attached to glasses or a hat.

Study limitations and future directions

Case study design

Due to the case study design of the experiment and use of only one participant, generalization to the population of individuals with CP is not possible. It is likely that those with some experience head-tracking or head-based direct selection (as our participant, who used her nose on touchscreen devices, or with a mouth-stick) would be able to trial the system. Future studies could focus on the amount of head control necessary to use *Camera Mouse*.

We used a laboratory laptop, but *Camera Mouse* software could be loaded on any Windows-based device including most SGDs. This would be particularly beneficial, as then individuals could switch between access methods (e.g., eye-gaze, head-tracking, touchscreen) independently from on-screen keyboards and text-to-speech software.

Satisfaction ratings

The participant rated *Camera Mouse+* highly in most areas asked (all dimensions rated at least 3 out of 5), with the highest ratings (5/5) in weight and safety. However, with her brief exposure to *Camera Mouse*, this participant was unable to evaluate it in all of the contexts that she might typically need to use it. For example, infrared systems are designed to be used in diverse lighting conditions. *Camera Mouse* relies on some ambient light to properly track the user (Gips, 2017; Vojtech et al., 2020). Although this may be supplied by the laptop screen itself in darkness, it may not work in direct sunlight. In addition, the participant rated *Camera Mouse+* highly in all service-related questions (4/5 for service delivery, repairs/servicing,

follow-up services, professional services). *Camera Mouse* and the additional components used to make it a full text to speech system (Click-N-Type, Natural Reader) are all free software and therefore come without service. The participant may be reacting to the software being available on typical consumer-grade electronics, compared to her experience with proprietary software, but we did not probe further in this study. Specifically, we did not assess whether the participant would be able to set up the hardware (typical laptop) or software (*Camera Mouse*) independently. It is clear that a follow-up study in which a participant uses *Camera Mouse* in a variety of daily situations is needed.

Intelligibility ratings

Here we used transcriptions by unfamiliar listeners as a proxy for unfamiliar listeners that the participant may encounter throughout her daily life. However, with familiar listeners, she may be reaching much higher levels of intelligibility; with these individuals, she would likely choose to continue using natural speech. We intentionally chose context-less messages to try to measure the quality of the speech signal without additional listener cognitive contribution. Future research could utilize comprehensibility measurements to get a more realistic measurement of communication success.

Camera Mouse as part of a multimodal AAC strategy

Camera Mouse was not directly compared to an eye-gaze SGD. Future studies should compare the integration *Camera Mouse* into multimodal AAC strategies (e.g., SGDs, mobile devices, natural speech). The developers of *Camera Mouse* have recently published work combining head tracking with eye gaze, which may provide benefits to some individuals (Feng et al., 2021) However, the ability to maintain eye contact found in *Camera Mouse* (i.e., with head movements only) could be a benefit to others, and should be evaluated further. Similarly, speech supplementation strategies might be combined with *Camera Mouse* to require less text generation. Such research would provide guidance for individuals who use AAC and their treatment teams about the specific contexts that different modalities might best meet communication needs.

Acknowledgements

The authors would first like to thank the participant for her work; we are very grateful for her help and her expertise as an AAC user. We would further like to thank Amrita Nishtala and Saniya Shah for their contributions in the early design stages of this work.

Disclosure statement

No potential conflict of interest was reported by the author(s).

Funding

The work was supported by the National Institutes of Health - National Institute on Deafness and Other Communication Disorders under grants F31 DC014872 and F32 DC017637 (GJC) and the National Science Foundation under grant 1452169 (CES).

ORCID

Cara E. Stepp PhD  <http://orcid.org/0000-0002-8045-252X>
 Susan K. Fager PhD, CCC-SLP  <http://orcid.org/0000-0001-8125-5977>
 Define Abur PhD  <http://orcid.org/0000-0002-0410-8487>
 Gabriel J. Cler PhD  <http://orcid.org/0000-0002-5279-7336>

References

- Bates, R., & Istance, H. O. (2003). Why are eye mice unpopular? A detailed comparison of head and eye controlled assistive technology pointing devices. *Universal Access in the Information Society*, 2(3), 280–290. <https://doi.org/10.1007/s10209-003-0053-y>
- Beck, A., Fritz, H., Keller, A., & Dennis, M. (2000). Attitudes of school-aged children toward their peers who use augmentative and alternative communication. *Augmentative and Alternative Communication*, 16(1), 13–26. <https://doi.org/10.1080/07434610.012331278874>
- Betke, M., Gips, J., & Fleming, P. (2002). The camera mouse: Visual tracking of body features to provide computer access for people with severe disabilities. *IEEE Transactions on Neural Systems and Rehabilitation Engineering*, 10(1), 1–10. <https://doi.org/10.1109/TNSRE.2002.1021581>
- Beukelman, D. R., & Light, J. C. (2020). *Augmentative & alternative communication: Supporting children and adults with complex communication needs* (5th ed.). Paul H. Brookes Publishing Co., Inc.
- Binger, C., & Light, J. (2008). The morphology and syntax of individuals who use AAC: Research review and implications for effective practice. *Augmentative and Alternative Communication*, 24(2), 123–138. <https://doi.org/10.1080/07434610701830587>
- Bryen, D. N., Potts, B. B., & Carey, A. C. (2007). So you want to work? What employers say about job skills, recruitment and hiring employees who rely on AAC. *Augmentative and Alternative Communication*, 23(2), 126–139. <https://doi.org/10.1080/07434610600991175>
- Buxton, W., Foulds, R., Rosen, M., Scadden, L., & Shein, F. (1986). Human interface design and the handicapped user. *ACM SIGCHI Bulletin*, 17(4), 291–297. <https://doi.org/10.1145/22339.22386>
- Chung, Y., Behrmann, M., Bannan, B., & Thorp, E. (2012). Perspectives of high tech augmentative and alternative communication users with cerebral palsy at the post-secondary level. *Perspectives on Augmentative and Alternative Communication*, 21(2), 43–55. <https://doi.org/10.1044/aac21.2.43>
- Cockerill, H., Elbourne, D., Allen, E., Scrutton, D., Will, E., McNee, A., Fairhurst, C., & Baird, G. (2014). Speech, communication and use of augmentative communication in young people with cerebral palsy: The SH&PE population study. *Child: Care, Health and Development*, 40(2), 149–157. <https://doi.org/10.1111/cch.12066>
- Cooper, L., Balandin, S., & Trembath, D. (2009). The loneliness experiences of young adults with cerebral palsy who use alternative and augmentative communication. *Augmentative and Alternative Communication*, 25(3), 154–164. <https://doi.org/10.1080/07434610903036785>
- Dattilo, J., Estrella, G., Estrella, L. J., Light, J., McNaughton, D., & Seabury, M. (2008). “I have chosen to live life abundantly”: Perceptions of leisure by adults who use augmentative and alternative communication. *Augmentative and Alternative Communication*, 24(1), 16–28. <https://doi.org/10.1080/07434610701390558>
- Demers, L., Monette, M., Lapierre, Y., Arnold, D. L., & Wolfson, C. (2002). Reliability, validity, and applicability of the Quebec user evaluation of satisfaction with assistive technology (QUEST 2.0) for adults with multiple sclerosis. *Disability & Rehabilitation*, 24(1–3), 21–30. <https://doi.org/10.1080/09638280110066352>
- Demers, L., Weiss-Lambrou, R., & Ska, B. (2002). The Quebec user evaluation of satisfaction with assistive technology (QUEST 2.0): An overview and recent progress. *Technology and Disability*, 14(3), 101–105. <https://doi.org/10.3233/TAD-2002-14304>
- Fager, S. K. (2018). Alternative access for adults who rely on augmentative and alternative communication. *Perspectives of the ASHA Special Interest Groups*, 3(12), 6–12. <https://doi.org/10.1044/PERSP3.SIG12.6>
- Fager, S. K., Beukelman, D. R., Fried-Oken, M., Jakobs, T., & Baker, J. (2012). Access interface strategies. *Assistive Technology*, 24(1), 25–33. <https://doi.org/10.1080/10400435.2011.648712>
- Feng, W., Zou, J., Kurauchi, A., Morimoto, C. H., & Betke, M. (2021). Hgaze typing: Head-gesture assisted gaze typing. In *ACM Symposium on Eye Tracking Research and Applications* (pp. 1–11). Association for Computing Machinery.
- Garcia, J. M., & Dagenais, P. A. (1998). Dysarthric sentence intelligibility: Contribution of iconic gestures and message predictiveness. *Journal of Speech, Language, and Hearing Research: JSLHR*, 41(6), 1282–1293. <https://doi.org/10.1044/jslhr.4106.1282>
- Gips, J. (2017). *Camera Mouse 2018 user manual*. <http://www.camera-mouse.org/downloads/CameraMouse2018Manual.pdf>
- Gips, J., Betke, M., & Fleming, P. (2000). The Camera Mouse: Preliminary investigation of automated visual tracking for computer access. In *Proceedings of RESNA 2000*, (pp. 98–100).
- Higginbotham, D. J. (2009). In-person interaction in AAC: New perspectives on utterances, multimodality, timing, and device design. *Perspectives on Augmentative and Alternative Communication*, 18(4), 154–160. <https://doi.org/10.1044/aac18.4.154>
- Hodge, S. (2007). Why is the potential of augmentative and alternative communication not being realized? Exploring the experiences of people who use communication aids. *Disability & Society*, 22(5), 457–471. <https://doi.org/10.1080/09687590701427552>
- Hustad, K. C., Schueler, B., Schultz, L., & DuHadway, C. (2012). Intelligibility of 4-year-old children with and without cerebral palsy. *Journal of Speech, Language, & Hearing Research*, 55(4), 1177–1189. [https://doi.org/10.1044/1092-4388\(2011/11-0083\)](https://doi.org/10.1044/1092-4388(2011/11-0083))
- MATLAB. (2013). The MathWorks Inc.
- Miller, J. F., Andriacchi, K., & Nockerts, A. (2011). *Assessing language production using SALT software: A clinician's guide to language sample analysis*. SALT Software, LLC.
- Miller, J., & Iglesias, A. (2012). *Systematic analysis of language transcripts (SALT)*. SALT Software LLC.
- Ramirez, R. D. (2018). Evaluating the usability of three accessible smartphone designs by people with cerebral palsy. *The American Journal of Occupational Therapy*, 72(4_Supplement_1), 7211500063p1. <https://doi.org/10.5014/AJOT.2018.72S1-PO7009>
- R Core team. (2015). *R: A language and environment for statistical computing*. R Foundation for Statistical Computing.
- Schölderle, T., Staiger, A., Lampe, R., Strecker, K., & Ziegler, W. (2016). Dysarthria in adults with cerebral palsy: Clinical presentation and impacts on communication. *Journal of Speech, Language, & Hearing Research*, 59(2), 216–229. https://doi.org/10.1044/2015_JSLHR-S-15-0086
- Shikako-Thomas, K., Majnemer, A., Law, M., & Lach, L. (2008). Determinants of participation in leisure activities in children and youth with cerebral palsy: Systematic review. *Physical & Occupational Therapy in Pediatrics*, 28(2), 155–169. <https://doi.org/10.1080/01942630802031834>
- Smith, A., Thruston, S., Light, J., Parnes, P., & O’Keefe, B. (1989). The form and use of written communication produced by physically disabled individuals using microcomputers. *Augmentative and Alternative Communication*, 5(2), 115–124. <https://doi.org/10.1080/07434618912331275096>
- Vojtech, J. M., Hablani, S., Cler, G. J., & Stepp, C. E. (2020). Integrated head-tilt and electromyographic cursor control. *IEEE Transactions on Neural Systems and Rehabilitation Engineering*, 28(6), 1–1. <https://doi.org/10.1109/TNSRE.2020.2987144>
- Yorkston, K. M., Beukelman, D. R., & Hakel, M. (1996). Speech Intelligibility Test for Windows.
- Yorkston, K. M., Beukelman, D. R., Smith, K., & Tice, R. (1990). Extended communication samples of augmented communicators II: Analysis of multiword sequences. *Journal of Speech and Hearing Disorders*, 55(2), 225–230. <https://doi.org/10.1044/jshd.5502.225>