40 Years of Eye Typing: Challenges, Gaps, and Emergent Strategies

ALEESHA HAMID, University of Cambridge, UK
PER OLA KRISTENSSON, University of Cambridge, UK

Gaze interaction enables users to communicate through eye tracking, and is often the only channel of effective and efficient communication for individuals with severe motor disabilities. While there has been significant research and development of eye typing systems, in the context of augmentative and alternative communication (AAC), there is no comprehensive review that integrates the key findings from the variety of aspects that constitute the complex landscape of gaze communication. This paper presents a detailed review and characterization of the literature and aims to consolidate the disparate efforts to provide eye typing solutions for AAC users. We provide a systematic understanding of the components and functionalities that underpin eye typing solutions, and analyze the interplay of the different facets and their role in shaping the user-experience, accessibility, performance, and overall effectiveness of eye typing technology. We also identify the major challenges and highlight several areas that require further research attention.

CCS Concepts: • **Human-centered computing** → **HCI design and evaluation methods**; **Accessibility technologies**; **Accessibility design and evaluation methods**.

Additional Key Words and Phrases: eye typing, AAC, gaze communication, review

1 INTRODUCTION

Text entry is integral in people's lives and even more so when an individual is completely reliant on such a method for communication. Nonspeaking individuals with motor disabilities rely on augmentative and alternative communication (AAC) methods to communicate with their speaking partners. For some users, such AAC devices take the form of eye tracking devices that allow users to type on a keyboard by their gaze. This process is generally called *eye typing*.

It is extremely challenging to design an effective and efficient eye typing system that allows AAC users to achieve their goals. As a result, eye typing has been researched for many years and a review of eye typing called "Twenty years of eye typing: systems and design issues" [67] is now approximately 20 years old. Given the tremendous amount of research since then, and the fact that the research tends to focus on improvements to isolated aspects or specific functional components of an eye typing system, it is timely to review the state-of-the-art and understand current challenges and gaps.

In this paper, we systematically review and characterize the current literature on eye typing. This allows us to paint a picture of the current eye typing technology landscape and highlight challenges and gaps that still require research attention. The foundation of this review is an insight

---

Authors' Contact Information: Aleesha Hamid, University of Cambridge, Cambridge, UK, ah2242@cam.ac.uk; Per Ola Kristensson, University of Cambridge, Cambridge, UK, pok21@cam.ac.uk.

that emerged carrying out this work—successful eye typing requires a *systems perspective*, enabling designers and researchers to think holistically about not only an isolated aspect or individual component of an overall eye typing system but to consider an eye typing system in its entirety, including the needs and preferences of individual AAC users and their speaking partners, their strategies for appropriating the technology, and the unique design opportunities the collective insights accumulated in the field give rise to.

2 METHODOLOGY

2.1 Search Strategy and Study Criteria

In order to assemble a preliminary set of the literature for review, an initial Scopus search was conducted with the following query:

(TITLE-ABS-KEY (eye AND typing) OR TITLE-ABS-KEY (gaze AND communication) AND TITLE-ABS-KEY (augmentative AND alternative AND communication) OR TITLE-ABS-KEY (disabilities))

In Scopus search syntax, TITLE-ABS-KEY refers to the title, abstract and keywords and the logical operators OR and AND were used to narrow down search results. The combination of these search criteria in this query aimed to retrieve documents that are related to eye typing, gaze communication, augmentative and alternative communication, and disabilities. The results were also limited to documents in English. This query returned 358 results on 14th August 2023. The titles and abstracts, and in some cases the full-text, were then reviewed for relevancy. 109 documents of the 358 were found to be relevant to the use of eye typing for AAC users. We then expanded our coverage by carrying out a second Scopus search with the query:

(TITLE-ABS-KEY (eye AND typing) OR TITLE-ABS-KEY (gaze AND typing) OR TITLE-ABS-KEY (gaze AND communication))

This was done to ensure that papers which focused on the design and development of eye typing systems and performance were not excluded if they did not specifically mention the potential application of this technology for AAC users. This query returned 4705 results, out of which 245 were found to be relevant after a manual review. Combining the results of the two search queries and removing the duplicates resulted in a total of 286 documents. The full-texts were then studied and further assessed for eligibility, and 119 documents were included in this review.

While there is a diversity of technology and methods in the field of eye typing, in this review we have excluded papers focusing on Augmented Reality (AR) and Virtual Reality (VR) as well as electrooculography (EOG) or electroencephalogram (EEG) based eye typing methods. The primary objective of this review is to provide an in-depth examination of conventional eye typing methods, specifically those involving eye trackers and on-screen keyboards as the primary means of text or command input. While these alternative methods and technologies are valid and valuable, they represent specific technological approaches that differ from the eye typing methods we are primarily exploring and may have distinct applications, challenges and user groups. Therefore our selection criteria prioritise papers that directly align with our research objectives in order to provide a comprehensive and in-depth overview of this particular area. However, we acknowledge that these technologies present intriguing avenues for future research and exploration in the broader field of assistive technologies and AAC. Researchers and practitioners interested in EOG, EEG, AR, or VR may find value in dedicated literature reviews focusing on those technologies.

2.2 Data Extraction and Analysis

We first organised all qualifying literature into a spreadsheet by metadata (publication type, date, authors, keywords etc.) and relevant extracted information such as research questions, participant

characteristics, research methods, key findings and future research areas. We then conducted an iterative thematic analysis by assigning codes to each paper and clustering them through affinity diagramming using Miro, an online visual collaboration tool. This process was multi-categorical, since each paper could be assigned multiple codes and therefore be relevant to multiple sub-categories. This was because we were interested in understanding and analysing the range and nature of research in terms of technological design and development as well as research and evaluation methods.

2.3 Key Metrics

Text entry rate is commonly measured in words per minute (WPM), where a word is defined as five characters, including space and punctuation. Some studies report characters per minute (CPM) or characters per second (CPS). When we report these metrics they are only provided for indicative purposes as a direct comparison of these quantitative measures is not possible due to the variability in participant characteristics, experimental set up, procedures, and so on. We return to this discussion of variability and heterogeneity of AAC research later in this paper.

The accuracy of eye typing systems is measured using both uncorrected error rates and corrected error rates, where the former is typically measured as the minimum number of character-level insertions, deletions, and substitutions necessary to transform a response text into a stimulus text, divided by the number of characters in the stimulus text.

3 THE EYE TYPING TECHNOLOGY LANDSCAPE

Successful eye typing relies on a system capable of realizing its key functions: eye tracking, gaze estimation, selection methods, keyboard layout and interface design, feedback mechanisms, and issues of language. The aim of this section is to shed light on the intricate interplay of these key functions by explaining their roles and research efforts directed towards them.

3.1 Eye Tracking Technology and Gaze Estimation

Eye tracking technology is a foundational aspect of any eye typing system. Previously, low-tech solutions, such as communication boards, were the only option for enabling communication through gaze for people with motor disabilities. Eye-gaze transfer (e-tran) boards, for example, typically have clusters of letters printed on a transparent plastic board. The board is held up by the communication partner who closely observes the gaze pattern of the AAC user in order to determine what they are trying to convey. The AAC user selects each letter through two eye gestures, one to indicate which cluster of letters contains the target letter, and a second to indicate its position within the group. While such low-tech solutions have been vital in offering a communication option to AAC users, they have several limitations: (1) they result in a low communication bandwidth as users often have to make multiple eye gestures to select a single letter, which can take between 8–12 seconds [94]; (2) they impose a high cognitive load on communication partners, who have to remember which characters have been selected and infer the user's intended words; (3) they do not enable the user to indicate or correct errors; and (4) their design relies on an interpreter and thus limits the privacy and autonomy of the AAC user [8, 107].

High-tech solutions relying on eye tracking technology can counter several of these limitations. The fundamental objective of eye tracking with respect to eye typing systems is to monitor and record the movement and position of the user's gaze in order to bridge the gap between their intent and their desired actions, such as typing text or controlling the interface. Conventional eye typing setups typically rely on remote eye tracking systems which enable the users to interact with an on-screen keyboard without any physical contact with the eye tracking device. While many commercial gaze communication systems use wearable or head-mounted eye trackers, such

solutions are often bulky and heavy and therefore unsuitable for prolonged use as required by people with motor disabilities, and they also restrict their ability to make eye contact or observe the environment around them. Other methods such as electrooculography (EOG) have also been used to track eye movements. It is an obtrusive method as it requires electrodes to be mounted to the user's temples which can lead to discomfort and is thus not practical for extended usage. Brain-computer interfaces (BCIs) that leverage electroencephalogram (EEG) signals have also been used for gaze detection. However, such systems are quite complex, require extensive setups and are often very sensitive to noise, which leads to a lower accuracy and difficulties in utilizing such systems outside controlled environments. While we acknowledge the diversity of methods related to eye tracking, we limit the scope of this review to conventional eye typing systems, specifically those relying on remote eye tracking and on-screen keyboards.

Most commercial eye typing systems use infrared (IR) based eye trackers, as this approach provides a high degree of accuracy. However, these systems tend to be expensive, since this technique requires a high resolution camera and specialised hardware. Moreover, IR based eye trackers are limited to certain environments as they do not work well outdoors or in daylight or under other conditions that interfere with infrared light.

Recently, these drawbacks have given rise to an interest in developing low-cost eye tracking systems that can be used in a range of lighting conditions and environments and in a wider range of devices. Researchers have developed eye typing systems that use ordinary cameras or webcams for eye tracking. Since these systems use visible light instead of infrared light and usually have a lower resolution than commercial eye tracking cameras, the accuracy of this method is generally lower compared to IR based eye trackers [107].

The accuracy of the eye tracking system refers to the offset between the gaze point estimated by the system and the point where the user is actually looking. When this offset is large it becomes difficult to hit very small targets. Most eye tracking systems typically report the spatial resolution of their systems as being between 0.5 and 1 degree visual angle, which is sufficient to hit targets larger than 20×20 pixels at a distance of 50 cm [28]. However, several interactive elements in standard user interfaces tend to be smaller than that. Thus the use of zoom functions and multi-stage selection has been explored as a way to provide access to smaller targets [92], and several applications have been specifically designed to support inaccurate gaze input. For example, GazeTalk [3] uses very large screen buttons, and StarGazer [27] augments the selection method with continuous zoom. San Agustin et al. [86] carried out an evaluation of a low-cost gaze tracking system based on a webcam with built-in infrared light mounted close to the user's eye with the GazeTalk and StarGazer applications, which showed that participants were able to achieve an entry rate of 6.56 WPM, comparable with the 6.26 WPM entry rate achievable with a commercial gaze tracking system [29]. This suggests that such a webcam-based setup can effectively be used to interact with interfaces that have large targets, as the noise tolerance reduces the difference between the performance of different eye trackers. A study that used deliberate miscalibration to understand the threshold of tracking errors found that the GazeTalk and Dasher [100] interfaces were robust to eye tracking accuracy [36]. A study evaluating the accuracy and precision of gaze tracking using a standard video camera reported an average accuracy with sufficient spatial resolution for interaction with applications that have a high tolerance to noise, or with interfaces that can provide zoom functionality as part of the selection [28].

Emerging technologies and algorithms have also made it possible to narrow the performance gap between both types of eye trackers by enabling alternative methods of gaze estimation. The most popular approaches for gaze estimation are 3D model-based or feature-based methods that extract local features of the eye such as the pupil, eye corners etc. However, these approaches require highly accurate feature detection and are prone to errors. Appearance-based methods are

gaining popularity, as they make use of the whole image content of the eyes without explicit local feature extraction. Additionally, they are not based on known parameters of feature extraction, setup and lighting conditions and are therefore more flexible.

Prior work has studied the feasibility of developing an eye typing system using a standard webcam [59] and found an appearance-based method potentially sufficient. Accuracy may be improved by using information on the user's eye movement history [2], which resulted in 16.2 CPS, comparable to a conventional infrared eye-gaze input system [29].

Recent systems have explored deep learning algorithms to improve the accuracy of gaze estimation [54]. Such an approach enabled the design of a text entry method that circumvents the need to estimate gaze angles directly by dividing the gaze into nine directions [106].

Despite these and other efforts, accuracy remains a challenge to be addressed in all gaze based systems, and across all different eye tracking technologies. The lack of accuracy stems not only from hardware limitations but also from biological characteristics of the eye itself— miniature eye movements such as tremors, drifts and microsaccades mean that the eye is rarely absolutely still [22] and this problem is exacerbated for AAC users that tend to have even less gaze control and may exhibit more involuntary movements. Therefore, calibration, which is the process of aligning the system's gaze estimates with the user's actual gaze, is an extremely important part of using eye typing systems. Most eye trackers, IR-based or webcam-based, require calibration before use, which can be time consuming and inconvenient. Further, even with explicit user calibration, eye trackers are still prone to fluctuating accuracy and drift during use, and the systems often require multiple recalibrations [24].

Several approaches have been devised to tackle the issue of extensive calibration. One way is to carry out online calibration of model-based eye trackers to improve the sensitivity of the system towards head movement [76]. Another system is EyeAssist [42], which relies on extracted fixation data only in conjunction with signal processing filters and a unique one-time calibration that is subject and session independent. Yet another approach is the use of pursuit-based interfaces for eye typing [98]. In such systems, the user follows a moving target with their gaze in order to make a selection. Since the action is activated based on the fitting of the relative motion trajectory instead of fixed gaze coordinates on the interface, these systems can be operated without a personal calibration and can therefore overcome issues related to low accuracy of gaze data. Several such pursuit-based text entry systems have been developed, including SMOOVS [61], yielding 2.9 to 3.4 WPM, which was later improved in terms of its entry rate by using word prediction, yielding 4.5 WPM [103]. A later study [102] reported 4.7 WPM. Further variants of pursuit-based text entry have been explored, such as a redesign of SMOOVS [1], EyeTell [6], and SPEye [79]. All report comparable entry rates of 3.41 WPM, 1.27 WPM and 1.15 WPM respectively.

A final aspect to consider in this subsection is that most of the eye typing systems and eye tracking technologies require two-dimensional gaze motion. For certain AAC users, such as those with Locked-In Syndrome (LIS), this is often not possible. Cross et al. [17] designed a one-dimensional interface relying on vertical eye movements alone. Non-disabled participants had an entry rate of 4.7 WPM without prediction, and 11.36 WPM with prediction.

3.2 Selection Methods

Selection is a particularly challenging aspect of gaze-based communication because eyes are primarily sensory organs that are, in this case, also being used for control. One of the fundamental issues in eye typing is that the system must be able to accurately determine whether the user is looking at a specific region to sample sensory information or to activate the key selection. This is often referred to as the Midas touch problem [37]. Traditional eye typing tackles this conundrum by using a dwell timeout: the user must fixate on a desired key for a certain duration in order

to ensure the system knows the user intended to select that specific key. We call this approach dwell-based eye typing.

3.2.1 Dwell-based Eye Typing. While dwell-based eye typing is the most common method for selection, as it is relatively simple to design and implement, the performance is bounded by the dwell timeout duration. As a result, the average typing speed with dwell-based systems is around 5–10 WPM [63]. Reducing dwell times gives rise to a speed-accuracy trade-off: a longer dwell time prevents false selections while a shorter dwell time increases entry rate, but at the cost of more false selections. Dwell times can be straining for the user and result in eye fatigue, as its use forces the eyes to fixate on targets, ignoring the fact that the eyes are primarily sensory organs and not control organs [104]. In addition, dwell times make the act of entering each individual letter or character a high level task in itself, which can disrupt the flow of text entry and is counter-intuitive for users, as writing typically involves transmitting thoughts which are formed as words, sentences and phrases [48].

A longitudinal study investigating the effect of allowing users to adjust the dwell duration found that the average self-selected dwell time of the participants decreased over the course of the sessions, and the text entry rate significantly increased [63]. The participants were able to achieve 19.9 WPM, which is the fastest entry rate reported for a dwell-based eye typing system. The reason for the fast entry rates was the eventual low dwell time that participants were ultimately able to use. Further research later found that it is likely that the dwell time itself is the primary limiting factor towards faster entry rates [82].

Much research has thus focused on exploring mechanisms for mitigating the impact of dwell time. One approach is to adjust the dwell time according to the exit time, which is defined as the time interval between the moment a key is selected and the moment the gaze leaves that key [93]. Another approach is to dynamically adjust the dwell time of each key based on the likelihood that it will be selected next, and its position relative to other likely keys [70]. An evaluation demonstrated that this approach could increase entry rate by 16.67% while maintaining error rate [70]. Another variant of dynamically adjusting dwell time using a probabilistic method increased entry rate at no cost in error rate for both able-bodied users and users with spinal cord injuries [75]. Another variant adjusts dwell times by pupil size [99], as pupil dilation occurs when users are making a decision and shortly after as well, as a reflection of post-decisional consolidation of the selected outcome [91]. Currently, the practical efficacy of this approach is unclear as pupil size is sensitive to factors that may be difficult to control outside an experimental situation [99]. Finally, it is possible to adjust dwell time by adapting dwell times based on entry rate, allowing the system to potentially adapt dwell times to situations when the user is effective or tired [73].

3.2.2 Dwell-free Eye Typing. Dwell-free eye typing has been proposed as a promising alternative to traditional dwell-based typing. In dwell-free eye typing, the user writes by looking at each desired letter in sequence and the system then automatically decodes the continuous eye trace and converts it into the intended text [50]. Given that this approach eliminates the need for dwell timeouts through the use of statistical decoding, it has the potential to be much faster than dwell-based alternatives and this performance potential was experimentally demonstrated through a human performance estimate of 46 WPM using a simulated 'perfect' decoder [50]. A few years later this system design was implemented and released as a commercial system through Tobii-Dynavox [51, 52] and later studied in a deployment study with six users with motor disabilities [48]. This system works similar to a speech recognition system in that it translates observations of gaze points into hypotheses of the intended words using a statistical decoder. It allows the system to both filter out unwanted letter selections and generate words even though not all letters of a word have been explicitly indicated by the user. The user can write a single word, several words, or

sentences as the system does not operate on the word level. When the user gazes at the text area the system presents the decoded text to the user [51, 52].

Several variants of dwell-free eye typing have been proposed in the literature. Filteryedping [74] uses a letter-level filtering approach to filter out unwanted selected letters. Able-bodied participants achieved 15.95 WPM, while participants with disabilities achieved 7.60 WPM. Compared to a dwell-based baseline, it was on average 37% faster and led to fewer errors. A limitation is that it cannot insert missing letters. GazeTry [60] is a variant of a dwell-free system that relies on the Moving Window String Matching algorithm in order to handle missing-letter errors. Another similar approach acknowledges the need for a system to be robust to all common text-entry errors and uses the LCSMapping algorithm [58] to improve performance. EyeSwipe [55] uses word gestures like gesture typing [53, 105] in combination with a reverse crossing technique to allow dwell-free input at the word level. Able-bodied participants had an entry rate of 11.7 WPM. Swipe&Switch [56] switches between different regions to delimit word gestures, thus indicating the start or end of a word and enabling text editing commands. Results indicate a higher text entry rate when compared to EyeSwipe. GlanceWriter [18] is designed to circumvent the need for extra operations to mark the start and end of a word, which is a common issue among word-based dwell-free systems. This system probabilistically determines the intended letters and uses gaze dynamics to infer the starting and ending characters. The performance of this system increased from 6.49 WPM to 10.89 WPM compared to EyeSwipe and reduced errors.

Sarcar et al. [88] proposed EyeK, which enables dwell-free typing on a keyboard by inferring character selection through a particular interaction gesture whereby the user must look within the border of their desired key, then outside it, and then back inside it. With this approach, users were able to reach text entry rates up to 8.8 WPM.

Wobbrock et al. [101] created EyeWrite, a system that allows users to enter text through gaze gestures among the four corners of an on screen square, and Isokoski [34] proposed a method based on the Minimal Device Independent Test Input Method (MDITIM) stroke alphabet, where users gaze at off-screen targets placed at the edge of the screen. This particular approach was designed to mitigate false selections as a result of the Midas Touch problem, while also helping to conserve the display area required by the keyboard. Eye gestures were also utilized in the Eye-S [80] system, in which letters are drawn through sequences of fixations on specific parts of the screen called hotspots. Morimoto and Amir [69] introduced the concept of context-switching, through the duplication of keyboard. pEYEwrite [95] is a hierarchical pie menu system that allows users to enter text by gazing at regions of the menu that contain their intended letter.

Saccadic latency, a 200–250 ms duration that precedes the initiation of a saccade has also been explored as a mechanism for selection [45]. After the user triggers a fixation, the system works backwards and locates the point of fixation during saccadic latency, meaning the intended character is the one that the user was looking at right before the saccade. An initial study indicates an average typing speed of 27.1 characters per minute (CPM) for able-bodied users.

It has been observed that some more efficient systems may require a dedicated training period [78] or may increase cognitive load to maintain a high text entry rate [31]. The possibility of combining dwell-free and dwell-based approaches has been explored through a multi-selection technique [5] whereby the user is only required to fixate on the first and last letter of each word. A comparison of this technique with a dwell-based system showed that novices were able to achieve 63% faster text entry rates, without any influence on the error rate.

3.3 Keyboard Layout and Interface Design

Both the QWERTY and the alphabetical keyboard layouts are affected by the *centrality bias*, which means keys located in the center of the display have an increased susceptibility to accidental

selection. This occurs as a result of the natural tendency to look towards the center of the screen [10]. As a way of mitigating this issue, some eye typing systems use circular keyboard layouts so that all the keys are equidistant from the central fixation point, and therefore equally easy to access and select. Familiarity with keyboard shape and the presence of feedback significantly influence user performance [77].

Eyeboard [73] is a gaze based text entry system where letters are arranged in two zones, with the central zone containing the most frequently occurring letters and characters in order to reduce visual search time. It yielded 8.25 WPM in a study, and comparison with existing interfaces. pEYEs [33] is a dwell-free text entry system where letters are grouped into sectors of a pie and letter are selected by first selecting a letter group and thereafter an intended letter. The keys can also use a hexagonal arrangement of clusters and tiles [1] [61]. Another variant of two-stage selection is to manage selection access through a tree structure and arrange letter keys around a rectangular text output area [16]. Yet another variant is SliceType [9] which does not rely on hierarchical selection but instead uses word prediction to merge and reallocate space to the letter keys that are likely to be selected. The two-stage selection text entry method Side Keyboard [30] places triangular keys to the left and right sides of the text box, with the left side keys being letter groups which, upon selection, trigger the opening of the right panel that contains the letters from the group as individual keys. This design makes it possible for the user to correct errors without having to select a separate key to go back to the previous selection stage. Systems such as Dasher [100] and StarGazer [27] use dynamic interfaces to facilitate their dwell-free selection mechanisms and are therefore also able to overcome space limitations on the screen and manage a noisy eye tracking signal. GazeTalk [29] overcomes the limitations of inaccurate gaze data by only having 12 active buttons on the display, enabling possible use with low-resolution hardware.

Another interface challenge is word and phrase prediction, which may force the user to frequently shift their gaze from the keyboard to a list of suggestions, thereby possibly reducing efficiency and increasing cognitive load [62]. This observation can motivate designs which allow the user an easy way to navigate back and forth among predictions [68]. However, Sengupta et al. [90] found that bringing the suggestions closer to the visual attention of the user did not have a significant impact on text entry performance. AVIN [108] is a dwell based eye typing notepad application designed with a three-layer layout, organized as two rings surrounding a central text box. The outer ring contains the letters of the alphabet, arranged in alphabetical order, and the inner ring displays a continuously updated set of word suggestions. Finally, a study has shown that it is possible to significantly reduce AAC users' selection times by presenting screen elements at users' preferred locations of the display [38].

3.4 Feedback Mechanisms

AugKey [20] improves visual throughput in eye typing systems by augmenting keys with a prefix to allow continuous text inspection and hence faster error detection. This design also uses augmented suffixes to show the three most probable words that will appear in the word prediction list if the focused letter is selected, thereby helping the user to decide if the next selection will be another key or if they will switch their gaze to the word list. Experimental results comparing AugKey to dwell-based keyboards without augmented feedback, with and without prediction, showed that it improved text-entry rates in both conditions.

In a hybrid interface combining dwell time and pursuit movement [102] as well as in solely dwell-based systems [66], the highest text-entry rate was achieved for a combination of visual and auditory feedback, as compared to other conditions such as just visual feedback, just auditory feedback, or no feedback. Interestingly, this finding is only partially applicable when considering dwell-based systems with short dwell times [64]. In this case, spoken feedback results in slower

text entry rates and more errors due to spoken feedback varying in duration depending on the selected letter, disrupting the typing rhythm. However, vibrotactile feedback can produce results close to that of auditory feedback for dwell-based eye typing [65].

However, in eye typing systems that use gaze gestures or pursuit based interfaces, the different feedback modalities such as visual, haptic or auditory were comparable in terms of task completion time, error rate and user experience [40, 44]. One study also indicates that the keyboard layout may affect the efficacy of visual feedback [77].

Different variations and placements of visual feedback have also been studied to understand the role of feedback in mitigating specific usability issues, such as how a system responds to gaze aversion [25]. Another use of feedback exploits phenomena known as the gap effect and overlap condition [19]. The gap effect consists of a reduction in the mean saccadic latency when the visual stimulus at the current fixation point is removed before the presentation of a second stimulus at a different location. Conversely, if the first visual stimulus is maintained even after the second stimulus is presented, the mean latency increases. A possible explanation for this gap-overlap effect is that the disappearance of the first stimulus helps the user disengage their attention and move their gaze to the new stimulus faster and therefore it is more difficult to shift gaze if the first stimulus is maintained. Results from a pilot experiment demonstrated that a feedback mechanism based on the gap-overlap effect improved eye typing performance and user experience.

3.5 Language

While the majority of work in eye typing has been done for text entry in English, the existing research efforts directed towards implementing on-screen keyboards in other languages show that this requires specific understanding of the language characteristics which results in design considerations that cut across the different functional components.

The current method of using standard keyboard layouts such as QWERTY to input text in other languages typically requires multiple key presses to enter single characters due to larger character sets, which results in slow text entry and higher cognitive load. Arranging the keys according to the frequency of occurrence in the particular language can improve text-entry rates, as seen through the evaluation of dedicated keyboards for Vietnamese [71], Brazilian Portugese [83], Hindi [87] and Japanese [72].

4 CHALLENGES AND GAPS

4.1 Research Methods

While there has been a significant amount of work on the development of eye typing systems for AAC users, in most cases their evaluation has been based on studies with able-bodied participants. This is often due to problems accessing and recruiting AAC users and other factors, such as their lack of mobility, or the fatigue associated with certain disabilities and health conditions that make the evaluation process impractical. As a result, researchers are faced with a trade-off between doing any evaluation at all and reduced ecological validity. Previous discussion of user representation in accessibility research has highlighted numerous instances where certain insights have been overlooked or inaccurate conclusions have been drawn as a result of studying non-representative users [89]. Hornof and Cavender [32] noted the importance of their evaluation with users with cerebral palsy to understand the extent of difference in capabilities between them and the able-bodied participants in their study. Istance et al. [35] carried out two user studies using gaze gestures and dwell time selections to investigate the extent to which the performance of able-bodied participants could be representative of people with physical disabilities. Their studies included three groups of participants: (1) cerebral palsy (CP); (2) muscular dystrophy (MD); and

(3) a control group of able-bodied participants. Across both tasks, the groups of participants with disabilities performed significantly different than the control group, highlighting the risk that eye typing research can lack ecological validity if potential solutions are not evaluated with the target user group.

In some cases, studies that rely primarily on able-bodied participants also attempt to verify the findings in some way with members of the actual target group, although this is often through informal case studies with individuals [70], or very short and limited sessions in controlled lab settings. While this kind of engagement with the target user group can lead to useful feedback, long-term studies in deployment environments are likely necessary to ensure eye typing systems are fit for purpose for real-world use by AAC users. For instance, factors that may influence the text entry rate and overall performance, such as fatigue or drift in eye tracking calibration over time, may be misunderstood in short term studies. In the general text entry domain this need to "walk the last mile" [46] has been noted. It has also been observed through research that has explicitly studied (non-eye typing) text entry with able-bodied users [13, 43, 84] that real-world, or "in the wild", studies have led to unique insights that could not be obtained from lab studies alone.

In the eye typing domain, a recent deployment study [48] on a commercial implementation of dwell-free eye typing [50] reports results from in situ use of dwell-free eye typing with six AAC users completely reliant on eye typing for communication in their home. This study allowed the identification of several barriers to effective and efficient use of dwell-free eye typing, and yielded eight design implications for ensuring such systems are fit for purpose when deploying to AAC users.

Carrying out long-term studies of everyday eye typing use by AAC users would also enable the collection of data for a corpus of longitudinal, real-world AAC data that would greatly benefit researchers in this field as it would capture important trends, patterns and contextual information. Given the challenges in carrying out such studies in an ethical and secure way, researchers have so far put together a simulated AAC corpora of conversational messages through crowdsourcing [97]. Corpora from the target group would increase the extent to which this data can be representative. A protocol and open source tool named the SpeakFaster Observer [14] has recently been presented to allow measurements of everyday conversational text entry by gaze typing user. The initial case study with an AAC user with ALS and consenting conversational partners yielded a rich dataset, providing insights into entry rate, keystroke savings, patterns of utterance repetition and reuse, and the temporal dynamics of conversational turn-taking in gaze communication.

4.2 Performance Metrics and Usability Criteria

There is an underlying assumption that the performance metrics and usability criteria that have been considered important when studying text entry systems with able-bodied users are equally appropriate for eye typing systems designed for AAC users. Hence, many eye typing systems are evaluated with the primary objective of increasing throughput, or entry rate, typically while maintaining a low error rate, rather than improving the level of engagement of communication partners, the ease with which a system can be used or adapted for use by particular individuals, the fatigue of the AAC user, or their perceptions of autonomy or concerns for privacy.

This approach is reflective of the tendency to conceptualize disability through a medical perspective, which leads to the design of assistive technology that focuses solely on the functional limitations of the AAC user without accounting for the social aspect of interpersonal communication and other circumstantial or environmental barriers to inclusion. The design of AACrobat [26] is an inspiring example of research motivated by a social model of disability instead—the system is designed as a groupware system where all communicators share the burden of facilitating effective communication. The design guidelines synthesized from a formative study with users with ALS and

their communication partners are broadly related to user autonomy, engaging the communication partners, and user privacy.

Existing research supports the need to reframe current evaluation methods to consider the entire range of factors that can influence the efficacy of eye typing solutions for AAC users. For example, there is significant evidence of the positive impact of partner instruction and on the communication of AAC users [12, 41, 85], and findings that indicate that an effective AAC system requires a commitment from all social partners [7], training, and regular use of the device [11]. Through interviews with people with ALS and their communication partners, Kane et al. [39] found that specific communication difficulties, such as conversational pacing or personality expression, can also arise as a result of the transition to using an AAC device, which is often the case as a result of the gradual progression of diseases like ALS.

There is an opportunity to develop and use metrics that are specific to text entry through gaze input, in addition to standard metrics such as words per minute or character error rates. Currently, there is relatively little research on metrics that directly include the gaze behavior itself. Majaranta et al. [64] used a metric termed "number of Read Text Events per character", which measured the frequency with which a user directed their gaze towards the text field. While this metric does not directly relate to standard metrics, such as accuracy, there is a direct link as people are more likely to look at the typed text when they are uncertain about its accuracy. Aoki et al. [4] later suggested the use of the metric "attended keys per character", which measures the number of keys attended for each typed character. This metric is based on guidelines for efficient manual work that works on the principle of conserving human energy—although the energy cost of a single eye movement may be small, the cumulative cost of all unnecessary fixations may be high, and this cumulative effect can have an impact on user experience and workload. The attended keys per character metric was highly correlated to erroneous selections, since whenever a key is accidentally fixated upon there is a risk of such a fixation becoming a false activation as a result of the Midas touch problem.

4.3 Variability and Heterogeneity of Eye Typing Research

We observe that studies of eye typing systems for AAC users tend to be characterized by a significant degree of heterogeneity, as each study has a distinct cohort of participants, choice of evaluation metrics and particular research objectives. While each individual study contributes valuable insights, it is very difficult to synthesize and compare findings in order to identify broadly applicable robust design principles or strategies to guide future work.

Even within the set of studies conducting evaluations with able-bodied proxy-users there is a lot of variation in participant characteristics which can influence outcomes. For example, one study with able-bodied participants found that age had a significant effect on eye typing performance—as age increased, the typing speed decreased and error rate increased [15].

The procedure and duration of the studies can also influence performance as several studies have observed a clear learning effect with eye typing systems, as user performance increases over time and with more repetitions [15]. It has been noted that this learning effect could represent the reinforcement of motivation, which has been shown to be an important factor impacting communication among AAC users [57]. Familiarity with keyboard layout can also positively impact performance [77]. In contrast, learning a novel or custom layout introduces a steep learning curve and may be daunting prospect for an AAC user.

The hardware and physical set up is frequently different in studies that in addition use different eye trackers and set ups, and some studies physically restrict head movement while in other studies the researchers merely instruct the participants to stay as still as possible. As one example of the effect such changes can have, Mott et al. [70] report that the error rates obtained were higher compared to the error rates of participants in previous studies, and the authors noted that this

could be attributed to the fact that they used a smaller sized screen in their experiments, which can lead to more errors as a result of poor eye tracking calibration [81].

The specific research objectives of each individual study are also inconsistent. One study may be designed to investigate the effect of modifying a particular component in isolation on overall performance. However, this carries the implicit assumption that components are not coupled in the system. In practice, they frequently are, making it impossible to generalize such results because of the interplay between the many factors that gives rise to an overall eye typing system's behavior. Kristensson and Müllners [49] note a similar issue in the design of intelligent text entry systems, and termed this research error "short-circuit evaluation". Such a system design error occurs when the merits of a single design choice are determined through an evaluation that uses a complete system without fully understanding the implications of each individual design parameter on overall system performance. For example, a study that is designed to investigate the effect of adjustable dwell time by comparing two systems would typically control for other variables, such as keyboard layout, feedback mechanism, and so on across their experiments. While this is done to ensure a high level of internal validity, it limits the extent to which the quantitative findings can be used as a benchmark and their broader applicability for the purpose of building better systems, as such controlled factors have a tendency to change when people build new systems.

4.4 Diversity of User Needs

The target user group of people with disabilities is far from a homogenous group. A study assessing eye typing performance of participants with Cerebral Palsy and Muscular Dystrophy demonstrated that there is a significant difference in performance between the two groups [35]. There is also a large amount of variation between individual ability and limitations, even within the same nominal category of disability [21]. Even within the same individual the abilities may change over time and may fluctuate as a result of other factors such as time of day, level of fatigue, and other factors relating to their specific health conditions.

Subjective feedback on user experience and preferences also exhibits a high degree of variation. Qualitative findings from an extensive study with almost five hours of eye typing per participant reveal large discrepancies in the way individual users experienced the causes of workload [82]. For some participants the feelings of frustration were the most significant contributors to the experienced workload while other participants reported temporal demand or cognitive demand and the intensive focus required of the task as the primary contributor. It was also observed that there was a significant variation when participants were asked to rank their level of eye fatigue. The study also found quantiative individual differences as a result of user strategy.

Given the diversity in users' needs and preferences, and the influence of an individual user's strategy on outcomes, eye typing systems should support user control of parameters and features, as well as adaptation, in order to improve overall user experience and performance by tailoring it for the individual.

4.5 Towards a Systems Approach to Eye Typing

An emerging insight from this review is the lack of a coherent systems approach that takes into account the many factors that govern the eye typing experience and links them through all stages of the design, ranging from individual concepts, developing new technologies, designing new interaction mechanisms, and ultimately ensuring deployed eye typing systems allow AAC users to achieve their goals. We believe it may be useful to learn from systems engineering, sometimes called systems thinking [23], which give rise to insights such as (1) systems are complex—they have many dependencies and some of them include the user and any speaking partner; (2) systems operate at multiple levels, ranging from sensing gaze, inferring fixations, and calibration, but also the fact
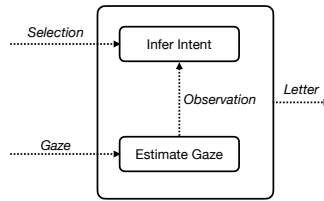
Fig. 1. An example of a high-level function structure model for an eye typing system.

that users desire their eye typing system to allow them to seamlessly interact on social media and enable them to fill out forms; (3) systems are tightly coupled—it is frequently not sufficient to study individual components in isolation; and (4) systems give rise to emergent properties that can only be understood by modeling the entire system.

As a practical indication of what this may look like for eye typing, in design engineering [47, 49] a design is viewed as an operating point in a multidimensional design space. Setting an operating point thus means making a series of trade-off decisions. Typically there is no optimal operating point, or optimal design, as many design considerations trade off against each other. However, there is an optimal set of trade-off decisions. One useful way to guide the design, at an early stage, is to make the role of the functions in a system explicit. Note that a *function* (*what* we need to carry out) is distinct from a *function carrier* (*how* we are going to do it).

An example of a high-level function structure for the function Type Key is shown in Figure 1. It decomposes this function into two key subfunctions: Infer Intent and Estimate Gaze. The dashed arrows represent *signals* that show the flow of information between functions. In this case Type Key receives two signals, *Selection* and *Gaze*, and responds by outputting the signal *Letter*. Importantly, no design decision at this stage has been taken on *how* to realize these functions or represent the signals. Yet, we can still parameterize this function structure into controllable and uncontrollable parameters. Controllable parameters are design parameters we can optimize and tune. For example, Infer Intent may have parameters such as the dwell timeout, the design of the dwell timeout and so on. Uncontrollable parameters are parameters that affect system outcomes but are not under the direct influence of the designer. For example, the user's strategy [49] when entering text cannot be directly controlled. By such function modeling and parameterization it is possible to carry out computational studies to tease out design implications that show how functions and parameters interrelate *before* carrying out user studies. As has been previously pointed out [47], this allows researchers to go into user studies *knowing what to look for*, which is particularly useful when carrying out studies with AAC users to ensure such observations are as informative as possible.

5 CONCLUSION: REFLECTIONS AFTER TWENTY YEARS

It is fascinating to note how the research into eye typing has progressed since the influential review "Twenty years of eye typing: systems and design issues" [67] was published over two decades ago. "Twenty years of eye typing" highlights the issues of eye tracking accuracy, extensive calibration, interface design, and feedback as areas requiring further attention. While much of the work in the following years has addressed these aspects, the relative pace of progress in different functional components of eye typing systems have not exhibited the same rate. For example, in terms of feedback, there has been considerably more research into the effect of different feedback modalities on performance in dwell-based systems and not so much on dwell-free systems using gestures or smooth-pursuit movements. What seems to be lacking is a systems approach that links together the
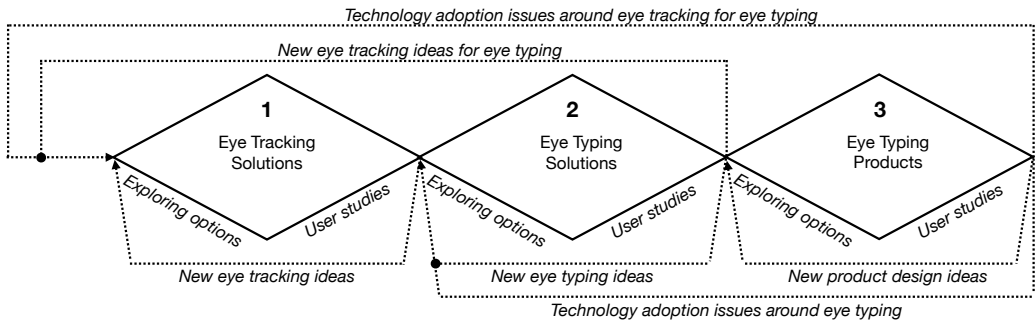
Fig. 2. A process model of field and industry progress in realizing eye typing products for users with motor disabilities.

individual components into a system, including AAC users and their speaking partners, that can be studied at the point of deployment to give rise to design implications and solution principles that enable us to both ideate fruitful ideas for new interaction mechanisms and solutions to existing barriers and challenges an AAC user faces in their daily lives with eye typing.

Figure 2 illustrates the process of realizing an eye typing system as three interrelated processes of divergence (exploring options) and convergence (narrowing down options through user studies). Phase 1 and Phase 2 have received extensive attention, however, Phase 3 has not been explored in much detail. This results in the feedback loops from Phase 3 to Phase 1 and 2 to be weak, illustrating a potential disconnect between the lived experience of AAC users relying on eye typing and the fundamental research activities that attempt to provide solutions.

In conclusion, there has been tremendous progress in eye typing in the last 20 years. While some of the challenges that existed 40 years ago still remain, we are optimistic that progress will continue. Among all the possible research trajectories highlighted in this review, we suggest some may warrant particular attention. First, exploring a systems approach to eye typing can enable us to more effectively ensure technology and interaction ideas relate to AAC users' current barriers and challenges. Second, the rapid advancement of generative AI, such as the use of large language models, represents both an opportunity and a challenge in the field of eye typing for AAC. These technologies may be able to improve entry rates and lower cognitive load and fatigue, but will also have a profound impact on the way users interact with AAC devices and their perceptions of autonomy and self-expression [96].

REFERENCES

[1] Yasmeen Abdrabou, Mariam Mostafa, Mohamed Khamis, and Amr Elmougy. 2019. Calibration-free text entry using smooth pursuit eye movements. In *Proceedings of the 11th ACM Symposium on Eye Tracking Research & Applications*. 1–5.

[2] Kiyohiko Abe, Shoichi Ohi, and Minoru Ohyama. 2007. An eye-gaze input system using information on eye movement history. In *Universal Access in Human-Computer Interaction. Ambient Interaction: 4th International Conference on Universal Access in Human-Computer Interaction, UAHCI 2007 Held as Part of HCI International 2007 Beijing, China, July 22-27, 2007 Proceedings, Part II 4*. Springer, 721–729.

[3] Hirotaka Aoki, John Paulin Hansen, and Kenji Itoh. 2008. Learning to interact with a computer by gaze. *Behaviour & Information Technology* 27, 4 (2008), 339–344.

[4] Hirotaka Aoki, John Paulin Hansen, and Kenji Itoh. 2009. Learning gaze typing: what are the obstacles and what progress to expect? *Universal Access in the Information Society* 8 (2009), 297–310.

[5] Tanya Bafna. 2018. Gaze typing using multi-key selection technique. In *Proceedings of the 20th International ACM SIGACCESS Conference on Computers and Accessibility*. 477–479.

[6] Tanya Bafna, Per Bækgaard, and John Paulin Paulin Hansen. 2021. EyeTell: Tablet-based Calibration-free Eye-typing using Smooth-pursuit movements. In *ACM Symposium on Eye Tracking Research and Applications*. 1–6.

[7] Rita L Bailey, Howard P ParetteJr, Julia B Stoner, Maureen E Angell, and Kathleen Carroll. 2006. Family members' perceptions of augmentative and alternative communication device use. (2006).

[8] Richard Bates, Michael Donegan, Howell O Istance, John Paulin Hansen, and K-J Räihä. 2007. Introducing COGAIN: communication by gaze interaction. *Universal Access in the Information Society* 6 (2007), 159–166.

[9] Burak Benligiray, Cihan Topal, and Cuneyt Akinlar. 2019. Slicetype: fast gaze typing with a merging keyboard. *Journal on Multimodal User Interfaces* 13 (2019), 321–334.

[10] Darrell S Best and Andrew T Duchowski. 2016. A rotary dial for gaze-based pin entry. In *Proceedings of the ninth biennial acm symposium on eye tracking research & applications*. 69–76.

[11] David Beukelman. 1991. Magic and cost of communicative competence. *Augmentative and alternative communication* 7, 1 (1991), 2–10.

[12] Maria Borgestig, Jan Sandqvist, Richard Parsons, Torbjörn Falkmer, and Helena Hemmingsson. 2016. Eye gaze performance for children with severe physical impairments using gaze-based assistive technology—A longitudinal study. *Assistive technology* 28, 2 (2016), 93–102.

[13] Daniel Buschek, Benjamin Bisinger, and Florian Alt. 2018. ResearchIME: A mobile keyboard application for studying free typing behaviour in the wild. In *Proceedings of the 2018 CHI Conference on Human Factors in Computing Systems*. 1–14.

[14] Shanqing Cai, Subhashini Venugopalan, Katrin Tomanek, Shaun Kane, Meredith Ringel Morris, Richard Cave, Robert Macdonald, Jon Campbell, Blair Casey, Emily Kornman, et al. 2023. SpeakFaster Observer: Long-Term Instrumentation of Eye-Gaze Typing for Measuring AAC Communication. In *Extended Abstracts of the 2023 CHI Conference on Human Factors in Computing Systems*. 1–8.

[15] Marco Caligari, Marica Giardini, Ilaria Arcolin, Marco Godi, Stefano Corna, and Roberto Colombo. 2021. Writing with the eyes: the effect of age on eye-tracking performance in non-disabled adults and a comparison with bimanual typing. *Computational intelligence and neuroscience* 2021 (2021).

[16] Hubert Cecotti, Yogesh Kumar Meena, and Girijesh Prasad. 2018. A multimodal virtual keyboard using eye-tracking and hand gesture detection. In *2018 40th Annual international conference of the IEEE engineering in medicine and biology society (EMBC)*. IEEE, 3330–3333.

[17] Michael Cross, Leping Qiu, Mingyuan Zhong, Yuntao Wang, and Yuanchun Shi. 2022. One-Dimensional Eye-Gaze Typing Interface for People with Locked-in Syndrome. In *Adjunct Proceedings of the 35th Annual ACM Symposium on User Interface Software and Technology*. 1–3.

[18] Wenzhe Cui, Rui Liu, Zhi Li, Yifan Wang, Andrew Wang, Xia Zhao, Sina Rashidian, Furqan Baig, IV Ramakrishnan, Fusheng Wang, et al. 2023. GlanceWriter: Writing Text by Glancing Over Letters with Gaze. In *Proceedings of the 2023 CHI Conference on Human Factors in Computing Systems*. 1–13.

[19] Antonio Diaz Tula, Andrew TN Kurauchi, and Carlos H Morimoto. 2013. Facilitating gaze interaction using the gap and overlap effects. In *CHI'13 Extended Abstracts on Human Factors in Computing Systems*. 91–96.

[20] Antonio Diaz-Tula and Carlos H Morimoto. 2016. Augkey: Increasing foveal throughput in eye typing with augmented keys. In *Proceedings of the 2016 CHI conference on human factors in computing systems*. 3533–3544.

[21] Mick Donegan, Jeffrey D Morris, Fulvio Corno, Isabella Signorile, Adriano Chió, Valentina Pasian, Alessandro Vignola, Margret Buchholz, and Eva Holmqvist. 2009. Understanding users and their needs. *Universal Access in the Information Society* 8 (2009), 259–275.

[22] T Andrew Duchowski. 2017. *Eye tracking: methodology theory and practice.* Springer.

[23] Chris Elliott and Peter Deasley. 2007. Creating systems that work: Principles of engineering systems for the 21st century. *R Acad Eng* 293074 (2007).

[24] Anna Maria Feit, Shane Williams, Arturo Toledo, Ann Paradiso, Harish Kulkarni, Shaun Kane, and Meredith Ringel Morris. 2017. Toward everyday gaze input: Accuracy and precision of eye tracking and implications for design. In *Proceedings of the 2017 Chi conference on human factors in computing systems*. 1118–1130.

[25] Misahael Fernandez, Florian Mathis, and Mohamed Khamis. 2021. GazeWheels: Recommendations for using wheel widgets for feedback during dwell-time gaze input. *it-Information Technology* 63, 3 (2021), 145–156.

[26] Alexander Fiannaca, Ann Paradiso, Mira Shah, and Meredith Ringel Morris. 2017. AACrobat: Using mobile devices to lower communication barriers and provide autonomy with gaze-based AAC. In *Proceedings of the 2017 ACM Conference on Computer Supported Cooperative Work and Social Computing*. 683–695.

[27] Dan Witzner Hansen, Henrik HT Skovsgaard, John Paulin Hansen, and Emilie Møllenbach. 2008. Noise tolerant selection by gaze-controlled pan and zoom in 3D. In *Proceedings of the 2008 symposium on Eye tracking research & applications*. 205–212.

[28] John Paulin Hansen, Javier San Agustin, and Henrik Skovsgaard. 2011. Gaze interaction from bed. In *Proceedings of the 1st Conference on Novel Gaze-Controlled Applications*. 1–4.

[29] John Paulin Hansen, Kristian Tørning, Anders Sewerin Johansen, Kenji Itoh, and Hirotaka Aoki. 2004. Gaze typing compared with input by head and hand. In *Proceedings of the 2004 symposium on Eye tracking research & applications*. 131–138.

[30] Katarzyna Harezlak, Pawel Basek, and Pawel Kasprowski. 2022. Side Keyboard–the New Approach for Eye-typing. *Procedia Computer Science* 207 (2022), 3348–3357.

[31] Takahiro Hayashi and Reo Kishi. 2014. Workload Evaluation of Gaze-Writing Systems. *Sensors and Materials* 26, 7 (2014), 529–538.

[32] Anthony J Hornof and Anna Cavender. 2005. EyeDraw: enabling children with severe motor impairments to draw with their eyes. In *Proceedings of the SIGCHI conference on Human factors in computing systems*. 161–170.

[33] Anke Huckauf and Mario H Urbina. 2008. Gazing with pEYEs: towards a universal input for various applications. In *Proceedings of the 2008 symposium on Eye tracking research & applications*. 51–54.

[34] Poika Isokoski. 2000. Text input methods for eye trackers using off-screen targets. In *Proceedings of the 2000 symposium on Eye tracking research & applications*. 15–21.

[35] Howell Istance, Stephen Vickers, and Aulikki Hyrskykari. 2012. The validity of using non-representative users in gaze communication research. In *Proceedings of the Symposium on Eye Tracking Research and Applications*. 233–236.

[36] Kenji Itoh, Hirotaka Aoki, and John Paulin Hansen. 2006. A comparative usability study of two Japanese gaze typing systems. In *Proceedings of the 2006 symposium on Eye tracking research & applications*. 59–66.

[37] Robert JK Jacob. 1991. The use of eye movements in human-computer interaction techniques: what you look at is what you get. *ACM Transactions on Information Systems (TOIS)* 9, 2 (1991), 152–169.

[38] DV Jeevithashree, Kamalpreet Singh Saluja, and Pradipta Biswas. 2019. A case study of developing gaze controlled interface for users with severe speech and motor impairment. *Technology and Disability* 31, 1-2 (2019), 63–76.

[39] Shaun K Kane, Meredith Ringel Morris, Ann Paradiso, and Jon Campbell. 2017. " At times avuncular and cantankerous, with the reflexes of a mongoose" Understanding Self-Expression through Augmentative and Alternative Communication Devices. In *Proceedings of the 2017 acm conference on computer supported cooperative work and social computing*. 1166–1179.

[40] Jari Kangas, Oleg Špakov, Poika Isokoski, Deepak Akkil, Jussi Rantala, and Roope Raisamo. 2016. Feedback for smooth pursuit gaze tracking based control. In *Proceedings of the 7th Augmented Human International Conference 2016*. 1–8.

[41] Jennifer Kent-Walsh and David Mcnaughton. 2005. Communication partner instruction in AAC: Present practices and future directions. *Augmentative and alternative communication* 21, 3 (2005), 195–204.

[42] Anwesha Khasnobish, Rahul Gavas, Debatri Chatterjee, Ved Raj, and Sapna Naitam. 2017. EyeAssist: A communication aid through gaze tracking for patients with neuro-motor disabilities. In *2017 IEEE International Conference on Pervasive Computing and Communications Workshops (PerCom Workshops)*. IEEE, 382–387.

[43] Andreas Komninos, Mark Dunlop, Kyriakos Katsaris, and John Garofalakis. 2018. A glimpse of mobile text entry errors and corrective behaviour in the wild. In *Proceedings of the 20th International Conference on Human-Computer Interaction with Mobile Devices and Services Adjunct*. 221–228.

[44] Anne Köpsel, Päivi Majaranta, Poika Isokoski, and Anke Huckauf. 2016. Effects of auditory, haptic and visual feedback on performing gestures by gaze or by hand. *Behaviour & Information Technology* 35, 12 (2016), 1044–1062.

[45] Kentaro Kotani, Yuji Yamaguchi, Takafumi Asao, and Ken Horii. 2010. Design of eye-typing interface using saccadic latency of eye movement. *International Journal of Human-Computer Interaction* 26, 4 (2010), 361–376.

[46] Per Ola Kristensson. 2009. Five challenges for intelligent text entry methods. *AI Magazine* 30, 4 (2009), 85–85.

[47] Per Ola Kristensson, James Lilley, Rolf Black, and Annalu Waller. 2020. A design engineering approach for quantitatively exploring context-aware sentence retrieval for nonspeaking individuals with motor disabilities. In *Proceedings of the 2020 CHI Conference on Human Factors in Computing Systems*. 1–11.

[48] Per Ola Kristensson, Morten Mjelde, and Keith Vertanen. 2023. Understanding Adoption Barriers to Dwell-Free Eye-Typing: Design Implications from a Qualitative Deployment Study and Computational Simulations. In *Proceedings of the 28th International Conference on Intelligent User Interfaces*. 607–620.

[49] Per Ola Kristensson and Thomas Müllners. 2021. Design and analysis of intelligent text entry systems with function structure models and envelope analysis. In *Proceedings of the 2021 CHI Conference on Human Factors in Computing Systems*. 1–12.

[50] Per Ola Kristensson and Keith Vertanen. 2012. The potential of dwell-free eye-typing for fast assistive gaze communication. In *Proceedings of the symposium on eye tracking research and applications*. 241–244.

[51] Per Ola Kristensson, Keith Vertanen, and Morten Mjelde. 2018. Gaze based text input systems and methods. US Patent 10,082,864.

[52] Per Ola Kristensson, Keith Vertanen, and Morten Mjelde. 2020. Gaze based text input systems and methods. US Patent 10,551,915.

[53] Per-Ola Kristensson and Shumin Zhai. 2004. SHARK2: a large vocabulary shorthand writing system for pen-based computers. In *Proceedings of the 17th annual ACM symposium on User interface software and technology*. 43–52.

[54] Keiichiro Kubo, Satoru Shibata, Tomonori Karita, Tomonori Yamamoto, and Shenglin Mu. 2022. Eye-interface system using convolutional neural networks for people with physical disabilities. In *4th EAI International Conference on Robotic Sensor Networks*. Springer, 81–91.

[55] Andrew Kurauchi, Wenxin Feng, Ajjen Joshi, Carlos Morimoto, and Margrit Betke. 2016. EyeSwipe: Dwell-free text entry using gaze paths. In *Proceedings of the 2016 chi conference on human factors in computing systems*. 1952–1956.

[56] Andrew Kurauchi, Wenxin Feng, Ajjen Joshi, Carlos H Morimoto, and Margrit Betke. 2020. Swipe&switch: Text entry using gaze paths and context switching. In *Adjunct Proceedings of the 33rd Annual ACM Symposium on User Interface Software and Technology*. 84–86.

[57] Janice Light and David McNaughton. 2014. Communicative competence for individuals who require augmentative and alternative communication: A new definition for a new era of communication? , 18 pages.

[58] Yi Liu, Bu-Sung Lee, and Martin J McKeown. 2016. Robust eye-based dwell-free typing. *International Journal of Human–Computer Interaction* 32, 9 (2016), 682–694.

[59] Yi Liu, Bu Sung Lee, Andrzej Sluzek, Deepu Rajan, and Martin Mckeown. 2016. Feasibility analysis of eye typing with a standard webcam. In *Computer Vision–ECCV 2016 Workshops: Amsterdam, The Netherlands, October 8-10 and 15-16, 2016, Proceedings, Part II 14*. Springer, 254–268.

[60] Yi Liu, Chi Zhang, Chonho Lee, Bu-Sung Lee, and Alex Qiang Chen. 2015. Gazetry: Swipe text typing using gaze. In *Proceedings of the annual meeting of the australian special interest group for computer human interaction*. 192–196.

[61] Otto Hans-Martin Lutz, Antje Christine Venjakob, and Stefan Ruff. 2015. SMOOVS: Towards calibration-free text entry by gaze using smooth pursuit movements. *Journal of Eye Movement Research* 8, 1 (2015).

[62] Päivi Majaranta. 2012. Communication and text entry by gaze. In *Gaze interaction and applications of eye tracking: Advances in assistive technologies*. IGI Global, 63–77.

[63] Päivi Majaranta, Ulla-Kaija Ahola, and Oleg Špakov. 2009. Fast gaze typing with an adjustable dwell time. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems*. 357–360.

[64] Päivi Majaranta, Anne Aula, and Kari-Jouko Räihä. 2004. Effects of feedback on eye typing with a short dwell time. In *Proceedings of the 2004 symposium on Eye tracking research & applications*. 139–146.

[65] Päivi Majaranta, Poika Isokoski, Jussi Rantala, Oleg Špakov, Deepak Akkil, Jari Kangas, and Roope Raisamo. 2016. Haptic feedback in eye typing. *Journal of Eye Movement Research* 9, 1 (2016).

[66] Päivi Majaranta, I Scott MacKenzie, Anne Aula, and Kari-Jouko Räihä. 2003. Auditory and visual feedback during eye typing. In *CHI'03 Extended Abstracts on Human Factors in Computing Systems*. 766–767.

[67] Päivi Majaranta and Kari-Jouko Räihä. 2002. Twenty years of eye typing: systems and design issues. In *Proceedings of the 2002 symposium on Eye tracking research & applications*. 15–22.

[68] Maria Laura Mele, Damon Millar, and Christiaan Erik Rijnders. 2015. Beyond direct gaze typing: A predictive graphic user interface for writing and communicating by gaze. In *Human-Computer Interaction: Interaction Technologies: 17th International Conference, HCI International 2015, Los Angeles, CA, USA, August 2-7, 2015, Proceedings, Part II 17*. Springer, 66–77.

[69] Carlos H Morimoto and Arnon Amir. 2010. Context switching for fast key selection in text entry applications. In *Proceedings of the 2010 symposium on eye-tracking research & applications*. 271–274.

[70] Martez E Mott, Shane Williams, Jacob O Wobbrock, and Meredith Ringel Morris. 2017. Improving dwell-based gaze typing with dynamic, cascading dwell times. In *Proceedings of the 2017 CHI Conference on Human Factors in Computing Systems*. 2558–2570.

[71] Minh Hoa Nguyen, Thi Duyen Ngo, Nguyen Ba Hung, Can Van Mao, Hai-Dang Kieu, and Thanh Ha Le. 2023. On-screen keyboard controlled by gaze for Vietnamese people with amyotrophic lateral sclerosis. *Technology and Disability* Preprint (2023), 1–13.

[72] Kohichi Ogata and Shuto Inoue. 2021. Effectiveness of an eye-gaze controlled typing system for the japanese syllabary. *International Journal of Innovative Computing, Information and Control* 17, 2 (2021), 369–381.

[73] Prateek Panwar, Sayan Sarcar, and Debasis Samanta. 2012. EyeBoard: A fast and accurate eye gaze-based text entry system. In *2012 4th International Conference on Intelligent Human Computer Interaction (IHCI)*. IEEE, 1–8.

[74] Diogo Pedrosa, Maria da Graça Pimentel, and Khai N Truong. 2015. Filteryedping: A dwell-free eye typing technique. In *Proceedings of the 33rd annual acm conference extended abstracts on human factors in computing systems*. 303–306.

[75] Jimin Pi, Paul A Koljonen, Yong Hu, and Bertram E Shi. 2020. Dynamic Bayesian adjustment of dwell time for faster eye typing. *IEEE Transactions on Neural Systems and Rehabilitation Engineering* 28, 10 (2020), 2315–2324.

[76] Jimin Pi and Bertram E Shi. 2019. Task-embedded online eye-tracker calibration for improving robustness to head motion. In *Proceedings of the 11th ACM Symposium on Eye Tracking Research & Applications*. 1–9.

[77] Patrik Pluchino, Valeria Orso, Giulia Bonamigo, and Luciano Gamberini. 2021. Influence of keyboard layout and feedback type in eye-typing tasks: a comparative study. In *Proceedings of the 32nd European Conference on Cognitive Ergonomics*. 1–7.

[78] Marco Porta. 2015. A study on text entry methods based on eye gestures. *Journal of Assistive Technologies* 9, 1 (2015), 48–67.

[79] Marco Porta, Piercarlo Dondi, Alice Pianetta, and Virginio Cantoni. 2021. Speye: A calibration-free gaze-driven text entry technique based on smooth pursuit. *IEEE Transactions on Human-Machine Systems* 52, 2 (2021), 312–323.

[80] Marco Porta and Matteo Turina. 2008. Eye-S: a full-screen input modality for pure eye-based communication. In *Proceedings of the 2008 symposium on Eye tracking research & applications*. 27–34.

[81] Kari-Jouko Räihä. 2015. Life in the fast lane: Effect of language and calibration accuracy on the speed of text entry by gaze. In *Human-Computer Interaction–INTERACT 2015: 15th IFIP TC 13 International Conference, Bamberg, Germany, September 14-18, 2015, Proceedings, Part I 15*. Springer, 402–417.

[82] Kari-Jouko Räihä and Saila Ovaska. 2012. An exploratory study of eye typing fundamentals: dwell time, text entry rate, errors, and workload. In *Proceedings of the SIGCHI conference on human factors in computing systems*. 3001–3010.

[83] Guilherme MA Ramos, Raiza Hanada, Maria da Graça C. Pimentel, and Cesar AC Teixeira. 2017. A word-prediction eye-typing approach for Brazilian Portuguese entries using geometric movements. In *Proceedings of the 35th ACM international conference on the design of communication*. 1–6.

[84] Shyam Reyal, Shumin Zhai, and Per Ola Kristensson. 2015. Performance and user experience of touchscreen and gesture keyboards in a lab setting and in the wild. In *Proceedings of the 33rd Annual ACM Conference on Human Factors in Computing Systems*. 679–688.

[85] MaryAnn Romski, Rose A Sevcik, Lauren B Adamson, Ashlyn Smith, Melissa Cheslock, and Roger Bakeman. 2011. Parent perceptions of the language development of toddlers with developmental delays before and after participation in parent-coached language interventions. (2011).

[86] Javier San Agustin, Henrik Skovsgaard, Emilie Mollenbach, Maria Barret, Martin Tall, Dan Witzner Hansen, and John Paulin Hansen. 2010. Evaluation of a low-cost open-source gaze tracker. In *Proceedings of the 2010 symposium on eye-tracking research & applications*. 77–80.

[87] Sayan Sarcar and Prateek Panwar. 2013. Eyeboard++ an enhanced eye gaze-based text entry system in Hindi. In *Proceedings of the 11th Asia Pacific Conference on Computer Human Interaction*. 354–363.

[88] Sayan Sarcar, Prateek Panwar, and Tuhin Chakraborty. 2013. EyeK: an efficient dwell-free eye gaze-based text entry system. In *Proceedings of the 11th asia pacific conference on computer human interaction*. 215–220.

[89] Andrew Sears and Vicki Hanson. 2011. Representing users in accessibility research. In *Proceedings of the SIGCHI conference on Human factors in computing systems*. 2235–2238.

[90] Korok Sengupta, Raphael Menges, Chandan Kumar, and Steffen Staab. 2019. Impact of variable positioning of text prediction in gaze-based text entry. In *Proceedings of the 11th ACM Symposium on Eye Tracking Research & Applications*. 1–9.

[91] HM Simpson and Shirley M Hale. 1969. Pupillary changes during a decision-making task. *Perceptual and Motor Skills* 29, 2 (1969), 495–498.

[92] Henrik Skovsgaard, Julio C Mateo, and John Paulin Hansen. 2011. Evaluating gaze-based interface tools to facilitate point-and-select tasks with small targets. *Behaviour & Information Technology* 30, 6 (2011), 821–831.

[93] Oleg Špakov and Darius Miniotas. 2004. On-line adjustment of dwell time for target selection by gaze. In *Proceedings of the third Nordic conference on Human-computer interaction*. 203–206.

[94] Sarah M Swift. 2012. *Low-tech, eye-movement-accessible AAC and typical adults*. San Jose State University.

[95] Mario H Urbina and Anke Huckauf. 2010. Alternatives to single character entry and dwell time selection on eye typing. In *Proceedings of the 2010 symposium on eye-tracking research & applications*. 315–322.

[96] Stephanie Valencia, Richard Cave, Krystal Kallarackal, Katie Seaver, Michael Terry, and Shaun K Kane. 2023. "The less I type, the better": How AI Language Models can Enhance or Impede Communication for AAC Users. In *Proceedings of the 2023 CHI Conference on Human Factors in Computing Systems*. 1–14.

[97] Keith Vertanen and Per Ola Kristensson. 2011. The imagination of crowds: conversational AAC language modeling using crowdsourcing and large data sources. In *Proceedings of the 2011 Conference on Empirical Methods in Natural Language Processing*. 700–711.

[98] Mélodie Vidal, Andreas Bulling, and Hans Gellersen. 2013. Pursuits: spontaneous interaction with displays based on smooth pursuit eye movement and moving targets. In *Proceedings of the 2013 ACM international joint conference on Pervasive and ubiquitous computing*. 439–448.

[99] Xiang Sheng Wang. 2014. Optimizing Gaze Typing with Pupil Size for People with Severe Motor Disabilities. *Applied Mechanics and Materials* 644 (2014), 1330–1333.

[100] David J Ward, Alan F Blackwell, and David JC MacKay. 2000. Dasher—a data entry interface using continuous gestures and language models. In *Proceedings of the 13th annual ACM symposium on User interface software and technology*. 129–137.

[101] Jacob O Wobbrock, James Rubinstein, Michael W Sawyer, and Andrew T Duchowski. 2008. Longitudinal evaluation of discrete consecutive gaze gestures for text entry. In *Proceedings of the 2008 symposium on Eye tracking research &*

applications. 11–18.

[102] Zhe Zeng, Elisabeth Sumithra Neuer, Matthias Roetting, and Felix Wilhelm Siebert. 2022. A One-Point Calibration Design for Hybrid Eye Typing Interface. *International Journal of Human–Computer Interaction* (2022), 1–14.

[103] Zhe Zeng and Matthias Roetting. 2018. A text entry interface using smooth pursuit movements and language model. In *Proceedings of the 2018 acm symposium on eye tracking research & applications*. 1–2.

[104] Shumin Zhai. 2003. What's in the Eyes for Attentive Input. *Commun. ACM* 46, 3 (2003), 34–39.

[105] Shumin Zhai and Per-Ola Kristensson. 2003. Shorthand writing on stylus keyboard. In *Proceedings of the SIGCHI conference on Human factors in computing systems*. 97–104.

[106] Chi Zhang, Rui Yao, and Jinpeng Cai. 2018. Efficient eye typing with 9-direction gaze estimation. *Multimedia Tools and Applications* 77 (2018), 19679–19696.

[107] Xiaoyi Zhang, Harish Kulkarni, and Meredith Ringel Morris. 2017. Smartphone-based gaze gesture communication for people with motor disabilities. In *Proceedings of the 2017 CHI Conference on Human Factors in Computing Systems*. 2878–2889.

[108] Xianjun Sam Zheng, Stuart Goose, Joeri Kiekebosch, and James Jeng-Weei Lin. 2011. AVIN (Assisted visual interactive notepad): a novel interface design to expedite the eye writing experience. In *Universal Access in Human-Computer Interaction. Users Diversity: 6th International Conference, UAHCI 2011, Held as Part of HCI International 2011, Orlando, FL, USA, July 9-14, 2011, Proceedings, Part II 6*. Springer, 635–644.