A new prototype system addresses the limitations of language prediction and retrieval features found in current AAC devices.

BY ANNALU WALLER, TOM GRIFFITHS, AND PER OLA KRISTENSSON

'What I Think about When I Type about Talking': Reflections on Text-Entry Acceleration Interfaces
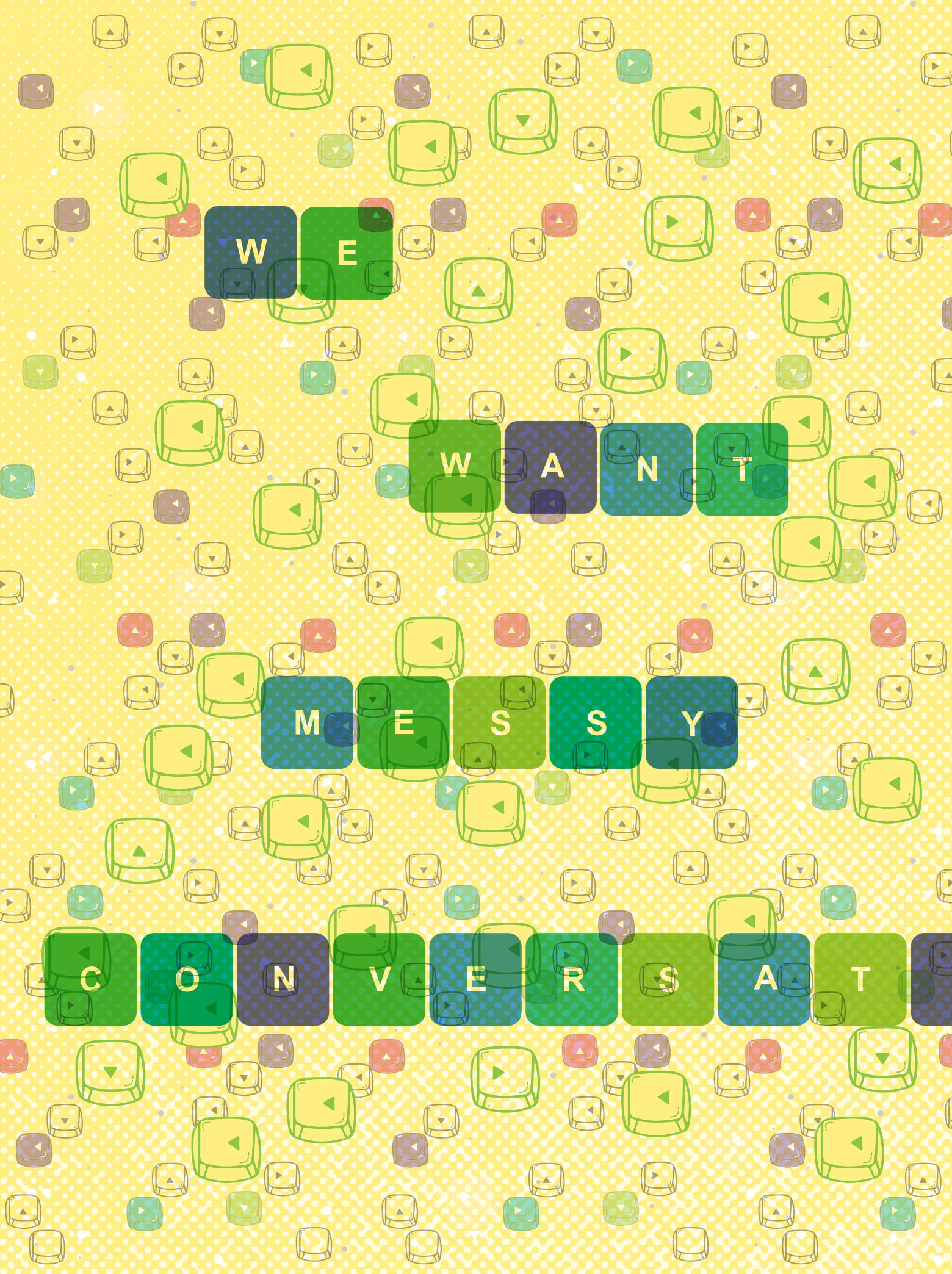
TODAY'S TEXT-ENTRY TOOLS offer a plethora of interface technologies to support users in a variety of situations and with a range of different input methods and devices.[16] Recent hardware developments have enabled remarkable innovations, such as virtual keyboards that allow users to type in thin air, or to use their body as a surface for text entry. Similarly, advances in machine learning and natural language processing have enabled high-quality text generation for tasks such as summarizing, expanding, and co-authoring. As these technologies rapidly develop, there has been a rush to incorporate them into existing systems, often with little thought for the inter-activity problems this may cause. The use of large language models (LLMs) to speed up text generation and improve prediction or completion models is becoming increasingly commonplace, with enormous theoretical efficiency savings;[29] however, the implementation of these LLMs into text-entry interfaces is crucial to realizing their potential.

By considering the perspective of "extreme" cases or "extraordinary" users, we may improve the design of these interfaces for all.[20] One such user group is those who use text-based augmentative and alternative communication (AAC) devices, which can enable literate, nonspeaking individuals with motor disabilities to communicate using text or synthesized voice output, increasing their participation in a range of life situations. Literate AAC users may be almost entirely dependent on text entry to facilitate their daily interactions, and therefore are greatly affected by iterations and advances in text-entry tools.

» **key insights**

■ **Text-entry interfaces are everywhere in mobile technology. However, these interfaces present challenges for both disabled and non-disabled typists. In particular, it is known that the potential time and effort savings of word prediction and sentence recall often go unrealized.**

■ **Considering the use cases of disabled typists can inform better interface design for all. Focusing on the particular physical and cognitive barriers of extraordinary users has implications for mainstream interface design.**

■ **We present the prototype GenieTalk system as a lens through which to study these barriers, discussing five key themes: intuition, uncertainty, mode switching, the nature of conversation, and authorship.**

WE WANT MESSY CONVERSAT

However, despite significant advances in the technologies on which these systems are based, this user group still faces many barriers to using them for interactive conversation.[7,32]

Communication using AAC systems is typically much slower than conversational speech, with output rates dependent on a variety of factors, including cognitive and physical abilities, personal preferences, and the type of access method used to interface with the AAC system.[36] Speed has been a long-acknowledged challenge in the field of AAC,[2] but output rates for AAC users employing direct-access techniques such as typing on an onscreen keyboard have changed little. Single-point typists (those who type with one isolated finger) or those using input devices such as mice, touchscreens, and eye-gaze control typically attain rates of 8 to 10 words per minute (WPM).[32] Users of switch scanning, where users navigate to letters or words and select them using a switch, average speeds of around 1.7 words per minute, orders of magnitude slower than conversational speech rates, which are typically between 120 and 185 WPM.[14] Several solutions have been proposed to increase the speed of AAC users' output, with the main focus being on techniques to enhance the text-entry rate. The most common of these are word and/or sentence prediction, and phrase storage and retrieval. Prediction typically encompasses ideas such as word completion, next-word prediction, and prediction of entire phrases or utterances. Analogs exist for non-disabled users, with word-prediction systems commonplace in mobile interfaces and sentence completion increasingly common in productivity software, with both learning from the user's previous input. Phrase storage and recall systems typically require the user to consciously save an utterance for reuse later.

When we review available AAC systems, a template emerges for how rate enhancement is often implemented. Word-prediction candidates are typically located in a static array, above the keyboard and below the sentence bar, which is the section of the interface where the user's message is constructed prior to speaking. Sentence-prediction candidates are usually limited to a single auto-complete function in the sen-

Providing AAC users with flexible ways to use and reuse language, while maintaining an acceptable rate and flow of conversation, is a problem that remains unsolved.

tence bar itself. When the user is a slow, single-point typist—typical of many AAC users—this can result in a high number of attention shifts to different areas of the screen, many of which will be unnecessary if either no suitable prediction candidate is available or a potentially suitable candidate is overlooked. Phrase storage and retrieval typically have a dedicated button to move the active phrase to the device's memory, with the user then required to move to a page named "conversation history," "stored phrases," or some variation of these terms to retrieve them at a later point. We hereafter refer to these concepts as *prediction* and *retrieval*, respectively.
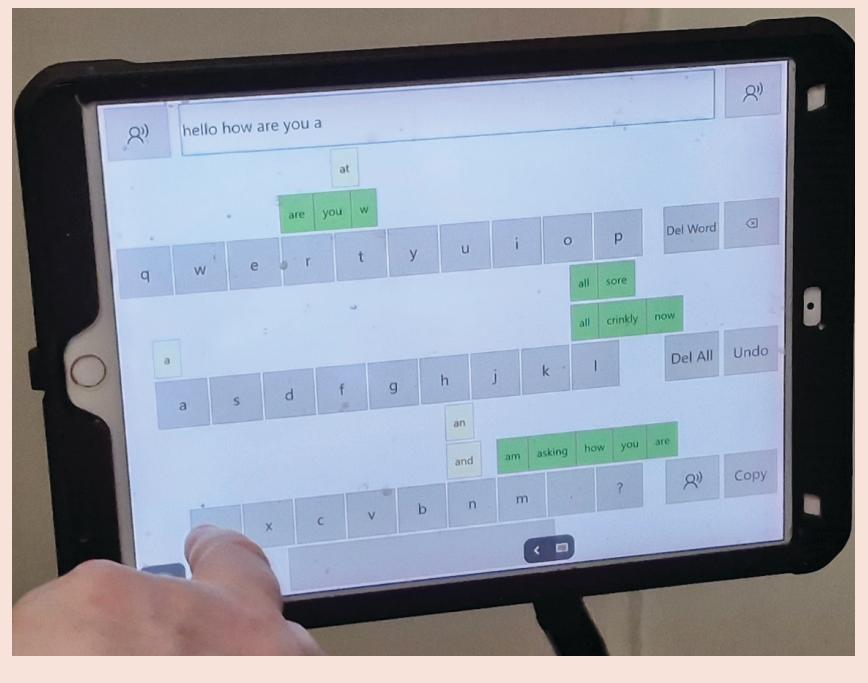
In this article, we engage in a reflective discussion on the design of these features within existing and future AAC systems. The starting point is a prototype interface designed to investigate context-aware AAC,[3] which we use as a lens for these discussions. Alongside our knowledge of existing AAC interfaces and systems, we use the prototype interface to reflect on why existing prediction and retrieval strategies do not provide their intended advantages.[1] One plausible explanation is that the way these features are currently offered to users does not meet their needs. Rate-enhancement features may be unused or abandoned by users if they are not sufficiently flexible to allow them to express exactly what they want to say. Waller and colleagues[33] observe that the design of AAC systems, including the implementation of prediction and retrieval strategies, mainly focuses on the expression of simple phrases of needs and wants, but is poorly suited to more complex and less predictable interactions. Here there is scope for innovative design solutions that address the need for flexible language prediction and retrieval strategies that improve AAC users' performance in conversation.

Currently, the AAC field is changing rapidly as a result of the opportunities and risks posed by generative AI and LLMs. While these undoubtedly have a role to play in supporting AAC users in achieving faster output rates, questions remain about how they might best support a user's conversation performance, which we define as being more than a measure of input/output rates. Instead, we borrow the definition of performance in conversational interac-

tion from Higginbotham and Caves,[11] who in turn expand on the work of John Todman.[26] These authors conceptualize performance in conversation as multi-dimensional, comprising elements of both speed and quality. Conversation performance is improved by AAC users being able to flexibly use language in interactions to achieve their conversational goals, with varying communication partners and in different contexts, while maintaining the speed, flow, and structure of a conversation. Todman observes that transactional communication, such as ordering in a restaurant or shop, includes highly predictable structures and content. Here, the performance deficit for an AAC user is relatively low, as the use of routine, pre-stored phrases will likely be sufficient. In such examples, speed can be prioritized over flexibility of personal expression. Conversation, by contrast, presents a greater performance gap between AAC users and speaking communication partners: Long "gaps" between conversational turns, short utterances, or repeated use of inflexible and potentially impersonal stored phrases are all known to affect the enjoyment of conversation for both sides.[9,26] Providing AAC users with flexible ways to use and reuse language, while maintaining an acceptable rate and flow of conversation, is a problem that remains unsolved. Conversation is a dynamic, uniquely human act; therefore, we take a human-centered approach to the question of enhancing text-entry rates to in turn improve performance in conversation—focusing on users' needs, evolving the design of rate-enhancement strategies at a functional level (the *what*) to ensure systems are fit for purpose, and thereafter translating these functions into technical solutions (the *how*).

Conversational narrative and pragmatics. Many theorists from the fields of disability, linguistics, and computer science cite storytelling as a key component of the human experience. The construction of "conversational narratives" helps shape our identity and personal continuity,[8] and the telling, retelling, refinement, and interpretation of these narratives is crucial to how we make sense of our experiences. So uniquely human is the need to tell stories, to relay narratives about our experiences, that creating human-like

Figure 1. A one-finger typist using GenieTalk to create text to be spoken using speech synthesis.

interactions with artificial intelligence (AI) agents is a foundational challenge for the AI movement.[28] Theorists such as Roger Schank[22,23] argue that the ultimate measure of human intelligence is the ability to recount experience in the form of a story that is appropriate to the conversation in terms of its topic and the manner in which it is shared. The advent of LLMs and general-purpose AI has further highlighted this discrepancy between humans and machines. LLMs trained on large corpora of (typically written) language enable the generation of human-like speech with a fraction of the effort previously required;[29] however, the generation of novel stories and narratives that are interesting, coherent, and engaging for human evaluators remains difficult for these systems.[5] In this sense, modern LLMs are still not passing the measures proposed by Schank.

Speaking individuals make use of conversational narratives frequently, retelling and finessing the retelling of experiences for different audiences—changing elements of the narrative on the fly, perhaps reusing chunks of language they think will have relevance to particular conversation partners or will better suit the context. They can, for example, take account of grounding, the shared understanding of a subject or

context that allows the short-cutting of description and clarification.[10] Typically this is an unconscious process, requiring no conscious effort to recall or reconstruct previously spoken utterances. But the speed at which conversation typically proceeds is a barrier to aided communicators adapting their narratives to suit new audiences.[25] In addition, available AAC technology does not help with this, so users are severely restricted in their ability to flexibly employ conversational narratives.[24] In particular, systems based on existing prediction and retrieval paradigms are ill-suited to the construction of conversational narratives. Their use tends to result in narratives being delivered as monologues, offering the user little or no opportunity to adapt them spontaneously to account for context, audience, or conversational flow.[21,33]

Three key challenges complicate the use of these systems. First, the presence of a narrative in the retrieval system is reliant on the user's having saved it, typically requiring an active extra step. Second, text offered by the retrieval system either fits or it does not, leaving the user with a choice: either use it anyway because it is "good enough" or rewrite the narrative from scratch, perhaps using prediction to increase speed.[35] Finally, there are known difficulties in

editing pre-stored text, with users often unable to change individual lexical items in a larger string without great difficulty.[21,31]

For interaction designers, addressing these barriers embedded in traditional prediction and retrieval strategies has the potential to enhance AAC users' performance in conversation by increasing both the speed and flexibility of output. Through finding ways to capture and flexibly reuse utterances and conversational narratives without the user's needing to deliberately

identify and store them, users may be enabled to pragmatically adapt or fine-tune elements of previous conversation to new scenarios.

GenieTalk

Here, we explore the use of a system called *GenieTalk* (Generating Expression through Narrative in Everyday Talk; Figure 1),[a] designed as an attempt

---

a An example of GenieTalk in use can be found at 14:00 in this video: https://youtu.be/ic1TT8cShCA.

Figure 2. The prototype GenieTalk interface in its launch state, showing the split QWERTY keyboard and an initial offering of prediction candidates.


Figure 3. The GenieTalk system displaying word- (yellow) and sentence- (green) prediction candidates after the user has pressed the initial letter s.

to address the issues with current prediction and retrieval paradigms. In particular, the system breaks from the established workflow of users needing to actively move to a different location or screen to store and subsequently retrieve previously typed utterances (phrases or sentences spoken via a speech synthesizer) within an interactive conversation.

The system uses an onscreen keyboard as its text-entry method. The prototype interface, shown in Figure 2, is based on a standard QWERTY layout, simplified to include only letter keys, the space bar, question mark, and full stop. In addition, a set of function keys (Speak, Stand-by, Delete Word, Backspace, Delete All, and Undo) are located to the right of the keyboard, and an utterance bar is at the top of the screen. Where the layout differs from a standard onscreen keyboard is that the three rows of letter keys are split, with a space in between each row large enough to offer a further two rows of predicted words or retrieved utterances. The maximum number of predicted-word and retrieved-utterance candidates are set in a config text file.

Word prediction uses a standard probabilistic language model, while previously typed utterances are retrieved using a bespoke probabilistic algorithm to identify relevant candidates. As an example of how the system displays word prediction, once the user has typed the letter *s*, the prediction candidates appear above the letters *e* (see), *a* (said), *h* (she), and *o* (some), shown in Figure 3. Prediction candidates are co-located with the next letter to be typed, a design decision based on observational studies showing that AAC users look at where they are typing.[6] Word and retrieval candidates are thus displayed close to the user's current focus of attention, with single-word predictions displayed in yellow and retrieved utterances in green (Figure 3).

Within established paradigms of utterance retrieval, the difficulties in editing retrieved text may be perceived as too great for the user, presenting them with two unsatisfactory options: either use the retrieved text as is and risk it being inappropriate or unsuitable for the new context, or retype the narrative from scratch. In the GenieTalk system, however, the user can
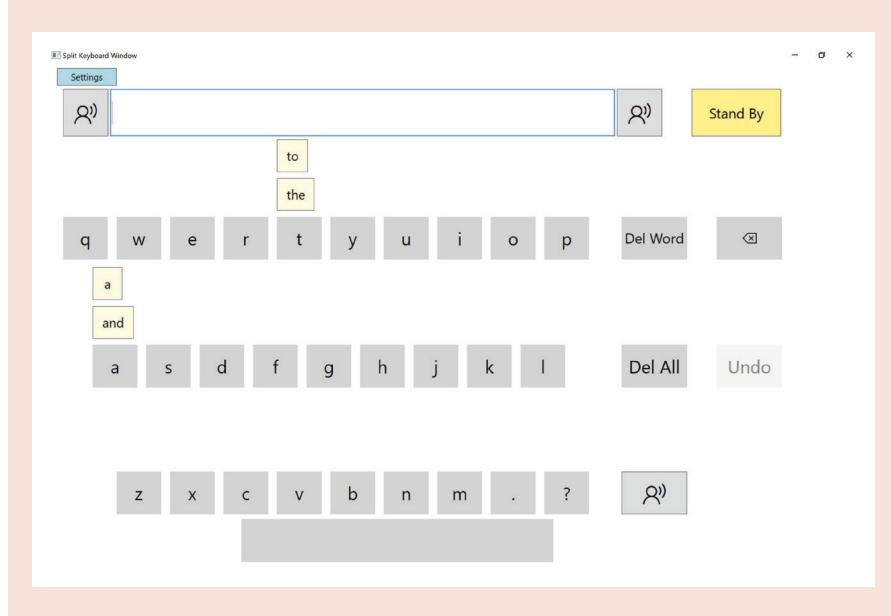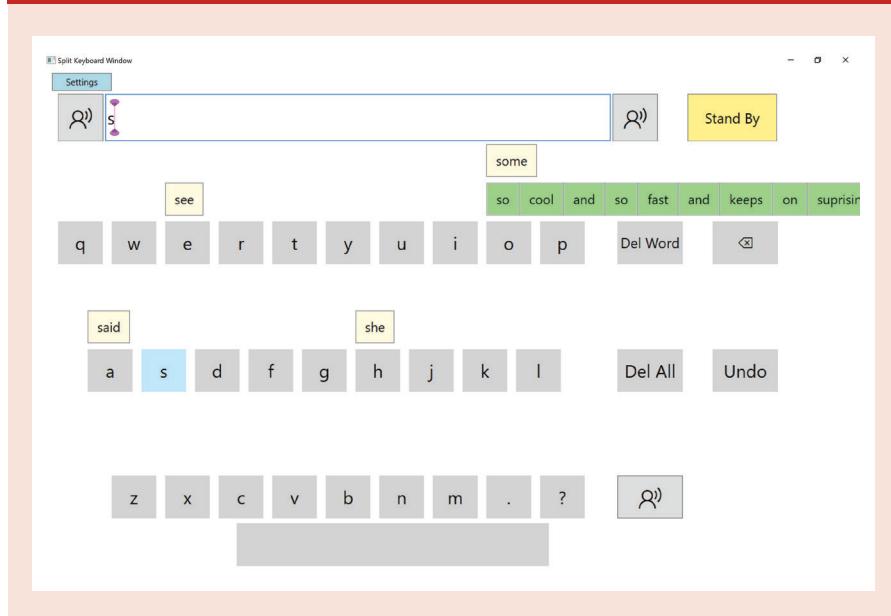
select any number of words from the retrieved utterance, allowing a more flexible approach.

For example, if the system offers the previously typed phrase "I went home last night in a taxi," the user could select all or part of the utterance by selecting the word that ends the relevant word sequence. By selecting "night," the system copies the first five words, "I went home last night," into the utterance bar, allowing the user to continue typing an alternative ending: ". . . and the cat had brought home a bird" (Figure 4). Because the GenieTalk system logs all spoken utterances, it can offer retrieval options that are tailored to the user, without the need for an additional step or conscious process to store them. The system therefore captures not only what the user types but also what they select to speak; hence, it captures the user's spoken utterances.
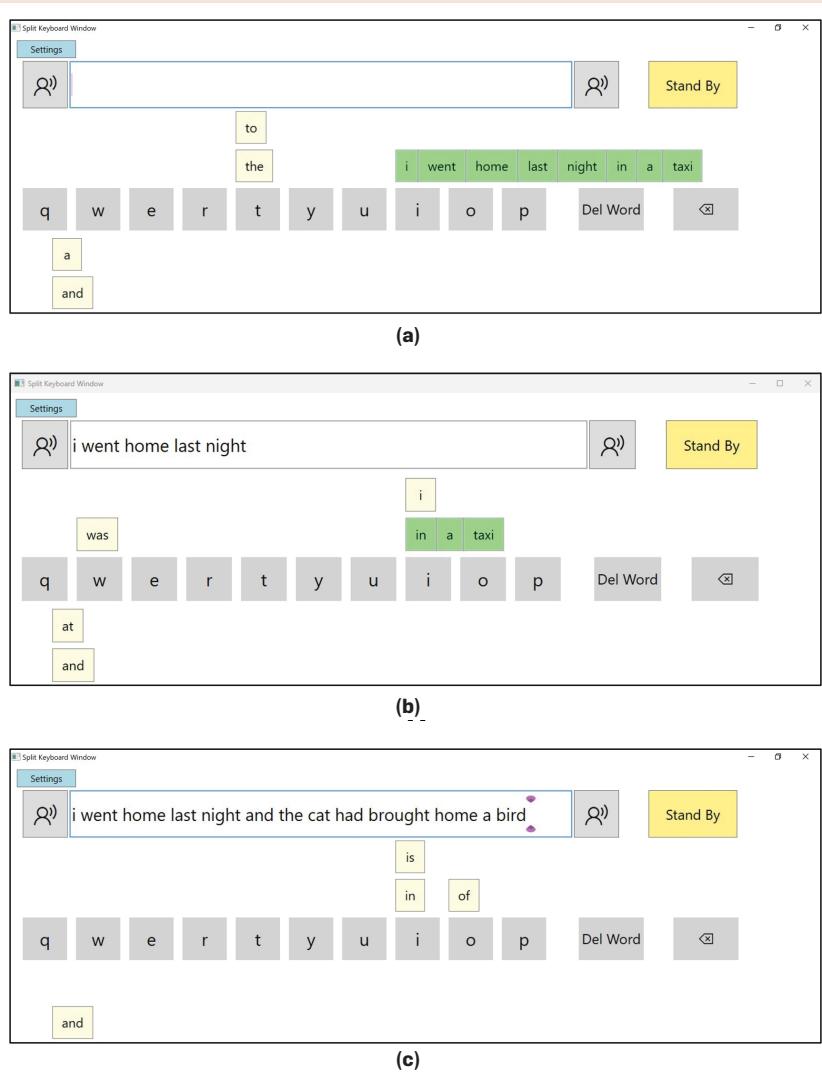
Framing the Discussion

The first author has used the system intermittently since 2019 to support her speech. The authors convened in May 2023 to discuss her experiences and to reflect on the insights from her use of the system over time. Conversations were recorded and transcribed by the second author, who created a synthesis of the discussion that was then revised and agreed on by the other authors.

First author positioning statement. The first author has been at the forefront of AAC technology research since the mid-1980s. Though her speech is generally intelligible to both familiar and unfamiliar communication partners, the environment in which interaction takes place can result in difficulties making herself understood. In these circumstances, she uses AAC strategies and devices to augment her speech.

The genesis of this research is in a serendipitous piece of autoethnographic study conducted by the first author. While researching her Ph.D.,[30] she lost her voice for three days. Following good AAC practice, she used a notebook and pen to write what she wanted to say. She very quickly realized that, instead of writing things from scratch, she could flip back and simply point to sentences or phrases that had been previously written. This experience revealed to her the nature and benefit of reusing parts

Figure 4. Screenshots showing the sequence of using part of a retrieved sentence and repurposing it for a different audience: a) the predicted sentence is offered to the user; b) the user selects the opening clause of the sentence only; c) the user completes the sentence with an alternative ending to suit their current interaction.

of previously communicated text in different conversations and contexts. The notebook became a personalized cache of narratives and phrases that could be called upon to enhance conversations and interactions. Where this method differed from other strategies available to AAC users at that time was that it did not require conscious storage of the information, merely remembering that a particular phrase or sentence had been used before. This led to a career-long research interest in conversational narrative in AAC.

As a clinician researcher, she reflects that she has often recommended tools such as word prediction and sentence retrieval as methods to speed up both voice output and text input, but does not use these strategies herself. Similarly, she has tried on several occasions to use voice dictation software, but found the retraining requirements of these for dysarthric speech (speech that is unclear due to disorders of the motor system or musculature) to be frustratingly high. A common thread emerges here: All of these strategies require extensive training and personalization, or at the very least conscious effort to recognize the potential need to reuse something and then store phrases or sequences of commands. In the prevailing AAC paradigm, phrase and sentence recall require the user to consciously save utterances once they

have been spoken, or to navigate and visually search a "history" or similar page. Likewise, macros in assistive technology (AT) software, which chain together groups of commands, require the foresight to sit and program the macro before use. From an AAC perspective, these methods do not lend themselves to the ebb and flow of natural conversation. In conversation, phrases and narratives that have been previously used are deployed by speakers "on the fly" and subsequently reused with small adjustments on repeated occasions.

Conversational narrative is often seen as being unique and therefore difficult to augment with predictive support. Early work in narrative and AAC[34] demonstrated that, although sharing past experience constitutes the bulk of conversation, it is not straightforward for AAC systems, due largely to the fact that sharing a story is an interactive process in which past experience is reformulated depending on context and conversation partner. The sharing of past experience is therefore largely ignored within the AAC literature and, consequently, within the design of AAC systems. Supporting users to reuse conversational narratives in a way that is interactive and intuitive remains complex and difficult with currently available systems and interfaces.

The first author feels a great sense of frustration, as both an occasional AAC user and a full-time AAC researcher, in understanding too well the capacity and possibilities that intelligent computing and user-centered design can offer. She also recognizes that AAC users need access not only to vocabulary but also to ways in which they can formulate, structure, and restructure the content of their conversation.

Reflecting again on her low-tech notebook, she notes a number of features that have inspired GenieTalk's core principles. First, there is no conscious need to recognize when an utterance or phrase should be stored for reuse. Second, the physical act of storing each phrase is embedded in its creation. Third, there is no conscious need to remember what has been stored or where it has been stored. This reduction in both the cognitive and physical effort required to reuse phrases is GenieTalk's core design principle.

Second and third author positionality. The second author is a linguist, researcher, and a clinician with a background in AAC and computer access for people with physical disabilities. In particular, he is interested in the acquisition of new methods of computer access and control for children with disabilities, and in the design of tools and systems that address the needs of both users and those working to support them. The third author is a professor of interactive systems engineering and the co-founder and co-director of the Centre for Human-Inspired Artificial Intelligence at the University of Cambridge. As a co-inventor of the gesture or "swipe" keyboard, on which he based his own Ph.D. thesis, he has significant experience in the design and use of text-entry systems. His research focuses on user-centered design and the development of systems that support users to interact with others in ways that are flexible, expressive, and creative.

Reflective Discussion

This section is organized under five themes that emerged from a review of the conversations between the authors. All highlighted quotes are from the first author reflecting on her experience with the GenieTalk system.

Intuition. We discussed the nature of intuitive systems: those where the information required is so well integrated into the workflow that it appears "like magic" almost before the user is aware that they need it. The first author reflected that GenieTalk feels intuitive because her own words are predicted within the system, without her needing to actively retrieve them or learn a method by which they can be retrieved.

For those with disabilities, minimizing the learning needs of a new system is a design imperative. In the field of AAC, the skills needed to successfully operate a device are grouped together under the domain of *operational competence*, a term coined in Janice Light's seminal 1989 paper[18] and subsequently summarized in a revision of this work as "skills in the technical operation of AAC strategies and techniques."[19] In clinical discussions about this area of competence, the emphasis is generally placed on the user's ability to learn or develop the skills to use an AAC system. Having to think about operational aspects of the device related to prediction and retrieval, alongside the content and objective of a communicative act, places high demands on the user. We reflected that the accepted view of operational competence ignores the responsibilities of system and interface designers to make operation as intuitive and easy as possible.

In this sense, GenieTalk offers an advantage to users. The familiarity of the QWERTY layout and the presentation of prediction candidates co-located with letters that the user will most likely type next reduces the need to actively learn a new way of interacting with a system to gain the advantages of rate enhancement. Prediction and retrieval are integrated into an existing workflow:

> There was no need to learn any other keyboard or navigation patterns, no need to switch attention or locate something in another part of the screen. When others look, they think, how do you manage all that? But because I am focusing where I am typing, everything else reduces in my vision. So my visual focus allows me to scan easily a very small number of options which are located near my point of focus.

We observed that the sense of whether a prediction system is intuitive is also linked to the efficiency and accuracy of prediction candidates. The first author felt that some of the advantages that GenieTalk offers relate to the personal nature of the prediction, which improves over time as it learns the user's individual utterance patterns. In this sense, personalization is closely allied to the feeling of intuition that the system offers. The first author reflected that the number of prediction candidates was almost always enough and was a comfortable number to scan. Our discussion highlighted that greater numbers of prediction candidates can start to feel as though they offer diminishing returns, with the prediction candidates becoming less accurate and the array taking longer to visually scan.[15,16] Better, more accurate prediction engines could offer the opportunity to further reduce the number of prediction candidates, making the system feel even more personalized.

Uncertainty. The opposite of an intuitive system is one that presents the user with a high degree of uncertainty about whether performing an action, or learning a different workflow, will assist them in completing their goal. All authors discussed their experiences with various text-entry systems on AAC devices and mainstream mobile keyboards. Similar to AAC systems, prediction interfaces in mainstream mobile technologies make an ask of users: that they invest time and effort in learning a new system with no guarantee of performance gains.[16] Optimizing their use of a system might result in performance gains, but these may be too small for the average user to be willing to invest the time. For people with expressive communication difficulties, this is once again magnified: Changing the way in which they use an AAC system or device may require more cognitive and physical investment. It is possible, even, that there may be a negative impact on overall performance. For AAC users, this ask is even greater when we add phrase retrieval or conversation history as options.

Faced with this choice, many users will simply abandon prediction and revert to the status quo of typing without any rate enhancement. People have a tendency, we reflected, to revert to what they know works for them—what is tried and tested. Users will frequently adopt inefficient or suboptimal ways of interacting with systems, purely because they know that these work for them.

Further uncertainty is introduced when we consider the constantly changing nature of software and hardware. AAC systems evolve and change at the behest of available technology, and moving between systems is a recognized challenge for users and those supporting them,[13] often requiring them to spend a lot of time customizing if they wish to move to a new device or software. Similarly, one can optimize the setup of an onscreen keyboard on a mainstream mobile device, but this will inevitably change when the time comes for an operating system upgrade or the switch to a new phone. The volatility of systems can therefore mean that any performance gains accrued from the use of a rate-enhancement method may be time-limited.

> **Supporting users to reuse conversational narratives in a way that is interactive and intuitive remains complex and difficult with currently available systems and interfaces.**

Viewed from the standpoint of being unsure about the likely eventual performance gains, and their likely duration, it is completely rational that users would reject the ask to learn a new system, perceiving this as a gamble with uncertain results. Behavioral economics tells us that people are generally risk averse and disinclined to try new things when they are uncertain or skeptical about the outcome.

There is an associated opportunity cost to all interactions by AAC users in that errors, when they occur, can take a long time to repair. This further decreases the motivation to try out new approaches and instead stick to tried-and-tested methods. In some cases, users opt to engage a skilled human partner rather than a powered AAC device:[32] It is a simpler method of getting their messages across and they accept the limitations associated with it, such as the risk of occasional misinterpretation.

The first author reflects that the uncertainty related to the potential gains from a new system is exacerbated for users with disabilities: We cannot know the exact physical and cognitive capabilities of an individual user when we design interfaces. Non-disabled users can target areas of the screen almost without thinking, whereas disabled users need to consciously think about how to target and make selections, resulting in a greater cognitive and physical load. This is an extra "layer" that non-disabled, speaking people do not have to contend with. The first author equates this with walking: Individuals without a physical disability do not have to think about walking, whereas those with movement disorders or other impairments of lower limb function need to put thought and effort into each step. Similarly, where non-disabled individuals might simply stand up and walk across a room to retrieve something, disabled people will need to compute whether the effort associated with this is "worth it" for the eventual outcome. Much the same is true for interactions with AAC systems.

Mode switching. Prediction and retrieval in current AAC systems requires the user to perform an explicit "mode switch." Put simply, if a user wishes to reuse a concept or string of words that they know they have used previously, they need to switch modes from typ-

ing to retrieval, which are distinctly different cognitive processes. When using current prediction and retrieval interfaces located on separate pages or screen areas, the user needs to know about the page or storage function and has to make a judgement on the likelihood of the mode switch resulting in their desired outcome. Returning to Light's competency framework for AAC use,[18,19] constructing and typing a message is primarily a linguistic act, whereas the storage and retrieval of said message is more in the domain of operational competence. Again, a mode switch is asking the user to gamble. Irrespective of the outcome, the meta-level consideration of the mode switch has taken the user out of their primary workflow: typing a message.

We reflected that there is a level of decision making for AAC users that does not apply to speaking participants in a conversation. This relates to whether what they are intending to type coincides with something they have typed before. If an AAC user wants to say something new, they will need to type it. However, if they have said something previously, then they must consider whether the system has saved this for them, or whether they have actively stored it previously. Too often, users ignore available strategies such as using conversation histories or phrase storage. The first author reflected on observing the regular expert user group at Dundee:

> There remains a frustration when I see a group member share a recent experience with a conversation partner, only to then retype the entire thing when a new partner arrives—even when they have access to conversation history.

Speaking people do not have to think about how to convert their ideas into speech. When they speak, they think about what they are saying, not the operational effort of forming words. AAC users, by contrast, are required to carry out and coordinate a lot of extra processes alongside formulating a message: remembering whether they have said something previously, and remembering the location of words or sentences in their vocabulary and the routes to retrieval.

GenieTalk is an attempt to provide the advantages of prediction and retrieval without the meta-cognitive load—it does not force the user to mode switch, integrating prediction into the typing workflow. When non-disabled people speak, they generate information, words, and phrases without needing to think about where they are stored. There is no need to consciously think about whether something has been said before; rather, the realization that this is a repeat utterance happens organically in the process of speaking. GenieTalk seeks to replicate this for AAC users by removing the need to mode switch to achieve the benefits of prediction and retrieval. We propose that integrating retrieval and prediction into the typing workflow supports this goal. As the first author reflected:

> When I'm typing, I'm thinking about what I'm saying. [With a conventional AAC system] as soon as I have to think about whether I've said something before, whether it is stored in my system, and where I can find it, that executive thinking actually impedes my interaction.

Parallels with this approach exist elsewhere in text input. Error-correction interfaces are often modeless, or at least better integrated into the user's active workflow; mode switching to a special "error correction mode" was seen by users as undesirable and cumbersome.[17]

Nature of conversation. Historically, AAC devices have done comparatively well at supporting formulaic, task-oriented interactions. These interactions, such as requesting things, making choices, or making a purchase from a shop, tend to be more predictable. Such transactional interactions lend themselves well to being supported by pre-stored phrases. Interactional conversation is different, being much less formulaic and almost entirely unpredictable: Its content and structure are dependent on two or more independent interlocutors, building on each other's turns. Conversation is therefore a dynamic, rather than a formulaic, interaction.

In HCI, an established principle is that users are trying to accomplish a goal when using a system. Conversational interaction differs from more predictable tasks in that the goal is co-constructed between the participants and may be constantly revised based on previous conversational turns. In many cases, the goal of conversation is simply the continuance of conversation for the social benefit of all participants.

The first author reflected on early attempts in the 1990s to create interfaces that offered users different types of communication, with reusable phrases in one part of the screen, narrative in another, and the facility to type new phrases in a third.[34] She reflected that, although systems like this were predicated on analyzing the anatomy of a conversation, they were not successful because this is not how people think about how they talk. Users do not necessarily have a conceptual model for conversation and certainly no two users would share the same one. Therefore, we cannot impose upon them the additional demand of thinking about what type of conversation they are engaging in. GenieTalk solves this problem by creating a seamless workflow that functions in the same way, irrespective of the type of interaction, while still offering access to methods of rate enhancement and recall.

Returning to the idea of conversational narrative, the first author observes that her experiences using GenieTalk are similar in nature to spoken use of narrative. When we recount stories to a communication partner, these are tailored to the audience, with chunks or structures reused or omitted according to the context and listeners. The novel sentence prediction in GenieTalk allows users to select parts of sentences for reuse, offering a more flexible approach to recounting and retelling.

Another acknowledged challenge for the AAC field is supporting users in "keeping up" with conversations. It is well-established that the length of time taken to construct utterances has a negative impact on users' participation in conversations: Pauses of 0.5 seconds and above are perceived as disrupting the flow of conversation, and a pause of more than 3 seconds creates an awkward silence.[12] Pauses in conversation are also known to negatively affect conversation partners' perception of AAC users.[27]

Authorship. The first author reflected that what excited her most when first us-

ing GenieTalk was that her own words and phraseology were being predicted, rather than those of a generic model:

My words appear at my fingertips: This is the first time I feel like I'm communicating with my voice, not with someone else's text.

In mainstream software, sentence completion is increasingly becoming the norm, and all authors reported using it. However, the first author observes that these predictions are not one's own words; they are generic and based on probability. We discussed how GenieTalk offers users the option to easily select from phrases based on their own usage, rather than needing to assess a bank of generic predictions or stored phrases to find something that is good enough. Often, prediction candidates that are pre-generated, or that are generated using an LLM, can contain a lot of useless or non-relevant options, which can make it feel like the system is guessing. Being able to recognize one's own phraseology reduces the effort needed to read unfamiliar text. This plays into the HCI principle that recognition is preferable to recall. In GenieTalk, the words are the users' own, engendering a feeling of agency, of being in control. The system only generates sentence-prediction candidates that have been typed by the user, thus avoiding jargon sentences, or sentences the user would never type themselves, that often characterize other prediction systems.

Ownership of one's language is key to self-expression and authenticity.[4] There are several linguistic levels that make up our authentic voice, which is how we feel ownership of what we are saying. Recent studies in digital voice banking and synthetic voice creation[6] have shown that users value not only the phonological characteristics of their voice, but also particular words and phrases that they use. If someone consistently uses a form of utterance that is non-standard, for example, this should be offered as a prediction candidate, rather than them being repeatedly offered the standard form. This foregrounds a user's authentic voice while also not attempting to make them fit a particular linguistic norm or predetermined model. GenieTalk returns the locus of control to users: They feel

GenieTalk offers users the option to easily select from phrases based on their own usage, rather than needing to assess a bank of generic predictions or stored phrases to find something that is good enough.

ownership of the text that is generated, allowing their authentic voice to come through in what they are saying. The need to retain personality and voice is important to users, particularly those with degenerative conditions.

Users may also wish to preserve idiosyncratic syntactic structures—another feature of GenieTalk. The first author recounted a conversation with another AAC user who felt that he would want to retain elements of his often idiosyncratic syntax, as these are a key part of how he communicates and how he wishes to be perceived. Drawing too heavily on community language corpora presents the risk that users will be forced into homogeneous or idealized speech patterns, or that prediction artifacts crowdsourced from elsewhere will make prediction more complicated to use and ultimately less personal.

With questions about how LLMs and other AI might support AAC users becoming more urgent, this question of authenticity and authorship felt important in this discussion, which contributes to another longstanding debate in AAC: the desire to avoid wanting to put words into people's mouths through the options in a vocabulary system. Sentence prediction from a bank of sentences can, by very definition, not cover everything a user might want to say. And though sentence generation, such as that used by LLMs in machine learning algorithms, can generate anything conceivable, the authentic voice of the individual user may still be absent. We propose that sentence reuse, as expressed in a system such as GenieTalk, allows the user to preserve an authentic voice, retaining authorship over their utterances while still enjoying the benefits of rate enhancement through prediction.

Conclusion

Prediction and retrieval offer AAC users the possibility of increased performance in conversation, of streamlining conversations, and of moving toward more symmetrical interactions between aided and speaking communication partners. However, we propose that such performance improvements are attainable only if prediction and retrieval are implemented in a way that does not disrupt the user's typical workflow, minimizing demands on them

while allowing them to retain a flexible, authentic voice. In this article, we discussed the first author's experience using the GenieTalk system, which was designed to address barriers in current prediction and retrieval systems. She felt that the system offers a method of prediction and retrieval that does not require conscious consideration of the need to store and retrieve utterances. We discussed that such a system may reduce the need to mode switch to locate prediction candidates, as they appear "in line" with the keyboard keys and at the location of the most likely next keystroke. The intent in designing GenieTalk was to reduce the effort involved in learning and using rate-enhancement systems. Validation of this through experimentation and user testing is important in developing this prototype system further.

GenieTalk was designed around the needs of users participating in spoken interactions, not generating blocks of written text. It was designed to support the unpredictability of conversation, rather than dealing in probabilistic predictions based on either known lists or LLMs. Indeed, recent research in this area by Valencia and colleagues[29] demonstrated that, while users were broadly happy with the prediction and expansion candidates offered by an LLM, they expressed concern that the LLM output in isolation did not reflect their personal communication style, personality, or identity. Also, to better support performance gains in conversation for AAC users, simply leveraging the ability of LLMs to increase output speed is insufficient. There is an equally important need to consider how prediction and retrieval candidates are presented to users and how they allow users to flexibly reuse their utterances. By focusing on a modeless design and on allowing users flexible access to reusable conversational narratives, the GenieTalk system allows users to continually repurpose their own words, phrases, and stories for less homogeneous output and greater emphasis on their authentic voice.

As the first author has described it, "We don't want formulaic written text: We want messy conversation."

Acknowledgments

References

1. Anson, D. et al. The effects of word completion and word prediction on typing rates using on-screen keyboards. *Assistive Technology 18*, 2 (2006), 146–154.
2. Arnott, J.L., Newell, A.F., and Alm, N. Prediction and conversational momentum in an augmentative communication system. *Commun. ACM 35*, 5 (May 1992), 46–57.
3. Black, R., Waller, A., and McKillop, C. Presentation matters: A design study of different keyboard layouts to investigate the use of prediction for AAC. *AAATE 2019 Conf. (Supplement 1 ed., Vol. 31)*. L. Desideri et al. (Eds.). IOS Press (2019), 141–142.
4. Broomfield, K. et al. A qualitative evidence synthesis of the experiences and perspectives of communicating using augmentative and alternative communication (AAC). *Disability and Rehabilitation; Assistive Technology* (2022), 1–15.
5. Callan, D. and Foster, J. How interesting and coherent are the stories generated by a large-scale neural language model? comparing human and automatic evaluations of machine-generated text. *Expert Systems 40*, 6 (2023).
6. Cave, R. and Bloch, S. Voice banking for people living with motor neurone disease: Views and expectations. *Intern. J. of Language & Communication Disorders 56*, 1 (2021), 116–129.
7. Curtis, H., Neate, T., and Vazqez Gonzalez, C.. State of the art in AAC: A systematic review and taxonomy. In *Proceedings of the 24th Intern. ACM SIGACCESS Conf. on Computers and Accessibility*. ACM (2022).
8. Grove, N. *Using Storytelling to Support Children and Adults with Special Needs: Transforming Lives Through Telling Tales*. Routledge (2013).
9. Higginbotham, D.J. In-person interaction in AAC: New perspectives on utterances, multimodality, timing, and device design. *Perspectives on Augmentative and Alternative Communication 18*, 4 (2009), 154–160.
10. Higginbotham, D.J. et al. The effect of context priming and task type on augmentative communication performance. *Augmentative and Alternative Communication 25*, 1 (2009), 19–31.
11. Higginbotham, D.J. and Caves, K. AAC performance and usability issues: The effect of AAC technology on the communicative process. *Assistive Technology 14*, 1 (2002), 45–57.
12. Jefferson, G. Preliminary notes on a possible metric which provides for a 'standard maximum' silence of approximately one second in conversation. In *Conversation: An Interdisciplinary Perspective*. Multilingual Matters (1989), 166–196.
13. Judge, S. et al. Attributes of communication aids as described by those supporting children and young people with AAC. *Intern. J. of Language & Communication Disorders 58*, 3 (2023), 910–928.
14. Koester, H.H. and Arthanat, S. Text entry rate of access interfaces used by people with physical disabilities: A systematic review. *Assistive Technology 30* (2018), 151–163.
15. Koester, H.H. and Levine, S. Model simulations of user performance with word prediction. *Augmentative and Alternative Communication 14*, 1 (1998), 25–36.
16. Kristensson, P.O. Next-generation text entry. *Computer 48*, 07 (2015), 84–87.
17. Kristensson, P.O., Mjelde, M., and Vertanen, K. Understanding adoption barriers to dwell-free eye-typing: Design implications from a qualitative deployment study and computational simulations. In *Proceedings of the 28th Intern. Conf. on Intelligent User Interfaces*. ACM (2023), 607–620.
18. Light, J. Toward a definition of communicative competence for individuals using augmentative and alternative communication systems. *Augmentative and Alternative Communication 5*, 2 (1989), 137–144.
19. Light, J. and McNaughton, D. Communicative competence for individuals who require augmentative and alternative communication: A new definition for a new era of communication? *Augmentative and Alternative Communication 30*, 1 (2014), 1–18.
20. Newll, A.F. and Cairns, A.Y. Designing for extraordinary users. *Ergonomics in Design: The Quarterly of Human Factors Applications 1*, 4 (1993), 10–16.
21. Savolainen, I. et al. Linguistic and temporal resources of pre-stored utterances in everyday conversations. *Child Language Teaching and Therapy 36*, 3 (2020), 195–214.
22. Schank, R. *Tell Me A Story: A New Look at Real and Artificial Memory, 1st ed.*, MacMillan Publishing (1990).
23. Schank, R. *Tell me a Story: Narrative and Intelligence, 3rd ed.* Northwestern Univ. Press (2000).
24. Soto, G., Solomon-Rice, P., and Caputo, M. Enhancing the personal narrative skills of elementary school-aged students who use AAC: The effectiveness of personal narrative intervention. *J. of Communication Disorders 42*, 1 (2009), 43–57.
25. Ter, N. Wal et al. Everyday barriers in communicative participation according to people with communication problems. *J. of Speech, Language, and Hearing Research 66*, 3 (2023), 1033–1050.
26. Todman, J. Rate and quality of conversations using a text-storage AAC system: Single-case training study. *Augmentative and Alternative Communication 16*, 3 (2000), 164–179.
27. Todman, J. and Rzepecka, H. Effect of pre-utterance pause length on perceptions of communicative competence in aac-aided social conversations. *Augmentative and Alternative Communication 19*, 4 (2003), 222–234.
28. Turing, A. Computing machinery and intelligence. *Mind LIX 236* (1950), 433–460.
29. Valencia, S. et al. 'The less I type, the better': How AI language models can enhance or impede communication for AAC users. In *Proceedings of the 2023 CHI Conf. on Human Factors in Computing Systems*. ACM (2023).
30. Waller, A. *Providing Narratives in an Augmentative Communication System*. Ph.D. thesis, University of Dundee (1992).
31. Waller, A. Communication access to conversational narrative. *Topics in Language Disorders 26*, 3 (2006), 221–239.
32. Waller, A. Telling tales: Unlocking the potential of AAC technologies. *Intern. J. of Language Communication Disorders 54*, 2 (2019), 159–169.
33. Waller, A. et al. Chronicles: Supporting conversational narrative in alternative and augmentative communication. In *Human-Computer Interaction INTERACT 2013*. P. Kotzé et al. (Eds.). Lecture Notes in Computer Science. Springer (2013), 364–371.
34. Waller, A. and Newell, A.F. Towards a narrative-based augmentative communication system. *Intern. J. of Language & Communication Disorders 32*, S3 (1997), 289–306.
35. Wray, A. Formulaic language in computer-supported communication: theory meets reality. *Language Awareness 11*, 2 (2002), 114–131.
36. Yang, B. and Kristensson, P.O. Tinkerable augmentative and alternative communication for users and researchers. *Design for Sustainable Inclusion CWUAAT 2023*. J. Goodman-Deane et al. (Eds.). Springer, Cham (2023).

Annalu Waller is a professor in the Department of Computing at the University of Dundee, Dundee, U.K.

Tom Griffiths is a lecturer in the Department of Computing at the University of Dundee, Dundee, U.K.

Per Ola Kristensson is a professor in the Department of Engineering at the University of Cambridge, Cambridge, U.K.

Watch the authors discuss this work in the exclusive *Communications* video. https://cacm.acm.org/videos/text-entry-acceleration