UNIVERSIDADE FEDERAL DE PERNAMBUCO

CENTRO DE INFORMÁTICA

PROGRAMA DE PÓS-GRADUAÇÃO EM CIÊNCIA DA COMPUTAÇÃO

MAYSA BORGES GAMA

Human-Computer Interaction guidelines for complex Augmentative and Alternative Communication systems with Large Language Model integration

RECIFE

2024

MAYSA BORGES GAMA

Human-Computer Interaction guidelines for complex Augmentative and Alternative Communication systems with Large Language Model integration

Trabalho apresentado ao Programa de Pós-graduação em Ciência da Computação do Centro de Informática da Universidade Federal de Pernambuco, como requisito parcial para obtenção do grau de Mestre Profissional em Ciência da Computação.

Área de Concentração: Mídia e Interação

Orientador(a): Judith Kelner

Coorientador(a): Jutta Treviranus

RECIFE

Catalogação na fonte
Bibliotecária: Luiza de Oliveira/CRB 1316

G184h Gama, Maysa Borges.
Human-computer interaction guidelines for complex augmentative and alternative communication systems with Large Language Model integration / Maysa Borges Gama .– 2024.
107 f.: il.

Orientadora: Judith Kelner.
Coorientadora:Jutta Treviranus.
Dissertação (Mestrado) – Universidade Federal de Pernambuco. Centro de Informática, Programa de Pós-graduação em Ciência da Computação, Recife, 2024.
Inclui referências e apêndices.

1. Interação humano-computador. 2. Comunicação aumentativa e alternativa. 3. Grandes modelos de linguagem. 4. Inteligência artificial. 5. Blissymbolics. I. Kelner, Judith. II. Treviranus, Jutta. III. Título.

006.7 CDD (23. ed.) UFPE - CCEN 2024 –86

Maysa Borges Gama

“Human-Computer Interaction guidelines for complex Augmentative and Alternative Communication systems with Large Language Model integration”

Dissertação de Mestrado apresentada ao Programa de Pós-Graduação em Ciência da Computação da Universidade Federal de Pernambuco, como requisito parcial para a obtenção do título de Mestre em Ciência da Computação. Área de Concentração: Mídia e Interação.

Aprovado em: 5 de fevereiro de 2024.

BANCA EXAMINADORA

____
Prof. Dr. Cleber Zanchettin
Centro de Informática/UFPE

____
Dra. Anna Priscilla Albuquerque Wheler
Google Canadá

____
Profa. Dra. Judith Kelner
Centro de Informática/UFPE
(**orientadora**)

ACKNOWLEDGEMENTS

I would like to express my gratitude to those who have supported me throughout this journey. Firstly, I acknowledge my advisors. To Prof. Judith Kelner, I am grateful for her support and guidance, which allowed me to overcome all the difficulties that I have encountered along the path. I am sincerely thankful for trusting me and allowing me to pursue my dreams. To Prof. Jutta Treviranus, from whom I learned so much. Her support allowed me to extend this project internationally and broadened my knowledge and perspectives.

Next, my appreciations to both research groups that I had the privilege and pleasure to collaborate with. To the Virtual Reality and Multimedia Research Group (GRVM) and the Inclusive Design Research Centre (IDRC) teams, I am truly grateful for the experience of working with them, which has contributed significantly to my growth and development both as a researcher and as an individual.

I want to extend my gratitude to the Blissymbolics Communication International and the Ottawa Foyers Partage teams for contributing to this research. I want to thank specially Treena, Samantha, Julie, and Shirley, whose collaboration was crucial and greatly enriched this work.

I am thankful to all my friends who encouraged, comforted and lent me a hand whenever I needed it. Most of all, I want to express my deepest love and gratitude to my partner and soulmate. Vitor, thank you for your fondness and unconditional support since our childhood. I hope we can continue motivating each other as we grow older together.

I want to acknowledge the provision of funding for this research to the Coordenação de Aperfeiçoamento de Pessoal de Nível Superior (CAPES) and the Global Affairs Canada Emerging Leaders in the Americas Program.

ABSTRACT

Some individuals with complex communication needs rely on Augmentative and Alternative Communication (AAC) devices for daily interactions. However, AAC users often struggle to communicate during real-time conversations, considering the high amount of physical and mental effort it takes to compose their messages. While users can prepare for conversations by adding new words and phrases to their communication devices, pre-programmed messages may not always cover everything that is necessary to express a person's intentions during a real-time conversation. To address this issue, researchers are exploring the use of artificial intelligence (AI), particularly large language models (LLMs), to improve AAC systems. However, it is crucial to carefully analyze the benefits and drawbacks before rushing to create new technological interventions. This study examines the existing contributions and lists opportunities and challenges for applying LLMs to AAC software, collaboratively with users. Additionally, the study addresses the usage of Blissymbolics, a language with grammar and vocabulary but no phonology, which plays an essential role in assisting individuals with congenital conditions in developing their literacy skills. Two secondary research studies were conducted to systematically and critically map the relevant literature on both LLMs and Blissymbolics uses for AAC. The primary research followed a co-design approach and engaged AAC users and familiar listeners who shared their expectations, concerns, and ideas about future LLM integration into AAC devices. The data collected from the primary and secondary research was used to generate comprehensive human-computer interaction guidelines for developing robust AAC systems that use AI language models. Moreover, the opportunities, challenges and risks involved are also discussed. Furthermore, this work applies the guidelines proposed and showcases their adaptability by developing and detailing a conceptual AAC system that uses LLMs.

Keywords: human-computer interaction; augmentative and alternative communication; large language models; artificial intelligence; blissymbolics.

RESUMO

Alguns indivíduos com necessidades complexas de comunicação dependem de dispositivos de Comunicação Aumentativa e Alternativa (CAA) para suas interações diárias. Contudo, usuários de CAA comumente enfrentam dificuldades ao se comunicar em tempo real, devido ao grande esforço físico e mental demandado para compor as mensagens. Mesmo que os usuários se prepararem com antecedência e incluam novas palavras ou frases em seus dispositivos de comunicação, as mensagens pré-programadas nem sempre suprem tudo o que é necessário para se expressar plenamente durante uma conversa. Tendo em vista essa problemática, pesquisadores têm explorado o uso de Inteligência Artificial (IA), particularmente os Grandes Modelos de Linguagem (GML), para aprimorar os sistemas CAA. Entretanto, é crucial analisar cuidadosamente os benefícios e desvantagens antes do desenvolvimento desenfreado de novas intervenções tecnológicas. O presente estudo analisa as contribuições existentes e realiza um levantamento das oportunidades e desafios na aplicação de GML para softwares CAA, em colaboração com os usuários. Além disso, o estudo aborda o uso de Blissymbolics, uma linguagem com gramática e vocabulário, mas sem fonologia, que desempenha um papel essencial em auxiliar indivíduos com condições congênitas a se alfabetizarem. Duas pesquisas secundárias foram conduzidas para mapear sistematicamente e criticamente a literatura relevante sobre ambos os usos de GML e Blissymbolics na CAA. Seguindo a abordagem de Design Colaborativo, a pesquisa primária envolveu usuários e ouvintes familiares de CAA, os quais compartilharam suas expectativas, preocupações e ideias sobre futuras integrações de GML em dispositivos CAA. Os dados coletados das pesquisas primária e secundárias foram utilizados na criação de diretrizes para o desenvolvimento de sistemas robustos de CAA que aplicam modelos de linguagem e IA. Em conjunto, as oportunidades, desafios e riscos envolvidos foram também discutidos. Ademais, este trabalho aplicou as diretrizes propostas e demonstrou sua adaptabilidade através da criação de um sistema CAA conceitual que utiliza GML.

Palavras-chave: interação humano-computador; comunicação aumentativa e alternativa; grandes modelos de linguagem; inteligência artificial; blissymbolics.

LIST OF FIGURES

LIST OF TABLES

LIST OF ABBREVIATIONS AND ACRONYMS

AAC	Augmentative and Alternative Communication
AI	Artificial Intelligence
AT	Assistive Technology
BCI	Brain-Computer interface
BERT	Bidirectional Encoder Representations from Transformer
Bliss	Blissymbolics
CCN	Complex Communication Needs
GPT	Generative Pretrained Transformer
HCI	Human-Computer Interaction
IBGE	Brazilian Institute of Geography and Statistics
LLM	Large Language Model
ML	Machine Learning
NLG	Natural Language Generation
NLP	Natural Language Processing
PCS	Picture Communication Symbols
PI	Prompt Injection
RoBERTa	Robustly Optimized BERT

CONTENTS

1 INTRODUCTION

Taking a phone call, writing a text message, ordering food or joining a group conversation are seen as effortless and unconscious parts of daily activities. However, communication and interaction are not experienced as simple tasks by everyone. Individuals with severe speech and physical limitations often struggle to express themselves without proper support, leading to major issues such as social exclusion and unemployment. Rick Creech's – a young man with cerebral palsy – shares his description of how it feels like to be unable to speak:

> If you want to know what it is like to be unable to speak, there is a way. Go to a party and don't talk. Play mute. Use your hands if you wish but don't use paper and pencil. Paper and pencil are not always handy for a mute person. Here is what you will find: people talking; talking behind, beside, around, over, under, through, and even for you. But never with you. You are ignored until finally you feel like a piece of furniture. (BEUKELMAN and MIRENDA, 2020 *apud* MUSSELWHITE and ST. LOUIS, 1988, p. 104)

Those with complex communication needs (CCN) often rely on forms of Augmentative and Alternative Communication (AAC) for their everyday needs. AAC interventions can be unaided and do not require any equipment, or aided, which requires an external tool or technological device. Unaided forms include vocalizations and gestures such as word approximations, eye blink codes, facial expressions or nodding. Aided AAC consists of low-technology (e.g., communication boards) and high-technology options, such as computer-based systems with speech-generation and mobile technologies (Beukelman and Mirenda, 2020).

Individuals with speech limitations who use AAC often struggle with real-time communication. While spoken communication rates range between 125 and 185 words per minute, high-tech AAC can fluctuate between 2 words per minute for scanning interfaces and 8–10 words for direct selection. However, the AAC context is more complex than merely providing accessibility features. Several factors could impact the user's experience and technology adoption, ranging from insufficient personalization alternatives to a lack of proper multidisciplinary support (Waller, 2019).

This scenario is more challenging when considering individuals with congenital conditions (e.g., cerebral palsy, Down syndrome, autism spectrum disorder) who often face significant difficulties developing literacy skills – as opposed to those with acquired conditions (i.e., as a result of an accident or illness). Consequently, each person develops their own language structure, according to their unique capabilities, that deviates from standard grammatical rules. An approach in AAC that is often applied is using telegraphic language to reduce labor and ease communication, for example, uttering “Give phone” to express “Can you give me my phone?”.

Blissymbolics[1] is an invented graphic language composed of semantic symbols that allow boundless possibilities of combinations to generate new concepts. It is exclusively written and universal; it has no phonetics and is not associated with any spoken language. Thus, this minority language is extremely valuable in assisting people with physical and speech impairments to communicate, especially those with developmental conditions. Indeed, Bliss has been used to support children with speech limitations for more than 50 years, in countries such as Brazil, Canada, Sweden and South Africa (Blissymbolics Communication International, 2023?a).

Considering the challenges faced by AAC users, researchers have started exploring the use of artificial intelligence (AI) technologies. The recent advances in large language models (LLMs) offer a promising opportunity to improve AAC systems and support individuals with CCN. For instance, AI language models could reduce the cognitive and physical efforts required to compose messages by assisting with personalized and context-sensitive text generation. In this way, it could allow users to effectively communicate while expressing their unique personalities.

In this study, two systematic literature mappings were conducted in order to identify, analyze and synthesize the existing works on (i) LLMs for AAC and (ii) Blissymbolics in high-tech AAC. Moreover, AAC users were engaged in this study through a co-design approach. Two adults with speech limitations and two familiar listeners with collaborated by sharing their concerns, expectations, and ideas about future uses of AI for AAC. The findings from the literature and co-design sessions were categorized, organized and combined to generate human-computer interaction guidelines.

Overall, the guidelines detail three main principles for building complex AAC systems with AI language model integration: Reliable, Customizable, and Adaptable. The existing topics within each principle are thoroughly detailed, and a discussion of the research opportunities, challenges, and risks was also provided. Finally, a conceptual AAC system was elaborated as a proof of concept of the proposed guidelines.

[1] Blissymbolics language is distinct from other commonly used pictographic symbols in that it has a flexible structure, grammar, and rules. This unique characteristic enables learners to develop their literacy skills and eventually apply their learning to other spoken languages. Unlike other pictographic symbols that only facilitate quick face-to-face communication, Blissymbolics provides a more comprehensive framework for language development.

This work is structed as follows. Section 1 details the research problem, goals and questions. Next, section 2 contextualizes the readers on assistive technologies, specifically augmentative and alternative communication, as well as the Blissymbolics language. The research design is explained in section 3. Section 4 presents the secondary research, which includes two systematic literature mappings. Subsequently, section 5 describes the primary research study, which individuals who have CCN by following an inclusive design approach. Then, section 6 details the proposed guidelines built and provides a discussion about the existing challenges and risks, and other research opportunities. Section 7 presents the proof of concept, in which a conceptual AAC system was brainstormed based on the guidelines previously presented. The discussion is detailed in section 8, and the conclusion in section 9, comprising our contributions and future work.

1.1 RESEARCH MOTIVATION

According to the Brazilian Institute of Geography and Statistics (IBGE, 2023), in 2022, there were 18.6 million people with disabilities (aged 2 years or more) in Brazil, corresponding to 8.9% of the Brazilian population, in which a great part could benefit from the use of improved high-tech AAC interventions. Additionally, the IBGE has different classifications for the following impairments: visual, auditive, physical (upper and lower limbs), mental and multiple (more than one disability). We would like to bring attention to the lack of a clear classification for speech limitations – which might be included under the mental impairments category – and the absence of reports on communication needs and the use of AAC.

Individuals with speech limitations are highly diverse as their communication needs and capabilities are particular for each person. Regardless, they share some general challenges with real-time communication, such as major cognitive and physical load when using AAC devices, anxiety when interacting with non-familiar listeners (i.e., unknown people or individuals that are not used to interact with AAC users), and often dependency on human interpreters (Kane *et al*., 2017; Valencia *et al*., 2020).

Commonly, AAC users have to prepare in advance for interactions with non-familiar listeners by adding new words and phrases to their communication devices. Still, pre-programmed messages can not always cover everything that is necessary to express a person's intentions during a real-time conversation. Communications and their interpretations rely on more than just grammar, they are highly dependent on pragmatics (e.g., time, context, social cues). Thus, even highly skilled AAC users can feel pressured as simple sentences can take a few minutes to be composed or they may struggle to engage in group interactions (Kane *et al.*, 2017).

One promising approach to assist AAC users is the use of AI technologies, which have increasingly become popular in the last few years and have been widespread to different platforms. However, those applications are based on large amounts of data and statistical processing, leading them to disfavor users that are considered “outliers.” This phenomenon can be observed, for example, in voice-based personal assistants like Amazon's Alexa, which was proved to perform poorly when processing voice commands of users with an accent (Harwell, 2018) or speech limitations (Ballati; Corno; De Russis, 2018). On top of that, AI-based applications pose severe privacy and security risks to all users and can be particularly harmful to marginalized communities (Weidinger *et al.*, 2021).

The use of LLMs for enhancing AAC software is a promising approach. Individuals with CCN could benefit from generative AI technologies, for instance, to reduce efforts when composing sentences. However, researchers, developers and designers must cautiously analyse all benefits and risks, rather than rushing to create novel technologies. The final goal should not be to carelessly implement AI to the AAC domain, but to responsibly create meaningful tools that will assist individuals with CCN in achieving their full potential.

Higginbotham and Caves (2002) affirm that augmented communication is human-computer interaction (HCI) used to achieve human-human interaction. Hence, critically analyzing and discussing HCI within the AAC field is a crucial step in improving not only communication per se but all the other instances that are affected by it – such as autonomy, interpersonal relationships and self-esteem. Additionally, HCI for AI-based applications has exclusive particularities (Xu *et al.*, 2023), and its implications for AAC software should be carefully evaluated.

Moreover, it is essential to be mindful that AAC stakeholders (e.g., communicators, family members, educators and caregivers) are experts with their unique experiences and must be involved in the early development stages of any technology that will directly affect their everyday activities. Especially individuals with CCN should be at the forefront as active decision-makers in any initiative that affects their well-being. Blackstone, Williams and Wilkins (2007, p. 193) underline the key principles for research in AAC, whereas the first one is the importance of engaging people who rely on AAC to participate in research and practices in the field and emphasizes that “the most important voices are often the hardest to hear”.

1.2 RESEARCH GOAL AND QUESTIONS

Rather than hastily developing novel AI interventions, this work aims to take a step back and provide AAC and AI designers, developers, and researchers with comprehensive human-computer interaction guidelines and reflections for integrating reliable, customizable, and adaptable AI language models into high-tech AAC devices.

For this, we examine the existing contributions and, collaboratively with AAC users and familiar listeners, list opportunities and challenges in applying AI technologies to AAC software. Additionally, this work will address the use of Blissymbolics which, as a language with no phonology, plays an important role in bridging language and communication for children and adults with CCN. Thus, we delineate two main topics: (i) applying LLMs[2] for AAC and (ii) the usage of Blissymbolics language in the high-tech[3] AAC domain.

To achieve the research goal mentioned, a total of three questions were formulated. The first two were used as sub-questions to address the third and main research question. This approach was taken to ensure that all topics and aspects of research would be thoroughly explored and analyzed.

(Q1) What are the benefits and challenges, if any, related to the development of an AAC software with LLM application?

(Q2) What are the benefits and challenges, if any, of integrating Blissymbolics into an AAC software with LLM application?

(Q3) What is needed to build a robust AAC system with AI language model integration for highly diverse users?

Robustness is essential in complex systems, which encompass many components that interact with each other. In this work, "robust" is considered as the characteristic of a system that effectively and adequately integrates the multiple aspects and properties that ensure the users' needs are met.

---

[2] LLMs integrate transformers architecture due to their exceptional ability to capture complex patterns and long-range dependencies in data. Unlike traditional fundamental models, transformers excel at handling sequential data, making them indispensable for tasks like natural language understanding and generation. Additional information on this subject can be found in subsection 4.1.

[3] The differences between no-tech, low-tech and high-tech AAC are detailed on the next section (section 2).

2 CONTEXTUALIZATION

In this section, two topics will be introduced as means of framing and explaining the main context of this research. The first topic is about augmentative and alternative communication and assistive technologies. It will be explained what it is, how it works, and examples of it will be provided. The second topic will focus on Blissymbolics' history and how it works.

2.1 AUGMENTATIVE ALTERNATIVE COMMUNICATION AND ASSISTIVE TECHNOLOGIES

The American Speech-Language-Hearing Association (c2023) explains that augmentative and alternative communication represents all ways of communication besides talking, as Augmentative means to add to someone's speech and Alternative means to be used instead of speech. AAC can be used by any person, of all ages, that faces temporary or permanent limitations with speech or language skills, for a shorter or longer period of time. For example, people who have laryngitis or just had surgery and can not talk may use AAC for a brief period, while others may use AAC throughout their lives.

AAC can be no-tech, low-tech or high-tech. Examples of no-tech options include gestures and facial expressions, sign language, writing, and drawing. Examples of communication using low-tech AAC include basic tools, such as books and communication boards used for pointing at pictures and written words. Lastly, high-tech options include using an application on a tablet or using dedicated computers.

High-tech Assistive Technologies (AT) assist in improving the individual's independence and well-being; some examples are hearing aids, electric wheelchairs, and communication aids (World Health Organization, 2024) – such as high-tech AAC devices. AT for individuals with speech and language limitations often include built-in speech-generating options. Those are known as Speech Generating Devices or Voice Output Communication Aids, which can output digitized or synthesized speech. Digitized speech is natural human speech that has been recorded using a microphone and stored for later reproduction. In contrast, synthesized speech uses linguistic rules-based algorithms to translate the user input (e.g., letters, words, symbols) into machine-generated speech, and thus is not dependent on a fixed set of pre-stored messages (Augmentative Communication Inc., c2023).

High-tech AAC can be categorized into dedicated devices, which are made solely for AAC and are usually costly but more reliable; and non-dedicated devices, which are “off the shelf” products that can be easily purchased and have AAC applications installed later. In Figure 1, we portray examples of dedicated and non-dedicated AAC devices. The first illustration (top left) represents a dedicated device, which was created based on the Accent® 1400 – developed by the PRC-Saltillo company and has its prices starting from US$ 7,595.00 (PRC-Saltillo, c2023). Whereas the non-dedicated device illustration (bottom right) is based on an iPad – from the Apple company, with prices starting from US$ 449.00 for the 10th generation (Apple, 2023a) – and running a mobile application inspired by the TouchChat app – also created by the PRC-Saltillo group, priced at US$ 299.99 (Saltillo, c2023).

Figure 1 – Illustration of dedicated and non-dedicated AAC devices, based on real products.

Source: Author.

Every AT for communication aid is different, with its own benefits and drawbacks. Still, Curtis, Neate and Gonzalez (2022) identified that some features are predominantly applied in high-tech AAC, such as customization (55.2%) and automation (46.6%). One notable example of a customizable AAC mobile app is the Proloquo2Go for iPhone and iPad (created by the company AssistiveWare), which features 49 children and adult’s text-to-speech voices and 12,000 words. In contrast to tools that do not offer any sort of speech-generation personalization, these options allow the user to properly reflect their personhood.

Proloquo2Go was featured in an article published on Apple’s website (Apple, 2023b). They share the story of Jay Ashburn (see Figure 2), a 9-year-old boy diagnosed with autism who uses – since the age of 4 – the tool every day to communicate with his family, teachers, and friends. It helps users to build phrases or sentences, learn related words, and expand language and grammar. Additionally, Jay also uses the app to practice and improve his speech by repetitively pressing the button until he can speak the message himself.

Jay's mother, Meghan Ashburn, explains that he learns in a general third-grade class, and his classmates also use Proloquo as one way of interacting with him. She adds that "It's a big deal, because he's nonspeaking and they tend to get separated". Ashburn believes that "The awareness of AAC among the public will increase in the next five years."

Figure 2 – Photos of Jay Ashburn and his mother Meghan. On the left, he uses an iPad app to communicate with his mother. On the right, a close-up photo of him using the app Proloquo2Go.

Source: (Apple, 2023b).

The popularity and affordability of capacitive touchscreens (e.g., tablets, smartphones) – alongside the convenience of downloadable apps and the "democratization" of software development – evoked a decentralization of high-tech AAC (Light and McNaughton, 2012) and reflected, consequently, in the input methods used. Previously, mechanical input methods were prominently used between the 70s (100%) and 90s (71.4%). Examples of mechanical inputs are keyboards, switches, trackballs, and joysticks. However, in the 2020s, this scenario shifted as tactile input rose to 43.3% and mechanical decreased to 22.4%. Yet, output options for AAC predominantly center on audio with 62.8%, and visual with 27.9% (Curtis; Neate; Gonzalez, 2022).

Elsahar *et al.* (2019) categorized AAC special-input devices based on the human body signals generated to control the interface, that is, body movements, respiration, phonation, or brain activities. The five categories created were: (i) Imaging methods (e.g., eye gazing, head-pointing); (ii) Mechanical and Electromechanical methods (e.g., keyboards, switches); (iii) Touch-activated methods (e.g., touchscreen, touch membrane keyboards); (iv) Breath-activated methods (e.g., microphones, low-pressure sensors); (v) Brain-Computer Interface (BCI) methods (e.g., invasive and non-invasive options). These input methods can be used independently or in combination with others for multi-modal access. Figure 3 illustrates five input options, examples include one device for each of Elsahar's categories.

Figure 3 – Illustration of five different special-input devices being used, respectively: eye-gaze, foot switches, sip and puff switch, touchscreen, and BCI.

Source: Author.

Selection methods for AAC can be either direct, when the user can voluntarily select the desired option from a set of choices; or indirect, as in scanning methods, where the user is presented with dynamic options (changed automatically or deliberately) to select within a particular and adjustable time frame. Direct selecting methods are used with devices such as a keyboard or touchscreen. Conversely, indirect selections can be controlled using devices such as single switches, since they are often suitable for individuals who lack voluntary controls. Such diversity of input and selection options is indispensable since each individual with CCN has their unique capabilities, and thus, are able to choose whichever tool better suits their needs.

Communication is not an isolated activity, but rather a process that encompasses different forms of interaction throughout everyone's daily experiences (Light and McNaughton, 2012). Therefore, tools that support communication should provide integrated access beyond face-to-face conversations, and include other forms such as written, telecommunication and virtual interactions.

Additionally, the usability of AAC products are often related to the number of environments in which the device is fully functional (Elsahar *et al*., 2019). Hence, designers and researchers must attend to different aspects when developing new solutions for communication aid. For example, a dedicated computer with eye-tracking technology as an input option might restrict its usage to close environments, since it is not portable and requires specific conditions to operate appropriately.

Similarly, a non-dedicated mobile device might increase portability and reduce development costs. Regardless, if not based on research evidence and built collaboratively with AAC users, they may not appropriately suit the needs of individuals with CCN. Likewise, integrating novel technologies into AAC, such as LLMs, could potentially assist in reducing typing effort, improving customization and adaptability, supporting user autonomy, and more. However, such systems should not be carelessly developed but instead carefully analyzed to consider the existing benefits, drawbacks and risks, if any.

2.2 BLISSYMBOLS

Blissymbols (also called Blissymbolics or simply Bliss) is a writing system developed by Charles K. Bliss, initially called Semantography (Bliss, 1965). He aspired to create a universal written language that allowed communication between speakers of different languages. Blissymbols do not have phonetics and are composed of basic symbols representing concepts and allowing a variety of combinations for generating new characters.

2.2.1 History of Blissymbols

Charles K. Bliss, a Jew born in 1897, moved to different countries after being released from a Nazi concentration camp in 1939. During his time in Shanghai (China), the Chinese ideograms caught Bliss's attention and inspired him to develop a new intuitive and universal writing system in 1949.

As a visionary who experienced violence and segregation, Bliss wanted his creation to be used as a bridge to connect people from different places and bring peace. Despite years of attempts, his utopian idea was never noticed. However, decades later, at Canada's leading center for children with CCN, Shirley McNaughton first learned about Semantography in a book called Signs and Symbols Around the World (Helfman, 1967). As a special educator, she saw the system as an opportunity to facilitate communication with children with cerebral palsy.

In 1971, with a team of interdisciplinary clinicians, McNaughton initiated an evaluation project at the Ontario Crippled Children's Centre (now known as Holland Bloorview Kids Rehabilitation Hospital), where they applied Blissymbols within the educational program for children within the preschool to grade one levels. She started to communicate with her students, using Bliss as an augmentative and alternative communication approach.

Okrent (2009) narrates this story in her book with rich details. She further explains how, firstly, she did not understand the importance of Bliss and asked why Blissymbols were necessary for the children, when it was possible for persons like Stephen Hawking to use text to speech technology devices and communicate using words directly. McNaughton explained that contrary to the children with cerebral palsy:

> Stephen Hawking was an adult when he lost the ability to speak. (...) He already knew how to use English to express himself. He already knew how to read. (...) What good is English text to a child who can't read yet? And if a child can't speak and can't move, how do you teach them to read? How do you know what they know, what they understand? (Okrent, 2009, p. 269)
>
> They had little boards with pictures on them — a picture of a toilet, a picture of some food, all needs-based pictures — I went through a year just asking them yes-or-no questions: "Would you like to do this? Would you like to do that?" But they couldn't initiate anything themselves. They seemed to understand what was said to them, and, more important, they seemed to have something to say. "You could just tell with the twinkle in their eye or something." (Okrent, 2009, p. 270)

As described in the podcast The Symbols of Bliss (2018), McNaughton says the first experience communicating with her students using Blissymbolics "was the most exciting thing in the world." Children rapidly grasped Bliss and started communicating, leading to some emotional episodes that can be seen in the documentary movie Mr. Symbol Man (1974).

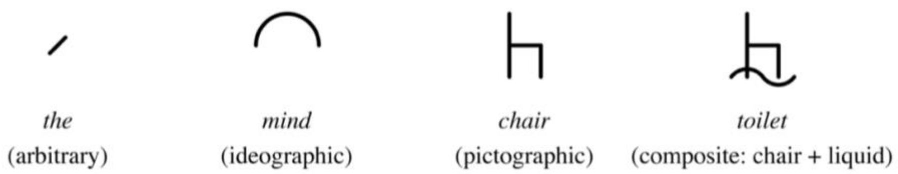
Children with complex speech needs could now self-express, communicating their thoughts and feelings to their parents and peers. Furthermore, Blissymbolics allowed them to explore their creativity, either telling jokes or creating art (e.g., poems). For the first time, children with severe impairments did not need to rely on a fixed set of pictographic options. Before Bliss, they depended on needs-based pictography, that is, communication with only pictures and drawings of basic necessities (e.g., toilet, food, shower).

However, the application of Blissymbols to children who lacked functional speech was not the original intention of C.K. Bliss. This led to several years of conflict that were resolved in 1982, through Charles Bliss granting an exclusive worldwide license for using and publishing Blissymbols for persons with communication, language and learning difficulties, to Blissymbolics Communication International, with Shirley McNaughton as its program director. The Section 3 of this work underwent personal revision by Shirley McNaughton to ensure our findings and description reflect her real story.

2.2.2 How Blissymbols works

Approximately 900 basic characters are derived from Semantography, and there are over 5,000 symbols currently. Unlike the Latin alphabet, each Bliss character represents an abstract or concrete concept. The characters can be arbitrary, ideographic, pictographic, or composite (see Figure 4).

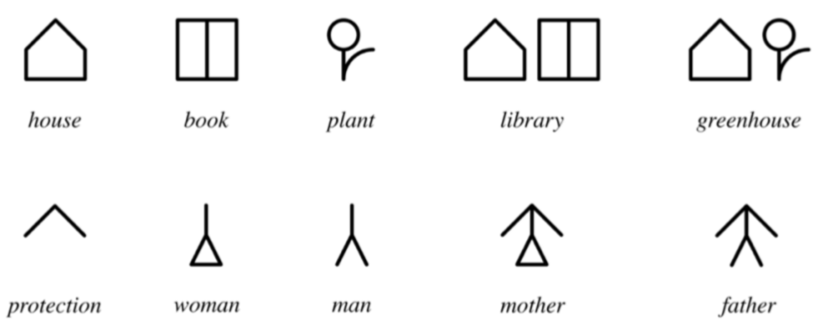
Figure 4 – Examples of Blissymbols types of characters.



Source: Author, based on Blissymbolics Communication International (2004).

The symbols may appear alone with their basic semantic meanings, combined with others to form new words, or merged to extend concepts (see Figure 5).

Figure 5 – Examples of Blissymbols characters.



Source: Author, based on Blissymbolics Communication International (2004).

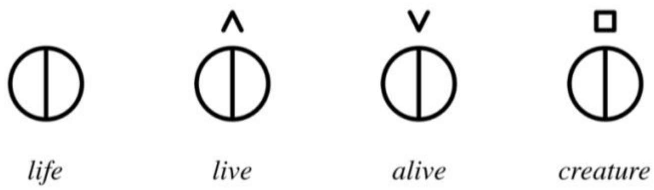
Blissymbols are constructed inside a matrix divided by two main lines, the skyline (top) and the earth line (base). Most symbols are placed between the skyline and the earth line, but their positioning in the matrix can also be related to their meaning (see Figure 6).

Figure 6 – Blissymbols matrix and examples of characters' positioning in the matrix.

Source: Author, based on Blissymbolics Communication International (2004).

Indicators are used as markers for grammatical or semantic modification, allowing for transforming symbols into verbs, adjectives, adverbs, or indicating objects or plurals (see Figure 7).

Figure 7 – Examples of Blissymbols' markers usage.



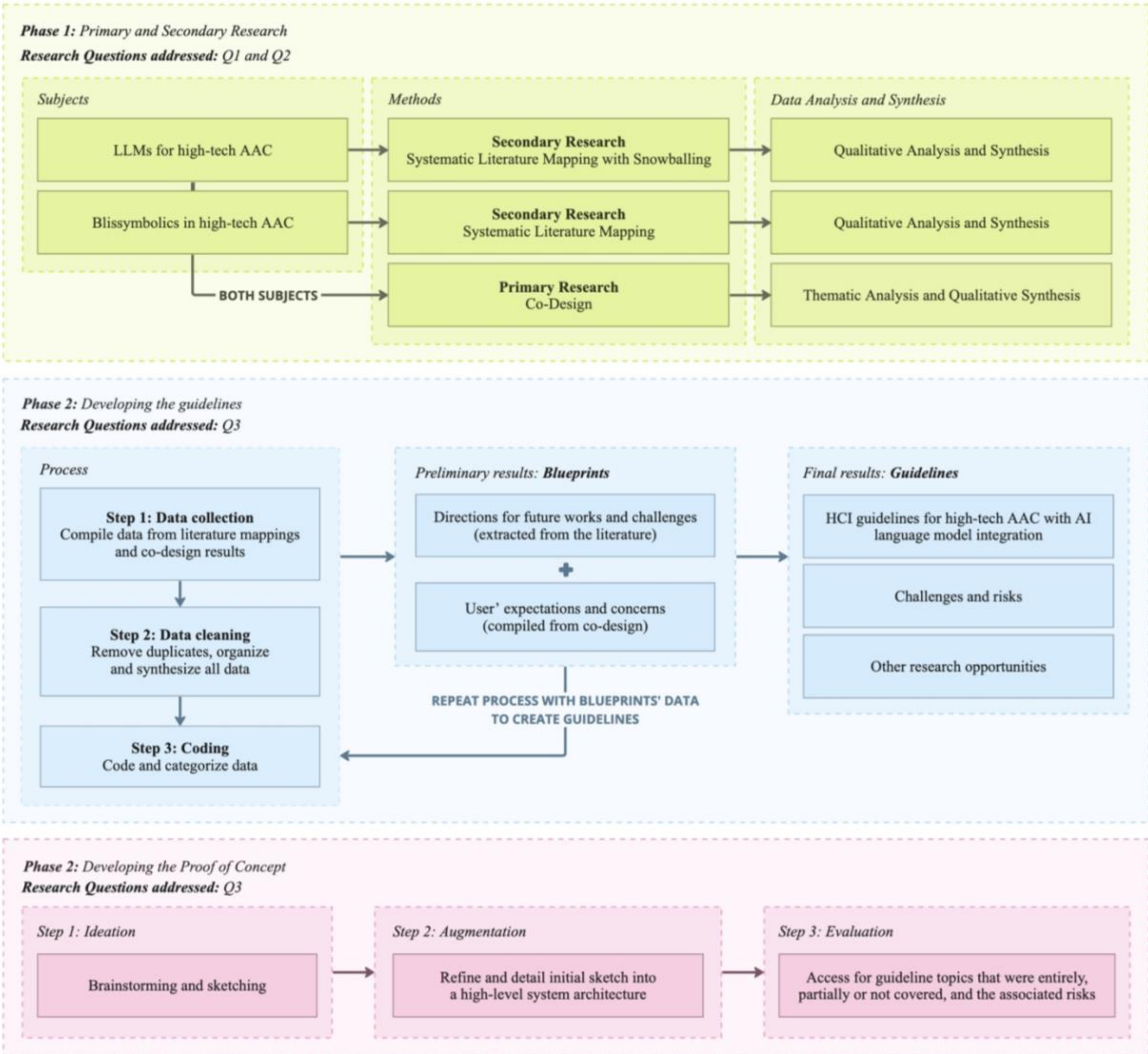
Source: Author, based on Blissymbolics Communication International (2004).

It should be emphasized that, regardless of the standard rules, the conceptual nature of the symbols allows for unique ways of expressing the information. The same idea might be represented using different symbols or altering the order of a compound character. For instance, in Hungary, the order of the symbols was changed to reflect Hungarian word order, while in Israel, symbols were written from right to left (Okrent, 2009). The rules described here are just a brief summary of the most significant corpus of Blissymbols rules that are regulated by Blissymbolics Communication International (2004).

3 RESEARCH METHOD

To achieve this study's main goal and address all questions defined in subsection 1.2, the research method was divided into three phases. The first phase involved conducting both primary and secondary research to establish the foundation of this work. The second phase focused on developing the guidelines, and the third phase involved applying them to create the Proof of Concept for this work. The contents of each research phase are detailed in the following paragraphs, and Figure 8 shows the complete research method structure.

Figure 8 – Research method structure.



Source: Author.

Phase 1 addressed the first and second research questions (Q1 and Q2) and sought to cover the two main topics of this work (i.e., LLMs for AAC and Blissymbolics in high-tech AAC). Consequently, the results of this phase provided the groundwork for the subsequent phases. For that, data was collected from: (i) the relevant literature, by conducting two secondary research covering each research topic separately; and (ii) alongside AAC users, covering both research topics simultaneously.

The secondary research comprised two systematic literature mappings in which similar methods were applied for data collection, analysis and synthesis, ensuring consistency. The studies follow a qualitative approach and provide readers with a synthesized and comprehensive analysis of the relevant literature. Three aspects are evaluated and discussed in each work mapped: strengths, weaknesses, and directions for future works drawn from it.

The primary research followed an Inclusive Design approach, using the co-design method, thematic analysis, and qualitative synthesis of the findings. Four collaborators were recruited, including individuals with communication needs, familiar listeners, and an expert in the Blissymbolics language. Throughout the four sessions course, co-designers shared their concerns and expectations, discussed the benefits of the Bliss language and brainstormed ideas for future AI and AAC applications.

Phase 2 addressed the third and main research question (Q3). The outcomes from the previous phase were used to generate the proposed guidelines, following a three-step process. Two blueprints, termed as such, were created as a preliminary result of individually collecting, cleaning and coding the data from the primary and secondary studies. The first blueprint presented directions for future work extracted from literature, while the second included user's expectations and concerns compiled from the co-design.

After repeating the same process using data extracted from both blueprints, the final guidelines were generated, and additionally, the existing challenges, risks, and opportunities for future work were identified. The guidelines' results presented three principles: reliable, customizable, and adaptable. Each principle covered seven specific topics that were individually reviewed and synthesized. This work provides a detailed analysis of all 21 specific topics identified.

Phase 3 addressed the third and main research question (Q3). The proposed guidelines were used as the foundation for developing a conceptual AAC system with LLMs integration. A three-step process was followed to create this Proof of Concept, which included ideation, augmentation, and evaluation.

4 SECONDARY RESEARCH: SYSTEMATIC LITERATURE MAPPINGS

In this section, the two secondary research are detailed along with their results. The works are split into two distinct topics: (i) LLMs for high-tech AAC, which were analyzed by conducting a systematic literature mapping followed by a snowball sampling (in subsection 4.1); and (ii) Blissymbolics in high-tech AAC, which were also analyzed by conducting a systematic mapping and, additionally, documenting extra resources for the readers (in subsection 4.2).

The two systematic mappings discussed in this section utilized a qualitative analysis and synthesis method. However, the first mapping employed a descriptive approach to present the findings, where the works identified were concisely summarized and examined. In contrast, the second mapping adopted a more discursive approach, where the results were reported by thoroughly summarizing, comparing and discussing the works.

4.1 LARGE LANGUAGE MODELS FOR AAC

Much has been researched on how to translate the nuances and complexities of human language and knowledge into a form that the computer could understand. The natural language processing (NLP) field focuses on allowing a computer to recognize and analyze, modify, or generate natural human language. Russell and Norvig (2010, p. 16) explain that "Understanding language requires an understanding of the subject matter and context, not just an understanding of the structure of sentences". Hence, in order to build a robust intelligent agent able to understand natural language, it should have a general knowledge about the world, its objects and relations, rather than merely reproducing grammatical rules.

Older NLP iterations were notably laborious as they required a detailed analysis of linguistic rules which significantly increased its computational complexity. On the other hand, modern NLP systems apply machine learning (ML) and statistical language models to process the user's input. Such models are trained by collecting large text corpora for training data and, then, the system identifies patterns and makes predictions accordingly (Senott *et al.*, 2019).

More recently, with improvements in computational power and the development of the transformer model (Vaswani *et al.*, 2017), NLP advanced notably with language models capable of generating content that is indistinguishable from human-written text (Clark *et al.*, 2021). Models such as the Bidirectional Encoder Representations from Transformer (BERT) (Devlin *et al.*, 2019) and the Generative Pretrained Transformer (GPT) (Brown *et al.*, 2020) are built using a huge amount of data (e.g., 3,3 billion words for BERT's training dataset) which allows them to perform various tasks including text generation and summarization, question answering, translation, and writing programming codes; such models can also be fine-tuned for better targeted results.

Another promising approach to overcome some NLP challenges, especially in the AAC field, is integrating context-awareness (Higginbotham *et al.*, 2012; Senott *et al.,* 2019). For example, the application can detect if the user is in a coffee shop, in the morning, and offer different communication options, similar to "Can I get a cup of coffee with double sugar and double cream?", based on the users' preferences. In this scenario, the system provided adaptable options based on the user's environmental data collected, such as location, time and prior language use or personal preferences.

OpenAI recently introduced a similar functionality, called "Custom Instructions", to their popular tool ChatGPT. It became available for free in 09/08/2023, and allows users to (i) provide personal information about themselves to enhance the responses and (ii) specify how they would like the bot to reply (OpenAI, 2023). Still, in the high-tech AAC field, contextual input is yet an unexplored input modality (Curtis; Neate; Gonzalez, 2022)

Unfortunately, a huge gap in AI research for AAC is the lack of a robust dataset by and for the AAC community (Curtis; Neate; Gonzalez, 2022). For instance, Google's Project Euphonia (Google, 2021?) has tried to partially tackle this issue within the speech recognition domain, which faces major challenges. Google showcased in a promotional video their ML-based tool which allows users to train the system to recognize their own unique voice pattern. The system's capability was demonstrated by accurately recognizing the voice of an individual with dysarthria after recording several voice messages (Google, 2019).

Applying NLP technologies to the AAC field is not something new. As a matter of fact, NLP has been long and widely used in the AAC field for assisting with keyboard arrangement, speech recognition and word completion and prediction (Higginbotham *et al.*, 2012). The last one has been mainly employed to, theoretically, increase keystroke savings and reduce physical load, as it allows the user to quickly select the desired word instead of typing each letter. However, Koester and Levine (1996) identified that word prediction alone can increase AAC user's cognitive load and slow communication; since it shifts their attention from typing and building the message, to also scanning the suggested options.

Despite that, several works highlight the advantages of exploring contextual information and artificial intelligence in AAC, considering its exponential growth and availability (Higginbotham *et al.*, 2012; Elsahar *et al.*, 2019; Senott *et al.*, 2019). This could assist not only with output generation, but also in enhancing the system's customization and adaptability – consequently, helping to better suit the system to each user's needs. Senott *et al.* (2019, p. 14) states that "AAC systems and devices powered by various AI tools hold potential to help give people with CCN enhanced pathways to solve the participation challenges they face when their speech and/or language capabilities do not allow them to fulfil their communication needs."

With these opportunities in mind, this study sought to search, analyze and critically discuss the existing works that applied state-of-the-art large language models to the AAC domain. Hereby, this work aims at:

1) synthesizing and critically discussing the current relevant works that explored the application of LLMs to high-tech AAC;
2) identify the strengths and weaknesses of each work, and detail directions for future works.

4.1.1 Method

Considering the novelty of this topic, two different methods were applied for data collection. Firstly, the PRISMA framework proposed by Moher *et al.* (2009) was used, which consists of 4 phases: Identification, Screening, Eligibility and Included. Second, a frontward and backward snowballing sampling was applied to identify other relevant works that could not be found on the initial search. The complete data search and collection framework is illustrated in Figure 9.

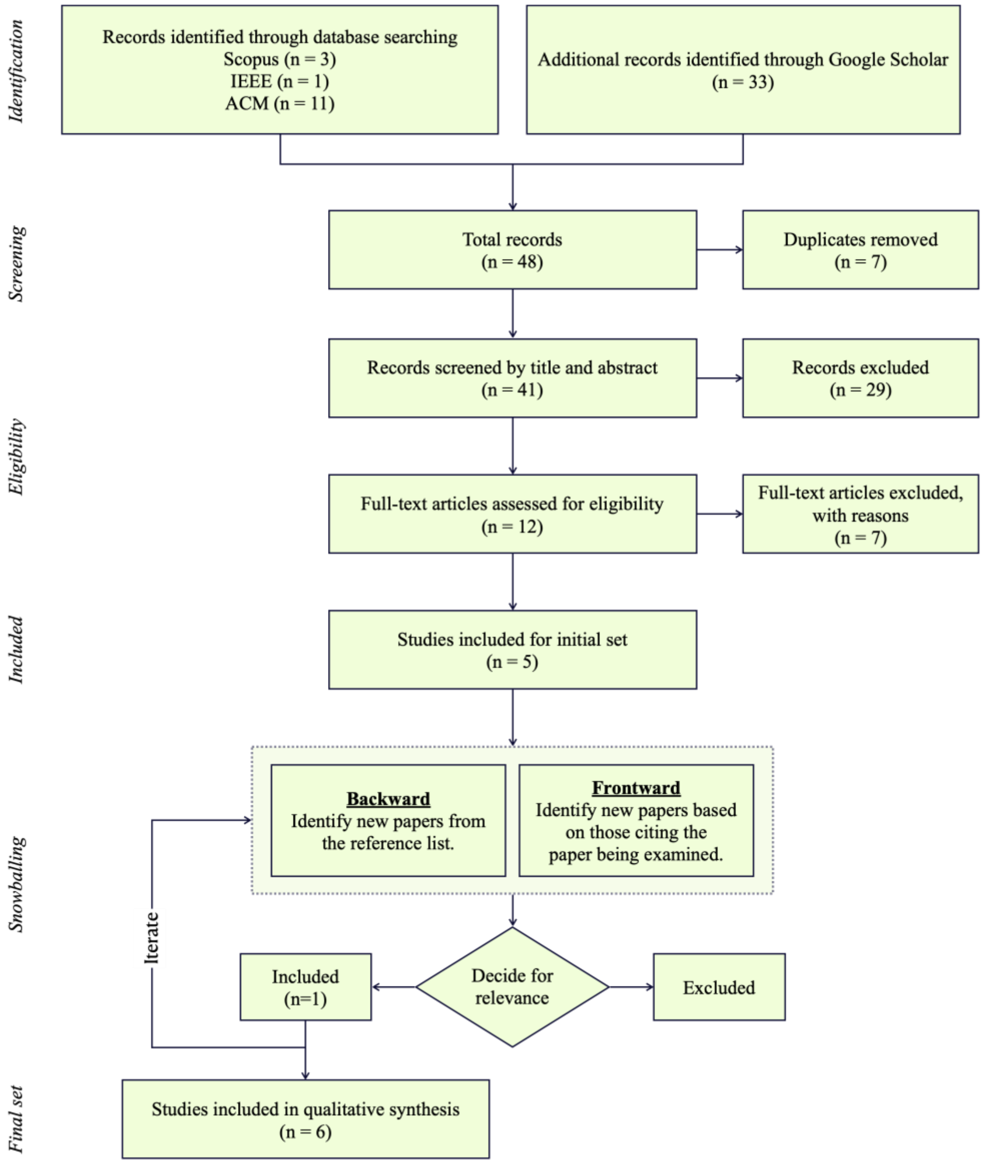
For the Identification phase, the search comprised both the full-length terms and their respective abbreviations. Two queries were used in all database searches: (i) `("large language models" AND "alternative and augmentative communication");` (ii) `(LLMs AND AAC).`

On 02 October 2023, data was extracted from three Digital Libraries: Scopus, IEEE and ACM. On the same day, additional data was collected using Google Scholar search engine. For that, results were sorted by relevance and not restricted by publication time or type. Applying both queries in Google Scholar resulted in a total of 3,773 papers. However, the data collection was restricted to the first three pages results, since from page four onward no significant literature was noticed.

In light of the newness of this subject, the data collection was not restricted by year and followed a search within all metadata. However, several criteria were defined to ensure the selection of relevant papers. Firstly, papers included must experiment and analyze the application of LLMs within the high-tech AAC field, such information should be explicitly included in the work's title or abstract.

Secondly, when retrieving multiple articles from the same author and about the same work, only one of the papers was included in the study. Doctoral or master's theses, books, book chapters, and preprints were not included. Additionally, the works should be open access or accessible through our university's resources.

Figure 9 – Flow diagram of the systematic mapping, combining the PRISMA framework (Moher *et al.*, 2009) and a snowball sampling.



Source: Author.

4.1.2 Results

After selecting the initial set of papers extracted from digital libraries, we conducted a detailed analysis of all the papers referenced within each selected article or those referencing it. Following this snowballing process, several articles underwent evaluation; however, only one successfully met the criteria for inclusion detailed in subsection 4.1.1 and advanced through the "Decide for relevance" phase. Therefore, this study is limited by the fact that only one article could be found using the snowballing method. Most articles simply referred to the same related works, which were already retrieved in the initial set of papers (n = 5).

We speculate that this might have happened due to the limited number of works available on this topic, as the use of LLMs is relatively new. Considering that, although a timeline restriction was not defined for the data collection process, all results publication years were either 2022 (n = 2) or 2023 (n = 4). Secondly, high-tech AAC may not receive much attention when compared to other fields. Consequently, future research should reevaluate the existing literature on this topic, as an increase in the number of works is expected. The final set of papers for this mapping is listed in Table 1.

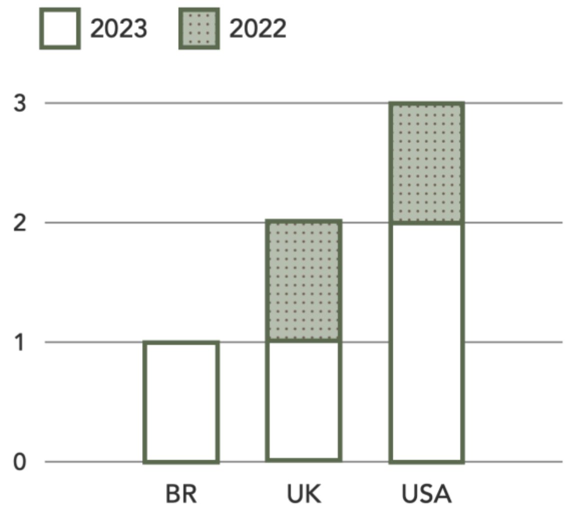
Table 1 – Final set of papers.

ID	Article	Title
1	Shen *et al.* (2022)	KWickChat: A Multi-Turn Dialogue System for AAC Using Context-Aware Sentence Generation by Bag-of-Keywords
2	Valencia *et al.* (2023)	“The less I type, the better”: How AI Language Models can Enhance or Impede Communication for AAC Users
3	Cai *et al.* (2022)	Context-Aware Abbreviation Expansion Using Large Language Models
4	Yusufali, Goetze and Moore (2023)	Bridging the Communication Rate Gap: Enhancing Text Input for Augmentative and Alternative Communication (AAC)
5	Pereira *et al.* (2023)	An Augmentative and Alternative Communication Synthetic Corpus for Brazilian Portuguese
6	Liu and Smith (2023)	Adapting Transformer Language Models for Predictive Typing in Brain-Computer Interfaces

Source: Author.

As shown in Figure 10, the studies retrieved took place either in the United States of America (n = 3), United Kingdom (n = 2) or Brazil (n = 1). Interestingly, two of the three works from the US were carried out by the Google Research team. It was also noticed that this team contributed significantly to the field as they represented the majority of papers found during the frontward snowballing process. Despite these works being quite relevant for high-tech AAC solutions, only few explored the use of LLMs and those were included in this study.

Figure 10 – Publication year and country distribution.



Source: Author.

Following the second objective, as outlined in subsection 5.1, Table 2 highlights the strengths of each work, while Table 3 outlines their weaknesses or drawbacks. It is noteworthy to mention that while most of the information described below was taken directly from the evaluated articles, some of the final notes were derived from our interpretation and analysis of the works.

Table 2 – Table of work's strengths (continue)

ID	Strengths
1	1) Detailed reporting of the system design, model and evaluation; 2) Context-awareness is built upon both the user's background information and conversation history.
2	1) Explored the use of LLMs in generating replies for targeted situations; 2) Detailed reporting of the evaluation results; 3) Tested with real users and gathered valuable insights; 4) Voiced concerns from AAC users about the use of AI in AAC tools.

Source: Author.

Table 2 – Table of work's strengths (conclusion)

ID	Strengths
3	1) Detailed reporting of the experiments; 2) The highly abbreviated input might not be easily decoded by other people, which increases privateness; 3) The abbreviation expansion allows the user to increase control over the desired output, as opposed to prompting for text generation; 4) Experimented with both prompting and fine-tuning using 4 different datasets.
4	1) Detailed reporting of their four contributions (i.e., user behavior modeling, selection interfaces for word-prediction, investigation of human response time, fine-tuning of LLMs for AAC text-entry); 2) First author is an AAC user who conducted informal experiments to test their work; 3) Tested the quality of generated output using two models across seven different datasets for fine-tuning.
5	1) Tackles a major gap in AAC research by using LLMs to generate synthetic text data for an AAC corpus focused on the Brazilian Portuguese language; 2) Collected initial set of sentences from real AAC stakeholders (i.e., speech therapists, psychologists, and parents of children with CCN).
6	1) Considered the models' size impact in a real system environment; 2) Evaluated different models; 3) Experiments conducted with a dataset created by individuals with Amyotrophic Lateral Sclerosis.

Source: Author.

Table 3 – Table of work's weaknesses and drawbacks.

ID	Weaknesses and drawbacks
1	No AAC users were involved in the process.
2	1) Poor evaluation of the generated sentences; 2) Tested only using few-shot prompting.
3	1) No AAC users were involved in the process; 2) It was not discussed the applicability of such highly abbreviated text in different situations, which can be very useful for daily activities (e.g., answering simple questions), but might not perform similarly in more complex conversations.
4	Poor reporting and discussion of the results.
5	1) The data collected was not directly created by real AAC users; 2) Seemed to be focused solely on children, specifically for a classroom environment; 3) The issues of generating a potential biased or limiting dataset after the data cleaning phase were not considered or discussed.
6	1) Lack of examples on the tasks and models' suggestions; 2) Poor discussion on the applicability and potential downsides; 3) Context was only considered at a word level.

Source: Author.

4.1.3 Discussion

As a result of the first aim of this mapping, the selected articles will be synthesized and critically discussed as follows. After that, we provided directions for future works, which were drawn based on each study's results. Those can be found in Table 4. Similar to the previous subsection, some of the information present in Table 4 is based on our interpretation and analysis of the works.

Aiming at reducing the communication rate gap between AAC users and non-AAC users, Shen *et al.* (2022) created a dialogue system called KWickChat (Keyword Quick Chat). Their work was the first published article that explored the use of LLMs for the high-tech AAC domain. Their memory-efficient and open-domain model (fine-tuned GTP-2 language model) incorporates contextual information and conversation history with "bag-of-keywords" as input.

That is, the KWickChat generates several suggested sentences based on the keywords input by the user, their personal information (previously provided by the user and stored in the system) and using information from the current conversation. Additionally, the authors foresee the possibility of fine-tuning the model with data from a specific user. This could allow the system to learn the user preferences and conversational behaviors and, consequently, generate more accurate suggestions. Further, the authors experiment with auto-complete for word prediction and analyze the effects of the amount of sentences displayed to the user. Despite that, the system was evaluated using an envelope analysis and human judges for semantic analysis, but not with AAC users as they affirmed having considered the ethical issues in testing an incomplete tool.

Valencia *et al.* (2023) conducted a study with 12 AAC users that tested and reviewed live-generated suggestions using a LLM-based interactive prototype. Their work draws from the previous KWickChat project and, instead of the bag-of-keyword model, they propose what they call "speech macros". Similarly, their reply-based approach also focuses on prompting a LLM to expand short inputs into full sentences, but targeted to certain pre-defined circumstances (i.e., extending a reply, replying with background information, turning words into requests).

Valencia's prototype was built upon Google's LaMDA language model and remotely tested by real users that pointed out some of the benefits and limitations. They reported that the users were excited about reducing the amount of effort required to compose sentences and tended more to the functionality of turning words into requests. On top of that, users brought crucial insights into improving the system and incorporating processes to which they were already used to (e.g., using shortcuts).

However, several things were reported as being drawbacks of the prototype and its generated outputs: (i) sometimes it used an inappropriate tone; (ii) replies were too abrupt and sounded “robotic”; (iii) it made false assumptions and used arbitrary or over specific information; (iv) a “safety layer” in the original model might have restricted conversations in more sensitive topics (i.e., smoking); (v) suggested sentences did not reflect the users personality and style.

Users were also concerned about privacy, loss of control, and the societal stigma of automatic-generated replies. As reported by previous studies, the amount of effort invested into communication and message composition is closely related to building and maintaining relationships (Kelly *et al.*, 2017). Valencia's work brings in-depth analysis of AAC users' perception on language models for high-tech AAC, which can serve as a valuable guide for future research in this field.

To accelerate text-entry,the work of Cai *et al.* (2022) introduces an abbreviation expansion system enhanced by LLMs and using a context-aware approach. Similar to Shen *et al.* (2022), their tool uses the conversational context to expand the users’ abbreviated input into a full phrase. For instance, if someone questions “Where is the dog?” the AAC user can type “iipitb” and the system will expand it to “It is playing in the backyard”. Compared to the previous works described, this approach allows the user’s to have more control over the desired output and increases privacy, as the abbreviations can not be easily decoded by other people.

The authors simulated typing noise to evaluate the system’s tolerance to typos and compared its performance using few-shot prompting and fine-tuning. Their quantitative-only evaluation showed that the fine-tuned LLMs could accurately expand 48-77% of abbreviations that were replies to initial turns of dialogs, with a keystroke saving rate of 73-77% for correctly predicted phrases. Lastly, they highlight that, in conversational settings, context-awareness is crucial for an effective abbreviated text entry.

The work of Yusufali, Goetze and Moore (2023) encompasses different issues within the same scope. The authors contribute by: (i) introducing a general user behavior modeling for predictive text-entry, based on HCI laws; (ii) modeling interfaces for text input with word-prediction; (iii) investigating the human response time according to the number of choices displayed; (iv) fine-tuning LLMs for AAC text-entry. Informal experiments were conducted by the main author, who is an AAC user.

For the last listed contribution, their fine-tuning of the BERT and RoBERTa (Robustly Optimized BERT) models included corpora from: TV episodes and shows, spontaneous telephone conversations, magazines, newspapers, academic texts, blogs, Wikipedia, Reddit, and more. In addition, they used an AAC corpus, which consists of 6,000 imagined messages from users who imagined themselves using AAC devices.

For the interface layout evaluation, their results showed that radial layouts outperformed other text-entry interfaces (such as grid or list distribution) as it offers closer movement distance but, consequently, increases the risk of error selection. For the LLM results, the authors show that fine-tuning the model using conversational corpora produces higher accuracy compared to generic ones (e.g., Wikipedia). They highlight the importance of choosing the right corpus for fine-tuning according to each application, as domain-specific corpora could be suitable for certain tasks and helpful for long-term AAC. The best results in their work were achieved by fine-tuning the model using the AAC corpus, which emphasizes the urgency for a robust and large-scale AAC text dataset.

Pereira et al. (2023) uses LLMs to generate a corpus of synthetic text for AAC in Brazilian Portuguese. Despite the recent growth in AAC tools usage, most of them are yet centered on the English language, and thus, leaving other languages underrepresented. The authors explain that translation is not always an effective approach considering the semantic and grammatical differences of both languages. To fill this gap, the authors collect sentence samples from 18 professionals (i.e., speech therapists, psychologists, and parents of children with CCN). The initial set included 667 unique sentences that the professionals considered to be the most commonly used by AAC users. Later, another set of 203 sentences were collected for testing purposes.

For the vocabulary, the authors used the Brazilian Portuguese version of a pictogram-based dataset that includes a set of keywords for each pictogram. Using few-shot prompting, they utilize the GPT-3 language model to augment the initial set of sentences which are used as examples for generating new phrases with similar structure. Their text cleaning phase included removing sentences: (i) containing offensive content; (ii) with higher perplexities; and (iii) with less than three or more than 10 tokens.

It should be noted that such a strategy for cleaning could eventually bias the corpus and limit conversations on sensible topics, as happened in Valencia’s work. Along with the safety and protection of children who are AAC users, their privacy and autonomy should also be considered. Another important topic that could be addressed by future studies is the average number of tokens in AAC corpus sentences and if, somehow, if it has any impact on the communication rates of experienced AAC users.

Their results showed that word, stop-word and bigram with more frequency included “eu quero” (i.e., “I want”), which indicates that their corpus might be centered on expressing wants or desires. They conclude that their three-step method (i.e., data collection, few-shot corpus augmentation, and cleaning) generated a corpus semantically similar to the initial set of sentences created by the professionals and, consequently, can be useful for ML tasks and research in the AAC domain for Brazilian AAC users.

Considering a typing task in BCI, Liu and Smith (2023) evaluate several pretrained language on their ability to accurately predict the user’s intended input. While their work did not experimented with current larger models (e.g., GPT-3, BERT), the authors explain that they opted for analysing models that would be large enough to be effective on several tasks (as supported by prior works), but small enough to run on a single machine for deployment in a real BCI system. The authors tested the performance of several models, including GPT and GPT-2. The last one has 1.5 billion parameters has and was considered a large model when it was fully released in November 2019. For comparison, the currently used GPT-3 has 175 billion parameters.

Their evaluation considered the model’s performance in several tasks including word prediction, context length, and input noise. The experiments were conducted with two datasets, the first contains 2,082 anonymized messages created by individuals with Amyotrophic Lateral Sclerosis; while the second is a collection of 260 hours of transcripted two-sided telephone conversations.

The results showed that the most difficult task for all language models was predicting the first characters in words. On the other hand, predicting the later characters of a word was an easier task and the models performed expressively better than the unigram baseline model, particularly when longer context lengths were available. Across all metrics, GPT and GPT-2 overperformed the other models tested but were significantly impacted by the presence of noise.

Despite having opted for a quantitative-only analysis for this work, the authors further celebrate that their collaborators have started to test the language models with real BCI users and hope to see its future results.

From all the works described in this study, only two of six articles did not mention engaging with any AAC stakeholder and opted for a quantitative approach. We consider this to be a positive outcome even if most did not engage the end users from early-stage development.

Whereas, the majority of work involved professionals or AAC users, either as a researcher in the project or as participants.

Table 4 – Table of directions for future works drawn from the studies.

ID	Directions for future works
1	Future studies could (i) evaluate the impact of the “golden reply” versus the system-generated and semantically equivalent reply, and (ii) experiment with the system learning from the user’s preferences and unique communication behavior.
2	1) Enhance the user control over the final output and increase customization in LLMs-generated output by allowing the user to, for example, adjust the message tone; 2) Pre-stored information can help with context-awareness and personalization, such as medical details, relationship situation with conversational partner, work information and environmental data (e.g., location, time, date); 3) New AAC applications should not disconsider what already works for AAC users, but integrate their existing strategies into novel solutions, such as pre-storing messages when preparing for future conversations or using the shorthands they are used to.
3	1) It is important to consider typo tolerance in LLMs-based text generation/expansion; 2) Considering their results, fine-tuning overperformed prompting; 3) Currently, improving accuracy comes at the expense of increasing latency which might increase hardware requirements to be managed.; 4) User Interface approaches could offer alternatives to partial selection in incorrect predictions, instead of resorting to fully sequential text-entry.
4	1) In their informal experiments, radial layouts outperformed grid and list distribution, but increased the risk of error selection; 2) For word-prediction, alphabetical ordering resulted in higher communication rates compared to probabilistic ordering; 3) Select the appropriate dataset for specific applications; 4) The field inevitably requires a robust AAC corpus, since even tests using a fictitious AAC user dataset outperformed other generic corpus.
5	Their proposed method using few-shot prompting with the GPT-3 model for text augmentation proved to be effective for generating sentences semantically similar to those used by AAC users. This clears the way for building novel and robust AAC datasets, especially for users whose first language is not English.
6	1) Consider the model’s size deployment and its applicability on real devices; 2) Consider the performance impact under noise input; 3) Experiment with and take advantage of language models’ abilities on dialog-level context.

Source: Author.

Even though only six articles exploring LLMs for AAC software were found, we consider all of the works identified here to be an outstanding and valuable resource to AAC researchers, designers and developers. Not only that, but these works also open doors to novel research in this field. We hope to see more projects in the future exploring this topic, especially involving and giving agency to AAC users as decision-makers throughout the process.

4.2 BLISSYMBOLS AND HIGH-TECH AAC

Analyzing and critically evaluating previous works is a crucial step prior to generating suggestions for new applications in any field. Hence, a systematic mapping was conducted focusing on the existing literature that seeks to assist Blissymbols communicators and learners using high-tech AAC. Given the information available at the moment this work was carried out, no other systematic mapping or review was found on the same topic.

This systematic mapping aims at:

1) documenting and critically discussing the works and projects executed across a wide range of sources that seek to contribute to the communication of Bliss users;
2) providing the readers (researchers, designers, and developers) with a synthesized listing of the relevant literature on human-computer interaction for Bliss users;
3) draw directions for future works from the existing literature.

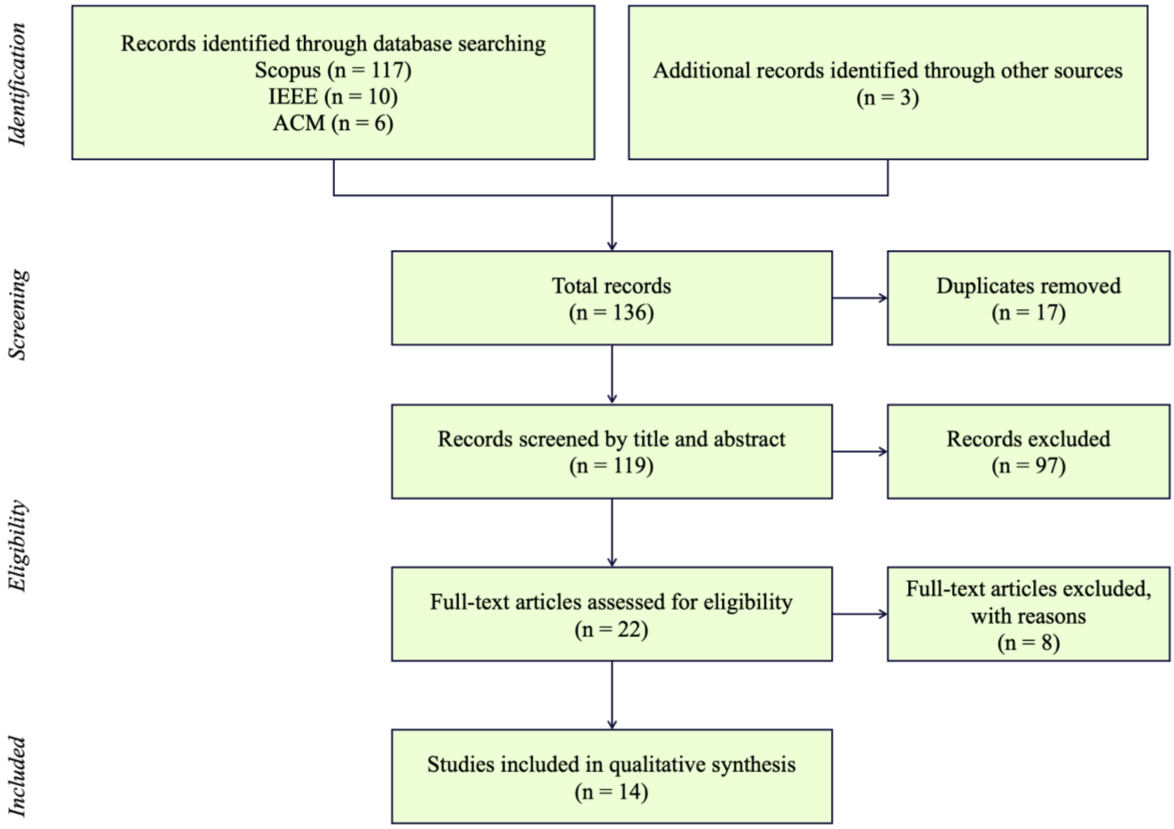
4.2.1 Method

Similar to the previous mapping, the framework for this study was grounded on the guidelines of PRISMA. For the Identification phase, it is essential to encompass a wide range of keyword variations when searching for articles. Thus, three different ways of referring to Blissymbols were considered: (i) Bliss symbols; (ii) Blissymbols; (iii) Blissymbolics. Then, they were used to build the final query used in all database searches: `(“bliss symbols” OR blissymbols OR blissymbolics)`.

Three well-known Digital Libraries were used to extract data (Scopus, IEEE and ACM) in 19 May 2023, and done within title, abstract and keywords. Additional records were collected during a background search on the field (n = 2) or received by e-mail from external collaborators (n = 1). The complete data search and collection framework can be seen in Figure 11.

Several criteria were defined to ensure the selection of relevant papers that met the aims for this mapping. Firstly, the works included should not use the symbols exclusively as a language option for the system without bringing new contributions to Bliss communicators. Secondly, they must be in some way related to the fields of computer science or human-computer interaction; that is, articles about other topics (e.g. language development and acquisition, cultural variations, teaching methods) should not be included. Third, records discussing solely issues related to encoding and the Unicode standard were excluded. Lastly, only one article was included when retrieving multiple studies from the same authors that discussed the same project.

Figure 11 – Flow diagram of the systematic mapping (following PRISMA 2009).



Source: Author.

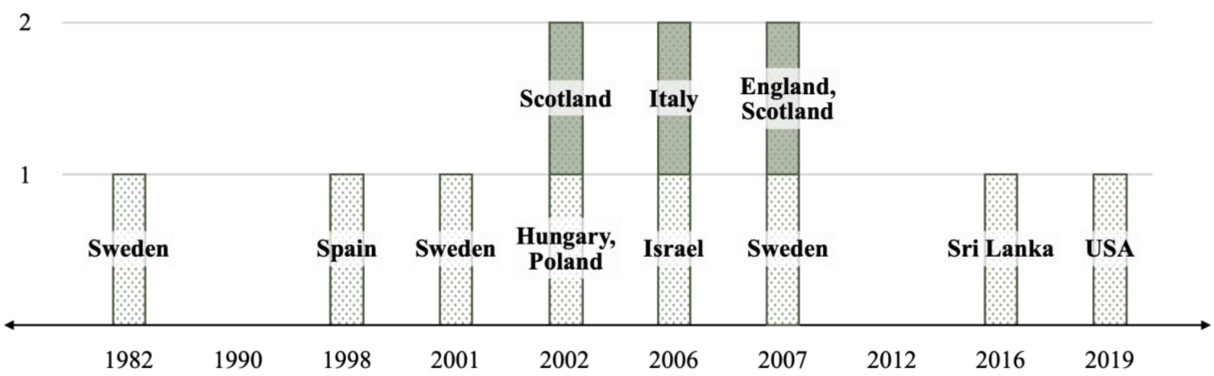
A restriction timeline was not defined for the articles selected, considering that Blissymbolics usage was prominently widespread between 1970 and early 2000. However, this can also be considered a constraint to the present systematic mapping since some relevant articles might not have been digitally archived and computerized.

4.2.2 Results

Following the aims of this systematic mapping defined previously, the eleven articles identified through database search were cataloged and analyzed. Accordingly, some of the topics addressed are: (i) the distribution of publication's year and author's country; (ii) the Bliss input and system evaluation methods chosen; (iii) the number of users engaged in the process; (iv) and other information such as strengths and weaknesses, (v) and directions for future works. In this subsection, the results are described and critically discussed. Furthermore, the three additional resources identified through sources other than database searches are discussed in the following subsection.

As shown in Figure 12, Sweden and Scotland are the only countries with multiple publications on the topic, with four recurring authors such as Annalu Waller, Sheri Hunnicutt, Rolf Carlson, and Björn Granström. This disparity is presumably an effect of the ongoing popularity and dissemination of Bliss among ACC users and children with special needs in Sweden. The Blissymbolics website states, "In Sweden, around 2-300 persons yearly are attending Bliss instructional courses, primarily arranged by the Swedish Agency for Special Needs Education and Schools" (Blissymbolics Communication International, 2023?b).

Figure 12 – Graphical distribution of the number of publications per year and the author's country.



Source: Author.

The publication years range from 1982 to 2019, in which only 2002, 2006 and 2007 had two publications each. As pointed out previously, Blissymbolics' usage was more prominent between the '70s and early '00s, and we believe that several relevant papers might be unavailable in digital format. Efforts for manual search and computerization of these resources should be highly encouraged. Additionally, not all relevant works can be found exclusively in digital database libraries, as some have not been publicized by mainstream academic means (e.g., journals or conferences); those works are detailed as additional resources.

As described in Table 5, the methods chosen for input ranged from using existing hardware (Ferm; Amberntson; Thunberg, 2001) and software (Waller and Jack, 2002; Gatti and

Matteucci, 2006; MacKay, Wills, and Waller, 2007), to custom-made interfaces (Olaszi, Koutny and Kálmán, 2002; Netzer and Elhadad, 2006; Hunnicutt and Magnuson, 2007) or newly developed devices (Carlson, Granström and Hunnicutt, 1982; Seneviratne and Basnayakege, 2016).

Table 5 – Input methods and evaluation (continue)

Article	Bliss input method	Evaluation
Waller and Jack (2002)	Bliss-Word system, which is built "on top of" Microsoft Word.	The accuracy of translation was tested with 20 sentences using three sizes of source text.
Olaszi, Koutny and Kálmán (2002)	Custom graphical interface with twelve symbols displayed at a time, with an automatic delay-based cursor.	The system was installed at the HCMC and continuously improved from 1994 to 1996.
Netzer and Elhadad (2006)	Custom digital lexicon, search can be done by keyword, or by semantic/graphic component.	Keystroke savings, comparing the number of choices needed to generate a set of sentences with their system.
Carlson, Granström and Hunnicutt (1982)	A 500-symbol Bliss board with multilingual text-to-speech system	Not mentioned.
Gatti and Matteucci (2006)	An AAC communication software based on Bliss called Bliss2003.	Quantitative analysis for prediction of training error and generalization error.
Coltell *et al.* (1998)	Special input devices are handled as the mainstream input ones.	User tests, not concluded by the time of publication.
Ferm, Amberntson and Thunberg (2001)	Touch Talker™ overlay combining 54 basic Blissymbols with Minspeak strategy.	Case study with continuous adjustments to fit each user particularity.
Hunnicutt and Magnuson (2007)	Selection of preprogrammed messages, a grammar-guided writing or free selection of symbols.	Beta user tests for adjustments.
Sohail and Traum (2019)	No user interface. Each Bliss symbol, word or character, was given a unique 4-5 digit numeric ID for input.	Quantitative analysis using BLEU algorithm.

Source: Author.

Table 5 – Input methods and evaluation (conclusion)

Article	Bliss input method	Evaluation
MacKay, Wills, and Waller (2007)	Dasher, a text-entry interface driven by continuous two-dimensional gestures.	Not mentioned.
Seneviratne and Basnayakege (2016)	Headband with built-in accelerometer used as cursor to point to a smart Bliss board using head movements.	Gradual evolution of the system development followed by testing it with an user.

Source: Author.

Coltell *et al.* (1998) prototype was compatible with a wide range of special input devices (e.g., hand-switches, mouth-switches, foot-switches, head-pointer switches) and handled them as mainstream input devices (e.g., keyboard or mouse), creating an adaptive and customizable system. On the other hand, Sohail and Traum (2019) did not develop a user interface but applied a numeric ID for symbol input.

Regarding the evaluation methods chosen, most works that used NLP predictive approaches opted for testing the accuracy of the translation task. The results generated were compared either using different sizes of word sources (Waller and Jack, 2002) or tested using the BLEU algorithm (Sohail and Traum, 2019). However, both work results showed significant errors, indicating the necessity for major improvements in the use of NLP for the Bliss writing system prior to introducing it for widespread utilization.

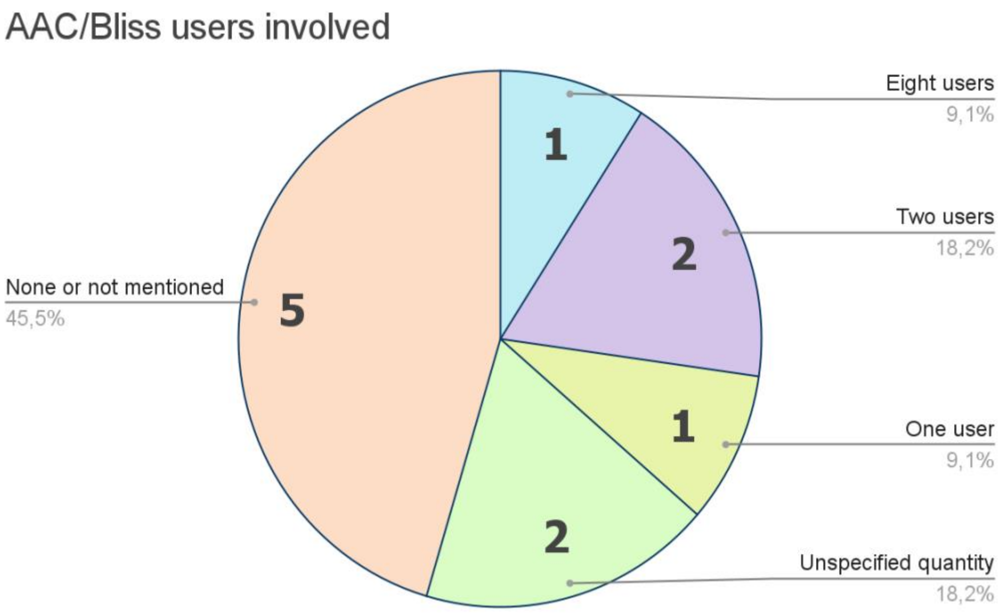
Other works that opted for quantitative analysis evaluated their system using error prediction (Gatti and Matteucci, 2006) or checking the presence of keystroke savings when constructing sentences (Netzer and Elhadad, 2006). Works that followed a qualitative approach conducted either user tests (n = 3) or a case study (n = 1), whereas only the last one stands out for having a detailed reporting of the process and its results (i.e., Ferm, Amberntson and Thunberg 2001). In contrast, the user tests conducted were not sufficiently described.

When collaborating with users, none of the projects stated having AAC or Bliss communicators as active decision-makers guiding the project from the beginning. Regardless, both Olaszi, Koutny and Kálmán (2002) and Ferm, Amberntson and Thunberg (2001) engaged with users for a longer period, allowing them to constantly adjust their system to the users' needs and preferences.

As pointed previously, it is critical that individuals with CCN are engaged early in the research and development process, as they are the experts and should not be considered merely as testers. Technology must be used as a tool and people should guide the process to ensure it serves their needs, not the other way around. As emphasized by Light and McNaughton (2013), one of the pillars to create meaningful AAC interventions is to work collaboratively with AAC users and their families to identify and evaluate the technologies that fit the needs, skills, and preferences of the individual who relies on it.

In contrast, as illustrated in Figure 13, most studies did not involve users at all during the development and opted for a quantitative analysis of the system (Waller and Jack, 2002; Netzer and Elhadad, 2006; Sohail and Traum, 2019) or did not mention conducting an evaluation phase (Carlson, Granström and Hunnicutt, 1982; MacKay, Wills, and Waller, 2007).

Figure 13 – Graphical representation of the number of users involved per project.



Source: Author.

Down below, the eleven articles are detailed according to the following information gathered: (i) a brief summary of the work; (ii) strengths; (iii) weaknesses and drawbacks; (iv) suggestions for future works, drawn from each study. A complete table including all the results can be found in Appendix A as a synthesized resource for the readers, which results from the second aim detailed in subsection 4.2. In contrast to the previous study presented in subsection 4.1.2, there are few instances where the directions for future works described here were directly referenced in the studies. Most suggestions outlined below are primarily based on our interpretation and analysis of the evaluated articles.

I. Waller and Jack (2002)

About: Bliss to English translation system that adjusts the output to generate a grammatically correct translated sentence instead of using a literal translation of each symbol from the gloss.

Strengths: Pioneer in using predictive techniques to translate Bliss sentences into grammatically correct English rather than traditional laborious techniques of natural language using rule-based syntactic parsing.

Weaknesses and drawbacks: Some concerning errors occurred, such as the insertion of inappropriate words and incorrect translations.

Directions for future works: Users might not input all the symbols required to generate a grammatically correct utterance. Hence, a translation system should not only guarantee that the gaps within the sentence are appropriately filled but, more importantly, without altering the meaning.

II. Olaszi, Koutny and Kálmán (2002)

About: A symbol-to-speech system called Blissvox with a linguistic module that transforms the strings of isolated words into grammatically correct Hungarian sentences.

Strengths: 1) Detailed reporting of the system design, technical aspects and usage; 2) System requires only a simple PC and a synthesizer kit; 3) System was used by a Bliss user for an extended period; 4) System was improved continuously during the 2 years of experimentation.

Weaknesses and drawbacks: 1) The linguistic module and its results are not entirely detailed in this paper; 2) Brief description of the experimentation phase.

Directions for future works: 1) It is crucial to fully evaluate the hardware that will be used for the system and favor cheaper, easy to install and more reliable options; 2) The English-oriented nature of Bliss symbols imposed challenges for translating to Hungarian and using language processing should be preferred but thoroughly evaluated.

III. Netzer and Elhadad (2006)

About: NLG-based (Natural Language Generation) system for Bliss users using semantic authoring, whereas the available choice of symbols dynamically changes according to the syntactic and semantic content of the selected symbols.

Strengths: 1) Addresses the context challenge by adding pre-defined, yet flexible, inputs into the system and enabling the assignment of default values to each conversation; 2) The intervention provides early feedback as it occurs during the process of composing the input sequence; 3) Implementation of a Bliss lexicon for both Hebrew and English.

Weaknesses and drawbacks: 1) Evaluation tests were conducted with non-AAC users; 2) Evaluations showed no improvement in keystroke saving.

Directions for future works: 1) When processing a Bliss sentence, real-time feedback of the output result is preferred; 2) Extra caution when parsing highly abbreviated telegraphic sentences, which can be incorrectly deciphered and misinterpreted since conversations rely heavily on pragmatics.

IV. Carlson, Granström and Hunnicutt (1982)

About: Details the development, structure and operation of both the software and the device for the BlissTalk project. Its framework uses syntactic parsing to create a multilingual text-to-speech system.

Strengths: 1) Detailed reporting of the system design, technical aspects and usage; 2) Uses a multilingual system.

Weaknesses and drawbacks: 1) No AAC or Bliss users engagement was mentioned; 2) No evaluation was conducted.

Directions for future works: 1) Text-to-speech systems should offer ways of improving the naturalness of synthetic speech (e.g., allowing pause insertion for prosody); 2) Provide access to voice output customization (i.e., different tunes based on age and gender); 3) Importance of building adaptive models that are able to follow the user's development.

V. Gatti and Matteucci (2006)

About: An automatic scansion system for Bliss sentences composition focusing on AAC users with motor impairments. It is called CABA2L and uses an Auto-Regressive Hidden Markov Model.

Strengths: 1) Detailed reporting of the system design, model and technical aspects; 2) Analyzes both human-computer interaction and technical strategies; 3) Adopts a hybrid prediction approach that allows to model the user's unique language behavior.

Weaknesses and drawbacks: Poor description of the graphical interface and input method chosen.

Directions for future works: Language prediction systems must take into account that each verbally impaired person develops their own syntactic model according to their capabilities, and so applying pure statistical or syntactical approaches may be inadequate.

VI. Coltell *et al.* (1998)

About: COMBLISS, a conceptual model of a communication system that operates interactively using speech synthesis, graphic and textual representations of Bliss symbols, and includes settings management options for tutors.

Strengths: 1) System is compatible with multiple special input devices that are handled as the mainstream ones; 2) Evaluation tests were conducted with eight AAC users.

Weaknesses and drawbacks: 1) No description of the graphical interface; 2) Included a user profile for tutors (more access and operative privileges), but did not discuss the AAC user privacy and autonomy issues related; 3) Poor results due to the prototype limitations; 4) Unfinished evaluation by the time of publication; 5) Reported that the AAC user must acquire enough experience in the system to improve the usage performance.

Directions for future works: 1) Offering compatibility with a wide range of assistive devices might prevent the system from being exclusively used by a favored group; 2) The system must be easy to install or customize by the end-users; 3) Whenever opting to include different profiles for tutors that allow for more access and operative privileges, it is crucial to consider the AAC user's privacy and autonomy.

VII. Ferm, Amberntson and Thunberg (2001)

About: Details the development of a Minspeak™ Swedish-language phrase-based application using Blissymbolics. The work was carried out in the form of case studies with two experienced Bliss users and their tutors. The application turned out not to be sufficient for the communicatively advanced study participants, who then required a second word-based vocabulary to maximize voice output communication aid use.

Strengths: 1) Detailed narrative of the case study and results; 2) Included an option that allowed the user to easily program personal private vocabulary; 3) The word-based application was easily incorporated into one of the user's daily activities as they were not required to plan and program new vocabulary in advance; 4) Improved the reading and writing abilities of one user; 5) Overall general positive results with both users involved.

Weaknesses and drawbacks: 1) They reported that one of the users was a perfectionist who wanted to correctly construct every sentence, so he had difficulties using the word-based application as it slowed down his communication; 2) Reported having issues with the sound and voice characteristics of one of the user's voice output communication aid; 3) No description of the interface.

Directions for future works: 1) Offer options that allow the user to independently customize and access private vocabulary; 2) The system must be able to adapt to the user's unique language behavior and ability, as some may opt to use highly telegraphic communication while others might prefer a long and grammatically advanced sentence; 3) Caution when opting for a system that requires the user to memorize codes to access the desired output, as this might increase the learning curve and cognitive load, especially for novice users.

VIII. Hunnicutt and Magnuson (2007)

About: A grammar-guided email writing tool for AAC users called Grammar-Guided Writing (GGW) where the user can choose preprogrammed messages and email phrases or select graphic signs freely without grammar guidance.

Strengths: 1) Detailed reporting of the system design, interface, technical aspects and usage; 2) Multi-language system; 3) The systems operate using a fixed word order, and each word in the output is under the control of the users to preserve their intention; 4) Beta tests were conducted with two Bliss users; 5) The user can learn to access the grammatical structures gradually using their stepwise introduction.

Weaknesses and drawbacks: No reporting or analysis of the beta tests was conducted.

Directions for future works: Adding an introductory mode to the AAC and Bliss systems can be helpful to guide novice users and ease the learning curve. However, it should also allow expert users to customize the system to their preferences and needs.

IX. Sohail and Traum (2019)

About: Based on previous work (Waller and Jack, 2002), the authors build a proof of concept for a new translating system from Blissymbols to English.

Strengths: Added the possibility to handle alternate spellings and unseen conjugations (e.g., past tense, gendered words).

Weaknesses and drawbacks: Results show some errors from either the language model or due to the Blissymbolics structure.

Directions for future works: A text-to-speech context-aware component can be important for supporting the users to be more independent of a human interpreter.

X. **MacKay, Wills, and Waller (2007)**

About: Implementation of Bliss communication in the Dasher system, which is a text-entry interface that requires few gestures and is adaptable to different AAC devices.

Strengths: 1) Open-source system available for download; 2) Detailed reporting of the system design.

Weaknesses and drawbacks: 1) No AAC or Bliss users engagement mentioned; 2) No evaluation mentioned.

Directions for future works: The absence of a Unicode standard for Bliss was reported to be an issue. However, it is essential to evaluate if there is an alternative option within the technology available nowadays to address this problem without limiting the language.

XI. **Seneviratne and Basnayakege (2016)**

About: A system that consists of a desktop-based interface displaying a Bliss Board and headband device with an accelerometer that detects the users' head movements to select the symbols.

Strengths: 1) Experimentation with a novel device for head interaction; 2) The system does not require the tutors to concentrate on each child at a time, as opposed to traditional methods.

Weaknesses and drawbacks: 1) Poor reporting of the evaluation; 2) Accuracy and usability of the device developed were not measured.

Directions for future works: Consider the use context and prefer a scalable AAC system to different situations rather than exclusive to one environment.

4.2.3 Discussion

From the works described in the previous subsection, some experimented with natural language prediction techniques to a translation system for Blissymbols. Instead of having a literal translation of each symbol from the gloss or using traditional laborious approaches (e.g., rule-based syntactic parsing), Waller and Jack (2002) were the pioneers in applying statistical prediction using a word trigram to translate Bliss.

Likewise, Sohail and Traum (2019) draw on the previous work and add the possibility of handling alternate spellings and conjugations. Their proof-of-concept language model identified nouns in the sentence using WordNet and the training set was composed of 15 bliss utterances from children’s books. However, both studies share the same drawbacks as they reported concerning translation errors during the evaluation phase.

A user might communicate using telegraphic language and compose sentences that omit some characters, as each individual with speech limitations develop their own way of communicating. For example, instead of asking “Where is the cat?” or “Have you seen the cat?” the communicator can simply input “Where cat” and the system should be responsible for generating comprehensible outputs. However, the grammatical correction of such telegraphic sentences should not be met at the expense of altering the user's intention and message.

With this in mind, Gatti and Matteucci (2006) adopts a hybrid prediction approach that allows to model the user's unique language behavior. The authors propose an automatic scansion system for sentence composition that receives the last selected symbol from a BLISS2003 hardware, calculates the most probable four symbols according to the established requirements, and scans them in a graphical interface.

On the other hand, a skilled communicator can opt to use grammatically complex and long sentences. In this context, a highly simplified system that offers few options may be burdensome as it can increase the user's mental and physical overload to compose the desired outputs and, consequently, affect their real-time conversations. This circumstance is outlined in the work of Ferm, Amberntson and Thunberg (2001), who experimented with a semantic compaction system in both phrase and word level.

In addition, the authors described that during the case study a disagreement took place between the communicator and his parents and educators, who could not agree on whether he should be allowed to use swear words to express negative feelings. Later in the development, the authors added the possibility to insert private vocabulary. This emphasizes the importance of considering the user's privacy and autonomy when developing AAC software. In particular, when projects include different profiles for communicators and their tutors, and allow the last one to have more access and operative privileges, such as the work of Coltell *et al.* (1998).

Coltell *et al.* (1998) highlights that the users are not technical specialists, and so the system should be easy to install or customize and explains the importance of offering compatibility with a wide range of assistive devices. The authors states that:

> If this most sophisticated systems needs highly qualified people to operate, or special installing or operating conditions, or the mass production cost is too high; the disability-aid initiatives cannot include it in their budgets, and this complex or expensive system can only be issued to a restricted handicapped population.” (COLTELL *et al.*, 1998, p. 640)

MacKay, Wills, and Waller, 2007 partially addresses this issue by using the Dasher system to create a Bliss communication tool. Dasher is an open-source text-entry interface that requires few gestures and is adaptable to different AAC devices. Contrarily, Seneviratne and Basnayakege (2016) created a new device based on head movement interactions. The hardware is a smart headband coupled with an accelerometer that captures the users head movements as cursor, allowing them to interact with a virtual Bliss board. However, the evaluation and discussion of such systems are poorly described.

Asynchronous conversations are a great way of communication for individuals with speech impairments, as they allow for constructing the message at their own pace and ease the tension existent within real-time conversations with non-familiar people. Considering this, Hunnicutt and Magnuson (2007) create a grammar-guided email writing tool for AAC users.

The Bliss Talk was one of the early hardware systems for communication-aid created by Carlson, Granström and Hunnicutt (1982), it used a rule-based parsing approach and was a multilingual symbol-to-speech electronic communication board. Similarly, the project Blissvox (Olaszi; Koutny; Kálmán, 2002) created a communication tool for the Hungarian language that required only a simple PC and a synthesizer kit. Despite providing a brief description of the experimentation phase, the Blissvox system was reported to have been continuously improved during the 2 years of research and development with an AAC user. On the other hand, the Bliss Talk system did not report having any AAC users engaged nor detailed having conducted any evaluations.

To avoid issues related to telegraphic parsing, Netzer and Elhadad (2006) propose an intervention during the process of sentence composition by using an approach called Semantic Authoring. For this, their system's available choice of symbols dynamically changes according to the syntactic and semantic content of the selected symbols. However, their evaluation tests were only conducted with non-AAC users and showed no improvement in keystroke saving.

The information gathered from the eleven articles included are not generalizable as each project has its own strengths and weaknesses. The systems developed are extremely diversified with several functionalities overlay; some aimed at a translation and symbol-to-speech systems, while others developed tools for assisting with writing. The most present weaknesses are related to the evaluation phase, such as not engaging with AAC users, brief description or not conducting an evaluation.

4.2.4 Additional resources

As mentioned before, some resources for building high-tech AAC systems are not found on mainstream academic platforms. In light of that, additional materials are described in this subsection. They were sent by external collaborators and are not included in the previous results as their format differs from the regular articles included in the systematic research.

The first resource is Lundälv (2012), it showcases the state-of-the-art of Blissymbolics in ICT (Information and Communication Technologies) platforms for the 2012 year. Nonetheless, most works described are either discontinued, non-updated or lack public information. The three projects displayed under the category "AAC software with specific support for Bliss" are detailed in Table 6 according to their name, description (ipsis litteris as in LUNDÄLV, 2012), and situation as of May 2023.

Table 6 – Projects described in Lundälv (2012) and their current situation.

Project	**Description**	**Situation**
Symbol for Windows	A comprehensive AAC package with a fairly well updated Blissymbol vocabulary, and special features for Bliss support	Discontinued.
Mind Express	This communication software includes Blissymbolics support (though hardly no public information provided), medium well updated vocabulary, and an editor for new symbols etc.	It is being constantly updated, but still lacks information on the Bliss function.
MisterBLISS	Italian web based or locally installed cross-platform dedicated Bliss software package. Published 2010. So far only in Italian.	Still the same as described.

Source: Created by the author, based on Lundälv (2012).

The second material is a paper briefly documenting the BlissWord tool, that was used as an input method in Waller and Jack (2002). It is an extension for Microsoft Word allowing for inputting symbols into the system. It allows the user to access the symbols by browsing: (i) the categories; (ii) for the components that create the symbol (e.g., the symbol for Ostrich is a combination of the symbols for Bird and Legs); (iii) English gloss; or (iv) the next word prediction.

The third and last resource is a master's thesis (Shearer, 2021) where the author creates a system that predicts a novel Bliss word meaning based on its composition. For example, when combining the symbols for "rodent", "little", and "pig", the system might predict the meaning as “hedgehog”, “porcupine” or “guinea pig”. For that, the project uses the symbol's English gloss to search for its hyponyms in WordNet, then compares articles on the Simple Wikipedia website that contain such words, gives each word a score and shows the top ranked results. Unfortunately, the project did not engage AAC or Bliss communicators in the process. The author explains certain gaps in the system such as the word mismatches between Simple Wikipedia and WordNet, and the inability of the system to consider the structural combination of the words.

5 PRIMARY RESEARCH: CO-DESIGN

5.1 METHOD

The present study involves working with a minority community that is vulnerable and already highly misrepresented and underrepresented. Consequently, we opted for an inclusive design approach, which focuses on the perspective of marginalized individuals, rather than following current trends of designing only for the mass-population (Pullin *et al.*, 2017). Opting for an inclusive approach allows us to look closely at each individuals needs and wants, instead of excluding their idiosyncrasies and generalizing their experiences.

Solutions that follow a "one-size fits all" approach might create a huge mismatch in what is experienced by their users – who are plural and can have diverse backgrounds, such as cultures, age, gender or language. We argue that the user should not have to accommodate themselves to the technology, but the other way around. The inclusive design approach has its roots in interaction design, which is not centered on how people interact with technology, but rather on individuals interacting with one another through technology (Inclusive: The film, 2015).

This idea of human-human interaction mediated by technology is deeply related to the AAC field and, considering the AAC community, the challenge of embracing plurality is even greater as individuals with CCN are highly diverse; and diversity increases complexity. The field of AAC echoes the principle of "nothing about us without us" (Blackstone; Williams; Wilkins, 2007), which emphasise the importance of including AAC stakeholders in the development of anything that is related to them. This resonates with the inclusive design approach, as it focuses on designing *with* rather than *for* individuals.

Interaction and inclusive design are usually described as exploratory methods, as "they do not seek to test or validate solutions, rather they use design to illuminate and provoke new insights that might indirectly inform or inspire solutions in the future" (Pullin *et al.*, 2017, p. 31). This emphasizes that these approaches are suitable to achieve our research goals of building suggestions for future works, as mentioned in subsection 1.2.

Our primary research followed a co-design approach, which is commonly used in inclusive design research. For this, the Virtual Reality & Multimedia Research Group (Federal University of Pernambuco, UFPE) partnered with the Inclusive Design Research Center (Ontario College of Arts and Design University, OCAD), mediated by Canada's Emerging

Leaders in America program. This research was submitted for OCAD's Research Ethics Boarder and got an approval on 2023/09/12 under the number 102400.

Six research questions were formulated to guide the collaborative process of identifying challenges and opportunities for applying LLMs and Blissymbolics to high-tech AAC with individuals who have CCN and familiar listeners.

1) Considering their current AAC device, what already works for AAC users?
2) Considering their current AAC device, what does not work for AAC users?
3) What expectations do AAC users and familiar listeners have about integrating AI and LLMs into AAC devices?
4) What concerns do AAC users and familiar listeners have about integrating AI and LLMs into AAC devices?
5) What ideas do AAC users and familiar listeners have about integrating AI and LLMs into AAC devices?
6) What are the benefits for individuals with CCN, if any, of integrating Blissymbolics into a LLM-based AAC system?

Previously, we planned to use two different methods to accommodate each individual's needs. Initially, we tried applying what we called asynchronous and synchronous co-design activities. The asynchronous co-design activities happened over email exchange with collaborators with speech limitations. Co-designers were not required to reply to the email within a time limit or following a certain structure, and were invited to communicate independently, at their own pace, using their preferred tone and grammar, and exchanging only information that they felt comfortable sharing.

On the other hand, the synchronous co-design activities were carried oust remotely and individually with either a familiar listener or a Bliss expert. In both asynchronous and synchronous activities co-designers were invited to share their experiences, expertise and ideas for the subject. However, we noticed that engaging the AAC users in the asynchronous activities proved to be a challenging task. This is due to the fact that it was hard to maintain the conversations, as progress was slowed due to the delay in response time. Finally, we opted to turn all activities into synchronous co-design session.

Each session took about one and a half hours and one participant required two sessions to complete the discussion. Besides the Zoom platform for video conferencing, we used the Miro tool during the sessions. However, as Miro does not offer robust accessible options for individuals with mobility limitations, the tool was used for displaying the main conversation topics for prompting the conversation, and as a form of transparent notetaking. As shown in Figure 14, the board's structure was not fixed, and they could be adjusted to better fit each session's conversation pace.

Figure 14 – Miro board used for the co-design sessions.

Source: Author.

During each session, it was emphasized that nothing would be shared or published without the AAC users' approval, as this work would be done *with*, and not *for*, them. After the data collection phase, the researcher watched each session recording and, accordingly, organized the Miro board to ensure that all relevant information was included.

Thereafter, those boards were sent back to each participant so that they could add or correct any information. Collaborators who use AAC could not interact with Miro due to lack of accessibility support for those with mobility limitations. To mitigate this problem, different forms of visualization were sent via email, which allowed them to evaluate the material and reply with their thoughts on it.

After conducting a thorough analysis of the co-design notes using thematic analysis (Adams; Lunt; Cairns, 2008), the key findings were highlighted in a comprehensive report, which is detailed in subsection 5.3. A draft of this report was previously sent to the collaborators for their final review and feedback. This collaborative review process – for both the Miro board and the findings' report – helped to ensure that all information shared was approved by the AAC users and stakeholders involved in the study.

5.2 CO-DESIGNERS

This study focus on recruiting a small group of highly skilled co-designers that are prone to collaborate with in-depth information. The goal is not to seek saturation or generalization in data but to gather enough information to understand the individual's experience and guide the development of future works.

Prospective co-designers were recommended to us by two collaborators non-profit organizations: (i) Blissymbolics Communication International, which is a charitable organization that holds the perpetual, worldwide rights for the use and publication of Blissymbolics (Blissymbolics Communication International, 2023?c); and (ii) Ottawa Foyers Partage, that provides quality and innovative support to persons with disabilities in Ottawa, Canada (Ottawa Foyers Partage, c2023).

Five individuals demonstrated interest on collaborating in this project, they were three AAC users, one familiar listener and one Blissymbolics expert. Unfortunately, during the co-design process, one of the AAC users faced difficulties accessing his communication device which could not be solved before the deadline for this research. Consequently, he could not attend the sessions and the activities were carried out with the other four collaborators.

All co-designers were adults, resided in Canada, and opted – in the Consent Form – to have their names included in this research with their ideas associated with it. Consequently, the names displayed is this research are the individual's real names. Table 7 details each collaborator's name, group (AAC user, familiar listener and/or Blissymbolics expert) and information.

The recruited individuals with CCN had cerebral palsy, which affected their verbal speech production. Each non-speaking collaborator had their caregivers or family members to support their participation by helping them access the computer or clarify their message. Staff caregivers were not considered as participants for this research, but they often contributed with their perspective which were included in the results.

Table 7 – Research co-designers' information.

Name	Group	About
Samantha Millar	AAC user	Samantha is a cheerful 30-year-old who loves listening to audiobooks and driving her pontoon boat.
Julie Millar	Familiar listener	Julie is a retired registered nurse and mother of two, one of them being Samantha.
Treena Guy	AAC user	Treena is a very kindhearted and loving woman who enjoys interacting with others.
Shirley McNaughton	Bliss expert and familiar listener	Shirley is a well-known advocate for individuals with speech limitations (see subsection 2.2.1)

Source: Author.

In preparation for the session, co-designers were provided with the main topics that would be discussed, in this way they could prepare in advance if wished. The researcher also asked if: (i) collaborators had any preferred style of communication, for example, if they are comfortable with open-ended questions or if they preferred only yes or no questions; or (ii) if there was anything that we could provide or support with to make the meeting as comfortable as possible. In addition, co-designers were encouraged to suggest new topics for our conversation.

5.3 FINDINGS

The results discussed bellow were organized into four categories according to each research question they answer: (i) "What works and what doesn't", for the first and second research questions; (ii) "Expectations and concerns", for the third and forth research questions; (iii) "Ideas", for the fifth research question; (iv) "Benefit of Blissymbols", for the sixth research question.

5.3.1 What works and what doesn't

Samantha and Treena shared their experiences using AAC devices and highlighted what works and what does not work for them. Other collaborators also contributed to this topic by sharing their perspectives as familiar listeners and supporters. Overall, collaborators shared more negative than positive aspects of their experience using AAC devices.

Samantha pointed out how her layout worked greatly for her. She prefers face-to-face communication and uses two head switches for static column scanning with a multiple pages layout. For instance, her page called "Urgency" includes sentences such as "It's time for my medicine" or "Can you get me a drink of water?", while "Question starters" features words such as "who", "what", and "where". However, Samantha mentioned preferring her low-tech options, as they keep the listener engaged in the conversation as they scan the text with her.

Treena was eager to share that she could use her device to make phone calls, send text messages and control the television. Additionally, she mentioned that one positive aspect of it was being able to use her eyes, as she used eye-gazing for input.

Co-designers shared different opinions when questioned about self-expression using AAC devices. Treena mentioned that her personality is more straightforward, and thus it fits well with what she is able to do with her device. On the other hand, Samantha complained about "missing her personality". When she was younger, her family used to manually program her favourite jokes on the device. However, they couldn't continue doing it and she can not do all the programming by herself. Her mother stated that "That's missing from her, like, that side of her is missing from her computer. It's not there. Her personality is not really there."

Collaborators also shared their frustrations, some regarded the device limitations such as requiring assistance to move, since the device was unable to control the wheelchair. Other issues emerged during social interactions, Treena mentioned how some people would make fun of her way of communicating without her AAC device. However, when communicating with her device, nonfamiliar listeners were usually inpatient and would not wait for her to finish typing.

Another uncomfortable situation voiced was when people slipped behind them and started to watch what they were doing on the device or look at what was being typed. Moreover, AAC users were annoyed by people that would try to guess what they wanted to say before they could finish composing their message.

Shirley and Julie, familiar listeners, brought to attention that while some AAC users were independent, others required a companion and those had less control of their lives and struggled to speak for themselves. Consequently, most relied on their caregivers to program the devices, which often was challenging and required much manual labor. Moreover, users often struggle to adjust to new devices, as they must learn a new system that may not be compatible with their previous one. This often increases the learning curve and the demand for support in customizing it.

It was also mentioned how most individuals with CCN are highly dependent of their device or letterboard to communicate and, when it is not available, they often can not express themselves. On top of that, not all devices are portable, and even among those that are, not all prove to be suitable or adaptable to each individual's needs. For example, Julie mentioned that Samantha would like to use an iPad to communicate, but it doesn't work for her as Apple's scanning option is very limited and can not be adapted to her needs. Julie explained that the iPad works great if the user uses direct-access input (e.g., touchscreen), but not for block-scan users. It uses an item per item scanning, which makes it very tedious and challenging for indirect-access users.

AAC users also pointed out how hard was for them to access the internet as most websites and applications are not accessible. Samantha shared that she enjoyed watching videos and reading articles on her device, but oftentimes it was extremely frustrating as things online are usually not consistent and most interactions are designed only for mouse users. In addition, aggressive use of pop-ups and adds impeded her from having access to the internet, even when having ad-blockers enabled. Consequently, when trying to use the internet by herself, she constantly gets "stuck" on something and needs further assistance to proceed.

5.3.2 Expectations and concerns

Co-designers shared their expectations and concerns regarding a possible future integration of AI language models into AAC devices and software. As a familiar listener, Shirley stated that she expects AI to "exceed what we can do now", by acting as a companion. Moreover, she hopes that AI responds to the needs of each individual users rather than generalizing from the masses.

When asked about her concerns, Shirley voiced that she was concerned about the users wasting their time with incorrect suggestions generated by the language model. On top of that, she was apprehensive about the current rush for creating new AI applications and worried that technology would advance more quickly than Blissymbolics and AAC users' knowledge. She highlighted the importance of building meaningful things instead of rushing to create new technology.

Regarding sentence generation features, Shirley emphasized that AI should avoid making interpretations from the bigger populations of users and questioned: "How does the AI identifies the user's intention? Is the user telling a story or describing something? Did it happen yesterday or a month ago? How can the technology generate accurate suggestions when the context is more complex?”

Julie and Samantha, familiar listener and AAC user respectively, shared that they expected AI technology to assist with message composition and make it faster to get the message across. However, they hope it can anticipate the user’s sentence without necessarily directing the conversation but offering the paths to it. Regarding their concerns, Samantha did not point out any, but rather manifested her excitement for future applications.

On the other hand, Julie stated being worried about privacy, especially audio capturing features. She emphasized that even though it was a huge issue, this would not discourage her from choosing such technologies, as she would evaluate the risks opposed to benefits of using them. Julie also mentioned that there is not any sensible information on Samantha’s device that would impede her from making use of AI.

Furthermore, Julie expressed her concern, as a parent, about AAC users relying too much on the technology to direct their thought process. She explained that, instead of using your creativity, having AI lead the conversation path might generate an effect similar to losing the ability to do mental calculations when depending excessively on digital calculators.

Nevertheless, Treena, AAC user, mentioned that she would like AI to help her “find the right words” and considers that a sentence generation feature would be helpful, especially considering a grammatical correction layer. Besides that, she wishes for an AI assistant that could also help her with movements, for example, by commanding “take me to the kitchen” and having her wheelchair automatically controlled. Similarly to Julie, Treena also considers privacy to be her number one priority and worried about audio capturing.

5.3.3 Co-designer’s ideas

We sought to expand the co-designers’ expectations mentioned above and emphasize their ideas for future AI technologies in the AAC domain. None of the concepts detailed here were created by the researchers, all ideas were brainstormed by each collaborator during the co-design activities.

Firstly, the familiar listener Shirley’s raised the following question: “what AI can do that the human companion can not?” She highlighted two crucial points that, in her opinion, language models could contribute to AAC: consistency and total memory. Shirley explained that AAC users often need support from people that tent to come and go off their life, for instance, a support staff might be changed. On the other hand, the AI companion must be consistent and should always be available for the user.

On top of that, computers are able to easily store and retrieve information. That “total memory” ability might facilitate communication by relating previous information to the current context. For example, suppose an AAC user has previously solved a communication problem with an individual in a particular context. In that case, the system should remember it and have that information ready for use in the same or similar context.

Moreover, Shirley mentioned how technology should also assist the user in group conversations. For instance, by supporting them in situations where the conversation topic has changed without allowing the user to express themselves. She pointed out that developers should consider two key-factors for building novel AAC systems: (i) individualized, that is, to design thinking about each user’s particularities rather than generalizing and reducing their options; (ii) learning and communication modes, which might allow the user to expand their knowledge and learn continuously rather than solely communicating passively.

As an AAC user, Samantha mentioned how, sometimes, she prepares a message for an upcoming conversation, but its flow does not necessarily end up following what she predicted. In this situation, her preprogrammed message can not fully cover what is now being discussed. For that reason, she pointed out that AI could adapt her preprogrammed phrases to the new context.

The same applies to updating the information available in other preprogrammed texts. For instance, Samantha has a short introduction about herself prepared for whenever she meets someone new. However, data such as her age needs to be updated annually, which is not done automatically by the system but has to be manually adjusted by her. That is why, when Samantha talks about her dog, she cleverly opted to mention its birthday instead of age. This prevents her from having to manually correct that information each year. Julie, familiar listener, further explains their idea for having AI adapt and provide paths to the conversation:

> Julie Millar: “Sometimes she (Samantha) is a little bit telegraphic, especially if she's speaking to people who don't know her. Because by the time she gets all of the words that go along with the thought, to make it a proper sentence, it's taken time. So if she gave you the main idea, kind of thing, you know, have the AI be able to to get her down that path. (...) To help her along with that sentence structure, to make it faster, to help her, hmm…”
> Researcher: “Like, to extend her main idea?”
> Julie Millar: “It can extend the main idea, but extend it to put it into the right context. So if she wanted to talk about, I don't know, she wanted to talk about her dog. She might just say ‘Riley’ and then the AI would then anticipate. Okay, so want to you tell somebody about your dog? Or do you want to go down, like, give her the paths that she can go. So it's not directing her exactly, but it's helping, it's making it quicker for her and adapting to some of the conversations.”

Moreover, Julie shared her idea of an AI that could assistant Samantha when navigating the internet and helping to make it more accessible for her. As an example, it could automatically detect, close website pop-ups and avoid ads. She mentioned that it could also help her with commands automation, which she currently does manually – such as for allowing Samantha to login in websites. Besides that, Julie stated that Samantha often faces difficulties when trying to use the automations she prepared due to connection issues. As the automations do not execute as expected in such situations, she believes that AI could help by predicting the slow connection and controlling the automation's commands so that they are executed in sequence.

Treena, AAC user, discussed a wide array of ideas regarding the use of AI for AAC devices. Firstly, she mentioned how a massive issue for her is being unable to independently control or move things using her device. She believes that an AI virtual assistant could help with mobility and, consequently, increase her independence, as she would be able to go out by herself. She wished to use such technology to command specific movements, such as opening doors and windows or moving her wheelchair.

Afterward, Treena remarked that such technologies should also be easily accessed from bed. Thereby, she could use them to ask her staff for help, without the need to be physically placed on her wheelchair to access verbal communication. One burdensome activity for Treena is communicating with her medical staff, as she usually requires someone to accompany and clarify her message. She suggested that having an AAC device with AI language model embed could assist in communicating more independently with her doctors.

Treena believes that LLMs in AAC could ultimately assist when talking with unfamiliar listeners and, potentially, help her with dating and going out. These technologies could help to connect and keep in touch with her friends and family. Moreover, Treena mentioned that such systems should help to promptly come up with phrases and allow her to defend herself more quickly. She recounted how some people rudely opt to ignore her and ask something to her staff instead of directly talking to her; she wanted to immediately tell them: “You can talk to me about this.”

5.3.4 Benefit of Blissymbols

Previous knowledge of Blissymbolics was not required from the collaborators recruited. However, all of them have had experience with the Blissymbolics language. Therefore, individuals with CCN were invited to share: (i) their prior experience with the language; (ii) if they still used Bliss today; and (iii) if they could somehow benefit from Bliss nowadays.

Samantha did not have formal Bliss training, because by the time she was in school, Bliss was not being used by the school board anymore. Her mother, Julie, explained their experience with the PCS (Picture Communication Symbols), which was being used at Samantha's school, and Blissymbols:

> Julie Millar: "I just didn't like the PCS symbols that they were trying to present initially for her. I don't think a picture says a thousand words. It doesn't teach literacy. I think that you can use it to help, for sure, face-to-face communications. I had trouble with it. It doesn't have a syntax, it doesn't have a language. It doesn't teach a syntax, doesn't teach people how to put words together. So when she was younger, I went out and actively sought out Shirley McNaughton. I met with her and really, I don't know, it just made sense to me. Bliss made sense."

The Bliss language was part of Samantha's learning path and helped develop her literacy skill, but nowadays she doesn't use it very often. Similarly, Treena stated that Bliss helped her communicate with people. She recounted how Bliss played a crucial part in her life, as it empowered her to break free from an institution that had been weighing her down.

Treena also mentioned how Bliss helped her friend, Justin Clark, who is well-known for having fought for independence and transformed the disability rights in Canada. As many individuals with cerebral palsy, Justin lived most of his life in an institution without any control of his own life decisions. When he was 20 years old, Justin took his parents to court to prove that he was a mentally competent adult, who had the right to make decisions about himself. For the first time in a Canadian court, someone testified using a Blissboard. Justin communicated by pointing to symbols, which were then read aloud. In 1982, he won the right to leave the institution – the same that Treena mentioned being freed from – and make his own decisions about his future (Canadian Broadcasting Corporation, 2018).

After recalling her dear friend's story, Treena added that she still has her Blissboard, but rarely uses it. She explained that learning Bliss helped with language and communication, but she eventually switched to communicating in English. For her, Blissymbols now works as a "backup", which means she only uses her Blissboard to communicate whenever her high-tech communication device is not available.

When asked about her thoughts on Blissymbolics in future AAC applications, as a Bliss expert, Shirley replied that Blissymbols usage could be foundational. She explained that parents often come to her saying: "I want my child to be like their speaking peers, to read and communicate like them." Accordingly, future uses of Bliss should accommodate the children and parents' hopes. Users should be given the opportunity to develop their literacy skills and further decide if they want to keep using the symbols or switch to other languages. Shirley emphasized that Bliss should be provided as an option, but never imposed.

Shirley reminiscent how individuals with CCN are now relying on current AAC software which receded to using pictures, drawings and icons; however, images are not a language. She clarified how import it is that people understand Blissymbolics, its importance and strategies. Regarding the last one, she explains that there are different Bliss strategies and that future works should accommodate each users' preferences.

Different uses of Bliss can be narrowed down to: (i) word for word association of Blissymbol, and (ii) telegraphic short form. The first one might help users to better communicate in certain situations, such as telling a short story. On the other hand, the second option is useful for quickly communicating a need. Shirley explains that telegraphic speech might help in certain contexts to promptly convey their message, but it does not meet the most children and parents' hopes.

Future AAC software using Bliss must accommodate the strategy according to each individual's needs and preferences. Shirley recalled some experiences with AAC users who preferred to use telegraphic speech, while others opted for using all Bliss potential. Firstly, she told the story of a Bliss user who used telegraphic form to communicate with caregivers. The person explained her that, as the caregivers were only prepared to provide people with two minutes of attention, thus nothing more than that would be offered back to them.

Shirley also shared two stories of Bliss users who refused to use telegraphic communication. One translated the Bible using Bliss full capabilities, while the other user wanted their Bliss communication to be as complete and perfect as English, by using, for example, past tense. Moreover, Shirley mentioned that Bliss used in combination with AI could help to bridge languages, as intended by Charles Bliss, as the language focus on meaning and not on phonetics.

6 HCI GUIDELINES FOR HIGH-TECH AAC WITH AI LANGUAGE MODEL INTEGRATION

To build this work's main contribution and answer the third and main research question (see subsection 1.2), we sought to define what is needed to build a robust AAC software with AI language model integration for highly diverse users. For that, our previous studies were consolidated by mapping the findings from the literature and co-design sessions. Thereafter, the preliminary results were combined to develop HCI guidelines that will lead designers, developers, and researchers into creating meaningful AI and AAC interventions. This section will detail the guidelines' development process and discuss its outcomes.

6.1 PROCESS

The guidelines' development stems from data collected from this work's secondary and primary research. As detailed in Table 8, our three-step process underwent two rounds in order to generate the final result.

Table 8 – Guidelines' development process.

Step	First round	Second round
Data Collection	Extract findings from the secondary and primary studies individually.	Merge data from both preliminary results' blueprints (i.e., first round results).
Cleaning	Remove duplicates, organize, and synthesize data, individually for each study.	Remove duplicates, organize, and synthesize all data.
Coding	Code and categorize data, individually for each study.	Code and categorize all data.
Results	Preliminary Results: Blueprint 1, Directions for Future works and Challenges (extracted from the literature); Blueprint 2, Users' Expectations and Concerns (compiled from co-design).	Final result: HCI guidelines for high-tech AAC with AI language model integration.

Source: Author.

Both rounds followed the same process, which included (i) a data collection step for individually compiling or merging the data to be processed in each round, (ii) a data cleaning step for removing duplicates, organizing and synthesizing, and (iii) a coding step for labelling and categorizing each information.

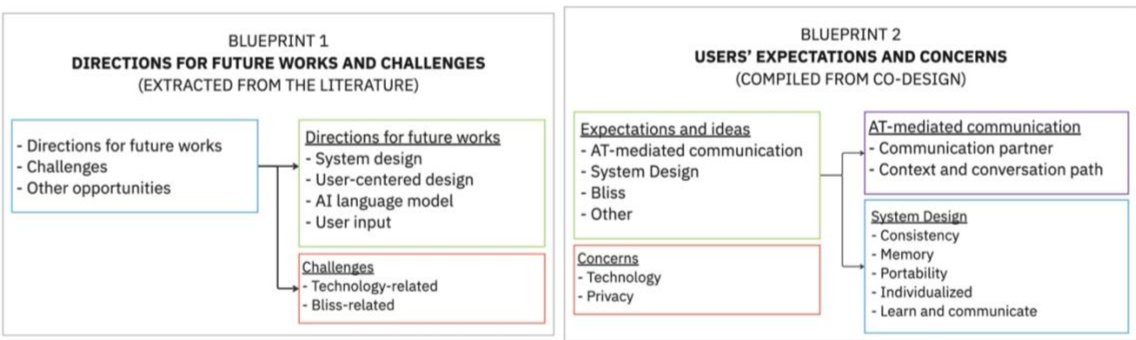
During the first round, data from the two systematic mappings and co-design findings were individually handled to generate the preliminary results, or what we called blueprints. Those blueprints individually reflect (i) the opportunities and challenges extracted from the literature and (ii) the users' expectations and concerns reported in the co-design activities. Both blueprints were then combined to develop the final HCI guidelines, which underwent the same data cleaning and coding process during the second round.

6.2 RESULTS

As shown in Figure 15, both blueprints had distinct category structures according to their outcomes. The systematic mappings focused in discussing the suggestions and difficulties identified in the literature. Hence, their categorization included directions for future works, challenges and other opportunities. Apart from the last group, each one had their specific topics. For instance, directions for future works concerned: (i) system design, (ii) user-centered design, (iii) AI language model, and (iv) user input.

Accordingly, the primary research was centered around the users' perspectives regarding the integration of AI technologies into AAC devices. Therefore, mapping the co-design findings resulted in two main categorizations: expectations and ideas, and concerns. The first group included topics such as: (i) AT-mediated communication, which was then subdivided into "communication partner" and "context and conversation path"; (ii) system design, which was subdivided into consistency, memory, portability, individualized, and learn and communicate; (iii) bliss; (iv) other. The last category was sorted according to the users' concerns about technology and privacy.

Figure 15 – Blueprints compiled from the secondary and primary research.



Source: Author.

Both blueprints were then combined to generate the final guidelines. Each blueprint result will not be individually detailed as they contain the same information as the guidelines and other topics discussed later in the next subsections. The full blueprint extracted from the literature is provided in Appendix B, and the blueprint resulting from the co-design findings is available in Appendix C.

As shown in Table 9, the main guidelines were created after combining the directions for future works identified in the literature, with the users' expectations and ideas. Consequently, merging the challenges identified in the literature with the users' concerns allowed us to determine some risks associated with integrating LLMs technology to AAC devices. Lastly, the "other opportunities" group was combined with some ideas brainstormed during the co-design sessions, this resulted in identifying new research opportunities for future works.

Table 9 – Categories combined and their results.

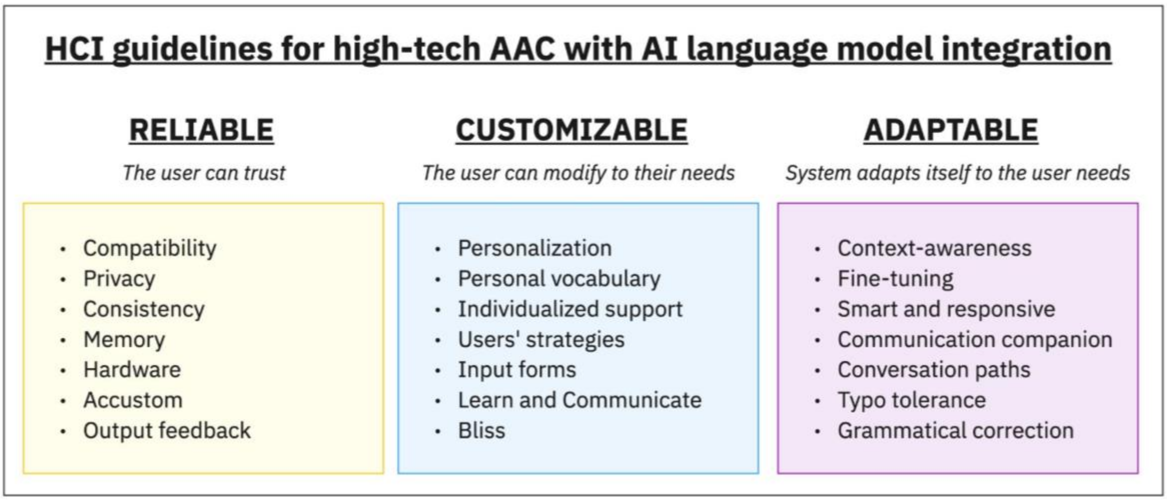
Blueprint 1 Categories	**Blueprint 2 Categories**	**Categories combination result**
Directions for future works	Expectations and ideas	HCI guidelines for high-tech AAC with AI language model integration
Challenges	Concerns	Challenges and risks
Other opportunities	Expectations and ideas	Other research opportunities

Source: Author.

Our HCI guidelines introduce three principles for creating robust AI interventions for high-tech AAC. In other words, all AAC software that targets an LLMs integration must be: reliable, customizable and adaptable. By *reliable*, we emphasize on offering a system that the user can always trust and rely on. On the other hand, *customizable* is related to having a system that allows the user to easily modify it to their needs, preferences and unique capabilities. Lastly, *adaptable* is associated to a system that adapts itself to the user's needs, according to their preferences, environment and situation.

As illustrated in Figure 16, all three principles are further divided into seven categories each. As discussed earlier, these groups were not defined arbitrarily, but rather based on the AAC literature and in the perception of AAC stakeholders involved in our research. It should be noted that some information extracted from the previous blueprints might fall into more than one principle. However, for better organization and visualization, we opted to attribute only one category to each information. The subsequent subsections will delve into the principles according to each one of their subcategories.

Figure 16 – Guidelines for high-tech AAC with AI language model integration.



Source: Author.

It must be noted that our guidelines are meant to be flexible and should not be followed rigorously but adjusted according to each context. For instance, if aiming at building a system exclusively for assisting with communication, the readers might benefit more from the *Conversation paths* or *Communication companion* topics rather than the *Learn and communicate*. Alternatively, if, due to project requirements, it is not feasible to follow certain essential directions (e.g., provide maximum support for diverse input or selection methods), the developers could benefit from the guidelines by increasing their awareness of the implications and seeking to mitigate the risks. We invite AAC developers, designers and researchers to apply our guidelines and accommodate them to their project's reality while weighing the pros and cons of their adaptations.

6.2.1 Reliable

Compatibility. This concerns the ability of a system to provide seamless access by multiple means. In the context of high-tech AAC, this topic is extremely fundamental as users rely on their preferred input form, which is directly related to their physical limitations. Systems that do not offer robust and extensive compatibility will exclude users that depend on certain input forms. For instance, during our co-design sessions, one of the co-designers mentioned how they wanted to but could not use the iPad as it was not suited to their needs. They further explained that the device did not offer the appropriate support for indirect selection users.

Providing ample hardware compatibility allows the user to rely on a system that will constantly allow them to interact according to what works best for them at the moment. Besides that, software and interface compatibility are crucial to ensure that the system is compatible with its different versions and components. Furthermore, this can help improve the system's architecture flexibility and scalability.

Privacy. Considering the current ubiquity of technology and malicious uses of personal data, addressing this topic is indispensable for any HCI discussion. Privacy becomes even more complex in the context of AAC where most individuals with CCN are highly dependent on their caregivers and family members for most of their daily activities.

AAC developers, researchers and designers must deeply evaluate the implications of their particular AI interventions and guarantee the data usage process transparency. For example, it is important to question: (i) What user data will be collected?; (ii) How will their data be stored and retrieved?; (iii) How will users be able to access and control their personal data usage?; (iv) Will this system retrieve data from the users' communication partners? If yes, how will their privacy be maintained? To ensure that future AI works will make responsible and ethical use of data, it is crucial to build on top of established regulations, such as the European AI regulatory framework (European Union, 2023).

Regardless of the concern for user privacy, it was observed during our primary research that this would not prevent users' from adopting novel AI interventions in high-tech AAC. Co-desginers were clear about how they would thoroughly evaluate the risks and benefits of using these technologies. They also recalled how they use applications daily despite being known for collecting and selling users' data.

Besides that, privacy is not only related to personal data, but it can also be associated with the users' autonomy. This is a critical issue for AAC users who, until the 80's in Canada, were yet considered as mentally incompetent and unable to control their own life decisions – as witnessed in Justin Clark's story (see subsection 5.3.4).

A reliable AAC system should promote the user's control over their own communication. We recall the work of Ferm, Amberntson and Thunberg (2001), in which a research participant had to confront his parents and educators as they did not allow him to express negative feelings by using swear words on his communication device. AAC developers and designers should offer options that allow users to express themselves while maintaining their privacy and autonomy.

Another privacy-related and uncomfortable situation commonly faced by AAC users is having other people approaching from behind and observing how they use the device. In this case, creative strategies could be applied to increase the users' privacy, such as allowing them to quickly express their feelings or using input forms that are hard to decipher, similar to the work of Cai *et al.* (2022).

Privacy is a broad term that encompasses different aspects from ethical data usage to ensuring the users' autonomy. As LLMs for high-tech AAC is yet a novel topic, we hope to see more research and further discussions around this topic in future works.

Consistency. During the co-design sessions, Shirley McNaughton emphasized that an AI system should be able to exceed what humans can do. With that in mind, she expects that consistency could be one of the benefits of applying such technologies to AAC. As detailed in subsection 5.3.3, she further explains that humans are not consistent, which means that AAC users are constantly having to adapt to new situations.

For example, a staff can unexpectedly leave their job after several years providing care to the same individuals with CCN. After that, they will now have to readapt to new staff reassigned to them. Contrariwise, the technology can, and should be, consistent so that users can always depend on it regardless of external conditions. Because of that, we opted to include consistency as one of the main pillars to develop a reliable system.

Creating a consistent system not only means that it should always be available to users, but also that it must be consistent between software updates and interface components – similarly to the compatibility topic. Besides avoiding any drastic changes, developers should address consistent error handling and response time.

Memory. Equally to the consistency topic, "total memory ability" was also pointed out by Shirley as being a benefit of AI technologies which transcends human capabilities. AAC developers, designers and researchers should explore the multiple possibilities related to storing and retrieving data, especially when associated with LLMs usage.

As mentioned in subsection 5.3.3, this could largely assist communication by relating and adapting previously stored information to the users' current environmental conditions and conversation partner. For instance, the system could assist users in quickly generating simple messages, such as replying to questions about their preferred hobbies or favourite films; that information could have been previously answered and could now be easily retrieved and adapted if needed.

In the context of generative AI, it can be hard to ensure that outputs generated will fulfil the consistency requirements. Therefore, we propose taking advantage of the system's memory capabilities. Even though developers might be unable to fully control the output that will be generated, the system could identify whether this same request has been made before and offer the previously selected option.

Hardware. As contextualized in subsection 2.1, there are two types of AAC devices: dedicated and "off-the-shelf" options. AAC developers should carefully consider which options they wish to work on, as this will directly reflect on the systems' usability and reliability. Dedicated devices are costly, but more reliable as they offer a robust and targeted hardware solution. On the other hand, "off-the-shelf" products (e.g., an iPad) can be a cheaper alternative, but often more limited.

During our co-design sessions, both AAC users mentioned that they would prefer portable devices, but none of the available options suited their needs. Currently, mobile solutions only work for a targeted group of users. With this in mind, we want to bring attention to the urgency to opt and work towards creating more portable and reliable alternatives for diverse users.

Accustom. This topic regards the importance of easing the adaptation period for novel users that are changing AAC devices. During our co-design sessions, it was mentioned that accustoming from an older device to a new one was an arduous task that demands the user to relearn and adapt to their new environment. Besides that, one of the recruited AAC users wanted to engage in the activities but was not able to participate as they were going through the long and burdensome process of changing to a new device, which can take a couple months. We bring attention to the importance of offering options that will support the novel user and mitigate their adaptation process.

Output feedback. For developing a reliable AI-enhanced communication system, the user must be able to trust that the system will not change or have control over their message. Ideally, AI and AAC software could provide real-time feedback over the output results. In this way, the user will have increased control over the final outcome.

Considering LLMs prompting, providing this option could save keystrokes as it might reduce the need for correcting requests. However, considering the current computation power available, this could reflect in increased latency and hardware requirements.

Moreover, AAC developers could experiment with incorporating multimodal feedback and explore diverse options of visual, auditory, and haptic feedback to provide a better experience for users with diverse sensorial and cognitive needs.

6.2.2 Customizable

Personalization. Individuals with CCN constitute a highly diverse group as each AAC user has their unique abilities and needs. Consequently, personalization is the foundation of AAC and must be the main priority in any inclusive product development. Personalization is a broad term that encompasses from interface design to device hardware. Ensuring that your project is greatly personalizable and easy to operate is guaranteeing that your users' will be able to customize it to meet their particular needs.

We highlight the crucial role of avoiding a "one size fits all" approach which progressively excludes more users. Especially in the AAC field, researchers, designers and developers must focus on understanding the unique and complex cases instead of designing only for the masses. Following well-established guidelines can help designers create more accessible interfaces, such as the Web Content Accessibility Guidelines (Web Accessibility Initiative, 2023).

On top of that, we emphasize the importance of understanding the uniqueness of AAC users and engaging them during the development process of any tool that will directly affect their everyday activities and relations. Having individuals with CCN directly guide the design process might lead to more effective and inclusive AAC solutions.

In AI's context, personalization can ensure that the user will have more control over the final message. Users could personalize the generated output to fit their needs or preferred communication style by, for example, personalizing different options as tone, voice, emotion, or message type. Another option would be to enhance the generated messages based on the user's prestored information and communication behaviour. On a similar note, – as mentioned in subsection 4.1 – the ChatGPT tool allows users to enhance the responses by providing information about themselves or specifying how they would like the bot to reply.

Personal vocabulary. As previously discussed in the privacy topic, developers have a role to play in ensuring the users' privacy and autonomy. For that, this topic was included to discuss and emphasize the importance of exploring the possibilities of a personal vocabulary component. If easy to operate, this can allow the user to independently control their own unique ways of communicating.

This option can be especially useful for ensuring the privacy of preteen and adolescent AAC users, who might face difficulties when depending on their caregivers or family members to decide how they will communicate – as noted in the work of Ferm, Amberntson and Thunberg (2001). Instead of limiting the communication options, developers must focus on allowing users to autonomously expressing their unique personalities and preferred communication style.

Besides that, children and adult users can also benefit from independently customizing their communication options to meet their needs. For instance, during our co-design sessions, it was mentioned how Samantha enjoyed making jokes when she was a child but needs support to program the jokes on her device. After she got older, her family could not keep up with this activity and she could not do all the programming by herself. Because of that, Julie lamented that currently Samantha's personality is missing from her computer. Considering this, an easily accessible and customizable personal vocabulary option could assist adult users in independently expressing themselves.

In the context of AI intervention, this becomes more complex as data used for training can directly affect the users' expressivity by, for instance, restricting them from talking about specific topics. This situation was observed in the work of Valencia *et al.* (2023), where one of the participants wanted to talk about "smoking" but was impeded from doing it, possibly due to a security layer in the model or it's training data.

Individualized support. Considering that individuals with CCN have unique abilities, any tool for AAC users must offer an individualized support approach. This means that the assistance offered should be compatible with the user's necessities. Rather than treating all users equally, its essential to respond to the needs of individual users. Are they a novice or expert user? What are their literacy skills? What other personal, cultural, or linguistic particularities are relevant to take into account? Such information could be collected to provide a personalized experience and enable the system to advance with the user.

User's strategy. Each AAC user develops their own strategies which should be considered as opposed to disregarding what already works for users. For that, AAC researchers, designers and developers should carefully understand what strategies are used and how they could accommodate them to a novel AAC product. This topic is deeply related to the "Accustom" category, as accommodating existing strategies might ease the burdensome adaptation time to a novel device. Moreover, engaging AAC users during the planning and development process could help identify those strategies and provide valuable insights.

Input forms. While some AAC users may communicate in a telegraphic form, others prefer to use complex and long sentences. A customizable AAC system should allow users to accommodate their preferred communication style by offering different input forms that meet their needs. Apart from word-to-word input, other existing forms in high-tech AAC include abbreviations and shorthand, symbols, numeric codes, selecting pictographs, and more.

Researchers, developers, and designers could experiment with novel input forms and possibilities, especially considering the opportunities with LLMs. However, such systems must not limit the users' alternatives and force them to experiment only with the developer's desired input form. Instead, users should be able to freely choose and personalize whichever input form works best for them.

Learn and communicate. During our co-design sessions, Shirley mentioned that it would be interesting to explore the opportunity of including multiple modes to support the users' needs in the context of an AAC system with AI intervention. She explained that the system could have two modes, one for communication and another for learning. Her idea aligns with our views about avoiding the creation of context-exclusive AAC systems for use in particular circumstances, which could, for instance, create an unreasonable requirement to switch machines or learn how to operate multiple software instead of relying on only one device throughout their day.

Such a learn-and-communicate system could allow children with CCN to gradually improve their communication and literacy skills according to their particular needs. Existing high-tech AAC systems might also unintentionally assist with learning. For example, as detailed in subsection 2.1, the Proloquo2Go app was used by a child with autism not only to communicate, as intended, but also to train and improve his speech. On top of that, considering the use of LLMs, there are great opportunities to provide an AI-enhanced learning mode for AAC users, and we hope that future research will explore this exciting possibility.

Bliss. As a language that focuses on meaning rather than phonetics, Blissymbolics is a valuable asset to individuals with CCN. Thus, Bliss could be present in novel AAC systems to democratize learning by providing formal language education to children with congenital speech limitations. As described in subsections 2.2.1 and 5.3.4, the Blissymbolics language has helped many individuals with CCN develop their literacy skills and independently communicate their needs.

Rather than merely providing pictographic-based communication, high-tech AAC could offer an introductory mode to novel Bliss users and guide them into a gradual and personalized learning path. With that, children with CCN could learn syntax and semantics and unlock their full potential instead of relying on needs-based communication.

Nevertheless, Bliss must be offered as a choice and never imposed. In other words, after users acquire their language proficiency, they can choose whether to keep using Bliss in their daily lives or if they prefer to switch to communicating in another language. As mentioned during our co-design activities, AAC users have learned to express themselves using Bliss, but they barely use it today after switching to English.

The Bliss language can support either word-per-word associations or telegraphic and shortened communication forms. An AAC system with Bliss support should allow users to personalize their preferred communication form and accommodate the language according to their needs. Besides assisting with training and language acquisition, Bliss could also help to bridge language, as intended by its creator, since it is exclusively written and uses conceptual symbols.

However, Blissymbolics is, unfortunately, a minority and low-resource language. Considering that and the inexistence of a Unicode standard, training an AI to communicate in Bliss is currently a challenging task. We hope to see future efforts to digitalize Bliss resources into a computer-readable format and generate enough data to train a language model in Bliss.

6.2.3 Adaptable

Context-awareness. Context-awareness is related to the ability of a system to operate and adapt according to its current environmental context without explicit user intervention. A context-aware software can collect data through different sources such as sensors, the internet, device status, or pre-stored information. Thus, these systems can increase usability and effectiveness by quickly reacting and dynamically modifying their behaviour according to the constant changes in the users' environment (Baldauf; Dustdar; Rosenberg, 2007).

As Curtis, Neate and Gonzalez (2022) revealed, using contextual information as input for AAC is yet an unexplored modality. Nonetheless, as identified in subsection 4.1, novel works in the field of AAC that experimented with LLMs started to apply or suggest using contextual information to enhance the quality of AI's generated output. AAC developers should strive to create more adaptable systems that reduce cognitive and physical efforts and the constant need for visual vigilance. A robust AAC system should independently modify itself according to the users' current context without requiring them to adjust it manually.

Contextual information could be retrieved from (i) the environment, such as location and time, or (ii) the current conversation, which could include capturing the dialogue and identifying the conversation partner. Adding those data with pre-stored and personalized user information could provide a powerful source of information for AI-based AAC systems.

During our co-design sessions, AAC users described several scenarios where context-awareness could be highly beneficial for them. For instance, users mentioned how they wanted a system that was able to put their ideas in the right context. Samantha explained that oftentimes, she prepares in advance for a conversation, but it might follow a different path than she predicted, and her pre-programmed texts now seem a bit off the subject.

Additionally, AAC users usually have messages prepared that they frequently access, such as an introduction text or telling a story. However, some information, including the users' age, changes annually and could be easily updated by the system, but currently requires manual intervention. An AAC system with AI language model integration could automatically adapt the users' messages to their new conversation or environmental context.

Fine-tuning. Developing a deep learning model from scratch is a costly and time-consuming process that requires great computational power and huge datasets for performing unsupervised training. When performing tasks to a specific domain with limited access to data, a feasible option is to use existing pre-trained models and fine-tune it, that is, perform supervised training using a smaller labeled dataset for the desired task.

Another option is to use few-shot prompting, where an LLM is provided with a description of the specific task and a few demonstrations to induce the desired output (Brown *et al.*, 2020). Correspondingly, one-shot prompting is the same scenario but uses only one example, while in zero-shot prompting, just a description of the task is provided without offering any examples. In those cases, the new knowledge is not encoded in the pre-trained model, which occurs when fine-tuning it.

As identified in subsection 4.1, novel works are exploring the use of fine-tuning and prompting to leverage LLMs integration to AAC software. In this context, Cai *et al.* (2022) identified that fine-tuning pre-trained models overperformed zero and few-shot prompting for abbreviation expansion tasks, increasing keystroke savings and potentially speeding up conversations in the AAC domain. Similarly, Yusufali, Goetze and Moore (2023) fine-tuned pre-trained models using eight datasets. They observed that the smaller domain-specific synthetic AAC corpora achieved higher communication rates than other generic corpora.

Hence, when targeting better output results, future studies could opt for fine-tuning pre-trained models instead of using few or zero-shot prompting. Additionally, works must choose an appropriate dataset for fine-tuning models that will be used in the AAC domain according to any specific tasks to be performed.

Developers must be aware that choosing inappropriate training datasets, either for pre-training or fine-tuning, might reflect in incorrect or even harmful outputs. As the pre-training process normally uses large unlabeled data for unsupervised training, the system will learn and make correlations independently, which reduces human intervention and can pose severe risks for the end users. As observed in the work of Valencia *et al.* (2023), a limiting LLM might restrict the users' conversation and impede them from expressing themselves properly.

There is an urgency for generating an open-source and robust AAC dataset (Curtis, Neate and Gonzalez, 2022; Yusufali, Goetze and Moore, 2023), especially considering underrepresented languages, such as Brazilian Portuguese (Pereira *et al.*, 2023). The work of Vertanen and Kristensson (2011) proposed a large dataset of surrogate AAC messages that were crowdsourced from individuals who imagined themselves as AAC users. While their corpus might be a good provisional alternative to the problem, it might not reflect the complexity of AAC and diversity of user's abilities and communication styles.

As Pereira *et al.* (2023), Vertanen and Kristensson's work data generated tends towards needs-based communication as both were based on information collected from non-AAC users. Moreover, datasets that do not contain long-term two-sided conversations, if used for training purposes, could generate models that struggle to relate previous information, deviate from the current context, misinterpret the user's intent or limit communication and self-expression.

Ideally, a robust AAC dataset should include real data from users with diverse communication styles, cognitive abilities and language skills. However, as pointed out by Vertanen and Kristensson, generating a large corpus of genuine AAC messages is a huge challenge considering the complexity and ethical implications. A long-term solution for AAC software with LLM integration could be to gradually fine-tune their models based on the data generated from the users’ preferences, system usage and communication behaviour over time.

In the future, such data could be used to generate a dataset of AAC dialogues of real users instead of relying on synthetically generated data or corpus generated from individuals who pretended to be AAC users. Such data extracted from real users must definitely be ethically sourced and properly handled, ensuring full anonymization and that no personal or sensitive information is included before dissemination.

Smart and responsive. Smart Systems have the ability to learn, dynamically adapt and make decisions based on the data received, transmitted, or processed, thereby allowing them to improve their capabilities. Such smart systems are designed with an emphasis on the human, with the knowledge, capabilities, and values of the system shaped by the people involved (Romero *et al.*, 2020).

A highly adaptable and responsive smart system should consider and reply to the user's individual development. This is especially critical when considering a learning mode where users can develop their literacy skills; such a system could accompany each learner's progress and provide proper support according to their abilities. Additionally, they should distinguish novice from expert users and responsively adapt to their expertise.

As discussed in the "Context-awareness" topic, even though AAC systems could dynamically adapt to each user's needs and behaviours to reduce user intervention, those changes must follow user consent and not force them into accepting new adaptations without providing control over it.

Communication companion. During our co-design sessions, collaborators expressed that they expected future AI integrations in AAC to work as a communication companion and help them communicate more effectively. AAC users often require assistance to help them clarify their message; likewise, AI could help them independently put their ideas in the right context.

For example, Treena mentioned how she often needs help communicating during medical appointments. She expressed that AI-based AAC could help her communicate independently with the medical staff. Similarly, in the work of Valencia *et al.* (2023), participants were excited about the opportunity to share their medical information and have AI assist during healthcare consultations.

Furthermore, participants shared uncomfortable situations when communicating with unfamiliar listeners, that is, individuals who are not familiar with AAC users. They mentioned how they wanted AI to help them communicate better and meet new people, including helping them find romantic partners.

AAC users also voiced the importance of AI assisting them in coming up with phrases more quickly. One co-designer mentioned how she wanted to defend herself more quickly in situations where people were rude to her, either by ignoring her or by making fun of her way of communicating. Additionally, AAC users expected AI to help not only with face-to-face communication, but also make it easier to keep in touch with friends and family.

Conversation paths. As the mother of an AAC user, one of our co-design participants expressed concerns about individuals relying too heavily on technology to direct their thought processes or influence their decisions. She proposed that AI integrations in AAC could anticipate the user's sentence without necessarily directing the conversation but rather offering the paths to it. Additionally, this would increase the user's control over the final output and conversation flow.

Typo tolerance. Often, AAC users have some form of physical limitation, which might increase the occurrence of typos. Typing error occurrence can be exceptionally high in some input methods, such as eye-gaze (Hansen *et al.*, 2004). Thus, novel AAC systems must include forms to mitigate this. In the context of AI-based applications, it is essential to include noise data during the model's training phase, similar to the work of Cai *et al.* (2022).

Grammatical correction. Since each verbally impaired person develops their own syntax model, their message might lack formal grammatical rules and require further clarification. Considering this, AAC developers must also consider the importance of including grammatical corrections in their applications, which could potentially enhance communication and independence by reducing the need for clarification, especially for users with limited literacy skills. However, such correction must not come at the expense of altering the message's meaning.

6.2.4 Challenges and risks

After analyzing both systematic mappings results and co-design findings, two different types of challenges were identified: technology-related and Bliss-related challenges. The technology-related challenges and risks identified were associated with: relying on the semantically equivalent rather than obtaining the intended message; AI interpretations and the challenges of highly abbreviated telegraphic sentences; incorrect suggestions and the cost of improved accuracy; the quick advancing of technology; and concerns of AI controlling the conversation. Moreover, we compare and detail additional risks from the work of Weidinger *et al.* (2021) and other cybersecurity threats.

The works of Shen *et al.* (2022) and Valencia *et al.* (2023) proposed a prompt-based approach to expand users' few input words into full sentences. In both cases, the user relies on the language model to output their desired message accurately, also known as the "golden reply." However, considering this context, the application might not generate the perfect output planned by the user, and thus, the user will have to opt for a reply that is closer to their intended message. It is essential that developers and researchers thoroughly evaluate whether there are any risks for the users in opting for the semantically equivalent reply instead of achieving their desired golden reply. For example, in such cases, users' self-expression and communication style might be constrained by the technology's capability.

Similarly, another risk identified is depending on language models' limited interpretation ability, which may underperform in the AAC domain, where users have unique ways of communicating and expressing their thoughts. Some AAC users communicate in a highly abbreviated telegraphic form, which poses a significant challenge to the technology, mainly when it performs correlations based on non-AAC users' data. Such a problem could be partially mitigated by fine-tuning models with a robust labelled AAC dataset of non-fictional users, yet unavailable.

Issues related to the poor performance of language models for AAC were voiced out by our co-design participants, who were concerned about users' wasting their time with incorrect suggestions. Slow communication rate is an ongoing issue for AAC users that could potentially be aggravated by the low efficiency of LLMs in this specific domain. At the same time, improving the model's accuracy might come at the expense of increasing latency and hardware requirements (Cai *et al.*, 2022), especially when considering improving the users' control over the final output by, for instance, implementing real-time output feedback.

Besides that, participants were concerned about the rampant advancement in LLMs, which could pose more risks than benefits for individuals with CCN if created heedlessly. Collaborators worried about users' relying too much on technology to direct their thought process instead of thinking for themselves and hoped that AI could assist in predicting the users' intentions rather than controlling the conversation.

Weidinger *et al.* (2021) delineates six types of risks of harm from language models. Firstly, they explain the dangers related to Discrimination, Exclusion and Toxicity. A language model can provide unfair discrimination and representation by perpetuating stereotypes and social biases. Furthermore, it has the potential to disseminate toxic language and incite hate or offend individuals.

Similarly to what was discussed before, Weidinger's work explains that language models can perform more poorly for some social groups than others and harm disadvantaged groups, such as individuals with CCN. For this reason, our co-design collaborators expressed their concern about future technologies developing AI that makes interpretations from the mass population rather than from the particular context of AAC users.

The second risk presented by Weidinger's work is Information Hazards and discusses the risks of language models leaking private data or correctly inferring sensitive information. Concern with user's privacy was a very prevalent topic during our primary research. In the context of AAC, this could potentially be harmful to the users', who constantly depend on their communication devices for their daily activities. For instance, LLMs could potentially leak the user's private information or conversations dialogs. This concern was also voiced out by the participants of Valencia's study, where one user revealed to worry if their proposed system would suggest phrases that could reveal how he truly felt about someone, or embarrass him by revealing that he had talked badly about somebody (Valencia *et al.* 2023).

The third risk discussed in Weidinger's work is Misinformation Harms, which explains that LLMs can provide false or misleading information and consequently create less well-informed users. In the context of AAC, such technologies could harm users by inappropriately making false assumptions or altering their message intention.

The authors further details that such misinformation could erode trust in shared information or lead users to perform unethical or illegal actions. These risks arise due to language models' inability to distinguish between factually correct and incorrect information, which is rooted in the limitations of the statistical methods used for learning how to represent language. The fourth risk is Malicious Uses, which could be used by product developers who try to use language models to cause harm by, for instance, using it to create personalized scams.

The fifth risk is a specific use case for conversational agents that directly interact with human users; the authors call it Human-Computer Interaction Harms. When developed under poor product design decisions or training objectives, such conversational agents could lure users into resenting it as human-like, consequently making them overestimate its capabilities and use it in unsafe ways. Such risks and harms must be intensely studied and considered by developers who wish to create conversational agents for AAC, especially considering the education domain, where users will be actively interacting with such technologies, which, if poorly crafted, could affect their development.

The sixth and last risk delineated by Weidinger *et al.* (2021) is Automation, Access, and Environmental Harms, which extends to different AI domains and is concerned with the high environmental costs of language models and the unequal access and benefits of these models to different groups.

Another essential topic for discussion in the LLMs risks domain is its vulnerability to attacks. Similarly to computer programs, LLMs are designed to follow human instructions, and thus are susceptible to virtual attacks. Next, we focus on detailing a specific type called Prompt Injection (PI), that can either be direct or indirect. The first one occurs when malicious users directly prompt the model to ignore its original instructions, which can be used to bypass its content restrictions, or obtain access to its original instructions, as demonstrated by Perez and Ribeiro (2022).

On the other hand, indirect PI attacks can be mediated by LLMs and remotely executed by external attackers who wish to take advantage of well-intentioned users. As demonstrated by the work of Greshake *et al.* (2023), attackers can exploit LLMs-integrated applications and indirectly prompt it for malicious uses, either using: (i) passive methods, which occurs by retrieval (e.g., asking the model to fetch information from a web page that contains a hidden malicious prompt); (ii) active methods, by actively delivering the injection prompts to the model (e.g., sending emails with prompts that can be automatically processed by an email client with LLMs integration); (iii) user-driven injections, which can be achieved by luring the user into entering malicious prompts (e.g., using social engineering to make them copy and paste a malicious prompt without their knowledge); (iv) or hidden injections, where the attacker can disguise the malicious prompts and indirectly manipulate the model to fetch a larger payload from an external source (e.g., by hiding prompts in images or encoded in python programs that the model is instructed to run).

Users are already susceptible to virtual attacks and scams while browsing the internet and using social media. However, such attacks could be aggravated with LLMs usage, as individuals might overestimate and trust excessively on such technologies, especially when considering their ability to generate convincing content that can be indistinguishable from human-written text (Clark *et al.*, 2021). Attackers could, for instance, take advantage of LLMs capabilities to extract users' personal information, produce and disseminate convincing scams, gain access to victims' systems, spread malwares, manipulate users (e.g., by generating and propagating disinformation or polarized content), or induce Denial-of-Service attacks that unables the system from working properly (Greshake *et al.*, 2023).

Similarly to the mass population of users, AAC users are also vulnerable to such attacks. Nevertheless, considering a LLM-based AAC application, AAC users could be the ones who are most affected by such attacks as they would depend on such unpredictable systems for their everyday chores and social interactions. For instance, PI could allow attackers to circumvent original instructions and lead the system into generating inaccurate, offensive, unintended, or biased messages.

PI attacks also have the potential to cause privacy breaches, which can be highly dangerous for users who rely on systems that can store not only their personal data but also most of their conversations and environmental information. Additionally, AAC users could also be vulnerable to virtual manipulations since attackers could use the system to coerce or deceive them or alter their preprogrammed messages. Consequently, all these risks and insecurities discussed could minimize the users' trust in such systems and directly affect their expression and autonomy.

For the Bliss-related challenges, it was firstly identified issues related with the absence of a Bliss Unicode standard, which might increase the complexity of the already arduous task of translating Bliss into a computer-readable format. However, Bliss is a highly flexible language that allows for boundless combinations of symbols. Thus, fitting it into a set of requirements might come at the expense of limiting the language's capabilities.

Another challenge identified by the literature is the English-oriented nature of Bliss, which can pose challenges in translation tasks and make it arduous to support users who might need to switch to languages other than English. As mentioned in subsection 2.2.2, Bliss can be adapted to fit certain language structures. Considering this, developing a Bliss and AI application might require supporting different "types" of Bliss language. Otherwise, AAC users from non-English speaking countries could be excluded from adopting such technology.

We echo Greshake's perspective that the current race for AI implementation is not following adequate guardrails and safety evaluations. Moreover, we worry that imprudently creating and wide-spreading LLMs-based applications could be potentially harmful to groups that are already marginalized and underrepresented.

6.2.5 Other research opportunities

As thoroughly discussed in the high-tech AAC literature, there is an urgency for open-access AAC corpora. Novel AI applications for individuals with CCN must consider the uniqueness of AAC and avoid relying on a corpus collected from the larger population. Considering the huge issue of data inequality in NLP, multiple fields could benefit from research that focuses on generating robust open-access text datasets from real AAC users and, especially, recognizing languages other than English.

Another research opportunity is to explore different User Interface strategies for partial selection in incorrect predictions, as pointed out by Cai *et al.* (2022). This could assist users in achieving their desired output without having to prompt the language model again until the expected reply is offered or relying on full-sentence typing. Moreover, the user's unique utterances could also be explored by future research as a quick input method for such AI-based AAC systems.

During our co-design sessions, AAC users also pointed out research fields that they hope could be explored by future works, including assistance with mobility, device automation, and web accessibility. One participant mentioned how she desired that future technology could invest in AI-enhanced digital assistants that could assist individuals with mobility limitations to move independently. For instance, they wanted such digital assistants to help them go out without human assistance or help them open doors and windows.

Another participant mentioned that she often needs support to create automation on her device, such as for logging in to a website. On top of that, sometimes those automations do not work as expected, especially when facing connection issues. Thus, she hoped that AI applications could offer AAC users the proper support to not only create those automation but also to ensure that they will execute as required. Besides that, users recounted many of the issues faced when navigating the internet, which is often inaccessible to individuals with mobility limitations. Therefore, users wished for AI assistants that could help them navigate the internet independently and smoothly by, for instance, removing intrusive ads and pop-ups.

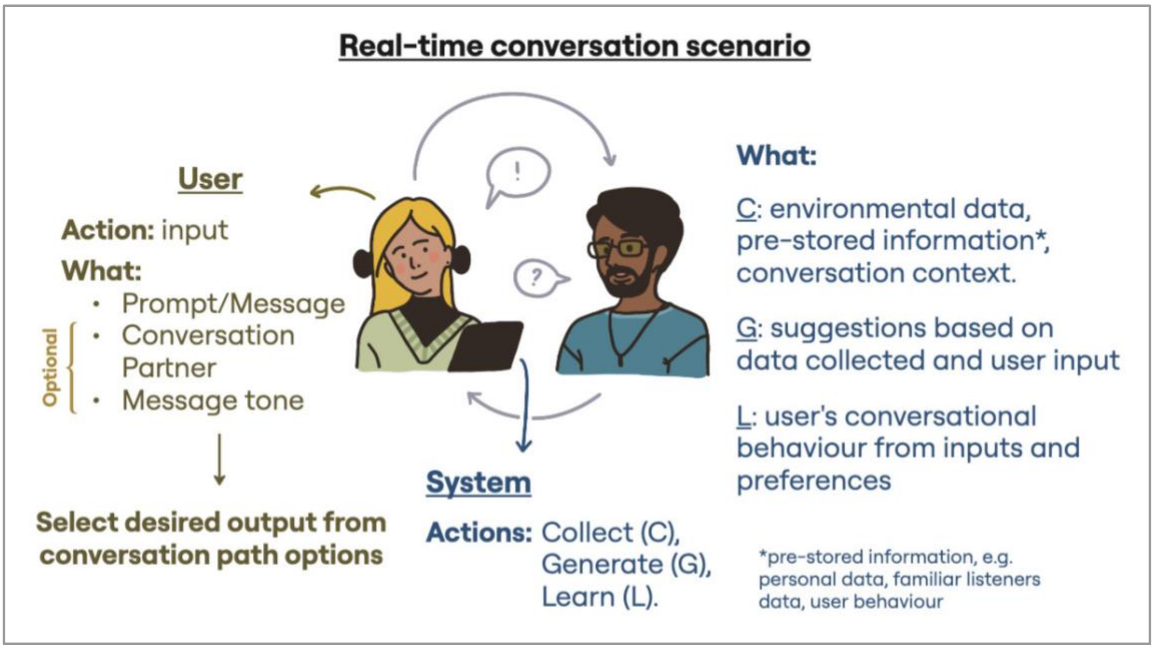
Multiple opportunities are yet to be explored in ethically, responsibly and safely using AI and LLMs to assist AAC users. Future works should consider both the advantages and disadvantages while aiming to create more inclusive and meaningful AAC projects alongside and guided by the ones that will benefit from it.

7 PROOF OF CONCEPT

Based on the previously detailed guidelines, we conceptualized an AAC software that includes a LLM integration. As illustrated in Figure 17, the first initial sketch of the system focused on a real-time conversation scenario. In this first concept, each agent (i.e., user or system) had their actions with its attributes. That is, the user's only action is *input*, and the attributes inputted could be: (i) a prompt (i.e., a short-length text that will be expanded by the LLM) or a full-length message; (ii) conversation partner, that could be identified to generate targeted output; (iii) define the message tone (e.g., surprised, enthusiastic, serious). Those attributes would be used by the system to generate different conversation paths to the current context, in which the user could select their preferred option.

On the other hand, in this initial sketch, the system included three actions: collect, generate and learn. The *collect* action gathers the user's current environmental data (e.g., time and location), any pre-stored data (e.g., user's personal data, familiar listener's information, user behaviour), and the conversation context. Next, the *generate* action creates suggestions based on the data collected and the user's input. Lastly, the *learn* action allows the system to learn the user's conversational behaviour from their inputs and preferences.

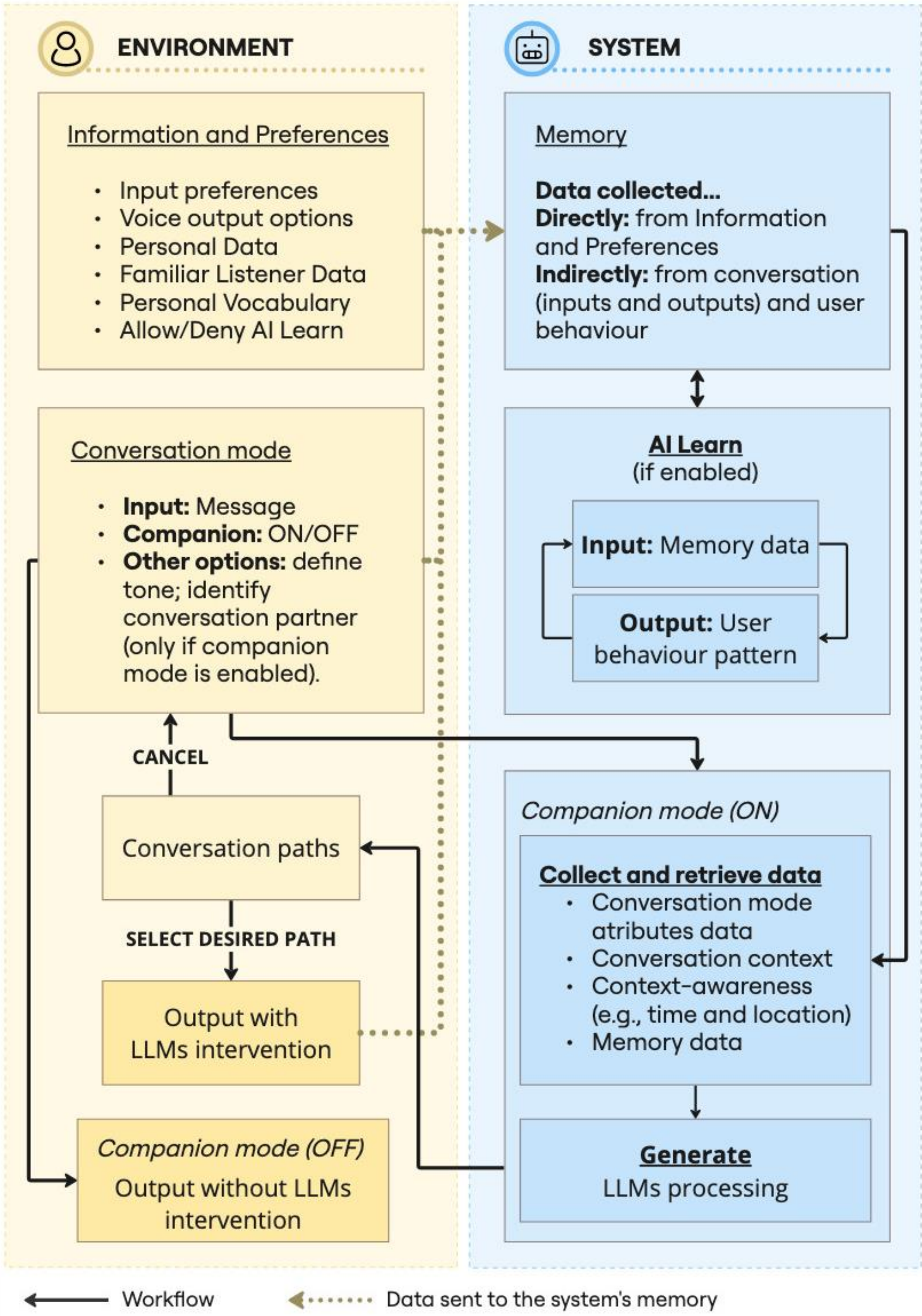
Figure 17 – Initial sketch of the conceptual system's functionalities during a real-time conversation scenario.



Source: Author.

Afterwards, the previous idea was polished and detailed by creating a high-level architecture diagram of the conceptual system, as shown in Figure 18. The following paragraphs will describe the hypothetical software and then compare its attributes with this work's guidelines. In addition, the subcategories that were not integrated into this system will also be discussed, demonstrating the guidelines' adaptability to different contexts and the importance of acknowledging the points that are not covered.

Figure 18 – High-level architecture diagram of the conceptual system.



Source: Author.

The conceptual system's high-level architecture (see Figure 19) was split into two settings: the *environment* that directly interacts with the user and the *system* that manages information and assists when requested. Following the initial sketch, the system's three actions (i.e., collect, generate and learn) were included and extended. In the environment, the user can define preferences and feed pieces of information to the system, including:

1) Desired input options, such as language, format (e.g., word-to-word, symbols, numeric codes), selection method (i.e., direct or indirect), and input method (e.g., eye-gazing, keyboard, switches).
2) Preferred audio output customization from a diverse voicebank.
3) User's personal data that they wish to provide to the system (e.g., profile, hobbies, medical records, work-related information).
4) Information about familiar listeners (e.g., family, friends, co-workers).
5) Personal vocabulary that can be used to easily create preprogrammed notes, including jokes, stories, and other messages that reflect their personhood.
6) Choose whether the system is allowed or not to dynamically learn from the user's data and conversation behaviour, including the information provided directly (e.g., personal or listeners data, preferences) or indirectly from conversations.

During a real-time conversation scenario, the environment's conversation mode allows the user not only to input their message but also to define its tone, identify the conversation partner and select if they want to use the "Companion mode." If enabled, this mode allows the message and its attributes to be processed by a LLM that will generate possible conversation paths suggestions according to the current conversation context, context-aware data (e.g., time and location), and data retrieved from the system's memory. Furthermore, this mode must be offered as an option, the user can still input full sentences to be outputted without any LLM intervention.

Considering these diverse possibilities, the system might be able to generate more accurately targeted results that meet the user's true intention. This could potentially assist users in effectively communicating while reducing their cognitive and physical loads, especially when enough data is available for the system to learn their behaviour pattern from information stored in the system's memory.

In this conceptual system, we tried to apply the previous guideline's principles. The topic of *Compatibility* was partially addressed by offering multiple input alternatives, such as diverse input methods and forms, selection methods and language options. Similarly, the

Privacy topic was only covered by allowing the users to voluntarily decide if they want the AI to train on their personal data and to process their messages. Privacy was lightly addressed in this system architecture, but as a crucial topic, it deserves more in-depth discussions, including how the user data will be stored and managed.

The topics of *Consistency* and *Memory* were explored by designing an environment that allows users to provide ample information about themselves and their context. This enables the system to store and retrieve data to generate reliable suggestions based on the user's personality, environment and previous scenarios. For instance, some questions might be repeated constantly during interactions, such as "What do you do on weekends?" This system could automatically retrieve the appropriate reply if the user previously answered it.

On the other hand, the topic of *Hardware* was not included as, for this conceptual system, we focused solely on software. Still under the *Reliable* principle, the topics of *Accustom* and *Output feedback* were not included in the system architecture but were brainstormed for discussion. An idea for tackling both *Accustom* and *Individualized support* is to provide an introductory mode to the system. This could allow users to gradually get used to the system and reduce their learning curve while also transparently collecting information on their unique skills and abilities. On top of that, ideally, the system should provide real-time *output feedback*. This could reduce the need for correction when LLMs fail to suggest the desired output. However, offering this option could increase hardware requirements and latency.

Considering the *Customizable* principle, *Personalization* was covered by offering different options for input, output customization (e.g., voicebank and message tone), and multiple forms of providing detailed information about the user and their context. The *Personal vocabulary* topic was directly included to allow users to preprogram messages that express their personality and communication style through, for instance, jokes, stories or most-used vernacular terms.

Aside from the idea of an introductory mode, the *Individualized support* topic was partially addressed by collecting the user's personal information (e.g., profile) and offering an option that allows the system to dynamically learn from this data and their behaviour to offer personalized support. Contrarily, only previously identified *User's strategies* were considered, such as using customized numeric codes as input form. If further developed, this conceptual model must accompany robust studies on existing AAC users' strategies so as not to disregard what already works for them.

For *Input forms*, besides only word-to-word, we also considered different input forms such as *Bliss*. AAC systems should offer support for multiple input options and languages. The symbols could be, for instance, labeled and offered as shorthand option to either access a preprogrammed message or input them to an LLM that expands the message considering the user's data, behaviour and context. Like all languages, Blissymbolics requires some knowledge of its rules to communicate. Thus, a formal Bliss introduction and training should be offered to novel Bliss users. Furthermore, using symbols as an input method would make it hard for observers to interpret the message, enhancing privacy and making users feel more comfortable when composing their sentences.

In the *Adaptable* principle, *Context-awareness* was included in the Companion mode – which also tackles the *Communication companion* and *Smart and responsive* topics. If enabled, the system would collect the users' environmental data and associate information to generate real-time adaptable messages. Moreover, the user's time and location could help the system identify and suggest messages commonly used in a specific scenario.

Similarly, the *Conversation paths* topic was directly incorporated by designing a LLM-based interaction that allows it to generate multiple contextual-based suggestions. However, the feasibility of such model must be evaluated toughly to ensure that the outputs are relevant, and that the AI does not lead the conversation but helps the user to find the right words.

The *Fine-tuning topic* was only partially addressed by allowing the system to learn from the user data and behaviour. However, this would not be a short-term solution to the lack of an appropriate AAC dataset, as the system would still require enough data for training properly. *Typo tolerance* and *Grammatical correction* could both be tackled by choosing an appropriate model and, if necessary, perform training on noise data.

8 DISCUSSION

This work identified the benefits and challenges of applying LLMs to high-tech AAC systems alongside AAC users and familiar listeners. It provides readers with comprehensive and versatile guidelines, an extensive analysis, and reflection on the context. Additionally, considering its importance, Blissymbolics usage was also addressed. In the following paragraphs, the main outcomes of this work will be discussed according to the research questions detailed in subsection 1.2.

Benefits and challenges of an AI and AAC software. By mapping the literature and engaging AAC users in the research, this work identified several benefits for applying AI technologies, specifically LLMs, to AAC software. When cautiously designed, such models could potentially help AAC user in effectively communicating and expressing their unique personalities while still reducing cognitive and physical efforts.

In the literature on LLMs for AAC, some works contributed by exploring with prompt and abbreviation expansion systems or evaluating and developing synthetic datasets for AAC. The works detailed during the systematic mapping presented in subsection 4.1 are a valuable resource for future research that attempts to explore the possibilities of generative AI for assisting individuals with CCN. Our mapping attempted to synthesize, provide readers with details on each works' strengths and drawbacks, and draw directions for future works.

The literature mainly mentions that a benefit of LLMs for AAC would be to help reduce the time to compose a message. In contrast, the challenges identified included keeping users' personal expressivity, issues with privacy and autonomy, inaccurate predictions, lack of appropriate datasets, and societal stigma of AAC systems.

In the same way, our primary work's findings supported the same benefits and challenges identified in the literature. Additionally, the familiar listeners involved in our research also shared their concerns about the current rush to create AI applications. They emphasized the importance of building meaningful tools rather than hurrying to create new technologies.

On top of that, co-design participants shared many ideas on how AI could be implemented to successfully support AAC users, including retrieving previous replies to save efforts when composing for repeated interactions, assisting during group interactions, helping come up with phrases more quickly to defend themselves, adapting their message according to the context, and soothing communication with unfamiliar listeners.

Benefits and challenges of integrating Blissymbolics into an AAC software with LLM application. Reflecting on the results of the systematic literature mapping conducted in subsection 4.2 and the co-design findings detailed in subsection 5.3.4, we identify that the Blissymbolics languages have a crucial role in assisting individuals with CCN in expressing themselves. For this reason, future works could explore Bliss for AI language model usage in AAC, especially considering users who are developing their literacy skills and applications for educational domains.

Blissymbols can be used at any developmental level and are especially effective with individuals of high cognitive ability. As a non-spoken language with grammar and vocabulary, Bliss is able to accommodate the structure of a spoken language when desired. In this context, AI applications to AAC could accompany the learners' gradual development, provide personalized support, and, if required, assist the user in transferring their knowledge to other languages.

Nevertheless, data inequality is an ongoing issue in AI applications and the NLP field faces immense inequality in the development of novel technologies across the world's languages (Blasi; Anastasopoulos; Neubig, 2021). The majority of NLP works are focused on mainstream over-served languages, in which LLMs are consequently more efficient. The transfer of NLP developments from a high-resource language (e.g., English) to low-resource and minority languages – such as Blissymbolics – and other non-Latin writing scripts is a challenging, but valuable task.

Building a robust AAC system with AI language model integration for highly diverse users. The context of AAC is exceptionally complex and requires researchers to cautiously evaluate the diverse aspects included. This work attempted to thoroughly analyze and synthesize the various factors that contribute to or withhold AAC users from communicating and expressing their personalities. Rather than presenting impracticable solutions to this context, this work emphasizes the importance of perceiving technology as a tool and AAC users as experts in their own experiences.

This work proposes a comprehensive set of guidelines for researchers, developers, and designers that strive to create reliable, customizable, and adaptable AAC systems using LLMs. The guidelines cover different aspects that were either identified in the literature or emphasized by AAC stakeholders during the co-design activities, including privacy concerns, personalization, individualized support, input forms, output feedback, context-awareness, fine-tuning models, and more. Not only that, but this work also carefully reflects on the existing risks in this context and proposes other research opportunities.

We emphasize the importance of using the guide as an adaptable support for building robust AAC systems rather than using its principles as strict rules. As demonstrated in section 7, the guidelines should be adjusted according to each project's context, and missing points should be assessed for risks.

9 CONCLUSION

The following subsections will describe this work's main contributions (subsection 9.1) and limitations (subsection 9.2). Additionally, potential paths for future work and improvements will be outlined.

9.1 CONTRIBUTION

Both the primary and secondary research conducted drives this work's major contributions. Collecting information from the literature and by directly engaging with AAC stakeholders allowed us to gather multiple perspectives on the topic and draw straightforward directions for future research. This work not only discusses the existing opportunities and positive aspects, but it also extensively reflects on the complexity of AAC and highlights the importance of involving individuals with CCN as experts in AAC research.

While somewhat similar, our work differs from Valencia *et al.* (2023). Valencia's work focuses on capturing the views of AAC users on their prototype of an LLM-enhanced AAC system, but we take a step back; our work provides readers with a general perspective on AAC users' expectations, concerns and ideas about AI implementations in AAC.

Moreover, we addressed the importance of including multiple languages, especially Blissymbolics, which has helped many individuals with CCN due to congenital conditions to develop their literacy skills and express themselves instead of depending on needs-based communication. Even as crucial as Blissymbolics is, as identified in our systematic literature mapping, the Bliss community has received little attention in the last decade. Our work contributes simultaneously to AI, AAC and Blissymbols literature.

This work also contributes to the development of robust high-tech AAC with AI interventions by proposing guidelines for creating reliable, customizable and adaptable systems. Other crucial topics addressed by this work were the challenges and risks of such technologies and other research opportunities identified throughout our work development.

Our guidelines' versatility was demonstrated in section 7, where we developed a conceptual system with LLM integration. Besides explaining the system's functionalities, we discussed the guidelines' contents that were followed, and the risks related to those that were not. Considering this work's contributions from both the primary and secondary research conducted, we hope it becomes a helpful resource for AAC and AI researchers, developers and designers.

9.2 LIMITATIONS

Three limitations were identified in this study. Firstly, our systematic literature mapping on LLMs for AAC was conducted amidst the high popularity of the topic; because of that, the study can become quickly outdated as novel research on the field is constantly being conducted and published. The same does not apply to our mapping on Bliss for high-tech AAC, which has not seen many contributions in the last years.

Second, our co-design sessions included only four participants, two of whom were individuals with CCN due to cerebral palsy. While we were able to conduct in-depth conversations with highly skilled AAC users and familiar listeners who extensively collaborated on this project, we were only able to engage with a small sample of users. Interacting with more diverse users, such as individuals with various cognitive abilities and conditions, could have allowed us to gather more information and different perspectives on the topic.

Lastly, the guideline's development process considered only information collected from this work's systematic mappings and co-design findings. While we sought to support the development of robust AAC systems following a broader approach, the topics detailed might not cover all aspects within the multifaceted context of high-tech AAC.

9.3 FUTURE WORK

We foresee three paths for improving the current work in the future. Firstly, we would like to refine the proposed guidelines with help from AAC and AI experts and users. Collaborating with diverse stakeholders could assist with overcoming the limitations mentioned above. With this, we expect that the guidelines will cover more diverse perspectives and contexts. For instance, the guidelines could be updated to include suggestions on identifying and accommodating different users, especially throughout the various stages of their lives. Secondly, the systematic literature mapping on the use of LLMs for AAC should be updated to reflect the field’s progress. Furthermore, we plan on evaluating the feasibility and possibilities for enhancing the conceptual system presented in this work alongside specialists in the field.

REFERENCES

ADAMS, A.; LUNT, P.; CAIRNS, P. (2008). **A qualitative approach to HCI research.** In: Cairns, Paul and Cox, Anna (Ed) Research Methods for Human-Computer Interaction. Cambridge, UK: Cambridge University Press, chapter 7, pp. 138–157. Available at: <https://oro.open.ac.uk/11911/> Retrieved 03 nov. 2023.

American Speech-Language-Hearing Association. **Augmentative and Alternative Communication.** [*S.l.*]: ASHA, c2023. Retrieved from: <https://www.asha.org/public/speech/disorders/aac/> Accessed 10 sep. 2023.

ANDREASEN, P. N.; WALLER, A.; GREGOR, P. **BlissWord-Full access to Blissymbols for all users.** In: Proceedings of the 8th Biennial Conference of the International Society for Augmentative and Alternative Communication. [*S.l.*],1998. p. 167-168.

Apple. **AssistiveWare ushers in next generation of augmentative and alternative communication technology.** [*S.l.*]: Apple, 2023b. 17 May 2023. Retrieved from: <https://www.apple.com/newsroom/2023/05/assistiveware-ushers-in-next-generation-of-aac-technology/> Accessed 16 sep. 2023.

Apple. **Buy iPad 10th generation.** [*S.l.*]: Apple, 2023a. Retrieved from: <https://www.apple.com/shop/buy-ipad/ipad> Accessed 15 sep. 2023.

Augmentative Communication, Inc. **Section 5: Types of AAC Devices.** [*S.l.*]: ACI, c2023. Retrieved from: <https://www.augcominc.com/whatsnew/ncs5.html> Accessed 13 sep. 2023.

BALDAUF, M.; DUSTDAR, S.; ROSENBERG, F. A survey on context-aware systems. Int. J. Ad Hoc Ubiquitous Comput., Inderscience Publishers, Geneva 15, CHE, v. 2, n. 4, p. 263–277, jun 2007. ISSN 1743-8225. Available at: <https://doi.org/10.1504/IJAHUC.2007.014070>. Retrieved 20 Dec 2023.

BALLATI, F.; CORNO, F.; DE RUSSIS, L. **"Hey Siri, do you understand me?":** Virtual Assistants and Dysarthria. Intelligent Environments 2018. [*S.l.*]: IOS Press, 2018. p. 557-566.

BEUKELMAN, D. R.; MIRENDA, P. **Augmentative & alternative communication:** supporting children & adults with complex communication needs. 5th ed. Baltimore. Paul H. Brookes Pub. Co, 2020.

BLACKSTONE, S. W.; WILLIAMS, M. B.; WILKINS, D. P. **Key principles underlying research and practice in AAC.** [*S.l.*]: Augmentative and alternative communication, Taylor & Francis, v. 23, n. 3, p. 191–203, 2007.

BLASI, D.; ANASTASOPOULOS, A.; NEUBIG, G. Systematic inequalities in language technology performance across the world's languages. CoRR, abs/2110.06733, 2021. Available at: <https://arxiv.org/abs/2110.06733>.

BLISS, C. K. **Semantography** (Blissymbolics): A simple system of 100 logical pictorial symbols, which can be operated and read like 1+2=3 in all languages. 2nd ed. Sydney: Semantography (Blissymbolics) Publications, 1965.

Blissymbolics Communication International. **About BCI.** [*S.l.*]: BCI, [2023?c]. Accessed from: <https://www.blissymbolics.org/index.php/about-bci> Retrieved 26 oct. 2023.

Blissymbolics Communication International. **The Fundamental Rules of Blissymbolics**: creating new Blissymbolics characters and vocabulary. (PDF). [*S.l.*]: BCI, 2004. Available at: <https://www.blissymbolics.org/images/bliss-fundamental-rules-2020-06-16.pdf> Retrieved 20 apr. 2023.

Blissymbolics Communication International. **Where is Bliss used?** [*S.l.*]: BCI, [2023?b]. Retrieved from: <https://www.blissymbolics.org/index.php/where-is-bliss-used> Accessed 24 jun. 2023.

Blissymbolics Communication International. **Who uses Bliss?** [*S.l.*]: BCI, [2023?a]. Retrieved from: <https://www.blissymbolics.org/index.php/who-uses-bliss> Accessed 24 jun. 2023.

BROWN, T.; MANN, B.; RYDER, N.; SUBBIAH, M.; KAPLAN, J. D.; DHARIWAL, P.; NEELAKANTAN, A.; SHYAM, P.; SASTRY, G.; ASKELL, A.; AGARWAL, S.; HERBERT-VOSS, A.; KRUEGER, G.; HENIGHAN, T.; CHILD, R.; RAMESH, A.; ZIEGLER, D.; WU, J.; WINTER, C.; HESSE, C.; CHEN, M.; SIGLER, E.; LITWIN, M.; GRAY, S.; CHESS, B.; CLARK, J.; BERNER, C.; MCCANDLISH, S.; RADFORD, A.; SUTSKEVER, I.; AMODEI, D. **Language models are few-shot learners.** LAROCHELLE, H.; RANZATO, M.; HADSELL, R.; BALCAN, M.; LIN, H. (Ed.). In: Advances in Neural Information Processing Systems. Vancouver, Canada. Curran Associates, Inc., 2020. v. 33, p. 1877–1901. Available at: <https://proceedings.neurips.cc/paper_files/paper/2020/file/1457c0d6bfcb4967418bfb8ac142f64a-Paper.pdf>.

CAI, S.; VENUGOPALAN, S.; TOMANEK, K.; NARAYANAN, A.; MORRIS, M.; BRENNER, M. **Context-aware abbreviation expansion using large language models.** In: CARPUAT, M.; MARNEFFE, M.-C. de; RUIZ, I. V. M. (Ed.). Proceedings of the 2022 Conference of the North American Chapter of the Association for Computational Linguistics: Human Language Technologies. Seattle, United States: Association for Computational Linguistics, 2022. p. 1261–1275. Available at: <https://aclanthology.org/2022.naacl-main.91>. Retrieved: 02 October 2023.

Canadian Broadcasting Corporation. **How Justin Clark's fight for independence transformed disability rights in Canada.** Canada: CBC, 25 November 2018. Available at: <https://www.cbc.ca/radio/sunday/november-25-2018-the-sunday-edition-with-michael-enright-1.4911588/how-justin-clark-s-fight-for-independence-transformed-disability-rights-in-canada-1.4911590> Retrieved 29 Nov 2023.

CARLSON, R.; GRANSTRÖM, B.; HUNNICUTT, S. **Bliss communication with speech or text output.** In: IEEE International Conference on Acoustics, Speech, and Signal Processing. [s.n.], 1982. Available at: <https://api.semanticscholar.org/CorpusID:35188651>. Retrieved 19 may 2023.

CLARK, E.; AUGUST, T.; SERRANO, S.; HADUONG, N.; GURURANGAN, S.; SMITH, N. A. **All that's 'human' is not gold**: Evaluating human evaluation of generated text. CoRR, abs/2107.00061. [*S.l.*: s.n.], 2021. Available at: <https://arxiv.org/abs/2107.00061>.

COLTELL, O.; LLACH, J.; SANZ, P.; RUIZ, C.; CARRERES, D. **Alternative communication interface for severely handicapped people based on a multimedia human-computer interaction system.** In: POBIL, A. Pasqual del; MIRA, J.; ALI, M. (Ed.). Tasks and Methods in Applied Artificial Intelligence. Berlin, Heidelberg: Springer Berlin Heidelberg, 1998. p. 638–647. ISBN 978-3-540-69350-5. Available at: <https://doi.org/10.1007/3-540-64574-8_450> Retrieved 19 may 2023.

CURTIS, H.; NEATE, T.; GONZALEZ, C. V. **State of the art in AAC**: A systematic review and taxonomy. In: Proceedings of the 24th International ACM SIGACCESS Conference on Computers and Accessibility. New York, NY, USA: Association for Computing Machinery, 2022. (ASSETS '22). ISBN 9781450392587. Available at: <https://doi.org/10.1145/3517428.3544810>.

DEVLIN, J.; CHANG, M.; LEE, K.; TOUTANOVA, K. BERT: pre-training of deep bidirectional transformers for language understanding. CoRR, abs/1810.04805. [*S.l.*: s.n.], 2019. Available at: <http://arxiv.org/abs/1810.04805>. Retrieved 28 sep. 2023.

ELSAHAR, Y.; HU, S.; BOUAZZA-MAROUF, K.; KERR, D.; MANSOR, A. **Augmentative and alternative communication (AAC) advances**: A review of configurations for individuals with a speech disability. [*S.l.*]: Sensors, v. 19, n. 8, 2019. ISSN 1424-8220. Available at: <https://www.mdpi.com/1424-8220/19/8/1911>.

European Union. European Commission, Digital Strategy. **Regulatory framework proposal on artificial intelligence.** 15 November 2023. Available at: <https://digital-strategy.ec.europa.eu/en/policies/regulatory-framework-ai> Accessed 12 Dec 2023.

FERM, U.; AMBERNTSON, B.; THUNBERG, G. **Development and evaluation of a Minspeak™ application using Blissymbols**: Experiences from two case studies. Augmentative and Alternative Communication, Taylor Francis, v. 17, n. 4, p. 233–244, 2001. Available at: <https://doi.org/10.1080/aac.17.4.233.244>. Retrieved 19 may 2023.

GATTI, N.; MATTEUCCI, M. **CABA2L a Bliss predictive composition assistant for AAC communication software.** In: SERUCA, I.; CORDEIRO, J.; HAMMOUDI, S.; FILIPE, J. (Ed.). Enterprise Information Systems VI. Dordrecht: Springer Netherlands, 2006. p. 277–284. ISBN 978-1-4020-3675-0. Available at: <https://doi.org/10.1007/1-4020-3675-2_33>. Retrieved 19 may 2023.

Google. **Project Euphonia:** Helping everyone be better understood. [*S.l.*]: Google, 2019. 1 video (3 min). Retrieved from: <https://www.youtube.com/watch?v=OAdegPmkK-o>. Accessed 28 sep. 2023

Google. **Project Euphonia.** [*S.l.*]: Google, [2021?]. Retrieved from:<https://sites.research.google/euphonia/about/ >. Accessed 28 sep. 2023

GRESHAKE, K.; ABDELNABI, S.; MISHRA, S.; ENDRES, C.; HOLZ, T.; FRITZ, M. **Not what you've signed up for**: Compromising real-world LLM-integrated applications with indirect prompt injection. In: Proceedings of the 16th ACM Workshop on Artificial Intelligence and Security. New York, NY, USA: Association for Computing Machinery, 2023. (AISec '23), p. 79–90. ISBN 9798400702600. Available at: <https://doi.org/10.1145/3605764.3623985>. Retrieved 28 Dec 2023.

HANSEN, J. P.; TØRNING, K.; JOHANSEN, A. S.; ITOH, K.; AOKI, H. **Gaze typing compared with input by head and hand.** In: Proceedings of the 2004 Symposium on Eye Tracking Research & Applications. New York, NY, USA: Association for Computing Machinery, 2004. (ETRA '04), p. 131–138. ISBN 1581138253. Available at: <https://doi.org/10.1145/968363.968389>. Retrieved 22 Dec 2023.

HARWELL, Drew. **Why some accents don't work on Alexa or Google Home.** [*S.l.*]: The Washington Post, 19 jul. 2018. Retrieved from: <https://www.washingtonpost.com/graphics/2018/business/alexa-does-not-understand-your-accent/> Accessed 24 jun. 2023.

HELFMAN, E. S. **Signs and Symbols Around the World.** [*S.l.*]: Othrop, Lee & Shepard Co., 1967.

HIGGINBOTHAM, D. J.; CAVES, K. **AAC Performance and Usability Issues:** The Effect of AAC Technology on the Communicative Process. [*S.l.*]: Assistive Technology, v. 14, n. 1, p. 45–57, 30 jun. 2002.

HIGGINBOTHAM, D. J.; LESHER, G. W.; MOULTON, B. J., ROARK, B. **The application of natural language processing to augmentative and alternative communication.** [*S.l.*]: Assistive Technology, Taylor Francis, v. 24, n. 1, p. 14–24, 2012. Available at: <https://doi.org/10.1080/10400435.2011.648714>.

HUNNICUTT, S.; MAGNUSON, T. **Grammar-guided writing for AAC users.** [*S.l.*]: Assistive Technology, Taylor Francis, v. 19, n. 3, p. 128–142, 2007. PMID: 17937055. Available at: <https://doi.org/10.1080/10400435.2007.10131870>. Retrieved 19 may 2023.

Inclusive: The film. Directed by Miao Wang. United States: Microsoft Design in partnership with Cinelan, 2015. 1 movie (21 min). Available at: <https://www.youtube.com/watch?v=mkggZjUY7SY> Accessed 03 nov. 2023.

Instituto Brasileiro de Geografia e Estatística (IBGE). PNAD Contínua. **Pessoas com deficiência têm menor acesso à educação, ao trabalho e à renda.** Rio de Janeiro, 2023. Agência IBGE Notícias. Retrieved from: <https://agenciadenoticias.ibge.gov.br/agencia-noticias/2012-agencia-de-noticias/noticias/37317-pessoas-com-deficiencia-tem-menor-acesso-a-educacao-ao-trabalho-e-a-renda> Accessed 18 oct. 2023.

KANE, S. K.; MORRIS, M. R.; PARADISO, A.; CAMPBELL, J. **"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. [*S.l.*: s.n.], 2017. p. 1166-1179. Available at: <https://www.microsoft.com/en-us/research/publication/at-times-avuncular-and-cantankerous-with-the-reflexes-of-a-mongoose-understanding-self-expression-through-augmentative-and-alternative-communication-devices/> Retrieved 10 jun. 2023.

KELLY, R.; GOOCH, D.; PATIL, B.; WATTS, L. **Demanding by design**: Supporting effortful communication practices in close personal relationships. In: Proceedings of the 2017 ACM Conference on Computer Supported Cooperative Work and Social Computing. New York, NY, USA: Association for Computing Machinery, 2017. (CSCW '17), p. 70–83. ISBN 9781450343350. Available at: <https://doi.org/10.1145/2998181.2998184>.

KOESTER, H. H.; LEVINE, S. P. **Effect of a word prediction feature on user performance.** [*S.l.*]: Augmentative and Alternative Communication, v. 12, p. 155–168, 1996. Available at: <https://doi.org/10.1080/07434619612331277608>.

LIGHT, J.; MCNAUGHTON, D. **Putting people first**: Re-thinking the role of technology in augmentative and alternative communication intervention. [n.d.] Augmentative and Alternative Communication, Taylor Francis, v. 29, n. 4, p. 299–309, 2013. PMID: 24229334. Available at: <https://doi.org/10.3109/07434618.2013.848935> 10 jun. 2023.

LIGHT, J.; MCNAUGHTON, D. **The changing face of augmentative and alternative communication**: Past, present, and future challenges. [*S.l.*]: Augmentative and Alternative Communication, Taylor Francis, v. 28, n. 4, p. 197–204, 2012. PMID: 23256853. Available at: <https://doi.org/10.3109/07434618.2012.737024>.

LIU S.; SMITH D. A. **Adapting Transformer Language Models for Predictive Typing in Brain-Computer Interfaces.** arXiv preprint arXiv:2305.03819. [*S.l.*: s.n.], 5 May 2023. Available at: <https://arxiv.org/abs/2305.03819>. Retrieved: 02 October 2023.

LUNDÄLV, M. **Access to Blissymbolics in ICT state-of-the-art and visions.** Presented by Stephen van Tetzchner, ISAAC Research Symposium, Pittsburgh, USA, 2012. Available at: <http://www.meloton.com/filer/Bliss-ICT.pdf> Retrieved 20 jun. 2023.

MACKAY, D.; WILLS, S.; WALLER, A. **Bliss-Dasher–Efficient semantic writing with any muscle.** [s.n.]: Cambridge, UK, 2007. Available at: <http://www.inference.org.uk/dasher/development/bliss/abi/BlissDasher.pdf>. Retrieved 19 may 2023.

MOHER, D.; LIBERATI, A.; TETZLAFF, J., ALTMAN, D. G.; PRISMA Group. **Preferred reporting items for systematic reviews and meta-analyses**: The PRISMA statement. [*S.l.*]: American College of Physicians. Annals of Internal Medicine, v. 151, n. 4, p. 264–269, 2009. Available at: <https://doi.org/10.7326/0003-4819-151-4-200908180-00135>.

MR. SYMBOL MAN. Directed by Bruce Moir and Bob Kingsbury. Australia and Canada. Briarcliff Manor, N.Y.: Benchmark Films, 1974. 1 documentary movie (50 min) Available at: <https://archive.org/details/mrsymbolman> Accessed 20 apr. 2023.

NETZER, Y.; ELHADAD, M. **Using semantic authoring for Blissymbols communication boards.** In: Proceedings of the Human Language Technology Conference of the NAACL, Companion Volume: Short Papers. USA: Association for Computational Linguistics, 2006. (NAACL-Short '06), p. 105–108. Available at: <https://dl.acm.org/doi/10.5555/1614049.1614076>. Retrieved 19 may 2023.

OKRENT, A. **In the land of invented languages.** New York: Spiegel & Grau, 2009. *E-book.*

OLASZI, P.; KOUTNY, I; KÁLMÁN S. L. **From Bliss Symbols to Grammatically Correct Voice Output**: A Communication Tool for People with Disabilities. [*S.l.*]: International Journal of Speech Technology 5, 49–56 (2002). Available at: <https://doi.org/10.1023/A:1013682632553>. Retrieved 19 may 2023.

OpenAI. **Custom instructions for ChatGPT.** [*S.l.*], 2023. Retrieved from: <https://openai.com/blog/custom-instructions-for-chatgpt>. Accessed 28 sep. 2023

Ottawa Foyers Partage. **About Us**: Empowering people with disabilities to thrive since 1982. Canada: OFP, c2023. Retrieved from: <https://ottawafoyerspartage.org/about/> Accessed 26 oct. 2023.

PEREIRA, J.; NOGUEIRA, R.; ZANCHETTIN, C.; FIDALGO, R. **An augmentative and alternative communication synthetic corpus for brazilian portuguese.** In: 2023 IEEE International Conference on Advanced Learning Technologies (ICALT). [*S.l.*: s.n.], 2023. p. 202–206. Available at: <https://doi.org/10.1109/ICALT58122.2023.00066> Retrieved: 02 October 2023

PEREZ, F.; RIBEIRO, I. **Ignore Previous Prompt**: Attack Techniques For Language Models. [*S.l.*: s.n.], 2022. Available at: <https://doi.org/10.48550/arXiv.2211.09527> Retrieved 28 Dec 2023.

PRC-Saltillo. **Accent 1400.** [*S.l.*]: PRC-Saltillo, c2023. Retrieved from: <https://store.prc-saltillo.com/accent-1400> Accessed 15 sep. 2023.

PULLIN, G.; TREVIRANUS, J.; PATEL, R.; HIGGINBOTHAM, J. **Designing interaction, voice, and inclusion in AAC research.** [*S.l.*]: Augmentative and Alternative Communication, Taylor Francis, v. 33, n. 3, p. 139–148, 2017. PMID: 28675073. Available at: <https://doi.org/10.1080/07434618.2017.1342690>. Retrieved 03 nov. 2023.

ROMERO, M.; GUÉDRIA, W.; PANETTO, H.; BARAFORT, B. **Towards a characterisation of smart systems**: A systematic literature review. Computers in Industry, v. 120, p. 103224, 2020. ISSN 0166-3615. Available at: <https://www.sciencedirect.com/science/article/pii/S0166361520300270>. Retrieved 21 Dec 2023.

RUSSELL, S. J.; NORVIG, P. **Artificial intelligence**: a modern approach. 3rd ed. Upper Saddle River, New Jersey: Prentice-Hall, 2010.

Saltillo. **TouchChat App**: Discover the Joy of Communication. [*S.l.*]: Saltillo, c2023. Retrieved from: <https://touchchatapp.com/> Accessed 15 sep. 2023.

SENEVIRATNE, I. R. K.; BASNAYAKEGE, C. J**. Smart Bliss Board System for Multiple Disabilities.** Journal of Engineering and Technology of the Open University of Sri Lanka (JET-OUSL), v. 4, n. 1, 2016. Available at: <http://repository.ou.ac.lk/handle/94ousl/1005>. Retrieved 19 may 2023.

SENNOTT, S. C.; AKAGI, L.; LEE, M.; RHODES, A. **AAC and Artificial Intelligence (AI).** [*S.l.*]: Topics in language disorders, NIH Public Access, v. 39, n. 4, p. 389–403, 2019. Available at: <https://doi.org/10.1097/tld.0000000000000197>.

SHEARER, A. M. **Determining the meaning of Bliss Words.** 41 p. Master's Dissertation (M.Sc. in Applied Computing) - University of Dundee, Scotland, UK. September 2021

SHEN, J.; YANG, B.; DUDLEY, J. J.; KRISTENSSON, P. O. **Kwickchat**: A multi-turn dialogue system for aac using context-aware sentence generation by bag-of-keywords. In: 27th International Conference on Intelligent User Interfaces. New York, NY, USA: Association for Computing Machinery, 2022. (IUI '22), p. 853–867. ISBN 9781450391443. Available at: <https://doi.org/10.1145/3490099.3511145>. Retrieved: 02 October 2023.

SOHAIL, U.; TRAUM, D. **A Blissymbolics translation system.** In: CHRISTENSEN, H.; HOLLINGSHEAD, K.; PRUD'HOMMEAUX, E.; RUDZICZ, F.; VERTANEN, K. (Ed.). Proceedings of the Eighth Workshop on Speech and Language Processing for Assistive Technologies. Minneapolis, Minnesota: Association for Computational Linguistics, 2019. p. 32–36. Available at: <https://aclanthology.org/W19-1705>. Retrieved 19 may 2023.

THE SYMBOLS OF BLISS. Presenter: Michael Symmons Roberts. Participants: Michael Symmons Roberts, Peter Zein, Shirley McNaughton, and Brian Stride. Producer: Geoff Bird. Executive Producer: Jo Meek. [*S.l.*]: A Sparklab production for BBC Radio 4, 29 dec. 2018. *Podcast.* Available at: <https://www.bbc.co.uk/programmes/b08y26qf> Accessed 20 apr. 2023.

VALENCIA, S.; CAVE, R.; KALLARACKAL, K.; SEAVER, K.; TERRY, M.; KANE, S. K. **"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. New York, NY, USA: Association for Computing Machinery, 2023. (CHI '23). ISBN 9781450394215. Available at: <https://doi.org/10.1145/3544548.3581560>. Retrieved: 02 October 2023

VALENCIA, S.; PAVEL, A.; MARIA, J. S.; YU, S. G.; BIGHAM, J. P.; ADMONI, H. **Conversational agency in augmentative and alternative communication.** In: Proceedings of the 2020 CHI Conference on Human Factors in Computing Systems. New York, NY, USA: Association for Computing Machinery, 2020. (CHI '20), p. 1–12. ISBN 9781450367080. Available at: <https://doi.org/10.1145/3313831.3376376> Retrieved 10 jun. 2023.

VASWANI, A.; SHAZEER, N.; PARMAR, N.; USZKOREIT, J.; JONES, L.; GOMEZ, A. N.; KAISER, L.; POLOSUKHIN, I. **Attention is all you need.** Long Beach, CA, USA: 31st Conference on Neural Information Processing Systems (NIPS 2017). CoRR, abs/1706.03762, 2017. Available at: <http://arxiv.org/abs/1706.03762>. Retrieved 28 sep. 2023.

VERTANEN, K.; KRISTENSSON, P. O. **The imagination of crowds**: Conversational AAC language modeling using crowdsourcing and large data sources. In: Proceedings of the Conference on Empirical Methods in Natural Language Processing. USA: Association for Computational Linguistics, 2011. (EMNLP '11), p. 700–711. ISBN 9781937284114. Available at: <https://dl.acm.org/doi/10.5555/2145432.2145514>. Retrieved 04 Jan 2024.

WALLER, A. **Telling tales: unlocking the potential of AAC technologies.** [*S.l.*]: International Journal of Language & Communication Disorders, v. 54, n. 2, p. 159–169, 30 dec. 2018. Available at: < https://doi.org/10.1111/1460-6984.12449> Retrieved 02 jan. 2024.

WALLER, A.; JACK, K. **A predictive Blissymbolic to English translation system.** In: Proceedings of the Fifth International ACM Conference on Assistive Technologies. New York, NY, USA: Association for Computing Machinery, 2002. (Assets '02), p. 186–191. ISBN 1581134649. Available at: <https://doi.org/10.1145/638249.638283>. Retrieved 19 may 2023.

Web Acessibility Iniciative. **Web Content Accessibility Guidelines 2 (WCAG 2)**: Overview. 5 October 2023. Available at: <https://www.w3.org/WAI/standards-guidelines/wcag/> Accessed 15 Dec 2023.

WEIDINGER, L.; MELLOR, J.; RAUH, M.; GRIFFIN, C.; UESATO, J.; HUANG, P.; CHENG, M.; GLAESE, M.; BALLE, B.; KASIRZADEH, A.; KENTON, Z.; BROWN, S.; HAWKINS, W.; STEPLETON, T.; BILES, C.; BIRHANE, A.; HAAS, J.; RIMELL, L.; HENDRICKS, L. A.; ISAAC, W.; LEGASSICK, S.; IRVING, G.; GABRIEL, I. **Ethical and social risks of harm from language models.** CoRR, abs/2112.04359. [*S.l.*: s.n.], 2021. Available at: <https://arxiv.org/abs/2112.04359>.

World Health Organization. **Assistive technology.** Geneva: WHO, 2 Jan. 2024. Retrieved from: <https://www.asha.org/public/speech/disorders/aac/> Accessed 13 jan. 2024.

XU, W.; DAINOFF, M. J.; GE, L.; GAO, Z. **Transitioning to human interaction with AI systems: New challenges and opportunities for HCI professionals to enable human-centered AI.** [*S.l.*]: International Journal of Human–Computer Interaction, Taylor & Francis, v. 39, n. 3, p. 494–518, 2023.

YUSUFALI, H.; GOETZE, S.; MOORE, R. **Bridging the communication rate gap**: enhancing text input for Augmentative and Alternative Communication (AAC). GAO, Q.; ZHOU, J.; DUFFY, V. G.; ANTONA, M.; STEPHANIDIS, C. (Ed). In: HCI International 2023 – Late Breaking Papers. 25th International Conference on Human-Computer Interaction (HCII 2023). Copenhagen, Denmark. Lecture Notes in Computer Science, v. 14055. Jul 2023. . Springer Cham . ISBN 978303135681023-28. Available at: <https://doi.org/10.1007/978-3-031-48041-6_29> Retrieved: 02 October 2023.

APPENDIX A – **FULL TABLE OF WORKS ON BLISSYMBOLICS AND AAC**

Article	Country	Bliss input method	Evaluation	Users engaged	About	Strengths	Weaknesses and drawbacks	Recommendations extracted
Waller and Jack (2002)	Scotland	Bliss-Word system (Andreasen et al. 1998), which is built "on top of" Microsoft Word.	The accuracy of translation was tested with 20 sentences using three sizes of source text.	None	Bliss to English translation system that adjusts the output to generate a grammatically correct translated sentence instead of using a literal translation of each symbol from the gloss.	Pioneer in using predictive techniques to translate Bliss sentences into grammatically correct English rather than traditional laborious techniques of natural language using rule-based syntactic parsing.	Some concerning errors occurred, such as the insertion of inappropriate words and incorrect translations.	Users might not input all the symbols required to generate a grammatically correct utterance. Hence, a translation system should not only guarantee that the gaps within the sentence are appropriately filled but, more importantly, without altering the meaning.
Olaszi et al. (2002)	Hungary	Custom graphical interface with twelve symbols displayed at a time, with an automatic delay-based cursor.	The system was installed at the HCMC and continuously improved from 1994 to 1996.	One 17 year old boy with cerebral palsy	A symbol-to-speech system called Blissvox with a linguistic module that transforms the strings of isolated words into grammatically correct Hungarian sentences.	1) Detailed reporting of the system design, technical aspects and usage; 2) System requires only a simple PC and a synthesizer kit; 3) System was used by a Bliss user for an extended period; 4) System was improved continuously during the 2 years of experimentation.	1) The linguistic module and its results are not fully detailed in this paper; 2) Brief description of the experimentation phase.	1) It is crucial to fully evaluate the hardware that will be used for the system, and favour cheaper, easy to install and more reliable options; 2) The English-oriented nature of Bliss symbols imposed challenges for translating to Hungarian, and using language processing should be preferred but thoroughly evaluated.
Netzer and Elhadad (2006)	Israel	Custom digital lexicon, search can be done by keyword, or by semantic/graphic component.	Keystroke savings, comparing the number of choices needed to generate a set of sentences with their system.	None	NLG-based (Natural Language Generation) system for Bliss users using semantic authoring, whereas the available choice of symbols dynamically changes according to the syntactic and semantic content of the selected symbols.	1) Addresses the context challenge by adding pre-defined, yet flexible, inputs into the system, and enabling the assignment of default values to each conversation; 2) The intervention provides early feedback as it occurs during the process of composing the input sequence; 3) Implementation of a Bliss lexicon for both Hebrew and English.	1) Evaluation tests were conducted with non-AAC users; 2) Evaluations showed no improvement in keystroke saving.	1) When processing a Bliss sentence, real-time feedback of the output result is preferred; 2) Extra caution when parsing highly abbreviated telegraphic sentences, which can be incorrectly deciphered and misinterpreted since conversations rely heavily on pragmatics.
Carlson et al. (1982)	Sweden	A 500-symbol Bliss board with multi-lingual text-to-speech system	Not mentioned.	None by the time of publication	Details the development, structure and operation of both the software and the device for the BlissTalk project. Its framework uses syntactic parsing to create a multi-lingual text-to-speech system.	1) Detailed reporting of the system design, technical aspects and usage; 2) Uses a multi-lingual system.	1) No AAC or Bliss users engagement was mentioned; 2) No evaluation was conducted.	1) Text-to-speech systems should offer ways of improving the naturalness of synthetic speech (e.g., allowing pause insertion for prosody); 2) Provide access to voice output customization (i.e., different tunes based on age and gender); 3) Importance of building adaptive models that are able to follow the user's development.
Gatti and Matteucci (2006)	Italy	An AAC communication software based on Bliss called Bliss2003.	Quantitative analysis for prediction of training error and generalization error.	Collaborated with two clinics for verbal impairments.	An automatic scansion system for Bliss sentences composition focusing on AAC users with motor disabilities. It is called CABA2L and uses an Auto-Regressive Hidden Markov Model.	1) Detailed reporting of the system design, model and technical aspects; 2) Analyzes both human-computer interaction and technical strategies; 3) Adopts a hybrid prediction approach that allows to model the user unique language behaviour.	Poor description of the graphical interface and input method chosen.	Language prediction systems must take into account that each verbally impaired person develops their own syntactic model according to their capabilities, and so applying pure statistical or syntactical approaches may be inadequate.
Coltell et al. (1998)	Spain	Not detailed. Special input devices are handled as the mainstream input ones.	User tests, not concluded by the time of publication.	Eight users from a collaborator institution.	COMBLISS, a conceptual model of a communication system that operates interactively using speech synthesis, graphic and textual representations of Bliss symbols, and includes settings management options for tutors.	1) System is compatible with multiple special input devices that are handled as the mainstream ones; 2) Evaluation tests were conducted with eight AAC users.	1) No description of the graphical interface; 2) Included a user profile for tutors (more access and operative privileges), but did not discuss the AAC user privacy and autonomy issues related; 3) Poor results due to the prototype limitations; 4) Unfinished evaluation by the time of publication; 5) Reported that the AAC user must acquire enough experience in the system to improve the usage performance.	1) Offering compatibility with a wide range of assistive devices might prevent the system from being exclusively used by a favoured group; 2) The system must be easy to install or customize by the end-users; 3) Whenever opting to include different profiles for tutors that allow for more access and operative privileges, it is crucial to consider the AAC user's privacy and autonomy.

Article	Country	Bliss input method	Evaluation	Users engaged	About	Strengths	Weaknesses and drawbacks	Recommendations extracted
Ferm et al. (2001)	Sweden	Touch Talker™ overlay combining 54 basic Blissymbols with Minspeak strategy.	Case study with continuous adjustments to fit each user particularity.	Two Bliss users	Details the development of a Minspeak™ Swedish-language phrase-based application using Blissymbolics. The work was carried out in the form of case studies with two experienced Bliss users and their tutors. The application turned out not to be sufficient for the communicatively advanced study participants, who then required a second word-based vocabulary to maximize voice output communication aid use.	1) Detailed narrative of the case study and results; 2) Included an option that allowed the user to easily program personal private vocabulary; 3) The word-based application was easily incorporated into one of the user's daily activities as they were not required to plan and program new vocabulary in advance; 4) Improved the reading and writing abilities of one user; 5) Overall general positive results with both users involved.	1) They reported that one of the users was a perfectionist who wanted to correctly construct every sentence, so he had difficulties using the word-based application as it slowed down his communication; 2) Reported having issues with the sound and voice characteristics of one of the user's voice output communication aid; 3) No description of the interface.	1) Offer options that allow the user to independently customize and access private vocabulary; 2) The system must be able to adapt to the user's unique language behaviour and ability, as some may opt to use highly telegraphic communication while others might prefer a long and grammatically advanced sentence; 3) Caution when opting for a system that requires the user to memorize codes to access the desired output, as this might increase the learning curve and cognitive load, especially for novice users.
Hunnicutt and Magnuson (2007)	Sweden	Selection of preprogrammed messages, a grammar-guided writing or free selection of symbols.	Beta user tests for adjustments.	Two Bliss users	A grammar-guided email writing tool for AAC users called Grammar-Guided Writing (GGW) where the user can choose preprogrammed messages and e-mail phrases or select graphic signs freely without grammar guidance.	1) Detailed reporting of the system design, interface, technical aspects and usage; 2) Multi-language system; 3) The systems operate using a fixed word order, and each word in the output is under the control of the users to preserve their intention; 4) Beta tests were conducted with two Bliss users; 5) The user can learn to access the grammatical structures gradually using their stepwise introduction.	No reporting or analysis of the beta tests was conducted.	Adding an introductory mode to the AAC and Bliss systems can be helpful to guide novice users and ease the learning curve. However, it should also allow expert users to customize the system to their preferences and needs.
Sohail and Traum (2019)	United States of America	No user interface. Each Bliss symbol, word or character, was given a unique 4-5 digit numeric ID for input.	Quantitative analysis using BLEU algorithm.	None	Based on previous work (Waller and Jack, 2002), the authors build a proof of concept for a new translating system from Blissymbols to English.	Added the possibility to handle alternate spellings and unseen conjugations (e.g., past tense or gendered words).	Results show some errors from either the language model or due to the Blissymbolics structure.	A text-to-speech context-aware component can be an important component for supporting the users to be more independent of a human interpreter.
MacKay et al. (2007)	England, Scotland	Dasher, a text-entry interface driven by continuous two-dimensional gestures.	Not mentioned.	Not mentioned.	Implementation of Bliss communication in the Dasher system, which is a text-entry interface that requires few gestures and is adaptable to different AAC devices.	1) Open-source system available for download; 2) Detailed reporting of the system design.	1) No AAC or Bliss users engagement mentioned; 2) No evaluation mentioned.	The absence of a Unicode standard for Bliss was reported to be an issue. However, it is essential to evaluate if there is an alternative option within the technology available nowadays to address this problem without limiting the language.
Seneviratne and Basnayakege (2016)	Sri Lanka	Headband with built-in accelerometer used as cursor to point to a smart Bliss board using head movements.	Gradual evolution of the system development followed by testing it with an user.	Applied at the Prithipura Infant Home. The number of users was not mentioned.	A system that consists of a desktop-based interface displaying a Bliss Board and headband device with an accelerometer that detects the users' head movements to select the symbols.	1) Experimentation with a novel device for head interaction; 2) The system does not require the tutors to concentrate on each child at a time, as opposed to traditional methods.	1) Poor reporting of the evaluation; 2) Accuracy and usability of the device developed were not measured.	Consider the use context and prefer scalable AAC systems to different situations rather than exclusive to one environment.

APPENDIX B – BLUEPRINT 1: DIRECTIONS FOR FUTURE WORKS AND CHALLENGES (EXTRACTED FROM THE LITERATURE)

Recommendations

Subcategories	Topics covered
System Design	(1) Compatibility with assistive devices; (2) Hardware requirements; (3) Avoiding context-exclusive systems; (4) Introduction to Bliss.
User-centered	(1) Highly customizable; (2) Private vocabulary; (3) Smart and responsive; (4) Individualized syntactic model; (5) Considering existing strategies; (6) Privacy and autonomy.
AI language model	(1) Grammatical correction; (2) Output feedback (real-time); (3) Context-awareness and personalization; (4) Fine-tuning and adequate dataset and size.
User Input	(1) Support multiple options; (2) Typo tolerance.

Challenges

Subcategories	Topics covered
Technology-related	(1) "Golden reply" VS semantically equivalent; (2) Incorrectly deciphered and misinterpreted highly abbreviated telegraphic sentences; (3) Improved accuracy: increased latency and hardware requirements.
Bliss-related	(1) Absence of a Bliss Unicode standard; (2) English-oriented nature of Bliss.

Other opportunities

Topics covered
(1) Necessity for AAC corpora; (2) UI strategies for partial selection in incorrect predictions.

APPENDIX C – BLUEPRINT 2: USER'S EXPECTATIONS AND CONCERNS (COMPILED FROM CO-DESIGN)

Expectations and ideas	
1) AT-mediated communication	
Subcategories	**Topics covered**
Communication Partner	(1) Make easier to help keep in touch and connect with friends, family and partners; (2) Help communicating with medical staff and caregivers more independently; (3) Help talking with unfamiliar listeners; (4) Help coming up with phrases more quickly (defend yourself).
Context and conversation path	(1) Put users idea in the right context; (2) Update incorrect information according to the context; (3) Anticipate sentence without directing the conversation, but offering the paths to it.
2) System Design	
Subcategories	**Topics covered**
Consistency	Always available.
Memory	(1) Store and retrieve previous communication interactions; (2) Related previous information to facilitate communication.
Portability and adaptability	(1) Prefer portable options; (2) Facilitate adapting from previous devices.
Individualized	Provide individualized support that respond to the needs of individual users.
Learn and communicate	Provide separate modes for learning and communication.
3) Bliss	
(1) Bridging languages using Bliss, as it focusses on meaning and not on phonetics; (2) Accommodate Bliss and communication to each users' needs; (3) Bliss as a choice.	
4) Other	
(1) Have an AI that helps users navigate the internet and make it more accessible; (2) Have an AI that helps with automation of commands (e.g., for login in a website) regardless of connection issues; (3) Have an AI that helps users with going out independently by, for example, moving chair with computer, help opening doors and windows.	
Concerns	
Subcategories	**Topics covered**
Technology	(1) Wasting the users time with incorrect suggestions; (2) Technology advancing more quickly than the knowledge of Blissymbolics and AAC users; (3) AI interpretations made from the bigger population of users; (4) Concern about users relying on the technology to direct their thought process.
Privacy	Concerns with privacy, especially audio capturing features.