Designing a Tool that automatically translates Makaton signs from live video streams into written English

Sokratis Karkalas
s.karkalas@derby.ac.uk
University of Derby
Derby, United Kingdom

Omotayo Emmanuel Omoyemi
o.omoyemi1@unimail.derby.ac.uk
University of Derby
Derby, United Kingdom

ABSTRACT

This paper presents the design process for a tool intended to automatically translate Makaton signs from live video streams into written English. Makaton, a communication system utilising signs and symbols to assist individuals with communication difficulties, requires an effective translation system to improve accessibility. To guide the design of this tool, we conducted a thematic analysis of information collected from experts in the field. This analysis identified key themes and needs, which were used to develop design specifications for the tool. The output was evaluated and validated through member checking, confirming that the results accurately reflected the experts' knowledge, experiences, and needs. Feedback from participants was incorporated to refine the design specifications, ensuring they were grounded in practical application and user expectations. The refined specifications will inform the subsequent development of a prototype. This paper outlines the methodological approach, including thematic analysis, member checking, and the formulation of design specifications. It highlights the role of iterative feedback in enhancing the tool's relevance and effectiveness, and underscores the importance of user-centered design in assistive technology development. The study provides a framework for future research and development in this area.

CCS CONCEPTS

• **Software and its engineering → Requirements analysis**; **Software design engineering**.

KEYWORDS

Makaton Translation, Assistive Technology, Sign Language Recognition, User-Centered Design, Thematic Analysis, Member Checking, Design Specifications, Communication Aids, Qualitative Research, Accessibility Tools

ACM Reference Format:
Sokratis Karkalas and Omotayo Emmanuel Omoyemi. 2024. Designing a Tool that automatically translates Makaton signs from live video streams into written English. In *2024 13th International Conference on Software and Information Engineering (ICSIE) (ICSIE 2024), December 2–4, 2024, Derby, United Kingdom.* ACM, New York, NY, USA, 9 pages. https://doi.org/10.1145/3708635.3708647

ICSIE 2024, December 2–4, 2024, Derby, United Kingdom

ACM ISBN 979-8-4007-1776-5/24/12
https://doi.org/10.1145/3708635.3708647

1 INTRODUCTION

Effective communication is a fundamental human right, yet it remains a significant challenge for individuals with substantial difficulties in verbal communication. This group encompasses a diverse range of individuals affected by developmental, acquired, or progressive conditions, as well as those with temporary speech impairments. The spectrum of these conditions includes people of all ages with varying degrees of disabilities that impact their ability to speak or produce intelligible speech [1, 2, 9].

The challenges faced by individuals with severe verbal communication difficulties are multifaceted and affect nearly every aspect of their lives [41]. These difficulties can severely impact emotional and psychological well-being, social integration, educational and professional opportunities, and overall quality of life. Individuals with communication impairments often experience frustration, isolation, and dependence, which can lead to mental health issues such as anxiety, depression, and low self-esteem [25, 30]. Their ability to form and maintain relationships, participate in community activities, and feel included in society can be significantly hindered. Furthermore, limited communication skills can constrain educational achievements and career prospects, leading to reduced economic independence and social mobility. Overall, these challenges can substantially diminish an individual's quality of life, affecting their ability to engage fully in life and reach their potential.

In addition to these challenges, individuals with communication difficulties often face significant obstacles in text-based communication. The rise of digital communication methods - such as smartphones, social media, and video conferencing - has introduced new barriers for those who struggle with verbal or written communication [40]. Writing text messages can be particularly challenging for individuals with verbal communication difficulties, as it requires not only the ability to formulate coherent thoughts but also to express them clearly in writing [26]. This can be compounded by cognitive or literacy issues, motor impairments, or a lack of familiarity with the technology used. Consequently, these individuals may experience further marginalisation in digital spaces, where effective written communication is essential for participation and engagement.

The relevance of addressing these communication challenges is particularly important today due to several contemporary factors. The shift toward digital communication has not always been accompanied by advancements in accessibility, leaving many platforms inadequate for users with communication difficulties. The COVID-19 pandemic further highlighted the critical need for effective communication tools, as remote interactions became more prevalent [36, 49]. Individuals with communication impairments

faced exacerbated isolation and barriers in accessing essential services. This underscores the urgent need for innovative solutions that bridge communication gaps and facilitate greater inclusivity.

Augmentative and Alternative Communication (AAC) systems have long been employed to address the communication challenges faced by individuals with significant speech and language difficulties [3, 26]. AAC encompasses a broad range of methods, from unaided systems like sign language and gestures to sophisticated electronic devices that generate speech [14, 19]. These systems are tailored to meet diverse needs, allowing for effective communication across various contexts [21]. AAC is utilised in educational settings, healthcare, and daily living to support individuals with developmental disabilities, acquired conditions (e.g., stroke, traumatic brain injury), and progressive neurological diseases (e.g., ALS, Parkinson's). By providing alternative communication means, AAC helps individuals maintain social connections, express needs, and participate more fully in their communities.

Among the AAC systems, manual signing is often noted for its expressive potential [13, 32, 44]. It allows for a rich, nuanced form of communication capable of conveying complex ideas, emotions, and subtleties. Sign languages like British Sign Language (BSL) possess their own grammar, syntax, and nuanced expressions, similar to spoken languages. However, the effectiveness of manual signing presupposes that both the sign user and their communication partners are familiar with the specific sign language or signing system. This requirement poses a challenge for widespread communication. To address this, manual signing is frequently used in conjunction with other methods, such as text-based communication, speech-generating devices, or visual aids like PECS [7, 15, 20, 47].

This paper explores the potential of an automated Makaton-to-English translation tool as a means to bridge the communication gap between Makaton users and those unfamiliar with the system. Makaton [10, 17] was selected for this study due to its advantages over BSL [44, 45] and Signalong [16], particularly its accessibility and integration with speech. By combining signs with symbols and spoken language, Makaton offers a supportive communication method that is easier to learn and use for various needs. This automated translator aims to reduce the learning burden for English speakers, fostering more inclusive interactions without requiring mutual familiarity with the communication system.

2 PREVIOUS WORK

Augmentative and Alternative Communication technologies have a long history, with early roots in the 17th century. In 1680, George Dalgarno introduced a "talking glove" in his book *Didascalocophus, or the Deaf and Dumb Man's Tutor*, mapping letters of the alphabet onto a glove to facilitate communication through touch. This concept was later adapted by Alexander Graham Bell in 1873 to teach a congenitally deaf boy, marking an early step in gesture-based communication. In the mid-20th century, Norbert Wiener, the founder of cybernetics, further advanced these ideas with his invention of the "hearing glove", which converted sound vibrations into tactile sensations [33]. These foundational efforts laid the groundwork for the development of gesture-based communication tools - assistive technologies that leverage gesture recognition systems to interpret physical movements such as hand signs, body gestures, and facial expressions, translating them into digital outputs like text or synthesised speech.

One of the early contemporary successes in the domain of gesture-based communication tools involved systems utilising accelerometers and motion sensors to track hand movements and recognize specific gestures. A notable example is the "PowerGlove" developed by Mattel in 1989, which, though initially designed for gaming, laid the groundwork for later assistive technologies. Another important system was "CyberGlove", introduced in the 1990s, which used sensors embedded in a glove to measure hand positions and movements [43]. These systems often relied on predefined gesture libraries, where each gesture was mapped to a corresponding word, phrase, or command, enabling the recognition of specific gestures and translating them into digital outputs such as text or speech. These early innovations demonstrated the potential for technology to support individuals with speech and language impairments, contributing significantly to the development of more sophisticated AAC systems.

Another major advancement in gesture-based communication tools is the integration of computer vision technology. Systems like those presented by [29, 46] illustrate this progress by capturing and processing sequences of images of users' gestures, enabling real-time recognition and translation. Unlike earlier sensor-based approaches, computer vision-based tools can interpret a wider range of gestures, including subtle fine motor movements and facial expressions, resulting in a more refined and nuanced communication experience.

Despite their promise, gesture-based communication tools face several challenges. One key limitation is the accuracy and reliability of gesture recognition, particularly in uncontrolled environments with variable lighting or background conditions. Additionally, the complexity of gestures in sign languages - where meaning can be influenced by subtle changes in movement, hand orientation, or facial expressions - makes it difficult for these systems to achieve the level of expressiveness and fluency required for natural communication [39]. Furthermore, while these tools are designed to assist users in conveying their messages, they often require both the user and the communication partner to have some level of familiarity with the system or the underlying gesture vocabulary. This dependency can limit the broader applicability of such tools in spontaneous, everyday interactions where both parties may not share the same knowledge or communication skills.

In response to these limitations, systems like Mak-Messenger [38], emerged as simpler, more user-friendly alternatives. These solutions are less complex in design and implementation, with a small learning curve, making them accessible to users with limited familiarity with gesture vocabulary. By focusing on predefined inputs and straightforward functionality, Mak-Messenger allows users to communicate effectively without the need for extensive training or prior knowledge, thus making it more suitable for spontaneous, everyday interactions where both parties may not share the same communication skills. Users input signs via a touchscreen or predefined symbol selection, which the app then translates into written English, enabling seamless communication with non-Makaton users. However, while Mak-Messenger provided a groundbreaking solution for text-based communication, it had its limitations.

The reliance on manual input of signs or symbols meant that real-time communication was not fully supported. Users had to know and select the correct symbols for their messages to be translated, which could be time-consuming and potentially cumbersome for more complex conversations. Additionally, the absence of live video processing capabilities limited its use to pre-defined signs and symbols, making it less adaptable to spontaneous, dynamic interactions. The development of Mak-Messenger highlighted the growing need for more sophisticated tools that could handle real-time, dynamic communication without requiring extensive input from users.

Building on the need for real-time communication, FingerChat [38] represents a significant advancement. This system recognises finger spelling by capturing video of the user's hand movements and analysing them with computer vision techniques to translate these gestures into text. Although FingerChat excels in precision and is particularly useful for spelling names and technical terms, it is limited to finger spelling and does not encompass the full complexity of sign language, such as gestures and facial expressions. Additionally, its performance is optimised in controlled environments with consistent lighting.

Other more recent developments in AAC technology include systems like CoughDrop, MyChoicePad, Widgit Go and Sign4Me. CoughDrop[1] is a commercially available and relatively recent AAC app that offers versatile communication support. Although not specifically designed for Makaton, it allows users to create custom communication boards incorporating Makaton symbols. The app's customizable library enables the addition of relevant signs and symbols, supporting text-to-speech and symbol-to-speech translation, making it an effective tool for users to communicate more easily. MyChoicePad[2] is a language and communication development app that integrates Makaton symbols and signs. It is designed to help users learn Makaton and use it for communication. The app includes a vast library of Makaton signs and symbols, paired with visual aids and learning tools to support users in improving their communication skills. While primarily educational, it also serves as a communication aid for Makaton users. Widgit Go[3]: Similar to MyChoicePad, Widgit Go is an app that supports communication and learning through symbol-based interactions. It includes a range of symbol sets, including those compatible with Makaton. Users can create communication grids that translate Makaton symbols into text or speech, enhancing communication for those with speech and language difficulties. Sign4Me[4] is an app designed for American Sign Language (ASL) that demonstrates a similar functionality to what could be achieved with a Makaton-specific translation tool. Sign4Me converts ASL signs into written text or speech, facilitating communication between sign language users and non-signers. A Makaton-specific app modeled after this approach could provide analogous features, translating Makaton signs into English in real-time and thereby enhancing communication for Makaton users.

In summary, the evolution of gesture-based communication tools, from the early "talking glove" to modern AAC apps, demonstrates a steady progression in assistive technologies aimed at improving communication for individuals with speech and language impairments. Early innovations like sensor-based systems and computer vision technology have advanced the field significantly, offering more sophisticated recognition of gestures, finger spelling, and even facial expressions. However, challenges remain, particularly in the areas of real-time communication, accuracy in uncontrolled environments, and the complexity of recognising the full spectrum of sign languages. Furthermore, solutions like Mak-Messenger and FingerChat, while useful for specific tasks, lack adaptability for dynamic, everyday interactions.

The research gaps identified - particularly in the need for real-time processing and broader, more intuitive gesture recognition - highlight the opportunity to design a more accessible and efficient tool. Developing a new translation tool that can automatically convert Makaton signs into English text would address many of these limitations. By offering a system that is less dependent on predefined gestures and more adaptable to spontaneous interactions, such a tool could greatly enhance communication for people with significant speech and language difficulties, bridging the gap between sign-based and verbal communication in a variety of contexts.

This research contributes to the expanding field of automated translation technologies by specifically addressing AAC users' needs. Building on studies highlighting the limitations of current approaches, it explores potential advantages of real-time, dynamic translation. Positioned within a broader context, it acknowledges the demand for domain- and language-independent systems, and machine learning integration [11, 12, 23, 42, 50]. This work also recognises ALMs' potential for nuanced communication support [24] and underscores opportunities for AAC advancements toward more responsive interaction tools.

3 METHODOLOGY

This study follows a qualitative research design, employing semi-structured interviews [31], thematic analysis [35, 48], and member checking [6, 34] to inform the design of the tool. The methodology is structured into four key phases: data collection through interviews, thematic analysis, evaluation through member checking, and the production of design specifications.

3.1 Gaining Insights to Inform Further Design

The initial two phases of the design process involve conducting semi-structured interviews with experts followed by a thematic analysis. This approach was selected for two primary reasons: the team's lack of expertise in the subject matter and the desire to avoid preconceptions and biases. When a design team lacks specialised knowledge, it is crucial to gain insights from experts early in the process. Experts offer reliable, detailed, and context-specific information that aids in formulating pertinent questions and aligning the design process accordingly. Their input is invaluable for identifying potential pitfalls and challenges that may not be apparent to non-experts [5]. Moreover, incorporating expert insights from the outset helps to mitigate the influence of the team's own assumptions and biases on the design process. By grounding our initial understanding in expert knowledge, we establish a more objective and accurate foundation for the design.

[1]https://www.mycoughdrop.com/
[2]https://www.mychoicepad.com/
[3]https://www.widgit.com/products/widgit-go/
[4]https://specialedtechcenter.com/product/sign-4-me/

As mentioned above, the methodology in this part consists of semi-structured interviews and a thematic analysis. This combination was selected due to its effectiveness in obtaining comprehensive insights. Semi-structured interviews allow for the in-depth exploration of key topics, offering the flexibility to pursue interesting points raised by the expert. This method generates detailed qualitative data, providing profound insights into the expert's knowledge and experience, while ensuring that all relevant topics are covered and allowing for the emergence of new and unexpected information. Thematic analysis, on the other hand, offers a structured approach for identifying, analysing, and reporting patterns within qualitative data. This technique aids in generating themes that reveal underlying issues, needs, and opportunities related to the design of the tool. By identifying these key themes, actionable insights can be derived to inform design decisions and strategy.

3.2 Refinement and Validation of Insights

To ensure the accuracy and relevance of the identified themes and to validate the findings from the initial analysis, member checking was employed as a crucial evaluation step. Member checking involves returning the analysed data and findings to the original participants for their review and feedback. This process helps to confirm that the themes and insights accurately reflect the participants' experiences and perspectives.

3.3 Design Specification Development

Based on the validated themes, the final phase involves developing design specifications for the translation tool. This phase focuses on translating the identified themes and insights into practical and actionable requirements for the tool.

This comprehensive approach ensures that the design specifications are well-grounded in expert knowledge and user feedback, providing a solid foundation for the development of the tool.

4 IMPLEMENTATION

4.1 Conducting Semi-Structured Interviews

4.1.1 Recruitment of Experts. Recruiting experts for semi-structured interviews involves several key steps to ensure the selected individuals possess the requisite knowledge and experience relevant to the study. The initial step was to define the criteria for expertise, determining the specific knowledge, skills, and experience necessary for qualification. These criteria included academic qualifications, professional experience, contributions to the field, and recognized status within professional organizations.

Expertise Criteria

The criteria established for selecting experts were as follows:

- **Skills**: Proficiency in special education, child development, communication, teaching, and teacher training. The inclusion of teacher training was due to its common expectation in more experienced practitioners.
- **Educational Background**: A minimum of a master's degree in subjects related to education or special education.
- **Experience**: At least five years of experience engaging with students using sign languages.
- **Qualifications**: Fully qualified teachers with accreditation from established national or international professional bodies.
- **Contributions**: Published research or professional work in the area of communication aids or educational technology.

Identification of Potential Experts

Using the aforementioned criteria, we utilized LinkedIn, a professional networking platform, to identify potential experts. The search query was designed to find profiles containing the exact phrase "qualified teacher" and any one of the specified terms: special needs, special education, sign language, autism practitioner, or autism teacher. The Boolean search query used was:

```
"qualified teacher" AND
(
"special needs" OR
"special education" OR
"sign language" OR
"autism practitioner" OR
"autism teacher"
)
```

Additional Filters and Selection Process

Additional filters were applied to further refine the search:

- **Locations**: United Kingdom, England
- **Industry**: Education, Primary and Secondary Education
- **Profile Language**: English

We manually reviewed the first 100 profiles and selected 10 candidates based on our defined criteria of expertise.

Outcome

All selected candidates were approached for an interview, resulting in two positive responses.

4.1.2 Development of the Interview Guide. The interview guide used for these interviews was intentionally designed to incorporate flexibility. Although a set of prepared questions was established, the interviewer was allowed the discretion to deviate from these questions, enabling them to ask follow-up questions that probe deeper into responses or to explore new topics that might emerge during the conversation. This approach was intended to facilitate a more natural and open-ended dialogue, encouraging interviewees to provide more detailed and nuanced responses, ultimately leading to richer insights. Given the objective of exploring complex issues, understanding experiences, and gaining insights that require a deeper understanding, a semi-structured format was adopted in the design of the interview guide.

The topics covered are grouped into the following categories:

- **Understanding Makaton and its use**: This section addresses background knowledge and experience with Makaton, as well as the challenges associated with its use.
- **Technical feasibility and challenges**: This section addresses the key challenges in translating Makaton signs into English, including variability in signing speed, style, and the physical abilities of individuals performing Makaton signs. It also examines existing technologies and methods

for real-time sign recognition, as well as potential limitations of current solutions.

- **User Experience Considerations**: This section addresses issues related to balancing the accuracy of sign recognition with the usability of the tool, such as ensuring a smooth and responsive user experience. It also explores the type of feedback the tool should provide to users, as well as considerations for accessibility.
- **Ethical and Practical Considerations**: This section explores privacy concerns and potential strategies for mitigation, assesses the tool's usability for diverse users to ensure inclusivity, and evaluates the anticipated impact on the Makaton-using community, including potential risks and benefits.
- **Potential Improvements and Innovations**: This section discusses technological and methodological advancements that could significantly enhance the accuracy or usability of the tool, as well as additional functionalities that would add value to the process.
- **Final Thoughts**: This section is intended to facilitate an open discussion on topics not covered in the previous sections. It addresses important considerations for the design and development of a Makaton translation tool, as well as potential resources and research that could be valuable for the project.

4.1.3 Interview Process. The interview guide outlined in the previous section was shared with the interviewees prior to the interview. This approach was intended to allow interviewees to prepare their thoughts and deliver more considered and detailed responses, which is particularly beneficial for complex or technical subjects, as is the case with this work. By providing the guide beforehand, interviewees could feel more at ease, leading to a more relaxed and productive conversation. Additionally, knowing the topics in advance helps interviewees focus their responses on the most relevant aspects, resulting in more targeted and valuable insights. This preparation also enhances the efficiency of the interview process by ensuring that both the interviewer and interviewee are aligned on the discussion topics.

The guide was not shared in its entirety; instead, a broad outline of the topics was provided to the interviewees. This approach allowed them to understand the areas of focus without being constrained by specific questions. We aimed to avoid having interviewees adhere too rigidly to the guide to preserve the opportunity for exploring unexpected insights that might emerge during the conversation. Additionally, this strategy was intended to minimize bias, as advance knowledge of the questions could lead interviewees to tailor their responses in a manner they believe to be more favorable or relevant, potentially skewing the data.

Along with the interview invitation, participants were provided with instructions outlining the interview process. These instructions included an overview of the interview's purpose, format, and preparation guidelines, including a broad outline of the topics to be reviewed. Additionally, contact details for the interviewers and a confidentiality disclaimer were provided.

The interviews were conducted with two interviewers, each assigned clearly defined roles to prevent overlap. This approach enhanced the accuracy and depth of the data collected while also reducing bias through the cross-checking of interpretations and observations during and after the interviews.

Two interviews were conducted, each lasting approximately one hour, and both were recorded.

4.1.4 Data Recording and Transcription. Before the interviews began, the session structure was explained to the participants, and their consent was obtained for recording the meetings. Both interviews were conducted online via Microsoft Teams, which was also employed to record and transcribe the sessions. The resulting MP4 and VTT files were manually reviewed and edited to address transcription errors using Subtitle Edit[5], an open-source software designed for the creation, editing, and synchronization of subtitles, including VTT files. The version utilized was 4.0.6. Subsequently, the edited VTT file was processed with a custom Python script to merge successive subtitles from the same speaker and turn the transcription into a dialog format in a text file.

4.2 Performing Thematic Analysis

For this part of the study, an inductive thematic analysis was conducted to systematically analyse and interpret the interview data. This approach facilitates the identification of patterns, themes, and key concepts directly from the data, without the constraints of pre-established categories or theoretical frameworks. By centering the analysis on the content of the interviews and allowing themes to emerge naturally, this method maintains a strong connection to the participants' perspectives and lived experiences. The inductive process enables the identification of nuanced insights that are deeply rooted in the data, ultimately guiding the development of design specifications that align closely with the actual needs and communication requirements of the tool's intended users.

4.2.1 Familiarization with the Data. Prior to coding, the human coders engaged in a familiarization process with the full interview transcripts. This involved reading through the entire text, including both the interviewers' questions and the participants' responses, to gain a comprehensive understanding of the context and content of the interviews. Familiarization with the full transcript allowed the coders to better grasp the nuances of the interviewees' responses, the flow of the conversation, and the specific prompts that elicited certain responses.

After this familiarization process, the text corresponding to the interviewers was removed from the transcripts to focus the subsequent analysis solely on the content provided by the interviewees. This step ensured that the coders had a thorough understanding of the data before beginning the coding process, while also allowing for a targeted analysis of the participants' responses.

4.2.2 Generating Initial Codes. The transcriptions from both interviews were combined sequentially to facilitate a unified analysis. The combined data were then subjected to independent coding by two human coders using NVivo v14[6] and by an AI language model (GPT-4). Each coder independently analyzed the text to identify codes-discrete labels representing key ideas, concepts, or patterns

[5]https://www.nikse.dk/
[6]https://lumivero.com/products/nvivo/

within the data. The human coders provided context-sensitive interpretation and ensured nuanced understanding, while GPT-4 offered a different perspective and comprehensive processing, potentially identifying aspects that the human coders might have overlooked.

The initial codes generated by both human coders and GPT-4 were subsequently reviewed and refined by the human coders, who merged them into a consolidated list. This review process was critical for validating the AI-generated codes and ensuring that the final coding scheme accurately reflected the data's context and content. The involvement of multiple coders and the inclusion of AI in the initial coding phase were intended to enhance the reliability and depth of the analysis, with inter-coder agreement and systematic review procedures ensuring methodological rigor.

Notably, there was one instance where a code identified by GPT-4 was a misinterpretation of the interview content. However, this specific code was later deemed by the human coders to be a valuable concept worth incorporating into the design of the tool. This is the label about safeguarding.

After the initial coding, a cross-analysis was performed to compare the codes generated by each coder. This process involved identifying commonalities, discrepancies, and unique insights across the different coding outputs. A consensus was reached through discussion, resulting in the consolidation of the codes into a single comprehensive list. The final list of codes is as follows:

- **Core Vocabulary Focus**: The need to prioritize basic communication needs by focusing on core vocabulary (e.g., greetings, requests, rejections) to build a foundation for further communication development.
- **Communication Systems**: Emphasis on using Makaton as a primary tool but also considering other communication systems like Signalong and ensuring the tool can handle variations in sign language.
- **Communication Challenges and Support**: Addressing the challenges of signing inaccuracies, inconsistency, and the merging of different systems by providing tools like visual symbols, photos, and context-aware features to support accurate communication.
- **Cultural and Contextual Sensitivity**: The importance of incorporating cultural and religious vocabulary to make the tool relevant to diverse users and their specific communication needs.
- **Safeguarding and Vulnerability Awareness**: Incorporating features to recognize and flag potential safeguarding issues, especially for vulnerable users, such as those at risk of abuse.
- **Educational Integration and Platforms**: Exploring the potential of integrating the tool with educational platforms and existing visual communication resources.
- **Motivation for Communication**: Identifying and leveraging what motivates learners to communicate, such as personal needs, sensory tools, engaging activities, and visual stimuli.
- **Trial and Adaptation in Real Contexts**: The necessity of trialing the tool in practical settings (e.g., schools, youth centers) and adapting it based on user feedback and real-world use.
- **Sensory Processing Considerations**: Awareness of sensory processing differences and ensuring the tool can pause or adjust based on user behavior, avoiding overwhelming sensory input.
- **Technological Integration**: The potential for the tool to interact with technological devices like iPads and computers, allowing non-verbal users to sign and navigate content.
- **User-Centered Design and Flexibility**: Ensuring the tool is designed with the user's needs and contexts in mind, offering flexibility to adapt to different sign language systems and individual communication styles.
- **Consultation and Collaboration with Experts**: Highlighting the importance of consulting well-trained Makaton users, educators, and other stakeholders to ensure the tool effectively meets the needs of its users.
- **Scalability and Long-Term Vision**: Starting small with a basic, functional tool and planning for gradual expansion to include more features, systems, and cultural considerations.

To enhance the understanding of the data and assist the design team in visualizing the relationships between the final codes identified, a mind map was automatically generated using Mapify[7]. This mind map is illustrated in Figure 1.

4.2.3 Defining and Naming Themes. Through collaborative discussion and consensus-building, the team synthesised the individual codes into a set of overarching themes that encapsulate the core ideas from both interviews. The goal was to ensure that these final themes accurately represented the nuances of the data while providing a cohesive set of insights. The identified themes are as follows:

(1) **Foundational and Inclusive Communication**:
   - **Codes**: Core Vocabulary Focus, Communication Systems, Communication Challenges and Support
   - **Theme Summary**: This theme focuses on building a solid communication foundation by incorporating core vocabulary and ensuring the tool is inclusive of multiple communication systems like Makaton and Signalong. It also addresses the need to overcome communication barriers, such as inaccuracies in signing.

(2) **Cultural Sensitivity and Safeguarding:**
   - **Codes**: Cultural and Contextual Sensitivity, Safeguarding and Vulnerability Awareness
   - **Theme Summary**: This theme emphasizes the importance of adapting communication tools to fit cultural, religious, and contextual needs, while also integrating features that address safeguarding concerns for vulnerable users.

(3) **Educational Integration and Motivation**:
   - **Codes**: Educational Integration and Platforms, Motivation for Communication
   - **Theme Summary**: This theme explores the integration of the communication tool into educational settings and platforms. It also considers the motivational aspects of communication, focusing on what drives learners to engage in communication.

[7] https://mapify.so/

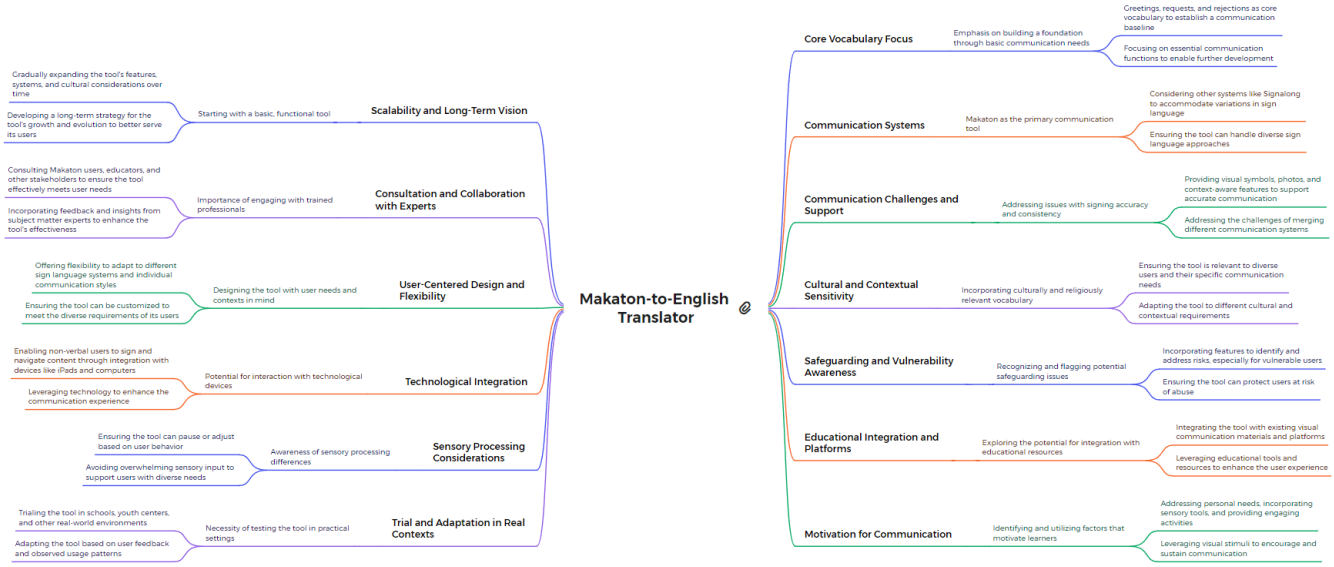

Figure 1: Mind Map of Initial Codes

(4) **Practical Implementation and User-Centered Design**:
- **Codes**: Trial and Adaptation in Real Contexts, Sensory Processing Considerations, User-Centered Design and Flexibility
- **Theme Summary**: This theme addresses the practical aspects of implementing the tool in real-world contexts, ensuring it is adaptable, user-centered, and sensitive to sensory processing differences.

(5) **Scalability and Future Vision**:
- **Codes**: Technological Integration, Scalability and Long-Term Vision, Consultation and Collaboration with Experts
- **Theme Summary**: This theme looks at the long-term vision for the communication tool, including its scalability, potential technological integrations, and the importance of consulting experts throughout its development.

4.2.4 Evaluation - Member Checking. To validate the themes derived from the interviews and ensure that the design specifications accurately reflect the needs of the intended users, we employed member checking [4, 8, 18, 22, 27] as a key validation strategy. Member checking, also known as participant validation, involves returning the initial findings to the original participants to solicit their feedback on the accuracy and relevance of the themes identified during the analysis.

In this study, after the thematic analysis of the interview transcriptions was completed, the resulting themes and corresponding design specifications were shared with the interview participants. Participants were provided with a summary of the key themes and were asked to review and comment on whether these themes accurately captured their experiences and needs as discussed during the interviews. They were also invited to suggest any modifications or additional considerations that might have been overlooked during the initial analysis.

The feedback obtained from this process played a crucial role in refining the design specifications for the tool. Participants generally affirmed the relevance of the identified themes, with some providing additional insights that were incorporated into the final design. This iterative process ensured that the final design specifications were not only rooted in the data but also resonated with the lived experiences and expectations of the end-users. The member checking process thus contributed to the credibility and trustworthiness of the research findings, enhancing the overall validity of the study.

4.2.5 Producing the Design Specifications. This report translates the final themes into design specifications by breaking down each theme into actionable guidelines and requirements for the development of the translation tool. These requirements are intended to serve as a comprehensive guide for the tool's development, ensuring it meets the diverse needs of its users and supports effective communication across different systems and contexts.

(1) **Core Vocabulary and Multi-System Integration**
- **Core Vocabulary List**: The tool must include a comprehensive core vocabulary list, covering essential daily interactions and needs (e.g., greetings, requests, rejections).
- **Multi-System Support**: The tool should support multiple communication systems (e.g., Makaton, Signalong, British Sign Language) and recognize variations in signs.
- **Accuracy Handling**: Implement functionality to handle inaccuracies and inconsistencies in signing. Provide real-time feedback and user correction options.

(2) **Cultural Sensitivity and Safeguarding**
- **Cultural and Religious Vocabulary**: Include customizable options for cultural and religious vocabulary to cater to diverse user backgrounds.
- **Safeguarding Features**: Incorporate mechanisms to detect and flag potential safeguarding issues, such as distress signals or abuse indicators.

- **Contextual Adaptation**: Ensure the tool can adapt to different communication contexts, including personal care, cultural practices, and social interactions.

(3) **Educational Integration and Motivation**
- **Educational Content**: Integrate the tool with educational platforms and resources. Consider partnerships with organizations that provide relevant content (e.g., Singing Hands UK).
- **Motivational Features**: Design engaging elements such as visual aids, interactive content, games, and sensory-friendly features to motivate learners.
- **Personal Needs Focus**: Address personal needs and interests of learners to enhance motivation, including food, activities, and sensory preferences.

(4) **Practical Implementation and User-Centered Design**
- **Real-World Application**: Ensure the tool is practical for use in various real-world settings such as schools, youth centers, and community environments.
- **Sensory Processing Considerations**: Include features that accommodate users with sensory processing differences, such as the ability to pause and adapt based on user input and sensory needs.
- **User-Friendly Interface**: Design an intuitive interface that allows easy customisation and adaptation by users and caregivers.

(5) **Scalability and Future Development**
- **Scalable Architecture**: Build the tool with a scalable architecture to support future updates and expansions.
- **Technological Integration**: Plan for the integration of advanced technologies (e.g., AI for recognizing diverse signing variations) to enhance the tool's functionality.
- **Expert Consultation**: Establish a framework for ongoing consultation with experts in communication systems, technology, and user experience to refine and improve the tool.

(6) **Error Handling and User Feedback**
- **Error Detection**: Implement error detection mechanisms to identify and correct issues with sign recognition and translation.
- **User Feedback System**: Provide options for users to give feedback on inaccuracies or issues, allowing continuous improvement of the tool.

(7) **Customisation and Adaptability**
- **Customisable Settings**: Allow users to customise settings based on their specific communication needs and preferences.
- **Adaptability for Different Users**: Ensure the tool can adapt to the needs of different users, including those with varying levels of communication skills and sensory sensitivities.

(8) **Resource Integration**
- **Additional Resources**: Integrate resources such as visual communication symbols, educational materials, and training guides to support users and caregivers.
- **Training and Support**: Provide access to training materials and support resources to help users and caregivers effectively utilise the tool.

5 CONCLUSIONS

This paper presents the design process for a tool intended to automatically translate Makaton signs from live video streams into written English. This study addresses a critical need for accessible communication solutions, particularly in today's digital era, which presents additional challenges and further marginalizes individuals with severe verbal communication difficulties.

The study employs a user-centered design methodology [28, 37], integrating several key components to ensure the development of an effective translation tool. It begins with semi-structured interviews with experts in the field, allowing for an in-depth exploration of their knowledge and insights regarding the needs of Makaton users. Thematic analysis is then used to systematically identify and interpret recurring themes and patterns from the interview data. To validate and refine these findings, member checking is utilised, involving expert feedback to ensure that the identified themes accurately reflect their experiences and needs. This iterative process culminates in the development of detailed design specifications, informed by both thematic analysis and expert feedback. This approach emphasises a comprehensive understanding of user requirements, ensuring that the design specifications are both practical and responsive to the needs of end-users.

A significant contribution of this paper is the innovative methodology employed. The study utilises a user-centered design approach to derive specifications for a tool that translates Makaton into written English in real-time. This methodology ensures that the tool is technologically advanced and deeply aligned with the needs and preferences of individuals with severe verbal communication difficulties. The real-time translation capabilities, combined with a rigorous user-centered design process, represent a significant advancement in the field of AAC, offering a more accessible and practical solution for those facing communication challenges.

The next phase of this research involves using the developed design specifications to create a functional prototype of the translation tool. This prototype will undergo iterative testing and refinement cycles, allowing for continuous feedback and adjustments to enhance its performance and usability. This iterative redesign process will ensure that the final tool meets the practical needs of its users and achieves its intended goals.

Looking ahead, we envision this solution serving as a valuable framework for the development of similar software tools within the field of AAC. By applying the principles and methodologies established in this study, future projects can benefit from a well-defined approach to designing and refining accessible communication technologies. This work not only advances Makaton translation tools but also provides a model for streamlining the design and development of other assistive technologies aimed at improving communication accessibility.

REFERENCES

[1] Ruth Huntley Bahr and Elaine R Silliman. 2015. *Routledge handbook of communication disorders.* Routledge Abingdon.
[2] David Beukelman and Janice Light. 2020. Augmentative and alternative communication: Supporting children and adults with complex communication needs. (2020).
[3] David R Beukelman, Pat Mirenda, et al. 1998. *Augmentative and alternative communication.* Paul H. Brookes Baltimore.
[4] Linda Birt, Suzanne Scott, Debbie Cavers, Christine Campbell, and Fiona Walter. 2016. Member checking: a tool to enhance trustworthiness or merely a nod to

validation? *Qualitative health research* 26, 13 (2016), 1802–1811.
[5] Lucienne T.M. Blessing and Aditi Chakrabarti. 2009. *DRM, a Design Research Methodology.* Springer. https://link.springer.com/book/10.1007/978-1-4419-0936-0
[6] Michelle M Byrne. 2001. Evaluating the findings of qualitative research. *AORN journal* 73, 3 (2001), 703–703.
[7] Marjorie H Charlop-Christy, Michael Carpenter, Loc Le, Linda A LeBlanc, and Kristen Kellet. 2002. Using the picture exchange communication system (PECS) with children with autism: Assessment of PECS acquisition, speech, social-communicative behavior, and problem behavior. *Journal of applied behavior analysis* 35, 3 (2002), 213–231.
[8] John W Creswell and Dana L Miller. 2000. Determining validity in qualitative inquiry. *Theory into practice* 39, 3 (2000), 124–130.
[9] Jack S Damico, Nicole Müller, and Martin John Ball. 2010. *The handbook of language and speech disorders.* Wiley Online Library.
[10] Ioannis Deliyannis, Christina Simpsiri, and Plateia Tsirigoti. 2008. Interactive multimedia learning for children with communication difficulties using the Makaton Method. In *International Conference on Information Communication Technologies in Education.* 10–12.
[11] Eleni Efthimiou, Stavroula-Evita Fotinea, Thomas Hanke, John R. W. Glauert, R. Bowden, Annelies Braffort, Christophe Collet, Petros Maragos, and François Lefebvre-Albaret. 2012. Sign Language technologies and resources of the Dicta-Sign project. https://api.semanticscholar.org/CorpusID:54101032
[12] Eleni Efthimiou, Galini Sapountzaki, Kostas Karpouzis, and Stavroula-Evita Fotinea. 2004. Developing an e-Learning Platform for the Greek Sign Language. In *International Conference on Computers for Handicapped Persons.* https://api.semanticscholar.org/CorpusID:1276842
[13] Karen Emmorey. 2001. *Language, cognition, and the brain: Insights from sign language research.* Psychology Press.
[14] Susan Fager, Lisa Bardach, Susanne Russell, and Jeff Higginbotham. 2012. Access to augmentative and alternative communication: New technologies and clinical decision-making. *Journal of pediatric rehabilitation medicine* 5, 1 (2012), 53–61.
[15] Michelle Flippin, Stephanie Reszka, and Linda R Watson. 2010. Effectiveness of the Picture Exchange Communication System (PECS) on communication and speech for children with autism spectrum disorders: A meta-analysis. (2010).
[16] Judy Graves. 2000. Vocabulary needs in augmentative and alternative communication: A sample of conversational topics between staff providing services to adults with learning difficulties and their service users. *British Journal of Learning Disabilities* 28, 3 (2000), 113–119.
[17] Nicola Grove and Margaret Walker. 1990. The Makaton Vocabulary: Using manual signs and graphic symbols to develop interpersonal communication. *Augmentative and Alternative Communication* 6, 1 (1990), 15–28.
[18] Melissa Harper and Patricia Cole. 2012. Member checking: Can benefits be gained similar to group therapy. *The qualitative report* 17, 2 (2012), 510–517.
[19] D Jeffery Higginbotham, Howard Shane, Susanne Russell, and Kevin Caves. 2007. Access to AAC: Present, past, and future. *Augmentative and alternative communication* 23, 3 (2007), 243–257.
[20] Patricia Howlin, R Kate Gordon, Greg Pasco, Angie Wade, and Tony Charman. 2007. The effectiveness of Picture Exchange Communication System (PECS) training for teachers of children with autism: a pragmatic, group randomised controlled trial. *Journal of child Psychology and Psychiatry* 48, 5 (2007), 473–481.
[21] Jeanne M Johnson, Ella Inglebret, Carla Jones, and Jayanti Ray. 2006. Perspectives of speech language pathologists regarding success versus abandonment of AAC. *Augmentative and Alternative Communication* 22, 2 (2006), 85–99.
[22] Lori E Koelsch. 2013. Reconceptualizing the member check interview. *International journal of qualitative methods* 12, 1 (2013), 168–179.
[23] Rakesh Kumar Attar, Vishal Goyal, and Lalit Goyal. 2023. State of the Art of Automation in Sign Language: A Systematic Review. *ACM Transactions on Asian and Low-Resource Language Information Processing* 22, 4 (April 2023), 1–80. https://doi.org/10.1145/3564769
[24] Russell Lang, Laurie McLay, and Mandy Rispoli. 2023. Advanced language models: Potential to improve augmentative and alternative communication for individuals with intellectual and developmental disabilities. *Advances in Neurodevelopmental Disorders* 7, 3 (2023), 481–484.
[25] James Law, Sheena Reilly, and Pamela C Snow. 2013. Child speech, language and communication need re-examined in a public health context: a new direction for the speech and language therapy profession. *International Journal of Language & Communication Disorders* 48, 5 (2013), 486–496.
[26] Janice Light and David McNaughton. 2014. Communicative competence for individuals who require augmentative and alternative communication: A new definition for a new era of communication? , 18 pages.
[27] Yvonna S Lincoln. 1985. *Naturalistic inquiry.* Vol. 75. Sage.
[28] Martin Maguire. 2001. Methods to support human-centred design. *International journal of human-computer studies* 55, 4 (2001), 587–634.
[29] Sotiris Malassiotis, Niki Aifanti, and Michael G Strintzis. 2002. A gesture recognition system using 3D data. In *Proceedings. First International Symposium on 3D Data Processing Visualization and Transmission.* IEEE, 190–193.
[30] Jane McCormack, Linda J Harrison, Sharynne McLeod, and Lindy McAllister. 2011. A nationally representative study of the association between communication impairment at 4–5 years and children's life activities at 7–9 years. (2011).
[31] Sharan B. Merriam. 2009. *Qualitative Research: A Guide to Design and Implementation.* Jossey-Bass.
[32] Diane C Millar, Janice C Light, and Ralf W Schlosser. 2006. The impact of augmentative and alternative communication intervention on the speech production of individuals with developmental disabilities: A research review. (2006).
[33] Mara Mills. 2011. On disability and cybernetics: Helen Keller, Norbert Wiener, and the hearing glove. *differences* 22, 2-3 (2011), 74–111.
[34] Janice M Morse, Michael Barrett, Maria Mayan, Karin Olson, and Jude Spiers. 2002. Verification strategies for establishing reliability and validity in qualitative research. *International journal of qualitative methods* 1, 2 (2002), 13–22.
[35] Janice M Morse and Carl Mitcham. 2002. Exploring qualitatively-derived concepts: Inductive—deductive pitfalls. *International journal of qualitative methods* 1, 4 (2002), 28–35.
[36] Minh Hao Nguyen, Jonathan Gruber, Jaelle Fuchs, Will Marler, Amanda Hunsaker, and Eszter Hargittai. 2020. <? covid19?> Changes in Digital Communication During the COVID-19 Global Pandemic: Implications for Digital Inequality and Future Research. *Social media+ society* 6, 3 (2020), 2056305120948255.
[37] Don Norman. 2013. *The design of everyday things: Revised and expanded edition.* Basic books.
[38] James Ohene-Djan, Robert Zimmer, James Bassett-Cross, Andrew Mould, and Ben Cosh. 2004. Mak-Messenger and Finger-Chat, communications technologies to assist in teaching of signed languages to the deaf and hearing. In *IEEE International Conference on Advanced Learning Technologies, 2004. Proceedings.* IEEE, 744–746.
[39] Sylvie CW Ong and Surendra Ranganath. 2005. Automatic sign language analysis: A survey and the future beyond lexical meaning. *IEEE Transactions on Pattern Analysis & Machine Intelligence* 27, 06 (2005), 873–891.
[40] Parimala Raghavendra, Lareen Newman, Emma Grace, and Denise Wood. 2013. 'I could never do that before': effectiveness of a tailored I nternet support intervention to increase the social participation of youth with disabilities. *Child: care, health and development* 39, 4 (2013), 552–561.
[41] Robert J Ruben. 2000. Redefining the survival of the fittest: communication disorders in the 21st century. *The Laryngoscope* 110, 2 (2000), 241–241.
[42] Stephanie Stoll, Necati Cihan Camgöz, Simon Hadfield, and R. Bowden. 2018. Sign Language Production using Neural Machine Translation and Generative Adversarial Networks. In *British Machine Vision Conference.* https://api.semanticscholar.org/CorpusID:52288950
[43] David J Sturman and David Zeltzer. 1994. A survey of glove-based input. *IEEE Computer graphics and Applications* 14, 1 (1994), 30–39.
[44] Rachel Sutton-Spence and Bencie Woll. 1999. *The linguistics of British Sign Language: an introduction.* Cambridge University Press.
[45] Ruth Swanwick. 2010. Policy and practice in sign bilingual education: development, challenges and directions. *International Journal of Bilingual Education and Bilingualism* 13, 2 (2010), 147–158.
[46] Nobuhiko Tanibata, Nobutaka Shimada, and Yoshiaki Shirai. 2002. Extraction of hand features for recognition of sign language words. In *International conference on vision interface.* 391–398.
[47] Matt Tincani and Kathryn Devis. 2011. Quantitative synthesis and component analysis of single-participant studies on the Picture Exchange Communication System. *Remedial and Special Education* 32, 6 (2011), 458–470.
[48] Anthony G Tuckett. 2005. Applying thematic analysis theory to practice: A researcher's experience. *Contemporary nurse* 19, 1-2 (2005), 75–87.
[49] Althea Z Valentine, Sophie S Hall, Emma Young, Beverley J Brown, Madeleine J Groom, Chris Hollis, and Charlotte L Hall. 2021. Implementation of telehealth services to assess, monitor, and treat neurodevelopmental disorders: systematic review. *Journal of Medical Internet Research* 23, 1 (2021), e22619.
[50] Tony Veale, Alan Conway, and Bróna Collins. 1998. The Challenges of Cross-Modal Translation: English-to-Sign-Language Translation in the Zardoz System. *Machine Translation* 13 (1998), 81–106. https://api.semanticscholar.org/CorpusID:3901531