COMPA: Using Conversation Context to Achieve Common Ground in AAC

Stephanie Valencia
University of Maryland
sval@umd.edu

Jessica Huynh
Carnegie Mellon University
jhuynh@cs.cmu.edu

Emma Y Jiang
Carnegie Mellon University
eyjiang@andrew.cmu.edu

Yufei Wu
Carnegie Mellon University
yufeiwu2@andrew.cmu.edu

Teresa Wan
Carnegie Mellon University
xinyuewa@andrew.cmu.edu

Zixuan Zheng
Carnegie Mellon University
zzheng3@andrew.cmu.edu

Henny Admoni
Carnegie Mellon University
henny@cmu.edu

Jeffrey P Bigham
Carnegie Mellon University
jbigham@cs.cmu.edu

Amy Pavel
The University of Texas at Austin
apavel@cs.utexas.edu

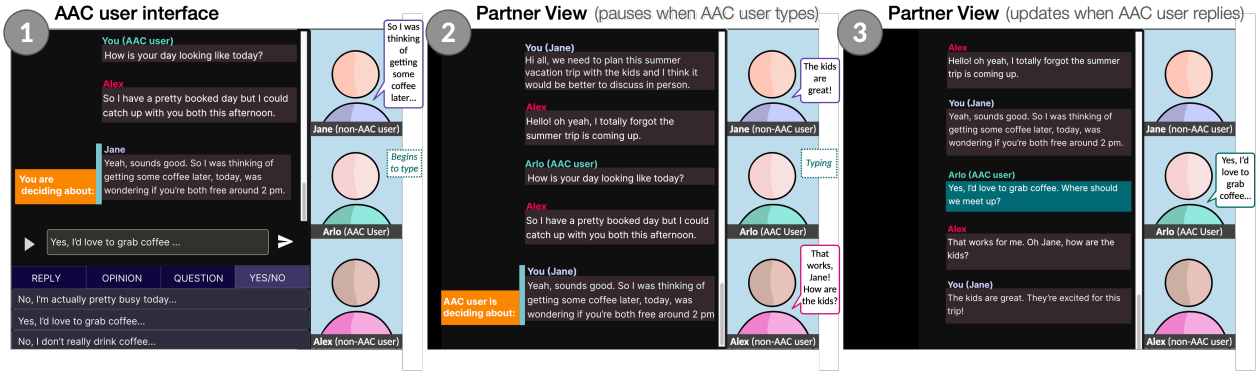

Figure 1: While AAC users type, co-occurring talk by other conversation participants may change the topic of the conversation faster than AAC users can respond. COMPA is an AAC browser extension that can be used side-by-side a video call. Using COMPA, AAC users and non-AAC communication partners (CPs) can view live captions and notifications during their conversation. (1) AAC users can use COMPA to type and receive contextual phrase starters. (2) While an AAC user is typing a reply, CPs see a paused snapshot of the conversation. Even if CPs talk, the live transcript won't update while an AAC user is replying. (3) When an AAC user shares their message, it is read out loud and printed below it's intended context. The live transcript resumes, including speech that happened while the AAC user typed.

ABSTRACT

Group conversations often shift quickly from topic to topic, leaving a small window of time for participants to contribute. AAC users often miss this window due to the speed asymmetry between using speech and using AAC devices. AAC users may take over a minute longer to contribute, and this speed difference can cause mismatches between the ongoing conversation and the AAC user's response. This results in misunderstandings and missed opportunities to participate. We present COMPA, an add-on tool for online group conversations that seeks to support conversation partners in achieving common ground. COMPA uses a conversation's live transcription to enable AAC users to mark conversation segments they intend to address (Context Marking) and generate contextual starter phrases related to the marked conversation segment (Phrase Assistance) and a selected user intent. We study COMPA in 5 different triadic group conversations, each composed by a researcher, an AAC user and a conversation partner (n=10) and share findings on how conversational context supports conversation partners in achieving common ground.

CHI '24, May 11–16, 2024, Honolulu, HI, USA

ACM ISBN 979-8-4007-0330-0/24/05.
https://doi.org/10.1145/3613904.3642762

CCS CONCEPTS

• **Human-centered computing**; • **Accessibility systems and tools**;

KEYWORDS

AAC, accessibility, communication, language models

ACM Reference Format:

Stephanie Valencia, Jessica Huynh, Emma Y Jiang, Yufei Wu, Teresa Wan, Zixuan Zheng, Henny Admoni, Jeffrey P Bigham, and Amy Pavel. 2024. COMPA: Using Conversation Context to Achieve Common Ground in AAC. In *Proceedings of the CHI Conference on Human Factors in Computing Systems (CHI '24), May 11–16, 2024, Honolulu, HI, USA.* ACM, New York, NY, USA, 21 pages. https://doi.org/10.1145/3613904.3642762

1 INTRODUCTION

Speech-generating Augmentative and Alternative Communication (AAC) devices are used by individuals to enhance or substitute their vocal communication. AAC users' comments in conversation can often be over a minute delayed as AAC users need to input text into their device before it is synthesized into speech. Therefore, speaking rates for AAC users are slower than people that use speech to talk. Asymmetrical speeds of communication between AAC and their conversation partners (CPs) can make conversations challenging as being able to reply to a topic while it is still relevant requires the correct timing. This time asymmetry can cause misunderstandings or make an AAC user's comments appear out-of-context or be hard to interpret [14, 32]. It has also been reported that AAC users initiate conversation less and appear more passive [21, 52]. Additionally, AAC users often miss opportunities to participate in group conversations due to fast topic changes [31, 49].

One method of addressing time asymmetry in conversation is to create common ground. There are multiple resources people draw from in conversation to avoid misunderstandings and maintain common ground [6]. Individuals can add reference words, longer descriptions, or connecting ideas to their utterances to make them more specific and clear to their CPs. While adding more words is easy for people who use speech, it requires more effort and time for AAC users and therefore creates an even more delayed response. As such, it is common for AAC users' responses to be shorter and lack contextual references in real-time [11]. Additionally, in group conversations with many speakers, topics change quickly and misunderstandings and ambiguous messages can be harder to resolve [13, 34, 35].

To facilitate AAC user participation and mutual understanding, prior work has explored context-aware predictive models [9, 37], re-using a CP's noun phrases [52], a CP's collaborative input through a companion app [10], and other tools such as visual cues, moving objects, and typing status indicators [10, 39, 50]. These predictive models or knowledgeable communication partners cannot always exactly anticipate what an AAC user wants to say since there can be an expansive set of replies to a conversation [48]. However, little work has explored how to improve mutual understanding of what the current conversation context is [7].

We conducted two formative studies to identify clear design goals for a system that could support group conversations. The first study focused on analyzing two group conversations between an AAC user, a CP, and an interviewer. This first analysis of two conversations suggested that the time at which AAC users start typing can indicate the portion of the conversation they are referring to. In addition, while AAC users occasionally provided ambiguous responses, conversational partners who knew the AAC user well were able to clarify AAC user responses to the interviewer. We used these initial results to conduct a participatory design workshop to gather a wider understanding of the challenges and opportunities around establishing mutual understanding. Multiple stakeholders participated in this two day workshop including two AAC users, three technologists and three communication scientists. An idea that emerged from the workshop included having a visualization of the conversation's transcript real-time, finding ways to cue partners about a user's intent, and in line with prior work, increase partners' awareness of an AAC user's typing status [39].

We introduce COMPA, an add-on tool for AAC and non-AAC users that creates a shared context of the current conversation to support AAC users and their conversation partners participating in remote conversations. COMPA transcribes the conversation in real-time, then enables AAC users to pause the transcript to respond to a conversation topic. COMPA marks the conversation comment (*i.e.*, **Context Marking**) that the AAC user aims to respond to (*e.g.*, *"I would like to adopt a dog."*) such that conversation partners can later understand the context of an AAC user's response (*e.g.*, *"You should!"*) even if the conversation has moved on. COMPA further uses the real-time transcript to create context-based starter phrases (i.e. **Phrase Assistance**) that the AAC user can use to quickly ground their response in the prior conversation (*e.g.*, *"About adopting a dog,..."*). Inspired by our formative study, COMPA also enables AAC users to use **intent** to further signal to partners how they aim to respond, and guide the phrase assistance options. To explore how AAC users could use COMPA in conversation, we implemented COMPA as a browser extension to enhance remote conversations as an initial use case.

We evaluated COMPA in triadic conversations between a researcher and pairs of AAC and non-AAC CPs. Participants joined the researcher for an online video call to talk about three distinct themes related to planning a summer vacation (as in Figure 1). All participants experienced three COMPA

versions, each with different features. We collected AAC users and their CPs perceptions and observable participation metrics. Most AAC users and CPs expressed that they wanted to use some version of COMPA in future conversations. AAC users expressed that COMPA leveled the playing field and provided structure to the conversation and could particularly be useful when talking with partners unfamiliar to AAC. CPs expressed that COMPA helped with the conversation flow and to orient them to what their AAC partner wanted to say.

This work contributes:

- Design opportunities to support common ground in group conversations derived from our formative conversation analysis and participatory workshop.
- COMPA, an add-on AAC system to support group conversation by using the conversation context.
- A user study that demonstrates how context marking and phrase assistance can be useful for AAC and non-AAC conversation partners.

2 BACKGROUND AND RELATED WORK

When AAC users are typing, co-occurring discussion by other conversation participants (CPs) may change the topic of the conversation faster than what AAC users can respond to. As a result, AAC users may be discouraged from participating or their responses may be out-of-context and thus misunderstood. In this section we outline the current known challenges related to speaking-in-time in AAC-based and multi-party conversations, and discuss techniques that have been proposed to alleviate the out-of-context problem.

Challenges with Context and Speaking in Time

Albert Robillard, Professor of Sociology and an AAC user, explained the out-of-context problem he experienced as a patient in an intensive care unit [32]. Dr. Robillard was often cut-off mid-sentence when someone left his room before he had finished his message and when he wanted to resume speaking he experienced the following: *"Usually the interaction has moved along so far that when I address an old topic my conversants have a hard time seeing the relevance of what I am saying. It takes so much effort to spell out what I am saying I could not easily recycle the topic by saying "You know what we were speaking about a little while ago, the X topic." I could only, because of time and energy, speak directly to a former topic. The speaking out of context would generate many complaints and confusion."*

Grounding in conversation is the process of establishing mutual understanding [6]. One way CPs achieve grounding is by clearly establishing the entity they are referring to, then further elaborating (ex. *"The dog, he just bit me"*) [6]. However, using such references comes with a cost of additional words, which can worsen the out-of-context problem due to the slower communication speeds of AAC users [6, 36, 49]. In Dr. Robillard's words: *"...I did not have the temporal dimension to say, "You know what we were talking about before," as a method of reintroducing a topic I was talking about."*

The cost of grounding is dependent on the medium of communication [6] and little is known about how specific AAC interface designs may impact establishing a common ground or mutual understanding among speakers. In this work we propose different ways to support establishing common ground despite time differences in the response time among AAC and CPs.

Current Technical and Social Solutions

In order to achieve mutual understanding, conversation partners must be engaged with each other's contributions. Having a CP understand what an AAC user is trying to say is a large part of communication; in a study with children, children noted that they "liked when their communication partners tried to figure out what they wanted to say" and that premature topic change is an issue [27]. The AAC user must also figure out if their CP is understanding what they are saying, and adjust their efforts accordingly or modify their talking style based on the situation that they are in [51]. Close CPs who know each other well, seem to predict or assume what AAC users are trying to say, while unfamiliar CPs may ask more yes and no questions which lead to participation asymmetry, but this behavior could change in multi-party conversations [47]. CPs can learn how to more effectively communicate with AAC users through instruction, and this instruction can benefit communication [22]. These results encourage efforts in helping CPs be better CPs and in providing assistive tools that will help both the AAC user and the CP to facilitate conversation.

There are several aspects of conversation such as turn-taking and conversational context that CPs speaking with AAC users can be made more aware of [39, 43, 44]. Several systems have been made to improve CP awareness; an early AAC system, Lightwriter used a partner-facing screen to support communication and AACrobat [10] used a communication partner companion app to share the AAC users' typing status and communication preferences. When CPs have access to the AAC user's entire interface, they are able to support the AAC user in expanding their messages or anticipating what they want to say; however, there is the issue of AAC user privacy and autonomy. Therefore, there has been work with status indicators to still relay information to a CP while still preserving privacy [10]. Other indicators include visual feedback cues [39]. In this work we explore raising a partner's awareness about the topics that have been addressed in a conversation, when an AAC user wants to participate, what context they are speaking to, and what possible action they are intending to take. All of these are

important aspects that support mutual understanding but have not yet been explored in AAC systems.

Assistance tools for Group Conversations

Prior work has studied and designed systems to improve remote conversations as this medium provides the ability for more convenient additional visualizations. Basic visualizations can consist of transcripts or summaries [40] or word clouds [19], which can be viewed on the same interface the conversation is already taking place on. These basic visualizations may not be enough, as participants in one study noted that they wanted an indication of social cues (camera on) and 78% of participants indicated that these social cues were crucial in this setting [33]. Visualizations can help participants contextualize the conversation and stay in time [40] and functionalities such as being able to pause a transcript [23, 26] have appeared to be useful.

Although visualization can bring a lot to a conversation, there are also some potential downsides. Feedback during a conversation can be distracting [41]. Further work needs to be done on the best way to use and represent social cues [39]. Topic shift is also another large issue during meetings; participants can be confused by constant topic shifts even if word clouds [19] or other visualizations are created [5]. At the same time, speed at which a conversation develops has also been found to make a conversation less accessible for people who are deaf or hard of hearing and use sign language [26]. In this work we explore how different ways to visualize and pause conversation transcripts impact conversation between AAC users and CPs.

Conversation Support through Predictive Models

Various context-aware sentence prediction strategies have been explored to increase the speaking rate with AAC devices [4, 38, 45]. Some systems like AACrobat allow CPs to recommend word completions in real-time [10]. Utterance-based systems utilize pre-set responses that can be easily accessed [25, 42]. Automated systems can predict possible responses in conversation; however, since dialogue is variable, a small percentage of responses can be predicted successfully [28]. KWickChat adds the additional information of a person's persona to generate responses [38]. However, it is still difficult to generate a limited set of options that will contain a response that is exactly what an AAC user wishes to say [48]. In this work we explore how an AI-based Large Language model can assist AAC users in retrieving starter phrases they can use in conversation in real-time to address different topics in conversation.

Prior work has found that when using AI composition tools, there should be functionalities in place for AAC users to edit pre-created phrases such that it fits their style [48]. However, if there is too much editing, it may be easier for the AAC user type the input themselves. There also seems to be an upper bound to saving keystrokes [46]. These works suggest there should be more focus on providing various styles of speaking rather than predicting large parts of future text. As such, we study how language models can be used to facilitate the addition of context to AAC users' utterances, instead of replacing them completely. In particular we explore how providing short starter phrase suggestions (maximum 6 words), that are open-ended and related to the current conversation and the AAC user's intended action (*i.e.*, reply, ask a question, give an opinion) compare to AI suggestions that are only fitted to help reply to a specific conversation.

3 FORMATIVE WORK

To better understand breakdowns in common ground during conversation and discover opportunities to address them we carried out two formative studies. First, we analyzed how AAC users and a close communication partner maintained common ground while being interviewed by an unfamiliar researcher. We then carried out a two-day virtual participatory design workshop to explore the design space where language-driven technologies could help support communication challenges. We recruited participants for both studies using organization mailing lists and social media interest groups. Both studies were approved by our institution's IRB.

Conversation Interaction Analysis

Prior literature indicates people utilize references and reuse each other's words to achieve common ground [6, 35], we sought to understand how AAC and non-AAC communication partners collaborate with each other to maintain common ground despite speaking rate asymmetries. One author carried out two interviews, each with an adult expert AAC user and a CP of their choice to collect group conversations that included both AAC and non-AAC users.

Participants. Four participants took part in the interviews: two adult expert AAC users (A1 and A2) and two conversational partners (C1 and C2). C1 was A1's mother and C2 was A2's father. Both AAC users used a Tobii Dynavox with switch scanning, and A2 additionally used a joystick to access the device. A1 had 20 years of experiencing using an AAC device, and A2 had 32 years of experiencing using an AAC device. AAC users only had disabilities associated with speech production, and did not have disabilities associated with language use or processing.

Method. Video and audio recordings were collected. One interview lasted 28 minutes, the second was 38 minutes long, and each touched upon general topics about the AAC user

and their partner's background and communication experiences. Our analysis was guided by the following three questions: (1) Could we identify the relevant conversation context an AAC user is addressing in a conversation? (2) if identified, where is this relevant context located in time and in relationship to the AAC user's turn? (3) how do AAC and non-AAC users collaborate with each other to solve misunderstandings? Two researchers analyzed the data by annotating the frequency of each speaking event and then transcribing each turn taken by the AAC user, the CP, and the interviewer. Researchers manually labelled all instances when the AAC user communicated using their AAC device. For each AAC user response, the researchers annotated the relevant context in the prior conversation that the AAC user was responding to. We annotated 34 AAC user verbal responses (*e.g.*, speaking turns in which the AAC user used their device to communicate) in total along with each response's corresponding conversation context. We also annotated the start and end time of any typing event by the AAC user to calculate how the typing start time related to the relevant conversational context.

Findings. Our analysis suggested that the conversation context an AAC user wants to speak to is related to their typing activity. In 27 out of 34 instances, the conversation context an AAC user was responding to was located in the turn before they started typing. In these cases, AAC users started typing a message after another speaker finished their turn and other speakers waited for the AAC user to finish typing and sharing their message before inserting new topics into the conversation. Due to our interview format, many AAC responses were connected to questions but not every AAC user participation in the interview was a direct answer.

In four separate instances we observed that the conversation's context overlapped with the AAC user's typing activity, as other spoke while they typed or an AAC user started typing while another speaker shared a long comment or long question. In two separate occasions, AAC users started typing right after the conversational context they wanted to reply was mentioned but then decided to erase their comment due to others changing the conversation topic while the AAC user was typing. Finally in an additional separate ocassion, an AAC user typed and spoke their messages in separate turns, elaborating their ideas through separate short sentences. Often other speakers spoke in between these turns. When AAC users spoke in separate turns, the relevant conversation context was contained in their prior utterances and not in their partner's speaking turns.

In the conversations we analyzed, CPs knew the AAC users well so they could often clarify messages for the AAC user when their messages were delayed or were ambiguous responses, such as *"yes"* after many comments in the group had been shared. Partners often guessed what the AAC user meant to say but other times they needed to confirm with the AAC user that their clarification was correct. This finding demonstrates that CPs can better support AAC users' communication when they have more context about what an AAC user intends to say.

While our conversation interaction analysis is limited in the sense that only two cases were thoroughly examined, we brought these examples and observations to a workshop with AAC experts to gather a wider understanding of the challenges around establishing mutual understanding. Our findings are based on small amounts of data and a bigger data exploration is needed to understand if these behaviors are common.

Participatory Ideation Workshop

Methods. We designed a fully remote participatory workshop with two AAC users (A3 and A4), three technologists and three communication researchers. An online workshop using Zoom made our study more accessible to all our participants who were located across the United States. A3 has 35+ years of experience using AAC devices, uses multiple devices depending on the scenario, and accesses all devices with direct touch input. For the workshop, A3 used the PRC Touch Talker with Minspeak. A4 has 20+ years of experience using AAC devices and uses the Tobii Dynavox with eye tracking. AAC users only had disabilities associated with speech production, and did not have disabilities associated with language use or processing. The workshop was split into two sessions, one per day. On day one, participants shared challenges related to communication breakdowns and staying in time and brainstormed potential solutions through three writing exercises. Writing exercises involved using an shared virtual board to post pressing challenges of interests and respond to three prompts: *imagine a technology that (1) helps you behind the scenes, (2) helps your communication partner be a better partner, and (3) is proactive and responsive.* We chose these prompts to elicit conversation and thoughts from participants based on opportunities for technology to support AAC user's agency, suggested in prior work [49]. Participants used the generated insights for discussion and were encouraged to comment on and up-vote each other's ideas during and after the first session.

Between the sessions, two authors visually summarized the challenges from the online discussion board by themes and labelled each individual idea (see Figure 2). On day two, workshop participants used the generated summary and a graph (Figure 2) to discuss challenges and ideas and further ideate specific solutions. A3 and A4 were asked to identify challenge areas to prioritize and share how they could or could not add value. All participants were then asked to

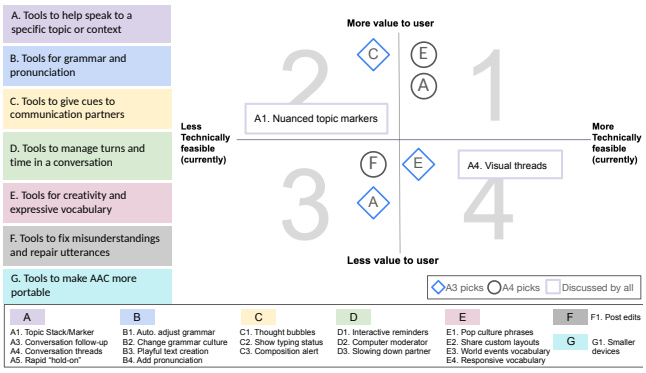

Figure 2: On the second day of the workshop, we employed a slide that condensed all the ideas put forth on the initial day. Utilizing a graph featuring two dimensions—value to the user and technical feasibility—we encouraged discussion and information sharing. This approach facilitated the establishment of a common understanding of our conversations and provided a convenient reference for AAC users to navigate challenges and solutions. Using the slide as a visual aid, we tasked A3 and A4 with ranking ideas on a spectrum ranging from "will add more value" to "will add less value." Following their input, all workshop participants engaged in a collective discussion, exploring how certain ideas might be more technically feasible than others.

reflect on what ideas were the most technically feasible. Afterwards, participants were split into two teams. Each team included one AAC user, and at least one technologist and one communication researcher. Teams were invited to create a shareable slide describing a proposed solution, what need it solves, what the improved experience would be, and how the idea could go wrong. We provided these dimensions to facilitate ideation. Video and audio recordings were collected during the two-days of the workshop. Data collected was analyzed through inductive qualitative research, moving from specific observations of topics discussed during the workshop to broader themes identified through thematic analysis [2].

Findings. AAC users and other workshop participants explained how improved AAC systems should be (1) conversation-aware to assist with message composition and (2) give conversation partners more feedback and context clues about what an AAC user's intended action towards a topic is. More importantly, AAC users highlighted how individuals unfamiliar with AAC could be trained to be better partners through visual cues that remind them to wait as a change in partner behavior could go a long way in making their group conversations better.

Conversation-awareness. Participant A3 suggested having the AAC device's interface dynamically changed based on questions they get asked. The participant envisioned their own AAC user interface adapting to the current conversation and update itself to support them in easily retrieving the most relevant responses suitable for the situation. A3 explained how an AAC system that is aware of the current conversation could help an AAC user navigate to their vocabulary pages faster: *"if you ask AAC user a question. Hey, how's it goin'? a device change display screen itself. Changes to feels area..."*

Partner Feedback and context clues. Workshop participants also shared ideas around monitoring the conversation turns and using this information to remind partners to pause (see Figure 2). This could be either through programmed interactive reminders or a "computer moderator" that could act as a referee or mediator between non-AAC and AAC participants, and would be in the "look-out" for cues a partner misses. Other ideas involved a topic marker that would visually establish when a current topic wanted to be revisited by an AAC user or having a one-word reminder of what the topic of conversation was once the AAC user started typing. We discussed how it would be technically challenge to distinguish nuances among topics, but perhaps visualizations of conversation threads could be easier to achieve.

Building on these discussions, communication researchers explained that misunderstandings often arise in conversation not because the topic is unknown but because it is hard to link the topic to an intended action: *"sometimes when we know the [communicative] action we do not need to know the topic, what matters is action and people can do action with different words. Sometimes it's more important to know what the person was doing than what it was said."* Therefore, knowing the intention or stance a person has towards a topic can help others better understand their contributions in conversation.

4 DESIGN GOALS

Our two formative studies indicated possible directions to support three goals (**G1-G3**):

G1: Create a Shared Conversational Context. Communication partners who know the AAC user well, familiar CPs, draw from their shared experiences to solve misunderstandings efficiently together. As group conversations might include both familiar and unfamiliar CPs, shared experience is not always present and there needs to be a new way to resolve misunderstandings. A shared conversation context that all group participants have access to could alleviate this issue, which is an important design goal we consider. Additionally, we consider that the shared conversational context should be responsive to what is happening in a real-time conversation, inspired by A3 who, in the workshop, highlighted the need for AAC devices to adapt to better support

them in communicating in a timely manner.

G2: Provide Awareness of AAC User Intent. Through our formative ideation workshop, we learned that having a way for the AAC user to specify their stance or intention towards a topic can help their CPs provide better support and explanations on their behalf when needed, or even improve the word and phrase suggestions the device provides. Additionally, CPs can benefit from more signaling [39] to be reminded of when and how AAC users want to participate. COMPA needs to provide awareness of when the user wants to participate and provide a way for users to specify how they wish to participate. Making an AAC users intention salient could also create more specific cues for their conversation partners.

G3: Support AAC User & Partner Initiated Alignment. CPs who know the AAC user well often act as mediators between unfamiliar partners and AAC users [1, 49]. Similarly, as mentioned in the workshop, conversation would greatly improve if non-AAC users were patient, learned to wait and slowed down. Tools designed to support mutual understanding between AAC and non-AAC users should be shared, similar to groupware explored in prior work [10]. Our third design goal is to create AAC and partner interfaces that can support their mutual adjustment and learning process to communicate with each other.

5 COMPA

COMPA is a system that supports creating a shared understanding of conversational context and awareness of AAC user intent (Figure 1). To support shared conversational context, COMPA displays a **real-time transcript** of the conversation, which automatically **pauses** the shared transcript when the AAC user starts to type their response. COMPA **notifies** both the AAC user and the conversation partner (CP) when the transcript has been paused. To support awareness of AAC user intent, AAC users can select an **intent** (reply, opinion, question, yes/no answer), which is added to the pause notification. The transcript and intent are additionally leveraged to create **starter phrases** based on the current conversational context to let the AAC user easily add context to their comment.

COMPA was realized as a Chrome browser extension to support AAC users and their CPs in remote meetings. Many AAC devices are used to directly browse the web or interface via Infrared switches with computers. We developed COMPA following keyboard accessibility guidelines (*e.g.*, tab navigable) and in dark mode to reduce eye strain.

Remote meetings using Google Meet were selected as COMPA's first use case due to the availability of high-quality real-time audio streams from computer microphones to enable transcription, and built-in access to a shared screen to support AAC user and partner focused conversation support.

Interfaces

COMPA has an **AAC interface** and a **CP interface** (Figure 1). The AAC interface includes both the live transcript of the current conversation on the transcript panel and a text composition panel that enables the AAC user to compose a message and share the message out loud using text-to-speech. The CP interface includes the automated live transcript of the current conversation which pauses when an AAC user is typing and updates after the AAC user shares their message.

Transcript Panel. The transcript panel displays the conversation speakers and live transcriptions of the ongoing dialogue captured using automatic speech recognition. The AAC user can pause the transcript panel by starting to type or by pressing a pause button when they want to speak. After they finish speaking, the tool will insert their utterance at the point in the conversation where they paused, and any conversation that has happened since then will populate below their utterance. The transcript on the CP's tool will pause and continue the same way (Figure 1). On both the AAC user's and CP's interface, a set of previous utterances in the conversation will be highlighted to **mark a context** the AAC user may be speaking to and to notify the CP that the AAC user is typing (Figure 3). This serves to further ground and establish the relevance of the AAC user's message. AAC users and CPs see the same context mark in their own interface views. COMPA currently highlights the last or current turn closest to the time that the AAC user presses the pause button or starts typing.

Text Composition Panel. The text composition panel consists of a pause button to control the transcript panel view, a text input message window where the AAC user can type using their AAC device or preferred keyboard, and a submit button that will read the AAC users message out-loud and print it on the transcript panel after it is activated.

COMPA provides **starter phrase suggestions**. Starter phrases are defined as a sentence fragment that starts off a valid response, but is neither complete nor overly specific, based on the previous context of the conversation. The starter phrases can conversationally act as bridge between what the AAC user wants to say and the current conversation context. These phrases are generated by prompting ChatGPT [29] with manually created examples of starter phrases to various conversations, then providing the context of the current conversation for ChatGPT to start responding to. Three phrases are generated at each time the AAC user starts typing below the text input window (Figure 3 (V2 and V3), and Figure 1).

COMPA Version	Transcript Behavior	Starter Phrase Assistance	Partner Notification
1) Pause Signaling (PS)	Pause	None	"user is typing"
2) Context Marking (CM)	Pause + marks context	Conversation-specific	"user is typing about:" [context]
3) Intent Grounding (IG)	Pauses + marks context + user selects intention	Conversation & Intent- specific	"user has [intention] about:" [context]

Table 1: COMPA versions explored three different grounding strategies: Pause Signaling, Context Marking, and Intent Grounding.

COMPA provides three phrase suggestions as to not overwhelm the AAC user with many options while still providing a variety of possible responses.

COMPA also includes **intent** buttons an AAC user can select to indicate how they want to respond in the conversation. These intent buttons are informed by common communicative intents commonly used in conversation [30]. The intents are condensed into four broad intents: reply, opinion, question, and yes/no. This is done to not overwhelm the AAC user with more than four choices. These groups are chosen as most responses can fall into one of these categories: reply (ex. statement-non-opinion), opinion (ex. statement-opinion), yes/no (ex. yes answers, agree/accept, response acknowledgement), and questions (ex. yes-no-question, open-question) [20].

The AAC user can select an intent prior to their response that will affect how starter phrases are generated. Once an intent is selected, instead of the generic message of "(the AAC user) is typing about", the CP will see a message such as "(the AAC user) has an opinion about". An exhaustive list of these messages can be found in Figure 3 and Table 1.

COMPA's Features and Versions

Three versions of COMPA were created to study how different features and changes to each interface's panels and different grounding strategies impact communication for AAC and CPs. Each version provides different context grounding strategies, phrase assistance types to the AAC user, and different types of feedback notifications to the CP, (Table 1 and figure 3):

- Version 1 (Pause Signaling): contains only the context grounding functionalities, pauses the transcript and displays general typing notifications to the CP.
- Version 2 (Context Marking): contains starter phrases for the AAC user that are conversation-specific in addition to the context grounding functionalities.
- Version 3 (Full version with Intent Grounding): contains context grounding functionalities and intents that the AAC user can select to generate intent-specific partner notifications in addition to starter phrase suggestions that are both conversation and intent specific.

6 EVALUATION STUDY

A total of ten participants, five AAC users (3 women) along with a communication partner (CP) of their choice were recruited for this study using organization mailing lists and social media interest groups (Table 2). Each AAC user and their CP were given a pre-study survey, which asked about each participant's communication strategies with their partner as well as each participant's perceived participation time during group conversations. Specific questions can be found in Appendix A.

The two-hour remote study session was completed through Google Meet. Each participant, AAC user and CP, was asked to connect from their own computer and from a separate room if they were in the same household to minimize any text-to-speech error. A researcher and author of this work acted as the study moderator, also actively participating in the conversation as a third party. Using a within-subjects experimental design, all participants used all three versions of COMPA. The order of the conditions was counterbalanced to prevent acclimation bias. We designed the evaluation study to simulate a conversation among three people: the AAC user, the CP, and the researcher. Our study set up resembles the triadic interaction shown in Figure 1. This study was approved by our institution's IRB.

Procedure

The remote user study consisted of four main parts for each interface condition: (1) an overview of COMPA's interface version, (2) a short tutorial, (3) a conversational task: planning a summer vacation, and lastly (4) an online questionnaire post-condition.

Tutorial. After a general explanation of the features, the researcher performed a storytelling tutorial task for the pair of participants so that they are familiar with the version's interface. The storytelling task consisted of the researcher asking the AAC user to use the features of the corresponding COMPA version to respond with the word "green" and

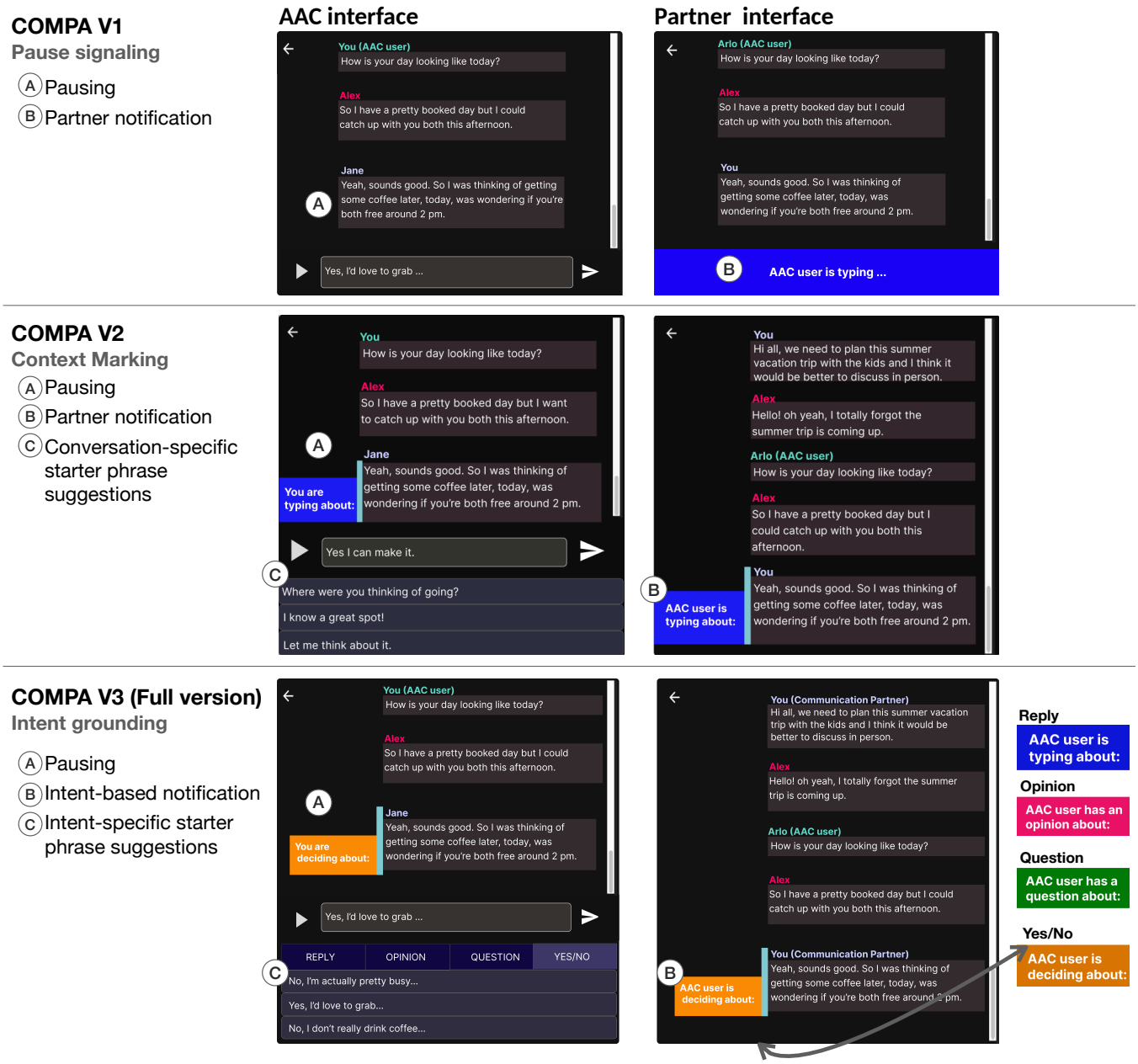

Figure 3: We created three versions of COMPA. Each version added a new layer of context information and phrase assistance type. Testing these three different versions would allow us to evaluate the contribution of each grounding strategy to the conversation experience.

a selected phrase starter (when applicable) upon hearing the same "green" keyword in the story narration. The CP is also invited to monitor their COMPA view and describe the changes they perceive on their screen.

Conversational task. After the tutorial, the researcher, AAC user, and CP participated in a structured conversational task of planning a summer vacation together. All three actively participated in the summer vacation discussion. The summer vacation planning task was split into three sub-tasks: (1) decide on the location for the vacation, (2) decide on the trip activities, and lastly (3) discuss who else to invite to the trip. The group completed each sub-task with a different version of COMPA. The order of these sub-tasks remained constant

during each run of the study. The summer vacation task was chosen to simulate a natural conversation among the triad where distinct sub-task or sub-topic (*i.e.*, choosing a location, activities, and people to invite) could be discussed under one umbrella theme of vacation planning.

Each sub-task lasted about 10 minutes and the researcher guided the conversation in a specific structure to maximize uniformity across participants. First, the researcher instructed participants on the items to be discussed (location, activities, or people to invite) and then proceeded with a direct question to the AAC user to hear their input first. Then, the researcher asked a direct question to the CP followed by one general question to the group. The researcher also made at least one big topic change, by sharing an unrelated anecdote, and by purposefully making common errors unfamiliar partners make with AAC users such as speaking while the AAC user is typing or asking two questions in a row. Additionally, the researcher actively participated in the conversation as a third member who would help plan the vacation and suggested locations, activities, and people to invite.

Post-condition surveys. Proceeding the use of each version of the tool, the AAC user and CP received different post-condition surveys. The AAC user was asked how their experience was using the tool, as well as how effective they thought the tool was in helping them communicate or participate more. The CP was asked questions targeting their awareness of the functionalities on the tool and their ability to support their partner. Each were also asked about the specific functionalities in the version of COMPA they are using. The specific questions can be found in Appendix B. The researcher stayed on the call while participants filled in the post-condition surveys, and allowed time for participants to voice open-ended feedback.

To conclude the study, the AAC user and CP received the same set of questions that asked them to rank the three COMPA versions in order of usefulness for them and for their partner (Appendix B). Using a take-home survey, participants were queried about their likelihood of utilizing the tool during meetings and were also prompted to identify additional scenarios where the tool could be applied (Appendix C).

Analysis and Data Collection

The remote study sessions (video and audio) were recorded, and they, along with the interaction logs of COMPA's interfaces, and survey responses from both AAC and CP participants were reviewed to extract both subjective and objective metrics. The subjective metrics included perceived sense of agency, that is if participants felt they participated in the conversations as much desired; perceived ease of input or effort required, and perceived awareness of screen changes of the CP. Objective measures included the number of turns taken by each participant and length of AAC user contributions per turn (words used per turn). To understand how the different COMPA features impacted the AAC users' participation the type of communicative functions used during each AAC user's turn are labeled (e.g., an initiation, stating a preference related to the task, a reply). We compared these metrics across the three COMPA versions. These metrics are all factors that can impact the effectiveness of communication and mutual understanding.

Study Limitations

While the study lasted two hours and included tutorials for each version of COMPA, participants in this study all experienced COMPA for the first time. Thus, it is possible that with long-term use of COMPA, results may vary. Similarly, we used a conversational task to elicit a natural three person conversation; sharing and collaborative decision-making was limited to one big theme of planning a summer vacation. Casual chat allowed for more free-form conversation where we could see more topic shifts; meetings would have more of a clear agenda. Future work is needed to understand how different COMPA features may or may not support different types of conversations (*e.g.*, personal v.s. work-related, virtual v.s. in-person).

In order to provide a safe space for the study, we chose a small group setting with three participants: the AAC user, their familiar CP, and the researcher. There are certain actions that are more present in conversations with unfamiliar partners, such as CPs talking without waiting for AAC users, that COMPA could alleviate, but we did not want to put AAC users in those situations with CPs they did not know. Given the exploratory nature of our study, it's essential to interpret our results within the specific scope of our research. Subsequent investigations, involving the extended and consistent use of our tool and a more extensive participant pool, could yield additional insights into features that contribute to mutual understanding between AAC and non-AAC users in real-time conversation.

7 RESULTS

COMPA was well received across the diverse set of study participants; in the post study survey, four out of five pairs of AAC users and CPs would use some version of COMPA in their online meetings (see Figure 6). COMPA was successfully installed by all study participants. However, AC1 had trouble accessing COMPA directly with their AAC device, as it restricted Google Chrome browser access. AC1 and CP1 still completed all study tasks with the help of an assistant who operated COMPA on their behalf. AC1 and CP1 completed the take-home survey to share their opinions on COMPA's features. We only include AC1's and CP1's pre-survey and take-home survey results about communication

AAC User	Age	AAC Device/Access Method	AAC Use	AAC Years	Speaking Partner
AC1	38	Tobii I 15 /Single Switch Scanning	Full time	32	CP1, Sister
AC2	53	Various text-based apps/Direct keyboard input	Part-time	20	CP2, Daughter
AC3	58	Speech Assistant, NextUp Talker, or Words+ apps/ EMG switch, SCATIR switch and EZ Keys	Full time	10	CP3, Spouse
AC4	45	Predictable and Speechify apps/Direct keyboard input	Part-time	40	CP4, Sister
AC5	48	Proloquo4Text app/Direct keyboard input	Full time	30	CP5, Colleague

Table 2: Our eight study participants include five AAC users (3 women) and five corresponding non-AAC communication partners. AAC users had 10 or more years of experience using AAC and use a diverse set of AAC devices and access methods. Some AAC users can use some speech to communicate. AAC users in our study only had disabilities associated with speech production, and did not have disabilities associated with language use or processing (*e.g.*, aphasia).

styles and feedback on COMPA as a whole. Our participants used different types of AAC devices and access methods and we noticed that each preferred different features of COMPA, and used its features differently.

Benefits of COMPA's Features

Participants assessed the utility of COMPA's features by ranking them from least to most useful (Figure 5). This scale allowed for a straightforward comparison of participants' preferences for each feature. Overall, AAC users leaned towards versions one and two of COMPA, which did not include intents—either due to their initial experience with COMPA or the increased intricacy of COMPA V3. AC1 and AC2 expressed equal favor for all COMPA versions, while AC3 leaned towards the full version (V3) for its inclusion of intents. In contrast, AC6 distinctly preferred COMPA V2, finding the inclusion of intention buttons to introduce unnecessary complexity. On the other hand, communication partners (CPs) favored versions two and three of COMPA, appreciating the context marking and intent-based notifications. Broadly, each feature served a unique purpose for every user, underscoring the importance of customization. AC2 noted: *"it would be great if the app allows users to choose different versions depending on the communication partners and the specific situation they find themselves in"*.

Context Grounding. COMPA's context grounding features consists of both transcript pausing and context marking within the transcript (*i.e.*, marking what turn an AAC user is replying to). Participants liked the transcript pausing and context marking functionalities (Figure 5, COMPA V2), but mostly felt that having the context marking was more useful than having the transcript pause by itself. AC3 shared that *"the pauses are well designed to keep [their] responses in context"* and that it *"seemed easier to time [their] responses"*. In general, context marking helped re-ground conversations when AAC users had comments about a previous topic or utterance. In the following example, CP5 shared an idea they had for COMPA while AC5 was typing a response to the previous comment made by the researcher:

RESEARCHER: Yeah the Texas *[transcript error]* bit will not work either if I mistype... [**this comment was marked and paused**]

CP5: *[starts to talk about another idea for COMPA]*

AC5: *[pauses transcript after researcher speaks]* the transcript just said Texas for [researcher] when [they] said text-to-speech

After AC5's comment, the discussion circled back to the limitations of text-to-speech for a short period of time without causing a misunderstanding.

CP3 and AC3 found the context grounding features to be especially useful for mutual understanding. CP3 shared: *"knowing the context of a response is especially important when solving problems together."* Participants used a wide range of non-verbal signals to understand when an AAC user wanted to participate. Typing sounds and some vocalizations cued CPs and helped them generally know when the AAC user wanted to speak and what part of the conversation they were responding to. CP4 noted that since they *"understand [AC4's] pattern of speech more so than others...it wouldn't help me have better interaction, but others who have a more difficult time understanding, it would be beneficial to them to know when [AC4] has something to say and what [AC4] has to say"*. However, AC4 noted that *"people may not pay attention [to the transcript pausing]"*.

Live Transcript. Having access to the transcript of the conversation in real-time also highlighted opportunities to include additional functionalities like translation. Two of our

AAC participants (AC1, AC2) are bilingual and both suggested embedding translation features into the transcription. The transcription would also often misspell proper names that were not of English origin. AC2 appreciated the transcript itself as well: *"As a person who speaks English as a second language, I appreciate the transcription feature embedded in the AAC app".*

Partner Notifications. In general all AAC users found partner notifications to be one of the most useful features to them and CPs preferred general typing notifications slightly more to intent-specific ones (Figure 5). AC2 commented that both general and intent-specific notification would help *"the conversations partners know when I am typing, so they would wait".*

Communication partners found the intent-specific notifications helped the flow of the conversation. CP3 shared *"the boxes like reply and opinion helped the conversation...It seemed like AC3 was able to move faster in conversation and it was easier to understand.".* CP4 also liked having more information about the AAC user's intention:

> *"I liked that we could anticipate what type of response [AC4] was going to make "* –CP4

CP4 shared that because she is used to communicating with her sister, the intent-specific notifications did not really change their ability to communicate with each other: *"we interacted about the same as usual...I feel the options on how to respond would benefit people who are not used to communicating with someone with AAC".*

CP2 found COMPA V2's notifications to be helpful because these appeared next to the turn AC2 wanted to reply to. CP2 found the intent-specific partner notifications to be unneeded: *"I enjoyed knowing what the AAC user is replying to so I know what to expect and what kind of a response I am waiting for, however I think it is excessive and unnecessary to know what their desired intention is."*

Starter Phrases. AAC users opinions about the phrase assistance types were varied; both general starter phrases (available in COMPA V2) and intent-specific starter phrases (COMPA V3) were both rated on average to be slightly appropriate for the conversation (2.75 (σ = 0.96) and 2.75 (σ = 1.3) respectively). General starter phrases were used two times and intent-specific starter phrases were used six times. Three out of 5 AAC users found the intent-specific starter phrases as the most useful for them (Figure 4). General starter phrases were only used twice by AC4. AC4 used the phrases to reply to two specific questions and modified the phrases using the text input window:

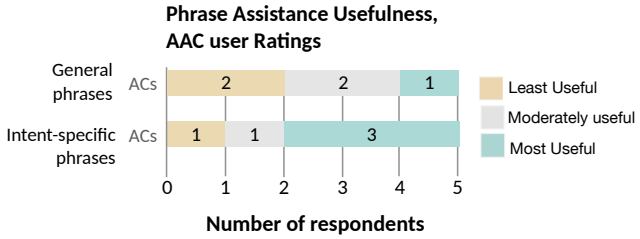

Figure 4: When asked about phrase assistance types and their usefulness, three participants (AC1, AC2 and AC4) rated intent-specific starter phrases to be most useful to them. AC5 consistently rated both phrase assistance types as the feature least useful to them. AC5 provided additional comments sharing that she prides herself in her vocabulary and the phrases provided are limited.

RESEARCHER: I was wondering if you all do any research about like, before you decide to go somewhere. If it's accessible or if it has some accessible spots before you decide to go?

AC4: *[pauses transcript by* commencing typing: "y", *then stops and* identifies phrase they want *and selects it. The phrase selected is "I actually like looking for accessible spots."* AC4 goes back to *editing her message.]*

AC4: yes I actually like looking for accessible spots.

AC2 and AC3, who had very different typing speeds and used different typing methods, did not use any of the general starter phrases, but preferred the intent-specific starter phrases. AC3, who controls a switch with his muscles to type, shared: *"I didn't feel a good fit with the suggestions."* AC2, who uses a keyboard, shared she did not use them because she *"mostly typed without looking at them."* Overall, five intent-specific phrases were utilized by three AAC user participants. Two users shared they preferred the intent-specific phrases, including AC2 who used the intent-specific phrases in 33% of her turns when having a conversation using COMPA 3.

AC2 who used the most intent-specific phrases, selected a phrase and edited it in 3 out of 4 occasions. In one example when the researcher was talking about seafood, AC2 selected the reply intent-specific phrase: *"I'm not a big seafood fan, but..."* and edited it to *"I m a big seafood fan".*

Overall, it was difficult for the starter phrases to capture the exact wording AAC users wanted. AC5 and AC2 reported not paying attention to the phrases, while AC3 recommended *"[phrase] suggestions could combine context with word[s] typed".*

Intention Buttons. AC2 and AC3 found the intention buttons were appropriate for the conversation, while AC5 was neutral and AC4 only found the intention options to be

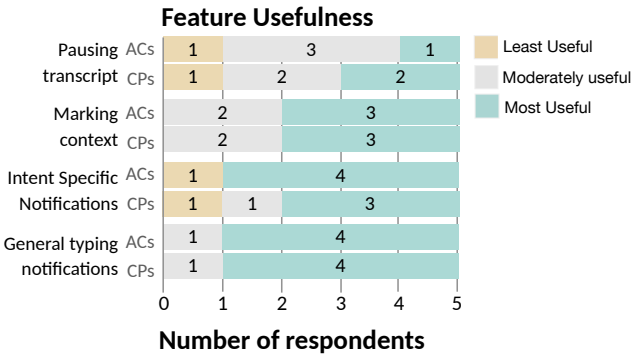

Figure 5: All study participants (N=10), 5 AAC users (ACs) and 5 communication partners (CPs), rated all of COMPAs features after experiencing all versions of the tool. CPs rated pausing the transcript, marking the context, and general typing alerts as the most useful features to help them better communicate with their AAC partners. All AAC users rated the general typing alerts and the intent-specific typing alerts as either highly or moderately useful features.

slightly appropriate. AAC users selected the intention buttons 7 times in total to change their partner notification message (see Figure 3, COMPA V3 for a list of the intent-specific CP notifications). Even though AC4 did not use any intent-specific phrase suggestions, they used the intent buttons to indicate their intention to their partner. Similarly, AC3 used an intention button to indicate he had a question after the researcher summarized the plans they all had made for their zoo trip. AC3 chose the question intent button and then proceeded to type a question of his own. The researcher had forgotten an item, so he asked: *"Did you forget my bears?"* The "reply" intention was more often selected than all the other intentions (5 times in total), followed by the opinion intention and the question intention, chosen one time each.

Although the intentions were not used frequently, AC3 noted that *"'using intents could be very useful with practice"*. CP4 made additional comments on how this feature may *"benefit people who are not used to communicating with someone with AAC"*. However, AC4 thought that *"the version with the categories is way too complex"*.

AC5 and CP5 noted that the four intents chosen were not all-encompassing and suggested implicit bias that an AAC user might be passive in the conversation. CP5 shared: *"I think that sometimes using those carrier phrases you have would alert the communication partner to what the intent is...however it also narrows the intent to suggest that someone would respond to only these ways and [AC5] does not only respond in the ways that are presented to her."*

Using COMPA

In general, both AAC users and CPs had positive experiences using COMPA. COMPA did not heavily detract from and slightly helped the ability of the AAC user to communicate what they wanted in the conversation by requiring slight additional effort in some ocassions. AC2, AC3 and AC5 reported slight additional effort or less to use COMPA, while AC4 reported some additional effort when using COMPA V2 and extreme additional effort when using its full version. When using COMPA, AAC users reported being often (AC3, AC4) or always (AC2, AC5) able to say what they wanted (COMPA V1: 4 (σ = 0.82), COMPA V2: 4.5 (σ = 0.58), COMPA V3: 4 (σ = 0.82)). COMPA also made it slightly easier than usual to communicate with their partner (V1: 3.5 (σ = 0.58), V2: 3.25 (σ = 0.5), V3: 3.5 (σ = 1.3)). AC2 found that communication with their partner while using COMPA V1 and V2 remained relatively consistent with their typical experience, and became slightly easier when using COMPA V3. AC3 reported it was slightly easier to communicate using V1 and V2 and markedly improved when using COMPA V3. In contrast, AC4 reported it was easier to communicate with COMPA V1 and V2 but harder to communicate when using COMPA V3.

Amount and Quality of Participation. On average AAC users took more turns when using COMPA's full version (COMPA 3) but used more words per turn when using COMPA 1. In general, CPs took more turns in conversation than AAC users. The distribution of turns taken by AAC users and CPs was comparable when participants used the full version of COMPA (Table 3). Perceived participation is also important to consider when evaluating COMPA. AAC users thought that they participated on average slightly more in the conversation than usual (3.5 (σ = 1.3), 3.5 (σ = 0.58), 4.25 (σ = 0.96). CPs also rated the AAC user's participation to be slightly more than usual (average 3.25 (σ = 0.5)) for all versions). AAC users perceived amount of participation matched with the observed number of turns quantified (Table 3). AC2 and AC3 said that their participation was definitely more than usual (rated 5) when using version 3 of COMPA, and they thought that their participation increased from versions 1 and 2. Interestingly, CPs often did not quantify AAC user increased participation in the same way the AAC users did. Most CPs rated that the AAC user participated about the same amount while AAC users felt that they participated more. Pair 5 noted that AC5 participated the same amount as usual, and AC5 noted that COMPA did not *"enhance what i already do to communicate"*.

Partner Awareness. One of COMPA's design goals is to facilitate mutual understanding between the AAC users and CPs. It is important that COMPA achieves this goal while not introducing additional hindrances. CPs did not find it difficult

	AAC users				Communication Partners			
COMPA Version	**AC turns**	**Stdev**	**Words per turn**	**Stdev**	**CP turns**	**Stdev**	**Words per turn**	**Stdev**
V1	**9**	2.31	**11.54**	13.12	**11**	6.06	**25.72**	12.39
V2	**11.75**	2.87	**7.60**	3.49	**19**	9.20	**14.13**	4.95
V3	**14.75**	5.74	**6.27**	2.75	**24.50**	13.96	**16.48**	3.89

Table 3: Average turns taken and words per minute spoken by turn. The average number of turns taken by AAC users and CPs was comparable when participants used the full version of COMPA.

to keep track of screen changes during the conversation. CP2 and CP4 reported finding it easy to keep track of screen changes during the conversation, while CP3 and CP5 rated this ability as neutral. CPs reported knowing more often what part of the conversation the AAC user wanted to respond to with COMPA V2 and V3 than COMPA V1 (V1: 3.25 (σ = 1.3), V2: 4.25 (σ = 0.96), V3: 4.25 (σ = 0.96)).

CP2 and CP3 almost always knew when AC2 or AC3 wanted to talk. CP4 often knew when AC4 wanted to talk with COMPA V2 and V3, and CP5 reported sometimes knowing when AC4 wanted to talk with all three versions of COMPA. In general CPs knew when AAC users wanted to participate in conversation (V1: 4 (σ = 1.2), V2: 4.25 (σ = 0.96), V3: 4 (σ = 0.82)).

Most CPs sometimes had an idea of how the AAC user wanted to respond (V1: 3.5 (σ = 0.58), V2: 3.75 (σ = 0.50), V3: 3.75 (σ = 0.50)). To improve partner awareness, CP1 recommended exploring different modalities for notification such as haptics or animations. CP4 reported often having an idea, while CP2, CP2, CP3, and CP5 reported either often or sometimes having an idea of how the AAC user wanted to respond.

COMPA Usage in Everyday Life AACs and CPs had ideas of where they would use COMPA in their everyday life. AC3 noted that their *"participation in dynamic group conversation is minimal [at work]. Biggest obstacle is responses are not tied to proper context"* and so COMPA *"would be a huge benefit during a group conversation on Zoom"*. AC2, AC4, and CP4 talked about using the tool for virtual conversations. AC1 wanted to *"'use the app along with my device because other people do not necessarily know if I am typing a respond to them, and I usually lose the conversation"*.

AAC users shared that it is often hard to switch between their AAC device and remote meeting software. Thus, they really liked that COMPA was a browser extension that could be easily added on to Google Meet and be accessed with their AAC software or keyboard. One pair of participants (AC2, AC4) also enjoyed the built-in text-to-speech with COMPA. Participants noted that having an auditory cue when the AAC user was speaking enhanced the experience compared to solely relying on text input. This was particularly valuable for them, given their tendency to use chats exclusively during remote meetings.

COMPA's Impact on Conversation Partner Support

In the pre-study surveys, participants had shared how they already had various strategies in place to facilitate communication. For example, AC1 said: *"My partner and I will look at each other for cues or share thoughts"* and AC5 added that *"[the CP] translates a lot of my verbal speech"*. CP2, CP3, CP4 noted that they will help the AAC user communicate if needed, to resolve misunderstandings or provide further elaboration. When using COMPA, CPs and AAC users made use of these strategies as well, being vigilant to various auditory cues and waiting to hear a response to facilitate the exchange. CP2 noted that *"Since I am able to hear my partner typing, I don't have to rely solely on COMPA and the typing flags to notice if my partner is communicating"*.

AC5 and CP5 had an insightful discussion on turn-taking and the need for mutual support and accommodations during communication. A notable disparity in perspectives emerged. In the following example, the researcher had on purpose started sharing a personal trip anecdote while AC5 was still typing, to simulate what could happen in a group conversation. CP5 promptly alerted the researcher, emphasizing that COMPA indicated AC5 was in the process of typing, and speaking simultaneously was not ideal:

CP5: I see that [the AAC user] is typing but we did not stop talking.

AC5: But I didn't wait for you to stop talking. I should have waited instead of type and wait to hit play.

CP5: I don't think you should have done anything...when is one on one is different but in a group we can just keep babbling...

AC5: but that is me talking over you which is no better than others talking over me.

While CP5 wanted everyone to stay quiet, and even suggested COMPA should *"perhaps mute everyone"*, AC5 did not want that. AC5 further shared that she does not usually expect any support from the individuals she is talking to: *"I am*

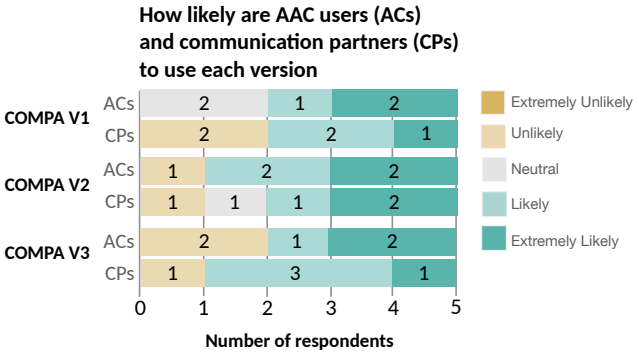

Figure 6: AAC users rated each in version in terms of how likely they would use it again if it were available to them. The second version received the most highest ratings from four AAC users (AC1-AC4), while CPs (CP1-CP4) chose the full version as their preferred option. AC5 and CP5 expressed they were either unlikely to use COMPA as it did not seem to contribute significantly beyond their usual methods of communication with each other.

never seeking support, just normal conversation with whoever I am talking to." CP5 challenged AC5 further:

CP5: Now, I am going to challenge you a little bit. You need people to wait. Or I need to wait to have a conversation with you.
AC5: I never expected that in my life so.
CP5: If I am ever going to have a good conversation with you, I need to wait, maybe it's more on me.
AC5: but it does depend on the context and person.

8 DISCUSSION AND FUTURE WORK

In this work, we studied COMPA in triadic conversations among a researcher, an AAC user and a conversation partner (CP) acquainted with the AAC user. Participants used three versions of COMPA, a browser extension that proposes different features to support mutual understanding during real-time conversations among AAC and non-AAC users. Based on this work, we reflect on how COMPA's features may support mutual understanding and participation, what future work is needed, who really needs support during conversation, and how we can move towards more embedded and widely available AAC accessibility features.

Creating a Shared Common Ground with COMPA

COMPA's context marking features were inspired by formative work and the idea that if we augment the shared conversational context and make it interactive (*e.g.*, with pausing capabilities, and added notifications), AAC users could speak to any topic despite other group members moving on to a new topic before they are done typing. In our study, we observed that in all group conversations, AAC users actively participated on all topics discussed by the group.

Our hypothesis was that COMPA's shared live transcript could help others keep track of context so that when AAC users spoke, everyone could reference the transcript as needed to disambiguate any misunderstandings. We found that in general COMPA facilitated communication between AAC and non-AAC communication partners. COMPA's shared live transcript and intent-specific notifications enabled partners to know when and to what an AAC user wanted to reply to. COMPA's features were generally easy of use and provided valuable contextual information to all conversation participants. The majority of AAC users indicated COMPA was easy to use and that they would use COMPA's features if available to them.

While COMPA presents novel features, COMPA only considered the text modality of shared context through the transcript as a metaphor for shared common ground. Our participants highlighted that they didn't solely rely on COMPA for their conversations; instead, they also incorporated non-verbal gestures, sounds, and the dynamics of their rapport. Supporting AAC users and their partners in group conversation requires a multi-modal approach that also acknowledges people's personal relationships.

There is an opportunity for future work to consider other ways to represent conversational common ground. For example, creating a meeting summary[24] for the cases when conversation has jumped between multiple topics, could reduce some screen space or making use of conversation thread visualizations explored in other Computer-Mediated-Communication tools [12, 18].

Communicative Intents as a New Signal

Guided by our formative work, we incorporated intent buttons into COMPA. As such, COMPA proposes this novel feature, and our study demonstrates that intention-specific notifications provide non-AAC speaking partners with useful information about an AAC users' stance which allowed partners to wait in expectation of an AAC users comment. For AAC users, intents were useful to find possible starter phrases to use and to clearly notify the group that they had a specific question to ask before the group resumed a specific topic. As a reflection of their usefulness, intent-specific phrases were rated as most useful than general phrase suggestions by three AAC user participants.

While, COMPA only considers 4 communicative intents, they served as an example to AAC users and CPs about other possible status indicators for AAC beyond typing notifications. The four intents chosen for COMPA were informed by commonly used intents in conversational chat [30]. We chose this limited set as a starting point to explore the concept with our specific casual conversational task of planning

a summer vacation. These intents may differ depending on the setting of the conversation; intents for meetings may be different than intents for casual conversation. As emphasized by AC5, presenting pre-defined limited intents in an interface can unintentionally convey a misleading message to an AAC user, suggesting an expectation for only specific intents from them. Even if the categories of intents available to a user are predicted by a model, it is important the user can customize and override or have easy access to a myriad of intents to avoid unintentional harmful messaging that all AAC users communicate in a similar way. Further work is needed to explore multiple ideas for adding intent buttons or intent indicators to AAC systems while considering user's agency and self-expression.

Starter Phrases vs Phrase Completion

This work introduces starter phrases as a new type of conversation assistance that aims to bridge the gap between the previous context and the current new context in a conversation with a short phrase. These starter phrases were inspired by much work done in the field of AAC and utterance-based speech-generating devices [42, 53]. Utterance-based AAC devices use full utterances or phrases people can quickly use in conversation. In this work we propose using starter phrases generated by a language model that consider the current conversational context close to the AAC user's typing time and, in COMPA V3, a user's selected intention. Unlike word prediction or full context-aware phrase generation [37], COMPA's starter phrases provide a starting point to linguistically reintroduce a former topic. We observed that general starter phrases were used twice and intent-specific phrases were used six times across participants. Used phrases contained key contextual words that helped the AAC user save time when responding to specific questions or comments.

Recent work in AI generated text support for AAC demonstrated that providing a full phrase, even if contextually-aware, was not enough [48]. AAC users see value in communicating their unique style and personal expression through their writing. In line with this prior work, participants in our study tended not to utilize starter phrase suggestions as frequently, likely for similar reasons. When they did use them, they edited the phrases to accurately convey their intended message. This highlights a potential trade-off between personal expression and the extra effort required for communication, a challenge that future Large Language Models (LLMs) could potentially address through personalization. Starter phrases were useful for two participants in our study and current advances in LLMs [3, 4, 38, 48] suggest a future where personalized models could adapt to an AAC user's context, speaking style and changing preferences over time.

Who Needs the Most Support?

Throughout this work, our focus was on understanding how to effectively support both AAC and non-AAC users in group conversation, considering the challenges posed by speaking rate asymmetries [15, 36]. COMPA features such as partner notifications, the live transcript pausing, and context marking were thought to support the non-AAC communication partner who might miss the context an AAC user may want to speak to. These features in turn would allow AAC users with more opportunities to revisit past topics despite their non-AAC user partners moving on. COMPA also supports AAC users through starter phrases and intent buttons that could be used to cue their CP when needed. AC5 and CP5, close friends and colleagues, had a deep moment of reflection about assistance and support while using COMPA. Initially, CP5 thought she should wait for AC5's response without changing topics. However, AC5 reminded CP5 that if that were the case, she herself also needed to wait for CP5 and not type while others were speaking. This realization underscored the mutual dependence for a productive conversation, but also the presence of unspoken assumptions about what AC5 needed that were not in line with her expectations.

In the past years, AAC research in HCI has moved from an AAC user-centric perspective towards improving the shared experienced of communication which is impacted by social factors. As such, researchers have discussed new AAC design goals such as maximizing agency [17, 49] and relational maintenance [8]. Similarly, groupware [10], shared tools that both AAC and non-AAC partners can use to share cues and feedback [39, 50], have proven beneficial. COMPA embraced these research directions by seeking to improve communication for all participants. Nonetheless, the insights from the discussions between AC5 and CP5 brought to the forefront the tendency to make assumptions about AAC users' preferences and expectations and the need to design for multiple paradigms and mindsets. It is imperative to design assistive tools for both AAC and non-AAC users that support them both in learning and creating new ways of communicating while also examining how tools might unintentionally impose additional challenges or unequal expectations among users. By critically reflecting on implicit assumptions we might be able to steer away from ableism in HCI research which often manifests as prejudice against disabled people's experiences and or erasure of their perspectives [16].

Broadening AAC Accessibility

COMPA was designed to be an add-on tool that does not replace a user's communication device, but rather provides additional features. Due to COMPA being a web browser extension, it can connect to virtual meeting rooms and be compatible with various types of AAC devices. As a result,

we were able to recruit a variety of AAC users who could install COMPA on their browsers and operate it with their preferred AAC setup at home. Many participants expressed they would like to use COMPA with Zoom and other remote meeting platforms. COMPA demonstrates it is possible to create AAC tools that are accessible to a wide variety of users. Many of the systems developed for AAC are either separate applications or completely new AAC systems [7]. Seeking to develop AAC "add-on" tools revealed the opportunity to embed AAC accessibility features into existing consumer software, such as remote meeting software.

9 CONCLUSION

In this work, we present COMPA, an AAC tool that complements an AAC user's current AAC device and provides context marking, starter phrases, and intent-specific notifications for communication partners. COMPA's design was informed by formative work and motivated by the "out of context problem": when AAC users are typing, other conversation participants may change the topic of the conversation faster than what AAC can respond to. This can cause AAC users' comments to be misunderstood, dismissed, or taken out of context, which can discourage AAC users from participating in conversation. A user study (N=10) indicated that COMPA can support both AAC and non-AAC users in having a better flow in their conversation. Participants said they would use COMPA for their virtual meetings and highlighted important opportunities for future work.

ACKNOWLEDGMENTS

J.H. was supported by the NSF Graduate Research Fellowship under Grant Nos. DGE1745016 and DGE2140739. The opinions expressed in this paper do not necessarily reflect those of that funding agency.

REFERENCES

[1] Sarah W. Blackstone and Barbara Collier. 2008. Communication assistants: Human supports for communication access. *Augmentative Communication News* 20, 3 (2008), 1–15.

[2] Virginia Braun and Victoria Clarke. 2006. Using thematic analysis in psychology. *Qualitative research in psychology* 3, 2 (2006), 77–101.

[3] Shanqing Cai, Subhashini Venugopalan, Katrin Tomanek, Shaun Kane, Meredith Ringel Morris, Richard Cave, Robert Macdonald, Jon Campbell, Blair Casey, Emily Kornman, Daniel E Vance, and Jay Beavers. 2023. SpeakFaster Observer: Long-Term Instrumentation of Eye-Gaze Typing for Measuring AAC Communication. In *Extended Abstracts of the 2023 CHI Conference on Human Factors in Computing Systems* (Hamburg, Germany) *(CHI EA '23)*. Association for Computing Machinery, New York, NY, USA, Article 397, 8 pages. https://doi.org/10.1145/3544549.3573870

[4] Shanqing Cai, Subhashini Venugopalan, Katrin Tomanek, Ajit Narayanan, Meredith R Morris, and Michael P Brenner. 2022. Context-Aware Abbreviation Expansion Using Large Language Models. *arXiv preprint arXiv:2205.03767* (2022).

[5] Senthil Chandrasegaran, Chris Bryan, Hidekazu Shidara, Tung-Yen Chuang, and Kwan-Liu Ma. 2019. TalkTraces: Real-time capture and visualization of verbal content in meetings. In *Proceedings of the 2019 CHI Conference on Human Factors in Computing Systems*. 1–14.

[6] Herbert H Clark and Susan E Brennan. 1991. Grounding in communication. (1991).

[7] Humphrey Curtis, Timothy Neate, and Carlota Vazquez Gonzalez. 2022. State of the Art in AAC: A Systematic Review and Taxonomy. In *Proceedings of the 24th International ACM SIGACCESS Conference on Computers and Accessibility* (Athens, Greece) *(ASSETS '22)*. Association for Computing Machinery, New York, NY, USA, Article 22, 22 pages. https://doi.org/10.1145/3517428.3544810

[8] Jiamin Dai, Karyn Moffatt, Jinglan Lin, and Khai Truong. 2022. Designing for Relational Maintenance: New Directions for AAC Research. In *Proceedings of the 2022 CHI Conference on Human Factors in Computing Systems*. 1–15.

[9] Adriane B Davis, Melody M Moore, and Veda C Storey. 2002. Context-aware communication for severely disabled users. In *Proceedings of the 2003 conference on Universal usability*. 106–111.

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

[11] Eric Friginal, Pamela Pearson, Laura Di Ferrante, Lucy Pickering, and Carrie Bruce. 2013. Linguistic characteristics of AAC discourse in the workplace. *Discourse Studies* 15, 3 (2013), 279–298.

[12] Carla F. Griggio, Arissa J. Sato, Wendy E. Mackay, and Koji Yatani. 2021. Mediating Intimacy with DearBoard: A Co-Customizable Keyboard for Everyday Messaging. In *Proceedings of the 2021 CHI Conference on Human Factors in Computing Systems* (Yokohama, Japan) *(CHI '21)*. Association for Computing Machinery, New York, NY, USA, Article 342, 16 pages. https://doi.org/10.1145/3411764.3445757

[13] Bronwyn Hemsley, Susan Balandin, Leanne Togher, and Disability Initiative. 2007. Communication in a Small Group: Exploring Successful Interaction. *PEC@ 2007* (2007), 101.

[14] DJ Higginbotham, K Fulcher, and J Seale. 2016. Time and timing in interactions involving individuals with ALS, their unimpaired partners and their speech generating devices. *The silent partner* (2016), 199–229.

[15] D Jeffery Higginbotham and DP Wilkins. 1999. Slipping through the timestream: Social issues of time and timing in augmented interactions. *Constructing (in) competence: Disabling evaluations in clinical and social interaction* 2 (1999), 49–82.

[16] Megan Hofmann, Devva Kasnitz, Jennifer Mankoff, and Cynthia L Bennett. 2020. Living disability theory: Reflections on access, research, and design. In *Proceedings of the 22nd International ACM SIGACCESS Conference on Computers and Accessibility*. 1–13.

[17] Seray B Ibrahim, Asimina Vasalou, and Michael Clarke. 2018. Design Opportunities for AAC and Children with Severe Speech and Physical Impairments. In *Proceedings of the 2018 CHI Conference on Human Factors in Computing Systems*. ACM, 227.

[18] Zainab Iftikhar, Yumeng Ma, and Jeff Huang. 2023. "Together but Not Together": Evaluating Typing Indicators for Interaction-Rich Communication. In *Proceedings of the 2023 CHI Conference on Human Factors in Computing Systems* (Hamburg, Germany) *(CHI '23)*. Association for Computing Machinery, New York, NY, USA, Article 724, 12 pages. https://doi.org/10.1145/3544548.3581248

[19] Ryo Iijima, Akihisa Shitara, Sayan Sarcar, and Yoichi Ochiai. 2021. Word Cloud for Meeting: A Visualization System for DHH People in Online Meetings. In *Proceedings of the 23rd International ACM SIGACCESS Conference on Computers and Accessibility*. 1–4.

[20] Dan Jurafsky, Liz Shriberg, and Debra Biasca. 1997. WS-97 Switchboard DAMSL Coders Manual. https://web.stanford.edu/~jurafsky/ws97/manual.august1.html

[21] Shaun K. Kane, Meredith Ringel Morris, Ann Paradiso, and Jon Campbell. 2017. At Times Avuncular and Cantankerous, with the Reflexes of a Mongoose: Understanding Self-Expression Through Augmentative and Alternative Communication Devices. In *Proceedings of the 2017 ACM Conference on Computer Supported Cooperative Work and Social Computing* (Portland, Oregon, USA) *(CSCW '17)*. ACM, New York, NY, USA, 1166–1179. https://doi.org/10.1145/2998181.2998284

[22] Jennifer Kent-Walsh, Kimberly A Murza, Melissa D Malani, and Cathy Binger. 2015. Effects of communication partner instruction on the communication of individuals using AAC: A meta-analysis. *Augmentative and alternative communication* 31, 4 (2015), 271–284.

[23] Walter S Lasecki, Raja Kushalnagar, and Jeffrey P Bigham. 2014. Helping students keep up with real-time captions by pausing and highlighting. In *Proceedings of the 11th Web for All Conference*. 1–8.

[24] Shixia Liu, Michelle X Zhou, Shimei Pan, Yangqiu Song, Weihong Qian, Weijia Cai, and Xiaoxiao Lian. 2012. Tiara: Interactive, topic-based visual text summarization and analysis. *ACM Transactions on Intelligent Systems and Technology (TIST)* 3, 2 (2012), 1–28.

[25] Kathleen F McCoy, Jan Bedrosian, and Linda Hoag. 2010. Implications of pragmatic and cognitive theories on the design of utterance-based AAC systems. In *Proceedings of the NAACL HLT 2010 Workshop on Speech and Language Processing for Assistive Technologies*. 19–27.

[26] Emma J McDonnell, Soo Hyun Moon, Lucy Jiang, Steven M. Goodman, Raja Kushalnagar, Jon E. Froehlich, and Leah Findlater. 2023. "Easier or Harder, Depending on Who the Hearing Person Is": Codesigning Videoconferencing Tools for Small Groups with Mixed Hearing Status. In *Proceedings of the 2023 CHI Conference on Human Factors in Computing Systems* (Hamburg, Germany) *(CHI '23)*. Association for Computing Machinery, New York, NY, USA, Article 780, 15 pages. https://doi.org/10.1145/3544548.3580809

[27] Hanne Sofie Midtlin, Kari-Anne B Næss, Tone Taxt, and Asgjerd Vea Karlsen. 2015. What communication strategies do AAC users want their communication partners to use? A preliminary study. *Disability and Rehabilitation* 37, 14 (2015), 1260–1267.

[28] Margaret Mitchell and Richard Sproat. 2012. Discourse-based modeling for AAC. In *Proceedings of the Third Workshop on Speech and Language Processing for Assistive Technologies*. 9–18.

[29] OpenAI. 2023. ChatGPT Large Language Model (gpt-3.5-turbo). https://chat.openai.com/chat

[30] E Margaret Perkoff. 2021. Dialogue Act Analysis for Alternative and Augmentative Communication. (2021).

[31] S. Pyszka. 2020. *Inside My Outside: An Independent Mind in a Dependent Body*. Independently Published. https://books.google.com/books?id=wmdrzQEACAAJ

[32] Albert B Robillard. 1994. Communication problems in the intensive care unit. *Qualitative Sociology* 17, 4 (1994), 383–395.

[33] Samiha Samrose, Daniel McDuff, Robert Sim, Jina Suh, Kael Rowan, Javier Hernandez, Sean Rintel, Kevin Moynihan, and Mary Czerwinski. 2021. Meetingcoach: An intelligent dashboard for supporting effective & inclusive meetings. In *Proceedings of the 2021 CHI Conference on Human Factors in Computing Systems*. 1–13.

[34] Antara Satchidanand, Jeff Higginbotham, Ann Bisantz, Naif Aldhaam, Ahmed Elsayed, Iman Carr, Ahmed A Hussein, and Khurshid Guru. 2021. "Put the what, where? Cut here?!" challenges to coordinating attention in robot-assisted surgery: a microanalytic pilot study. *BMJ open* 11, 7 (2021), e046132.

[35] Emanuel A Schegloff. 1997. Practices and actions: Boundary cases of other-initiated repair. *Discourse processes* 23, 3 (1997), 499–545.

[36] Jennifer M Seale, Ann M Bisantz, and David J Higginbotham. 2020. Interaction symmetry: Assessing augmented speaker and oral speaker performances across four tasks. *Augmentative and Alternative Communication* (2020), 1–13.

[37] Junxiao Shen, Boyin Yang, John J Dudley, and Per Ola Kristensson. 2022. 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* (Helsinki, Finland) *(IUI '22)*. Association for Computing Machinery, New York, NY, USA, 853–867. https://doi.org/10.1145/3490099.3511145

[38] Junxiao Shen, Boyin Yang, John J Dudley, and Per Ola Kristensson. 2022. 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*. 853–867.

[39] Kiley Sobel, Alexander Fiannaca, Jon Campbell, Harish Kulkarni, Ann Paradiso, Ed Cutrell, and Meredith Ringel Morris. 2017. Exploring the design space of AAC awareness displays. In *Proceedings of the 2017 CHI Conference on Human Factors in Computing Systems*. 2890–2903.

[40] Seoyun Son, Junyoug Choi, Sunjae Lee, Jean Y Song, and Insik Shin. 2023. It is Okay to be Distracted: How Real-time Transcriptions Facilitate Online Meeting with Distraction. In *Proceedings of the 2023 CHI Conference on Human Factors in Computing Systems*. 1–19.

[41] Yla R Tausczik and James W Pennebaker. 2013. Improving teamwork using real-time language feedback. In *Proceedings of the SIGCHI conference on human factors in computing systems*. 459–468.

[42] John Todman. 2000. Rate and quality of conversations using a text-storage AAC system: Single-case training study. *Augmentative and Alternative Communication* 16, 3 (2000), 164–179.

[43] Minh Hong Tran, Yun Yang, and Gitesh K Raikundalia. 2005. Supporting awareness in instant messaging: an empirical study and mechanism design. In *OZCHI'05: Proceedings of the 17th Australia conference on Computer-Human Interaction: Citizens Online: Considerations for Today and the Future*. Computer-Human Interaction Special Interest Group (CHISIG) of Australia.

[44] Minh Hong Tran, Yun Yang, and Gitesh K Raikundalia. 2009. Conversational awareness in text-based computer mediated communication. *Awareness systems: Advances in theory, methodology and design* (2009), 313–333.

[45] Keith Trnka, Debra Yarrington, John McCaw, Kathleen F McCoy, and Christopher Pennington. 2007. The effects of word prediction on communication rate for AAC. In *Human Language Technologies 2007: The Conference of the North American Chapter of the Association for Computational Linguistics; Companion Volume, Short Papers*. 173–176.

[46] Keith Trnka, Debra Yarrington, Kathleen McCoy, and Christopher Pennington. 2005. The Keystroke Savings Limit in Word Prediction for AAC. *University of Deleware* (2005).

[47] Meng-Ju Tsai. 2013. The effect of familiarity of conversation partners on conversation turns contributed by augmented and typical speakers. *Research in developmental disabilities* 34, 8 (2013), 2326–2335.

[48] Stephanie Valencia, Richard Cave, Krystal Kallarackal, Katie Seaver, Michael Terry, and Shaun Kane. 2023. "The less I type, the better": How AI Language Models can Enhance or Impede Communication for AAC Users. (2023).

[49] Stephanie Valencia, Amy Pavel, Jared Santa Maria, Seunga (Gloria) Yu, Jeffrey P. Bigham, and Henny Admoni. 2020. Conversational Agency in Augmentative and Alternative Communication. In *Proceedings of the 2020 CHI Conference on Human Factors in Computing Systems* (Honolulu, HI, USA) *(CHI '20)*. Association for Computing Machinery, New York, NY, USA, 1–12. https://doi.org/10.1145/3313831.3376376

[50] Stephanie Valencia, Mark Steidl, Michael Rivera, Cynthia Bennett, Jeffrey Bigham, and Henny Admoni. 2021. Aided Nonverbal Communication through Physical Expressive Objects. In *Proceedings of*

the 23rd International ACM SIGACCESS Conference on Computers and Accessibility. 1–11.

[51] Mary Wickenden. 2011. Whose voice is that?: Issues of identity, voice and representation arising in an ethnographic study of the lives of disabled teenagers who use Augmentative and Alternative Communication (AAC). *Disability studies quarterly* 31, 4 (2011).

[52] Bruce Wisenburn and D Jeffery Higginbotham. 2008. An AAC application using speaking partner speech recognition to automatically produce contextually relevant utterances: Objective results. *Augmentative and Alternative Communication* 24, 2 (2008), 100–109.

[53] Bruce Wisenburn and D Jeffery Higginbotham. 2009. Participant evaluations of rate and communication efficacy of an AAC application using natural language processing. *Augmentative and Alternative Communication* 25, 2 (2009), 78–89.

A PRE-STUDY QUESTIONS

The AAC user and CP have the same set of three questions.

(1) How do you rate your conversation when in group settings (when there is at least one other person and your AAC communication partner is present)? Please provide a short description of how you each participate in group conversations.
   - We participate equally in conversation
   - I participate more
   - They participate more
   - Other (please explain)
(2) Does **your partner** user any strategies (verbal or non-verbal) to make communication easier for you during group conversations? If so, please tell us more about what these are.
(3) Do **you** use any strategies (verbal or non-verbal) to make communication easier for your partner during group conversations? If so, please tell us more about what these are.

B POST-CONDITION QUESTIONS

The AAC user and CP have different post-condition questions. The base questions for the Version 1 of COMPA are listed below; Version 2 and Version 3 add additional questions. The questions for the AAC user are as follows, which focus on the usage of the tool and their perceived ability to effectively participate in the conversation:

(1) How much time would you say it took you to communicate using this tool?
   - Definitely less than usual (1)
   - Less than usual (2)
   - Around the same as usual (3)
   - More than usual (4)
   - Definitely more than usual (5)
(2) Compared to your everyday form of communication, how much additional effort did using this tool require?
   - No additional effort (1)
   - Slight additional effort (2)
   - Some additional effort (3)
   - Moderate additional effort (4)
   - Extreme additional effort (5)
(3) How would you rate your participation in this conversation?
   - Definitely less than usual (1)
   - Less than usual (2)
   - Around the same as usual (3)
   - More than usual (4)
   - Definitely more than usual (5)
(4) Were you able to pause the transcript at the point at which you wanted to talk?
   - Never (1)
   - Rarely (2)
   - Sometimes (3)
   - Often (4)
   - Always (5)
(5) Were you able to say what you wanted to say during this conversation?
   - Never (1)
   - Rarely (2)
   - Sometimes (3)
   - Often (4)
   - Always (5)
(6) How would you rate your ability to communicate with your partner?
   - Definitely harder than usual (1)
   - Slightly harder than usual (2)
   - Around the same as usual (3)
   - Slightly easier than usual (4)
   - Definitely easier than usual (5)
(7) How would you describe your conversation/interaction with your partner while using this tool?

The questions for the CP are as follows, which focus on their awareness during the conversation:

(1) How much time would you say it took you to communicate using this tool?
   - Definitely less than usual (1)
   - Less than usual (2)
   - Around the same as usual (3)
   - More than usual (4)
   - Definitely more than usual (5)
(2) Did you feel like you knew what part of the conversation your partner was responding to?
   - Never (1)
   - Rarely (2)
   - Sometimes (3)
   - Often (4)
   - Always (5)
(3) Did you feel like you knew when your partner wanted to talk?

- Never (1)
- Rarely (2)
- Sometimes (3)
- Often (4)
- Always (5)

(4) Did you have an idea of how your partner was planning to respond?
- Never (1)
- Rarely (2)
- Sometimes (3)
- Often (4)
- Always (5)

(5) How would you rate your ability to keep track of screen changes during the conversation?
- Very difficult (1)
- Difficult (2)
- Neutral (3)
- Easy (4)
- Very easy (5)

(6) How would you rate your AAC partner's participation in this conversation?
- Definitely less than usual (1)
- Less than usual (2)
- Around the same as usual (3)
- More than usual (4)
- Definitely more than usual (5)

(7) How would you rate your ability to support your partner during the conversation?
- More difficult than usual (1)
- Slightly more difficult than usual (2)
- Same as usual (3)
- Slightly easier than usual (4)
- Easier than usual (5)

(8) Did any of the tools affect your ability to support your partner? Please explain.

Both the AAC user and CP are given open-response areas to expand on these questions throughout the survey.

For Version 2, the AAC user is additionally asked:

(1) Were the starter phrases provided appropriate for the conversation?
- Inappropriate (1)
- Slightly inappropriate (2)
- Neutral (3)
- Slightly appropriate (4)
- Appropriate (5)

(2) What did you think about the starter phrases functionality?

For Version 3, the AAC user is additionally asked:

(1) Were the starter phrases provided appropriate for the conversation?
- Inappropriate (1)
- Slightly inappropriate (2)
- Neutral (3)
- Slightly appropriate (4)
- Appropriate (5)

(2) Were the intents provided appropriate for the conversation?
- Inappropriate (1)
- Slightly inappropriate (2)
- Neutral (3)
- Slightly appropriate (4)
- Appropriate (5)

C TAKE-HOME SURVEY QUESTIONS

After study completion, the AAC user and the CP are given the following questions to answer within the next few days:

(1) If we were to make this available online, would you use it in your meetings?
- Version 1 - alert others your are typing and pauses the transcript
- Version 2 - marking what you are referring to and giving you phrase suggestions
- Version 3 - marking what you are referring to and suggesting personalized phrases based on a desired intention

on a scale of:
- Extremely unlikely (1)
- Unlikely (2)
- Neutral (3)
- Likely (4)
- Extremely likely (5)

(2) Tell us a bit of why or why not you would use any of the versions of the application?

(a) Please rank these sets of features in order of usefulness to YOU as an AAC device user (or you as a communication partner):
- Pausing the conversation transcript when I type
- Marking what part in the conversation I am referring to
- General starter phrase suggestions
- Starter phrase suggestions by intents
- Alerting my partner that I am typing
- Alerting my partner about my intent (asking, deciding, etcetera)

on a scale from:
- Least useful to me (1)
- Somewhat useful to me (2)
- Most useful to me (3)

(b) Please rate these sets of features in order of what you think would be helpful for your communication partner (or for your AAC user partner):
- Pausing the conversation transcript when I type

- Marking what part in the conversation I am referring to
- General starter phrase suggestions
- Starter phrase suggestions by intents
- Alerting my partner that I am typing
- Alerting my partner about my intent (asking, deciding, etcetera)

on a scale from:

- Least helpful for my partner (1)
- Somewhat helpful for my partner (2)
- Most helpful for my partner (3)

(3) In what scenarios would you use this tool? Are there different scenarios in which you would use particular versions of the tool?

They are also asked to provide additional feedback with the questions: "How would you improve this? Did we miss anything? Anything you would like us to know?", and given additional space for other comments.