

# A Generative AI-Based Approach to Support Automated Utterance Generation for Different Conversational Contexts within AAC Systems

*Completed Research Paper*

**Daniel Konadl**

University of Regensburg  
Universitätsstraße 31,  
93051 Regensburg, Germany  
Daniel.Konadl@ur.de

## Abstract

*To foster their full integration into the mainstream of society, speech-impaired individuals need to be empowered to use vocal language in various daily life situations. AI-based AAC systems enable speech-impaired individuals to generate completed utterances for use in conversations. However, available solutions exhibit remarkable drawbacks. They do not meet all posed requirements and cannot provide completed sentences that match the expected styles of formal and informal conversational contexts. Therefore, design requirements for a Generative AI-based utterance composition approach were identified. It has been followed design science research to derive design principles and an instantiation of an AAC prototype that enables context-specific and user-individualizable articulation. The artifact was demonstrated and evaluated with formal and informal interactions commonly performed in the daily routine of visiting a restaurant. In particular, the importance of meeting all specific requirements and employing Generative AI, such as ChatGPT-4.0, to generate both formal and informal utterances, is exemplified.*

**Keywords:** Augmentative and Alternative Communication, Generative AI, Conversational Contexts, Automated Utterance Composition, Design Science Research

## Introduction

*“Perhaps the single quality most central to humanness is the ability to exchange thoughts, ideas, and feelings with others”* (Hourcade et al. 2004, p. 235). Speech is perceived as an individual’s most important instrument to get in contact with their surroundings (Fritzell 1996; Kane et al. 2017). However, individuals suffering from severe physical disabilities or brain injuries are not able to control their oral-respiratory musculature sufficiently for speech (Allen 2005; Vanderheiden 1983). This restriction to verbal communication and the *“separation from the mainstream of society”* (Hourcade et al. 2004, p. 235) holds true for people with damage to the vocal tract or other impairments (e.g., aphasia, autism, dyslexia) (Allen 2005).

To give speech-impaired individuals a voice and enable them to express themselves using vocal language, digital solutions for Augmentative and Alternative Communication (AAC) have emerged. These AAC systems enable the translation of an individual’s intended meaning into speech and subsequent voice outputs (Bradshaw 2013; Desai et al. 2014). One primary goal has thereby been to reduce the communication gap between speech-impaired individuals and their speech-capable interlocutors. To reach this goal, the extant AAC research has investigated different AI-based techniques that can automatically compose completed sentences from the single words and expressions speech-impaired individuals provide within their AAC systems (cf. Konadl et al. 2023). In this way, the communication gap could indeed be reduced by increasing the rate of words that speech-impaired individuals can contribute to conversations using AAC systems.

Nevertheless, AI-based AAC solutions still exhibit remarkable drawbacks regarding the utterances they generate. There is no comprehensive AAC system or tool available that meets all the requirements outlined in the scientific literature (see Table 2). However, extant literature points out the necessity of an AAC system that provides speech-impaired individuals with comprehensive support in various daily life conversations (Kane et al. 2017; Konadl et al. 2023). Beyond the extant research literature, the importance of such a comprehensive AI-based AAC system has been further underlined in this research through exchanges with speech-impaired individuals and representatives of the “Professional Association for the Hearing and Speech Impaired in Bavaria”. Concerning that, available AAC systems can provide completed utterances only for a limited number of situations as the underlying language models are trained on a limited number of daily life conversations with limited diversity (e.g., Dempster et al. 2010; Heo and Kang 2019; Kristensson et al. 2020). As a result, these language models struggle to generate completed utterances for domains and communication activities that speech-capable individuals usually can take part in. Moreover, even the fastest way of expressing oneself is of little help if the utterances proposed by an AAC system are perceived as distracting by speech-impaired individuals and their interlocutors. In particular, the style of an utterance may not fit the expected way of articulation in a formal or informal conversational context (Bedrosian et al. 2003; Hoag et al. 2004; Todman and Alm 1997). For example, an AAC system might suggest an utterance in which an authority figure like a doctor at the doctor’s office is addressed with colloquial language (e.g., “bro” or “buddy”) that is more suited to informal contexts. Hence, interlocutors may form negative perceptions towards speech-impaired individuals and avoid engaging in a conversation with them. The style of the utterances may also lead to speech-impaired individuals themselves avoiding conversations, as the utterances their AAC systems provide them do not reflect their personality and individual style of articulation (Shen et al. 2022; Tintarev et al. 2016; Valencia et al. 2023).

Therefore, this research aims to close these gaps. Based on Design Requirements (DRs) derived from the extant AAC research and theory, an instantiation of a functional AAC system prototype is proposed that integrates comprehensive functionalities to tackle the aforementioned drawbacks. The AI-based utterance composition approach is designed and constructed to combine single words and expressions into completed utterances in a wording that is appropriate for either formal or informal conversational contexts. Thereby, ChatGPT-4.0 as a Generative AI technique is employed because of its impressive capabilities in formulating human-like texts (cf. Feuerriegel et al. 2023; Radford et al. 2019; Teubner 2023). As a Generative Pre-trained Transformer (GPT), it combines large-scale architectures with vast amounts of textual training data, resembling contributions of humans on diverse daily life conversations, communication activities, and conversational styles (i.a., for formal and informal situations). However, the integration of ChatGPT-4.0 into an AAC system is not purposeful without additional and extensive adaptations, which have been realized in the design and implementation of the artifact. Furthermore, ChatGPT-4.0 alone is insufficient to meet all the DRs. Consequently, further AI and machine learning techniques were integrated and purposefully combined in order to create a comprehensive Generative AI-based utterance composition within the AAC system prototype. This proposed AI-based approach was evaluated with several communication activities commonly performed in the daily life routine of visiting a restaurant with a companion (cf. Ehlich and Rehbein 1972; Schindler 2013). Against this background, the following two Research Questions (RQs) are posed:

- **RQ1:** *How to design a Generative AI-based approach that can compose within an AAC system single expressions into completed utterances in formal and informal wordings?*
- **RQ2:** *How to instantiate and technically realize the Generative AI-based approach?*

The paper unfolds as follows: The following section provides the conceptual foundations and related work. Next, the research procedure following the design science approach (Hevner et al. 2004; Peffers et al. 2007) is described. The next section deals with the compilation of the Design Principles (DPs) that have been derived from the DRs. Then, the technical realization of the proposed AI-based approach, its demonstration to interactions in visiting a restaurant, and the results are presented. The paper proceeds with its contributions, potential limitations, and future research ideas. Afterwards, it closes with concluding remarks.

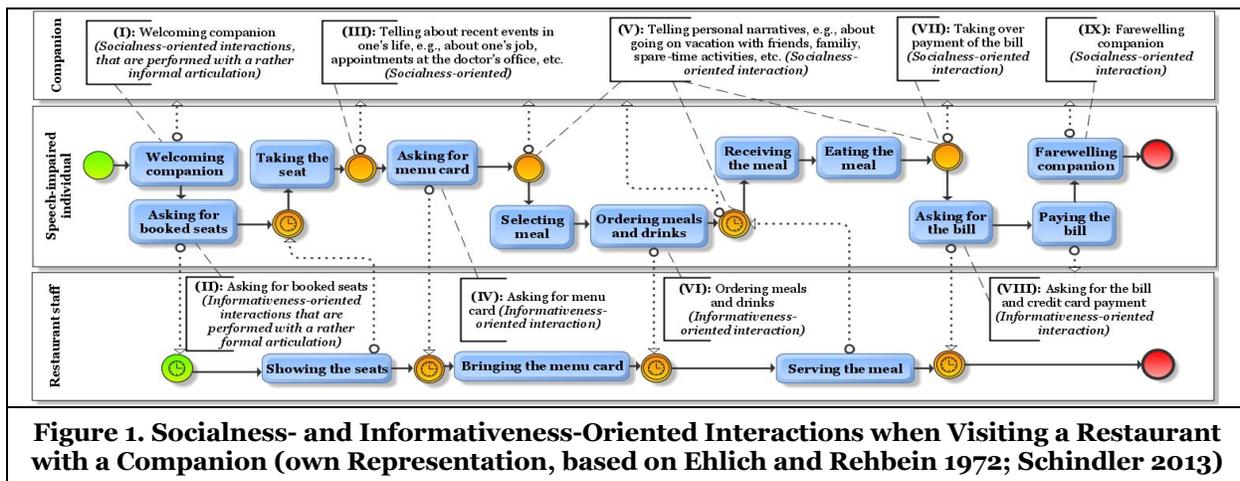
## Conceptual Foundations and Related Work

### Conceptual Foundations

Communication research and extant theories like the Gricean Maxims of Communication (Grice 1975) have investigated structures and components of daily life conversations (i.a., conversation phases, communica-

tion activities, and communication strategies) (Henne and Rehbock 2012). The way communication activities are performed is determined by different factors. In this regard, one of the most important factors is the conversational context (Bedrosian et al. 2003; Hoag et al. 2004; Todman and Alm 1997). Therefore, interlocutors use different articulations depending on where a conversation takes place and who they are speaking with. Articulations can differ in conversations with friends in restaurants compared to meetings with superiors at work. Furthermore, articulations are adjusted based on the complexity and cognitive demands of the conversation (cf. Cognitive Load Theory, Sweller 1988) and based on whether a social relationship has already been established with the conversation partner (social closeness to friends vs. social distance to strangers or authority figures).

Interactions with a friend (e.g., talking about recent events in one's life, telling personal stories, or thanking for a nice evening in a restaurant) are rather socialness-oriented and informal (cf. Todman and Alm 1997). Compared to that, interactions with waiters as authority figures in the restaurant (e.g., asking for the booked table, asking for the specials of the week, asking for payment) are rather informativeness-oriented and formal (cf. Bedrosian et al. 2003; Hoag et al. 2004). These differences are also reflected in the way interlocutors accommodate speech and the concrete wordings they apply within their utterances for the respective contexts (cf. Giles et al. 1991). In the German language, for example, interlocutors in an informal context tend to address each other by their first names and are more liable to apply contractions to words. Thereby, they may incorporate dialect, slang, and rather unsophisticated expressions into their utterances. In contrast, in a formal context, interlocutors are more likely to use the formal term of address (i.a., using last names) and employ more elaborated language constructs (i.a., euphemisms, precise terminology, forms of courtesy) (cf. Bedrosian et al. 2003; Budde 2012; Elspaß 2018; Hoag et al. 2004; Todman and Alm 1997). Speech-impaired individuals lack the fundamental capability to perform communication activities using their own vocal language in a style that fits the respective conversational contexts. Nonetheless, vocal language is an essential human capability for interacting with other individuals in various daily life routines like visiting a restaurant with a companion.



AAC encapsulates strategies, techniques, and tools that support individuals with speech impairments in expressing their thoughts, ideas, and feelings (Hourcade et al. 2004). In this context, AAC systems can be described as voice output communication aids or speech generating devices that provide speech-impaired individuals with the ability to express themselves using vocal language. High-tech AAC systems encompass the use of electronic devices (e.g., smartphones, tablet PCs) and their built-in peripherals (e.g., cameras, microphones) (Baldassarri et al. 2014; Laxmidas et al. 2021). The board within an AAC system contains the vocabulary that users can employ. Cognitive and language comprehension disorders may also require assigning pictograms to utterances or words to support the understanding of the related concepts. To contribute utterances to conversations, users of AAC systems can relate to pre-stored sentences (Alm and Higginbotham 2008), manually compose single words into sentences within the so-called speech composition line, or make use of keyboard capabilities and combinations of letters to express words and sentences (cf. Klauer et al. 2021; Kristensson et al. 2020). The utterance formed is then processed by the speech synthesis module which generates the voice output that is emitted to the conversation partner by the loudspeaker of the device. Voice outputs can then be interpreted and reciprocated by the interlocutor, allowing the conversation to progress by mutual and coordinated utterances until it reaches a targeted and desired state.

To help speech-impaired individuals compose completed sentences from single expressions, there have been proposed different AI-based approaches (i.a., abbreviation expansion, sentence prediction) (e.g., Dempster et al. 2010; Heo and Kang 2019) and recently also Generative AI (e.g., Cai et al. 2022; Valencia et al. 2023). As elaborated in the next two sections, Generative AI has yet not been harnessed to construct an AAC system with a comprehensive Generative AI-based utterance composition. Crucial aspects regarding the style of the utterances (e.g., matching expected and common styles of articulation in formal and informal contexts, including personal nuances and styles of articulation into the utterances) are thereby not met. Thus, in the following section, DRs derived from related work and extant theory are outlined, specifying what a Generative AI-based utterance composition needs to meet to be comprehensive and purposeful.

### ***Design Requirements of AI-Based Utterance Composition for AAC Systems***

To derive the DRs for the Generative AI-based utterance composition approach, an extensive search and consolidation of the research literature was conducted. Thereby, it was aligned with the procedure for conducting a systematic literature review as proposed by Vom Brocke et al. (2015). In the **first step**, a keyword search was performed on *ACM Digital Library*, *AIS Electronic Library*, *EBSCOhost*, *Emerald Insight*, *Google Scholar*, *IEEE Xplore*, *SAGE Publications*, and *SpringerLink*. According to Snyder (2019), the search strings for conducting a literature review should refer to words and concepts that are directly linked to the research questions. Thus, “*augmentative and alternative communication*”, “*artificial intelligence*”, “*sentence generation*”, “*generative artificial intelligence*”, and several combinations and abbreviations thereof were derived as search terms. To obtain even further keywords, iterative keyword refinement was performed (e.g., “*natural language generation*”, “*sentence composition*”, “*large language model*”). Additionally, it was related to the systematic literature review by Konadl et al. (2023), who have identified and assessed the current state of AI approaches applied within AAC systems.

As a result, *1,129 papers* could be initially collected whose titles, abstracts, and thematic foci were then investigated. In general, only papers dealing with the generation of completed utterances based on single words or expressions using an AI-based approach were retained. Papers that included further external devices, which hinder the portability of the AAC systems in daily life situations, were removed. In applying these two inclusion criteria to the *1,129 papers*, several decisions about the inclusion and exclusion of the papers were documented and validated with two additional researchers. To assess the coding and the agreement between the three raters, interrater reliability has been calculated using Fleiss’ Kappa (Fleiss et al. 2013). The rating procedure yielded Fleiss’ Kappa values of 0.8772 for the first inclusion criterion and 0.9072 for the second inclusion criterion, both indicating high levels of agreement (Fleiss et al. 2013). In the very few cases of disagreement, the coding was discussed to come to a consensus about including or excluding the respective paper. Next, duplicates and preprints were eliminated, and only papers written in English were retained. In the **second step**, forward and backward searches have been conducted, resulting in two additional relevant papers. As follows, a total of *19 relevant papers* could be derived.

Subsequently, the contents of these papers were investigated. Using qualitative content analysis and inductive category development, it could be carried out an abstraction from the basic data, while, at the same time, deriving a coherent image of these data (cf. Mayring 2000; Mayring 2002). To inductively extract the DRs from the papers, the notion of what constitutes a DR was applied. In doing so, a DR is considered as any specific condition, capability, or functionality that the intended artifact needs to implement in order to achieve its intended goals and purposes (cf. Möller et al. 2020; Walls et al. 1992). Based on that, the 19 relevant papers have been screened regarding any text statement that hinted at a DR. Several text statements hinting at a DR were consolidated, and based on them, corresponding DRs were deduced. These resulting DRs were then verified with two additional researchers. Thereby, a Fleiss’ Kappa value of 0.8871 could be obtained, indicating a high level of interrater agreement (Fleiss et al. 2013). In case of any disagreement about a DR, the respective DR was analyzed again, followed by discussions until a consensus was reached. Based on the attention and importance that the respective DRs have received within extant research literature, **nine DRs** have been derived that will be described in more detail in the following:

AAC systems enable speech-impaired individuals to articulate their thoughts, ideas, and feelings within daily life conversations. Thereby, if speech-impaired individuals must sense on their own for signs of the respective conversational context, this puts additional cognitive efforts on them. This, in turn, may result in decelerations of the conversational flow. This is underlined by Cognitive Load Theory in that the amount of information that can be processed within the working memory is limited (cf. Sweller 1988). Exceeding the cognitive capacity of speech-impaired individuals could hinder their full concentration on the ongoing

conversation. Therefore, **(DR1)** the AI-based utterance composition approach for AAC systems needs to incorporate an automated identification of formal and informal conversational contexts (e.g., Arnott and Alm 2013; Black et al. 2012; Fang et al. 2023). In this vein, **(DR2)** the applied language model must resemble formal and informal conversational styles to produce sentences for both contexts (e.g., Dempster et al. 2010; Shen et al. 2022; Valencia et al. 2023). There are individual styles of articulation in formal and informal contexts. For example, one may tend to apply in an informal context rather unsophisticated language and expressions (e.g., applying slang and dialect, calling the interlocutor “buddy” or “bro”). Compared to that, others may tend to apply a more reserved wording in an informal conversation (e.g., Cai et al. 2022; Tintarev et al. 2016; Valencia et al. 2023). Therefore, the utterance composition approach **(DR3)** needs to provide means to tailor the utterances towards the personal style of articulation. Furthermore, as Grice (1975) states in the “Maxim of Relation”, it is expected by interlocutors that several conversational contributions are appropriate at each stage. So, it is essential that **(DR4)** the utterance composition is adaptive based on the conversation history and that it incorporates the preceding turns of a conversation (e.g., Kristensson et al. 2020; Mitchell and Sproat 2012; Shen et al. 2022). This also requires **(DR5)** to generate statement and question sentences so that the speech-impaired individuals can apply appropriate conversational acting at different stages of a conversation (i.a., sharing ideas and information with others, asking for information from others) (cf. Grice 1975, and e.g., Black et al. 2012; Garcia et al. 2016; Tintarev et al. 2016). The utterance composition approach **(DR6)** must generate completed utterances that are grammatically correct and semantically tailored to the user input (e.g., Adhikary et al. 2021; Garcia et al. 2016; Mitchell and Sproat 2012). This is necessary to maintain the impression of communicative competency of the speech-impaired individuals towards their interlocutors because these errors transmit to the speech synthesis and are reflected within the voice output. Beyond that, these are functionalities that relieve speech-impaired individuals in composing completed utterances. In this way, they can reduce the mental efforts of speech-impaired individuals and increase their level of participation in conversations (cf. Sweller 1988).

Within AAC systems, it ultimately needs to remain the speech-impaired individual whose thoughts, ideas, and feelings are communicated. Thus, the generated utterances must be semantically tailored towards the words and expressions that the users provide. To ensure that speech-impaired individuals maintain complete control over the generated utterances, **(DR7)** the utterance composition approach must allow them to modify the utterances before they are outputted by means of voice (e.g., Mitchell and Sproat 2012; Shen et al. 2022; Valencia et al. 2023). Speech-impaired individuals using a digital AAC system must be able to keep up with the pace of daily conversations. Therefore, **(DR8)** the generated conversational contributions must be provided immediately. Since conversations are central actions in daily life that consist of multiple interaction steps, **(DR9)** the approach needs to be repeatable at any time (e.g., Arnott and Alm 2013; Cai et al. 2022; Dempster et al. 2010; Fang et al. 2023; Walsh 2010).

Design Requirements	Sources
<b>DR1.</b> Support the automated identification of the conversational context.	Arnott and Alm 2013; Bedrosian et al. 2003; Black et al. 2012; Fang et al. 2023; Hoag et al. 2004; Sweller 1988; Todman and Alm 1997
<b>DR2.</b> Incorporate a language model that can produce utterances that match the style of articulating in the respective conversational contexts.	Bedrosian et al. 2003; Dempster et al. 2010; Hoag et al. 2004; Shen et al. 2022; Tintarev et al. 2016; Todman and Alm 1997; Valencia et al. 2023
<b>DR3.</b> User-adaptive tailoring of the utterances to be generated towards individual styles of articulation and nuances of the speech-impaired individual.	Black et al. 2012; Cai et al. 2022; Grice 1975; Shen et al. 2022; Tintarev et al. 2016; Valencia et al. 2023
<b>DR4.</b> Adaptive sentence generation based on the conversation history.	Black et al. 2012; Dempster et al. 2010; Grice 1975; Kristensson et al. 2020; Mitchell and Sproat 2012; Shen et al. 2022; Sweller 1988
<b>DR5.</b> Support the automated generation of statement sentences and question sentences.	Black et al. 2012; Garcia et al. 2016; Grice 1975; Reddington and Tintarev 2011; Tintarev et al. 2016; Walsh 2010
<b>DR6.</b> Automated generation of completed sentences that are grammatically correct and semantically tailored to the provided user input.	Adhikary et al. 2021; Alm and Higginbotham 2008; Cai et al. 2022; Dempster et al. 2010; Garcia et al. 2016; Grice 1975; Mitchell and Sproat 2012; Shen et al. 2022; Sweller 1988; Valencia 2023
<b>DR7.</b> Provide flexible modification of the generated sentences before they are outputted by voice.	Adhikary et al. 2021; Alm and Higginbotham 2008; Black et al. 2012; Kristensson et al. 2020; Mitchell and Sproat 2012; Shen et al. 2022; Tintarev et al. 2016; Valencia et al. 2023
<b>DR8.</b> Provide generated sentences immediately.	Alm and Higginbotham 2008; Cai et al. 2022; Dempster et al. 2010; Fang et al. 2023; Heo and Kang 2019; Kristensson et al. 2020; Reddington and Tintarev 2011; Shen et al. 2022; Tintarev et al. 2016; Valencia et al. 2023
<b>DR9.</b> Be repeatable and to be triggered by the speech-impaired individual at any time.	Arnott and Alm 2013; Cai et al. 2022; Dempster et al. 2010; Fang et al. 2023; Shen et al. 2022; Valencia et al. 2023; Walsh 2010

**Table 1. Design Requirements of an AI-based Utterance Composition Approach**

## Assessment of Available AAC Systems on the Market

To assess the capabilities of the AAC systems available on the market against the background of the derived DRs, corresponding AAC system providers were first identified. For this purpose, the Google search engine as well as the search engines of the Google Play Store and the Apple App Store were drawn upon. By verifying several steps in the identification of the AAC systems with two additional researchers, the possibility of overlooking established tool providers whose systems offer an automated or an AI-based utterance composition could be tremendously reduced. The resulting set of tool providers was then consolidated and any disagreements about the tools were resolved. In this way, a broad spectrum of providers could be covered, and it could be taken an up-close look at the most established AAC systems in the market. Based on the derived DRs (see Table 1), a tool survey was designed to assess whether the available AAC systems cover the posed requirements. The tool survey was then validated and subsequently applied to the data collection procedure. To gather the data for assessing the AAC systems, sales representatives of the identified AAC system providers were contacted regarding the specific features. Additionally, demo versions of the AAC systems were installed to test them out. Based on the providers' responses and by independently testing the functionalities of the AAC systems against the posed DRs, detailed and comprehensive insights into the capabilities and drawbacks of these AAC systems could be gained. The table below presents the results of the tool assessment for the resulting set of eleven AAC system providers: *Assistiveware*, *Attainment Company*, *Avaz*, *Coughdrop*, *Livox*, *Prentke Romich Company*, *Smartbox*, *Symbotalker*, *Therapybox*, *Tobii Dynavox*, and *Willowtree*. Due to reasons of confidentiality posed by the providers, it was necessary to anonymize their answers. Thus, the providers are labeled below with #1–#11 and their order is randomized.

Design Requirements	Results of the Tool Assessment										
	#1	#2	#3	#4	#5	#6	#7	#8	#9	#10	#11
DR1	-	-	-	-	-	-	○	-	-	-	○
DR2	○	-	-	-	○	○	-	○	-	-	○
DR3	○	-	○	○	-	-	-	-	-	-	-
DR4	-	-	-	-	●	-	-	●	-	-	●
DR5	○	○	○	○	○	-	-	○	-	○	●
DR6	○	○	○	○	●	○	○	●	○	○	●
DR7	●	●	●	●	●	●	●	●	●	●	●
DR8	DR8 has <b>not</b> been included in the comparison because an immediate provision of the completed utterance is also dependent on the respective device the AAC system is applied on.										
DR9	●	●	●	●	●	●	●	●	●	●	●
<b>Table 2. Results of Assessing the Capabilities of Available AAC Systems against the Background of the Derived Design Requirements</b>									<b>Legend:</b> ● = DR is fully met ○ = DR is partly met - = DR is not met		

As has been revealed, there are AAC systems that provide an automated composition of completed sentences. However, essential requirements like context- and user-specific styles of articulation as well as the automated identification of the current conversational context are thereby not met (see DR1–DR3 in Table 2). Without AAC systems that offer speech-impaired individuals completed utterances in both situational and personal styles, while also being morphologically and grammatically correct, their separation from the mainstream of society will continue. Research literature (cf. Konadl et al. 2023) as well as the affected individuals, caregivers, and representatives (see Section “Introduction”) clearly point out the necessity of AAC systems that support both formal and informal contexts with appropriately-styled utterances. To contribute to dismantling the communication barriers that speech-impaired individuals still suffer from across various daily life conversations, this research aims to close this gap.

## Research Procedure

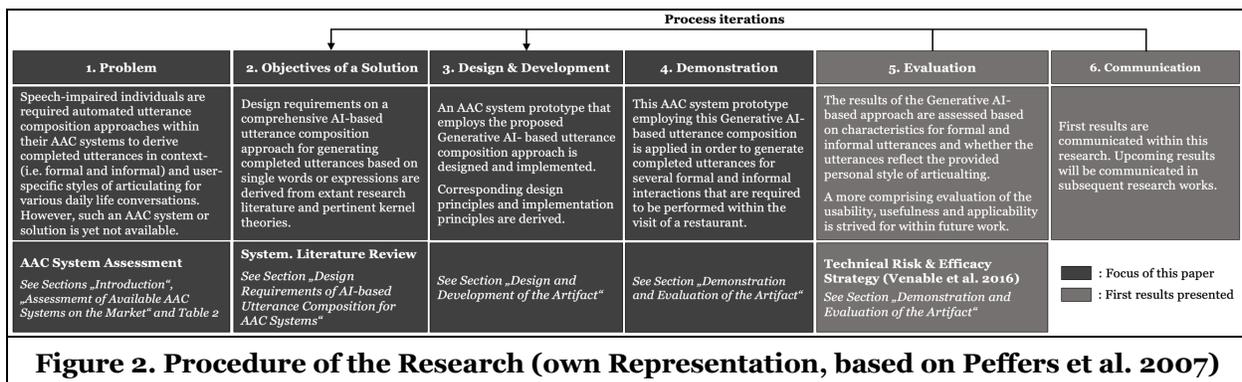
To accomplish the development of the approach that enables speech-impaired individuals to generate completed utterances in both formal and informal styles within an AAC system, Design Science Research (DSR) was applied (cf. Gregor and Hevner 2013; Hevner et al. 2004). In doing so, the procedure outlined by Peffers et al. (2007) was followed (see Figure 2).

As a first step, **(1) corresponding problems** of existing software tools and approaches were identified (see Section “Introduction” and Table 2). Available AAC systems are not capable of generating completed sentences for formal and informal conversational contexts. However, covering both contexts within AAC systems is indispensable for paving the way for speech-impaired individuals into the mainstream of society and for enabling them context- and user-specific articulation. Thus, the **(2) objective** is to combine, based

on the DRs, a set of machine learning techniques. Thereby, the Generative AI technique ChatGPT-4.0 is purposefully integrated to facilitate a comprehensive composition of completed utterances from single expressions as speech-impaired individuals provide them within AAC systems.

The third step contains the **(3) design and development** of the solution (see Section “Design and Development of the Artifact”). By relating to extant theory (Grice 1975; Sweller 1988) and by conducting a systematic literature review (cf. Vom Brocke et al. 2015), there could be identified the DRs the AI-based approach needs to meet. To close the gaps identified within step (1), it was focused on the design and technical realization of the Generative AI-based utterance composition. Beyond that, it was developed an AAC system prototype that employs this proposed AI-based approach. Regarding the design of the approach, there have been established DPs that are founded on the identified DRs. Thus, the artifact was established to provide speech-impaired individuals with a comprehensive utterance composition, enabling user-specific articulation and utterances that fit the expected style of articulation in formal and informal contexts.

**Steps (4) and (5)** deal with the **demonstration and evaluation** of the generated artifact. First, a proper setting was defined for rigorously demonstrating and evaluating the utility of the artifact. In this regard, it was related to Venable et al. (2016) who defined four evaluation strategies for DSR projects. In line with that, the DSR-project-specific evaluation strategy was aligned to the “Technical Risk & Efficacy Strategy” to rigorously demonstrate that the evaluand achieves its intended purpose (i.e., composing single expressions into completed utterances in a style that fits formal and informal conversational contexts and that reflects personal nuances of articulation). In addition, the aim is to show that meeting all the derived DRs (see Table 1) and purposefully harnessing ChatGPT-4.0 as a Generative AI technique are thereby imperative (cf. March and Smith 1995; Peffers et al. 2012). Thus, it was decided to apply one summative evaluation episode (cf. Venable et al. 2016). In doing so, the developed AAC system prototype equipped with the AI-based approach was applied to 61 communication activities of the common daily life routine of visiting a restaurant with a companion (cf. Ehlich and Rehbein 1972; Schindler 2013) (see Figure 1). The composed utterances were then assessed based on characteristics for the formal and informal context of the German language (cf. Budde 2012; Elspaß 2018) and whether the provided individual nuances of articulation are reflected within the generated sentences. Beyond that, it is planned to apply the “Human Risk & Effectiveness Strategy” (cf. Venable et al. 2016) and to investigate the constructed artifact within a large-scale empirical study in a controlled environment (e.g., within a laboratory experiment). In **step (6)**, the aim is to **document and communicate** the results of the research at hand.



With this procedure, the research also aligns with the guideline of Hevner et al. (2004). Regarding the design cycle, the artifact is presented as the result that has gone through the process of demonstration and evaluation (see Section “Demonstration and Evaluation of the Artifact”). Regarding the relevance cycle, several DRs from research literature and theory are identified (see Table 1) that guided the design of the artifact. In this way, the practical application of the generated artifact brought up several contributions for practice (see Section “Contributions to Theory and Practice”). In view of the rigor cycle, several techniques (ChatGPT-4.0, automated speech recognition, natural language processing) were applied to rigorously construct the artifact. There have been derived initial findings as contributions to theory, including both kernel theory (Gricean Maxims of Communication, Cognitive Load Theory) and design theory. To contribute to a rather general and abstract knowledge base – “Nascent design theory” (Gregor and Hevner 2013) – and to design a purposeful artifact in a comprehensible way, there have been deduced meta requirements and DRs (Gregor and Jones 2007; Heinrich and Schwabe 2014) for an automated (AI-based) utterance composition for AAC systems grounded in seminal works, which resulted in the DPs.

## Design and Development of the Artifact

### Design of the AI-Based Utterance Composition for AAC Systems

The meta requirements describe “*what the system is for*” (Gregor and Jones 2007, p. 325). In this way, meta requirements constitute solution objectives that are based on the clearly defined purpose and scope of the artifact and thus, are based on the problems the investigation addresses as discussed in the introduction. Therefore, it is for example identified as one meta requirement the objective of the proposed AI-based approach to empower speech-impaired individuals to express their needs, wishes, and wants by means of voice. Thereby, these meta requirements to be suitable for a class of artifacts are derived based on the current research literature (Gregor and Jones 2007; Heinrich and Schwabe 2014; Walls et al. 1992).

Subsequently, DPs have been synthesized based on the DRs that have been derived from related work (i.e., extant literature) and based on DRs that are inspired by theory (cf. Möller et al. 2020; Puroo et al. 2020). These DRs constitute specific conditions, capabilities, and functionalities that an artifact needs to implement in order to enable its intended users to fulfill certain tasks within a specific problem class (cf. Möller et al. 2020; Walls et al. 1992). The DRs that semantically correspond to each other constitute conceptual influences that need to be considered when proposing a DP. Thus, the derived DRs were carefully examined and DRs that are semantically related to each other have been consolidated into a respective DP (cf. Möller et al. 2020; Puroo et al. 2020). For example, DP3 was deduced based on DR2, DR6, and DR8, as it is necessary to automatically provide sentences that are both completed and correct. Hence, the applied language model needs to comprise both capabilities. In addition, because it is essential for the composition approach to provide the completed sentences immediately, this aspect was also considered within this DP.

These DPs are defined as generic and prescriptive statements, showing how to do something to achieve a certain goal (Gregor et al. 2020). In this way, the proposed DPs foster the transfer and the application of the design-related knowledge to similar problem classes beyond the specific problem class of Generative AI-based utterance composition within AAC systems. The disposed DPs fall into the category of action and materiality-oriented DPs that describe what an artifact should enable users to do and how the artifact should be built in order to do so (Chandra et al. 2015). Speech-impaired individuals (= users) are required to compose single expressions into completed sentences to communicate and collaborate within formal and informal contexts in daily life situations (= boundary conditions). Keeping the DRs of the artifact in mind, there have been derived **five DPs** regarding the proposed Generative AI-based approach.

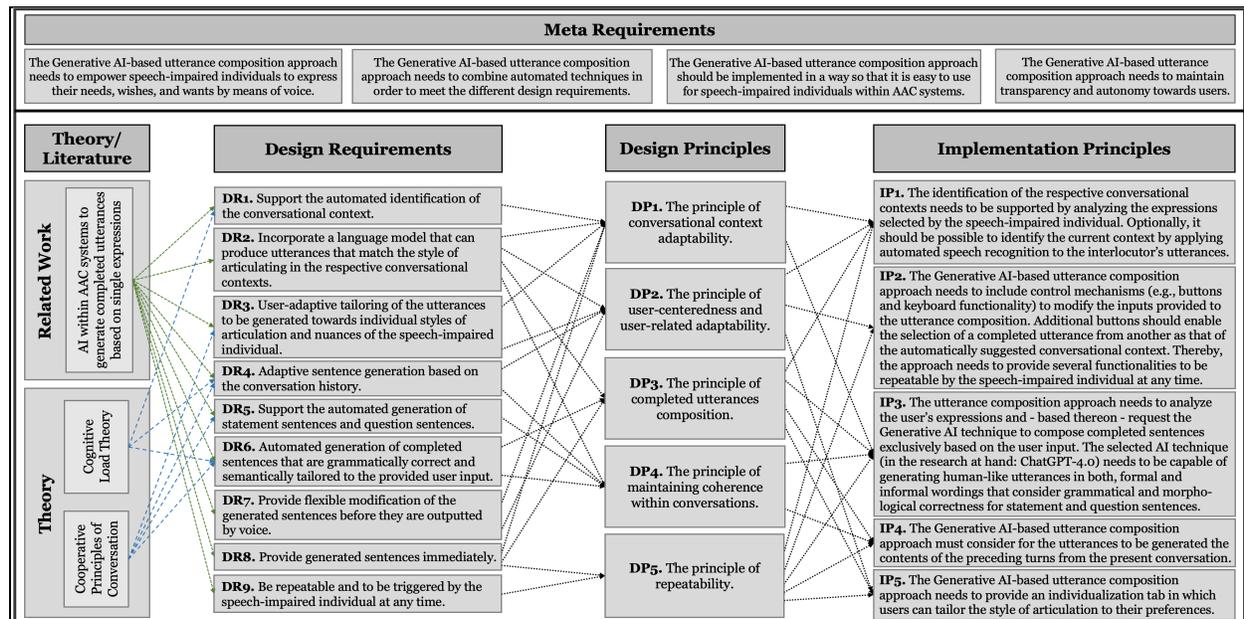


Figure 3. Derived Design of the Generative AI-Based Utterance Composition Approach

- **DP1: The principle of conversational context adaptability.** To enable speech-impaired individuals to apply situation-adaptive wordings within conversations, the Generative AI-based utterance comp-

position for an AAC system needs to support the tailoring towards the respective conversational context by sensing and proposing conversational contributions for the current context.

- **DP2: The principle of user-centeredness and user-related adaptability.** To consider individual styles of articulation and to give speech-impaired individuals control regarding the AI-generated conversational contributions, it is necessary to enable speech-impaired individuals to include personal nuances and expressions into the Generative AI-based utterance composition. Thereby, the approach needs to provide flexible means to modify a generated sentence before it is outputted by means of vocal language.
- **DP3: The principle of completed utterances composition.** The Generative AI-based utterance composition approach needs to include the user-selected, single expressions and form them into completed sentences by harnessing the relationships from a (large) language model. By that, these models need to cover various daily life situations, daily communication activities, and conversational contexts.
- **DP4: The principle of maintaining coherence within conversations.** To foster communicative effectiveness, the Generative AI-based utterance composition needs to take account of correctness (i.a., morphological, grammatical, spelling) and situational fit of the generated contributions (i.a., relatedness to previous utterances in a conversation, proposing the required sentence type).
- **DP5: The principle of repeatability.** As conversations consist of multiple interaction steps, the Generative AI-based utterance composition needs to provide repetition of several functionalities at any time and to enable speech-impaired individuals to immediately act and react within conversations.

These DPs are deduced from the DRs that are based on current research literature and theory. Gregor and Jones (2007) state that this reference to theory and literature discloses “*an explanation of why an artifact is constructed as it is and why it works*” (p. 328). Based on related work and theory, DRs were derived that an AI-based utterance composition approach to be employed in AAC systems should meet. These DRs offer guidance by designing the artifact and advising the DPs (Böckle et al. 2021; Gregor and Jones 2007). The DPs refer to at least one DR and serve as an abstract “blueprint” of the artifact (Böckle et al. 2021; Gregor and Jones 2007; Heinrich and Schwabe 2014). By establishing the DPs, it is ensured that they follow the value grounding (reference to the DRs) and explanatory grounding (DPs are based on current literature and kernel theories) (Heinrich and Schwabe 2014). Furthermore, based on the instantiation of the DRs and DPs within the generated artifact, Implementation Principles (IPs) are defined. In this way, it is supported “*the implementation in practice of an abstract, generic design method or development approach*” (Gregor and Jones 2007, p. 329) in view of an artifact for (Generative) AI-based utterance composition for AAC systems.

### **Technical Realization**

To address the drawbacks of prevailing research and existing AAC systems, the artifact was designed and developed in the programming language Python. There has resulted the Generative AI-based utterance composition approach. This proposed AI-based approach has been integrated into a more comprehensive AAC system prototype, enabling users to apply and interact with the approach. In the following, the instantiation of the DRs is described, as they depict the required features of the proposed AI-based approach in the most detailed way. Since the DPs and IPs depict a generic and prescriptive statement of how something should be done, they capture design-related knowledge. In this way, these IPs support the development of further IS (design) theories and new artifacts focusing on AI-based utterance composition for AAC systems. To reveal the association of the derived DPs and IPs to the DRs, their relations are pointed out as well.

Regarding the adaption of utterances to the current context within a conversation, it is essential to include a functionality within the AI-based approach that automatically identifies the conversational context. A technique that has been applied for this task within extant AAC research is automated speech recognition (e.g., Higginbotham et al. 2012; Kane et al. 2017; Neamtu et al. 2019). The Google cloud speech-to-text engine was chosen because of its impressive capabilities in correctly recognizing German expressions and sentences (Filippidou and Moussiades 2020). It records the utterances of the interlocutor and analyzes them for signs of formal (i.a., formal terms of addressing the interlocutor, addressing him/her with the last name) and informal expressions (i.a., informal terms of addressing the interlocutor, addressing him/her with the first name) (cf. Budde 2012; Elspaß 2018). Thereby, a manually defined list of common German expressions for the formal and informal contexts is applied. Automated speech recognition is yet not purposeful in every situation (e.g., noisy locations). Hence, a second functionality for identifying the context is implemented that applies common German expressions for formal and informal contexts to the expressions that the speech-impaired individual has selected (**DR1/DP1/IP1**). In this vein, it is mandatory that the harnessed language model can generate contributions for the respective conversational contexts. ChatGPT-

4.0 (OpenAI 2024) was chosen as its large language model comprises data on conversations from various daily life situations. Prompt engineering was applied to create and iteratively refine the input given to ChatGPT-4.0 until eliciting the desired task behavior (Teubner et al. 2023). ChatGPT-4.0 is thereby employed in the few-shot learning mode (Feuerriegel et al. 2023), as it is provided with examples of characteristics for formal and informal utterances of the German language (cf. Budde 2012; Elspaß 2018). In this way, it is aimed that the artifact reflects the provided characteristics for the formal and informal contexts within the generated utterances. As a result, it is derived a prompt that is sent to GPT-4.0 via its API, which makes ChatGPT-4.0 compose a completed utterance from the provided expressions in the wording for the respective conversational context (**DR2/DP1/IP1**, **DR2/DP3/IP3**). Conversational styles can contain specific nuances and expressions that vary for different contexts and individuals. For that reason, the generated artifact includes a separate tab area in which the users can enter expressions they are keen on applying given certain conversational contexts and situations. These expressions are incorporated into the prompts which are sent to GPT-4.0 for composing the utterances (**DR3/DP2/IP5**).

Low coherence of conversational contributions could lead to deviations in the course of a conversation. Speech-impaired individuals could be required to provide information and to ask questions when they lack information. Spelling errors as well as morphological and grammatical errors are directly reflected within the voice output, hindering the perception of communicative competency by others towards the speech-impaired individuals. To reflect their thoughts, ideas, and feelings, the generated utterances must be semantically tailored to the inputs that the speech-impaired individuals provide. These three requirements can all be met by the capabilities of ChatGPT-4.0. It maintains the coherence of the generated contributions as it automatically relates the contents of the utterances to be generated to the themes of the preceding turns within the ongoing conversation (**DR4/DP4/IP4**). For generating questions, the GUI includes a question mark symbol that adds via click a question mark to the speech composition line (**DR5/DP4/IP3**). The large language model applied within ChatGPT-4.0 is trained on large amounts of text data that have comprised statement and question sentences. In addition, the training procedure has adhered to the rules of correct spelling, grammar, and morphology (cf. Radford et al. 2019). Regarding semantically tailoring the generated sentences to the user input, the user-selected expressions are directly integrated to be essential units of the utterance to be composed within the prompt for requesting the utterance composition (**DR6/DP3/IP4**). In case that the generated sentences digress from the user input even after multiple repetitions of the generation procedure, the artifact provides the user with flexible and multiple means to adapt the contributions before voice output (**DR7/DP2/IP2**).

The generated utterances must be provided immediately to the speech-impaired individuals. Here, the proposed artifact implements multi-threading and immediate processing of information, resulting in an instant provision of the generated sentences. Furthermore, the generated conversational contributions for the detected context are directly inserted into the speech output line, fostering their immediate application (**DR8/DP1/IP1**). Conversations consist of multiple interaction steps so that several functionalities can be triggered and repeated by the speech-impaired individual any number of times (**DR9/DP5/IP1–5**).

The derived DPs and IPs as prescriptive design-related knowledge have all been considered within a concrete instantiation. Consequently, a developed AAC system prototype has resulted that employs the Generative AI-based utterance composition. Thus, the artifact is capable of providing for speech-impaired individuals context- and user-specific styles of articulation across various daily life conversations. To show the applicability of this instantiation, the next section proceeds with its demonstration and evaluation.

## Demonstration and Evaluation of the Artifact

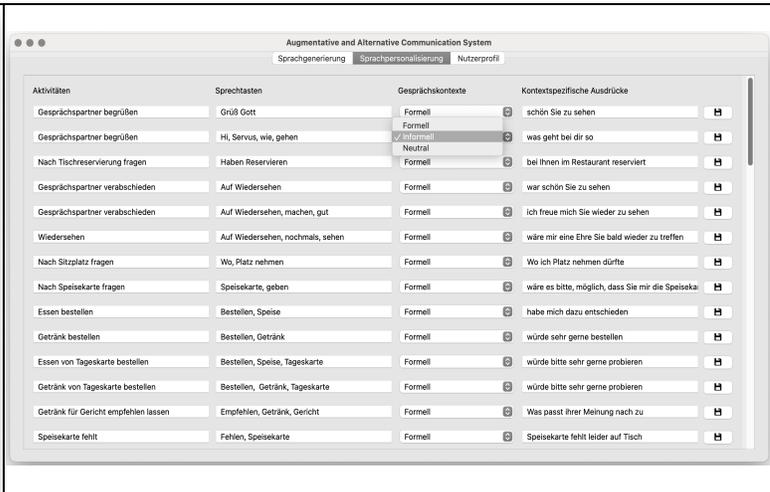
The demonstration and evaluation were aligned to the defined evaluation strategy (see Section “Research Procedure”) and to the routine of visiting a restaurant with a companion (as outlined in Figure 1). This situation was chosen as it is a commonly performed daily life activity, and collaborating and engaging within it requires individuals to perform several formal and informal interactions (cf. Ehlich and Rehbein 1972; Schindler 2013). Beyond that, it has also received considerable attention within extant AAC research (e.g., Kane et al. 2017; Obiorah et al. 2021; Valencia et al. 2023). In the following, there is described (a) the handling of the proposed AI-based approach. Then (b), the applied metrics are elaborated, before (c) relating to the two iterations back to the phase “Design & Development” and (d) to the reporting of the results.

(a). To cover a comprehensive set of communication activities, 121 German vocabulary entries were defined (i.e., conversation openers and closers, foods and drinks, daily activities and situations). Clicking the

button of a vocabulary entry leads to its insertion into the speech composition line. To detect the conversational context, the user can apply the speech recognition functionality to analyze the interlocutor’s utterance (see Figure 4). The “Sprachpersonalisierung” (speech personalization) tab (see Figure 5), enables the provision of personal nuances and expressions of articulation. After making any modifications to the selected expressions within the speech composition line (see Figure 4), the proposed AI-based approach can be triggered by clicking the “Satz generieren” (generate sentence) button. As a result, a completed sentence is generated and proposed for the current conversational context within the speech output line, so that it can be immediately applied. If another context is considered more suitable in a certain situation, the corresponding utterance for this context can be flexibly chosen. By clicking the speech output button, the sentence in the speech output line is voiced in the user-selected volume and speed.



**Figure 4. Speech Generation and Speech Output View**



**Figure 5. Speech and Articulation Personalization View within the proposed Artifact**

(b). Based on the 121 German vocabulary entries, 61 communication activities were defined for the daily life routine of visiting a restaurant with a companion (e.g., Ehlich and Rehbein 1972; Kane et al. 2017; Obiorah et al. 2021; Schindler 2013; Valencia et al. 2023). These 61 communication activities served as the basis for assessing the capabilities of the proposed AI-based approach (see Table 3 which illustrates the capabilities of the approach on a subset of nine communication activities from the total set of 61). On the one hand, it was investigated whether the peculiarities for articulating in formal and informal contexts are reflected within the composed sentences. For that reason, for both conversational contexts, the true positive and false positive rates are reported. The true positive rate counts the number of times in which the respective context is correctly reflected within the generated sentences. In contrast, the false positive rate reports the number of times in which the Generative AI-based approach generates sentences in the formal or informal context, despite it is expected to generate a sentence for the opposite context. The generated sentences were analyzed to determine whether they contained any characteristics of the formal or informal context (cf. Budde 2012; Elspaß 2018). Based on the expected and the determined contexts, the number of times the sentences reflected the correct or incorrect context was counted. On the other hand, it was investigated whether the generated sentences comprised a completed structure (i.e., correct word orders for statement and question sentences) and whether they were morphologically and grammatically correct (i.e., correct conjugation of verbs, correctly inserted filling words, and prepositions). Consequently, the completeness of the structure of the composed sentences and whether they adhered to the themes of previous conversational turns were also investigated. If any errors regarding these criteria have been observed within the utterances composed by the Generative AI-based approach, the corresponding number of errors is reported in (d).

Communication Activities	User Input (German)	GPT-Generated Sentences (German)	Translation to English
(I) Welcoming companion	Hallo, wie, gehen	Hi, wie geht's dir denn so? (Informal)	Hi, how are you doing?
(II) Asking for the booked seats	Wir, haben, reservieren	Grüß Gott, wir haben bei Ihnen im Restaurant einen Tisch reserviert. (Formal)	Excuse me, we have booked a table at your restaurant.
(III) Telling about recent events in one's life	Arbeit, stressig	Ich hab' aktuell echt gut zu tun in der Arbeit. (Informal)	There is currently much stress for me at work.
(IV) Asking for the menu card	Speisekarte, geben, ?	Entschuldigen Sie, aber könnten Sie mir bitte eine Speisekarte bringen? (Formal)	Excuse me, but could you please bring me a menu card?

(V) Telling personal narratives	Ich, fahren, bald, Urlaub, mit, Freunden	Ich werde bald mit meinen Kumpels in den Urlaub fahren. (Informal)	I will soon go on vacation with my buddies.
(VI) Ordering meals and drinks	Ich, Bestellen, Fleisch, Tageskarte	Ich würde bitte gerne das Fleischgericht von der Tageskarte bestellen. (Formal)	I would like to order the meat dish from the daily menu.
	Nehmen, Beilage, Pommes	Als Beilage würde ich bitte die Pommes nehmen. (Formal)	As a side I would appreciate to take the fries.
(VII) Taking over the payment of the bill	Ich, heute, übernehmen, Rechnung	Die Rechnung geht heut auf mich. (Informal)	Today's bill is on me.
(VIII) Asking for the bill and for credit card payment	Rechnung, bringen, ?, Ich, bezahlen, mit, Kreditkarte	Entschuldigen Sie, könnten Sie mir bitte die Rechnung bringen? Ich würde bitte gerne mit Kreditkarte bezahlen. (Formal)	Excuse me, could you please bring me the bill? I would like to pay by credit card please.
(IX) Farewelling companion	Tschüss, machen, gut, nochmals, sehen	Tschüss, mach's gut ich freu' mich dich bald wieder zu sehen. (Informal)	Bye, take care, I look forward to seeing you again soon.

**Table 3. Utterances Generated by the Generative AI-Based Composition Approach**

(c). During the demonstration and evaluation, there were observed cases where the prompt to compose a completed utterance based on the provided expressions was not followed correctly. Therefore, **two iterations back to the phase “Design & Development” within the DSR procedure** were performed. It turned out that, among other things, for the communication activities of greeting the conversation partner or ordering food or drinks, templates for writing a letter or instructions on how to perform these tasks have been harnessed. Instead, the prompt was expected to lead to the composition of a sentence to be uttered towards an interlocutor. In these cases, the extensive scope of the large language model employed by ChatGPT-4.0 seems leading to other relationships from the large language model being harnessed instead of strictly following this prompt. To mitigate this, it was implemented a semantical framing to limit the scope that is harnessed from the large language model. For each of the 61 communication activities, there was defined a specific prompt that is selected based on the expressions and single words that are commonly applied in the respective communication activity (e.g., the verb “order” within the selected expressions leads to a specific prompt for the respective communication activity and prepares ChatGPT-4.0 that the utterance to be generated is about ordering something). Since there may be communication activities that are not previously framed, it was added a default prompt to provide rudimentary utterance composition also for these cases. Proceeding with the *61 framed communication activities*, there were tested several variants of individual nuances and expressions of articulation in formal and informal conversational contexts. In doing so, it turned out that when individual nuances are semantically close to already selected vocabulary entries, this leads to duplicates within the generated utterances. Thus, it was implemented an automated mechanism that clears these duplicates. Employing these two iterations, the Generative AI-based approach could be improved, achieving accurate operation and properly meeting the posed DRs.

(d). In the following, this improved utterance composition approach was applied to the set of the *61 framed communication activities*. It was thereby found that in utterances for the formal context, the interlocutor is generally addressed by a formal term of address (cf. “Grüß Gott”, “Ihnen”, “Sie” in entries (II), (IV) and (VIII) in Table 3). These reflect more elaborated language and constructs like euphemisms, forms of courtesy, as well as a clear and precise terminology (cf. the German entries (VI) and (VIII) in Table 3). In contrast, the utterances for the informal context consistently address the interlocutor by an informal term (cf. “Hi”, “Tschüss” in entries (I) and (IX) in Table 3), contain contractions (cf. “geht’s”, “mach’s” in entries (I) and (IX) in Table 3), and in general a rather unsophisticated style of articulation (cf. “Ich hab’ aktuell echt gut zu tun in der Arbeit”, “Die Rechnung geht heut auf mich” in entries (III) and (VII) in Table 3). Regarding the formal context, three sentences contained characteristics of the informal context and were therefore incorrect. This results in a true positive rate of 58/61 and a false positive rate of 3/61. For the informal context, four generated sentences contained characteristics of the formal context, although these sentences were expected to be informal. This results in a true positive rate of 57/61 and a false positive rate of 4/61. Beyond that, it was assessed whether the proposed AI-based approach mixes both contexts. The Generative AI-based utterance composition was found to consistently adhere to the conversational style it was requested. Furthermore, all the provided individual nuances and expressions of articulation have been correctly considered within the generated sentences.

All in all, these results provide convincing evidence that the Generative AI-based approach can generate utterances for both conversational contexts, thereby confirming the general feasibility of the approach. Several utterances comprised a completed structure for the respective sentence types and were morphologically and grammatically correct. Only two minor grammatical errors could be observed as two verbs were not correctly conjugated. Adding a question mark to the speech composition line leads to generating questions instead of statement sentences (cf. entries (IV) and (VIII) within the “User Input” column of Table 3).

Except for three minor errors, the utterances consistently adhered to the themes of previous conversational turns. For example, when ordering food (cf. entry (VI) in Table 3), the meaning of the verb "take" is correctly considered given the previous utterance. The Generative AI-based approach automatically concludes that this means ordering fries as a side dish and not taking the sides in the sense of grabbing them.

Generally, the development of the artifact was based on the DRs, which are all met and technically realized (see Figure 3). The proposed AI-based approach was verified with 61 formal and informal communication activities of visiting a restaurant with a companion, validating its functionality and practical applicability.

## **Discussion**

### ***Contributions to Theory and Practice***

The implementation of the DRs and DPs has provided interesting results. ChatGPT-4.0 as a Generative AI technique turned out to be powerful to generate completed utterances from single words and expressions as speech-impaired provide them using their AAC systems. This not only leads to conversational contributions that are more quickly available within a conversation, but also relieves speech-impaired individuals of cognitive efforts when producing completed and grammatically correct sentences (i.e., inserting prepositions, conjugating verbs). As the demonstration and evaluation has shown, the generated conversational contributions consider the content and semantics of preceding turns within a conversation. Beyond that, the proposed AI-based approach can flexibly generate different sentence types (statement and question sentences). This empowers speech-impaired individuals to convey information when they are required and to ask questions when they need information within a conversation.

Even the fastest way of expressing is of little help if the utterances that an AAC system proposes are perceived as distracting by speech-impaired individuals and their interlocutors. The style of an utterance may not fit the expected way of articulation in a formal or an informal conversational context. As a contribution to practice, the proposed AI-based approach can suggest utterances for socialness-oriented (i.e., informal) as well as informativeness-oriented (i.e., formal) contexts (cf. Bedrosian et al. 2003; Hoag et al. 2004; Todman and Alm 1997). As Table 3 shows, the generated sentences clearly reflect wordings that match the common style of articulation of the German language (Budde 2012; Elspaß 2018). Moreover, the approach enables reflecting individual nuances one is keen on applying in formal and informal situations. This enables speech-impaired individuals to implement within utterances for formal situations the demanded level of contenance (i.e., non-familiar individuals or superiors may attach importance to maintaining distance because of being unfamiliar or due to their position). At the same time, they can demonstrate within their utterances for the informal contexts their closeness to familiar people like friends or family members.

As a further outcome of this DSR project, the investigation achieved theoretical contributions that go beyond the technical contribution (i.e., the artifact). The Generative AI-based approach and its capability of proposing completed utterances in a situation-specific articulation style empower speech-impaired individuals to comply with the Maxim of Modality. By complying with this Maxim, misunderstandings within a conversation can be reduced, leading to maintaining the flow of the conversation. Consequently, speech-impaired individuals will experience fewer breakdowns, making it easier for them to convey their thoughts, to be understood, and boosting their confidence in engaging with others using their AAC systems. Additionally, the compilation of the Generative AI-based utterance composition points out that there are deeply intertwined relationships of the Modality Maxim with the other Maxims of Grice (1975) in the AAC system context. The increased expressiveness provided by the proposed artifact could lead to speech-impaired individuals (deliberately) violating the Maxim of Relevancy or the Maxim of Quality with utterances that are not meaningful or not correct. Thus, investigating the opposing and amplifying effects of the Gricean Maxims in empirical settings could suggest refinements to context-specific communication styles, and insights into how speech-impaired individuals apply the Maxims (i.e., within negotiations with interlocutors).

Moreover, the research offers promising contributions to Cognitive Load Theory (cf. Sweller 1988). This theory addresses, among other things, the design of software system functionalities (cf. Hollender et al. 2010; Westphal et al. 2023). In the design and development of the proposed Generative AI-based approach, various AI and machine learning techniques (ChatGPT-4.0, automated speech recognition, natural language processing) were purposefully combined. In this way, speech-impaired individuals can be relieved in obtaining completed utterances that reflect situation- and user-specific styles of articulation. However, in light of the Cognitive Load Theory, it is not obvious in which circumstances speech-impaired individuals

will perceive this as a relief. The implemented supportive functionalities can take over several tasks typically performed during a conversation. But AAC systems that propose speech-impaired individuals context- and user-specific utterances exert for them also mental effort and cognitive load. Investigating the interaction effects of relief and cognitive load in empirical settings could suggest when the proposed AI-based approach should provide how much support. For example, there could be conversations where less support is deliberately preferred. Manually composing utterances can be harnessed as a means to demonstrate knowledge about certain topics to interlocutors (Valencia et al. 2023). Considering different situation-specific levels of support within the proposed AI-based approach could hence foster even more need-fitting communication.

The discussed results have shown the advantages of the proposed Generative AI-based approach for AAC systems compared to existing ones. Therefore, the DSR project is positioned in the group of improvements (new solutions for known problems) in terms of the DSR knowledge contribution framework of Gregor and Hevner (2013). DSR improvement projects contribute to both prescriptive theory, i.e., design theory (Gregor 2006), and descriptive theory, i.e., kernel theories such as the Gricean Maxims and Cognitive Load Theory as described above (cf. Gregor and Hevner 2013). By including ChatGPT-4.0 as a Generative AI technique within the utterance composition approach for AAC systems, this research also affects the quadrant of exaptations (known solutions extended to new problems) (Gregor and Hevner 2013). It is shown the proficiency of this technique for composing single expressions into completed utterances in different (formal and informal) context-specific styles. In this way, contributions to IS and DSR research are made alike. GPTs and Generative AI are proposed and proven as a suitable technique that enables situation- and user-specific styles of articulation for speech-impaired individuals.

Based on the DRs derived from literature and kernel theories, DPs were proposed. By applying them during the design and development of the artifact, followed by the demonstration and evaluation, an implicit empirical grounding of the DPs was achieved (Heinrich and Schwabe 2014). The DPs capture design-related knowledge and can thus support the development of further IS (design) theories and new artifacts. For example, by considering the principle of user-centeredness and user-related adaptability (DP2), the importance of including user-specific nuances into an AI-based utterance composition is highlighted. Since these expressions are directly reflected within the generated utterances, the alignment with user-centeredness leads to the communication being more personal and individualized for speech-impaired individuals. With the DPs, a first step is made towards contributing to design theory in terms of theory for design and action (Gregor 2006). It is complied with the conditions of March and Smith (1995) and Hevner et al. (2004), pointing out conditions in which a contribution to knowledge in design science has occurred: utility to a community of users, the novelty of the artifact and the persuasiveness of claims that it is effective.

### ***Limitations and Opportunities for Future Research***

There are also limitations to this research. Although this research encompasses a large set of investigations, researchers from other fields could probably identify even more DRs. However, the derived DRs are undoubtedly important for AI-based utterance composition within AAC systems. While this research proposes promising contributions to the Gricean Maxims and to Cognitive Load Theory, more in-depth investigations will suggest even more mature contributions.

To take the next steps toward a more mature design theory, the proposed AI-based approach will be evaluated in a formative environment (e.g., a laboratory experiment). Thereby, a homogenous sample of speech-impaired individuals will perform different formal and informal interactions with their companions and the staff of a restaurant. This allows for observing how speech-impaired individuals collaborate with the approach. Thereby, improvements can be made to it (whereby the DPs can be confirmed or adapted), before conducting a more elaborate evaluation in a natural setting (i.e., a field study in a restaurant) as a further part of the design cycle. There are several avenues for future research, for which the proposed AI-based approach constitutes the first essential step. Future studies can investigate the specific circumstances that lead to exceeding the cognitive load of speech-impaired individuals and avoiding context-specific articulation. Additionally, investigating in more detail the transferability of the Gricean Maxims to AAC-supported communication could shape daily life conversations for speech-impaired individuals even more inclusively.

### **Conclusion**

To enable speech-impaired individuals to plenary participation within daily life conversations, they must be empowered to communicate in formal and informal conversational contexts in an appropriate, situation-

and user-specific style. However, prior literature and existing solutions for automated utterance composition do not sufficiently cover the informal and formal contexts. Beyond that, the specific requirements deemed crucial within extant AAC literature for AI-based utterance composition have not been comprehensively met. Hence, this research has derived several DRs and DPs. To technically realize the DPs, several AI and machine learning techniques (ChatGPT-4.0, automated speech recognition, natural language processing) were purposefully combined into an automated utterance composition approach to close the revealed gaps (see RQ1 and RQ2). Demonstrating and evaluating a developed AAC system prototype employing this approach to 61 formal and informal communication activities of visiting a restaurant has indicated that covering all the derived DRs is essential. This coverage is crucial for ensuring that the proposed Generative AI-based approach is comprehensive and purposeful. To enable speech-impaired individuals to use context- and user-specific styles at the lowest possible level of effort, the approach integrates further supportive functionalities like an automated identification of the current conversational context. In this way, it fosters the immediate application of the utterances that match the current context of a conversation.

The investigation contributes to practice and research alike. The generated artifact empowers speech-impaired individuals to increase their expressiveness using formal and informal conversational contributions from their AAC systems and thus dismantling communication barriers that speech-impaired individuals suffer from. The research also highlights first contributions to design theory and kernel theory. Regarding the Gricean Maxims of Communication (Grice 1975), the compilation of the Generative AI-based utterance composition could reveal for AAC-supported communication intertwined relationships of the Modality Maxim with the other Maxims. Moreover, the research provides the basis to enrich Cognitive Load Theory based on the design and the concrete instantiation of the Generative AI-based approach. In this way, the proposed artifact is expected to relieve speech-impaired individuals from the cognitive load by obtaining coherent and grammatically correct utterances in context- and user-specific wordings.

## References

- Adhikary, S., Berger, J., and Vertanen, K. 2021. "Accelerating text communication via abbreviated sentence input," *Proceedings of the 59<sup>th</sup> Annual Meeting of the Association for Computational Linguistics*.
- Allen, J. 2005. "Designing desirability in an augmentative and alternative communication device," *Universal Access in the Information Society* (4:2), pp. 135–145.
- Alm, N., and Higginbotham, J. 2008. "Whole utterance approaches in AAC," *Augmentative and Alternative Communication* (24:3), pp. 235–254.
- Arnott, J., and Alm, N. 2013. "Towards the improvement of augmentative and alternative communication through the modelling of conversation," *Computer Speech & Language* (27:6), pp. 1194–1211.
- Baldassarri, S., Marco, J., Cerezo, E., and Moreno, L. 2014. "Accessibility evaluation of an alternative and augmentative communication (AAC) tool," *Proceedings of the 2014 International Conference on Universal Access in Human-Computer Interaction*, pp. 529–540.
- Bedrosian, J., Hoag, L., and McCoy, K. 2003. "Relevance and speed of message delivery tradeoffs in augmentative and alternative communication," *Journal of Speech Language and Hearing Research* (46).
- Black, R., Waller, A., Turner, R., and Reiter, E. 2012. "Supporting personal narrative for children with complex communication needs," *ACM Transactions on Computer-Human Interaction* (19:2), pp. 1–35.
- Böckle, M., Bick, M., and Novak, J. 2021. "Toward a Design Theory of User-Centered Score Mechanics for Gamified Competency Development," *Information Systems Management* (40:1), pp. 1–27.
- Bradshaw, J. 2013. "The use of augmentative and alternative communication apps for the iPad, iPod and iPhone: an overview of recent developments," *Tizard Learning Disability Review* (18:1), pp. 31–37.
- Budde, M. 2012. "Über Sprache reflektieren," *Kassel University Press*.
- Cai, S., Venugopalan, S., Tomanek, K., Narayanan, A., Morris, M., and Brenner, M. 2022. "Context-Aware Abbreviation Expansion Using Large Language Models," *Proceedings of the 2022 Conference of the North American Chapter of the Association for Computational Linguistics*.
- Chandra, L., Seidel, S., and Gregor, S. 2015. "Prescriptive Knowledge in IS Research: Conceptualizing Design Principles in Terms of Materiality, Action, and Boundary Conditions," *Proceedings of the 48<sup>th</sup> Hawaii International Conference on System Sciences*.
- Dempster, M., Alm, N., and Reiter, E. 2010. "Automatic generation of conversational utterances and narrative for Augmentative and Alternative Communication: a prototype system," *Proceedings of the 2010 Workshop on Speech and Language Processing for Assistive Technologies*, pp. 10–18.

- Desai, T., Chow, K., Mumford, L., Hotze, F., and Chau, T. 2014. "Implementing an iPad-based alternative communication device for a student with cerebral palsy and autism in the classroom via an access technology delivery protocol," *Computers & Education* (79), pp. 148–158.
- Ehlich, K., and Rehbein, J. 1972. "Zur Konstitution pragmatischer Einheiten in einer Institution: Das Speiserestaurant," *Linguistische Pragmatik*, pp. 209–254.
- Elspaß, S. 2018. "Sprachvariation und Sprachwandel," *Sprache in sozialen Gruppen*, pp. 87–107.
- Fang, Y., Xu, Y., Teng, Z., Jiang, Z., and Xiang, W. 2023. "SocializeChat: A GPT-Based AAC Tool for Social Communication Through Eye Gazing," *Proceedings of the 2023 ACM International Joint Conference on Pervasive and Ubiquitous Computing*, pp. 128–132.
- Feuerriegel, S., Hartmann, J., Janiesch, C., and Zschech, P. 2023. "Generative AI," *Business & Information Systems Engineering* (66), pp. 111–126.
- Filippidou, F., and Moussiades, L. 2020. "A benchmarking of IBM, Google and Wit automatic speech recognition systems," *Proceedings of the 16<sup>th</sup> IFIP WG 12.5 International Conference*.
- Fleiss, J., Levin, B., and Paik, M. 2013. "Statistical methods for rates and proportions," *John Wiley & Sons*.
- Fritzell, B. 1996. "Voice disorders and occupations," *Logopedics Phoniatrics Vocology* (21:1), pp. 7–12.
- Garcia, L., de Oliveira, L., and de Matos, D. 2016. "Evaluating pictogram prediction in a location-aware augmentative and alternative communication system," *Assistive Technology* (28:2), pp. 83–92.
- Giles, H., Coupland, N., and Coupland, J. 1991. "Accommodation theory: Communication, context, and consequence," *Cambridge University Press*.
- Gregor, S. 2006. "The nature of theory in information systems," *MIS Quarterly* (30:3), pp. 611–642.
- Gregor, S., and Hevner, A. 2013. "Positioning and Presenting Design Science Research for Maximum Impact," *MIS Quarterly* (37:2), pp. 337–355.
- Gregor, S., and Jones, D. 2007. "The Anatomy of a Design Theory," *Journal of the AIS* (8:5), pp. 312–335.
- Gregor, S., Chandra Kruse, L., and Seidel, S. 2020. "Research Perspectives: The Anatomy of a Design Principle," *Journal of the AIS* (21:6), pp. 1622–1652.
- Grice, H. 1975. "Logic and conversation," *Speech Acts, New York: Academic Press*, pp. 41–58.
- Heinrich, P., and Schwabe, G. 2014. "Communicating Nascent Design Theories on Innovative Information Systems through Multi-Grounded Design Principles," *Proceedings of the 9<sup>th</sup> International Conference on Design Science Research in Information Systems and Technology (DESRIST)*.
- Henne, H., and Rehbock, H. 2012. "Einführung in die Gesprächsanalyse," *de Gruyter*.
- Heo, Y., Kang, S. 2019. "Augmentative and Alternative Communication System Using Information Priority and Retrieval," *Proceedings of the 2019 International Conference on Big Data and Smart Computing*.
- Hevner, A., March, S., Park, J., and Ram, S. 2004. "Design Science in Information Systems Research," *MIS Quarterly* (28:1), pp. 75–105.
- Higginbotham, D., Leshner, G., Moulton, B., and Roark, B. 2012. "The application of natural language processing to augmentative and alternative communication," *Assistive Technology* (24:1), pp. 14–24.
- Hoag, L., Bedrosian, J., McCoy, K., and Johnson, D. 2004. "Trade-offs between informativeness and speed of message delivery in augmentative and alternative communication," *Journal of Speech, Language, and Hearing Research* (47), pp. 1270–1285.
- Hollender, N., Hofmann, C., Deneke, M., and Schmitz, B. 2010. "Integrating cognitive load theory and concepts of human–computer interaction," *Computers in Human Behavior* (26:6), pp. 1278–1288.
- Hourcade, J., Everhart Pilotte, T., West, E., and Parette, P. 2004. "A history of augmentative and alternative communication for individuals with severe and profound disabilities," *Focus on Autism and other Developmental Disabilities* (19:4), pp. 235–244.
- Kane, S., Morris, M., Paradiso, A., and Campbell, J. 2017. "At times avuncular and cantankerous, with the reflexes of a mongoose - Understanding Self-Expression through AAC Devices," *Proceedings of the 2017 Conference on Computer Supported Cooperative Work and Social Computing*.
- Klauer, A., Johannsen, F., and Leist, S. 2021. "BeMyVoice - An iPad-Based Helper for Speech Impaired Persons," *Proceedings of the 16<sup>th</sup> International Conference on DESRIST*.
- Konradl, D., Wörner, J., Luttner, L., and Leist, S. 2023. "Artificial Intelligence in Augmentative and Alternative Communication Systems - A Literature-Based Assessment and Implications of Different Conversation Phases and Contexts," *Proceedings of the 31<sup>st</sup> European Conference on Information Systems*.
- Kristensson, P., Lilley, J., Black, R., and Waller, A. 2020. "A design engineering approach for quantitatively exploring context-aware sentence retrieval for nonspeaking individuals with motor disabilities," *Proceedings of the 2020 CHI Conference on Human Factors in Computing Systems*.

- Laxmidas, K., Avra, C., Wilcoxon, C., Wallace, M., Spivey, R., Ray, W., Polsley, S., Kohli, P., Thompson, J., and Hammond, T. 2021. "CommBo: Modernizing Augmentative and Alternative Communication," *International Journal of Human-Computer Studies* (145).
- March, S., and Smith, G. 1995. "Design and natural science research on information technology," *Decision Support Systems* (15:4), pp. 251–266.
- Mayring, P. 2000. "Qualitative Content Analysis," *Forum Qualitative Social Research* (1:2).
- Mayring, P. 2002. "Introduction to qualitative social research," *Weinheim: Beltz*.
- Mitchell, M., and Sproat, R. 2012. "Discourse-Based Modeling for AAC," *Proceedings of the 3<sup>rd</sup> Workshop on Speech and Language Processing for Assistive Technologies*.
- Möller, F., Guggenberger, T., and Otto, B. 2020. "Towards a method for design principle development in information systems," *Proceedings of the 15<sup>th</sup> International Conference on DESRIST*.
- Neamtu, R., Camara, A., Pereira, C., Ferreira, R. 2019. "Using artificial intelligence for augmentative alternative communication for children with disabilities," *Proceedings of the 2019 IFIP Conference on HCI*.
- Obiorah, M., Piper, A., and Horn, M. 2021. "Designing AACs for people with aphasia dining in restaurants," *Proceedings of the 2021 CHI Conference on Human Factors in Computing Systems*.
- OpenAI. 2024. "ChatGPT-4.0," Accessed on 11.02.2024 [via this Link](#).
- Peppers, K., Tuunanen, T., Rothenberger, M., Chatterjee, S. 2007. "A design science research methodology for information systems research," *Journal of Management Information Systems* (24:3), pp. 45–77.
- Peppers, K., Rothenberger, M., Tuunanen, T., and Vaezi, R. 2012. "Design science research evaluation," *Proceedings of the 7<sup>th</sup> International Conference on DESRIST*.
- Purao, S., Kruse, L., and Maedche, A. 2020. "The Origins of Design Principles: Where do... they all come from?," *Proceedings of the 15<sup>th</sup> International Conference on DESRIST*.
- Radford, A., Wu, J., Child, R., Luan, D., Amodei, D., and Sutskever, I. 2019. "Language models are unsupervised multitask learners," *OpenAI Blog* (1:8).
- Reddington, J., and Tintarev, N. 2011. "Automatically generating stories from sensor data," *Proceedings of the 16<sup>th</sup> International Conference on Intelligent User Interfaces*.
- Schindler, E. 2013. "Kommunikative Ritualität im Gastronomiebereich," *Freiburger Arbeitspapiere zur germanistischen Linguistik, 11. Freiburg: Germanistische Linguistik, Universität Freiburg*.
- Shen, J., Yang, B., Dudley, J., and Kristensson, P. 2022. "Kwickchat: A multi-turn dialogue system for AAC using Context-Aware Sentence Generation by Bag-of-Keywords," *Proceedings of the 27<sup>th</sup> International Conference on Intelligent User Interfaces*.
- Snyder, H. 2019. "Literature review as a research methodology: An overview and guidelines," *Journal of Business Research* (104), pp. 333–339.
- Sweller, J. 1988. "Cognitive load during problem solving: Effects on learning," *Cognitive Science* (12:2).
- Teubner, T., Flath, C., Weinhardt, C., van der Aalst, W., and Hinz, O. 2023. "Welcome to the era of chatgpt et al.," *Business & Information Systems Engineering* (65:2), pp. 95–101.
- Tintarev, N., Reiter, E., Black, R., Waller, A., and Reddington, J. 2016. "Personal storytelling: Using Natural Language Generation for children with complex communication needs, in the wild ...," *International Journal of Human-Computer Studies* (92), pp. 1–16.
- Todman, J., and Alm, N. 1997. "Pragmatics and AAC approaches to conversational goals in natural language processing for communication aids," *Proceedings of the 35<sup>th</sup> Annual Meeting of the ACL*.
- Valencia, S., Cave, R., Kallarackal, K., Seaver, K., Terry, M., and Kane, S. 2023. "The less I type, the better: How AI Language Models can Enhance or Impede Communication for AAC Users," *Proceedings of the 2023 CHI Conference on Human Factors in Computing Systems*.
- Vanderheiden, G. 1983. "Non-Conversational Communication Technology Needs of Individuals with Handicaps," *Rehabilitation World* (7:2), pp. 8–12.
- Venable, J., Pries-Heje, J., and Baskerville, R. 2016. "FEDS: a framework for evaluation in design science research," *European Journal of Information Systems* (25:1), pp. 77–89.
- Vom Brocke, J., Simons, A., Riemer, K., Niehaves, B., Plattfaut, R., and Clevén, A. 2015. "Standing on the shoulders of giants: Challenges and recommendations of literature search in information systems research," *Communications of the AIS* (37:1), pp. 205–224.
- Walls, J., Widmeyer, G., and El Sawy, O. 1992. "Building an Information System Design Theory for Vigilant EIS," *Information Systems Research* (3:1), pp. 36–59.
- Walsh, T. 2010. "Utterance-based systems: organization and design of AAC interfaces," *Proceedings of the 12<sup>th</sup> International Conference on Computers and Accessibility*.
- Westphal, M., Vössing, M., Satzger, G., Yom-Tov, G. 2023. "Decision control and explanations in human-AI collaboration: Improving user perceptions and compliance," *Computers in Human Behavior* (144).